Materials Science and Engineering Core Curriculum

MSE Faculty

September 14, 2016

This working document includes all of course content that has been developed as part of our department initiative to digitize our course content. This particular document has everything in one place, but can be a little overwhelming. To get the pieces broken up in more useful ways, go to msecore.northwestern.edu.

Contents

Contents 1

I 301: Materials Science Principles 15

1 301 Laboratories 15

1.1 Laboratory 1: Crystal Structure, Imperfections, and Diffraction . . . 15
1.2 Laboratory 2: Electrical Conductivity & Optoelectronics . . . 19
1.3 Laboratory 3: Mechanical Properties of Metals . . . 22
1.4 Laboratory 4: Phase Diagrams and Transformations . . . 25
1.5 Laboratory 5: Mechanical Behavior of Ceramics and Polymers . . . 29
1.6 Discovery Lab 1: Cholesteryl Ester Liquid Crystals . . . 33

II 314: Thermodynamics of Materials 38

2 Introduction 38

3 Big Ideas of Thermodynamics 38

3.1 Classification of Systems . . . 38
3.2 Thermodynamic State Variables . . . 39
3.3 Process Variables . . . 39
3.4 Extensive and Intensive Properties . . . 41

4 The Laws of Thermodynamics 42

4.1 1st Law of Thermodynamics . . . 42
4.2 2nd Law of Thermodynamics . . . 42
4.3 Reversible and Irreversible Processes . . . 43
4.4 3rd Law of Thermodynamics . . . 44
4.5 Combined statement of 1st and 2nd Laws . . . 44

5 Thermodynamic Variables and Relations 45

5.1 Enthalpy . . . 45
5.2 Helmholtz Free Energy . . . 46
5.3 Gibbs Free Energy . . . 46
5.4 Other Material Properties . . . 46
5.5 Coefficient Relations . . . 47
5.6 General Strategy for Deriving Relations . . . 48

6 The Ideal Gas 50

6.1 Specific Heat of Ideal Gas . . . 50
6.2 Internal Energy of Ideal Gas U(T,P) . . . 51
6.3 Examples Using Ideal Gas . . . 51

7 Solids and Liquids 55

7.1 Example 4.13: ΔG for MgO at constant P . . . 55
8 Conditions for Equilibrium **60**
8.1 Equilibrium in thermodynamic system . . . 60
8.2 Finding equilibrium conditions . . . 62
8.3 Minima of energies as conditions of equilibrium . . . 62
8.4 Maxwell Relations . . . 64
9 From Microstates to Entropy **65**
9.1 Microstates and Macrostates . . . 65
9.2 Conditions for equilibrium . . . 68
10 Noninteracting (Ideal) Gas **70**
10.1 Ideal gas Law . . . 71
11 Diatomic Gases **71**
12 The Einstein Model of a Crystal **72**
12.1 Thermodynamic Properties of a Crystal . . . 72
13 Unary Heterogeneous Systems **73**
13.1 Phase Diagrams . . . 74
13.2 Chemical Potential Surfaces . . . 74
13.3 Calculation of $\mu(T, P)$ Surfaces . . . 75
13.4 Clasius-Clapeyron Equation . . . 77
14 Open Multicomponent Systems **81**
14.1 Partial Molal Properties of Open Multicomponent Systems . . . 81
14.2 The Mixing Process: Calculating changes in energy . . . 82
14.3 Graphical Determination of Partial Molal Properties . . . 84
14.4 Chemical Potential in Multicomponent Systems . . . 85
14.5 Properties of ideal solutions (e.g. ideal gas mixtures) . . . 85
14.6 Behavior of Dilute Solutions . . . 86
14.7 Excess Properties, relative to ideal . . . 87
14.8 Beyond the ideal solution: the "regular" solution . . . 88
15 Multicomponent Heterogeneous Systems **91**
15.1 Conditions for Equilibrium . . . 91
15.2 Gibbs Phase Rule . . . 92
16 Phase Diagrams **93**
16.1 Unary Phase Diagrams . . . 93
16.2 Binary Phase Diagrams . . . 94
16.3 Interpreting Phase Diagrams: Tie Lines . . . 94
16.4 The Lever Rule . . . 95
16.5 Applications of Phase Diagrams: Microstructure Evolution . . . 96
17 Review **96**
17.1 Types of $\alpha - L$ phase diagrams . . . 96
18 Thermodynamics of Phase Diagrams **97**
18.1 Calculation of $\Delta G_{mix}$from pure component reference states . . . 97
18.2 The Common Tangent Construction . . . 98
18.3 A Monotonic Two-Phase Field . . . 99
18.4 Non-monotonic Two-Phase Field . . . 100
18.5 Properties of a regular solution . . . 100
18.6 Misciblity gap: appearance upon cooling . . . 100
18.7 Spinodal decomposition: concave down region in ΔG drives unmixing and "uphill" diffusion . . . 100
19 Summary of 314 **102**
19.1 Chapters 2 and 3 . . . 102
19.2 Chapter 4 . . . 102
19.3 Chapter 5 . . . 102
19.4 Chapter 6 . . . 102
19.5 Chapter 7 . . . 103
19.6 Chapter 8 . . . 103

19.7 Chapter 9 . . . 104
19.8 Chapter 10 . . . 104
20 314 Problems 105
20.1 Phases and Components . . . 105
20.2 Intensive and Extensive Properties . . . 105
20.3 Differential Quantities and State Functions . . . 105
20.4 Entropy . . . 105
20.5 Thermodynamic Data . . . 106
20.6 Temperature Equilibration . . . 107
20.7 Statistical Thermodynamics . . . 108
20.8 Single Component Thermodynamics . . . 108
20.9 Mulitcomponent Thermodynamics . . . 109
21 314 Computational Exercises 109

III 315: Applications of Thermodynamics 112
22 Introduction 112
23 Gibbs-Duhem Equation 113
23.1 Degrees of Freedom . . . 115
23.2 Phase Rule . . . 115
23.3 Classifying Phase Diagrams . . . 117
24 Type I (Single Component) Phase Diagrams 118
24.1 The Ellingham-Richardson Diagram . . . 119
24.2 Two More Type I Phase Diagrams . . . 124
25 Type II (Binary) Phase Diagrams 125
25.1 Estimating Liquidus Lines . . . 126
25.2 Estimating Liquidus and Solidus Lines . . . 128
25.3 Estimating Solvus Lines . . . 129
25.4 Activity vs. Composition Plots . . . 131
25.5 Schematic Free Energy vs. Composition Diagrams . . . 137
25.6 Regular Solution Predictions . . . 141
25.7 Oxygen Partial Pressure vs. Composition Diagrams . . . 142
26 Type III (Ternary) Phase Diagrams 144
26.1 The Ternary Lever Rule . . . 146
26.2 Dealing with an Additional Degree of Freedom . . . 146
26.3 Liquidus Projection Diagrams . . . 146
26.4 Hummel's Rules . . . 148
26.5 Isothermal Sections . . . 157
26.6 Crystallization Paths and Microstructure Evolution . . . 160
26.7 Isothermal Sections of "Real" A-B-C Systems . . . 162
26.8 Some Technologically Important Type III Phase Diagrams . . . 165
27 Fick's Laws of Diffusion 169
27.1 Fick's First Law . . . 170
27.2 Random Walk Diffusion . . . 171
27.3 Steady State Diffusion . . . 172
27.4 Fick's Second Law . . . 173
28 Applications of Fick's Laws 175
28.1 Homogenization and Point Defect Relaxation . . . 175
28.2 Non-Infinite Systems . . . 177
28.3 Semi-Infinite Systems . . . 178
29 Atomistics of Diffusion 185
29.1 Interstitial Diffusion . . . 186
29.2 Vacancy Diffusion . . . 189
29.3 Interstitialcy Diffusion . . . 193
29.4 The Arrhenius Behavior of Diffusion . . . 194

30 Existence and Motion of Point Defects **194**
30.1 The Simmons-Balluffi Experiment . . . 195
30.2 Vacancy Formation in fcc Metals . . . 197
30.3 The Bauerle and Kohler Experiment . . . 197
30.4 Self-Diffusion and Vacancy Motion in fcc Metals . . . 200
31 Point Defects and Transport in Ceramics **201**
31.1 Kröger-Vink Notation . . . 201
31.2 Rules for Balancing Point Defect Reactions . . . 202
31.3 Point Defect Thermodynamics in Ceramics . . . 208
31.4 Brouwer Diagrams . . . 210
32 Electrical Conductivity **220**
32.1 Electronic Conductivity . . . 220
32.2 Ionic Conductivity . . . 221
33 315 Problems **223**
33.1 301 Computational Exercises . . . 232

IV 316-1: Microstructural Dynamics I **233**
34 Diffusion **233**
34.1 Review of the Basic Equations . . . 233
34.2 Mole Fractions and Volume Fractions . . . 235
34.3 Vacancy Diffusion Mechanism . . . 236
34.4 Kirkendall Effect . . . 238
34.5 The Interdiffusion Coefficient . . . 239
34.6 Connection to thermodynamics . . . 241
34.7 Summary of Diffusion in a Binary System . . . 242
34.8 Diffusion in Ternary Systems . . . 243
34.9 Crystal Defects and High Diffusivity Paths . . . 243
35 Dislocations **244**
35.1 Edge Dislocations . . . 248
35.2 Screw Dislocations . . . 250
35.3 The Burgers Circuit . . . 251
35.4 The $\vec{b} \times \hat{s}$ cross product . . . 253
35.5 Connection to the Crystal Structure . . . 254
35.6 Dislocation loops . . . 255
35.7 Dislocation Density . . . 256
35.8 Dislocation Motion . . . 256
35.9 Dislocation Energy . . . 258
35.10 Dislocation Line Tension . . . 261
35.11 Effect of Applied Stress . . . 262
35.12 Dislocation Multiplication . . . 263
35.13 Interactions Between Dislocations . . . 264
36 Thermodynamics of interfaces **267**
36.1 A Brief Review of the Thermodynamic Potentials . . . 267
36.2 Temperature and Chemical Potential Equilibrium . . . 268
36.3 The Dividing Surface . . . 270
36.4 Chemical Potential Expressions . . . 273
36.5 A practical example . . . 274
36.6 Curvature effects on T_m . . . 274
36.7 Size-dependent solubility . . . 277
37 Surfaces and Interfaces as Two Dimensional Defects **280**
37.1 Surface energy of a close packed plane . . . 282
37.2 Orientation Dependence of the Surface Energy . . . 283
37.3 Equilibrium Shape of Crystals . . . 284
38 Grain Boundaries **285**

38.1 Tilt Boundaries . . . 286
38.2 Twin Boundaries . . . 287
38.3 Twist Boundaries . . . 289
38.4 Interface Junctions . . . 289
38.5 Thermally activated migration of grain boundaries . . . 289
39 Three Phase Contact Lines **292**
39.1 Wetting . . . 293
39.2 Grain Boundary Junctions . . . 294
40 Interphase Interfaces **294**
40.1 Coherent Interfaces . . . 294
40.2 Semicoherent Interfaces . . . 295
40.3 Incoherent interfaces . . . 296
40.4 Case Study I: the Si-Ge system . . . 296
40.5 Case Study II: The Cu/Al system . . . 298
40.6 Second phase shape . . . 298
40.7 Elastic Effects . . . 298
40.8 Effects of Elastic Anisotropy . . . 300
41 Thin Film Growth **300**
42 Review Questions **302**
42.1 Diffusion . . . 302
42.2 Dislocations . . . 302
42.3 Solid/Liquid and Solid/Vapor Interfaces . . . 303
42.4 Grain Boundaries . . . 303
42.5 Interphase Interfaces . . . 304
42.6 Crystallization or Recrystallization . . . 304
42.7 MATLAB . . . 304
43 316-1 Problems **305**
43.1 Introduction . . . 305
43.2 Diffusion . . . 305
43.3 Stress and Strain . . . 311
43.4 Dislocation Structure . . . 311
43.5 Dislocation Interactions . . . 313
43.6 Interfacial Thermodynamics . . . 317
43.7 Surface and Interface Structure . . . 318
43.8 Wetting and Contact Angles . . . 319
43.9 Grain Boundaries . . . 320
43.10 Transformation Kinetics . . . 320
43.11 Interphase Interfaces . . . 320
44 316-1 Simulation Exercise **325**
44.1 Background . . . 325
44.2 Exercise . . . 329

V 316-2: Microstructural Dynamics II **333**
45 Nucleation **333**
45.1 Thermodynamics Review . . . 336
45.2 Thermodynamics of Pure Elements . . . 338
45.3 The Critical Nucleus . . . 338
45.4 Solidification . . . 340
45.5 Homogenous nucleation rate: . . . 343
45.6 Heterogeneous Nucleation . . . 344
45.7 Nucleation in a Binary System . . . 348
45.8 Effects of Strain Energy on Nucleation Barrier . . . 350
45.9 Heterogeneous Nucleation . . . 354
46 Coarsening Kinetics **357**

47 Precipitate Growth **359**
47.1 Growth of Planar Incoherent Interfaces . . . 359
47.2 Curvature Effects . . . 364
48 Case Study: Precipitation of Cu-rich Precipitates from Al: **367**
49 Case Study: Mineralization from Solution **373**
50 Case Study: Nanoparticle Nucleation and Growth **374**
50.1 Motivation . . . 374
50.2 Wulff Constuction . . . 375
50.3 Kinetic Control of Particle Shape . . . 375
50.4 Controlling Growth Rates . . . 376
51 Interface Stability **377**
52 Eutectic Solidification **382**
53 Eutectoid Reactions **386**
54 Spinodal Decomposition **388**
55 Review Questions **398**
56 Case study: Damascus Steel **399**
57 Matlab Example Scripts **401**
58 Lab Examples **401**
58.1 Coarsening . . . 402
58.2 Secondary Arm Spacing . . . 403
58.3 Surface to Volume Ratio . . . 404
58.4 Temperature Dependence: K . . . 405
58.5 LSW Theory . . . 406
58.6 Estimating D . . . 406
58.7 Estimating the Expected Coarsening Constant . . . 406
58.8 Heat treatment of Al alloys . . . 407
59 Appendix: Thermodynamic Data for Cu **408**
60 316-2 Problems **413**
60.1 General . . . 413
60.2 Laplace Pressure Derivation . . . 413
60.3 Homogeneous Nucleation . . . 413
60.4 Surface and Interface Effects . . . 415
60.5 Heterogeneous Nucleation . . . 416
60.6 Nucleation in a Binary System . . . 416
60.7 Spinodal Decomposition . . . 418
60.8 Constitutional Undercooling and the 'Mushy Zone' . . . 419
60.9 Coarsening . . . 420
60.10 Eutectic Solidification . . . 421
60.11 Eutectoid Transormations . . . 422
60.12 Transitional Phases . . . 423
60.13 TTT diagrams . . . 424
60.14 Mineralization . . . 424
60.15 Review Questions . . . 425

VI 332: Mechanical Behavior of Materials **427**
61 Introduction **427**
62 Stress and Strain **428**
62.1 Tensor Representation of Stress . . . 429
62.2 Tensor Transformation Law . . . 430
62.3 Principle Stresses . . . 433
62.4 Strain . . . 437
62.5 Small Strains . . . 437
62.6 Deformation Modes . . . 439

62.7 Representative Moduli . . . 442
62.8 Case Study: Speed of Sound . . . 442
63 Contact Mechanics **443**
63.1 Sign conventions . . . 443
63.2 Flat Punch Indentation . . . 443
63.3 Contact of Paraboloids . . . 450
63.4 Indentation with Berkovich Trips . . . 452
64 Fracture **454**
64.1 Fracture Modes . . . 455
64.2 Brittle behavior (fracture mechanics) . . . 455
64.3 Griffith Model and energy release rate . . . 458
64.4 Fracture Toughness of Materials . . . 462
64.5 Case Study: Fracture toughness of glassy polymers . . . 463
65 Matrix representation of Stress and Strain **468**
65.1 Compliance matrix . . . 469
65.2 Stiffness Matrix . . . 469
65.3 Symmetry Requirements . . . 469
66 Other Symmetry-Related Constitutive Relationships **474**
67 Crystal Systems and Point Groups **474**
68 Yield Behavior **476**
68.1 Critical Resolved Shear Stress . . . 477
68.2 Yield Surfaces . . . 478
68.3 Localized Deformation . . . 479
68.4 Deformation Twinning . . . 483
68.5 Deformation of Polymers . . . 484
69 Strengthening Mechanisms in Metals **486**
69.1 Work Hardening . . . 487
70 Linear Viscoelasticity **487**
70.1 Time and Frequency Dependent Moduli . . . 488
70.2 Torsional Resonator . . . 490
70.3 Viscoelastic Models . . . 491
71 Creep Behavior **494**
71.1 Creep in the Linear Viscoelastic Regime . . . 494
71.2 Nonlinear Creep: Potential Separability of Stress and Time Behaviors . . . 496
71.3 Use of empirical, analytic expressions . . . 496
71.4 Eyring Model of Steady State Creep . . . 497
71.5 High Temperature Creep in Metals . . . 499
71.6 Deformation mechanism maps . . . 503
72 Weibull Analysis of Failure **504**
73 Intrinsic and Extrinsic Fracture Toughness **505**
73.1 Extrinsic Toughening . . . 505
74 Materials Selection: A Case Study **510**
75 Appendix 1: Symmetry Relationships **514**
76 Appendix 2: MATLAB Code Examples **516**
77 Appendix 3: Introduction to Abaqus **516**
77.1 A First Look at Abaqus/CAE . . . 516
77.2 The Model Tree: . . . 517
77.3 Partitioning the Part . . . 519
77.4 Meshing with Uneven Density . . . 519
77.5 Data Visualization . . . 520
78 332 Problems **521**
78.1 Course Organization . . . 521
78.2 The Stress Tensor . . . 521
78.3 Strains . . . 522

78.4 Contact Mechanics . . . 522
78.5 Fracture Mechanics . . . 523
78.6 Matrix Representation of Stress and Strains . . . 524
78.7 Typical Moduli . . . 525
78.8 Yield Criteria . . . 525
78.9 Strengthening Mechanisms . . . 526
78.10 Viscoelasticity . . . 526
78.11 Nonlinear Viscoelasticity and Creep . . . 529
78.12 Nanoindentation . . . 531
78.13 Weibull Statistics . . . 532
78.14 Fracture Toughness . . . 532
78.15 General Knowledge . . . 532

79 332 Computational Exercises 532
79.1 MAT SCI 332 Report Guidelines . . . 532
79.2 FEA of a Cantilever Beam . . . 533
79.3 FEA of Stress Concentrators . . . 533

VII 331: Soft Materials 536

80 Introduction 536
80.1 Synthetic Polymers . . . 537
80.2 Understanding Polymer Chemistry . . . 539

81 Polymerization Reactions (week 1) 540
81.1 Step-Growth Polymerizations . . . 541
81.2 Interfacial Polymerizations . . . 542

82 Molecular Weight Distributions (week 2) 544
82.1 Number Average Molecular Weight . . . 545
82.2 Weight Average Molecular Weight . . . 545
82.3 M Evolution during Step Growth . . . 547

83 Non-Linear Step Growth Polymerizations (week 2) 551
83.1 Gelation . . . 551
83.2 Prepolymers . . . 552
83.3 Corothers Theory of Gelation . . . 552

84 Chain-Growth Polymerizations 554
84.1 Chain Growth Polymerization Mechanisms . . . 555
84.2 Reactive Species for Chain Growth Polymerization . . . 555
84.3 Initiation . . . 556
84.4 Termination Mechanisms . . . 557
84.5 Intramolecular Chain Transfer . . . 558
84.6 Polymerization of Dienes . . . 559
84.7 Living Polymerizations . . . 560
84.8 Vinyl Polymers and Tacticity . . . 560

85 Common Polymers 561
85.1 Chain Growth: Addition to a Double Bond . . . 561
85.2 Chain Growth: Ring Opening . . . 565
85.3 Step Growth Polymers . . . 566

86 Crosslinking 568

87 Copolymers 569

88 The Glass Transition 570
88.1 Free volume . . . 570
88.2 Enthalpy and Heat Capacity . . . 573
88.3 Differential Scanning Calorimetry (DSC) . . . 573

89 Structure of Amorphous Polymers 574
89.1 End-to-end Vector for Polymer Molecules . . . 575
89.2 Random Walk in One Dimension . . . 575

89.3 Average of a Function . . . 576
89.4 Averages for Random Walks . . . 577
89.5 Random Walks in 3 Dimensions . . . 578
90 Models of Chain Dimensions **581**
90.1 General Considerations . . . 582
90.2 Freely Jointed Chain Model . . . 583
90.3 Freely Rotating Chain Model . . . 583
90.4 Bond Angle Restrictions . . . 585
90.5 Characteristic Ratio . . . 585
90.6 Self-Avoiding Random Walks . . . 586
90.7 Deformation of a Single Polymer Molecule . . . 587
91 Elasticity **589**
91.1 Fundamental Definitions . . . 589
91.2 Rubber Elasticity . . . 593
92 Viscoelasticity **597**
92.1 Relaxation Modulus . . . 599
92.2 Boltzmann Superposition Principle . . . 599
92.3 Idealized Relaxation Curves . . . 601
92.4 Temperature Dependence . . . 602
92.5 Dynamic Moduli . . . 603
93 Semicrystalline Polymers **608**
93.1 Structural Hierarchy in Semicrystalline Polymers . . . 609
93.2 The Structural Repeat Unit . . . 609
93.3 Helix Formation . . . 609
93.4 Crystalline unit cells . . . 611
93.5 Stiffness of Polymer Single Crystals . . . 613
93.6 Chain Folding . . . 613
93.7 Spherulitic morphologies . . . 615
93.8 Birefringence . . . 617
93.9 Growth of a Lamellar Crystallite . . . 618
93.10 Density of Semicrystalline Polymers . . . 619
94 Liquid Crystals **620**
94.1 Nematic Liquid Crystals . . . 620
94.2 Nematic Liquid Crystals - Texture . . . 621
94.3 Liquid Crystal Displays . . . 621
95 Solutions and Blends **621**
95.1 Chemical Potentials and Free Energy of Mixing . . . 622
95.2 Ideal Entropy of Mixing for Polymers . . . 625
95.3 Idealized Enthalpy of Mixing . . . 625
95.4 Flory-Huggins Free Energy of Mixing . . . 627
95.5 The Coexistence (binodal) and Spinodal Curves . . . 628
95.6 Critical Point . . . 630
95.7 Spinodal Decomposition . . . 630
95.8 Phase Diagrams in Temperature Space . . . 631
95.9 Chemical Potentials . . . 632
95.10 Osmotic Pressure . . . 634
95.11 Equilibrium Swelling of a Neohookean Material . . . 640
96 Surfaces and Interfaces **641**
96.1 Equilibrium Contact Angle . . . 641
96.2 Polymer/polymer Interfacial Free Energy . . . 642
96.3 Block Copolymer Morphologies . . . 644
96.4 The Effect of Block Copolymers on Phase Behavior . . . 646
97 Case Study: Triblock Gels **647**
97.1 Introduction . . . 647
97.2 Thermoreversible Gelcasting of Ceramics . . . 648

97.3 Quantifying the Solid/Liquid Transition . . . 649
97.4 Characterizing the Glass Transition in PMMA domains . . . 651
97.5 Concentration of the gel Modulus . . . 653
97.6 Hydrogels: Water as the Solvent . . . 655
98 Summary **657**
98.1 Classification Scheme . . . 657
98.2 Molecular Weight Averages . . . 657
98.3 Polymerizations . . . 657
98.4 Chain Dimensions . . . 658
98.5 Crystallization . . . 659
98.6 Rubber Elasticity . . . 659
98.7 Viscoelasticity . . . 660
98.8 Solutions and Blends . . . 661
98.9 Interfaces . . . 661
99 Review Questions **662**
99.1 Polymer Synthesis . . . 662
99.2 Molecular Weight Distributions . . . 662
99.3 Gels and Elastomers . . . 662
99.4 Glass transition . . . 662
99.5 Chain dimensions of amorphous polymers . . . 663
99.6 Rubber Elasticity . . . 663
99.7 Viscoelasticity . . . 663
99.8 Semicrystalline Polymers . . . 663
99.9 Solution Thermodynamics . . . 664
99.10 Interfaces and Self-Assembly . . . 664
100 331 Problems **664**
100.1 Basic chemistry review . . . 664
100.2 Polymerization Reactions . . . 665
100.3 The Glass Transition . . . 667
100.4 Chain Dimensions . . . 668
100.5 Mechanical Properties . . . 669
100.6 Crystallinity . . . 673
100.7 Liquid Crystals . . . 675
100.8 Thermodynamics of Solutions and Blends . . . 675
100.9 Capillarity . . . 678
101 331 Computational Exercises **679**
102 Appendix 1: MATLAB scripts **681**
102.1 Symbolic Math . . . 681
102.2 Solution to a Nonlinear Equation . . . 681
103 Appendix 2: Data Tables **682**
103.1 Dynamic Moduli . . . 682
104 Other Resources **685**
104.1 Polymer Synthesis . . . 685

VIII 351-1: Quantum Mechanics **686**
105 Why Do We Need Quantum Mechanics? **686**
105.1 Classical (Planetary) Picture of the Hydrogen Atom . . . 686
105.2 The Bohr Atom . . . 686
105.3 Units and Magnitude of h . . . 687
106 Origins of Quantum Theory **688**
106.1 Black Body Radiation . . . 688
107 Origins of Schrödinger Equation **692**
108 Time-Independent Potentials: Wells and Barriers **696**
109 3-D Potentials **699**

110 The Hydrogen Atom **701**
111 Bonding **705**
112 Supplement (Bridge to Lecture 9.1) **709**
113 Modeling Metals **709**
114 Free Electron Gas **709**
115 Phonons: Quantized Lattice Vibrations / Elastic Waves **711**
116 Band Theory of Solids **714**
116.1 Approach: . . . 714
116.2 Ziman Model: . . . 715
116.3 What is the actual ψ? . . . 715
116.4 "Proof in 1-D": . . . 715
116.5 Kronig-Penney Model: . . . 715
116.6 Reduced Zone Scheme: . . . 716
116.7 More Generally: . . . 716
116.8 Feynman Model: . . . 716
117 Implications of Band Theory **717**
118 Semiconductors **720**
119 Current Flow in Semiconductors **722**
120 351-1 Problems **727**

IX 351-2: Solid State Physics **732**
121 Catalog Description (351-1,2) **732**
122 Course Outcomes **732**
123 351-2: Solid State Physics **732**
124 Principles of Semiconductor Devices **732**
124.1 Law of Mass Action . . . 733
124.2 Chemistry and Bonding . . . 733
124.3 Reciprocal Lattice and the Brillouin Zone . . . 734
124.4 Nearly Free Electron Model . . . 735
124.5 Two Level Model . . . 736
125 Semiconductor Band Diagrams **736**
125.1 P- and N-Type Semiconductors . . . 737
125.2 P-N Junctions . . . 737
125.3 Boltzmann Statistics: Review . . . 737
125.4 P-N Junction Equilibrium . . . 738
125.5 Charge Profile of the p-n Junction . . . 738
125.6 Calculation of the Electric Field . . . 739
125.7 Calculation of the Electric Potential . . . 740
125.8 Junction Capacitance . . . 740
125.9 Rectification . . . 741
125.10 Junction Capacitance . . . 742
126 Transistors: Semiconductor Amplifiers **742**
126.1 pnp Device . . . 743
126.2 pnp Device . . . 744
126.3 Amplifier Gain . . . 744
126.4 Circuit Configurations . . . 745
126.5 MOSFETs . . . 746
126.6 Oxide-Semiconductor Interface . . . 746
126.7 Depletion and Inversion . . . 746
126.8 Channel Pinch-Off . . . 747
126.9 Tunnel Diodes . . . 747
126.10 Gunn Effect . . . 748
126.11 Gunn Diodes . . . 748

127 Heterojunctions and Quantum Wells **749**
127.1 Quantum Wells . . . 749
127.2 Heterostructures and Heterojunctions . . . 751
127.3 Layered Structures: Quantum Wells . . . 751
127.4 Transitions between Minibands . . . 751
127.5 Quantum Dots . . . 752
127.6 Quantum Dot Example: Biosensor . . . 752
127.7 Quantum Cascade Devices . . . 753
128 Optoelectronic Devices: Photodetectors and Solar Cells **753**
128.1 I-V Characteristics . . . 754
128.2 Power Generation . . . 755
129 Lasers **761**
129.1 Emission Rate and Laser Intensity . . . 761
129.2 Two Level System . . . 762
129.3 Planck Distribution Law . . . 762
129.4 Transition Rates . . . 763
129.5 Two Level System . . . 763
129.6 Three Level System . . . 764
129.7 Lasing Modes . . . 767
129.8 Laser Examples . . . 767
129.9 Threshold Current Density . . . 768
129.10 Comparison of Emission Types . . . 770
129.11 Cavities and Modes . . . 770
129.12 Semiconductor Lasers . . . 771
129.13 Photonic Bandgap Materials . . . 771
130 Band Diagrams **772**
130.1 Band Diagrams . . . 772
130.2 Heterojunctions and the Anderson Model . . . 772
130.3 Band Bending at p-n Junctions . . . 772
131 Dielectric Materials **775**
131.1 Macroscopic Dielectric Theory . . . 775
131.2 Relation of ϵ to the Microscopic Structure . . . 776
131.3 Contributions to the Polarizability . . . 777
131.4 Polarization in Solids . . . 777
131.5 Calculation of the Average Dipole Moment . . . 778
131.6 Polarizability of Solids . . . 779
131.7 Dielectric Constant for a Solid . . . 779
131.8 Claussius Mossotti Relation . . . 779
131.9 Frequency Dependence of the Polarizability . . . 780
131.10 Quantum Theory of Polarizability . . . 781
131.11 Advanced Dielectrics: Ferroelectrics . . . 782
132 Phase Transitions **782**
132.1 Lattice Instabilities . . . 782
132.2 Curie Weiss Law . . . 783
132.3 Ferroelectric Phase Transitions: Landau Theory . . . 783
132.4 Other Instabilities . . . 786
132.5 Piezoelectrics . . . 786
133 Diamagnetism and Paramagnetism **787**
133.1 Paramagnetism . . . 788
134 Ferromagnetism **794**
134.1 Ferromagnetic Phase Transition . . . 795
134.2 Molecular Field . . . 795
134.3 Ferrimagnetic Order in Magnetic Oxides . . . 801
134.4 Domains and Walls . . . 803
134.5 Anisotropy of Magnetization . . . 804

134.6 Ferrimagnetic Ordering - Exchange Terms . . . 805
135 Electro-optic and Nonlinear Optical Materials 806
135.1 Frequency Doubling . . . 807
135.2 Nonlinearity in Refractive Index . . . 807
135.3 Electro-Optic Modulators . . . 807
135.4 Optical Memory Devices . . . 808
135.5 Volume Holography . . . 808
135.6 Photorefractive Crystals . . . 808
135.7 Photorefractive Crystals . . . 809
135.8 Phase Conjugation - All Optical Switching . . . 809
135.9 Acousto-Optic Modulators . . . 810
135.10 Integrated Optics . . . 810
135.11 Dielectric Waveguides . . . 811
135.12 Example: Phase Shifter . . . 811
135.13 Example: $LiNbO_3$ Phase Shifter . . . 811
136 Superconductivity 812
136.1 BCS Theory of Superconductivity . . . 812
136.2 Density of States and Energy Gap . . . 812
136.3 Heat Capacity of Superconductors . . . 813
136.4 Tunneling in Superconductor Junctions . . . 813
136.5 Semiconductor-Insulator-Semiconductor Junction . . . 813
136.6 Type I and Type II Superconductors . . . 814
136.7 High T_C Superconductors . . . 814
136.8 Cuprate Superconductors . . . 814
137 351-2 Problems 815

X 361: Crystallography and Diffraction 818
138 Symmetry of Crystals 818
138.1 Types of Symmetry . . . 818
138.2 Projections of Symmetry Elements . . . 820
139 Crystal Lattices 826
139.1 Indexing within a crystal lattice . . . 826
139.2 Lattices . . . 830
140 Stereographic Projections 837
140.1 Representing (hkl) Planes with Vectors . . . 840
141 Reciprocal Lattice 843
142 Representative Crystal Structures 851
143 Introduction to Diffraction 858
144 Absorption/Emission 864
144.1 Spectrometers . . . 864
144.2 X-ray Absorption . . . 872
145 Scattering 878
145.1 X-ray Scattering & Polarization . . . 878
145.2 Thomson Scattering by two electrons . . . 880
145.3 Compton Scattering . . . 886
146 Kinematic Theory 890
147 Diffraction 897
147.1 Laue condition for diffraction . . . 897
147.2 Single Crystal Diffractometer . . . 898
147.3 Diffraction Methods . . . 899
147.4 Structure Factor Examples . . . 903
147.5 Width of Diffraction Peaks (single crystal) . . . 906
147.6 Polycrystalline Aggregate (Powder sample) . . . 908

147.7 Rotating Crystal Method . . . 917
148 Temperature Effects **922**
148.1 Debye Theory of Lattice Dynamics/heat capacity — Corrected by Waller . . . 926
149 Thin Film X-ray Scattering **928**
150 Order-Disorder **942**
150.1 Order-disorder geometry . . . 942
150.2 Long-range order (W.L. Bragg and E.J. Williams) . . . 942
151 Special Topics in X-ray Physics: Extended X-ray Absorption Fine Structure (EXAFS) **945**
151.1 X-ray Absorption of Fine Structure (XAFS): an overview . . . 945
152 361 Problems **957**

XI 390: Materials Design **963**
153 Catalog Description **963**
154 Course Outcomes **963**
155 390: Materials Design **963**

XII 391: Process Design **964**
156 391 Problems **964**

XIII 396: Senior Projects **974**
157 Catalog Description **974**
158 Course Outcomes **974**
159 396: Senior Projects **974**
160 Writing a Literature Review **974**
160.1 General Advice . . . 974
160.2 Checklist . . . 975
160.3 Additional Information . . . 975
160.4 Assessment Rubric . . . 976

XIV References, Nomenclature and Index **978**
References **979**
Nomenclature **980**
Index **984**

Part I
301: Materials Science Principles

1 301 Laboratories

1.1 Laboratory 1: Crystal Structure, Imperfections, and Diffraction

1.1.1 Objective

The objectives of this lab are to gain a better understanding of crystal structures, how they are measured and/or identified and the types of defects present in crystalline materials.

1.1.2 Outcomes

Upon completion of the laboratory, the student will be able to:

1. Distinguish FCC, BCC and HCP structures.
2. Understand how - broadly - diffraction enables qualitative identification and quantitative measurement of repeating structures.
3. Describe edge and screw dislocations and how they might affect material properties.

Experimental Details Tennis balls simulate "metal atoms" formed into different structures. A laser pointer, ruler and tape measure are used to measure spacing of a two-dimensional pattern, simulating the application of x-ray diffraction.

1.1.3 Exercise 1: Benchtop Simulation of Crystal Structures

A. FCC and HCP space-filling structures Each group should have a box, several tennis balls, and small wooden balls.

1. Arrange the tennis balls in a single close-packed layer in the bottom of the box. Start with a line of balls parallel to the "Start here" arrows along one side of the box.
2. Add a second (partial) layer. Identify the three-fold symmetry (six atoms around one) in the first and second planes. Compare this to the HCP model, and you will notice that this hexagonal structure occurs in the basal (base) plane with Miller indices (0001).
3. Add a third plane. At this point in building your structure, you must choose between two different sites. What is the difference between picking between one site and another? Which sites correpond to which crystal structure? (FCC vs HCP)
4. Identify the three-fold symmetry associated with the close-packed planes in many directions. In all, there are 4 such planes, perpendicular to the 4 different cube diagonals. The Miller indices are (111), ( ), ( ) and ( ). There is only one set of close-packed planes in HCP, (0001). What are the consequences of FCC having multiple close-packed planes and HCP having only one? Think about the class discussion of dislocations.
5. Identify the FCC unit cell. It isn't easy…..it is easier to start with an FCC model and find the close-packed planes. They form the diagonals of the cube! That is, the family of {111} planes. (Look for the half-unit cell styrofoam ball models, which you may put together and take apart – carefully!!!)
6. Due to applied stress or deformation of a material, "mistakes" in stacking can occur. These "mistakes" are known as stacking faults. Create a stacking fault in the third plane of tennis balls with one side FCC and the other side HCP.

B. Interstitial Sites

1. You will notice gaps, or interstices, between the spherical atoms. These "interstitial" sites leave room for small solute atoms, and reinforce that that atomic packing factor (APF) is less than 1.
2. Identify two different interstitial sites: octahedral and tetrahedral. To aid with visualization, you may take the small wooden balls ("solute atoms") and place them in the intersitial sites. What are the coordination numbers for each of these interstitial sites?

C. Body Centered Cubic Space Filling Model Build a BCC lattice with tennis balls. Be careful about which balls touch. How do the interstices of BCC compare to that of the close packed structures from Part A?

1.1.4 Exercise 2: Diffraction (HW is the last part below)

Diffraction occurs when wavelength of the source and obstacle spacing are similar in size. The wavelength of x-rays, λ = 0.1-0.5 nm, is well suited to investigate atoms with d-spacings with the same order of magnitude. Constructive interference occurs when the Bragg law is satisfied: $n\lambda = 2dsin\theta$, where n is an integer and diffraction (3-D sample) occurs at an angle 2θ from the incident direction. The spacing between atom planes is d.

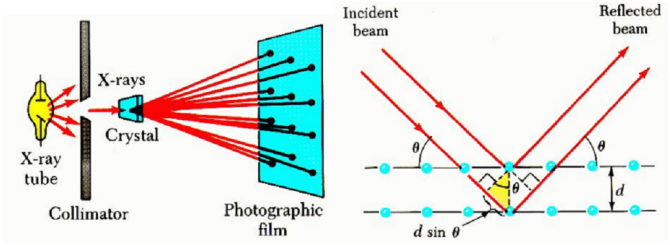


Figure 1.1: Diffraction condition

Visible light with wavelength about a thousand times larger (λ =0.4-0.7 μm) can be diffracted by small dots on a glass plate with a spacing similar to d. Optical diffraction of dot arrays is thus an analogy to crystal structure determination by x-ray diffraction.
Here, for a two-dimensional array, $n\lambda = dsin\theta$. This equation applies to the lab HW.

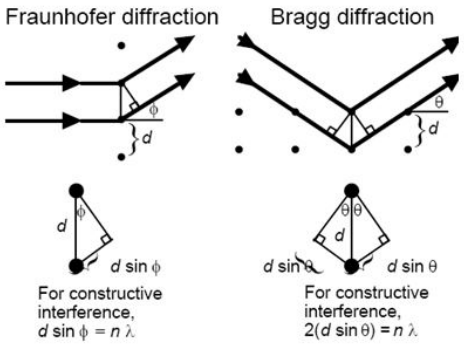
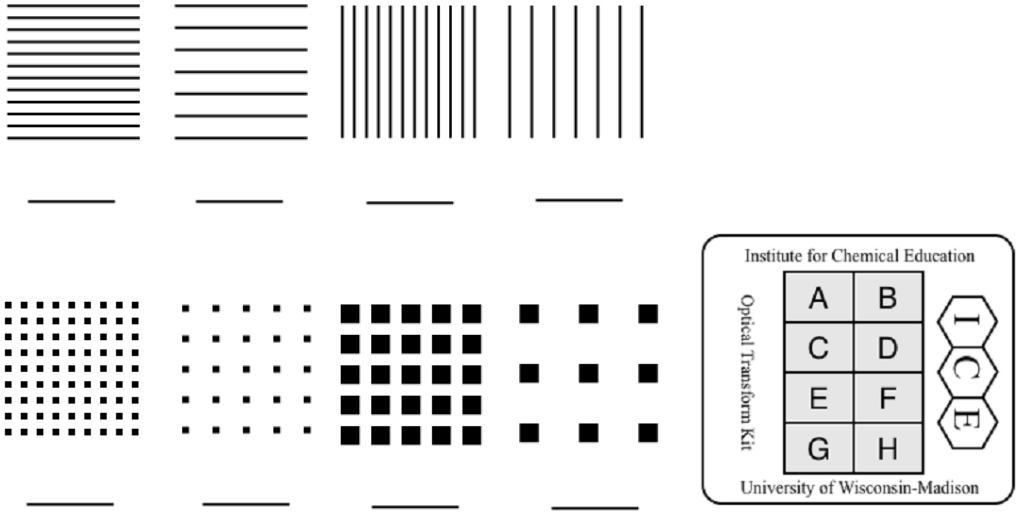


Figure 1.2: Figures from: The optical transform: Simulating diffraction experiments in introductory courses George C. Lisensky , Thomas F. Kelly , Donald R. Neu and Arthur B. Ellis J. Chem. Educ., 1991, 68 (2), p 91 DOI: 10.1021/ed068p91 Publication Date: February 1991

Two 35 mm slides are "samples" for this lab. Each slide contains eight patterns as indidated below and are designated A-H. Make sure the ICE text is on the right hand side, facing the laser, as shown below.

1. Discovery slide = green
2. Unit cell slide = black

Warm-up exercise: Label each of the patterns below with its location (a-h) on the slide to the right. You may check your conclusions by looking at the slide under the stereomicroscope. Don't spend too much time here- make sure you proceed to the quantitative measurements, below.



1.1.5 Lab Writeup

Due at the beginning of the next lab session.
For the lab writeup you will be using the black unit cell slide. See note about measurement accuracy, above. Record your measurements and calculations with the appropriate number of significant figures. Estimate the uncertainty in each computed value. (Assume your measurement tools are accurate, but with limited precision.)

1. Generate the diffraction pattern corresponding to unit cell (B) using a red laser. What wavelength is your laser? (How well do you know this value? How will you account for uncertainty? How might you measure the wavelength?)

2. Measure the distance between "sample" and "detector." (Note how changing this distance affects the pattern.)

3. Record the diffraction pattern (trace spots) on a sheet of paper, taped to the wall and use the spacing of your "detector" spots applied ot the Fraunhaurfer equation to determine the spacing of dots on the slide (the spacing in the "sample"). Is this a reasonable value? Be sure to consider the number of significant figures you should report in your answer.

4. Repeat the process to measure the spading for unit Cell (D) based on the corresponding diffraction pattern. Again, pay careful attention to the significance of your answer.

5. What if you used a green laser? Qualitatively, how would the diffraction pattern (displayed on the paper) change?

6. What type of packing arrangement is observed for Unit Cell (H)?

1.1.6 Models: Visualizing Defects

Look at the models at the front of the room and identify which model represents
(a) an edge dislocation
(b) a screw dislocation
(c) FCC
(d) HCP
(e) BCC

1.1.7 DEMO: Atomic Packing illustrated with the "Atomix Raft"

The steel ball raft is a 2-dimensional model of atomic packing. Observe the following features in the demonstration: vacancies, hexagonal-packed grains, square-packed grains, grain and twin boundaries. Similar structures can be seen in (micrographs of) copper and brass samples.

1.1.8 DEMO: "Atomic Trampoline" Amorphous vs. Polycrystalline Metal

The effect of defects and micro-structure on macroscopic properties:
Stacking the tennis balls provides an example of the ordered crystalline arrays that most metals form. Is it possible for metals to have a completely disordered amorphous structure? Yes. But typically these are many-component alloys.
In this demonstration, one base is aluminum (polycrystalline FCC), the second base has a 1/8"thick disc of amorphous metal alloy (Zr0.42Ti.14Cu.12Ni.10Be.22) glued to the aluminum. As the ball bearing bounces on the aluminum base, it loses kinetic energy to friction, heat and sound generation, and the formation of crystalline defects - slip and dislocation formation. The amorphous metal plate has a highly disordered structure – it is not a regular crystalline array – and these defects do not form. Hence less energy is transferred from the ball to the base. This explains why these alloys are used in state-of-the-art golf clubs.

1.2 Laboratory 2: Electrical Conductivity & Optoelectronics

1.2.1 Prelab

Before lab, please READ CAREFULLY both the lab handout and Callister Chapter 18, pages 666-689 & 694-695. The relevant concepts are: Ohm's Law, the difference between resistance (R) and resistivity (ρ), how conductivity changes with temperature in metals vs. semiconductors, bandgaps, and p-n junctions (pp. 694-695). Come in with questions.

1.2.2 Exercise 1

Conductivity and Resistivity In this exercise, we will determine resistance as a function of temperature for different materials. This behavior distinguishes semiconductors from metals. If the geometry of the sample is known, you may convert measured **resistance (R)** to the corresponding property, **resistivity (ρ)**.
A. Metals

1. Measure the series resistance of your meter and leads: connect the two leads, then turn the meter to the most sensitive resistance scale, which has a maximum value of 200 Ω. (Ignore the ►+ , which indicates the diode test function.) Note: the meter will read 1 if R is greater than the maximum of scale chosen. In that case, turn the dial to the next higher scale.
2. Measure the resistance R of the steel wire at room temperature. Record R in the table below.
3. Measure R at 77K. With the leads attached, immerse the coil in liquid nitrogen (T = 77 K = -196 °C). Record the value in the table.
4. Calculate resistivity and conductivity. The wire has a length of 1 meter and a diameter of 0.203 mm. Remember to subtract the series resistance from the measured value of R!
5. Compare your σ to that for "Steel alloy A36" in Table B.9 of the text; some deviation from the "book value" is expected. Why may your result differ from the book's?

	Temperature	R (Ω)	$\rho(\Omega m)$
Series resistance	RT (~298K)		
Steel Wire	RT		
	LN_2(77K)		
Graphite Composite	RT		
	$LN_2(77K)$		

Table 1.1: Resistance

Semiconductor: Thermistor

1. Measure the resistance of the metal oxide thermistor at three or more temperatures between 0°C and ~+80°C. Use water baths to expose the thermistor to these temperatures; measure the water bath temperatures with the thermocouple.
2. Plot ln(R) vs. 1/T in Kelvin; the slope = Eg/2k. Note that R ∝ 1/σ ∝ exp(Eg/2kT). What is the effective bandgap, Eg (in eV)? Hint: make sure k is in appropriate units. Typically 0.5 < Eg < 3 eV for semiconductors..

Temperature (K)			
Resistance (Ω)			

Table 1.2: Thermistor Resistance

B. Graphite Composite of Unknown Composite

1. Measure the resistance of the graphite composite pencil "lead" at room temperature. Record in the table on pg. 1.
2. Measure the resistance of the graphite composite cooled in liquid nitrogen. Observe what happens immediately when it is dunked into the LN2. If the clips freeze, the measurement will be inaccurate. Record in the table on pg. 1.
3. Calculate the resistivity of the graphite composite. The diameter of the rod is 0.7 mm. Measure the length between the leads. Does this material behave like a metal or a semiconductor?

1.2.3 Exercise 2

Optoelectronics In this exercise, we will determine the relationship between the current (I) through and voltage across (V) a light emitting diode (LED) and a photoresistor. We will correlate the I-V behavior with light emission (LED) or absorption (photoresistor).

A. Light Emitting Diode

A light emitting diode is an example of a p-n junction. It is based on a solid solution of GaP and GaAs ($GaAs_xP_{1-x}$). The emission wavelength depends on the composition of this solution, as well as any dopants present.

1. Set up the measurement apparatus. Attach the voltmeter clips to the alligator clips on the board, so it is in parallel with the diode (p-n junction) that you will also attach there. (Note, red = +) Attach the ammeter in series, across the two bare wire connections on the bottom of the board. (Connect red to red, so you obtain the correct sign for I.)
2. Forward Bias (V_{bias} >0). Do a quick check that as you turn up the voltage (rotate the post on the potentiometer) the diode eventually emits light. Start with the ammeter on the most sensitive scale (200 μA). Take enough data points (about 10 total) to determine the I-V curve from 0V <V_{bias} <+3V. Remember to note the units for both V,I.
3. Reverse Bias (V_{bias} <0) Connect the circuit in reverse bias: take the black lead with the alligator clip and attach it to the opposite side of the potentiometer. Again, measure the I- V response. The magnitude should very small; be sure to use the maximum sensitivity of the ammeter. Please plot the I-V curve based on both forward and reverse bias on one graph (use graphing software). Note that the voltage where the current increases rapidly, and where the light intensity increases, corresponds (approximately) to the bandgap energy. How would the curve change for a green LED? An IR LED?

	Reverse Bias (-V,-I)				Forward Bias (+V,+I)							
V				0								
I				0								
Light?												

Table 1.3: LED I-V Data

B. Photoresistor

This photoresistor is made of CdS, a II/VI compound intrinsic semiconductor (not doped). The bandgap E_g = 2.59 eV, which means that conduction electron/holes provided by thermal energy (kT) are very rare. However, light with energy E(eV) = hc/λ = 1.24/ λ= Eg (λ $\leq$ 0.48 μm) will promote valence electrons to create conduction electron/hole pairs. Visible light ranges from 0.4 to 0.7 μm. More light will provide more charge carriers (ne = nh), and increasing conductivity σ. A CdS element is the heart of light meters, etc.

1. Measure R in the dark. Shield the resistor from ambient light with your hand or a cardboard box. Record the resistance.
2. Measure R in ambient light. Take photoresistor out of shadow and record the resistance.
3. Observe how R changes with distance from a light source. Move the photoresistor closer to the ceiling lights. What would happen if you illuminated your photoresistor with monochromatic 1μm wavelength light?

	Dark	Light
Resistance (Ω)		

Table 1.4: Photoresistor resistance

C. Light Emission vs. Composition
Light emission is possible when charge carriers (electrons and holes) recombine across the bandgap of some semiconductors. The four LED's are solid solutions of GaAs (E_g = 1.4eV) and GaP(E_g=2.3eV). Varying the composition of the solid solution allows us to control the bandgap, and thus vary the color of emission.

1. Observe light emission from the four LEDs. The LEDs have differing ratios of GaAs and GaP- this ratio varies monotonically along the strip. Given what you know, which LED has the highest ratio of GaAs:GaP?
2. Observe the blue LED. This LED is made of a different material. Why might one have to switch to a different material to get the observed light emission? What property would need to be different?

1.2.4 Lab Writeup

One writeup per person (not per group) is due at the beginning of the next lab session.
Please turn in a full set of data tables, including calculated numbers like resistivity, and answer the questions/do the tasks in italics. The questions are repeated below for your convenience.

1. Compare your result for the σ of Steel alloy A36 to that in the book. Why might your value differ?
2. For the Thermistor, Plot ln(R) vs. 1/T in Kelvin; the slope = Eg/2k. Note that R α 1/σ α exp(Eg/2kT).
3. What is the effective bandgap, Eg (in eV)? Hint: make sure k is in appropriate units. Typically 0.5 < Eg < 3 eV for semiconductors.
4. Does the graphite composite behave like a metal or a semiconductor?
5. Plot the LED I-V curve based on both forward and reverse bias on one graph (use graphing software).
6. How would the curve change for a green LED? An IR LED? Note that the voltage where the current increases rapidly, and where the light intensity increases, corresponds (approximately) to the bandgap energy.
7. What would happen if you illuminated your photoresistor with monochromatic 1μm wavelength light?
8. Given what you know, which LED has the highest ratio of GaAs:GaP?
9. Why might one have to switch to a different material to get the observed light emission? What property would need to be different?

1.3 Laboratory 3: Mechanical Properties of Metals

1.3.1 Objective

The objectives of this lab are to explore how crystal structure and defects affect mechanical properties and to understand how these properties are measured.

1.3.2 Outcomes

Upon completion of the laboratory, the student will be able to:

1. Describe the effects of work-hardening and annealing in metal samples.
2. Describe a ductile-to-brittle transition temperature.
3. Compute the elastic modulus, yield strength, ultimate tensile strength and strain at fracture from load-displacement (stress-strain) curves.

1.3.3 Experiment

Parts A and B are in Room 2068, while parts C and D are in Room 1034. Safety**

A. Work hardening and annealing in copper - Use caution with torch and hot tubing - see below.* You are supplied a length of copper tubing. Grasp it near one end and bend it slightly (through about 20-30). Now bend it back to straighten the piece. You should feel considerable additional resistance to straightening. Bend and straighten the same place once again, and perhaps a third time. Resistance will continue to grow. Now bend an "original" section near the other end - notice how easy it is to deform the as-received Cu. You are feeling the effect of "work hardening" or "strain hardening". Plastic deformation is accomplished by dislocation motion. The force required to achieve plastic deformation also creates many additional dislocations that were not present initially. Dislocations move easily through structurally regular crystals. The dislocations created during the first deformation constitute numerous irregularities, and hence make the material more resistant to further deformation, in this case the straightening step. Dislocations are high energy defects that can be removed by annealing the material at Ta ~ 0.5Tm. The melting temperature of Cu is 1360 K, so annealing effects are rapid above 680 K or 407 C. The propane torch has a flame temperature greater than 1000 C.
**Keep the propane torch on the table at all times. Holding the Cu tubing with tongs or pliers, heat the "cold worked" region until it is red hot. Cool the entire length of copper and the pliers in water!! Now test the copper for yield force; you should be back to the original "soft" condition.

B. Rolling of brass As for copper above, brass is capable of significant work hardening. The purpose of this exercise is to quantify the effect of work hardening that was observed qualitatively with the copper tubing. Resistance to yield is here measured by hardness after different extents of plastic deformation by rolling.
You are provided with a set of brass samples that have been rolled to different thicknesses. For each:

1. Measure the thickness of the specimen.
2. Measure its hardness. (Use the Rockwell B scale: 1/16" ball indenter, 100 kg load)

Sample	**A**	**B**	**C**	**D**	**E**
Thickness					
%CW					
Hardness, R_B					

Table 1.5: Rolling of brass

C. Charpy impact testing Impact tests (also called Charpy tests) induce fracture under conditions of high rates (impact) at an intentional flaw (the notch). This method readily demonstrates the effect of temperature upon the fracture energy of A36 steel (carbon content of not more than 0.30 wt%) and on an Al alloy (2024, 4.4 wt. % Cu). Note: steel is a BCC metal, while Aluminum is FCC. Complete the table below.

Sample	Temperature (C)	Impact Energy (ft-lb)
A36 steel, RT	25C	
A36 steel, LN	-196 C	
2024 Al, RT	25 C	
2024 Al, LN2	-196 C	

Table 1.6

D. Tensile test of an Al alloy You will have an opportunity to observe a tensile test on a sample of an Al alloy. Notice inhomogeneous deformation (necking) before failure. The plot below indicates the tensile test results for two 2024 Al samples, one as-received and a second after heat-treatment. 2024 is an aluminum alloy containing small amounts of Cu. Both test samples had the same dimensions: width (w) = 6.35mm, thickness (t) = 1.59mm, and gauge length (L_o) = 12.7mm.

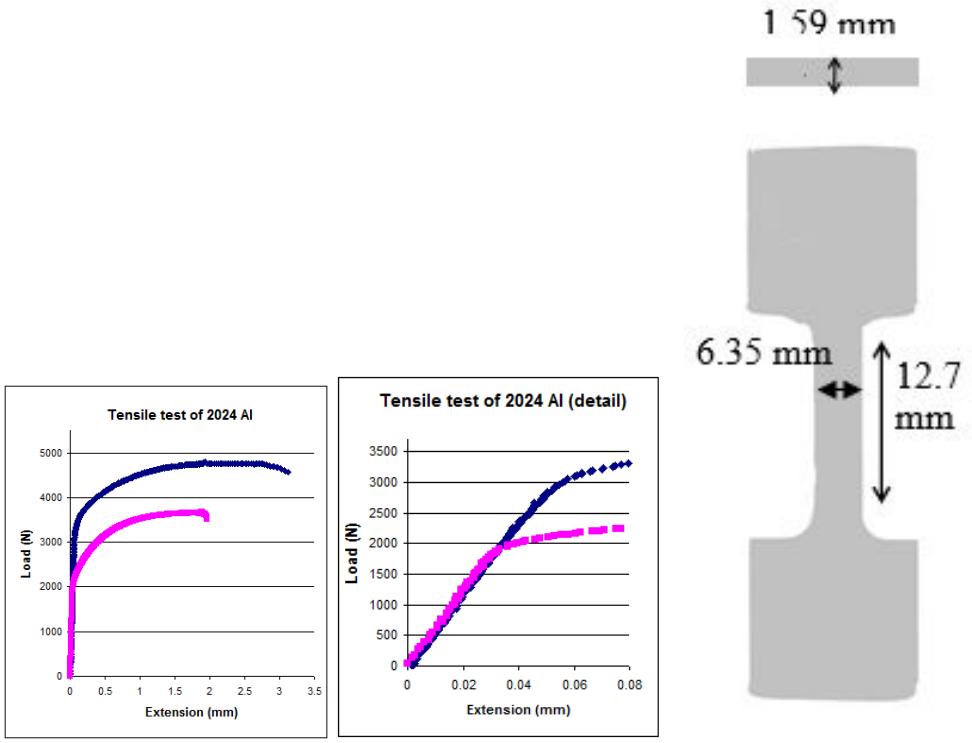


Figure 1.3: Tensile test of an Al alloy

1.3.4 Homework Problems

1. (Part A) Copper Work Hardening
   (a) Given the information in this handout, would you expect to be able to anneal work-hardened copper in boiling water?
   (b) Would you expect annealing temperatures for molybdenum to be higher or lower? Why?
2. (Part C) Charpy Impact testing
   (a) Why do the high and low temperature steel samples differ in appearance of the fracture region, while they appear similar for the aluminum? Note: this steel is BCC, while the aluminum is FCC.
3. (Part D) Tensile Testing: Convert load and displacement values from the graph to stress and strain to calculate the following. Show your work. Use engineering stress and strain (i.e. assume the cross-sectional area remains the same throughout the test).

(a) Elastic (Young's) modulus for 2024 Al.
(b) The 0.2% offset field strength for both the as-received sample and the annealed sample.
(c) The ultimate tensile strength for the both the as-received sample and the annealed sample.
(d) The amount of strain at fracture for both samples.

1.4 Laboratory 4: Phase Diagrams and Transformations

1.4.1 I. Eutectic salt solidification

Below is part of the phase diagram for the ammonium nitrate-sodium nitrate salt solution. The diagram shows a eutectic reaction at 21wt% $NaNO_3$.

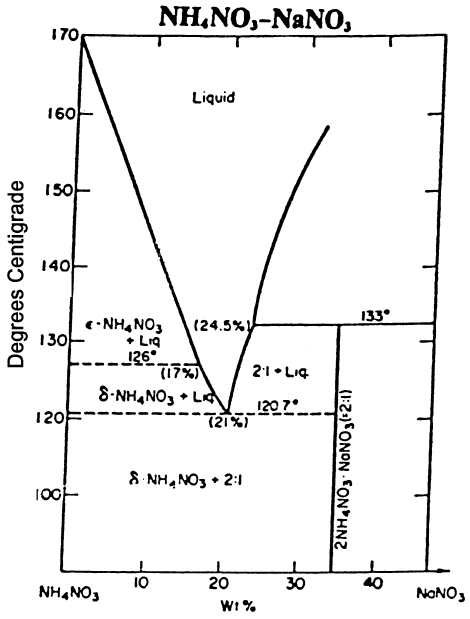


Figure 1.4: $NH_4NO_3 - NaNO_3$

NOTE: Phase diagrams are determined assuming equilibrium and thus predict equilibrium concentrations and phases. On heating, transitions occur close to equilibrium temperatures. This is not true for cooling (nucleation).

1.4.2 II. Steel rod

Many useful properties of iron and steel are determined by the crystal structure of the material, which in turn may be altered by heating and cooling. When steel is liquid, carbon dissolves readily in the iron. Even as a solid, carbon dissolves into the iron, forming a solid solution; however, the solubility of carbon in iron is limited by the crystal structure, based on the size of the interstitial sites available for carbon in the crystal structure. FCC gamma iron (austenite) can accommodate carbon at the edges of the crystal cell with little distortion. BCC alpha iron (ferrite) can accommodate carbon at the edge or face center of a cell; however the sites are smaller, and significant distortion of the crystal lattice occurs.
When slow cooled, the carbon has time to diffuse through the iron and form cementite Fe3C at dislocation sites; the steel remains ductile at room temperature.
When quenched, the carbon atoms are not given time to diffuse; they force iron atoms out of the usual places in the unit cell, resulting in a distortion of the crystal. This new structure is called martensite. It results in material that is hard– because dislocations cannot move easily through the distorted lattice, and brittle, for the same reason. If martensite is tempered, that is, heated to a temperature below the eutectoid transformation (usually this means between 250 – 600°C), the material will remain hard, but ductility and toughness will be enhanced. (Getting this right with the torch is a bit tricky – because the temperature is not well-controlled.)

1.4.3 III. Shape Memory wires (NiTi)

The wire on the table is NiTi (or NiTiNOL, named for the NiTi alloys developed at the Naval Ordinance Lab).

1. Start with the wire at room temperature.
2. Deform it gently, i.e. curl or bend it.
3. Warm it above the austenite transformation temperature, and observe how it recovers.

The low temperature phase (martensite) is deformable, but the alloy "remembers" its shape when it returns to the higher temperature (austenite) phase at T < 100°C.

1.4.4 IV. Bi-Cd binary eutectic alloy

We will examine metallographic sections of different compositions in the Bi-Cd system, which exhibits a eutectic at 39.7 wt.% Cd. Note that the Bi-Cd system has negligible solid solubility. Compare this to the binary eutectic Pb-Sn system which does have solid solubility, shown in Fig.5.7 of your text.

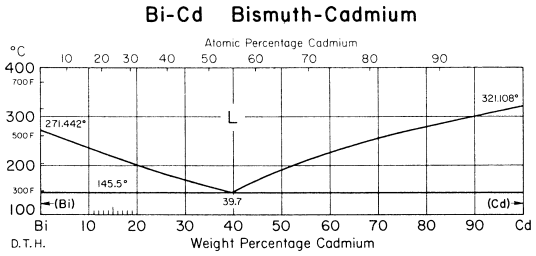


Figure 1.5: Bi-Cd Bismuth-Cadmium

HOMEWORK PROBLEM

1. **Sketch** each of the three microstructures corresponding to samples A-C.
2. **Label the composition of the alloy you were shown.**
3. **Label the phases present** (indicate which phase is etched dark, which phase is etched light).
4. For each sample, use the **lever rule** to predict amounts (wt%) of **primary phase vs. eutectic.**

1.4.5 V. Steel wire/carburized steel samples

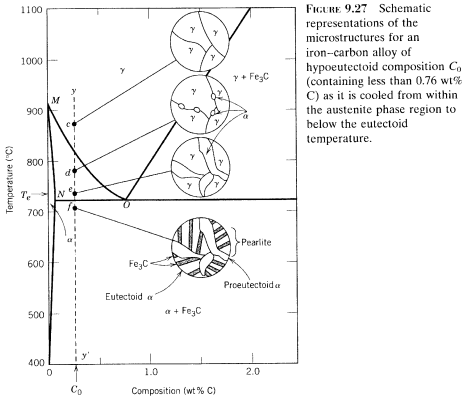


FIGURE 9.27 Schematic representations of the microstructures for an iron–carbon alloy of hypoeutectoid composition C_0 (containing less than 0.76 wt% C) as it is cooled from within the austenite phase region to below the eutectoid temperature.

Figure 1.6: Microstructures of hypoeutectoid iron-carbon alloy

A. _____wt% Cd
Amount primary phase: _____
Amount eutectic: _____
B. _____wt%Cd
Amount primary phase: _____
Amount eutectic: _____
C. _____wt%Cd
Amount primary phase: _____
Amount eutectic: _____

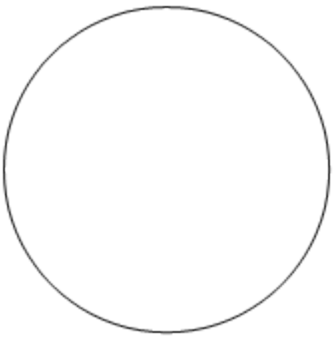


Figure 1.7: Part A

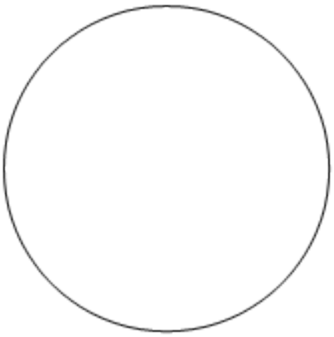


Figure 1.8: Part B

Phase Transformation When piano (~0.8% carbon steel) wire is heated (in this case by running a current through it) it transforms from BCC alpha iron to FCC gamma iron. The reverse is true on cooling. Why does the wire sag as it is cooled?

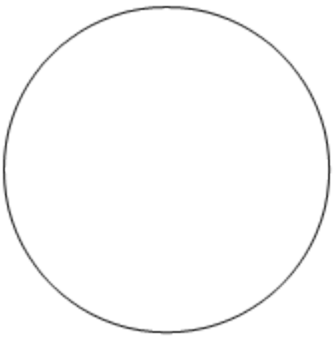
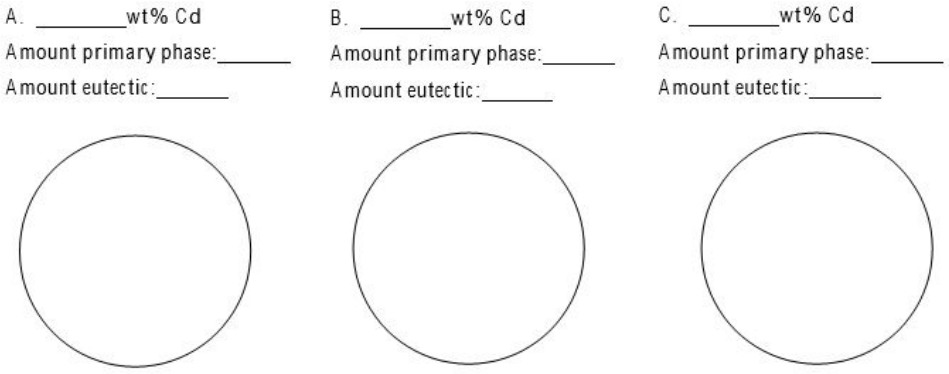


Figure 1.9: Part C

Name: _____

1.4.6 Homework Problems

1. Identify (circle) the eutectic phase transformation on the NH4NO3-NaNO3 phase diagram and indicate the temperature/composition at which it occurs.
2. Explain why steel quenched from high temperature which forms a martensite phase exhibits higher hardness and strength then slow cooled steel which forms a mixture of ferrite/cementite.
3.
   (a) Sketch each of the three microstructures you observed, corresponding to samples A-C.
   (b) Label the composition of the alloy you were shown.
   (c) Label the phases present, which is etched dark and which is etched light?
   (d) For each of the samples use the lever rules to predict the amounts (wt %) of the primary phase vs. the eutectic.
   (e) Are the samples hypo- or hyper eutectic?



4. When piano (~0.8% carbon steel) wire is heated (in this case by running a current through it) it transforms from BCC alpha iron to FCC gamma iron. The reverse is true on cooling. Why does the wire sag as it cools?

1.5 Laboratory 5: Mechanical Behavior of Ceramics and Polymers

1.5.1 Polymer 1. Crystallization of Isotatic Polypropylene

When cooled from the melt, this isomer, with R = CH3 groups all on the same side of the chain, crystallizes into spherulites. Note that crystallization occurs below Tm = 165 oC (undercooling).

1.5.2 Ceramics 1. Three-Point bend test (Bonus)

It is difficult to perform conventional tensile tests on most ceramic materials; they are brittle and break in the grips of the test apparatus. It is easier to test these materials in the bending mode, as illustrated below. Note that when loaded, the top of the material is in compression and the bottom is in tension.

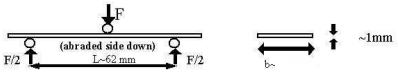


Figure 1.10: Three-point bend test

A. Measure the force F (weight of volume of water plus container) to fracture a pristine glass microscope slide.
B. Abrade a second microscope slide with 180 grit (a = 80 μm) paper. Abrasion should be in center of slide, and in the "b" direction (above). BE SURE YOUR HANDS AND SLIDE ARE DRY. Measure fracture force for slide with flaw dimension a; flaw on bottom (tensile side).
C. If time permits, repeat part b with abraded side up (under compression). This should be close to your answer in part a.
D. Glassblowers often use the trick of scribing a mark, then wetting it, to get the glass to break there. If you notice the maximum load is less when the abraded region of the slide is wet, you have observed stress corrosion cracking. The Si-O-Si network in the glass is broken by formation of Si-O-H bonds at the surface, and the material effectively "unzips."

Container weight = 118 g (Add to obtain total load)	Vol (ml)	Total load (g)
A. Pristine slide: vol. (ml)		
B. Flawed slide: vol. (ml)		
C. Compression: vol. (ml)		
D. Repeat b, but WET abraded region		

Table 1.7: Bend test results

1.5.3 Lab Homework

1. Calculate fracture force for experiments A and B.
2. Use equation 12.3a in Callister (see addendum) to calculate the "flexural strength" obtained with this three point bend experiment. Sample dimensions are d = 1 mm, b = 25 mm, and L = 62 mm between support points. Your answer should be in MPa, and you will need to convert Ff from g to N.
3. Use your result from experiment B and the following equation to calculate K_{IC} for the microscope glass material: $K_{IC} = Y\sigma_f\sqrt{\pi a}$, where a is the flaw size $(= 80\mu)$ and Y, dependent on sample & flaw geometry = 1. Compare your K_{IC} to the range predicted for ceramic materials in Table 8.1.
4. Use your K_{IC} and the result from experiment A to calculate flaw size a in the pristine glass slide (the slide you did not intentionally scratch).

1.5.4 Ceramics 2. Softening glass

Hold the glass rod in the flame of the propane torch. When the rod becomes hot enough, that is when the temperature exceeds the "glass transition temperature," T_g, the glass will flow and you can easily bend it. **BE CAREFUL NOT TO BURN YOUR FINGERS. Don't touch the HOT glass.**

1.5.5 Polymers 2: Nylon Polymerization

Nylon is formed by the reaction of two liquid monomers, here dissolved in immiscible liquids. These can react with each other, but not with themselves:

0.5 M hexamethylenediamine	$H_2N-(CH_2)_6-NH_2$	in 0.5 M NaOH
0.2 M sebacoyl chloride	$Cl-C-(CH_2)-C-Cl$	in hexane

The reaction occurs in a thin layer at the interface between the two liquids. By slowly pulling the film in this region, a long, continuous filament can be pulled from the beaker. NOTE THE BYPRODUCT IS HCl – so use gloves.

1.5.6 Polymers 3. Plastic Deformation of Polyethylene

1. With thumbs separated by about 0.5 cm, stretch the strip with a slow, steady force. The polyethylene will yield and deform plastically. Here the sliding of planes of atoms (crystals) is accompanied by uncoiling sections of the chain-like polymer molecules (amorphous regions).
2. You will notice that a thin portion or "neck" is formed and propagates during stretching. The plastically deformed polymer is anisotropic with different properties in different directions. With many chains or portions of chains in the stretch direction, this chain orientation means that covalent bonds are deformed or break when the film is further loaded in the stretch direction. Relatively weak van der Waals bonds control properties in the transverse directions.
3. Grasp a section of the necked polymer and pull in the stretch direction. The oriented film is stiff and strong (no easy yield or failure) in this direction.
4. Now grasp the same necked portion in the transverse direction and pull slowly. The materials will yield and "split", as weak secondary bonds are dominant in this direction.

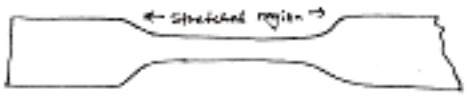


Figure 1.11: Necked polymer

Conclusion: You can "feel" the difference between strong covalent and weak van der Waals bonds when the chains are oriented.

1.5.7 Polymers 4. Plastic Deformation of Polyethylene

You are all familiar with "high density" (fc = 0.7) and "low density" (fc = 0.4) polyethylene. As PE at room temperature is a crystal/rubber combination, the elastic modulus of HDPE is larger. As there are more crystals, the yield strength σ_y of HDPE is larger as well.

- With increased yield strength (more crystallinity) comes reduced ductility.
- Somewhat similar to BCC metals, the yield strength of polymers is temperature dependent.

1. Test with your fingers the two flexure samples. Which is high density polyethylene? Which is low density polyethylene?
2. Cool the tensile samples in liquid nitrogen and repeat. A substantial change in ductility (and likewise an increase in yield stress) will be noticed.

1.5.8 Polymers 5. Rubber nails (Glass Transition Temperature (a))

Cool a rubber “nail” by holding in liquid nitrogen (with tongs!!). The material is now a glass. Use the glassy nail to join two pieces of (soft) balsa wood. Hints: grasp the nail with the tongs; align the nail with the grain of the wood; use repeated soft taps to avoid shattering your “nail.” Have everyone in your group contribute a nail.

1.5.9 Polymers 6. Rubber band (Glass Transition Temperature (b))

At room temperature, rubber is above its glass transition temperatures $T_g = -70C$, but at liquid nitrogen temperature, rubber is below T_g.
A stretched rubber band is cooled with liquid nitrogen. The shape is conserved as the macro-molecules are "frozen in", and the band is very brittle. As the temperature increases back toward room temperature and T_g is reached, the band recovers its natural undeformed shape.

1.5.10 Polymers 7. Happy/ Sad balls (Glass Transition Temperature (c))

Two polymer balls of different compositions are provided, along with hot water and LN2. DO NOT try to FREEZE the balls completely in LN2. Cool them gently, while moving around in a small amount of liquid, for 10-20 seconds. Otherwise, the balls will fracture.

1. Drop both balls onto a table at room temperature. Ball 1 is above its Tg and thus bounces back elastically (rubber elasticity). Ball 2 is near its T_g and dissipates all the kinetic energy without bouncing.
2. Heat Ball 2 and repeat the experiment. The ball is now above its T_g and bounces back elastically (rubber elasticity)
3. Cool Ball 2 and repeat the experiment. The ball is now below its T_g and bounces elastically (glass elasticity).
4. Cool Ball 1 and repeat the experiment. The ball will be glass elastic below T_g , "dead" near T_g and rubber elastic above T_g . Very cool!

What practical items are made from such rubbery elastomers? Rubber o-rings, used to seal surfaces – except when they get cold, below T_g, they become stiff and don't conform – hence resulting in leaks. (The Challenger Disaster was a consequence.)

E. Viscoelasticity (Stain rate dependence) Silly Putty is a polymer above its glass transition temperature which is thus viscoelastic at room temp. At high strain rates, it behaves elastically, while at low strain-rates, it is viscous.

1.5.11 Ceramics Demonstrations

III. Thermal Shock If you have ever poured hot water into a cold glass, or vice versa, you may have experienced the effect of the material's low resistance to thermal shock. ($TSR = \sigma_c k/E\alpha_1$). This, in turn, is due to the large thermal expansion coefficient and low thermal conductivity typical of most ceramic materials. These properties may be varied by changing the composition of the glass, as indicated below. (Of course, the cost changes, too.)

	Comp. (wt%)		Thermal expansion coefficient (linear), $\alpha_1(K^{-1})$	Thermal Shock Resistance
GlassSoda-lime (windows,bottles)Borosilicate (Pyrex)	70% SiO_2	Thermal Conductivity, k ($Wm^{-1}K^{-1}$) 11	8.5×10^{-6}	84
	10% CaO			
	15% Na_2O			
	80% SiO_2		4.0×10^{-6}	280
	15% B_2O_3			
	5% Na_2O			

Table 1.8: Thermal properties of glass

IV. Toughening plates As you will see in the laboratory exercise, ceramics have a high probability of fracturing when a tensile force is applied. This is due to the crack propagation through the material. One way to strengthen ceramics, then, is to introduce compressive forces in the outer portion of the piece, where cracks are likely to initiate. The compressive forces act to close surface cracks, rather than cause them to propagate. An example is tempered glass, which is formed by quickly cooling from above its glass transition temperature: the outside cools faster than the inside, and is put into compression by contraction of the slower cooling core. This technique is used in making car windshields.

A second method for producing ceramics with an outer layer in compression is to use layered materials. In this case, the outside layer has a lower thermal expansion coefficient than the inside, and again is put into compression as the piece forms.

V. Tough ceramic ball The material is Zr_2/CaO "alloy." At room temperature it contains ~10% of nonequilibrium tetragonal phase in a matrix of equilibrium monoclinic phase. Large local tensile stress near a crack tip triggers a tet→ mono transition with a 4% volume increase. The larger volume of new monoclinic crystals creates compressive stress that "pinches" the crack shut. K_{IC} is increased to about 10 MPa $m^{1/2}$.

VI. Other Ceramics Ceramics are refractrory, That is, they maintain structural integrity at high temperature, and may provide thermal and electrical insulation. Some are used as capacitors or resistors on computer motherboards. Some are piezoelectric, that is they may be used to transform mechanical energy into a voltage. Others are magnetic (Hard B $BaFe_{12}O_{19}$, $SrFe_{12}O_{19}$, used for small electric motor field magnets, Refrigerator door seals and posting magnets; Soft B $NiFe_2O_4$, $MnFe_2O_4$, transformers, inductors) or superconducting.

1.5.12 Addendum:

For a sample with a rectangular cross-section tested on a three-point bend apparatus, the tensile breaking stress, termed the flexural strength, σ_{fs}, is:

$$\sigma_{fs} = \frac{3F_bL}{2bd^2}$$

Here Fb is the load at break, L is the distance between supports, b is the width and d is the thickness of the rectangular cross-section.[1]

1. Callister, William D. Jr. Materials Science and Engineering: An Introduction, 6th Edition, ISBN 0-471-13576-3, John Wiley and Sons, Inc., New York, 2003.

1.6 Discovery Lab 1: Cholesteryl Ester Liquid Crystals

1.6.1 Pre-lab

Important: Be sure that you are wearing long pants and closed -toed shoes for this lab. Lab coats, safety glasses, and gloves will be provided.

Goal To investigate the relationship between temperature, composition, and reflected/trans-mitted colors of a cholesteryl ester liquid crystal.

Background Liquid crystals are organic compounds that are in a state between liquid and solid forms. They are viscous, jelly-like materials that resemble liquids in certain respects (viscosity) and crystals in other properties (light scattering and reflection). Liquid crystals must be geometrically highly anisotropic, as opposed to an isotropic liquid. Such properties, however, can be altered through thermal action (heat) or by the influence of a solvent.
Common liquid crystals are composed of derivatives of cholesterol, C27H46O. Materials that will be used in this lab for making the cholesteryl ester liquid crystals are cholesteryl benzoate, C34H50O2, cholesteryl pelargonate, C36H62O2, and cholesteryl oleyl carbonate, C46H80O3.
Cholesteric-phase liquid crystals contain molecules aligned in layers rotated with respect to one another, resembling helix structures (Figure 1.12). The distance between layers that have the same orientation is known as the pitch, p. At high temperature, the rotation angle from one layer to the next increases, hence the pitch is smaller. A color will be reflected when the pitch is approximately equal to the color's wavelength of light.

Figure 1.12: Cholesteric-phase liquid crystals (Figure courtesy of George Lindsey)

Pre-lab Questions

1. What are the safety concerns with each of the chemicals? What is the proper way to handle the chemical waste? Please read the Materials Safety Data Sheets (MSDS) for each chemical used in this lab beforehand.

2. Liquid crystals have anisotropic structures. In your own words, explain the term "anisotropy" and how will this affect the material properties of liquid crystals.

3. If the reflected color is green and changes to red, chemically/physically explain what has happened to the liquid crystal. Be specific. Does this correspond to a temperature increase or decrease? Why?

1.6.2 Lab Part 1

Objective Build a cholesteryl ester liquid crystals thermometer that covers the temperature range between 17-40C.

Laboratory Procedure

1. Based on Table 1, select 5 ratio combinations of the cholesteryl oleyl carbonate, the cholesteryl pelargonate, and the cholesteryl benzoate.
2. Measure out the alloted amounts of each of the cholesterol components and place them in a 10 ml vial.
3. Melt the solid in a smaple vial using a hair drier or heat gun. (Make sure the vial lid is off during the melting process because it can ruin the seal on the vial.)
4. Cut one piece of clear packing tape and lay the sticky side up on a table.
5. Take the vial of teh cholesteryl ester mixture, open it and transfer a small amount of the gel paste to the tape with a spatula. Spread the gel uniformly in the middle of the tape. Be sure to leave a centimeter clearance from the edge of the tape. Take the other piece of packing tape and place the sticky sides together. Smooth the paste evenly so there are no air bubbles.
6. Submerge the liquid crystal sandwich in a beaker of water and gently heat the water with a hot plate.
7. Measure and record the temperature at which the liquid crystals first change color. What color do you see first?
8. Measure and record the temperature at which the liquid crystals become colorless. What color do you see last?
9. Cool the sample by removing it from the water. Repeat steps 2-3 as necessary to list the order of colors observed.

Cholesteryl oleyl carbonate	Cholesteryl pelargonate	Cholesteryl benzoate	Transition range, *C*
0.65 g	0.25 g	0.10 g	
0.70 g	0.10 g	0.20 g	
0.45 g	0.45 g	0.10 g	
0.43 g	0.47 g	0.10 g	
0.44 g	0.46 g	0.10 g	
0.42 g	0.48 g	0.10 g	
0.40 g	0.50 g	0.10 g	
0.38 g	0.52 g	0.10 g	
0.36 g	0.54 g	0.10 g	
0.34 g	0.56 g	0.10 g	
0.32 g	0.58 g	0.10 g	
0.30 g	0.60 g	0.10 g	

Table 1.9: Cholesteryl liquid crystal mixtures and their transition temperatures. Different compositions give different color patterns over different temperature ranges.

Lab Writeup

1. Upon heating, when the liquid crystal sandwich undergoes a color change, what color do you see first (i.e., at coolest temperature)? What color do you see last (warmest)? What is the order of colors you see going from coolest to warmest? Explain what happens to the liquid crystal during this transition at molecular level.
2. What is the relationship between the mass of cholesteryl oleyl carbonate in the mixture and the transition temperature? What do you think might be the reason for this?

3. As the liquid crystal changes temperature it reflects different colors. Is the color change better observed over a black background or a white background? Why?
4. What are some of the applications you can think of with the temperature sensitive cholesteryl ester liquid crystals?
5. What happens if you heat 0.30 g of Cholesteryl oleyl carbonate, Cholesteryl pelargonate, Cholesteryl benzoate by themselves, will you observe the same behavior?

References

1. Lisensky, G (2007). Preparation of cholesteryl ester liquid crystals. Interdisciplinary Education Group, Retrieved May 26, 2008.
2. Lisensky, G, & Boatman, E (2005). Colors in liquid crystals. Journal of Chemical Education, 82, 1360A, Retrieved June 4, 2008.

1.6.3 Lab Part 2

Liquid crystal display (http://www.microscopyu.com/articles/polarized/polarizedlightintro.html)
One of the most common and practical applications of polarization is the liquid crystal display (LCD) used in numerous devices including wristwatches, computer screens, timers, clocks, and a host of others. These display systems are based upon the interaction of rod-like liquid crystalline molecules with an electric field and polarized light waves. The liquid crystalline phase exists in a ground state that is termed cholesteric, in which the molecules are oriented in layers, and each successive layer is slightly twisted to form a spiral pattern (Figure 1.13). When polarized light waves interact with the liquid crystalline phase the wave is "twisted" by an angle of approximately 90 degrees with respect to the incident wave. The exact magnitude of this angle is a function of the helical pitch of the cholesteric liquid crystalline phase, which is dependent upon the chemical composition of the molecules (it can be fine-tuned by small changes to the molecular structure).

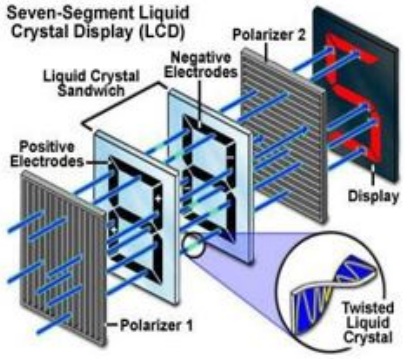


Figure 1.13: Liquid crystal display

An excellent example of the basic application of liquid crystals to display devices can be found in the seven- segment liquid crystal numerical display (illustrated in Figure 1.13). Here, the liquid crystalline phase is sandwiched between two glass plates that have electrodes attached, similar to those depicted in the illustration. In Figure 2, the glass plates are configured with seven black electrodes that can be individually charged (these electrodes are transparent to light in real devices). Light passing through polarizer 1 is polarized in the vertical direction and, when no current is applied to the electrodes, the liquid crystalline phase induces a 90 degree "twist" of the light that enables it to pass through polarizer 2, which is polarized horizontally and is oriented perpendicular to polarizer 1. This light can then form one of the seven segments on the display.
When current is applied to the electrodes, the liquid crystalline phase aligns with the current and loses the cholesteric spiral pattern. Light passing through a charged electrode is not twisted and is blocked by polarizer 2. By coordinating the voltage on the seven positive and negative electrodes, the display is capable of rendering the numbers 0 through 9. In this example the upper right and lower left electrodes are charged and block light passing through them, allowing formation of the number "2" by the display device (seen reversed in the figure).

Procedure Liquid Crystal Transition Temperature Measurement by Optical Characterization

1. Heat the liquid crystal mixtures prepared in last lab
2. Prepare sample slides for the 5 compositions of liquid crystals (LC) prepared in Lab 1
   (a) It will be helpful if you heat up your slide so that the LC mixture is still liquid when you carefully put a coverslip on top of it
   (b) While placing coverslip try to avoid trapping as many bubble as possible in your sample
3. TA/instructor will provide instructions on how to operate the optical microscope and record your data
   (a) The polarized light microscope is designed to observe and photograph specimens that are visible primarily due to their optically anisotropic character. In order to accomplish this task, the microscope must be equipped with both a polarizer, positioned in the light path somewhere before the specimen, and an analyzer (a second polarizer; see Figure 1.14), placed in the optical pathway between the objective rear

aperture and the observation tubes or camera port. It is a contrast-enhancing technique exploits the optical properties specific to anisotropy and reveals detailed information concerning the structure and composition of materials that are invaluable for identification and diagnostic purposes.

(b) The liquid crystalline phase exists in a ground state that is termed cholesteric where the molecules are oriented in layers where each successive layer is slightly twisted to form a spiral pattern. When polarized light waves interact with the liquid crystalline phase the wave is "twisted" by an angle of approximately 90 degrees with respect to the incident wave. This angle is a function of the helical pitch of the cholesteric liquid crystalline phase, which is dependent upon the chemical composition of the molecules (it can be fine-tuned by small changes to the molecules).

(c) References: http://www.microscopyu.com/articles/polarized/polarizedintro.html http://www.olympusmicro.com

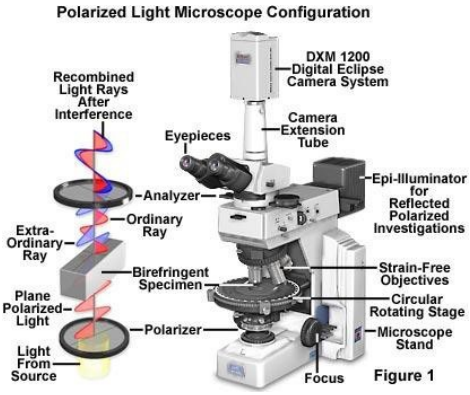


Figure 1.14: Polarized light microscope

4. You will generate movies of transitions observed in the liquid crystals with changing temperatures.

5. For the report please record observations of transition in structures from the different samples.

Part II
314: Thermodynamics of Materials

2 Introduction

The content and notation of these notes is based on the excellent book, "Thermodynamics in Materials Science", by Robert DeHoff [16]. The notes provided here are a complement to this text, and are provided so that the core curriculum of the undergradaute Materials Science and Engineering program can be found in one place (http://msecore.northwestern.edu/). These notes are designed to be self-contained, but are more compact than a full text book. Readers are referred to DeHoff's book for a more complete treatment.

3 Big Ideas of Thermodynamics

The following statements sum up some of foundational ideas of thermodynamics:

- **Energy** is conserved
- **Entropy** is produced
- **Temperature** has an absolute reference at 0 - at which all substances have the same entropy

3.1 Classification of Systems

Systems can be classified by various different properties such as constituent phases and chemical components, as described below.

- Chemical components
  - Unary system has one chemical component
  - Multicomponent system has more than one chemical component

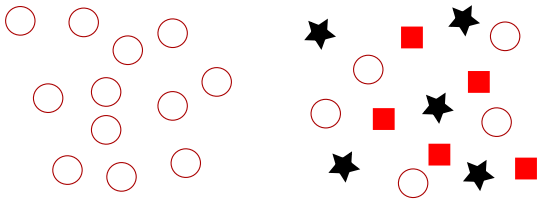


Figure 3.1: Unary system (left) and multicomponent system (right).

- Phases
  - Homogeneous system has one phase
  - Heterogeneous system has more than one phase

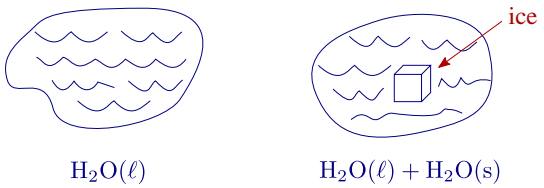


Figure 3.2: Homogeneous system (left) and heterogeneous system (right).

- Energy and mass transfer: closed and open systems:
  - An isolated system cannot change energy or matter with its surroundings.
  - Closed system can exchange energy but not matter with its surroundings.
  - Open system can exchange both energy and matter with its surroundings.

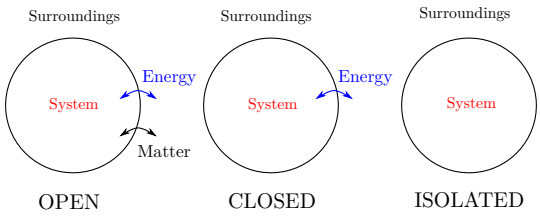


Figure 3.3: Open, closed and isolated thermodynamic systems.

- Reactivity
  - Non-reacting system contains components that do not chemically react with each other
  - Reacting system contains components that chemically react with each other

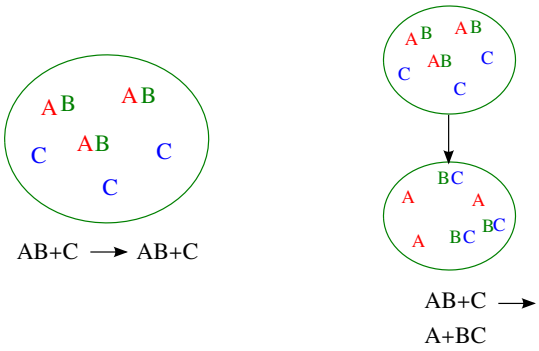


Figure 3.4: Non-reacting system (left) and reacting system (right).

Another classification is between simple and complex systems. For instance, complex systems may involve surfaces or fields that lead to energy exchange.

3.2 Thermodynamic State Variables

Thermodynamic state variables are properties of a system that depend only on current condition, not history. Temperature, T, is an obvious example, since we know intuitively that the temperature in a room has the same meaning to us, now matter what the previous temperature history actually was. Other state variables include the pressure, P, and the system volume, V. For example, two states A and A', each with their own characteristic values of P, V and T:

$$\begin{matrix} A \\ P, V, T \end{matrix} \quad \rightarrow \quad \begin{matrix} A' \\ P', V', T' \end{matrix} \tag{3.1}$$

The values of the pressure difference$\Delta P = P' - P$ = is independent of the path that the system takes to move from state A to A'. The same is true for ΔV and ΔT.

3.3 Process Variables

Work and heat are examples of process variables.
Work (W) - Done on the system. Note that our sign convention here is that work done ON the system is positive.
Heat (Q)- Absorbed or emitted. The sign convention in this case is that heat absorbed by the system is positive.
Their values for a process depend on the path.

3.3.1 Work Done by an External Force

The work done by an external force is the force multiplied by distance over which the force acts. In other words, the force only does work if the body it is pushing against actually moves in the direction of the applied force. In differential form we have:

$$\delta W = \vec{F} \cdot d\vec{x} \tag{3.2}$$

Note that $\vec{F}$ and $\vec{x}$ are vector quantities, but W is a scalar quantity. As a reminder, the dot product appearing in Eq. 3.2 is defined as follows:

$$\vec{A} \cdot \vec{B} = \left|\vec{A}\right| \left|\vec{B}\right| \cos\theta \tag{3.3}$$

where θ is the angle between the two vectors, illustrated in Figure 3.5.

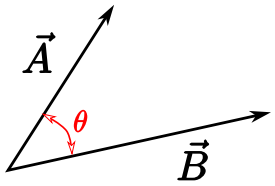


Figure 3.5: Definition of the angle θ between two vectors.

The total work is obtained by integrating δW over the path of the applied force:

$$W = \underbrace{\oint}_{path} \vec{F} \cdot d\vec{x} \tag{3.4}$$

For example, when we lift a from the floor to the table as illustrated in Figure 3.6 we increase the gravitational potential energy, U, of the barbell by an amount mgh, where:

- m = the mass of the barbell
- g = the gravitational acceleration (9.8 m/s^2)
- h_f = the final resting height of the barbell above the floor

Because the potential energy only depends only on the final resting location of the barbell, it is a state variable. The work we expend in putting the barbell there depends on the path we took to get it there, however. As a result this work is NOT a state function. Suppose, for example that we lift the barbell to a maximum height h_{max} and then drop it onto the table. The work, W, we expended is equal to mgh_{max}, which is greater than the increase in potential energy, ΔU. The difference between between W and ΔU is the irreversible work and corresponds to energy dissipated as heat, permanent deformation of the table, etc. We'll return to the distinction between reversible and irreversible work later.

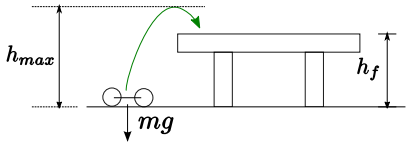


Figure 3.6: Work done on a barbell.

3.3.2 Work Done by an Applied Pressure

Another important example of work is the work done by an external force in changing the volume of the system that is under a state of hydrostatic pressure. The easiest and most important example to describe is a gas that is being compressed by a piston as shown in Figure 3.7. Equation 3.2 still applies in this case, but the force in this case is given by the product of the area and pressure:

$$\vec{F} = P\vec{A} \tag{3.5}$$

It may seem strange to write the area as a vector, but it is just reminding us that the pressure acts perpendicular to the surface, so the direction of $\vec{A}$ is perpendicular to the surface of the piston. Combining Eq. 3.5 with Eq. 3.2 gives:

$$\delta W = P\vec{A} \cdot d\vec{x} \tag{3.6}$$

The decrease in volume of the gas is given by $\vec{A} \cdot d\vec{x}$, so we can replace $\vec{A} \cdot d\vec{x}$ with $-dV$ to obtain:

$$\delta W = -PdV \tag{3.7}$$

This result is an important one that we will use quite a bit later on.

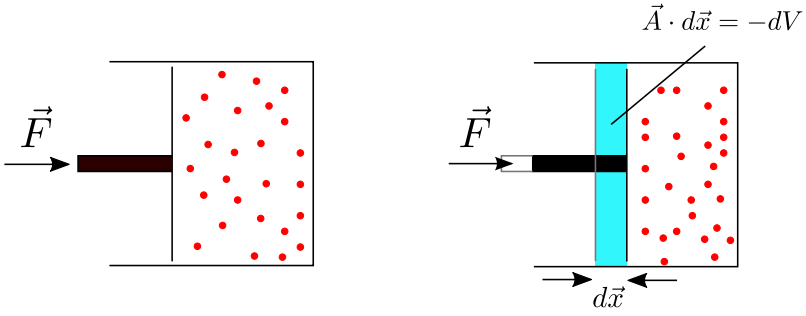


Figure 3.7: Work done by force on a piston.

3.4 Extensive and Intensive Properties

Extensive properties depend on size or extent of the system. Examples include the volume (V), mass (m) internal energy (U) and the entropy (S). Intensive properties are defined at a point, with examples including the temperature (T) and the pressure (P) and the mass density, ρ.
Extensive properties can be defined as integrals of intensive properties. For example the mass is obtained by integration of the mass density over the total volume:

$$m = \int_V \rho\left(\vec{r}\right) d\vec{r} \tag{3.8}$$

Also, intensive properties can be obtained from combinations of extensive properties. For example C_k, the molar concentration of component k, is the ratio of two extensive quantities: n_k the total number of moles of component k and the volume:

$$C_k \equiv \lim_{V \to 0} \frac{n_k}{V} \tag{3.9}$$

This approach works for the density as well, with the mass density, ρ, given by the ratio of the sample mass, m, to the sample volume, V:

$$\rho \equiv \lim_{V \to 0} \frac{m}{V} \tag{3.10}$$

4 The Laws of Thermodynamics

Outcomes for this section:

1. State the first, second and third laws of thermodynamics.
2. Write a combined statement of the first and second laws in differential form.
3. Given sufficient information about how a process is carried out, describe whether entropy is produced or transferred, whether or not work is done on or by the system, and whether heat is absorbed or released.
4. Quantitatively relate differentials involving heat transfer/production, entropy transfer/production, and work.

4.1 1st Law of Thermodynamics

The first law of thermodynamics is that energy is conserved. This means that the increase in the internal energy of the system is equal to the sum of the heat flow and the work done on the system and the heat flow into the system. In mathematical terms:

$$\Delta U = Q + W + W' \tag{4.1}$$

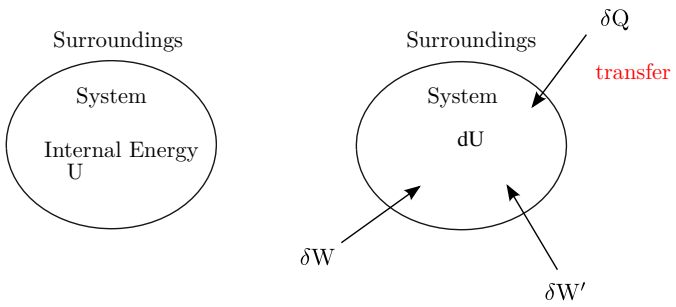


Figure 4.1: Conservation of energy

where:

- ΔU = Internal energy change of the system
- Q =Heat flow into the system
- W =Mechanical work done by surroundings on the system (P_{ext})
- qW' =All other work done on the system

Challenge: How do we evaluate Q, W?

$\underbrace{dU}_{\text{state function}} = \underbrace{\delta Q}_{\text{process variable}} + \delta W + \delta W'$

Integration to find dU does not seem advised - process variables are path dependent, destination U is path independent. δ : not an exact differential

4.2 2nd Law of Thermodynamics

The second law of thermodynamics states that entropy can be transferred or produced, but not destroyed:

$$\Delta S_{uni} \geq 0 \tag{4.2}$$

Here ΔS_{uni} is the change in entropy for the universe, which includes the system of interest and its surroundings:

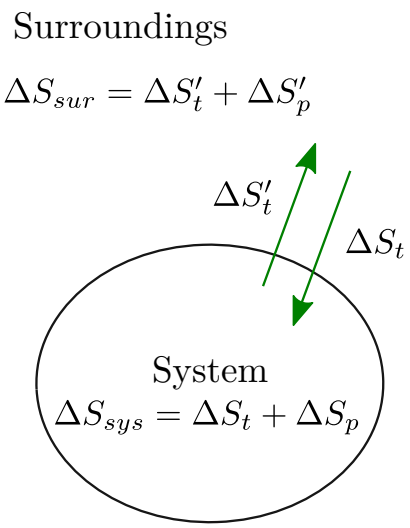


Figure 4.2: Transfer of entropy

$$= [\cancel{\Delta S'_t} + \Delta S'_p] + [\cancel{\Delta S_t} + \Delta S_p]$$

$$\Delta S_{uni} = \Delta S'_p + \Delta S_p \geq 0$$

Entropy change can only be positive in the universe.
For infinitesimal changes, $dS_{sys} = dS_t + dS_p$ where dS_{sys}can be negative.

4.3 Reversible and Irreversible Processes

Irreversible processes produce entropy, whereas reversible processes (an idealization) do not.
Entropy cannot be destroyed.
Constantly increase in entropy defines the "arrow of time." We recognize this intuitively when watching video in reverse.
Rate of Entropy Production ∝Separation from Equilibrium

4.3.1 Reversible Process

Processes are carried out very slowly (very near equilibrium) approach this ideal.
The imaginary situation of a perfectly reversible process is extremely useful for calculations.

- No entropy is produced
- No permanent changes in the universe

To compute differences in state functions, one can chose the simplest (reversible) path.

Heat Flow for Reversible Processes For a very slow process, $\Delta S_p \to 0$. Consider heat absorbed by system δQ_{rev}during that very small step.
$\oint \frac{\delta Q_{rev}}{T} = 0$ for any reversible, cyclic path. Therefore, $\frac{\delta Q_{rev}}{T}$ is a state function. We define $\delta Q_{rev} = TdS$, and $Q_{rev} = \oint TdS$, where S is the entropy.
Examine the implications by considering two process paths: 1 reversible and 1 irreversible.

Implications for heat transfer $\Delta S_{rev}[A \to B] = \oint \frac{dQ_{rev}}{T}$ ←we are interested in the amount of heat transferred
$\Delta S_{rev}[A \to B] = \Delta S_{irr}[A \to B]$ (these are state functions)
However, an irreversible process produces entropy:

$$\Delta S_{irr,t} + \Delta S_{irr,p} = \Delta S_{rev,t} \tag{4.3}$$

t :transferred , p : produced
$\Delta S_{irr,t} = \Delta S_{rev,t} - \Delta S_{irr,p}$ Entropy transfer is associated with heat transfer
$\Delta S_{irr,t} = \oint \frac{\delta Q_{irr}}{T} < \oint \frac{\delta Q_{rev}}{T}$.
For an isothermal process, $[Q_{irr}]_T < [Q_{rev}]_T$, which is the maximum heat absorbed in any process going from A to B.
Clarification:

"Rev" indicates how a process was carried out, i.e. Q_{rev}indicates heat transferred for a reversible process. There is no "reversible" or irreversible heat.

Why does the path matter for us?

Reversible - P,V,Q,W easier to calculate

4.3.2 Irreversible Process

Irreversible - P,V,T path independent, but Q,W harder to calculate

4.4 3rd Law of Thermodynamics

Third Law: There is an absolute lower limit to the temperature of matter, and the entropy of all substances is the same at that temperature.
For a cyclic process,

$$\Delta S_{cyc} = \Delta S_I + \Delta S_{II} + \Delta S_{III} + \Delta S_{IV} = 0$$

Experimentally, we find that $\Delta S_I + \Delta S_{II} + \Delta S_{III} = 0$. We conclude that $\Delta S_{IV} = 0$.
The entropy of all substances is the same at absolute zero.
Utility: $\Delta S_{II} = -(\Delta S_I + \Delta S_{III})$ Provides information about chemical reaction.

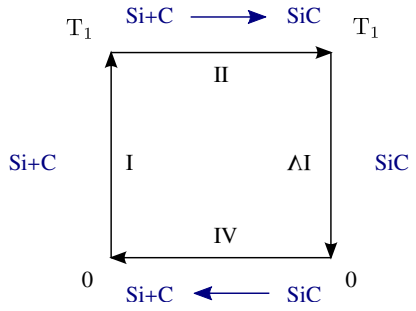


Figure 4.3: Cyclic process involving formation of silicon carbide, illustrating the third law of thermodynamics

The entropy of the compound in Figure is equal to the entropy of pure silicon and carbon at 0 K.

4.5 Combined statement of 1st and 2nd Laws

We begin with the following expression for δQ_{rev}:

$$\delta Q_{rev} = TdS \tag{4.4}$$

The expression for dU is:

$$dU = \delta Q_{rev} + \delta W + \delta W' \tag{4.5}$$

The work done on the system by an external pressure, V is included in dW:

$$\delta W = -PdV \tag{4.6}$$

These equations can be combined to give the following combined statement of the first and second laws:

$$dU = TdS - PdV + \delta W' \tag{4.7}$$

5 Thermodynamic Variables and Relations

Outcomes for this section:

1. Derive differentials of state functions (V, S, U, F, H, G) in terms of dP and dT.
2. Using the Maxwell relations, relate coefficients in the differentials of state functions.
3. Define the coefficient of thermal expansion, compressibility, and heat capacity of a material.
4. Express the differential of any given state function in terms of the differentials of two others.
5. Define reversible, adiabatic, isotropic, isobaric, and isothermal processes.
6. Calculate changes in state functions by defining reversible paths (if necessary) and integrating differentials for (a) an ideal gas, and (b) materials with specified α, β, and $C_P(T)$. (Multiple examples are given for ideal gas, solids, and liquids.)
7. Special emphasis (Example 4.13): calculate the change in Gibbs free energy when one mole of a substance is heated from room temperature to an arbitrary temperature at constant pressure.
8. Describe the origin of latent heat, and employ this concept in the calculation of changes in Gibbs free energy G.

5.1 Enthalpy

The enthalpy, H, is defined in the following way:

$$H \equiv U + PV \tag{5.1}$$

The total differential of the enthalpy is:

$$dH = dU + PdV + VdP \tag{5.2}$$

Now we eliminate dU by using Eq. 4.7:

$$dH = TdS + VdP + \delta W' \tag{5.3}$$

For a process at constant P, $dP = 0$ (isobaric), and if no other work is done $\delta W' = 0$. Under these conditions have:

$$dH_p = TdS_p = \delta Q_{rev,p} \tag{5.4}$$

where the 'p' subscripts indicate that the differential quantities are being evaluated at constant pressure. The enthalpy therefore measures the reversible heat exchange in an isobaric process ($\delta Q_{rev,p} = C_p dT_p$).

5.2 Helmholtz Free Energy

The Helmholtz free energy is defined in the following way:

$$F \equiv U - TS \tag{5.5}$$

In differential form this becomes:

$$dF = dU - TdS - SdT \tag{5.6}$$

$$dF = -SdT - PdV + \delta W' \tag{5.7}$$

For a process carried out at constant T (isothermal), $dT = 0$

$$dF_T = -PdV_T + \delta W'_T = \delta W_T + \delta W'_T = \delta W_{T,tot} \tag{5.8}$$

F_T reports the total (reversible) work done on the system.

5.3 Gibbs Free Energy

The Gibbs free energy, G, is defined in the following way.

$$G \equiv U + PV - TS = H - TS \tag{5.9}$$

$$dG = dU + VdP + PdV - TdS - SdT \tag{5.10}$$

$$= (\cancel{TdS} - \cancel{PdV} + \delta W') + VdP + \cancel{PdV} - \cancel{TdS} - SdT \tag{5.11}$$

$$dG = -SdT + VdP + \delta W' \tag{5.12}$$

For processes carried out under isothermal and isobaric conditions $(dT, dP = 0)$, then

$$dG_{T,P} = \delta W'_{T,P} \tag{5.13}$$

G reports the total work done other than mechanical work, e.g. chemical reactions and phase transformations.

5.4 Other Material Properties

5.4.1 Thermal expansion coefficient

Here we define some material properties that relate different thermodynamic state variables to one one another. The first is the thermal expansion coefficient, α, which describes the temperature dependence of the volume:

$$\alpha \equiv \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P \quad \left(\mathrm{K}^{-1}\right) \tag{5.14}$$

5.4.2 Compressibility

The compressibility, β, gives the relationship between the pressure and the volume :

$$\beta \equiv \frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \quad \left(\mathrm{Pa}^{-1}\right) \tag{5.15}$$

5.4.3 Heat Capacity

The heat capacity is the amount of energy needed to change temperature by some amount. Two different heat capacities are typically defined. The first of these, C_P, corresponds to a process performed at a constant pressure:, (J/mol-K)

$$\delta Q_{rev,P} \equiv C_P \underbrace{dT_P}_{\text{path specification}} \quad (\text{J/mol} \cdot \text{K}) \tag{5.16}$$

$$\delta Q_{rev,V} \equiv C_V \underbrace{dT_V}_{\text{path specification}} \quad (\text{J/mol} \cdot \text{K}) \tag{5.17}$$

Empirically, C_p is described by four parameters, a, b, c and d:

$$C_P(T) = a + bT + \frac{c}{T^2} + dT^2 \tag{5.18}$$

Note that C_P is necessarily larger than C_V. At constant P, the volume changes, so work is done on the surroundings. Therefore, more energy is needed to achieve the same ΔT. Below we show how We will show how C_V can be calculated from C_P.

5.5 Coefficient Relations

A variety of relationships exist between the different thermodynamic functions. These are a consequence of some basic mathematical relationships. Suppose that we have a function, Z of two variables X and Y:

$$Z = Z(X, Y) \tag{5.19}$$

The differential of Z s given by:

$$dZ = \left.\frac{\partial Z}{\partial X}\right|_Y dX + \left.\frac{\partial Z}{\partial Y}\right|_X dY \tag{5.20}$$

The differential potentials (dU, dH, dG, dG) all have the following

Table 5.1

$$dU = TdS - PdV \text{ (set } \delta W' = U)$$

$$T = \left(\frac{\partial U}{\partial S}\right)_V, -P = \left(\frac{\partial U}{\partial V}\right)_S \rightarrow \text{can now apply to H,F,G}$$

$$T = \left(\frac{\partial H}{\partial S}\right)_P, V = \left(\frac{\partial H}{\partial P}\right)_S$$

$$-S = \left(\frac{\partial F}{\partial T}\right)_V, -P = \left(\frac{\partial F}{\partial V}\right)_T$$

$$-S = \left(\frac{dG}{\partial T}\right)_P, V = \left(\frac{\partial G}{\partial P}\right)_T$$

5.5.1 Maxwell Relations

$$M = \left(\frac{\partial Z}{\partial X}\right)_Y, \; N = \left(\frac{\partial Z}{\partial Y}\right)_X$$

$$\left(\frac{\partial M}{\partial Y}\right)_X = \left(\frac{\partial N}{\partial X}\right)_Y \text{ Canchy-Riemann conditions}$$

$$dU = TdS - PdV \rightarrow \left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V$$

$$dH = TdS + VdP \rightarrow \left(\frac{\partial T}{\partial P}\right)_S = \left(\frac{\partial V}{\partial S}\right)_P$$

$$dF = -SdT - PdV \rightarrow \left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V$$

$$dG = -SdT + VdP \rightarrow -\left(\frac{\partial S}{\partial P}\right)_T = \left(\frac{\partial V}{\partial T}\right)_P$$

There are some examples.

$$V^\alpha = \left(\frac{\partial V}{\partial T}\right)_P \rightarrow V^\alpha = \left(\frac{\partial S}{\partial P}\right)_T$$

5.6 General Strategy for Deriving Relations

Equation of State Variables

- T, P: choose as independent variables
- V : choose as dependent variable
- S : dependent

Derive: $V(P,T)$ and $S(P,T)$
Energy Functions
U, H, F, G : Derive expressions in terms of T, P
One can then express any function in terms of any two independent variables.

5.6.1 Summary of Relations Derivation

$$Z = Z(X,Y)$$

$$dZ = \left(\frac{\partial Z}{\partial X}\right)_Y dX + \left(\frac{\partial Z}{\partial Y}\right)_X dY$$

$$\left(\frac{\partial M}{\partial Y}\right)_X = \left(\frac{\partial N}{\partial X}\right)_Y$$

$dU = TdS = -PdV$

$$= \left(\frac{\partial U}{\partial S}\right)_V dS + \left(\frac{\partial U}{\partial V}\right)_S dV$$

$$\Rightarrow T = \left(\frac{\partial U}{\partial S}\right)_V, -P = \left(\frac{\partial U}{\partial S}\right)_S; \left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V$$

$dG = -SdT + VdP$

$$\rightarrow \left(\frac{\partial S}{\partial P}\right)_T = \left(\frac{\partial V}{\partial T}\right)_P = V^\alpha \text{ (will use later)}$$

Another useful example:

$$dS = \left(\frac{\partial S}{\partial T}\right)_P dT + \left(\frac{\partial S}{\partial P}\right)_T dP \text{ (from dG)}$$

$$\left(\frac{\partial S}{\partial P}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_P = -V\alpha$$

$$\delta Q_{rev,P} = TdS_P = T\left(\frac{\partial S}{\partial T}\right)_P dT = C_P dT$$

$$(\delta Q_{rev,P} \equiv C_P dT_P)$$

$$\Rightarrow \boxed{\text{dS=}\tfrac{C_P}{T}\text{dT-V}\alpha\text{dP}}$$

It will be convenient to derive expressions for the differentials of S, U, V (and G, H, F) in terms of dT,dP.
Relate S to T, V

$$S(T,V) \rightarrow dS = MdT + NdV$$

$$dS = MdT + N(V^\alpha dT - V\beta dP)$$

$$dS = (M + NV^\alpha)dT - NV\beta dP$$

$$dS = \frac{C_P}{T}dT - V^\alpha dP$$

$$\frac{C_P}{T} = M + NV^\alpha$$

$$V^\alpha dP = N = \frac{\alpha}{\beta} = \left(\frac{\partial S}{\partial V}\right)_T$$

$$M + N = \frac{C_P}{T}, N = \frac{\alpha}{\beta}$$

$$M = \frac{C_P}{T} - \frac{(V^\alpha)^2}{\beta}$$

$$M = \frac{1}{T}\left(C_P - \frac{T\,(V^\alpha)^2}{\beta}\right) = \left(\frac{\partial S}{\partial T}\right)_V, \text{ and } N = \left(\frac{\partial S}{\partial V}\right)_T$$

$$dS = \frac{1}{T}\underbrace{\left(C_P - \frac{T\,(V^\alpha)^2}{\beta}\right)}_{M} dT + \frac{\alpha}{\beta}dV$$

$$M = \left(\frac{\partial S}{\partial T}\right)_V = \frac{C_V}{T} \text{ (4.40)}$$

$$\Rightarrow C_V = C_P - \underbrace{\frac{T\,(V^\alpha)^2}{\beta}}_{\text{positive quantity}} C_V < C_P \text{ as noted earlier}$$

Ideal Gases
Example 4.6:
Find change in T as a function of V, S with S held constant.
$dS = MdT + NdV$ Express in terms of P, T, then use Table 4.5 to solve for M, N
We found that $dS = \frac{C_V}{T}dT + \frac{\alpha}{\beta}dV_S$, $dT_s = -\frac{T_\alpha}{C_V\beta}dV_S$
One can integrate over change in volume to set ΔT.

6 The Ideal Gas

We begin to apply these principles by considering an ideal gas, which obeys the following constitutive relationship:

$$PV = nRT \tag{6.1}$$

Here n is the number of moles and R is the gas constant:

$$\begin{aligned} R &= 8.314\,(\mathrm{J/mol\cdot K}) \\ &= 1.987\,(\mathrm{cal/mol\cdot K}) \\ &= 0.08206\,(\mathrm{l-atm/mol\cdot K}) \end{aligned} \tag{6.2}$$

If n=1, then $PV = RT$, where $V = V_m$, where V_m is the molar volume.

Algorithm 1

Evaluate α, β, C_P, C_V, U, H for an ideal gas.

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P = \frac{P}{RT}\frac{\partial}{\partial T}\left(\frac{RT}{P}\right)_P = \frac{P}{RT}\cdot\frac{R}{P} = \frac{1}{T}$$

$$\beta = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T = -\frac{P}{RT}\frac{\partial}{\partial P}\left(\frac{RT}{P}\right)_T = -\frac{P}{RT}\left(\frac{-RT}{P^2}\right) = \frac{1}{P}$$

$$\therefore \frac{T\,(V^\alpha)^2}{\beta} = T\frac{RT}{P}\frac{1}{T^2}\cdot P = R$$

6.1 Specific Heat of Ideal Gas

$$C_V = C_P - \frac{T\,(V^\alpha)^2}{\beta} = C_P - R \tag{6.3}$$

We will show that $C_v = \frac{3}{2}R$ for a monatomic ideal gas. Therefore, $C_P = \frac{3}{2}R + R = \frac{5}{2}R$.
Example 4.8: Calculate the change in T when 1 mole of gas is compressed reversible and adiabatically (no heat flow) from $V_1$to V_2. $T(S,V)$
In Example 4.6, we found that $dS = \frac{C_V}{T}dT + \frac{\alpha}{\beta}dV$, which simplified to $dT_S = -\frac{T\alpha}{C_V\beta}dV_S$ for an isotropic process.
For an ideal gas, $\alpha = \frac{1}{T}$, $\beta = \frac{1}{P}$, so $dT_S = \frac{-T}{C_V}\frac{P}{T}dV_S = -\frac{P}{C_V}dV_S$. But $P = \frac{RT}{V}$

$$\Rightarrow dT_S = \frac{-RT}{C_V}\frac{dV_S}{V}.$$ Evaluate by integrating.

$$\int_{T_1}^{T_2}\frac{dT_S}{T} = \frac{-R}{C_V}\int_{V_1}^{V_2}\frac{dV_S}{V} \rightarrow \ln\left(\frac{T_2}{T_1}\right) = \frac{-R}{C_V}\ln\left(\frac{V_2}{V_1}\right)$$

$$\ln\left(\frac{T_2}{T_1}\right) = \ln\left[\left(\frac{V_2}{V_1}\right)^{-R/C_V}\right] \Rightarrow \frac{T_2}{T_1} = \left(\frac{V_1}{V_2}\right)^{\frac{R}{C_V}}$$

$$C_V = \frac{3}{2}R$$

$$T_2 = T_1\left(\frac{V_1}{V_2}\right)^{R/C_V} = T_1\left(\frac{V_1}{V_2}\right)^{2/3}$$

$\left(\frac{V_1}{V_2}\right) > 1$ for compression , so T increases.
If the gas expands ($V_2 > V_1$), then T decreases. One can also derive T(P) and P(V).

6.2 Internal Energy of Ideal Gas U(T,P)

$$dU = \underbrace{(C_p - PV^{\alpha})}_{i}\, dT + \underbrace{V(P\beta - T\alpha)}_{ii}\, dP \text{ (Table 4.5)}$$

$$i) C_P - PV\alpha = C_P - R = C_v$$

$$ii) P\beta - T\alpha = \frac{P}{P} - \frac{T}{T} = 0$$

$$\therefore dU = C_V dT, \text{ and } C_V = \tfrac{3}{2}R$$

One can then integrate to find that

$$\boxed{\Delta U = \tfrac{3}{2} R(T_2 - T_1)}$$

C_V is constant
For an ideal gas, the internal energy depends only on the temperature.

6.3 Examples Using Ideal Gas

Example 4.9: Free expansion of an ideal gas

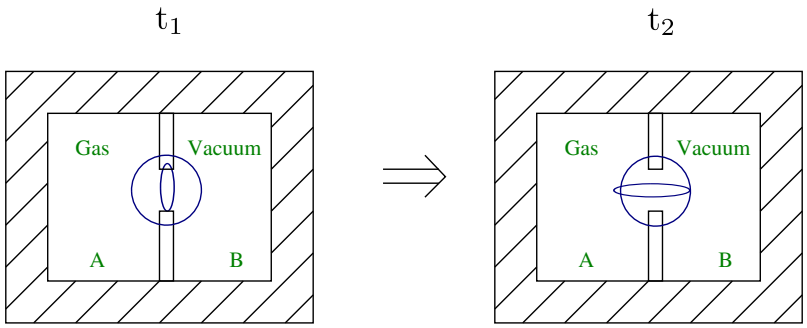


Figure 6.1: Free expansion of an ideal gas - the valve is opened at $t_2$and fills both chambers

Walls are rigid: $\delta W = 0$
Walls are insulating: $\delta Q = 0$
System is closed.
$\Rightarrow \Delta S_t = 0$, $\Delta U = 0 \Rightarrow \Delta T = 0$ (ideal gas)
The gas expands irreversibly.
Entropy is produced, but ΔS is not process dependent because S is a state function.
What is ΔS when 1 mole of gas expands freely to twice its volume? (recall $\Delta T = 0$).

$$dS = \tfrac{C_V}{T} dT + \tfrac{\alpha}{\beta} dV, dT = 0$$

$$PV = RT$$

$$\tfrac{P}{T} = \tfrac{R}{V}$$

$$dS = \tfrac{\alpha}{\beta} dV_T = \tfrac{P}{T} dV_T = \tfrac{R}{V} dV_T$$

$$\Delta S_T = \int_{V_1}^{V_2} \tfrac{R}{V} dV = R \ln\left(\tfrac{V_2}{V_1}\right) = R \ln\left(\tfrac{2V_1}{V_1}\right) = R \ln 2$$

$$\therefore \Delta S_T = 5.76 \text{J/mol-K}$$

This change for a reversible isothermal process is the same as for an irreversible free expansion in an isolated system. S is a state function.

Example 4.10a:
How much heat is absorbed/released when 1 mole of gas is compressed isothermally from 1 to 25 bar at 300K?
To find heat, use S: $\delta Q_{rev,T} = TdS_T$

$$dS = \frac{C_P}{T}dT - V^{\alpha}dP = -V^{\alpha}dP_T$$

The volume change is unknown, so use $V = \frac{RT}{P}$

$$dS_T = \frac{-R\cancel{T}}{P}\frac{1}{\cancel{T}}dP = -R\frac{dP}{P}$$

$$\Delta S_T = -R\int_{P_i}^{P_f}\frac{dP}{P} = -R\ln\left(\frac{P_f}{P_i}\right)$$

$$\rightarrow \Delta Q_{rev,T} = -RT\ln\left(\frac{P_f}{P_i}\right) = -(8.314J/mol-K)298K\ln(\frac{25}{1}) = -7.975\text{ kJ/mol}$$

Reversibility requires work/heat The piston in Figure 6.2 below may be withdrawn very slowly to maintain equilibrium.

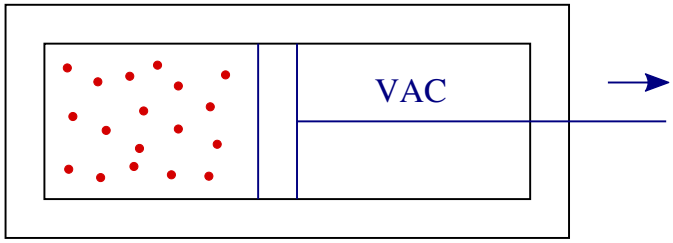


Figure 6.2: Piston-cylinder device

$$\Delta U = 0 \rightarrow \Delta T = 0$$

$$\Delta S = \int_{V_i}^{V_f}\frac{nR}{V}dV = nR\ln\left(\frac{V_f}{V_i}\right)$$

But no entropy can be produced, so heat flow is needed to maintain T (for free expansion, T decreases)
So actually, $\Delta U = Q + W = 0$
$\rightarrow W = -Q < 0$
The system does negative work! Heat flow is positive (inward).

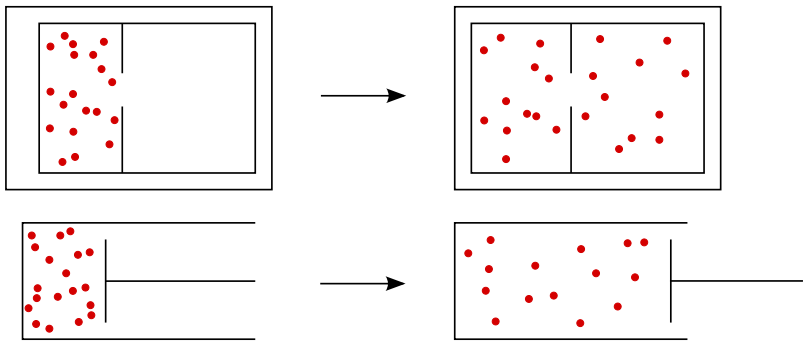


Figure 6.3: Irreversible Process (top) and reversible process (bottom)

Irreversible process	Reversible process
$\Delta Q = 0$	$\Delta Q \neq 0$
W=0	$W \neq 0$
$\Delta U = 0$	$\Delta U = 0$
$\Delta T = 0$	$\Delta T = 0$

Table 6.1: Irreversible vs reversible process characteristics

$$\Delta U = 0 \rightarrow W = -\Delta Q$$

$$W < 0, \Delta Q > 0$$

$$\rightarrow \Delta S_t > 0$$

What is the amount of heat required to raise the temperature to 1000K isobarically?

$$\delta Q_{rev,P} = TdS_p \equiv C_P dT_P$$

$$C_P = \frac{5}{2}R$$

$$\Delta Q_{rev,P} = C_P \Delta T = \frac{5}{2}R\Delta T = \frac{5}{2}(8.314 J/mol - K) \times (1000 - 298)K$$

Note that $dH = C_P dT + V(1 - T_\alpha)dP$. For $dP = 0$, $dH = C_P dT \rightarrow$ Enthalpy measures reversible heat flow in absence of work. This will be useful in the analysis of solids heated under conditions of constant pressure.

Some useful relations for heat flow

$$\delta Q_{rev,T} = TdS_T = dU_T - dF_T$$

$$\delta Q_{rev,T} = TdS_P \equiv C_P dT_P = dH_P \ (\delta W' = 0)$$

$$\delta Q_{rev,V} = TdS_V \equiv C_V dT_V = dU_V \ (\delta W' = 0)$$

Example 4.3: Relate S(P,V)

$$dS = MdP + NdV$$

$$dS = MdP + N(V^\alpha dT - V\beta dP) = (M - NV_\beta)dP + NV^\alpha dT$$

$$dS = -V^\alpha dP + \frac{C_P}{T}dT \rightarrow N = \frac{C_P}{V^\alpha T} = \left(\frac{\partial S}{\partial V}\right)_P$$

$$M - NV\beta = -V^\alpha$$

$$M = -V^\alpha + \frac{C_P}{\cancel{V}_\alpha T}\cancel{V}\beta$$

$$M = \frac{C_P\beta}{\alpha T} - V^\alpha = \left(\frac{\partial S}{\partial P}\right)_V$$

$$dS = \left(\frac{C_P\beta}{\alpha T} - V^\alpha\right)dP + \frac{C_P}{V^\alpha T}dV$$

Luijten examples

1. Q for ΔP at constant T
   Reversible, $\rightarrow \delta Q_{rev,T} = TdS$
   Use $S(T,P)$ like 4.10(a)
2. Find ΔT given ΔP at constant S
   Use $S(T,V)$, also idela gas law (~4.8)
3. Free expansion of ideal gas, find ΔS for ΔV
   $S(T,V)$ (4.9)
4. Adiabatic reversible compression of a solid.
   $T(S,P)$ $dS = 0$,ΔP given (4.14)
5. Change in G for heating at constant P
   $G(P,T)$, $dP = 0$ (4.13)
   $dG = -SdT + VdP$, need to find $S(T)$

$$PV = RT$$

1 mole of gas expands at an initial temperature constant S from $1l$ to $2l$. What is T_f? $(T_i = 300K)$

$$dS = MdT + NdP = 0$$

$$\left(\frac{C_P}{T} - \frac{(V^\alpha)^2}{P}\right) dT + \frac{\alpha}{\beta} dV = 0 \; \alpha = \frac{1}{T}, \beta = \frac{1}{P}$$

$$\frac{C_V}{T} dT = -\frac{\alpha}{\beta C_V} dV$$

$$\frac{dT}{T} = \frac{-\alpha}{\beta C_V} dV$$

$$\frac{dT}{\cancel{T}} = \frac{-P}{\cancel{T} C_V} dU$$

$$dT = -\frac{P}{C_V} dV$$

Use $PV = RT$ to get to two variables

$$P = \frac{RT}{V}, C_V = \frac{3}{2}R$$

$$dT = \frac{-RT}{C_V V} dV =$$

$$\frac{dT}{T} = \frac{-R}{3/2R} \frac{dV}{V} = \frac{-2}{3} \frac{dV}{V}$$

$$\int_{T_1}^{T_2} \frac{dT}{T} = -\frac{2}{3} \int_{V_1}^{V_2} \frac{dV}{V} \rightarrow \ln\left(\frac{T_2}{T_1}\right) = \ln\left(\frac{V_1}{V_2}\right)^{2/3}$$

$$\frac{V_1}{V_2} = \frac{1}{2}, T_2 = T_1 \left(\frac{1}{2}\right)^{2/3} = 189\text{K}$$

$$\text{Note } P_1 = \frac{300 \cdot R}{10^{-3} m^3} = 3 \times 10^5 R$$

$$P_2 = \frac{189R}{2 \times 10^{-3} m^3} = 0.95 \times 10^5 R$$

Pressure drops by threefold
$\downarrow PV \uparrow = RT \downarrow$
1 mole Fe is compressed reversibly. What is the change in T?
$P_1 = 1\text{atm}$, $P_2 = 10^5\text{atm}$

$$dS = \frac{C_p}{T}dT - V^{\alpha}dP$$

$\frac{dT}{T} = \frac{V^{\alpha}}{C_P}dP$ need exponential values
Does C_P vary with T?
If $\frac{dC_P}{dT} \approx 0$, then

$$\ln\left(\frac{T_2}{T_1}\right) = \frac{V^{\alpha}}{C_P}\Delta P$$

$$\alpha = 36 \times 10^{-6}\,\text{K}^{-1}$$

$$C_P \approx 45.3\,\text{J/mol} - \text{K}$$

Otherwise, $\int \frac{C_P}{T}dT = V^{\alpha}\Delta P$

$$V = \frac{3.1cm^3}{mol} = 7.1 \times 10^{-6}m^3/mole$$

$$T_2 = T_1 \times 1.006 = 300K$$

7 Solids and Liquids

7.1 Example 4.13: ΔG for MgO at constant P

Compute ΔG when 1 mole of MgO is heated from 300K to temperature T at constant pressure.

$$dG = -SdT + VdP$$

$$dP = 0$$

$$dG_P = -SdT_P = -S(T)dT_P$$

But $C_P$depends on temperature $\rightarrow$so does S.

$$dS = \frac{C_P}{T}dT - V^{\alpha}dP = \frac{C_P}{T}dT$$

$$\left(\frac{\partial S}{\partial T}\right)_P = \frac{C_P}{T}$$

$$S = S_o + \int \frac{C_P(T)}{T}dT$$

$$C_P = a + bT, S_o = 26.8J/mol - K$$

$$S = S_o + \int_{T_1}^{T_2} dT\left(\frac{a}{T} + b\right)$$

Evaluate expression, substitute data at the end

$$S = S_o + \int \frac{C_P(T)}{T}dT = S_o + \int_{T_1}^{T_2} dT\left(\frac{a}{T} + b\right)$$

$$S = S_o + a\ln\left(\frac{T_2}{T_1}\right) + b(T_2 - T_1)$$

$$T_1 = 300K, T_2 = T$$

$$S(T) = S(300K) + a\ln\left(\frac{T}{300}\right) + b(T - 300)$$

$$\equiv a\ln T + bT + \underbrace{[S_o - a\ln 300 - 300b]}_{\text{constant}}$$

$$C_P(T)_{MgO} = \underbrace{48.99}_{a} + \underbrace{3.43\times10^{-3}T}_{b} - \underbrace{\frac{11.34\times10^{-5}}{T^2}}_{\text{ignored}}(\frac{J}{mol-K})$$

$$\Delta G = \int_{300}^{T} dG = -\int_{300}^{T_f} S(T)dT = -\int_{300}^{T_f}(a\ln T + bT + C)dT$$

$$= [a(T\ln T - T) + \frac{1}{2}bT^2 + cT]_{T_f}^{300}$$

$$\Delta G = aT\ln T + \frac{1}{2}bT^2 + (C-a)T = -300a\ln 300 - \frac{1}{2}b(300)^2 - (C-a)300$$

↑Check this, see also sec 7.3 p. 180

$$a = 48.99\frac{J}{mol-K},\ b = 3.43\times10^{-3}\frac{J}{molK^2}$$

$$C = 26.8J/mol - K - a\ln 300 - 300b$$

Consider the following exothermic reaction (referred to as the Thermite reaction), where Aluminum reacts with iron oxide to form aluminum oxide and iron.

$$2Al + Fe_2O_3 \rightarrow Al_2O_3 + 2Fe$$

How hot will this system get as the reaction proceeds? We'll do this in two parts. First, we'll calculate the heat generated in the reaction, then we'll figure out how much this heat will increase the temperature of the system. Consider the change in enthalpy for each half-reaction:

$$\begin{array}{ll} 2Al + \frac{3}{2}O_2 \rightarrow Al_2O_3 & \blacksquare H_1 = -1675.7\frac{kJ}{mol} \\ Fe_2O_3 \rightarrow 2Fe + \frac{3}{2}O_2 & \blacksquare H_2 = 823\frac{kJ}{mol} \end{array}$$

$$2Al + \cancel{\frac{3}{2}O_2} + Fe_2O_3 \rightarrow Al_2O_3 + 2Fe + \cancel{\frac{3}{2}O_2}$$

$$\Delta H = (-1675.1 + 823)\frac{kJ}{mol} = -852.3\frac{kJ}{mol}$$

The minus sign implies that the reaction is exothermic. Now that we know how much heat is generated, we can figure out what the state of the system will be. Note that Fe melts at 1809 K and boils at 3343 K, and Al_2O_3 melts at 2325 K. We need to determine to what temperature the system will rise assuming that all heat generated by the reaction is absorbed by the system.
Step 1: Heat Fe and $Al_2O_3$to the melting point of Fe.

$$\Delta H_{Al_2O_3,1} = \int_{289}^{1809} C_P(T)dT = 183,649\ \text{J}$$

$\Delta H_{Fe,1}$ involves 3 sub-steps: (2 moles)

$$\uparrow \alpha \rightarrow \gamma : 2\int_{289}^{1187} C_{P,\alpha} dT + \Delta H_{\alpha\rightarrow\gamma}$$

$$\uparrow \gamma \rightarrow \delta : 2\int_{1187}^{1164} C_{P,\gamma} dT + \Delta H_{\gamma\rightarrow\delta}$$

$$\uparrow \delta \rightarrow l : 2\int_{1664}^{1809} C_{P,\delta} dT + \Delta H_{melt}$$

$$C_{P,\delta} = C_P, \alpha$$

Considering the three integrals of T and three latent heats,

$$\Delta H_{Fe,1} = 157,741 \text{ J (for 2 moles)}$$

So through step 1, we have $\Delta H_{Fe,1} + \Delta H_{Al_2O_3,1} = 341.390$ J

Step 2: Heat Fe and Al_2O_3 to melting point of Al_2O_3.

$$\Delta H_{Al_2O_3,2} = \int_{1809}^{2325} C_{P,Al} dT + \Delta H_{Al_2O_3}(S \rightarrow l) = (71,240 + 107,000) \text{ J}$$

$$\Delta H_{Fe,2} = 2\int_{1809}^{2325} C_{P,Fe(l)} dT = 43,178 \text{ J}$$

How much is left? $852,300 - 341,390 - 221,418 = 289,492$ J
Step 3: How much energy is needed to get to the boiling point of Fe?

$272,600$ J $\begin{cases} \Delta H_{Fe,3} = 2\int_{2325}^{3343} C_{P,Fe}(l) dT \\ \Delta H_{Al_2O_3,3} = \int_{2325}^{3343} C_{P,Al_2O_3}(l) dT \end{cases}$ leaving 16.892 J

How much Fe evaporates?

$$\frac{Q}{\Delta H_{Fe}(l\rightarrow g)} = \frac{16,892J}{340,159J} = 0.05\text{moles or } 5\%$$

$$\Delta H(T) = \Delta a T + \frac{\Delta b}{2} T^2 - \Delta C \frac{1}{T} + \frac{\Delta d}{3} T^3 - \Delta D$$

$$\Delta D = \Delta H(T_o) - [\Delta a T_o + \frac{\Delta b}{2} T_o^2 - \Delta C \frac{1}{T_o} + \frac{\Delta d}{3} T_o^3]$$

$$2Al + Fe_2O_3 \rightarrow 2Fe + Al_2O_3 \text{ Exothermic}$$

Compute the final state and temperature assuming adiabataic, reversible reaction at constant pressure.

$$dT, Q = 0, dP = 0$$

Need initial S, C_P
Exothermic implies ΔT from heat released.

$$dS = MdT + N\cancel{dP}$$

$$dH = TdS + VdP + \delta W'$$

$$2Al + \frac{3}{2}O_2 \rightarrow Al_2O_3 : \Delta H_1 = -1875\text{kJ/mol}$$

$$2Fe + \frac{3}{2}O_2 \rightarrow Fe_2O_3 \text{: } \Delta H_2 = -823.4\text{kJ/mol}$$

$$2Al + Fe_2O_3 \rightarrow 2Fe + Al_2O_3$$

$$2Al + \frac{3}{2}O_2 \rightarrow Al_2O_3 \text{ ; } \Delta H_1$$

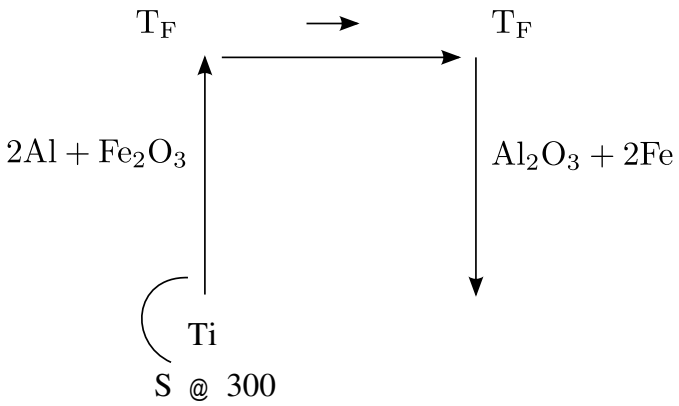


Figure 7.1: Thermite reaction

$$\underline{Fe_2O_3 \rightarrow 2Fe + \tfrac{3}{2}O_2; -\Delta H_2}$$

$$2Al + Fe_2O_3 + \cancel{\tfrac{3}{2}O_2} \rightarrow Al_2O_3 + 2Fe + \cancel{\tfrac{3}{2}O_2}$$

$$\Delta H_1 - \Delta H_2 = -852.3kJ/mol$$

$$\delta Q_{rev,P} = -TdS = dH$$

Heat is reversibly absorbed by the system.
$\Delta H_{Fe,1} = \int dH = \int_{298}^{1809} C_P dT = 157,741$J for 2 mol
Include ΔH due to increasing T as well as due to phase transitions. For Al_2O_3, just $\int C_P dT$
$\Delta H_{Al_2O_3} = 183,649$ J
For a total change of 341,390 J
Now heat $Al_2O_3$to melting point, and heat $Fe(l)$

$$\int C_P dT + \Delta H_{melt}$$

Step 2 takes: 221,418 J
852,300 J is provided by the reaction: $852,300 - 221,418 - 341,390 = 289,492$ J

$$dS = \frac{C_V}{T}dT + \frac{\alpha}{\beta}dV = \frac{3}{2}\frac{R}{T}dT + \frac{P}{T}dV = \frac{3}{2}\frac{R}{T}dT + \frac{R}{V}dV$$

Good for constant V: relate ΔS to T change.
For constant T: relate ΔS to volume change.
For constant S, dS=0

$$\left(\tfrac{T_2}{T_1}\right)^{3/2} = \left(\tfrac{V_1}{V_2}\right)\tfrac{T_2}{T_1} = \left(\tfrac{V_1}{V_2}\right)^{2/3}$$

$$\tfrac{T_2}{T_1} = \tfrac{V_2}{V_1}\tfrac{P_2}{P_1}$$

$$\left(\tfrac{P_2}{P_1}\right)^{3/2}\left(\tfrac{V_2}{V_1}\right)^{3/2} = \tfrac{V_1}{V_2} \rightarrow \left(\tfrac{P_2}{P_1}\right)^{3/2} = \left(\tfrac{V_1}{V_2}\right)^{5/2}$$

$$\tfrac{P_2}{P_1} = \left(\tfrac{V_1}{V_2}\right)^{5/3}$$

$$T_1 V_1^{\gamma-1} = T_2 V_2^{\gamma-1}$$

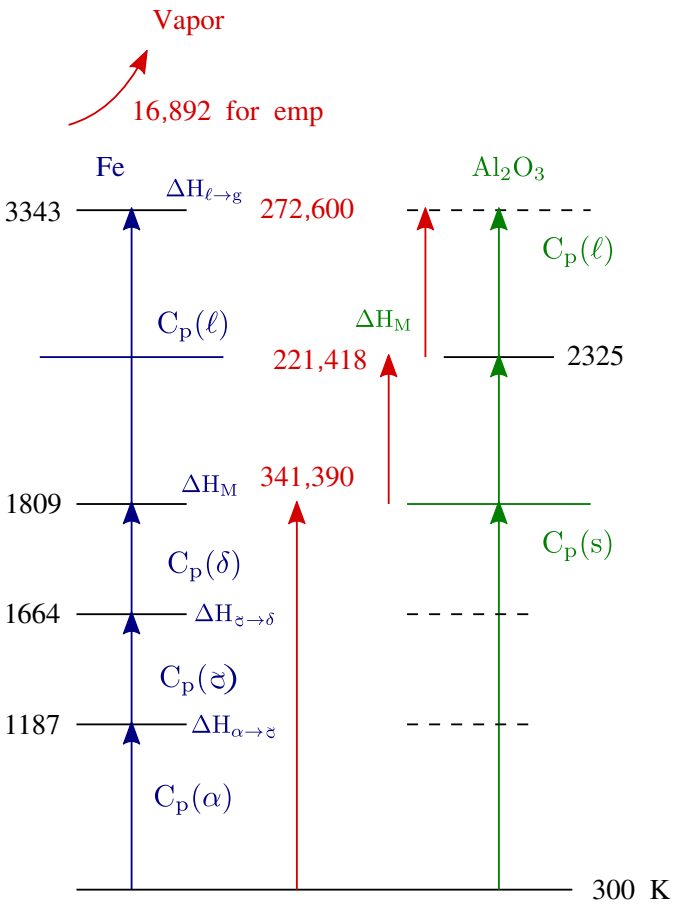


Figure 7.2: Changes in enthalpy in Fe and Al_2O_3

$$T_1^{\frac{\gamma}{1-\gamma}} P_1 = T_2^{\frac{\gamma}{1-\gamma}} P_2$$

$$P_1 V_1^{\gamma} = P_2 V_2^{\gamma}$$

$$\gamma = \frac{C_p}{C_V} = \frac{5/2}{3/2} = 5/3$$

$$\frac{5/3}{1-5/3} = \frac{5/3}{2/3} = 5/2$$

$$\gamma - 1 = 2/3$$

$$dU = (C_P - PV^{\alpha})dT + V(\rho\beta - T_{\alpha})dP$$

$$dU = (\frac{5}{2}R - R)dT + V(1-1)dP = \frac{3}{2}RT$$

$$dS = \frac{C_P}{T}dT - V^{\alpha}dP = \frac{5}{2}\frac{R}{T}dT - \frac{V}{T}dP$$

$$dS = \frac{5}{2}\frac{R}{T}dT - \frac{R}{P}dP$$

$$dS = \frac{C_V}{T}dT + \frac{\alpha}{\beta}dV = \frac{3}{2}\frac{R}{T}dT + \frac{P}{T}dV = \frac{3}{2}\frac{R}{T}dT + \frac{R}{V}dV$$

Constant P: $\Delta S = \frac{5}{2} R \ln\left(\frac{T_2}{T_1}\right)$
Constant T: $\Delta S = -R \ln\left(\frac{P_2}{P_1}\right)$
$\Delta S = -R \ln\left(\frac{V_2}{V_1}\right)$
Constant V: $\Delta S = \frac{3}{2} R \ln\left(\frac{T_2}{T_1}\right)$
Constant S: $\frac{P_1}{P_2} = \left(\frac{T_2}{T_1}\right)^{5/2} = \left(\frac{V_1}{V_2}\right)^{5/3}$

$$dU = -PdV + TdS$$

$$dU = \left(\frac{-\partial U}{\partial V}\right)_S dV + \left(\frac{\partial U}{\partial S}\right)_V dS \Rightarrow \left[\frac{\partial}{\partial S}\left(\frac{-\partial U}{\partial V}\right)\right]_V = \left[\frac{\partial}{\partial V}\left(\frac{\partial U}{\partial S}\right)_V\right]_S$$

Note that these are second derivatives of the internal energy, so they should be negative for equilibrium.

8 Conditions for Equilibrium

Outcomes for this section:

1. Derive the condition for equilibrium between two phases.
2. Show that at constant temperature and pressure, the equilibrium state corresponds to that which minimizes the Gibbs free energy G.
3. Explain why the evolution to a state of lower Gibbs free energy occurs spontaneously.

8.1 Equilibrium in thermodynamic system

For an isolated system ($\delta Q = 0, \delta W = 0$, n fixed, see)
Equilibrium $\longleftrightarrow$State of Maximum Enropy
Maximize S (function of multiple variables) and we will have determined the equilibrium state (regardless of past history).
How do we recognize that a system has reached equilibrium?

1. It is in a stationary state (i.e. not changing)
2. If the system is isolated from its surroundings, no further changes occur.

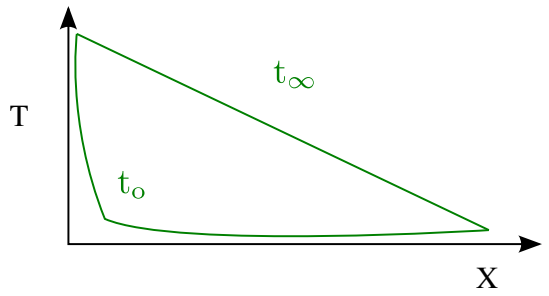


Figure 8.1

Counter example:
After $t \to \infty$, system is in steady state. Isolation of rod from heat sources/sinks would result in future evolution.
z varies monotonically with x,y - there is no global maximum. Constraints such as $y = y(x)$ may define a maximum in $z(x, y)$ subject to $y = y(x)$
Example 5.1

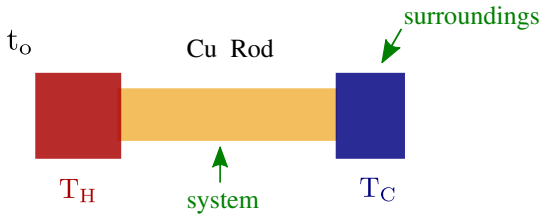


Figure 8.2

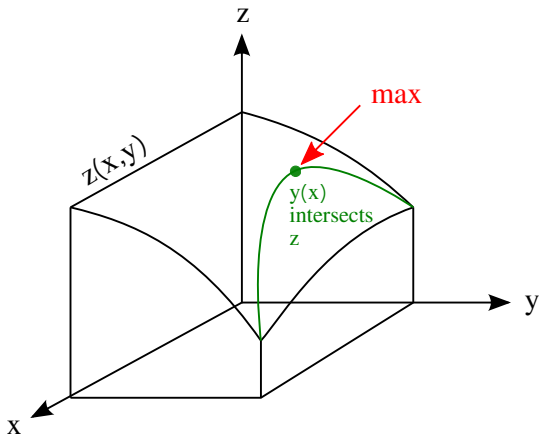


Figure 8.3

$z = xu + yv$
For simple cases, eliminate dependent variables. In general, use Lagrange Multipliers.
Challenge: minimize a funciton of several variables when relationships (constraints) exist between those variables.
Approach: Lagrange Multipliers

Minimize: (will be S)	Given these constraints (other state functions,variables)	
$f(x_1, x_2, ...x_n)$	$\Phi_1(x_1, x_2, ...x_n) = 0$	
function of n variables	$\Phi_2(x_1, x_2, ...) = 0$	$\Rightarrow$ m constraints
degrees of freedom=n-m	$\vdots$	
	$\Phi_m(x_1, x_2...) = 0$	

Define: $df + \lambda_1 d\Phi_1 + \lambda_2 d\Phi_2 + ... + \lambda_m d\Phi_m = 0$
λ_m : Lagrange multiplier
$dF = \frac{\partial F}{\partial x_1} dx_1 + \frac{\partial F}{\partial x_2} dx_2 + ... + \frac{\partial F}{\partial x_n} dx_n$
In general, $dF = \frac{\partial F}{\partial x_1} dx_1 + \frac{\partial F}{\partial x_2} dx_2 + ... + \frac{\partial F}{\partial x_n} dx_n$.
To guarantee that dF=0 for all $x_1, x_2...x_n$, the coefficients $\frac{\partial F}{\partial x_i}$ must all be zero.
i.e. $\frac{\partial F}{\partial x_i} + \sum_{k=1}^{m} \lambda_K \frac{\partial \Phi_K}{\partial x_i} = 0$
Example 5.1
$z = xu + yv$
Constraints

- $u = x + y + 12 \rightarrow \Phi_1 = x + y + 12 - u = 0$
- $v = x - y - 8 \rightarrow \Phi_2 = x - y - 8 - v = 0$

u, v: dependent
x, y: independent
$dz + \lambda_1 d\Phi_2 + \lambda_2 d\Phi_2 = 0$ Minimize z subject to constraints

$$= xdu + udx + ydv + vdy + \lambda_1(dx + dy - du) + \lambda_2(dx - dy - dv) = 0$$

$$= (x - \lambda_1)du + (y - \lambda_2)dv + (u + \lambda_1 + \lambda_2)dx + (v + \lambda_1 - \lambda_2)dy = 0$$

$$\left.\begin{array}{l} x - \lambda_1 = 0 \\ y - \lambda_2 = 0 \end{array}\right\} \Rightarrow \begin{array}{ll} \lambda_1 = x & u + \lambda_1 + \lambda_2 = 0 \\ \lambda_2 = y & v + \lambda_1 - \lambda_2 = 0 \end{array} \left.\begin{array}{l} \\ \end{array}\right\} \Rightarrow \begin{array}{l} u = -\lambda_1 - \lambda_2 \\ v = -\lambda_1 + \lambda_2 \end{array} \Rightarrow \begin{array}{l} u = -x - y \\ v = -x + y \end{array}$$

$$\begin{array}{l} x + y + 12 - u = 0 \\ \underline{x - y - 8 - v = 0} \end{array}$$

$$\begin{array}{ll} 2x + 4 - (-x - y) - (-x + y) = 0 & -1 - y - 8 - (-(-1) + y) = 0 \\ 4x + 4 = 0 \Rightarrow x = -1 & -10 - 2y = 0 \rightarrow y = -5 \end{array}$$

8.2 Finding equilibrium conditions

A unary 2-phase system is in equilibrium when the T, P and μ for both phases are equal.
Consider 2 phases α and β of one component.
$U = U^\alpha + U^\beta$, $V = V^\alpha + V_\beta$, $S = S^\alpha + S^\beta$ because these are extensive properties.
So $U^\alpha(S^\alpha, T_\alpha, n_\alpha)$ where n_α is the number of moles of α phase.
Then $dV^\alpha = T_\alpha dS^\alpha - P_\alpha dV^\alpha + \mu_\alpha dn_\alpha$, where $\mu \equiv \left(\frac{\partial V^\alpha}{\partial n_\alpha}\right)_{V_\alpha, S^\alpha}$
μ_α is the chemical potential of the α phase. To find the equilibrium condition, maximize S.

$$dS^\alpha = \frac{dU_\alpha}{T_\alpha} + \frac{P_\alpha}{T_\alpha}dV^\alpha - \frac{\mu_\alpha}{T_\alpha}dn_\alpha,\ dS^\beta = \frac{dV_\beta}{T_\beta} + \frac{P_\beta}{T_\beta}dV_\beta - \frac{\mu_\beta}{T_\beta}dn_\beta$$

$$dS = dS^\alpha + dS^\beta = \frac{dU_\alpha}{T_\alpha} + \frac{P_\alpha}{T_\alpha}dV^\alpha - \frac{\mu_\alpha}{T_\alpha}dn_\alpha + \frac{dV_\beta}{T_\beta} + \frac{P_\beta}{T_\beta}dV_\beta - \frac{\mu_\beta}{T_\beta}dn_\beta$$

If the entropy is maximized, then no changes will occur if the system is isolated (rigid walls, insulated).

$$dU = dU^\alpha + dV_\beta = 0 \rightarrow dU^\alpha = -dU^\beta \text{ (no heat is exchanged)}$$

$$dV = dV^\alpha + dV_\beta = 0 \rightarrow dV^\alpha = -dV_\beta \text{ (no work is done)}$$

$$dn_\alpha = -dn_\beta \text{ (particle number is conserved)}$$

Then $dS = \left(\frac{1}{T_\alpha} - \frac{1}{T_\beta}\right)dU^\alpha + \left(\frac{P_\alpha}{T_\alpha} - \frac{P_\beta}{T_\beta}\right)dV_\beta - \left(\frac{\mu_\alpha}{T_\alpha} - \frac{\mu_\beta}{T_\beta}\right)dn_\alpha = 0$
To guarantee dS=0, these coefficients must be zero independently.
Hence,

$$\boxed{T_\alpha = T_\beta, P_\alpha = P_\beta, \mu_A = \mu_B} \quad \text{Conditions for Equilibrium}$$

8.3 Minima of energies as conditions of equilibrium

Consider first the internal energy U in an m-component system.
$dU = TdS + PdV + \sum_i^m \mu_i dn_i$ where i : index of components
$dU = \delta Q + \delta W + \delta W'$
Consider a reversible process in an isolated system
$\delta Q_{rev} = TdS$ $\delta W_{rev} = -PdV \Rightarrow \delta W' = \sum_i \mu_i d\mu_i$ "chemical" work associated with reversible reactions
Claim: For a system of constant S, V, n_λ, the equilibrium state is a state of minimum internal energy U.
If this is not true, then work can be extracted and reinserted as heat. But then S would increase while nothing else changes. Therefore, energy cannot be extracted from a system under these conditions (unchanging/maximized $S_{v,n}$).

Derive conditions for equilibrium again $dU = 0$ (minimize U)

$dS = 0 \rightarrow dS^\alpha = -dS^\beta$
$dV = 0 \rightarrow dV^\alpha = -dV_\beta$
$dn = 0 \rightarrow dn_\alpha = -dn_\beta$
$dU^\alpha = T_\alpha dS^\alpha - P_\alpha dV^\alpha + \mu_\alpha dn_\alpha$
$dU^\beta = T_\beta dS^\beta - P_\beta dV_\beta + \mu_\beta dn_\alpha$
$dU = (T_\alpha - T_\beta)dS^\alpha - (P_\alpha - P_\beta)dV^\alpha + (\mu_\alpha - \mu_\beta)dn_\alpha = 0$
$\Rightarrow T_\alpha = T_\beta$, $P_\alpha = P_\beta$, $\mu_\alpha = \mu_\beta$ (Additional condition: second derivative)
$\Rightarrow$Thermal, mechanical, and chemical equilibium

U will decrease spontaneously if a system of constant S and V is not in equilibrium.
Enthalpy $H = U + PV$
Hemholtz Free Energy $F = U - TS$
Gibbs Free Energy $G = U + PV - TS$

Constant	Equilibrium
S, P	$dH = 0$
T, V	$dF = 0$
T, P	$dG = 0$
S, V	$dV = 0$

If T and P are held constant, a system will seek to minimize the Gibbs free energy, G. Consider two states connected by an isothermal and isobaric path:

$$G_1 = U_1 + PV_1 - TS_1$$

$$G_2 = U_2 + PV_2 - TS_2$$

$$\Delta G = \Delta U + P\Delta V - T\Delta S \text{ , and } \Delta U = Q + W + W'$$

Consider first a reversible path:
$W_{rev,T} = -P\Delta V$ and $Q_{rev,P} = T\Delta S$

$$\Delta G = [Q + W + W'] - W_{rev} - Q_{rev}$$

For any path: $\Delta G = \underbrace{(Q - Q_{rev})}_{<0} + \underbrace{(W - W_{rev})}_{<0} + \underbrace{W'}_{<0}$

In a system at constant T and P that is not in equilibrium, G will tend to decrease spontaneously to approach chemical equilibrium.

Example for two component system Two isolated systems are defined - each in equilibrium with separate environments.

$P_1V_1 = n_1RT_1$
$P_2V_2 = n_2RT_2$
} Let them come to equilibrium in isolation

$\Delta U_1 + \Delta U_2 = 0$

$$\tfrac{3}{2}n_1R(T_f - T_i) + \tfrac{3}{2}n_2R(T_f - T_i) = 0 \Rightarrow n_1T_f - n_1T_1 = -n_2T_f + n_2T_2 \Rightarrow T_f = \tfrac{n_1T_1 + n_2T_2}{n_1 + n_2}$$

We can also solve for P_f, V_i

$$P_f = \tfrac{(n_1+n_2)RT_f}{V_1+V_2} = \tfrac{R(n_1T_1+n_2T_2)}{V_1+V_2} \text{ and } V_1 = \tfrac{n_1RT_f}{P_f} \text{ , } V_2 = \tfrac{n_2RT_f}{P_f}$$

$$V_1 + V_2 = \tfrac{n_1RT_f}{P_f} + \tfrac{n_2RT_f}{P_f}$$

P_f	V
$n_1 + n_2$	T_f

Final

-S	U	V
H		F
-P	G	T

Figure 8.4: Thermodynamic square used to help determine thermodynamic relations

8.4 Maxwell Relations

For free energies:

- coefficients in opposite corners
- derivatives in adjacent corners
- minus signs apply to coeffiencients

For Maxwell Relations:
$\curvearrowleft = \curvearrowright$ i.e. $-\left(\frac{\partial S}{\partial P}\right)_T = \left(\frac{\partial V}{\partial T}\right)_P$
Note: This does not include the chemical potential. $\delta W' = \sum_i^m \mu_i dn_i$

Useful relations between S,T,P,V for an ideal monatomic gas $dS = \frac{C_p}{T}dT - V^\alpha dP = \frac{5}{2}\frac{R}{T}dT - \frac{V}{T}dP = \frac{5}{2}\frac{R}{T}dT - \frac{R}{P}dP$

$dS = \frac{C_V}{T}dT + \frac{\alpha}{\beta}dV = \frac{3}{2}\frac{R}{T}dT + \frac{P}{T}dV = \frac{3}{2}\frac{R}{T}dT + \frac{R}{V}dV$

Constant P: $\{\ \Delta S = \frac{5}{2}R\ln\left(\frac{T_2}{T_2}\right)$

Constant T: $\begin{cases} \Delta S = -R\ln\left(\frac{P_2}{P_1}\right) \\ \Delta S = -R\ln\left(\frac{V_2}{V_1}\right) \end{cases}$

Constant V: $\Delta S = \frac{3}{2}R\ln\left(\frac{T_2}{T_1}\right)$

Constant S: $\frac{P_2}{P_1} = \left(\frac{T_2}{T_1}\right)^{5/2} = \left(\frac{V_1}{V_2}\right)^{5/3}$

$dS = \frac{C_v}{T}dT + \frac{\alpha}{\beta}dV = \frac{3}{2}\frac{R}{T}dT + \frac{P}{T}dV = \frac{3}{2}\frac{R}{T}dT + \frac{R}{V}dV$

$0 = \frac{3}{2}\frac{R}{T}dT + \frac{R}{V}dV \rightarrow \frac{3}{2}\ln\left(\frac{T_2}{T_1}\right) = -\ln\left(\frac{V_2}{V_1}\right)$

$\left(\frac{T_2}{T_1}\right)^{3/2} = \frac{V_1}{V_2}, \left(\frac{T_2}{T_1}\right)^{5/2} = \left(\frac{V_1}{V_2}\right)^{5/3} = \frac{P_2}{P_1}$

One can show this using $\frac{P_1V_1}{T_1} = \frac{P_2V_2}{T_2}$ and substituting.

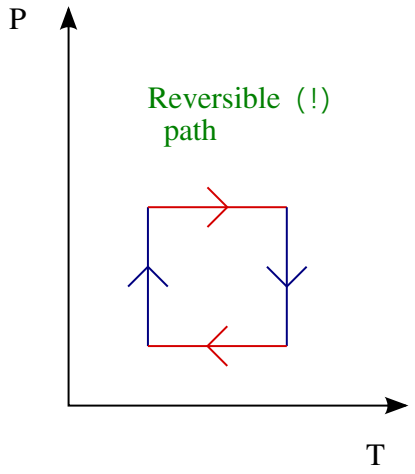


Figure 8.5

9 From Microstates to Entropy

Outcomes for this section:

1. Express the entropy S of the system in terms of equivalent microstates in the equilibrium (most probable) macrostate.
2. Given the energy levels of a system of particles, calculate the partition function.
3. Given the partition function, calculate the Helmholtz free energy, the entropy, the internal energy, and the heat capacity.

Destination: $S = k_B \ln \Omega$
where k_B =Boltzmann's constant, $8.317 \times 10^{-5} eV/K, 1.38 \times 10^{-23} J/K$ and Ω = number of equivalent microstates in (equilibrium) macrostate.
Matter is composed of atoms, and is therefore discrete. Contimuum thermodynamics arises from interactions between discrete components and interactions with the environment.
A statistical description is needed to connect the microscopic and macroscopic descriptions.

9.1 Microstates and Macrostates

Microstate: Specification of the state (*e.g.* energy level) of every atom at a particular point in time.
Example: 4 identical particles in 2 levels, pour 4 balls in 2 boxes.
If balls are placed randomly, which configuration is more likely?

# Microstates	Distribution
1	(4,0)
6	(2,2)
4	(1,3)

Table 9.1: Microstates and distribution of balls

Macrostate: collection of equivalent microstates. Equivalent means the same numbers of particles within each of the levels.
Example: the macrostate with 1L, 3R - (1,3)
We will use combinatorics to calculate:

1. The distribution function specifies the occupation of atomic states by identical particles, and thus the macrostate.

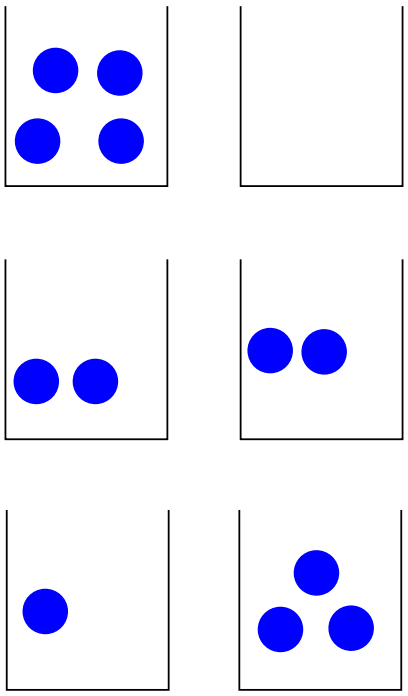


Figure 9.1

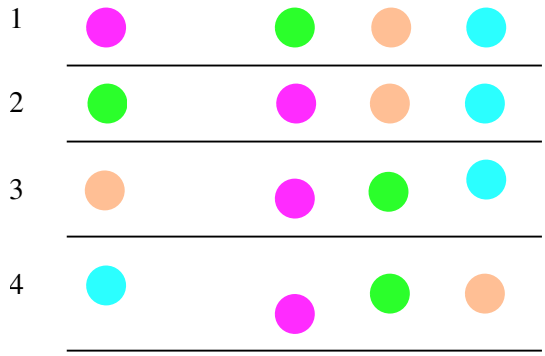


Figure 9.2: Macrostate (1,3)

2. Number of macrostates ≪number of microstates
3. If all microstates are equally likely, then some macrostates are much more likely. (e.g. (2,2))
   ⇒How do we see this?

How many microstates correspond to a given macrostate? N_0 : # of balls, r : # of boxes, n_i : balls in i^{th} box

$$\Omega = \frac{N_o!}{n_1!n_2!n_3!...n_r!} \qquad n! \equiv n(n-1)(n-2)...3\cdot 2\cdot 1$$
$$= \frac{N!}{\Pi_{i=1} n_i!}$$

Total number of possibilities is r^{N_o}, or $2^4 = 16$. There are 5 macrostates: (4,0) (1,3) (2,2) (3,1) (4,0)
Example: Four atoms in volume V. How many microstates are there, and how many macrostates?

- Divide each side into M compartments.
- Compartments are so small that M ≫4 and occupation is 0 or 1.

Number of Microstates	Distribution	
1	(4,0)	$\Omega_{4,0} = \frac{4!}{4!0!} = 1$
6	(2,2)	$\Omega_{2,2} = \frac{4!}{2!2!} = \frac{4\cdot3\cdot2\cdot1}{2\cdot1\cdot2\cdot1} = 6$
4	(1,3)	$\Omega_{1,3} = \frac{4!}{1!3!} = \frac{4\cdot3\cdot2\cdot1}{1\cdot3\cdot2\cdot1} = 4$

Table 9.2: Microstates corresponding to given macrostates

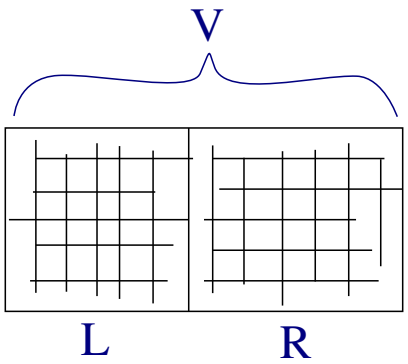


Figure 9.3

How many ways can we distribute 4 indistinguishable particles over 2M compartments?
Case I: Ω_I

$$\left(\begin{array}{c} 2M \\ 4 \end{array} \right) = \frac{(2M)!}{4!(2M-4)!} = \frac{2M\cdot(2M-1)(2M-2)(2M-3)}{4!} \approx \frac{(2M)^4}{4!}$$

Case II: How many configurations have all the particles on one side, i.e. M boxes?

$$\Omega_{II} = \left(\begin{array}{c} M \\ 4 \end{array} \right) \approx \frac{M^4}{4!}$$

Note that $\frac{\Omega_{II}}{\Omega_I} \approx \frac{1}{2^4} = \frac{1}{16}$, Ω_{II} is much less likely.
For N atoms, we would find $\frac{\Omega_{II}}{\Omega_I} = \frac{1}{2^N} \lll 1$
Even a state with 51% on the left would be highly improbable. $\frac{(1.98M)^N}{(2M)^N}$ for $N \sim N_A$ is vanishingly small.(Try plotting this).
100 - 3.790
1000 - 4×10^{-5}
10000 - 2×10^{-44}

9.1.1 Postulate of Equal Likelihood

All allowable microstates (energetically equivalent configurations) are equally likely.
$P_j = \frac{\Omega_j}{r^{N_0}}$is the probability of finding the system in macrostate j. The most probable macrostate is the one that represents the most microstates in Ω.

9.1.2 The Boltzmann Hypothesis (to be derived)

$S = k_B \ln \Omega$
$k_B = 1.38 \times 10^{-23} J/K$
Maximum in Ωcorresponds to a maximum in S.
How do we determine the most probable macrostate?
What are the conditions of equilibrium?

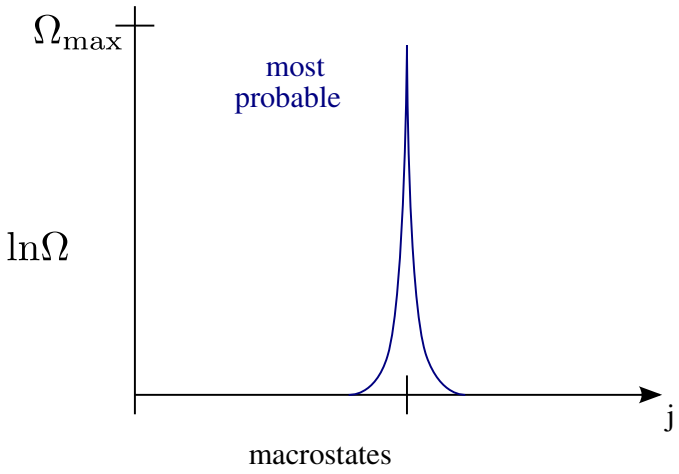


Figure 9.4

9.2 Conditions for equilibrium

We aim to minimize S, where $S = k \ln \Omega$, or

$$S = k_B \underbrace{\left(\frac{N_o!}{\Pi_{i=1}^{r} n_i!} \right)}_{\Omega}$$

What tools do we need to do so?
$\ln\left(\frac{x}{y}\right) = \ln x - \ln y$, $\ln(xy) = \ln x + \ln y$
$d \ln x = \frac{dx}{x}$
Stirling's Approximation: $\ln x! \approxeq x \ln x - x$
This is very accurate for large x. How large?

1. Express Ω as a sum over logarithms of occupation numbers.

$$\Omega = \frac{N!}{\Pi_{i=1}^{r} n_i!} \rightarrow \ln \Omega = \ln N! - \ln\left(\Pi_{i=1}^{r} n_i!\right)$$

Apply Stirling's Approximation, $\ln(\Pi x_i) = \sum \ln x_i$

$$\ln \Omega = \approx N \ln N - N - \sum_{i=1}^{r} (n_i \ln n_i - n_i)$$

Note also that $N = \sum_{i=1}^{r} n_i$ (summing occupation numbers gives the <u>total</u> particle number).
So

$$\ln \Omega \approx N \ln N - \cancel{N} - \sum_{i=1}^{r} n_i \ln n_i + \cancel{N}$$

$$= \sum_{i=1}^{r} n_i \ln N - \sum_{i=1}^{r} n_i \ln n_i$$

$\ln \Omega \approx \sum_{i=1}^{r} n_i (\ln N - n_i)$ Not too bad

2. Differentiate
$d \ln \Omega = \sum_{i=1}^{r} \left[dn_i (\ln N - \ln n_i) + n_i \left(\frac{dN}{N} - \frac{dn_i}{n_i} \right) \right]$

$= \sum_{i=1}^{r} (\ln N - \ln n_i) dn_i + \underbrace{\sum_{i=1}^{r} \left(\frac{n_i}{N}\right) \cancel{dN}}_{dN} - \underbrace{\cancel{\sum_{i=1}^{r} dn_i}}_{dN}$

$d\ln\Omega = -\sum_{i=1}^{r} \ln\left(\frac{n_i}{N}\right) dn_i$

3. Impose constraints

   For an isolated system, the total particle number and energy are constant as $\ln\Omega$ is maximized.

   $N = n_1 + n_2 + ...n_r = \sum_{i=1}^{r} n_i$, $dN = 0$

   $U = n_1\Sigma_1 + n_2\Sigma_2 + ... + n_r\Sigma_r = \sum_{i=1}^{r} n_i\Sigma_i$, $dU = 0$

   To impose constraints, we use Lagrange Multipliers.

   $d\ln\Omega + \lambda_1 dN + \lambda_2 dU = 0$

   $-\sum_{i=1}^{r} \ln\left(\frac{n_i}{N}\right) dn_i + \lambda_1 \sum_{i=1}^{r} dn_i + \lambda_2 \sum_{i=1}^{r} \Sigma_i dn_i = 0$

   $-\sum_{i=1}^{r} \left[\ln\left(\frac{n_i}{N}\right) dn_i + \lambda_1 - \lambda_2\Sigma_i\right] dn_i = 0 \Rightarrow$Gives r equations, e.g.

   $\ln\left(\frac{n_i}{N}\right) - \lambda_1 - \lambda_2\Sigma_i = 0 \rightarrow \frac{n_i}{N} = e^{\lambda_1} e^{\lambda_2\Sigma_i}$, $i = 1, 2, ..., r$

4. Use constraints to eliminate λ_1,λ_2

$\frac{n_i}{N} = e^{\lambda_1} e^{\lambda_2\Sigma_i}$
$\sum_{i=1}^{r} \frac{n_i}{N} = 1 \rightarrow \sum_{i=1}^{r} e^{\lambda_1} e^{\lambda_2\Sigma_i} = 1$
$e^{\lambda_1} = \frac{1}{\sum_{i=1}^{r} e^{\lambda_2\Sigma_i}} \rightarrow \frac{n_i}{N} = \frac{e^{\lambda_2\Sigma_i}}{\sum_{j=1}^{r} e^{\lambda_2\Sigma_j}} \equiv \frac{1}{Z} e^{\lambda_2\Sigma_i}$
The fractional occupation number of the i^{th}level.
So

$$\ln\Omega_i = -\sum_{i=1}^{r} n_i \left(\frac{n_i}{N}\right) = -\sum_{i=1}^{r} n_i \ln\left(\frac{e^{\lambda_2\Sigma_i}}{Z}\right)$$

$$= -\sum_{i=1}^{r} n_i(\lambda_2\Sigma_i - \ln Z)$$

$$\ln\Omega = -\lambda_2 \sum_{i=1}^{r} n_i\Sigma_i + \ln Z \sum_{i=1}^{r} n_i = -\lambda_2 U + N\ln Z$$

What is λ_2?
Consider response of system to the reversible absorption of an infinitesimal amount of heat δQ.
$d\ln\Omega = -\lambda_2 dU + \cancel{d(N\ln Z)}$
$dU = \delta Q_{rev} = TdS$
$d\ln\Omega = -\lambda_2 TdS$
So

$$S \propto \ln\Omega$$

$$S = \underbrace{-\frac{1}{\lambda_2 T}}_{k_B} \ln\Omega = k_B \ln\Omega \rightarrow \lambda_2 = -\frac{1}{k_B T}$$

$$\frac{n_i}{N} = \frac{e^{-\Sigma_i/kT}}{\sum_{i=1}^{r} e^{-\Sigma_i/kT}} \equiv \frac{1}{Z} \underbrace{e^{-\Sigma_i/kT}}_{\text{Boltzmann factor}}$$

Partition function

We can calculate macroscopic properties from the partition function.
Recall $\ln\Omega = -\lambda_2 U + N\ln Z$, $S = k\ln\Omega$
$(k\ln\Omega =)\ S = -\frac{U}{T} + k_B N\ln Z$
$-kT\ln Z = U - TS \equiv F$ Hemholtz Free Energy!
So

$$\boxed{F = -kTN\ln Z}$$

$$\frac{F}{N} = -kT\ln Z$$

Also, $S = -\left(\frac{\partial F}{\partial T}\right)_V = Nk\ln Z + NkT\left(\frac{\partial \ln Z}{\partial T}\right)_V$
Then

$$U = F + TS = NkT^2\left(\frac{\partial \ln Z}{\partial T}\right)_V$$

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = 2NkT\left(\frac{\partial \ln Z}{\partial T}\right)_V + NkT^2\left(\frac{\partial^2 \ln Z}{\partial T^2}\right)_V$$

Note that the pressure dependence of Z is needed to calculate V, H, G, C_P.

10 Noninteracting (Ideal) Gas

Conditions for ideal gases

- No interactions (forces/potentials) →only kinetic energy.
  $\sum = \frac{1}{2}mv^2 = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2)$
- The energy distribution of free particles is continuous.
  $\sum \to \int$
- For convenience, assume box of dimensions $L \times L \times L = V$

$$Z = \underbrace{\int_0^L \int_0^L \int_0^L dxdydz}_{V} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2kT}(v_x^2+v_y^2+v_z^2)} dv_x dv_y dv_z$$

$$= V \int_{-\infty}^{\infty} \overbrace{e^{-\frac{1}{2kT}mv_x^2}}^{\text{Gaussians}} dv_x \int_{-\infty}^{\infty} e^{-\frac{1}{2kT}mv_y^2} dv_y \int_{-\infty}^{\infty} e^{-\frac{1}{2kT}mv_Z^2} dv_z$$

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\tfrac{\pi}{\alpha}} \text{ (for } \alpha > 0\text{)}$$

$$\overbrace{e^{-\frac{1}{2kT}m v_x^2}}^{\alpha}$$

$$\therefore Z = V\left(\frac{2\pi kT}{m}\right)^{3/2}$$

Now we can easily derive many functions.

$$\ln Z = \ln V + \frac{3}{2}\ln\left(\frac{2\pi k}{m}\right) + \frac{3}{2}\ln T$$

,

$$\frac{\partial \ln Z}{\partial T} = \frac{3}{2}\frac{1}{T}$$

$$\Rightarrow F = -NkT\ln Z = -NkT\left[\ln V + \frac{3}{2}\ln\left(\frac{2\pi kT}{m}\right)\right]$$

$$U = NkT^2\left(\frac{\partial \ln Z}{\partial T}\right)_V = NkT^2\left(\frac{3}{2}\frac{1}{T}\right) = \frac{3}{2}NkT$$

$\boxed{\frac{U}{N}=\frac{3}{2}\mathrm{kT}}$ Energy per particle

$C_V = \left(\frac{\partial U}{\partial T}\right)_V = \frac{3}{2}NK = \frac{3}{2}\left(\frac{N}{N_A}\right)kN_A = \frac{3}{2}nR$

$\boxed{\frac{C_V}{n}=\frac{3}{2}\mathrm{R}}$ n: moles

10.1 Ideal gas Law

$$dF = -SdT - PdV + \delta W'$$

$$-P = \left(\frac{\partial F}{\partial V}\right)_T$$

$$P = -NkT\left(\frac{\partial \ln Z}{\partial V}\right)_T = NkT\left[\frac{\partial}{\partial V}\left(\ln V + \frac{3}{2}\ln\left(\frac{2\pi k}{m}\right) + \frac{3}{2}\ln T\right)\right]_T$$

$$\boxed{\text{P=NkT}\cdot\tfrac{1}{V}=\tfrac{nRT}{V}}$$

- Note that we have ignored interactions between particles/atoms. Attractive interactions lead to, e.g. condensation.
- We have also ignored repulsive interactions, i.e. atoms occupy finite volumes.

Including these leads to the van der Waals equation:

$$\left(P + \frac{a}{V^2}\right)(V - b) = RT$$

11 Diatomic Gases

Diatomic gases possess rotational and vibrational degrees of freedom.
The kinetic energy of a gas molecule with p degrees of freedom is
$\Sigma = \sum_{j=1}^{P} b_j v_j^2$, and $Z = V\Pi_{j=1}^{P}\sqrt{\frac{\pi kT}{b_j}}$
(e.g. $b = \frac{m}{2}$ for translations) For rotations $v \to \omega$, $M \to I$
Surprisingly, U and C_V do not depend on b_j:

$$\left(\frac{\partial \ln Z}{\partial T}\right)_V = \frac{P}{2T} \to U = NkT^2\left(\frac{P}{2T}\right) = \frac{P}{2}NkT$$

Get $\frac{1}{2}RT$ for each degree of freedom
$C_V = \frac{P}{2}Nk = \frac{P}{2}nR$
Observations:
Heat capacity measurement reveals the number of degrees of freedom.

- Monatomic: 3 translational
- Heteronuclear Diatomic: 3 translational + 2 rotational + 2 vibrational

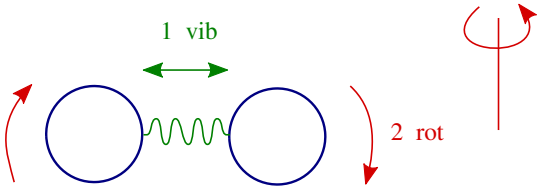


Figure 11.1: Diatomic gas

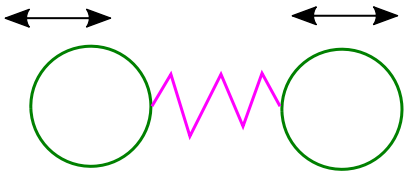


Figure 12.1: Vibration of atoms in harmonic potential

Figure 12.2: Binding to six nearest neighbors

12 The Einstein Model of a Crystal

How is the discrete nature of matter manifested in easily observable quantities?
Consider atoms in a simple cubic lattice

- Bounded to six nearest neighbors (vibrating in harmonic potential) $F = -kx$, $\omega = \sqrt{\frac{k}{m}}$
- 3 bonds/atom, 3N bonds in crystal

By solving the Schrödinger equation for the harmonic oscillator potential, one can show that
$\sum_i = (i + \frac{1}{2})\hbar\omega = (i + \frac{1}{2})h\nu \; i = 0, 1, 2...$
h: Planck's constant , ν : frequency , $\hbar : \frac{h}{2\pi}$
Energy is stored in vibration of atoms. To analyze, calculate the partition function for the vibrations:

$$Z = \sum_{i=0}^{n} e^{-\Sigma_i/kT} = \sum_{i=0}^{n} e^{-(i+\frac{1}{2})h\nu/kT} = e^{-\frac{1}{2}\frac{h\nu}{kT}} \sum_{i=0}^{n} e^{-ih\nu/kT}$$

Here n is the number of energy levels. If n is large, we can take the limit $n \to \infty$ and use
$\therefore \sum_{i=0}^{\infty} e^{-ih\nu/kT} = \frac{1}{1-e^{-h\nu/kT}}$, $Z = \frac{e^{-1/2\frac{h\nu}{kT}}}{1-e^{-h\nu/kT}}$

12.1 Thermodynamic Properties of a Crystal

Vibrations along each of three bonding axes contribute to energy:
$F = -3NkT \ln Z = 3NkT \left[\frac{h\nu}{2kT} + \ln(1 - e^{-h\nu/kT})\right]$
$= \frac{3}{2}Nh\nu + 3NkT \ln(1 - e^{-\frac{h\nu}{kT}})$
Note that the crystal has energy even at zero T!
$U = -\frac{\partial}{\partial\beta} \ln Z = kT^2 \frac{\partial \ln Z}{\partial T}$ $(\beta = kT)$ (per particle?)
$U = 3NkT^2 \frac{\partial \ln 2}{\partial T} = 2NkT^2 \frac{\partial}{\partial T}\left[-\frac{1}{2}\frac{h\nu}{kT} - \ln(1 - e^{-h\nu/kT})\right]$

$$U = \frac{3}{2}Nh\nu\left[\frac{1 + e^{-h\nu/kT}}{1 - e^{-h\nu/kT}}\right]$$

How does this change with temperature?
The heat capacity, $C_V = \frac{\partial U}{\partial T}$

$$C_V = \frac{3}{2}Nh\nu \cdot \frac{h\nu}{kT^2}e^{-h\nu/kT} \times \frac{(1 - e^{-h\nu/kT}) + (1 + e^{-h\nu/kT})}{(1 - e^{-h\nu/kT})^2}$$

$$C_V = 2Nk\left(\frac{h\nu}{kT}\right)^2 \frac{e^{-h\nu/kT}}{(1 - e^{-h\nu/kT})^2}$$

Classical limit $(3R\frac{J}{mol-K})$
The Einstein model captures the qualitative behavior of the heat capacity very well. (But something is missing).
Note here regarding information derived from heat capacity.

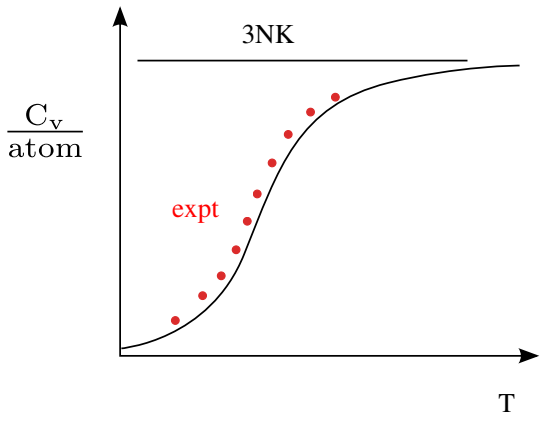


Figure 12.3: Heat capacity in an atom as a function of temperature

13 Unary Heterogeneous Systems

Outcomes for this section:

1. State the differential of the Gibbs free energy G as a function of S, T, V, and P.
2. State the conditions for two phases to coexist in equilibrium.
3. Explain how isobaric section of the chemical potential $\mu(T,\ P)$ as a function of temperature and pressure can be calculated from knowledge of the heat capacity and the entropy at room temperature.
4. Given $\mu(T,\ P)$ for multiple phases, determine the equilibrium phase at a given temperature.
5. Derive the Clausius-Clapeyron relation: express the general condition for determining the region of two-phase coexistence.
6. Calculate the ΔH given the $C_P(T)$, assuming a negligible contribution from thermal expansion.
7. Calculate $\Delta V^{\alpha\to\beta}$ assuming ideal gas behavior.
8. Integrate $\frac{dP}{dT} = \frac{\Delta S}{\Delta V}$ with the assumption that ΔS and ΔV depend only weakly on the temperature.
9. Given the $P-T$ diagram, qualitatively sketch $G(T)$ for a specified region.
10. Given the $P-T$ diagram, specify the sign of $\Delta V^{\alpha\to\beta}$.
11. State Trouton's rule and justify its validity qualitatively.
12. Sketch the $P-T$ diagram for a one-component system.

- Unary: single chemical component
- Homogeneous: 1 phase
- Heterogeneous: 2 phase
- Phases: Gas, Liquid, Solid; allotropes are distinct solid phases
- Phase boundary: limit of phase stability
  - multiple phases co-exist in equilibrium at a phase boundary
  - Boundary defined by Clasius - Clapeyron Equation
- Phase transformation: solid $\to$ liquid, liquid $\to$vapor, allotropic process of phase transformation determines microstructure

13.1 Phase Diagrams

Phase Diagram of Unary Heterogeneous System

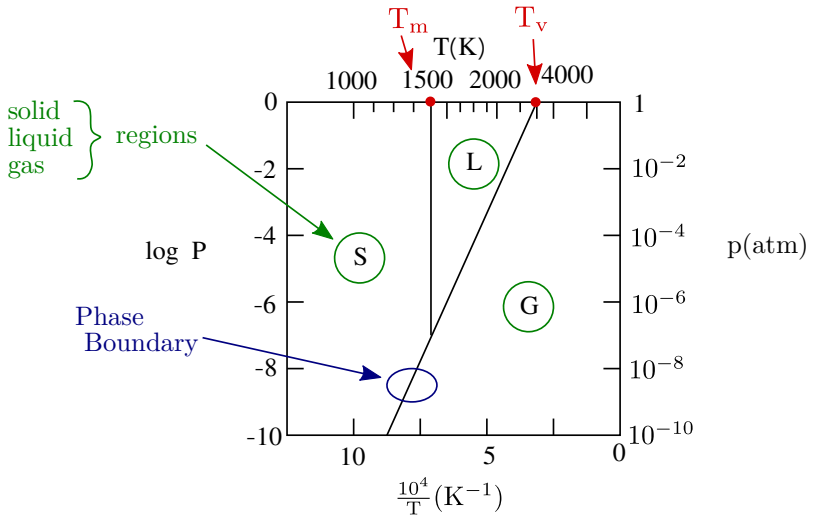


Figure 13.1: Phase Diagram Cu.

- Phase →solid, liquid, gas regions
- Unary → single component
- Heterogeneous → more than one phase
- (Cu) →component

The phase boundary defines the limits of phase stability. Phases co-exist at a phase boundary. A phase transformation occurs when crossing phase boundaries i.e. $s \to \ell, \ell \to v, s \to v$.

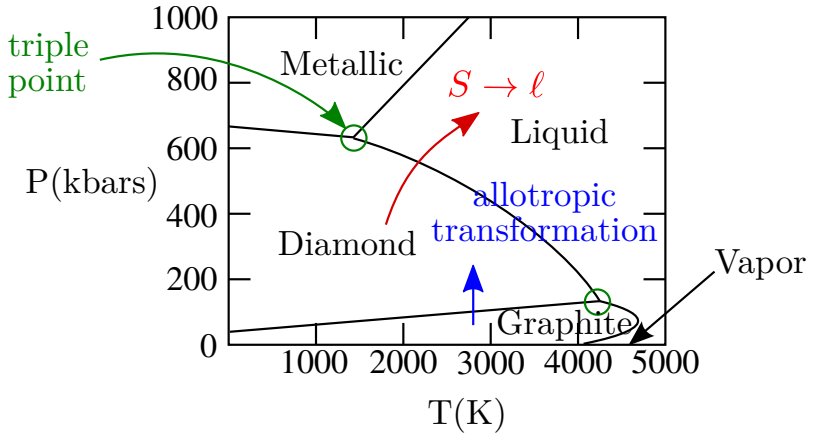


Figure 13.2: Phase Diagram of Unary Heterogeneous System (C)

13.2 Chemical Potential Surfaces

To find $\mu(T, P)$, integrate $d\mu = dG = -SdT + VdP$
knowns: C_P, expansion coefficient, compressibility depends on phase
Then repeat for each phase (solid α,solid β, liquid, etc.)
$\mu^L(T, P)$ and $\mu^S(T, P)$ (two surfaces) intersect at a line defined by $\mu^L = \mu^S$.

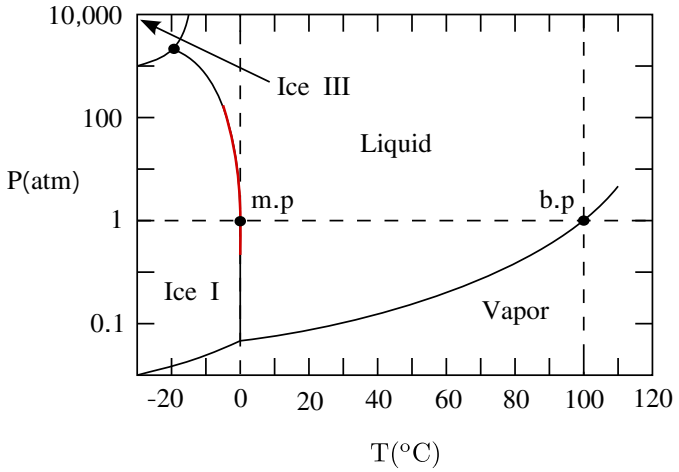


Figure 13.3: Componds can be treated as unary systems (H_2O)

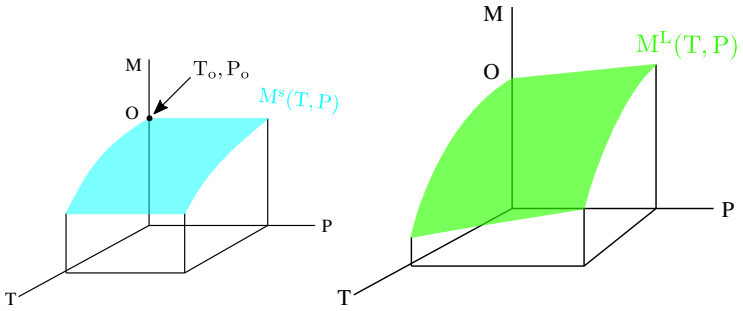


Figure 13.4: Solid α and liquid L

- μ^L and μ^S must be computed from same reference state (e.g. $\mu^S(T_o, \ell_o)$
- Equilibrium conditions are met at intersection
- $T^\alpha = T^L$, $P^\alpha = P^L$, $\mu^\alpha = \mu^L$

AB: solid + liquid
COD: solid +gas
EOF: liquid + gas
O: triple point S+ℓ+S
Relation between chemical potential and the Gibb's free energy (for unary systems).
G: intensive molar Gibbs free enrgy
$G' = nG$: Extensive
$dU' = TdS' - PdV' + \mu dn_1$, $\mu = \left(\frac{\partial u'}{\partial n}\right)_{S',V'}$
$G' = U' + PV' - TS'$
$dG' = dU' + PdV' + V'dP - TdS' - S'dT$
Susbstitute dU'
$dG' = \cancel{TdS'} - \cancel{PdV'} + \mu dn + \cancel{PdV'} + V'dP - \cancel{TdS'} - S'dT$
$dG' = -S'dT + V'dP + \mu dn$
$\mu = \left(\frac{\partial G'}{\partial n}\right)_{T,P}$
$\mu = \left(\frac{\partial G'}{\partial n}\right)_{T,P} = \left(\frac{\partial}{\partial n}(nG)\right)_{T,P} = G\left(\frac{\partial}{\partial n}\right)_{T,P} = G$

13.3 Calculation of $\mu(T,P)$ Surfaces

$d\mu^\alpha = dG^\alpha = -S^\alpha dT + V^\alpha dP$
Approach: compute isobaric sections by integrating $-S^\alpha dT$.

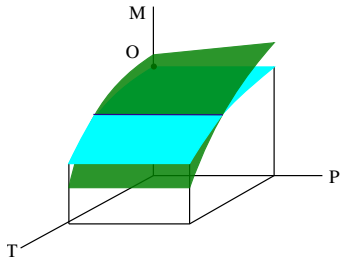


Figure 13.5: Intersection of $\mu^L(T,P)$ and $\mu^S(T,P)$

$$dS_P^\alpha = \frac{C_P^\alpha}{T}dT \Rightarrow S^\alpha(T) = S^\alpha(298) + \int_{298}^{T}\frac{C_P^\alpha(T')}{T'}dT'$$

$$\mu^\alpha(T) = \mu^\alpha(298) - \int_{298}^{T} S^\alpha(T')dT'$$

$$\mu^\alpha(T) = \mu^\alpha(298) - \int_{298}^{T}\left[S^\alpha(298) + \int_{298}^{T''}\frac{C_P^\alpha(T')}{T}dT'\right]dT''$$

To compare $\mu^\alpha$with μ^L, we need to calculate μ^L from the same reference point.
Connecitng $\mu^\alpha$with μ^L at the melting point T_m :

$$\mu^L(T) - \mu^\alpha(298) = \underbrace{\left[\mu^L(T) - \mu^L(T_m)\right]}_{\text{integral of } -S^L \text{from T to } T_m} + \underbrace{[\mu^\alpha(T_m) - \mu^\alpha(298)]}_{\text{just calculated this integral}}$$

$$\mu^L = \mu^L(T_m) - \int_{T_{Mm}}^{T}\left[S^L(T_m)^* + \int_{T_m}^{T''}\frac{C_P^L(T')}{T'}dT'\right]dT''$$

Plot $G(T) - G^\alpha(298)$ versus T for each phase α, L i.e. plot of G^αbegins at 0, do example where P=1 bar.
* Note the entropy changes with the phase change
$S^L(T_m) = S^\alpha(T_m) + \Delta S^{\alpha\to L}(T_m)$
ΔS : latent heat

$$S^L(T_m) = \underbrace{S^\alpha(T_m)}_{S_{298}^\alpha + \int_{298}^{T_m}\frac{C_P^\alpha(T)dT}{T}} + \underbrace{\Delta S^{\alpha\to L}(T_m)}_{\text{latent heat}}$$

$$G^L(T) - G^\alpha(298) = G^L(T) - G^L(T_m) + G^\alpha(T_m) - G^\alpha(298) =$$

$$= \int_{T_m}^{T} - \left\{\left[\underbrace{\overbrace{S^\alpha(298 + \int_{298}^{T_m}\frac{C_P^\alpha(T')}{T'}dT'}^{S^\alpha(T_m)} + \Delta S^{\alpha\to L}(T_m)}_{S^L(T_m)}\right] + \underbrace{\int_{T_m}^{T''}\frac{C_P^L(T')}{T'}dT'}_{S^L(T>T_m)}\right\}$$

$$+(-)\int_{298}^{T_m}\left[S^\alpha(298) + \int_{298}^{T''}\frac{C_P^\alpha(T)}{T'}dT'\right]dT''$$

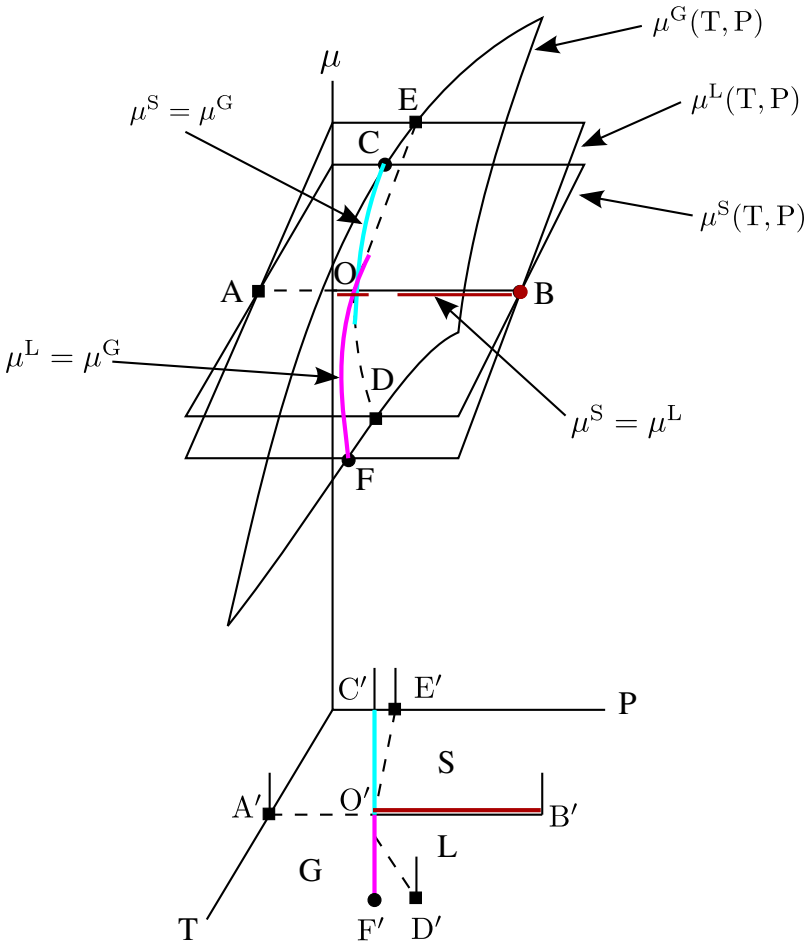


Figure 13.6: Superposition of Solid, Liquid, and Gas phase μ

13.4 Clasius-Clapeyron Equation

Goal: Determine two-phase coexistence lines in order to complete the P-T phase diagram
Areas: single phase
Lines: 2-phase
Intersections: 3 phases (triple point)
Consider co-existence of two phases in equilibrium:
$T^\alpha = T^\beta$, $P^\alpha = P^\beta$, $\mu^\alpha = \mu^\beta$
Phase changes change the amounts of two phases:
$d\mu^\alpha = -S^\alpha dT^\alpha + V^\alpha dP^\alpha$
$d\mu^\beta = -S^\beta dT^\beta + V^\beta dP^\beta$
But we require equilibrium:

$$dT^\alpha = dT^\beta \quad dP^\alpha = dT^\beta \quad d\mu^\alpha = d\mu^\beta$$

$$-S^\alpha dT + V^\alpha dP = -S^\beta dT + V^\beta dP$$

$$(S^\beta - S^\alpha)dT = (V^\beta - V^\alpha)dP$$

$$\Delta S^{\alpha\to\beta} dT = \Delta V^{\beta\to\alpha} dP$$

$\Delta S^{\alpha\to\beta}$: difference in molar entropy

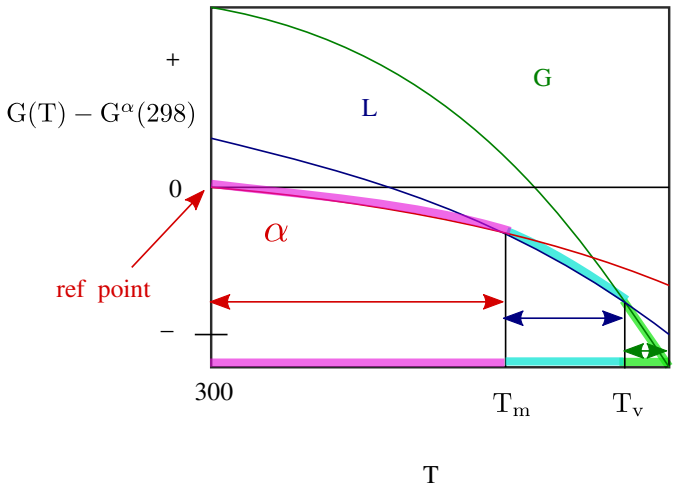


Figure 13.7: Determining the equilibrium state (at fixed P)

$\Delta V^{\alpha\to\beta}$: difference in molar volume

$\boxed{\frac{dP}{dT}=\frac{\Delta S^{\alpha\to\beta}}{\Delta V^{\alpha\to\beta}}}$ Classius-Claperyon Equation

$$\text{Slope} = \frac{\text{Entropy Change}}{\text{Volume Change}}$$

But how do we measure entropy change?
Instead, we measure the heat of fusion of vaporization isobarically under reversible conditions.
So, $Q^{\alpha\to\beta} = \Delta H^{\alpha\to\beta}$
In equilibrium,

$$G^\alpha = G^\beta \quad H^\alpha - TS^\alpha = H^\beta - TS^\beta \quad (T^\alpha = T^\beta$$

So $\Delta S^{\alpha\to\beta} = \frac{\Delta H^{\alpha\to\beta}}{T}$, and

$$\frac{dP}{dT} = \frac{\Delta H^{\alpha\to\beta}}{T\Delta V^{\alpha\to\beta}}$$

Can be integrated to get $P(T)$ for 2-phase coexistence.
We need $\Delta H(P,T)$, $\Delta V(P,T)$ from Chapter 4.

Step 1: Find $\Delta H(T,P)$

$$d(\Delta H^{\alpha\to\beta}) = d(H^\beta - H^\alpha) = dH^\beta - dH^\alpha$$

$$dH = C_P dT + V(1 - T_\alpha)dP$$

$$d\Delta H^{\alpha\to\beta} = \Delta C_P dT + \underbrace{\Delta[V(1 - T_\alpha)dP}$$

negligible for pressures up to $\sim 10^{-5}$bar

To find ΔH, we need the temperature dependence of the heat capacity. $C_P(T)$ is often given empirically as $C_P = a + bT + \frac{c}{T^2} + dT^2$ (See Appendix B,E)

$$\Delta C_P = \Delta a + \Delta bT + \frac{\Delta c}{T^2} + \Delta dT^2$$

Now we can integrate to get ΔH.

Step 2: Find $\Delta V^{\alpha\to\beta}(T,P) \equiv V^{\beta}(P,T) - V^{\alpha}(P,T)$ First consider the case of β as a vapor phase, and α as liquid or solid.

- $V^{\beta} \gg V^{\alpha}$
- Assume ideal gas behaviour.

$\Delta V^{\alpha\to\beta} = V^{\beta} - V^{\alpha} \approx V^{\beta} = \frac{RT}{P}$
Then for sublimation or vaporization to phase β,
$\frac{dP}{dT} = \frac{\Delta H^{\alpha\to\beta}}{T\Delta V^{\alpha\to\beta}} = \frac{\Delta C_P dT}{T\cdot\frac{RT}{P}} = \Delta C_P \frac{P}{RT^2} dT$ small, positive slope
Phase boundaries can be calculated based on knowledge of C_P. (or vice versa, as we will see)
For solid →solid transformations, (β is the stable phase at high T)

$$\frac{dP}{dt} = \frac{\overbrace{\Delta C_P^{\alpha\to\beta}(T)\, dT}^{\text{positive } (\Delta H,\Delta S>0}}{\underbrace{T\Delta V}_{\text{determines sign of slope}}}$$

ΔV is usually positive but can be negative (e.g. H_2O)
ΔV ~ constant over few 10's of atmosphere.
An approximate analysis of solid-solid phase boundary:
$\Delta S, \Delta H, \Delta V$ are determined by C_P, α, β, but suppose the variation with temperature is weak.
Since $\frac{dP}{dT} = \frac{\Delta S}{\Delta V}$ = constant, $P - P_o \approx \frac{\Delta S}{\Delta V}(T - T_o) \approx \frac{\Delta H}{\Delta V}\left(\frac{T-T_o}{T_o}\right)$
The slope $\frac{dP}{dT}$ is generally large since ΔV is small.

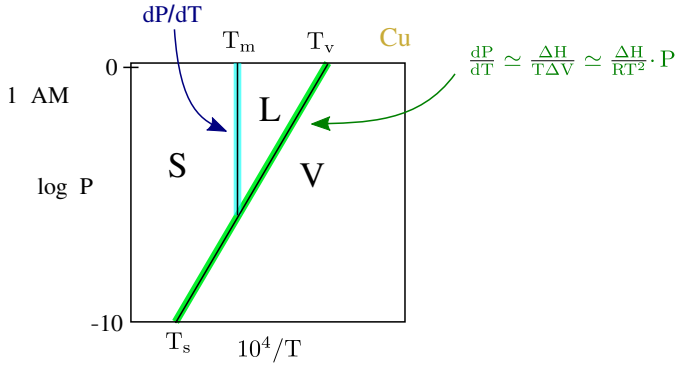


Figure 13.8

$$\frac{dP}{dT} \approx \frac{\Delta H}{T\Delta V} \approx \frac{\Delta H}{RT^2} \cdot P$$

$$\frac{dP}{P} \approx \frac{\Delta H(T)}{RT^2} dT$$

If we assume $\frac{d}{dT}\Delta H = 0$,

$$\ln\left(\frac{P}{P_o}\right) = -\frac{\Delta H}{R}\left(\frac{1}{T} - \frac{1}{T_o}\right)$$

$$\left|\frac{dP}{dT}\right|^{S\to L} > \left|\frac{dP}{dT}\right|^{S\to V}$$

Justification for Approximations Consider the enthalpy of transformation $d\Delta H^{\alpha\to\beta}(T) = \Delta C_p(T)dT$, and integrate

$$\Delta H(T) = \underbrace{\Delta H(T_o)}_{\text{Typically} \sim 100\text{kJ}} + \underbrace{\int_{T_o}^{T} \Delta C_P(T)dT}_{\text{Typically few kJ over few 100K}}$$

$\Delta H(T_o)$: enthalpy change at the transformation temperature.
Heat to drive phase change is larger than that needed to change the temperature.
Thus, $\Delta H(T) \approx \Delta H(T_o)$
For vaporization,

$$\ln\left(\frac{P}{P_o}\right) \approx -\frac{\Delta H(T_o)}{R}\left(\frac{1}{T} - \frac{1}{T_o}\right)$$

Trouton's Rule ΔS_{vap} is approximately the same for all materials. $\Delta S = \frac{\Delta H}{T} \to$ materials with high boiling points have larger ΔH_{vap}.

$$\ln\left(\frac{P}{P_o}\right) = -\left(\frac{\Delta H}{R}\right)\frac{1}{T} + \left(\frac{\Delta H}{RT_o}\right)$$

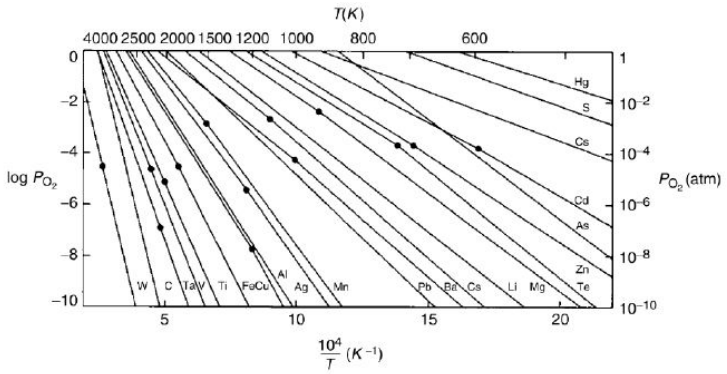


Figure 13.9

14 Open Multicomponent Systems

Outcomes for this section:

1. Given sufficient information about the dependence of a total property B or ΔB versus composition, find the partial molal properties as a function of composition via calculations of derivatives dB or $d\Delta B$.
2. Given the partial molal properties as a function of composition, calculate the total properties of the system via integration of the derivatives.
3. Given the differentials of state functions, write partial derivatives that define the partial molal properties.
4. What is B_k^0?
5. What is $\mu_k - \mu_k^0$ for (a) a component behaving ideally, and (b) in general?
6. Define the activity of a component.
7. Given $\mu_k(T, P, X_k)$, calculate changes in the partial molal properties.
8. State and justify/explain Raoult's law.
9. State and justify/explain Henry's law.
10. Explain how the activity coefficient γ can be used to describe the departure from ideal behavior in terms of "excess" quantities.
11. Define a "regular" solution and calculate the partial molal properties and total properties versus composition for binary mixtures.
12. Describe the driving forces for mixing and their origins in statistical mechanics.
13. Interpret changes in state functions with mixing in terms of ideal behavior and departures from ideality.

14.1 Partial Molal Properties of Open Multicomponent Systems

Volume V' is an extrinsic quantity that depends on amount of each component:
$V' = V'(T, P, n_1, n_2, ...n_c)$

$$C: \text{ \# components} \quad x_c: \text{ mole fraction} \quad n_k: \text{ \# moles of component k}$$

Then

$$dV' = \left(\frac{\partial V'}{\partial T}\right)_{P,n_m} dT + \left(\frac{\partial V'}{\partial P}\right)_{T,n_k} dP + \sum_{k=1}^{c} \left(\frac{\partial V'}{\partial n_k}\right)_{T,P,n_o \neq n_k} dn_k$$

We previously defined coefficients of dP, dT
For convenience, define $\bar{V}_K \equiv \left(\frac{\partial V'}{\partial n_k}\right)_{T,P,n_j \neq n_k}$ as partial molal volumes.
We can do the same for any other extensive property $B'(U', S', F', G', H')$

$$dB' = MdT + NdP + \sum_{k=1}^{r} \bar{B}_k dn_k$$

$\bar{B}_K \equiv \left(\frac{\partial B'}{\partial n_k}\right)_{T,P,n_j \neq n_k}$ $k = 1, 2, > ... <$
Consider a system (or sample) that we create by adding various amounts of the components sequentially (holding others constant).
If $dT = dP = 0$, $dK'_{T,P} = \sum_{k=1}^{C} \bar{V}_k dn_k$, which we could integrate.
But $\bar{V}_k$ depends on the composition, which changes continuously.

However, we can calculate changes in state functions using the simplest possible path: add all components simultaneously and in the proportions of the final mixture. →Intensive properties (T, P, X_k) are constant

$$V' = \sum_{k=1}^{C} \bar{V}_k dn_k = \sum_{k=1}^{C} \bar{V}_k \int_0^{n_k} dn_k = \sum_{k=1}^{C} \bar{V}_k n_k$$

and

$$B' = \sum_{k=1}^{C} \bar{B}_k n_k$$

Interpretation: We can calculate how much each component contributes to a given extensive quantity.

Gibbs - Duhem Equation: Evaluating Differentials

$$dB' = d\left[\sum_{k=1}^{C} (\bar{B}_k n_k)\right] = \sum_{k=1}^{c} d(\bar{B}_k n_k)$$

Apply product rule:
$dB' = \sum_{k=1}^{C} \bar{B}_k dn_k + \sum_{k=1}^{C} n_k d\bar{B}_k$, but dP, dT are held constant, and

$$dB' = \underbrace{MdT + NdP}_{0} + \sum_{k=1}^{C} \bar{B}_k dn_k$$

$$\rightarrow \sum_{k=1}^{C} n_k d\bar{B}_k = 0$$

Gibbs Duhem equation relates the different partial molal properties. To be shown: given one, the others can be computed. For two components: $n_1 d\bar{B}_1 + n_2 \bar{B}_2 = 0$

14.2 The Mixing Process: Calculating changes in energy

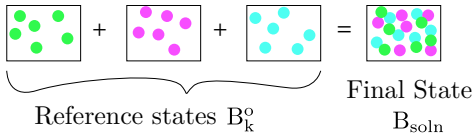


Figure 14.1: Mixing of three states into $B_{so\,ln}$

Reference and Final States have the same T and P.
B_k^o : value of property B for pure component, per mole.
$B'^o = \sum_{k=1}^{C} B_k^o n_k$: total value of B
$\Delta B'_{mix} - B'_{so\,ln} - B'^o$

$$\Delta B'_{mix} = \sum_{k=1}^{C} \bar{B}_k n_k - \sum_{k=1}^{C} B_k^o n_k = \sum_{C=1}^{k} \overbrace{(\bar{B}_k - \dot{B_k})}^{\Delta \bar{B}_k} n_k$$

$$\Delta B'_{mix} = \sum_{k=1}^{C} \Delta \bar{B}_k n_k$$

per mole

$$\Delta B = \frac{\Delta B'}{n_T} = \sum_{k=1}^{C} \Delta \bar{B}_k \frac{n_k}{n_T} = \sum_{k=1}^{C} \Delta \bar{B}_k X_k$$

How does this vary with changes in composition?

$$\Delta B'_{mix} = \sum_{k=1}^{C} \Delta \bar{B}_k n_k$$

$$d(\Delta B') = \sum_{k=1}^{C} (\bar{B}_k dn_k - B^o_n dn_k) = \sum_{k=1}^{C} \Delta \bar{B}_k dn_k$$

Alternatively, we could have written

$$d(\Delta B') = \sum_{k=1}^{C} (\Delta \bar{B} dn_k + n_k d\Delta \bar{B}_k) \rightarrow \sum_{k=1}^{C} n_k d\Delta \bar{B}_k = 0$$

We can also express these per mole:

$$d\Delta B_{mix} = \sum_{k=1}^{C} \Delta \bar{B}_k dX_k$$

$$\sum_{k=1}^{C} X_k d\Delta \bar{B}_k = 0$$

How do we find molal properties?

1. Measurement of total property B or ΔB versus composition.
2. Measurements of PMP $\bar{B}_k$versus composition.

We will first apply to two component systems.

$$B = \bar{B}_1 x_1 + \bar{B}_2 x_2, dB = \bar{B}_1 dx_1 + \bar{B}_2 x_2$$

$$dx_1 = -dx_2$$

$$x_1 + x_2 = 1\ , x_2 = 1 - x_1$$

$$dB = (\bar{B}_1 - \bar{B}_2) dX_1$$

$$\tfrac{dB}{dX_1} = \bar{B}_1 - \bar{B}_2$$

$$B = \bar{B}_1 x_1 + \left(\bar{B}_1 - \tfrac{dB}{dX_1}\right)(1 - X_1) = \bar{B}_1 - \tfrac{dB}{dX_1} + \tfrac{dB}{dX_1} X_1$$

Solve for $\bar{B}_1$

$$\bar{B}_1 = B + (1 - X_1)\tfrac{dB}{dX_1} \text{ and } \bar{B}_2 = B + (1 - X_2)\tfrac{dB}{dX_2}$$

Note also that
$\Delta B_1 = \Delta B + (1 - X_1)\frac{d\Delta B}{dX_1}$; $\Delta B_2 = \Delta B + (1 - X_2)\frac{d\Delta B}{dX_2}$ from Daltoff 8.2.1
Problem 8.3: Given behavior of volume change upon mixing $\Delta V_{mix} = 2.7 X_1 X_2^2$ $(L??/mol)$, derive expressions for partial molal volumes of each component as a function of composition.
$\Delta V_{mix} = 2.7 X_1 X_2^2 = a X_1 X_2^2\ a = 2.7$
To take derivative, express in terms of X_2 :
$\Delta V_{mix} = a(1 - X_2) X_2^2$
$\frac{d\Delta V_{mix}}{dX_2} = 2aX_2 - 3aX_2^2 = -\frac{d\Delta V_{mix}}{dX_1}$
Now use relations derived in lecture:
$\Delta \bar{B}_1 = \Delta B_{mix} + (1 - X_1)\frac{d\Delta B_{mix}}{dX_1}$ $(B \rightarrow V)$

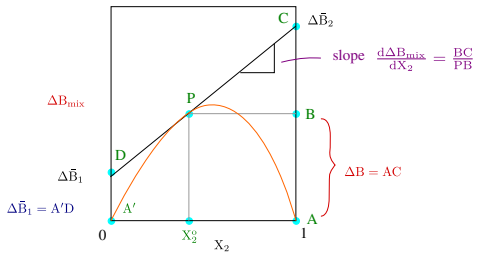


Figure 14.2

14.3 Graphical Determination of Partial Molal Properties

$\Delta\bar{B}_2 = \Delta B_{mix} + (1 - X_2)\frac{d\Delta B_{mix}}{dX_2}$
$\Delta\bar{B}_2 = AB + PB \cdot \frac{BC}{PB} = AB + BC = AC$
Evaluations of PMPs: given $\bar{B}_1(X) \rightarrow \bar{B}_2(X_1)$, B
Gibbs - Duhem: $X_1 d\bar{B}_1 + X_2 d\bar{B}_2 = 0$, $d\bar{B}_1 = -\frac{X_2}{X_1} d\bar{B}_2$
$\bar{B}_1 = \bar{B}_1(x = 0) + \int_0^{X_2} d\bar{B}_1 = B_1^0 + \int_0^{X_2} \left(\frac{-X_2}{X_1} d\bar{B}_2\right)$
$d\bar{B}_2 = \frac{d\bar{B}}{dX_2} dX_2$
$\therefore$Given $\bar{B}_2(X_2)$, we can compute $\bar{B}_1(X_2)$.
Then we can find $B = \bar{B}_1 X_1 + \bar{B}_2 X_2$
Example 8.2:
Given $\Delta\bar{H}_2$, find $\Delta\bar{H}_1$ and ΔH.
$\Delta\bar{H}_2 = aX_1^2$ Compute total derivative:

$$\frac{d\Delta\bar{H}_2}{dX_2} = 2aX_1 \underbrace{\frac{dX_1}{dX_2}}_{\hookrightarrow = -1} = -2aX_1$$

$$X_1 + X_2 = 1 dX_1 + dX_2 = 1$$

$$dX_1 = -dX_2$$

$$\Delta\bar{H}_1 = -\int_{X_2=0}^{X_2} \frac{X_2}{X_1} \frac{d\Delta\bar{H}_2}{dX_2} dX_2 = -\int_{X_2=0}^{X_2} \frac{X_2}{X_1}(-2aX_1) dX_2$$

$$\Delta\bar{H}_1 = 2a \int_0^{X_2} X_2 dX_2 = aX_2^2$$

$$\Delta H = X_1\Delta\bar{H}_1 + X_2\Delta\bar{H}_2 = X_1(aX_2^2) + X_2(aX_1^2) = aX_1X_2 \underbrace{(X_1 + X_2)}_{=1}$$

$$\Delta H = aX_1X_2$$

Relationships between partial molal properties: (F',G',H',S',U', their derivative, and coefficient relations)
Apply differential operator $\left(\frac{\partial}{\partial n_k}\right)_{P,T,n_j}$ to extensive quantity:
Example: Hemholtz free energy $F' = U' - TS'$
Differentiating,
$\left(\frac{\partial F'}{\partial n_k}\right)_{T,P,n_j} = \left(\frac{\partial U'}{\partial n_k}\right)_{T,P,n_j} - T\left(\frac{\partial S'}{\partial n_K}\right)_{T,P,n_j}$
By definition,
$\bar{F}_k = \bar{U}_k - T\bar{S}_k$
See De Hoff for more examples and for coefficient an Maxwell relations.

14.4 Chemical Potential in Multicomponent Systems

Given $\mu_k(T,\ P,\ X_k)$, one can find $B,\ B'$, $\bar{B}(T,P,X)$
The thermodynamic state is specified by C+2 variables:
$dU' = TdS' - PdV' + \sum_{k=1}^{C} \mu_k dn_k$, so $\mu_k \equiv \left(\frac{\partial U'}{dn_K}\right)_{S,V,n_j}$

From Chapter 4 then, $\delta W' = \sum_{k=1}^{C} \mu_k dn_k$ (necessary in open systems)
Examining dH', dF' and dG' , we find that
Given $\mu_k(T,P) = \bar{G}_k$, we can express all PMP's in terms of μ_k :

$$\bar{S}_k, \bar{V}_k, \bar{H}_k, \bar{U}_k, \bar{F}_k$$

$$\bar{S}_k = -\left(\frac{\partial \bar{G}_k}{\partial T}\right)_{P,n_k} = -\left(\frac{\partial \mu_k}{\partial T}\right)_{P,n_k}$$

$$\bar{V}_k' = \left(\frac{\partial \bar{G}_k}{\partial P}\right)_{T,n_k} = \left(\frac{\partial \mu_k}{\partial P}\right)_{T,n_k}$$

So

$$\bar{H}_k = \bar{G}_k + T\bar{S}_k = \mu_k - T\left(\frac{\partial \mu_k}{\partial T}\right)_{P,n_k} - P\left(\frac{\partial \mu_k}{\partial P}\right)_{T,n_k}$$

$$\bar{F}_k = \bar{U}_k - T\bar{S}_k = \mu_k - P\left(\frac{\partial \mu_k}{\partial P}\right)_{T,n_k}$$

μ_k is not measured directly

$\Rightarrow$measurement of $\mu_k(P,T)$ enables PMPs to be determined. Note that $\Delta\bar{B}_k$ can also be determined by substituting $\mu_k \to \Delta\mu_k = \mu_k - \mu_k^0$

Activity and Activity Coefficient We found that $\mu_k(T,P) = \left(\frac{\partial G'}{\partial n_k}\right)_{P,T,\mu_j} \equiv \bar{G}_k(T,P)$ but we do not know how to measure directly.
Define $\mu_k - \mu_k^0 = \Delta\mu_k \equiv RT\ln a_k$, where a_k is the activity of component k. Further we define the activity coefficient γ_k in $a_k = \gamma_k X_k$, where X_k is the mole fraction.
So

$$\Delta\mu_k = RT\ln(Y_k X_k)$$

Next step: derive this for an ideal solution.

14.5 Properties of ideal solutions (e.g. ideal gas mixtures)

Imagine removing the partitions in Figure 14.3 below and allowing the gases to mix while keeping T,P constant. The process is analogous to free expansion.

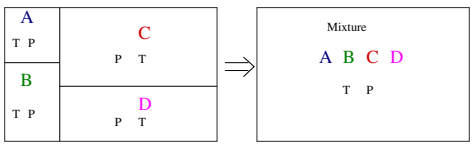


Figure 14.3: Mixture of ideal gases

Each component experiences a decrease in pressure from P to P_k. Each gas exerts, through collisions, a partial pressure on the wall of $P_k = X_k P$. The total pressure P is

$$P = \sum_{k=1}^{C} P_k = \sum_{k=1}^{C} X_k P = P\sum_{k=1}^{C} X_k = P$$

Consider the change in chemical potential for the component, K which we can find by integrating $d\mu_K$:

$d\mu_k = -\bar{S}_k dT + \bar{V}_k dP = \bar{V}_k dP$ since expansion is isothermal.
For an ideal gas,
$V' = n_T \frac{RT}{P} = \frac{RT}{P} = \frac{RT}{P}\sum_{k=1}^{C} n_k$,
for which we can find $\bar{V}_k = \left(\frac{\partial V'}{\partial n_k}\right)_{T,P,n_j} = \frac{RT}{P}$
So

$$\Delta\mu_k = \mu_k - \mu_k^0 = \int_P^{P_k} \bar{V}_k dP = \int_P^{P_k} \frac{RT}{P} dP = RT \ln \frac{P_k}{P} = RT \ln X_k$$

Since the activity $a_k = X_k$, we conclude that the activity coefficient $\gamma_k = 1$ for an ideal gas. (no interactions)
How do the partial molal quantities vary?

$\Delta\bar{G}_k = RT \ln X_k$

$\Delta\bar{V}_k = \left(\frac{\partial \Delta\mu_k}{\partial P}\right)_{T,n_k} = 0$

$\Delta\bar{S}_k = -\left(\frac{\partial \Delta\mu_k}{\partial T}\right)_{P,n_k} = -R \ln X_k$

$\Delta H_k = \Delta\mu_k + T\Delta\bar{S}_k = RT \ln X_k + T(-R \ln X_k) = 0$

$\Delta\bar{U}_k = \Delta\bar{H}_k - P\Delta\bar{V}_k = 0$

$\Delta\bar{F}_k = \Delta\bar{U}_k - T\Delta\bar{S}_k = RT \ln X_k$ The change in total volume, enthalpy, and internal energy is 0, as the changes in partial molal quantities is zero. However, entropy is produced, leading to decreases in the free energies $\Delta G_{mix}, \Delta F_{mix}$.

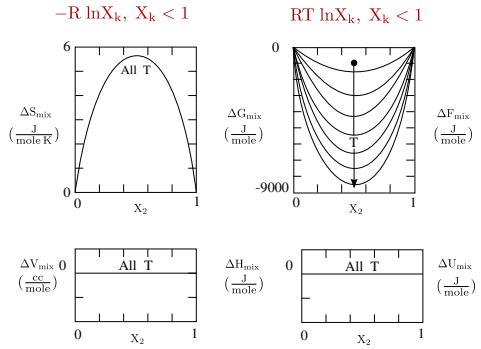


Figure 14.4: Properties of an ideal solution including $\Delta G_{mix}, \Delta S_{mix}, \Delta H_{mix}, \Delta V_{mix}, \Delta U_{mix}$ and ΔF_{mix}. Maximum temperature shown is 1400 K.

14.6 Behavior of Dilute Solutions

14.6.1 Case 1

Add a few atoms of component 2 (solute) to a large volume of component 1 (solvent).
The solvent atoms still behave as though they are in an ideal solution, regardless of interactions.
Raoult's Law: $\lim_{X_1 \to 1} a_1 = X_1$
The solute atoms interact primarily with solvent atoms, so the average properties will be proportional to X_2 for small X_2. (However, properties depend on solvent-solute combination.)
Henry's Law: $\lim_{X_2 \to 0} a_2 = \gamma_2^0 X_2 : \gamma_2^0(T, P)$ is Henry's Law constant (in particular range), the solute activity coefficient for a dilute solution.
Consider Solute A in Solvent B

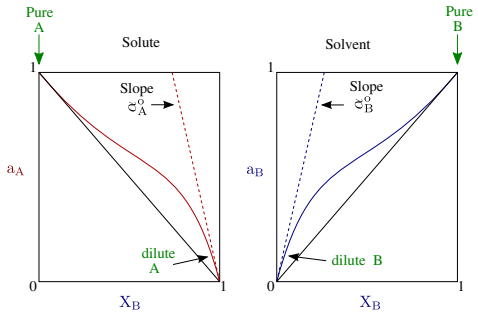


Figure 14.5

$\gamma_2 = e^{a_o X_1^2/RT}$ for $X_1 \to 1$, $\gamma_1^0 = e^{\alpha_0/RT}$

Example: HW 8.4

One mole of solid Cr_2O_3 at 2500K is dissolved in a large volume of a liquid Raoultian solution (at 2300K) of Al_2O_3 and Cr_2O_3(20%).

Calculate the resulting ΔH and ΔS given

$T_{m,Cr_2O_3} = 2538K$, $\Delta H_{m,Al_2O_3} = 107,500\sigma/mol - K$ at T_m(2324K)

$\Delta S_{m,Al_2O_3} = \Delta S_{m,Cr_2O_3}$

Recall that under isobaric, reversible conditions, $\Delta S^{\alpha\to P} = \frac{\Delta H^{\alpha\to\beta}}{T}$

$$\frac{\Delta H_{m,1}}{T_{m,1}} = \frac{\Delta H_{m,2}}{T_{m,2}} \Rightarrow \Delta H_{m,Cr_2O_3} = \Delta H_{m,Al_2O_3} = \Delta H_{m,Al_2O_3} \times \frac{T_{m,Cr_2O_3}}{T_{m,Al_2O_3}}$$

$$\Delta H_{m,Cr_2O_3} = 107,500\frac{J}{mol-K} \times \frac{2538K}{2324K} = 117,400J/mol - K$$

Note: $\Delta H = \Delta H^{S\to L} + \Delta H_{mix}$, but for ideal solution, $\Delta H_{mix} = 0$

Note: if heating of Cr_2O_3were required we read $C_P(T)$.

Next calculate the change in entropy due to 1) melting and 2) mixing.

"Large" liquid volume $\to$ no change in $X_{Cr_2O_3} \to \Delta\bar{H}_{Al_2O_3} = 0$

Focus on change in 1 mole of solid Cr_2O_3.

$$\Delta S = \Delta S_m + \Delta S_{mix}$$

$$\Delta S_m = \frac{\Delta H_m}{T_m} = \frac{117500}{2538} = 46.26J/K$$

$$\Delta S_{mix} = -R\ln X^F - (\cancel{-R\ln X^i}), \text{ where } X^i_{Cr_2O_3} = 1 \text{ (assume } \gamma^0 = 1)$$

$$\Delta S_{mix} = -R\ln(0.2)$$

$$\Delta S = 46.26 + 13.38 = 59.64J/K$$

14.7 Excess Properties, relative to ideal

$\Delta\mu_k = RT\ln(a_k) = RT\ln(Y_k X_k)$, so we can write

$$\Delta\bar{G}_k = \underbrace{RT\ln\gamma_k}_{\text{difference}} + \underbrace{RT\ln X_k}_{\text{ideal behavior}} = \Delta\bar{G}_k^{XS} + \Delta\bar{G}_k^{id}$$

If $\gamma_k > 1$, the excess free energy is positive.

If $\gamma_k > 1$, the excess free energy is negative.

Temperature and pressure derivatives of $\mu(P,T)$ give the remaining properties in terms of γ_k. (See table 8.5)

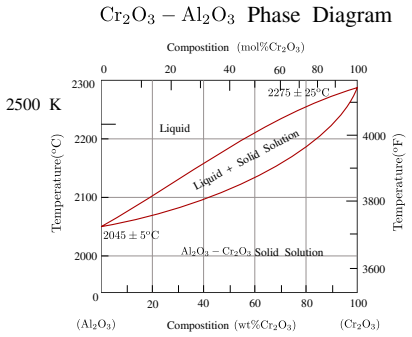


Figure 14.6: $Cr_2O_3 - Al_2O_3$ Phase Diagram

$$\left.\begin{aligned} \Delta H_k = \Delta\bar{G}_k - T\Delta S_k = \Delta\mu_k - T\left(\frac{\partial\mu_k}{\partial T}\right)_{P,\mu_k} &= -RT^2\left(\frac{\partial \ln\gamma_k}{\partial T}\right)_{P,n_k} \\ \Delta\bar{V}_k = \left(\frac{\partial\Delta\bar{G}_n}{\partial P}\right)_{T,n_k} = \frac{\partial}{\partial P}\left[(RT\ln\gamma_k + RT\ln X_R)\right]_{T,n_k} &= RT\left(\frac{\partial \ln\gamma_k}{\partial P}\right)_{T,n_k} \end{aligned}\right\} 0 \text{ for an ideal gas}$$

14.8 Beyond the ideal solution: the "regular" solution

1. The entropy of mixing is the same as for an ideal solution
2. The enthalpy of mixing is not zero, but depends on composition

Partial Molar $\Delta\bar{G}_k = \Delta\mu_k = \Delta\bar{H}_k - T\Delta\bar{S}_k$
Total Molar $\Delta G_{mix} = \Delta H_{mix} - T\Delta S_{mix}$
For ideal solution, $\Delta H_{mix} = 0$, For regular : $\Delta H_{mix} = \Delta H^{XS}_{mix}(X_2, X_3, ...X_c)$
For a regular solution, $\Delta S^{reg}_{mix} = \Delta S^{id}_{mix}$, so $\Delta S^{XS}_{mix} = 0$
$\Delta G^{XS}_{mix} = \Delta H^{XS}_{mix} - T\cancel{\Delta S^{XS}_{mix}}$
So $\Delta\bar{G}^{XS}_m = \Delta\bar{H}_k = RT\ln\gamma_k$
We see that γ_k can be evaluated from the heat of mixing: $\gamma_k = exp\left(\frac{\Delta\bar{H}_k}{RT}\right)$ (check in eq 8.102)
Let's look at the simplest binary example:
Let $\Delta H_{mix} = a_o X_1 X_2 = \Delta G^{XS}_{mix}$, where a_o is constant. This corresponds to example 8.1. The change in Gibbs Free energy.
$\Delta G_{mix} = \Delta G^{XS} + \Delta G^{id} = a_o X_1 X_2 + \underbrace{RTX_1\ln X_1 + RTX_2\ln X_2}_{\text{Table 8.3}}$
a_o determines the departure from ideal behavior arising from interactions between components.
A more realistic model (not symmetric and temperature dependent) of a binary system is
$\Delta H_{mix} = X_1X_2\left[\alpha_o(T)X_1 + \alpha_1(T)X_2\right] = \Delta G^{XS}_{mix}$
Let's examine the impact of interactions on ΔH_{mix}, ΔS_{mix}, and ΔG_{mix}, starting with the simpler regular solution.

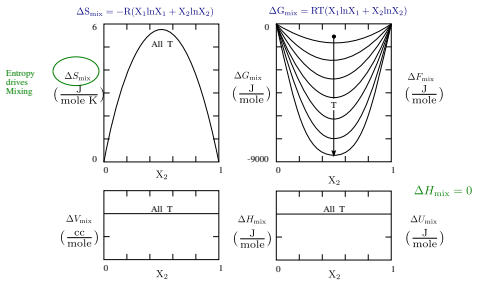


Figure 14.7: Properties of an ideal solution

$$\Delta\bar{S}_k = -R\ln\gamma_k - R\ln X_k - RT\left(\frac{\partial\ln\gamma_k}{\partial T}\right)_{P,n_k}$$

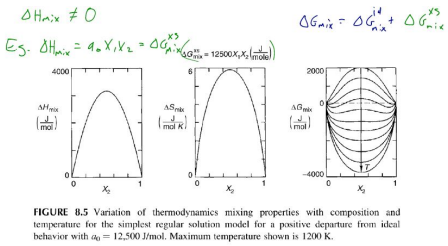


FIGURE 8.5 Variation of thermodynamics mixing properties with composition and temperature for the simplest regular solution model for a positive departure from ideal behavior with $a_0 = 12{,}500$ J/mol. Maximum temperature shown is 1200 K.

Figure 14.8: Properties of a regular solution

$$\Delta\bar{U}_k = RT^2\left(\frac{\partial \ln\gamma_k}{\partial T}\right)_{P,n_k} - PRT\left(\frac{\partial \ln\gamma_k}{\partial P}\right)_{T,n_k}$$

$$\Delta\bar{F}_k = RT\ln\gamma_k + RT\ln X_k - PRT\left(\frac{\partial \ln\gamma_k}{\partial P}\right)_{T,n_k}$$

For reference, not lecture
Gibbs Duhem Equations

$$X_1 d\Delta\mu_1 + X_2 d\Delta\mu_2 = 0$$

$$\Delta\mu_k = RT\ln(a_k)$$

$$X_1 d\ln\gamma_1 + X_2 d\ln\gamma_2 = 0$$

$$a_k = \gamma_k X_k$$

$$\Delta\mu = -\int_0^{X_2} \frac{X_2'}{X_1}\frac{d\Delta\mu_2}{dX_2'}dX_2'$$

$$\ln\gamma_1 = -\int_{X_1}^{X_2} \frac{X_2'}{X_1}\frac{d\ln\gamma_2}{dX_2'}dX_2'$$

$$\ln a_1 = -\int_0^{X_2} \frac{X_2'}{X_1}\frac{d\ln a_2}{dX_2'}dX_2'$$

$$\Delta\bar{S}_k = -\left(\frac{\partial\Delta\mu_k}{\partial T}\right)_{P,n_k}$$

$$\Delta\bar{H}_k = \Delta\mu_k + T\left(\frac{\partial\Delta\mu_k}{\partial T}\right)_{P,n_k}$$

$$\Delta\bar{G}_k = \Delta\mu_k = \Delta\bar{H}_k - T\Delta\bar{S}_k$$

$$\Delta G_{mix} = \Delta H_{mix} - T\Delta S_{mix}$$

For an ideal solution, $\Delta H_{mix} = 0$, for a regular solution $\Delta S^{reg}_{mix} = \Delta S^{id}_{mix}$

$$\Delta G^{XS}_{mix} = \Delta H^{XS}_{mix} - T\cancel{\Delta S^{XS}_{mix}}$$

$$\Delta S^{XS}_{mix} = 0$$

The entropy of mixing for a regular solution is the same as for an ideal solution (hence $\Delta S^{XS} = 0$)
The enthalpy of mixing for a regular solution deciates from that of an ideal solution ($\Delta H^{id} = 0, \Delta H^{XS} \neq 0$).

Example: HW 8.4 One mole of solid Cr_2O_3 at 2500 K is dissolved in a large colume of a liquid Raoultian solution (2300 K) of Al_2O_3 and Cr_2O_3(20%).
Calculate the resulting ΔH and ΔS given

$$T_{m,Cr_2O_3} = 2538K$$

$$\Delta H_{m,Al_2O_3} = 107,500\text{J/mol-K } @T_m(= 2324 \text{ K})$$

$$\Delta S_{m,Al_2O_3} = \Delta S_{m,Cr_2O_3}$$

Recall that under isobaric, reversible conditions, $\Delta S^{\alpha\to\beta} = \frac{\Delta H^{\alpha\to\beta}}{T}$

$$\frac{\Delta H_{m,1}}{T_{m.1}} = \frac{\Delta H_{m,2}}{T_{m,2}} \Rightarrow \Delta H_{m,Cr_2O_3} = \Delta H_{m,Al_2O_3} \times \frac{T_{m,Cr_2O_3}}{T_{m,Al_2O_3}}$$

$$\Delta H_{m,Cr_2O_3} = 107,500\frac{J}{mol-K} \times \frac{2538K}{2324K} = 117,400\frac{J}{mol-K}$$

Note: if heating of Cr_2O_3 were required, we need $C_P(T)$. Next calculate changes in entropy due to 1) melting and 2) mixing.
"Large" liquid volume $\rightarrow$ no change in $X_{Cr_2O_3} \rightarrow \Delta\bar{H}_{Al_2O_3} = 0$
Focus on change in 1 mole of solid Cr_2O_3

$$\Delta S = \Delta S_m + \Delta S_{mix}$$

$$\Delta S = \frac{\Delta H_m}{T_m} = \frac{117400}{2538} = 46.26J/K$$

$$\Delta S_{mix} = -R\ln X^F - (\cancel{-R\ln X^I}) \text{ , where } X^i_{Cr_2O_3} = 1$$

$$\Delta S_{mix} = -R\ln(0.2)$$

$$\Delta S = 46.26 + 13.38 = 59.64 \text{ J/K}$$

14.8.1 Mixtures of Real Gases: Fugacity

Describe "real" gases in terms fo their deviation from ideal behavior, starting with molar volume $\bar{V}_k$
Define $\alpha_k = \bar{V}_k - \frac{RT}{P}$ (note that $\alpha_k = 0 \leftrightarrow ideal$)
The change in chemical potential with mixing is then

$$\Delta\mu_k = \int_P^{P_k}(\alpha_k + \frac{RT}{P'})dP' = \int_P^{P_k}\alpha_k dP' + RT\ln\left(\frac{P_k}{P}\right)$$

Introduce fugacity f_k in

$$\boxed{\Delta\mu_k \equiv \text{RTln}\left(\frac{f_k}{P}\right)}$$

Note: $a_k = \frac{f_k}{P}$
Taking the exponential, we get

$$f_k = P_k exp\left[\int_P^{P_\mu}\alpha_k dP\right]$$

As $\alpha_k \rightarrow 0$, $f_k \rightarrow P_k$
Determination of fugacity (or a_k) $\leftrightarrow$finding $\mu(P,T)$
Key goal of 314-316 Sequence: Predict Microstructure

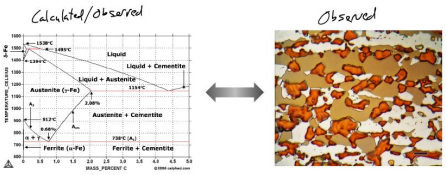


Figure 14.9: Calculated and observed microstructures

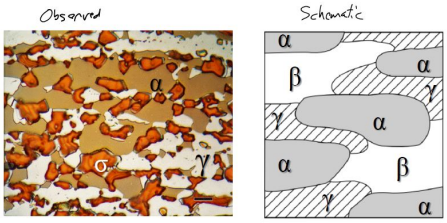


Figure 14.10: Multicomponent heterogeneous (multi-phase) system

15 Multicomponent Heterogeneous Systems

Outcomes for this section:

1. Write a combined statement of the first and second law of thermodynamics for multiple phases.
2. State the conditions for equilibrium for an arbitrary number of phases and components.
3. State the Gibbs phase rule and apply it to the description of unary and binary phase diagrams.
4. Given a single component phase diagram in terms of pressure and temperature, construct alternative representations that enable determination of phase fractions in terms of volume of mole fraction.
5. Given a two-component phase diagram in terms of pressure, temperature, and activity, construct alternative representations that enable determination of phase fractions in terms of volume or mole fraction.
6. Use the lever rule to compute phase fractions from temperature versus composition diagrams.

1. Define conditions for equilibrium
2. Derive Gibbs Phase Rule
3. Structure of 1 and 2 component phase diagrams
4. The Lever Rule

Consider $PV = RT$ (ideal gas)
There are two degrees of freedom: E.g. choose P, T (independent) $\rightarrow V$ (dependent) is determined
Equilibrium between 2 phases implies that the system is in equilibrium.
Homogeneous system: $dU' = \underset{(\delta Q)}{TdS'} - \underset{(\delta W)}{PdV'} + \underset{(dn)}{\sum_{k=1}^{C} \mu_k dn_k}$

sum over phases sum over components

Heterogeneous System: $dU'_{sys} = \sum_{\alpha=1}^{P} \left[T^\alpha dS'^\alpha - P^\alpha dV' + \sum_{k=1}^{C} \mu_k^\alpha dn_k^\alpha \right]$

15.1 Conditions for Equilibrium

$$T^\alpha = T^\beta \quad P^\alpha = P^\beta \quad \mu^\alpha = \mu^\beta$$

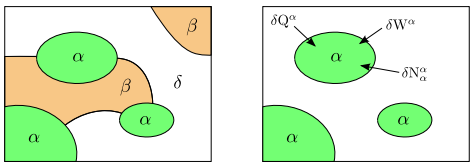


Figure 15.1: Multicomponent heterogeneous and homogeneous systems

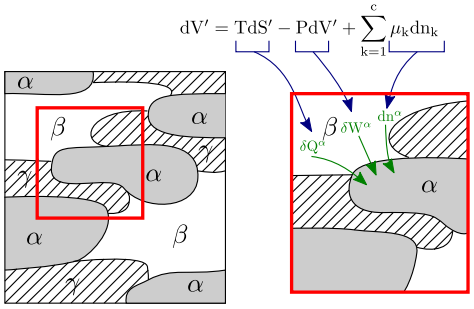


Figure 15.2: Multicomponent heterogeneous (multi-phase) system

likewise for all pairs of phases $\beta/\gamma, \gamma, ...P$, where P =number of phases
$T^I = T^{II} = ... = T^P$ $P-1$ independent equations
$P^I = P^{II} = ... = P^P$ $P-1$ independent equations

$$\left.\begin{array}{rl} \mu_1^I = \mu_1^{II} = ... = \mu_1^P & \text{〃} \\ \mu_2^I = \mu_2^{II} = ... = \mu_2^P & \text{〃} \\ \vdots & \\ \mu_C^I = \mu_C^{II} = ... = \mu_C^P & \text{〃} \end{array}\right\} \; C = \#\text{ components}$$

Equilibrium conditions generate $(C+2)(P-1) = n$ equations. These are constraints on the variables

15.2 Gibbs Phase Rule

$F = C - P + 2$
$F = m - n$ is the number of degrees of freedom
For n equations relating m variables. What is m?

$$\left.\begin{array}{ll} T^I, P^I, X_2^I, X_3^I ... X_C^I & 2 + C - 1 = C + 1 \\ T^{II}, P^{II}, X_2^{II}, X_3^{II} ... X_C^{II} & \text{〃} \\ \vdots & \vdots \\ T^P, P^P, X_2^P, X_3^P ... X_C^P & \text{〃} \end{array}\right\} \text{P sets of C+1 variables}$$

$$\therefore m = P(C+1)$$

$$m - n = P(C+1) - (C+2)(P-1) = \cancel{PC} + P - \cancel{PC} - 2P + C + 2$$

$$\boxed{\text{m-n=C-P+2}}$$

Gibbs Phase Rule

16 Phase Diagrams

Outcomes for this section:

1. Apply a common tangent construction to define the regions of phase stability in a G versus X diagram.
2. Use a single component $G - T$ diagram to compute $\Delta G_k^{0,\ \alpha\rightarrow\beta}$.
3. Given G versus X at different T, construct a phase diagram in terms of T and X.
4. Describe the influence of interactions (in a simple regular solution model) on the features of a phase diagram, including phase boundaries, their curvature, and regions of two-phase coexistence.
5. Describe the origins and consequences of a miscibility gap.

16.1 Unary Phase Diagrams

$C = 1$
$F = C - P + 2$
In single phase regions ($P = 1$), there are two degrees of freedom (e.g. T,P)

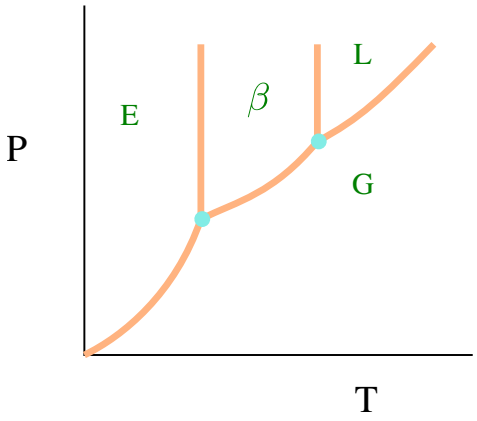


Figure 16.1: Unary (C=1) phase diagram

1. P and T required to specify state in single-phase region.
2. For 2-phase region, P fixes and T (and vice versa)
   1 degree of freedom ($F = 1 - 2 + 2 = 1$)
3. For triple point, $F = 0$

Only three phases are possible for a single component diagram.

Alternative Representations $P - V$, $T - X$, etc.
Not all useful variables (e.g. V,X) are thermodynamic potentials
2-phase coexistence lines →areas
To describe:

1. Increasing T, which increases V
2. Mixture of phases in regions with tie lines, average molar volume.
3. Triple points
4. Changes in enthalpy and phase fraction upon heating.

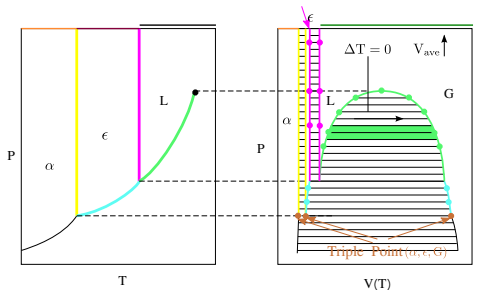


Figure 16.2: Alternate representations

16.2 Binary Phase Diagrams

$C = 2$
There are 5 possible phases: $\alpha, \epsilon, \beta, L, V$
$F = C - P + 2$
For a single phase region, $f = 3 \rightarrow$ Plot in P, T, $a_2$space. 2-phase regions are planes, 3-phase regions are lines, and 4-phase regions are points.

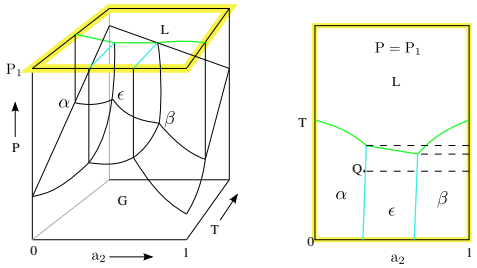


Figure 16.3: Binary (c=2) phase diagram

Alternative Representations: T vs. X In Figure 16.4 below, the 2-phase coexistence lines are areas.

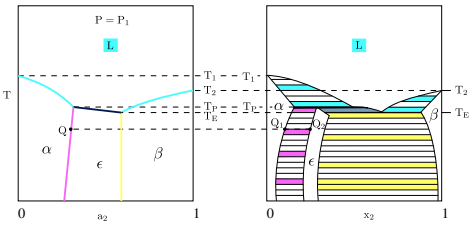


Figure 16.4: T vs X for binary phase diagram

16.3 Interpreting Phase Diagrams: Tie Lines

For plots with two degrees of freedom, a point on the diagram corresponds to a state of the system.

1. Which phases are present (and #) in equilibrium
2. The state of the phases (P,T, or other variables).
3. For plots with V or X, the relative amounts.

Example:
State P with average composition X_2^o at temperature T. The phase Σ is in equilibrium with the liquid ($T^\Sigma \neq T^L$).

$$n_T X_2^o = n^\Sigma X_2^\Sigma + n^L X_2^L \rightarrow X_2^o = \frac{n^\Sigma}{n^T} X_2^\Sigma + \frac{n^L}{n^T} X_2^L = f^\Sigma X_2^\Sigma + f^L X_2^L$$

where n= # of moles and f= fraction
In terms of X_2,

$$X_2^o = f^\Sigma X_2^\Sigma + (1 - f^\Sigma) X_2^L = X_2^L + f^\Sigma (X_2^\Sigma - X_2^o)$$

$$\therefore f^\Sigma = \frac{X_2^o - X_2^L}{X_2^\Sigma - X_2^L}$$

and

$$f^L = \frac{X_2^o - X_2^\Sigma}{X_2^L - X_2^\Sigma}$$

These fractions can be derived from the lengths of segments of a tie line.

16.4 The Lever Rule

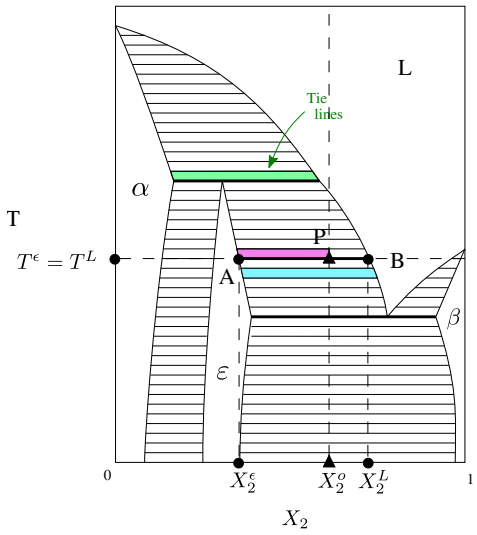


Figure 16.5: Tie lines on an X-T diagram

$$f^L = \frac{X_2^o - X_2^\Sigma}{X_2^L - X_2^\Sigma} = \frac{AP}{AB}$$

$$f^\Sigma = \frac{X_2^o - X_2^L}{X_2^\Sigma - X_2^L} = \frac{PB}{AB}$$

$$X_2^\Sigma - X_2^L = AB$$

$$X_2^\Sigma - X_2^o = AP$$

$$X_2^o - X_2^L = PB$$

X-T diagrams gives temperature, compositions, and amounts.

16.5 Applications of Phase Diagrams: Microstructure Evolution

Following an isotherm in the X-T diagram of Figure shows the following microstructures:

- T_o : homogeneous melt
- T_1 :$\alpha + L$ (nucleation)
- T_2 : increase in α phase
- T_3 : solidification - $\alpha + \beta$

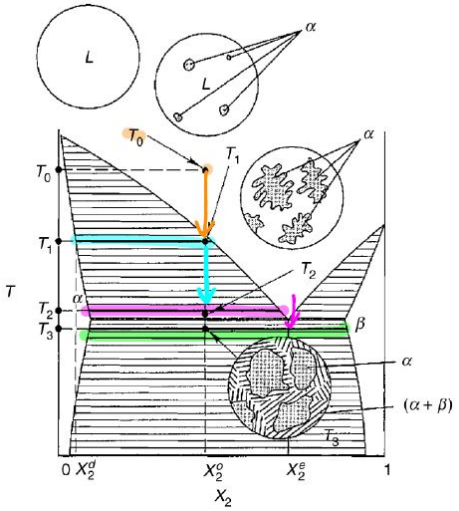


Figure 16.6: Microstructure Evolution

17 Review

$$\Delta G_{mix} = \Delta H_{mix} - TS^{id}_{mix}$$

$$\Delta G_{mix} = \Delta G^{XS}_{mix} + RT(X_1 \ln X_1 + X_2 \ln X_2)$$

$$\Delta G^{ideal}_{mix} = RT(X_1 \ln X_1 + X_2 \ln X_2)$$

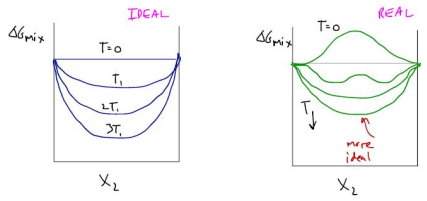


Figure 17.1: Gibbs free energy of mixing for ideal vs real mixtures

$\Delta H_{mix} \neq 0$: interactions change Gibbs free energy of mixing
$T \uparrow$: particle kinetic energy increases, influence of interactions decreases, behavior is more ideal.

17.1 Types of $\alpha - L$ phase diagrams

$\Delta H_{mix} = a_o X_1 X_2$
Note the curvature of the phase boundary in Figure 17.2 below - the sign follows that of a_o^2.

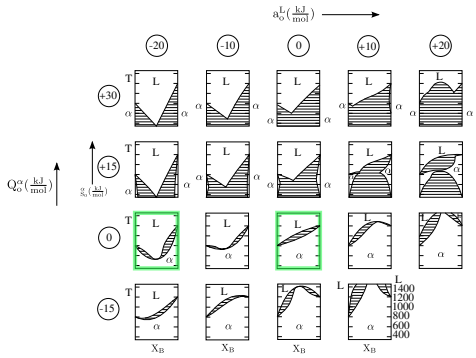


Figure 17.2: Patterns of phase diagrams can generated by only two phases, α and L, with the simplest regular solution model

18 Thermodynamics of Phase Diagrams

Goal: Generate phase diagrams from ΔG vs. X curves
Observe: how interactions (changing potential energy) and temperature (changing kinetic energies) influence characteristics of phase diagrams.
Approach: find common tangent lines between ΔG^{α}_{mix} and ΔG^{L}_{mix} to determine regions of phase stability.
Needed: comparison of solid and liquid reference states

18.1 Calculation of ΔG_{mix} from pure component reference states

We can plot ΔG_{mix} for a solid (α) or liquid (L) solution using solid or liquid reference states for components 1,2 (4 possibilities).
One cannot directly compare the following:

$$\left.\begin{aligned} \Delta G^{\alpha}_{mix} &= X^{\alpha}_1\left(\bar{G}^{\alpha}_1 - G^{o\alpha}_1\right) + X^{\alpha}_2\left(\bar{G}^{\alpha}_2 - G^{o\alpha}_2\right) \\ \Delta G^{L}_{mix} &= X^{L}_1\left(\bar{G}^{L}_1 - G^{oL}_1\right) + X^{\alpha}_2\left(\bar{G}^{L}_2 - G^{oL}_2\right) \end{aligned}\right\}$$ Need to change reference states

Indicate choice of reference states with brackets:

$$\left.\begin{aligned} \Delta G^{\alpha}_{mix} &= X^{\alpha}_1\left(\bar{G}^{\alpha}_1 - \boxed{G^{o\alpha}_1}\right) + X^{\alpha}_2\left(\bar{G}^{\alpha}_2 - \boxed{G^{o\alpha}_2}\right) \\ \Delta G^{L}_{mix} &= X^{L}_1\left(\bar{G}^{L}_1 - \boxed{G^{oL}_1}\right) + X^{\alpha}_2\left(\bar{G}^{L}_2 - \boxed{G^{oL}_2}\right) \end{aligned}\right\}$$ Reference states the same for $\alpha$and L phases

Rewrite:

$$\Delta G^{\alpha}_{mix} = X^{\alpha}_1\left(\bar{G}^{\alpha}_1 - G^{o\alpha}_1\right) + X^{\alpha}_2\left(\bar{G}^{\alpha}_2 - G^{o\alpha}_2\right) - X^{\alpha}_2\left(\bar{G}^{oL}_2 - G^{o\alpha}_2\right)$$

$$\Delta G^{\alpha}_{mix}(\alpha, L) = X^{\alpha}_1\left(\bar{G}^{\alpha}_1 - G^{o\alpha}_1\right) + X^{\alpha}_2 + X_2\left(G^{o\alpha}_2 - G^{oL}_2\right)$$

$$\Delta G^{\alpha}_{mix}(\alpha, L) = \Delta G^{\alpha}_{mix}(\alpha, \alpha) + X^{\alpha}_2 \Delta G^{oL\to\alpha}_2$$

Similarly,

$$\Delta G^{L}_{mix}(\alpha, L) = \Delta G^{L}_{mix}(L, L) + X^{L}_1 \Delta G^{o\alpha\to L}_2$$

What do these look like?

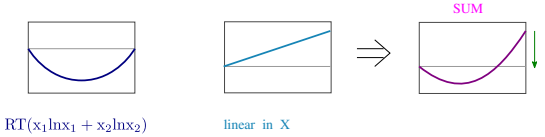


Figure 18.1

Next step: plot $\Delta G^{\alpha}, \Delta G^{L}$ together and find common tangent to determine regimes of phase stability.

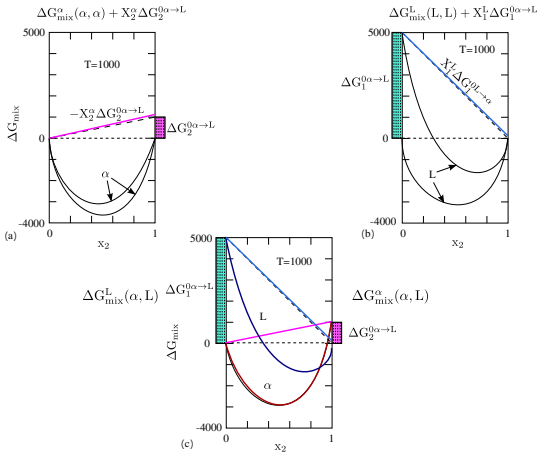


Figure 18.2

18.2 The Common Tangent Construction

The common tangent construction is a graphical represenation of two phase equilibrium for a given P,T. An example is illustrated in Figure 18.3. At this T and P, a solid of composition X_2^αcan be in equilibrium with a liquid of composition $X_2^L(T^\alpha = T^L, P^\alpha = P^L, \mu_1^\alpha = \mu_1^L, \mu_2^\alpha = \mu_2^L)$.

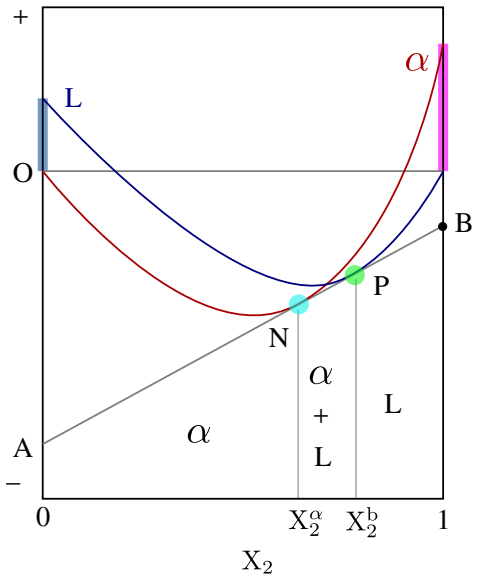


Figure 18.3: Common tangent construction

The intercepts reflect the conditions for chemical equilibrium.

A: $\left.\begin{array}{l}\Delta\mu_1^\alpha = \mu_1^\alpha(X_2^\alpha) - \mu_1^{o\alpha} \\ \Delta\mu_1^L = \mu_2^L(X_2^L) - \mu_1^{o\alpha}\end{array}\right\} \mu_1^\alpha(X_2^\alpha) = \mu_1^L(X_2^L)$

Chemical equilibrium of component 1 between α, L

B: $\left.\begin{array}{l}\Delta\mu_2^\alpha = \mu_2^\alpha(X_2^\alpha) - \mu_2^{o\alpha} \\ \Delta\mu_2^L = \mu_2^L(X_2^L) - \mu_2^{o\alpha}\end{array}\right\} \mu_2^\alpha(X_2^\alpha) = \mu_2^L(X_2^L)$

Chemical equilibrium of component 2 between α, L.

When are two-phase regions $(\alpha + L, \alpha + \beta)$ stable at intermediate compositions?

$\left.\begin{array}{l}X_2 < D : \alpha \\ X_2 > F : \beta\end{array}\right\}$ homogeneous mixtures of 1 and 2

$D < X_2 < F : \alpha + \beta$ has lower energy than homogeneous mixture of 1 and 2

$B : \Delta G$ for 1 & 2 in β phase

$A : \Delta G$ for 1 & 2 in αphase

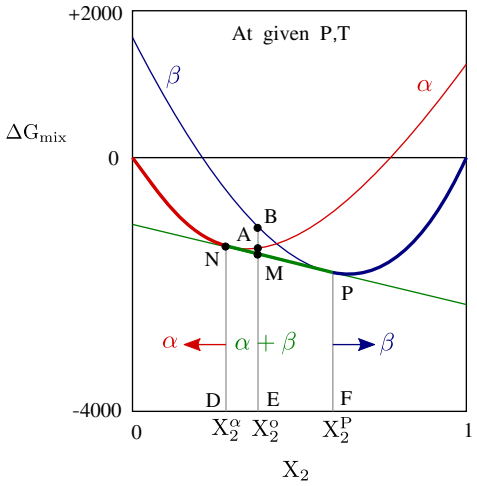


Figure 18.4

$M : \Delta G$ for $\frac{EF}{DF}$ of α and $\frac{DE}{DF}$ of β
$\Delta G_M < \Delta G_A < \Delta G_B$
Sequential transitions between multiple two-phase regions 3 solid phases, 1 liquid phase

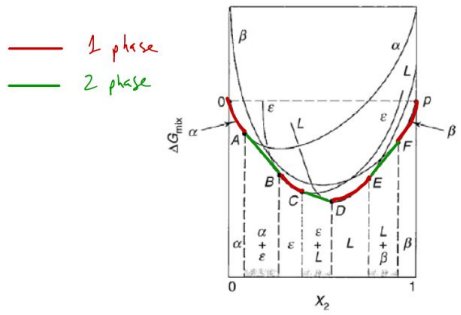


Figure 18.5

How are the single and two phase regions represented on a phase diagram?

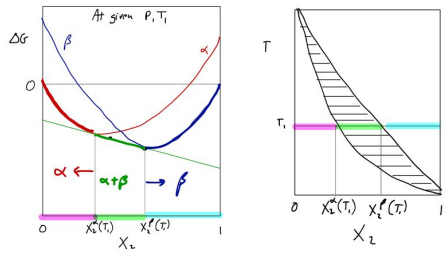


Figure 18.6: Representing single and two phase regions on a phase diagram

We need $\Delta G(X)$ for many temperature. Examine $G^\alpha(T)$ and $G^P(T)$.
To generate ΔG_{mix} vs. X curves, we need $G(T)$ of pure components to establish "hanging points" (the $\Delta G_i^{o\alpha\rightarrow\beta}$)
Next, examine changes inΔG_{mix}^α and ΔG_{mix}^L versus T to establish which phases are present.

18.3 A Monotonic Two-Phase Field

$a_o^\alpha = 1000 J/mol$ $a_o^L = -6000 J/mol$
$\Delta H_{mix} = a_o X_1 X_2$

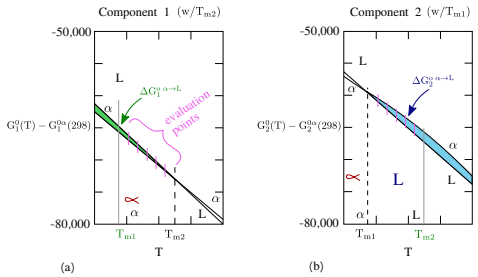


Figure 18.7: G(T) of pure components

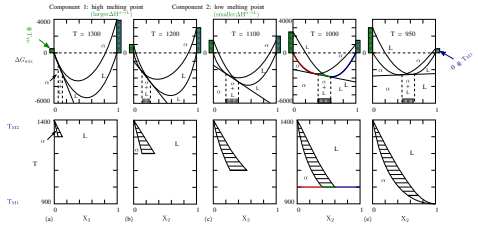


Figure 18.8

18.4 Non-monotonic Two-Phase Field

$a_o^\alpha = 6000 J/mol$
$a_o^L = -2000 J/mol$
$\Delta H_{mix} = a_o X_1 X_2$

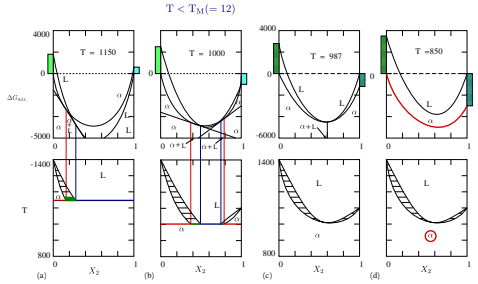


Figure 18.9

18.5 Properties of a regular solution

Recall: Regular - $\Delta G_{mix}^{id} + \Delta G_{mix}^{XS}$

18.6 Misciblity gap: appearance upon cooling

$\Delta H_{mix} = 13.7 X_1 X_2 kJ/mol$: below critical T, mixture is driven towards phase separation
$a_o^\alpha = 13,700 J/mol$

18.7 Spinodal decomposition: concave down region in ΔG drives unmixing and "uphill" diffusion

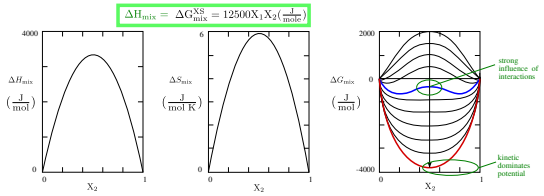


Figure 18.10: Variation of mixing properties with composition and temperature for the simplest regular solution model. Maximum temperature is 1200 K, and $a_o = 12,500 J/mol$

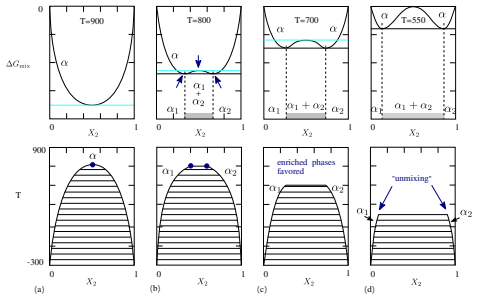


Figure 18.11: Miscibility gap

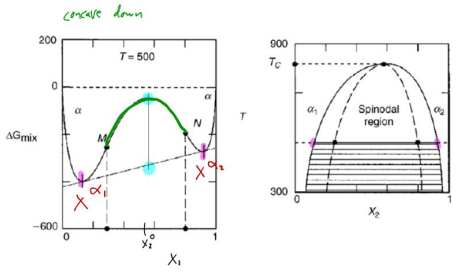


Figure 18.12: Spinodal decomposition

19 Summary of 314

The summary is presented in terms of a set of detailed learning outcomes, coupled to the chapters in Dehoff. These are things you should be able to do after completing 314.

19.1 Chapters 2 and 3

1. State the first, second and third laws of thermodynamics.
2. Write a combined statement of the first and second laws in differential form.
3. Given sufficient information about how a process is carried out, describe whether entropy is produced or transferred, whether or not work is done on or by the system, and whether heat is absorbed or released.
4. Quantitatively relate differentials involving heat transfer/production, entropy transfer/production, and work.

19.2 Chapter 4

1. Derive differentials of state functions (V, S, U, F, H, G) in terms of dP and dT.
2. Using the Maxwell relations, relate coefficients in the differentials of state functions.
3. Define the coefficient of thermal expansion, compressibility, and heat capacity of a material.
4. Express the differential of any given state function in terms of the differentials of two others.
5. Define reversible, adiabatic, isotropic, isobaric, and isothermal processes.
6. Calculate changes in state functions by defining reversible paths (if necessary) and integrating differentials for (a) an ideal gas, and (b) materials with specified α, β, and $C_P(T)$. (Multiple examples are given for ideal gas, solids, and liquids.)
7. Special emphasis (Example 4.13): calculate the change in Gibbs free energy when one mole of a substance is heated from room temperature to an arbitrary temperature at constant pressure.
8. Describe the origin of latent heat, and employ this concept in the calculation of changes in Gibbs free energy G.

19.3 Chapter 5

1. Derive the condition for equilibrium between two phases.
2. Show that at constant temperature and pressure, the equilibrium state corresponds to that which minimizes the Gibbs free energy G.
3. Explain why the evolution to a state of lower Gibbs free energy occurs spontaneously.

19.4 Chapter 6

1. Express the entropy S of the system in terms of equivalent microstates in the equilibrium (most probable) macrostate.
2. Given the energy levels of a system of particles, calculate the partition function.
3. Given the partition function, calculate the Helmholtz free energy, the entropy, the internal energy, and the heat capacity.

19.5 Chapter 7

1. State the differential of the Gibbs free energy G as a function of S, T, V, and P.
2. State the conditions for two phases to coexist in equilibrium.
3. Explain how isobaric section of the chemical potential $\mu(T,\ P)$ as a function of temperature and pressure can be calculated from knowledge of the heat capacity and the entropy at room temperature.
4. Given $\mu(T,\ P)$ for multiple phases, determine the equilibrium phase at a given temperature.
5. Derive the Clausius-Clapeyron relation: express the general condition for determining the region of two-phase coexistence.
6. Calculate the ΔH given the $C_P(T)$, assuming a negligible contribution from thermal expansion.
7. Calculate $\Delta V^{\alpha\to\beta}$ assuming ideal gas behavior.
8. Integrate $\frac{dP}{dT} = \frac{\Delta S}{\Delta V}$ with the assumption that ΔS and ΔV depend only weakly on the temperature.
9. Given the $P-T$ diagram, qualitatively sketch $G(T)$ for a specified region.
10. Given the $P-T$ diagram, specify the sign of $\Delta V^{\alpha\to\beta}$.
11. State Trouton's rule and justify its validity qualitatively.
12. Sketch the $P-T$ diagram for a one-component system.

19.6 Chapter 8

1. Given sufficient information about the dependence of a total property B or ΔB versus composition, find the partial molal properties as a function of composition via calculations of derivatives dB or $d\Delta B$.
2. Given the partial molal properties as a function of composition, calculate the total properties of the system via integration of the derivatives.
3. Given the differentials of state functions, write partial derivatives that define the partial molal properties.
4. What is B_k^0?
5. What is $\mu_k - \mu_k^0$ for (a) a component behaving ideally, and (b) in general?
6. Define the activity of a component.
7. Given $\mu_k(T, P, X_k)$, calculate changes in the partial molal properties.
8. State and justify/explain Raoult's law.
9. State and justify/explain Henry's law.
10. Explain how the activity coefficient γ can be used to describe the departure from ideal behavior in terms of "excess" quantities.
11. Define a "regular" solution and calculate the partial molal properties and total properties versus composition for binary mixtures.
12. Describe the driving forces for mixing and their origins in statistical mechanics.
13. Interpret changes in state functions with mixing in terms of ideal behavior and departures from ideality.

19.7 Chapter 9

1. Write a combined statement of the first and second law of thermodynamics for multiple phases.
2. State the conditions for equilibrium for an arbitrary number of phases and components.
3. State the Gibbs phase rule and apply it to the description of unary and binary phase diagrams.
4. Given a single component phase diagram in terms of pressure and temperature, construct alternative representations that enable determination of phase fractions in terms of volume of mole fraction.
5. Given a two-component phase diagram in terms of pressure, temperature, and activity, construct alternative representations that enable determination of phase fractions in terms of volume or mole fraction.
6. Use the lever rule to compute phase fractions from temperature versus composition diagrams.

19.8 Chapter 10

1. Apply a common tangent construction to define the regions of phase stability in a G versus X diagram.
2. Use a single component $G - T$ diagram to compute $\Delta G_k^{0,\,\alpha\rightarrow\beta}$.
3. Given G versus X at different T, construct a phase diagram in terms of T and X.
4. Describe the influence of interactions (in a simple regular solution model) on the features of a phase diagram, including phase boundaries, their curvature, and regions of two-phase coexistence.
5. Describe the origins and consequences of a miscibility gap.

20 314 Problems

20.1 Phases and Components

1. (2014) For each of the following thermodynamic systems, indicate the number of components, the number of phases, and whether the system is open or closed.

   (a) An open jar of water at room temperature (assume that the jar defines the boundaries of the system). Assume that the water molecules do not dissociate.

   (b) A sealed jar of water at room temperature.

   (c) A sealed jar of water with ice.

   (d) An open jar of water with NaCl entirely dissolved within.

   (e) If the jar is left open, in what ways might your description change?

   (f) How would your answer to (a) change if you take into account equilibrium between water, protons, and hydroxyl ions?

20.2 Intensive and Extensive Properties

2. (2014) Classify the following thermodynamic properties are intensive or extensive:

   (a) The mass of an iron magnet.

   (b) The mass density of an iron magnet.

   (c) The concentration of phosphorous atoms in a piece of doped silicon.

   (d) The volume of the piece of silicon.

   (e) The fraction by weight of copper in a penny.

   (f) The temperature of the penny in your pocket.

   (g) The volume of gas in a hot air balloon.

20.3 Differential Quantities and State Functions

3. (2014) Consider the function $z = 6x^2y^3\cos^2 u$.

   (a) Write down the total differential of z. Identify the coefficients of the three differentials in this expression as partial derivatives.

   (b) Demonstrate that three Maxwell relations (see section 2.3) hold between the coefficients identified under (a).

4. (2014) Why are state functions so useful in calculating the changes in a thermodynamic system?

5. (2014) Derive equation 4.41 starting from 4.34 and 4.31. Note that other equations listed in table 4.5 can be derived in a similar fashion.

20.4 Entropy

6. (2015) Following Section 3.6, compute the change in entropy in the formation of one mole of SiO_2 from Si and O at room temperature.

7. (2015) Consider an isolated system consisting of three compartments A, B, and C. Each compartment has the same volume V, and they are separated by partitions that have about. Initially, the valves are closed and volume A is filled with an ideal gas to a pressure P_0 at 298 K. Volumes B and C are under vacuum.

   (a) Calculate the change in entropy when the valve between compartment A and B is opened.

   (b) Calculate the change in entropy when the valve between compartment B and compartment C is opened.

   (c) Without considering the calculations above, how would you know that the overall change in entropy is positive?

   (d) What would you need to do to the system to restore the initial condition?

20.5 Thermodynamic Data

8. (2015) This problem requires you to find sources to look up the values of important materials parameters that will be used to compute thermodynamic functions.

   (a) Find values of the coefficient of thermal expansion for a metal, a semiconductor, an insulator, and a polymer. Provide the information below in your answer.

Material Type	Specific Material	α	Source (include page or link info)
Elemental Metal	*e.g.* Gold		
Semiconductor			
Insulator			
Polymer			

   (b) What is a common material with a negative α?

   (c) How is the coefficient of compressibility related to the bulk modulus?

   (d) Which metal has the highest bulk modulus at room temperature, and what is the value?

   (e) The heat capacity is an extensive quantity. Define the related intensive quantity.

   (f) What trend do you observe in elemental solids?

   (g) What is the smallest value you can find for a solid material? (Explain your search method, and cite your sources.)

9. (2014) The density of silicon carbide at 298 K and 1 atm is ~3.2 g/cm^3. Estimate the molar volume at 800 K and a pressure of 1000 atm. See tables 4.1 and 4.2 on page 61 of DeHoff for useful materials parameters.

10. (2015) The density of aluminum at 298 K and 1 atm (or "bar") is 2.7 g/cm3. Estimate the molar volume at 1000 K and a pressure of 1000 atm. See tables 4.1 and 4.2 on page 61 of DeHoff, and Appendix B, for useful materials parameters. Hint: break the problem into two steps, each corresponding to a path.

11. (2015) Use the car mileage dataset provided to do the following:

    (h) Create a second order polynomial fit to determine the coefficients for the mileage dataset online. Use the systems of equations we developed during discussion to help you solve for the coefficients. Write your polynomial coefficients down in your submitted assignment.

    (i) Using your curve of best fit, determine the optimal speed for driving that maximizes your mileage.

    (j) Identify an obvious failure of your model and comment on it below.

12. (2015) Answer the following questions using the heat capacity dataset and the following model:

$$C_p = a + bT + c/T^2 + dT^2$$

    (k) Use the system of equations derived in class to determine the coefficients a, b, c, d.

    (l) Give a possible Gibbs free energy function for bulk silicon using your heat capacity fit. The Gibbs free energy is related to the heat capacity through the following equation:

$$C_p = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_p$$

13. (2015) Compare the change in entropy for the specific examples below of isothermal compression and isobaric heating of gases and solids.

    (m) One mole of nitrogen (N_2) at 1000 K is compressed isothermally from 1 to 105 bar.

    (n) One mole of silicon at 300 K is compressed isothermally from 1 to 105 bar.

    (o) One mole of oxygen (O2) at 300 K is heated isobarically from 300 to 1200 K.

    (p) One mole of tungsten at 300 K is heated isobarically from 300 to 1200 K.

14. (2015) For each of the following processes carried out on one mole of a monatomic ideal gas, calculate the work done by the gas, the heat absorbed by the gas, and the changes in internal energy, enthalpy, and entropy (of the gas). The processes are carried out in the specified order.

    (q) Free expansion into the vacuum to twice the volume, starting from 300 K and 4 bar. Then,

    (r) Heating to 600 K reversibly with the volume held constant. Then,

    (s) Reversible expansion at constant temperature to twice the volume of the previous state. Then,

    (t) Reversible cooling to 300 K at constant pressure.

15. (2015) Consider one mole of a monatomic ideal gas that undergoes a reversible expansion one of two ways.

    (u) Under isobaric conditions, the gas absorbs 5000 J of heat in the entropy of the gas increases by 12.0 J/K. What are the initial and final temperatures of the gas?

    (v) Under isothermal conditions, 1600 J of work is performed, resulting in an entropy increase of 5.76 J/K and a doubling of the volume. At what temperature was this expansion performed?

16. (2015) In class we learned that the change in entropy of a material with temperature is given by:

$$S_2 - S_1 = \int_{T_1}^{T_2} \frac{C_p(T)}{T} dT \tag{20.1}$$

    In a prior homework, we fit the heat capacity to a polynomial, which we could then integrate. Now, we will numerically integrate the data points using the Trapezoid Rule discussed in class:

$$\int_{T_1}^{T_2} f(x) \approx (T_2 - T_1) \left[\frac{f(T_1 + f(T_2))}{2} \right] \tag{20.2}$$

    where the function $f(T)$ in our case is the right hand side of Equation , is simply the right hand side of Equation 20.1. Do this by creating a "FOR" loop in MATLAB that sums up all the trapezoids in the temperature range. Email your MATLAB script to the TA by the due date.

    (w) What is the difference in entropy at 300 K and 1300 K?

    (x) Previously, we determined that the heat capacity is given by:

$$C_p = 22.83 + 3.826x10^{-3}T - \frac{3.533x10^5}{T^2} + 2.131x10^{-8}T^2 \tag{20.3}$$

    Use Equation 20.1 to analytically solve for the change in entropy using Equation 20.3. Which method do you think is more accurate? Explain your reasoning.

20.6 Temperature Equilibration

17. (2015) 100 g of ice at 250 K is added to 100 g of water at 300 K, and the mixture is allowed to come to equilibrium in an isolated container at constant pressure. You may assume that C_p is constant for this problem (though it is not the same for water and ice) and that the melting point is 273 K.

    (a) What is the final temperature?

    (b) How much liquid is present?

    (c) How would your answer change if the initial liquid was 40% ethanol?

18. (2015) A square block of Al, initially at a uniform temperature of 300 K, is brought into contact with another block of aluminum, initially at a uniform temperature of 600K. Both blocks are of equal mass, and they are isolated at constant pressure while they come to equilibrium. The questions below assume equilibrium has been reached.

    (d) What do you know about the final temperatures of the Al blocks?

    (e) Will the final temperature(s) be 450 K? Justify your answer.

    (f) Find the final temperature.

20.7 Statistical Thermodynamics

19. (2015) DeHoff 6.3: Consider a system of two particles (A and B) that may each occupy any of the four energy levels ($\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$).

    (a) How many distinct microstates are there for this system?

    (b) List each of the microstates and indicate which microstates have the same energy.

    (c) How many macrostates are there?

    (d) List the most probable macrostates.

20. (2015) DeHoff 6.5 Variant: Consider the system consisting of 9 identical but distinguishable particles, each of which can be in any of three states. The respective energy levels of the states are $\epsilon_0 = 0, \epsilon_1 = \epsilon, \epsilon_2 = 2\epsilon$. The system has a temperature T.

    (e) Write the partition function for a single particle.

    (f) Calculate the average number of particles in each state.

    (g) Determine the number of configurations that have the following occupation numbers for the three states: $n_0 = 4$, $n_1 = 3$, $n_2 = 2$.

    (h) Calculate the entropy of the macrostate described by the occupation numbers above.

    (i) Calculate the internal energy.

    (j) Choose a different set of occupation numbers to give the same internal energy (*e.g.* (3,5,1)) and repeat your calculation of the entropy. Which macrostate is more likely?

21. (2015) DeHoff 6.7 variant: A System containing 500 particles and 15 energy levels is in the following macrostate: {14, 18, 27, 38, 51, 78, 67, 54, 32, 27, 23, 20, 19, 17, 15}. Estimate the change in entropyu when the system undergoes a process leading to the following changes in occupation numbers: {0, 0, -1, -1, -2, 0, 1, 0, 3, 2, -1, 1, -1, 0, -1}.

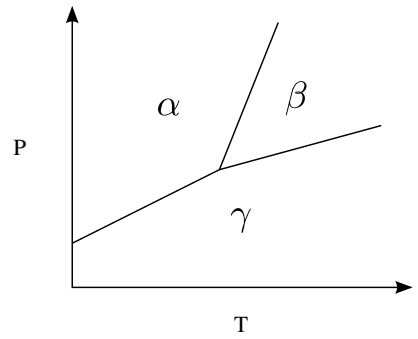
20.8 Single Component Thermodynamics

22. (2015) DeHoff 6.10: Compute the change in entropy when one mole of a monatomic ideal gas is compressed from an initial condition at 273K and 1 bar to 500K at 3.5 bar.

    (a) Calculate using the phenomenological thermodynamics of Chapter 4.

    (b) Calculate using the results of statistical thermodynamics. Hint: first calculate the initial and final volumes.

23. (2015) DeHoff 7.5: Sketch $G(T)$ for an element that the pressure corresponding to the triple point. Repeat the sketch for a pressure slightly above and slightly below the triple point.



24. (2015) DeHoff 7.6 See Lecture 17, last page.

25. (2015) DeHoff 7.7 variant: At what pressure will ice melt at -2°C?

26. (2015) DeHoff 7.8: At 1 atm pressure and below 1155 K, the ε form of titanium is stable; above 1155 K, the β form is the stable phase (ε becomes metastable). Given the following data:

    - $\Delta S^{\varepsilon\rightarrow\beta} = 3.43\,\mathrm{J/mol\cdot K}$ (This is the difference in molar entropy 3 between the phases).
    - The change in molar entropy upon melting is $9.02\,\mathrm{J/mol\cdot K}$.
    - T_m^{β}=1940 K.

    (c) Sketch G^{ε}, G^{β} and G^{ℓ} in the temperature range of interest.

    (d) What is T_m^{ε}?

    (e) There is a database of the Gibbs free energy of 78 pure elements as a function of temperature. The database can be found here: http://www.crct.polymtl.ca/sgte/unary50.tdb
    Find the ε phase of titanium (labeled as GHSERTI), the β phase (labeled as GBCCTI), and the liquid phase (labeled as GLIQTI) and repeat a and b using the empirical formulas. Compare your answers and comment on the accuracy of your assumptions.
    NOTE: The formula is written so that a program called Thermocalc can read them. Each free energy curve is a piecewise formula. The ";" separates the parts of the function over different temperature ranges. In addition, a "**" is the same as an exponent or "^".

20.9 Mulitcomponent Thermodynamics

27. (2015) DeHoff 8.1: Titanium metal is capable of dissolving up to 30 atomic percent oxygen. Consider a solid solution in the system Ti–O containing an atom fraction, $X_0 = 0.12$. The molar volume of this alloy is 10.68 cc/mol. Calculate the following:

    (a) The weight percent of O in the solution.

    (b) The molar concentration (mol/cc) of O in the solution.

    (c) The mass concentration (gm/cc) of O in the solution.

    (d) Use these calculations to deduce general expressions for weight percent, molar, and mass concentrations of a component in a binary solution in terms of the atom fraction, X_2, the molar volume, V, and the molecular weights, MW_1 and MW_2, of the elements involved.

28. (2015) DeHoff 8.4: Use the partial molal volumes computed in Problem 8.3 (worked out in class) to demonstrate that the Gibbs – Duhem equation holds for these properties in this system.

29. (2015) DeHoff 8.6: For an ideal solution it is known that, for component 2, $\Delta G_2 = RT\ln X_2$. Use the Gibbs – Duhem integration to derive corresponding relation for component 1.

30. (2015) One mole of solid Cr_2O_3 at 2500 K is dissolved in a large volume of a liquid Raoultian solution (also at 2500 K) of Al_2O_3 and Cr_2O_3 with $X_{Cr_2O_3}$= 0.2. Calculate the resulting changes in the total enthalpy and entropy given the following:

$$T_{m,Cr_2O_3} = 2538\,\mathrm{K};\ \Delta H_{m,Al_2O_3} = 107,500\,\mathrm{J/mol}\ \text{at}\ T_{m,Al_2O_3} = 2324\,\mathrm{K};\ \Delta S_{m,Al_2O_3} = \Delta S_{m,Cr_2O_3}$$

21 314 Computational Exercises

31. (2015) For this problem, you will be using MATLAB's symbolic solver (fzero) and function handles to find the zero of an equation. On last week's quiz we found that 89.1 grams of ice were necessary to cool a 1 kg block of Pb down to 300 K from 600 K. We will be plotting the change in temperature for both the Pb and ice. You will need the following parameters:

$$C_p^{Pb} = 0.1169 + 4.2x10^{-5}T\ \tfrac{\mathrm{kg}}{\mathrm{kg\cdot K}}$$
$$C_p^{H_2O} = 4.2\ \tfrac{\mathrm{kJ}}{\mathrm{kg\cdot K}}$$
$$\Delta H_{fus}^{H_2O} = 344\ \tfrac{\mathrm{kJ}}{\mathrm{kg}}$$
$$\tfrac{dH}{dt} = a\Delta T$$

Here a is a coefficient that controls the heat transfer in conduction, which we will assume to be $3.33x10^{-3}$ kJ/K · s, andΔT is the temperature difference between the two materials.

(a) Use the equations derived in class to plot the change in temperature for H2O and Pb. Assume dt=1 s and calculate the first 200 time steps. How do you know when the system is at equilibrium?

(b) Create a plot that shows the amount of water in the system as a function of time. At what time is all the ice gone?

(c) Create a plot that shows the total heat transfer occurring between the Pb and H2O. How can you tell when equilibrium is reached from this plot?

32. (2015) We will be putting together a program to help calculate phase diagrams of all sorts piece by piece. The first step is to create a MATLAB script that solves a system of two equations. The system is below:

$$G_s - X_s \frac{dG_s}{dX_s} = G_\ell - X_\ell \frac{dG_\ell}{dX_\ell} \tag{21.1}$$

$$\frac{dG_s}{dX_s} = \frac{dG_\ell}{dX_\ell} \tag{21.2}$$

where G_s and G_ℓ are given by the following expressions:

$$G_s\left(X_s\right) = \Omega_s X_s\left(1 - X_s\right) + RT\left[X_s \ln X_s + \left(1 - X_s\right)\ln\left(1 - X_s\right) + 200X_s - 400\left(1 - X_s\right)\right].$$

$$G_\ell\left(X_\ell\right) = \Omega_\ell X_\ell\left(1 - X_\ell\right) + RT\left[X_\ell \ln X_\ell + \left(1 - X_\ell\right)\ln\left(1 - X_\ell\right)\right].$$

Here G_s is the Gibbs energy of the solid phase, G_ℓ the Gibbs free energy of the liquid phase, R the gas constant (8.314 J/K), T the absolute temperature, X_s and X_ℓ are the compositions of the solid and liquid phase respectively, and Ω_s, Ω_ℓ are parameters to be defined later. We can rewrite Eqs. 21.1 and 21.2 as follows:

$$G_s - X_s \frac{dG_s}{dX_s} - G_\ell + X_\ell \frac{dG_\ell}{dX_\ell} = 0 \tag{21.3}$$

$$\frac{dG_s}{dX_s} - \frac{dG_\ell}{dX_\ell} = 0 \tag{21.4}$$

Create a MATLAB function that takes X_s, X_ℓ, T, Ω_s and Ω_ℓ as inputs and then create a script that uses the MATLAB command fsolve to calculate X_s and X_ℓ for T=700K, $\Omega_\ell = 1500\,\text{cal/mol}$ and $\Omega_s = 3000\,\text{cal/mol}$.

33. (2015) Now that we are able to solve for the composition of the solid and liquid at one point, we will improve our script to calculate it over a range of temperatures. Start with an initial guess for both the solid and liquid near zero and a temperature of 900 K. Determine the composition of the liquid and solid down to 1 K for each temperature using a for loop. Make sure to update your guess with the correct answer for the previous temperature to help your program converge. Repeat again starting from 600 down to 1 K, this time starting with an initial guess near 1, and plot your results. Use the following parameters to make the Gibbs energy more physical:

$$T_m^\alpha = 900\,\text{K}$$

$$T_m^\beta = 600\,\text{K}$$

$$\Delta H_f^\alpha = 2000\,\frac{\text{cal}}{\text{mol}}$$

$$\Delta H_f^\beta = 1300\,\frac{\text{cal}}{\text{mol}}$$

The liquid and solid free energies are given by the following expressions. (Note that the Gibbs energy for the solid phase has changed slightly and should be adjusted in your code. In these units $R = 1.987\,\text{cal/mol}\cdot\text{K}$

$$G_s\left(X_s\right) = \Omega_s X_s\left(1-X_s\right) + RT\left[X_s \ln X_s + \left(1-X_s\right)\ln\left(1-X_s\right) + X_s \Delta G_\beta^{\ell\to s} + \left(1-X_s\right)\Delta G_\alpha^{\ell\to s}\right].$$

$$G_\ell\left(X_\ell\right) = \Omega_\ell X_\ell\left(1-X_\ell\right) + RT\left[X_\ell \ln X_\ell + \left(1-X_\ell\right)\ln\left(1-X_\ell\right)\right].$$

As a reminder, the free energy change for the melting transition can be written in terms of the enthalpic and entropic contributions to the free energy:

$$\Delta G^{s\to\ell} = \Delta H^{s\to\ell} - T\Delta S^{s\to\ell}$$

34. (2015) Our phase diagram calculation is almost complete! We only have to find the equilibrium between the two solid phases left. To do that, we simply take the derivative of the Gibbs free energy of the solid phase and set it equal to zero. The equation becomes:

$$RT\ln\left(\frac{X_s}{1-X_s}\right) - \Omega_s\left(2X_s - 1\right)$$

Again, assume that Ω_s= 3000 cal/mol. This is easily done by creating a for loop that solves for the temperature at each composition between .01 and .99. Plot your results on the same figure from the previous homework and voila, your first phase diagram!

Now with your working code, replot the diagrams for the following interaction coefficients. You will have to change the range of temperatures for one of the sets below. You can figure out which one it is, if you think of the physical significance of the parameters.

(a) $\Omega_s = 3000\,\text{cal/mol} : \Omega_\ell = 0$

(b) $\Omega_s = 0 :\ \Omega_\ell = 3000\,\text{cal/mol}$

(c) $\Omega_s = 0;\ \Omega_\ell = 0$

(d) $\Omega_s = 3000\,\text{cal/mol} : \Omega_\ell = 3000\,\text{cal/mol}$

For each phase diagram, plot your results and describe how the changing interaction parameters changed the shape of the plot.

Part III
315: Applications of Thermodynamics

22 Introduction

"Thermodynamics" and "kinetics" are fundamental skill sets/tool boxes required of all materials scientists/engineers. Thermodynamics tells us which phase or assemblage of phases has the absolute lowest free energy, and therefore represents the equilibrium state. (Note: there may be other metastable states at higher energies.) Kinetics tells us much more: how fast those phases will form, and the paths they will take along the way. Together, thermodynamics and kinetics determine the phase assemblages/microstructures that can be obtained, and how to obtain them. In the materials science & engineering paradigm of Figure 22.1, thermodynamics and kinetics come primarily into play in the first "chain link" between "Processing" and "Structure."

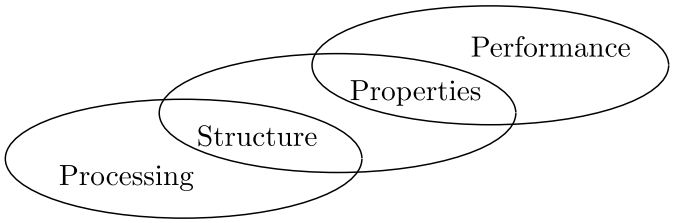


Figure 22.1: The Materials Science & Engineering Paradigm

The interplay between thermodynamics and kinetics can be illustrated with two case studies. The first involves the Fe-C phase diagram, which is introduced in virtually all introductory Materials Science & Engineering courses. A schematic of the Fe-rich end of this diagram is given in Figure 22.2.

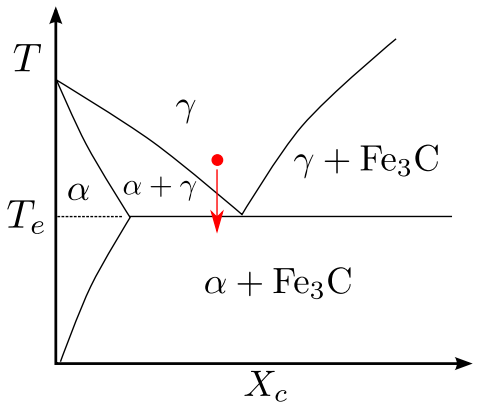


Figure 22.2: Schematic of the Fe-rich end of the Fe-C phase diagram

If we solutionize the eutectoid composition austenite (γ-phase) at the point indicated in Figure 22.2, and then cool it (follow the arrow) below the eutectoid temperature (thermodynamics), various microstructures can result depending upon the rate of cooling (kinetics). Very slow cooling can result in discrete coarse-grained phases (α-phase austenite and cementite, or Fe_3C). This assemblage of phases is not very strong or hard. On the other hand, by cooling more rapidly, we can produce a layered structure of austenite and cementite, referred to as pearlite for its "mother of pearl" appearance under the microscope. This microstructure is found to be quite strong and hard. This is a prime example of how the processing-structure "chain link" in Figure 1 can influence the resulting structure $\Leftrightarrow$ properties "chain-link" in the materials science & engineering paradigm of Figure 22.1.
The second example involves the oxidation of silicon to silicon dioxide through the reaction of equation 22.1.

$$\mathrm{Si(s) + O_2(g) \rightleftharpoons SiO_2(s)} \tag{22.1}$$

Later in this course we will learn about Richardson-Ellingham diagrams (for simplicity, these will be referred to as Ellingham diagrams). An Ellingham diagram is just a superposition of lines representing the free energy of

oxidation for a large number of metals. A schematic of the Ellingham diagram for just silicon is given in Figure 22.3.

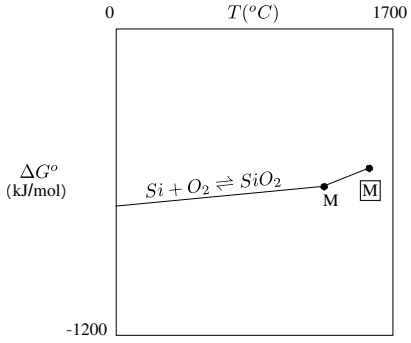


Figure 22.3: Schematic Ellingham diagram for the oxidation of silicon

The top left of the diagram is at zero degrees Centigrade and at zero of ΔG^o. (Note that the letter "M" and "M" inside a box refer to the melting points of the metal (Si) and oxide (SiO_2), respectively.) As we will see, the free energies of formation of all oxides, including silica, are strongly negative. This means that reactions like that in Equation 22.1 have a strong tendency to go to the right, namely, to produce their oxide at the expense of the corresponding metal. If we throw an iron bar out in the "elements," we know that it rusts (forms the oxide) quite readily. However, an aluminum object, in spite of having an even larger negative free energy of oxidation than either iron or silicon, will hardly corrode under the same conditions. That is owing to the formation of a coherent "passive" oxide film on the surface, through which diffusion is extremely slow. The same thing takes place on silicon, as illustrated in Figure 22.4.

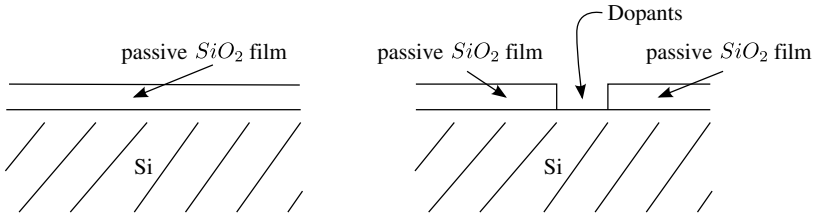


Figure 22.4: Passive oxide film formation on silicon, which can be removed in certain areas for dopant incorporation.

The passive SiO_2 film that forms is extremely important to the microelectronics at work in many computers. (Note: at present, silica is being replaced by oxides with larger dielectric constants, to ensure that the miniaturization necessary to keep extending Moore's Law, i.e., the number of transistors on a processor chip doubling every 18 months, can continue.) On the right side of Figure 22.4 part of the film has been intentionally removed (in a process known as photolithography) so that dopants can be introduced to change its electronic properties locally. This may take place by ion-implantation (using an ion beam) or it may take place by diffusion from a gaseous source. Again, this is a good illustration of the interplay of thermodynamics and kinetics in both processing $\Leftrightarrow$ structure and structure $\Leftrightarrow$ properties links in Figure 22.1.

In the first half of this text, we concentrate on the thermodynamics side of things, namely, the application of thermodynamics to the prediction and interpretation of phase diagrams. The level of presentation assumes two things: 1) You have had an introductory course in materials science and engineering, one that introduced simple phase diagrams, the phase rule, and the lever rule. 2) You have a background in the laws of thermodynamics, and especially in the area of solution thermodynamics. If you know the difference between Raoultian and Henrian solution behavior and have been introduced to the Regular solution model, you will be in a good position to follow along. If not, it is suggested that you spend some time reading about basic solution thermodynamics.

23 A Most Useful Equation: The Generalized Gibbs-Duhem Equation

From our knowledge of thermodynamics, most of us are familiar with the standard form of the Gibbs-Duhem equation 23.1.

$$0 = X_A d\mu_A + X_B d\mu_B \tag{23.1}$$

which is used extensively to describe the behavior of solutions, namely, it plays a big role in "solution thermodynamics." This equation only holds true for binary systems at fixed temperature and pressure, the X terms represent mole fractions, and the μ terms represent chemical potentials. A more general form of the Gibbs-Duhem equation can be derived as follows. Let's start with a binary system. The total internal energy (U) is given by equation 23.2.

$$U = TS - PV + n_1\mu_1 + n_2\mu_2 \tag{23.2}$$

where T is absolute temperature, P is pressure, the n terms represent the number of moles of each component, and the μ terms are chemical potentials, as in equation 23.1. However, from the First law of thermodynamics (equation 23.3) we know that the change of internal energy is a balance between heat *in* (δq) and work done *by* (or out of) our system (δw). For now and for the sake of simplicity, δw will be limited to PV-work only (at constant pressure δw becomes PdV).

$$dU = \delta q - \delta w = \delta q - PdV \tag{23.3}$$

The second law of thermodynamics (equation 23.4) tells us that the change of the entropy (S) of the system is always greater than the actual heat in (δq) divided by absolute temperature.

$$dS \geq \delta q/T \tag{23.4}$$

In fact, the entropy change is equal to the reversible heat in (δq_{rev}) divided by absolute temperature, as expressed in equation 23.5.

$$dS = \delta q_{rev}/T \tag{23.5}$$

If we now combine the first law (equation 23.3) and the second law (equation 23.4) for a closed system (no matter in or out, i.e., no $nd\mu$ terms), we obtain equation 23.6.

$$dU \leq TdS - PdV \tag{23.6}$$

or at equilibrium, using δq_{rev}, and equation 23.5 instead of equation 23.4, we get equation 23.7.

$$dU = TdS - PdV \tag{23.7}$$

Now let's consider an open system that can exchange matter with the environment. We will keep it a binary system for now (and for the sake of simplicity). With the μdn terms added, the combined first and second law equation 23.7 becomes equation 23.8.

$$dU = TdS - PdV + \mu_1 dn_1 + \mu_2 dn_2 \tag{23.8}$$

On the other hand, the total differential of total internal energy, equation 23.2, gives us equation 23.9.

$$dU = TdS + SdT - PdV - VdP + n_1 d\mu_1 + \mu_1 dn_1 + n_2 d\mu_2 + \mu_2 dn_2 \tag{23.9}$$

If we now subtract equation 23.8 from equation 23.9, we obtain a more complete form of the Gibbs-Duhem equation, at least for binary systems, given in equation 23.10.

$$0 = SdT - VdP + n_1 d\mu_1 + n_2 d\mu_2 \tag{23.10}$$

Although we will "generalize" this equation still further, this equation is powerful! The author refers to this equation as the "Swiss army knife" for understanding degrees of freedom and the phase rule, and also for classifying and interpreting phase diagrams of all kinds.

23.1 The Gibbs-Duhem Equation and Degrees of Freedom

For example, as we look at equation 23.10, how many total variables do we have? The answer is "four," including T (temperature), P (pressure), and two chemical potentials (μ_1, μ_2). But if we ask (for the case of a single phase) how many variables we need to control to establish thermodynamic equilibrium, the answer is "three." For example, if we fix temperature ($dT = 0$), pressure ($dP = 0$), and the chemical potential of component 1 ($d\mu_1 = 0$), then according to equation 23.10 the other chemical potential must also be fixed ($d\mu_2 = 0$). In other words, there are three degrees of freedom, which we refer to by the variable $\underline{F}$. Later we will understand degrees of freedom to be the number of thermodynamic variables that need to be fixed to establish equilibrium, or alternatively, the number of thermodynamic variables that can be independently varied without a change in the number of phases present in our system. We should also note in passing that we get a hint of the $(C + 2)$ term in the familiar Gibbs phase rule, equation 23.11.

$$\underline{F} = (C + 2) - \underline{P} \tag{23.11}$$

In the $(C + 2)$ term, C stands for the number of components ($d\mu_1, d\mu_2$) and the 2 stands for temperature (dT) and pressure (dP), or in other words, $(C + 2)$ is the total number of thermodynamic variables. Note that we have employed a different symbol to represent the number of phases ($\underline{P}$) to differentiate this from pressure (P). If we add another component (component 3), then we would have to add a $n_3\mu_3$ term to equation 23.10. Now let's see what happens when we have more than one phase, which means adding a second Gibbs-Duhem equation.

23.2 Using the Gibbs-Duhem Equation to Derive the Phase Rule

Let keep it simple by limiting ourselves to a single-component system (C=1). The relevant version of the Gibbs-Duhem equation in equation 23.10 would become equation 23.12,

$$0 = SdT - VdP + nd\mu \tag{23.12}$$

where we have dropped the subscripts for component 1, i.e., $n_1 d\mu_1 = nd\mu$. But now let's imagine that we have two phases, α and β, in equilibrium along the phase boundary in Figure 23.1.

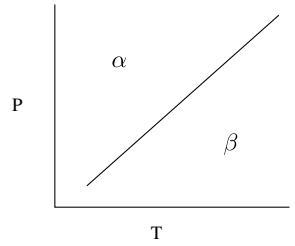


Figure 23.1: A two-phase boundary in a P-T phase diagram

Since the moles are distributed between the two phases, it is not necessarily true that $n^\alpha = n^\beta$, where the superscripts refer to the two phases. The same can be said of entropy or volume. Imagine ice floating in water. The two phases (solid, liquid) are in equilibrium, which we know by the fact that the temperature remains constant as long as there is any significant amount (moles, volume) of ice. But once all the ice melts, the water is free to rise in temperature. So what we need to do is to write two Gibbs-Duhem equations, one for the alpha phase (equation 23.13) and one for the beta phase (equation 23.14).

$$0 = S^\alpha dT - V^\alpha dP + n^\alpha d\mu \tag{23.13}$$

$$0 = S^\beta dT - V^\beta dP + n^\beta d\mu \tag{23.14}$$

Now let's combine these two equations to eliminate one thermodynamic variable, e.g., the chemical potential (dμ), to arrive at equation 23.15:

$$\left(\frac{S}{n}\right)^\alpha dT - \left(\frac{V}{n}\right)^\alpha dP = -d\mu = \left(\frac{S}{n}\right)^\beta dT - \left(\frac{V}{n}\right)^\beta dP \tag{23.15}$$

Reorganizing, equation 23.15 becomes equation 23.16:

$$\left[\left(\frac{V}{n}\right)^{\beta} - \left(\frac{V}{n}\right)^{\alpha}\right] dP = \left[\left(\frac{S}{n}\right)^{\beta} - \left(\frac{S}{n}\right)^{\alpha}\right] dT \tag{23.16}$$

Think for a minute about what equation 23.16 means. For a single-phase in a single-component system, equation 23.12 tells us that two thermodynamic potentials must be fixed in order to establish the equilibrium state. As pointed out above, in order to fix the chemical potential ($d\mu = 0$), we have to fix both temperature ($dT = 0$) and pressure ($dP = 0$). But now that we have two phases in equilibrium ($\mu^{\alpha} = \mu^{\beta}$), we only need fix one variable. In equation 23.16 if pressure is fixed ($dP = 0$), then temperature must also be fixed ($dT = 0$), or vice versa; so we have decreased the "degrees of freedom" by one by adding the second Gibbs-Duhem equation (the second phase in equilibrium with the first). (Note: It is assumed that the molar volumes and molar entropies of the two phases are constant.) In other words, the coexistence of two phases in thermodynamic equilibrium requires the writing of two Gibbs-Duhem equations--one for each phase (equations 23.13 and 23.14). This clearly demonstrates that "degrees of freedom" (F) equals the total number of thermodynamic variables, which equals the number of components plus 2 (for temperature and pressure) or ($C + 2$) minus the number of Gibbs-Duhem equations, which is equal to the number of phases in equilibrium (P). We have thereby derived the Gibbs Phase Rule in equation 23.11. Note that if we have three phases in thermodynamic equilibrium in a single-component system ($\mu^{\alpha} = \mu^{\beta} = \mu^{\gamma}$), we would have to add an additional Gibbs-Duhem equation 23.17:

$$0 = S^{\gamma} dT - V^{\gamma} dp + n^{\gamma} d\mu \tag{23.17}$$

From this equation and equations 23.13 and 23.14 we could write equations like equation 23.16 for each of the three phase boundaries meeting at what is known as a "triple point," as in the P-T diagram for water in Fig. 23.2. We could then eliminate one of the two remaining thermodynamic variables, e.g., temperature (dT), giving us equation 23.18:

$$\frac{\left[\left(\frac{V}{n}\right)^{\beta} - \left(\frac{V}{n}\right)^{\alpha}\right]}{\left[\left(\frac{S}{n}\right)^{\beta} - \left(\frac{S}{n}\right)^{\alpha}\right]} dP = dT = \frac{\left[\left(\frac{V}{n}\right)^{\gamma} - \left(\frac{V}{n}\right)^{\alpha}\right]}{\left[\left(\frac{S}{n}\right)^{\gamma} - \left(\frac{S}{n}\right)^{\alpha}\right]} dP \tag{23.18}$$

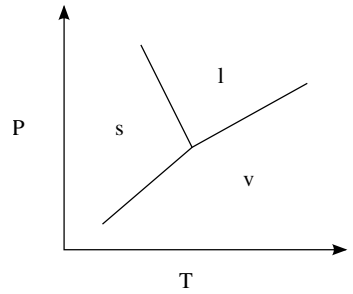


Figure 23.2: Schematic pressure-temperature phase diagram for H_2O.

Assuming that all molar volumes and molar entropies are constants and non-zero, there is only one possible solution to equation 23.18, namely that dP must be zero. By adding the third Gibbs-Duhem equation for the third phase (γ), we end up with zero degrees of freedom. In fact, the "triple point" of water is something you can look up in a handbook, 273.16K (0.01^{o}C) and 0.00604 atm, and can only be changed by increasing the number of components (for example, by doping water with salt, as is done to lower its freezing point at constant pressure, which we describe later).

You probably saw equation 23.16 in prior courses (e.g., chemical or materials thermodynamics), but in a slightly different form, as in equation 23.19:

$$\left(\frac{dP}{dT}\right)_{eq} = \frac{\left[\left(\frac{S}{n}\right)^{\beta} - \left(\frac{S}{n}\right)^{\alpha}\right]}{\left[\left(\frac{V}{n}\right)^{\beta} - \left(\frac{V}{n}\right)^{\alpha}\right]} \tag{23.19}$$

If we let S stand for molar entropy and V stand for molar volume, this equation becomes the well-known Clausius-Claypeyron equation 23.20:

$$\left(\frac{dP}{dT}\right)_{eq} = \frac{\Delta S^{\alpha\rightarrow\beta}}{\Delta V^{\alpha\rightarrow\beta}} = \frac{\Delta H^{\alpha\rightarrow\beta}}{T_{eq}\Delta V^{\alpha\rightarrow\beta}} \tag{23.20}$$

since in equilibrium the free energy difference is zero, $\Delta G^{\alpha\rightarrow\beta} = 0$, which means that $\Delta H^{\alpha\rightarrow\beta} = T_{eq}\Delta S^{\alpha\rightarrow\beta}$. As you already know, this is a powerful equation for P-T diagrams. Given the molar volume difference between alpha and beta phases, from the slope of their P-T equilibrium phase boundary at a chosen T_{eq} we can calculate the enthalpy (and entropy) of the phase transformation, or vice versa.
Before moving on, we must make one clarification regarding the number of components. It would seem that the number of components should be 2 for the H_2O system, one each for hydrogen and oxygen. However, if the ratio of hydrogen-to-oxygen remains constant for all phases in the system, namely H:O remains 2:1, then we can consider this as a one-component system.

23.3 Using the Generalized Gibbs-Duhem Equation to Classify All Phase Diagrams

In section 2.1 we spoke of the Gibbs-Duhem equation as the "Swiss army knife" of phase equilibrium thermodynamics. We have used it thus far to determine the degrees of freedom in a single-phase, one-component system. We have added second and then third phases in equilibrium (and therefore second and third Gibbs-Duhem equations) to derive Gibbs' Phase Rule. And we have used it to derive the Clausius-Clapeyron equation. Now we will use it to derive an overarching classification scheme for *all* phase diagrams. We acknowledge Professor Arthur Pelton of École Polytechnique Montreal as the originator of this powerful classification scheme [29].
First of all, let's generalize the Gibbs-Duhem equation to equation 23.21:

$$0 = SdT - VdP + \Sigma n_i d\mu_i = \Sigma Q_i d\phi_i \tag{23.21}$$

in which μ_i stands for the various thermodynamic "potentials" and Q_i stands for the corresponding "conjugate extensive variables." In Table 23.1 the thermodynamic potentials, whether thermal (T), mechanical (P) or chemical (μ_i), are "intensive," meaning that they do not depend upon the size of the "system" under consideration. For example, take copper at standard temperature and pressure (STP). A cube of copper 1 cm on a side has the same temperature, pressure and chemical potential as a cube of copper 1 m on a side. On the other hand, the conjugate variables are definitely "extensive," meaning that they clearly depend upon the size of the system. On going from the 1 cm cube of copper to the 1 m cube of copper all of these variables increase: volume, number of moles, and entropy (although the last is not as obvious).

ϕ_i (intensive thermodynamic potential)	Q_i (conjugate, extensive variable)
T (thermal)	S (entropy)
P (mechanical)	$-V$ (volume)
μ (chemical)	n (moles)

Table 23.1: Thermodynamic "potentials" vs. "conjugate extensive variables."

We are now in a position to understand Pelton's classification scheme for all phase diagrams. Schematic representations of the three types are given in Fig. 23.3. Type I diagrams are plots of one thermodynamic potential vs. another, in other words ϕ_i vs. ϕ_j. Type II diagrams are plots of a thermodynamic potential (ϕ_i) vs. a ratio of conjugate extensive variables (Q_j/Q_k). (Later we will prove that fixing a ratio of conjugate extensive variables is tantamount to fixing their thermodynamic potentials.) Type III diagrams are plots of one ratio of thermodynamic potentials vs. another, in other words Q_i/Q_k. vs Q_j/Q_k.
Schematic representations of "real" phase diagrams for each case are given in the second row of diagrams. For example, a conventional P-T diagram like that of water is a good example of a Type I diagram. However, you may have noticed a couple of anomalies in the other "representative" diagrams. For example, the schematic (and easy to recognize) Type II binary eutectic diagram does not have n_B/n_A as its x-axis. There is a good reason for this. Think of what happens if we let n_A go to zero. This would result in an infinite value of n_B/n_A. Instead, we use mole fraction ($X_B = n_B/(n_A + n_B)$), which is zero for $n_B = 0$ and unity for $n_A = 0$. Note in Fig. 23.3 that mole fraction can be easily related to the ratio of n_A/n_B, which is just the inverse of n_B/n_A. The other anomaly is that we seldom, if ever, see Type III ternary diagrams in rectilinear form, namely n_C/n_A vs. n_B/n_A. The reason is pretty

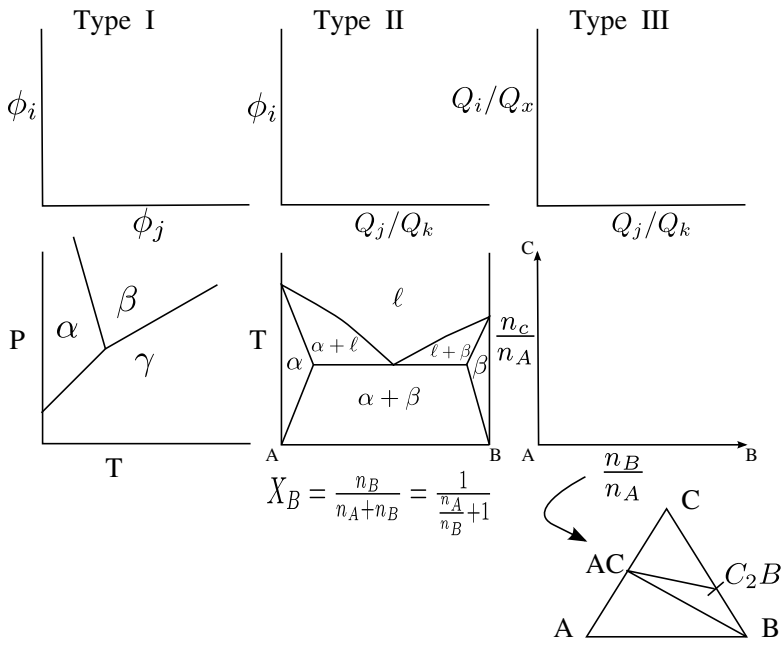


Figure 23.3: Pelton's classification scheme for phase diagrams.

obvious. Pure "end-member" A, using phase diagram parlance to be discussed later ($n_B = n_C = 0$), is at the origin of this plot, but pure end-members B and C are at infinity on the x- and y-axes, respectively. We can thank J. Willard Gibbs for introducing the now universally employed "Gibbs phase triangle" diagram, where the mole fraction of each component goes to unity in its respective corner. The Gibbs triangle diagram at the bottom right of Fig. 23.3 is a representative isothermal section of a "real" ternary phase diagram in the "subsolidus," meaning well below temperatures that would result in the formation of any liquid. Furthermore, this system exhibits negligible solid solubility, so the "end-member" and intermediate compounds (AC and C_2B) are the vertices of "tie-triangles." Such triangles are the hallmark of ternary phase diagrams, which we discuss in detail later.

24 Type I (Single Component) Phase Diagrams

Given the well-known phase diagram of water (see Fig. 23.2), Type I phase diagrams are mistakenly thought of as "unary" or single-component phase diagrams, but this is incorrect. Type I diagrams can be unary, binary or even higher. But they all have in common the plotting of one thermodynamic potential vs. another. They also share in common that the interpretation is the same for all Type I diagrams, as we will show. In fact, this is true for each category of diagrams; the rules of interpretation are identical within each type.
Figure 24.1 shows schematics of four different Type I diagrams. The first (a) is a repeat of the single-component $H_2OP - T$ phase diagram. The second (b) is a $T - \mu_{O_2}$ diagram of the two-component Ni-O system. The third diagram (c) is a slighty different version of the $Ni - O$ binary system. *This* is actually a $\mu_{O_2} - T$ diagram in disguise, as we later show. The top lines on this diagram are actually taken from a very special Type I diagram, known as the Richardson-Ellingham diagram (shown later). We will refer to this as the "Ellingham" diagram, and will spend quite a bit of time with it shortly. The final Type I phase diagram (d) is a $\mu_{S_2} - \mu_{O_2}$ diagram for the Cu-S-O ternary system. Such diagrams are referred to as "stability area" diagrams or "predominance area" diagrams. The descriptor, "stability area," is quite informative. It speaks to the fact that interpretation is identical for all Type I phase diagrams: areas represent single-phase regimes (or the "area"/range of thermodynamic potentials over which a given phase is solely "stable"); lines represent the combination of the thermodynamic potentials required for co-equilibrium of two phases; "triple points" (where three lines meet) indicate the thermodynamic conditions (potentials) where co-equilibrium of three phases occurs. Note that in going from a one-component system (H_2O) to a two component system ($Ni - O$), one of the thermodynamic potentials, in this case pressure (P =1 atm), must be held constant to arrive at a Type I diagram (equilibrium being determined by the two potentials on the axes). And in going further to a three-component system ($Cu - S - O$) two thermodynamic potentials, in this case pressure (P =1 atm) and temperature (T =1000K), must be held constant to arrive at a Type I diagram (equilibrium being determined by the two potentials on the axes.

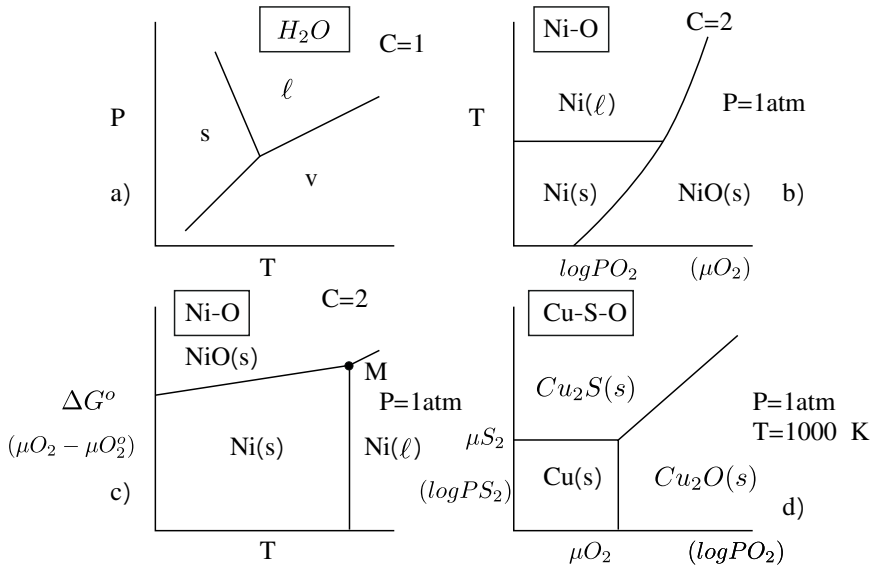


Figure 24.1: Schematic Type I Phase Diagrams for C=1, 2 and 3.

24.1 The Ellingham-Richardson Diagram

In 1944, it was observed (by Ellingham) that plots of standard free energy of oxidation of metals to oxides had essentially the same slope ($\Delta G^o \approx A + BT; B \approx const$) as long as the reaction was written per mole of oxygen gas $O_2(g)$ as in equation 24.1:

$$2\frac{x}{y}\mathrm{M(s)} + O_2(g) \rightleftharpoons \frac{2}{y}M_xO_y \tag{24.1}$$

where M in this equations represents a metal. For example, if we are dealing with the oxide MO ($x = y = 1$), equation 24.1 simplifies to equation 24.2:

$$2\mathrm{M(s)} + O_2(g) \rightleftharpoons 2MO(s) \tag{24.2}$$

But we might be dealing with a different metal, N, which forms the N_2O_3 ($x = 2$, $y = 3$), for which equation 24.1 becomes equation 24.3:

$$\frac{4}{3}N(s) + \mathrm{O_2}\,(g) \rightleftharpoons \frac{2}{3}N_2O_3 \tag{24.3}$$

Note that if we subtract equation 24.3 from equation 24.2, the oxygen term cancels out and we obtain equation 24.4:

$$2\mathrm{M}\,(\mathrm{s}) + \frac{2}{3}\mathrm{N_2O_3} \rightleftharpoons 2\mathrm{MO}\,(\mathrm{s}) + \frac{4}{3}\mathrm{N}\,(\mathrm{s}) \tag{24.4}$$

This is a powerful capability to determine whether, thermodynamically speaking, a given metal will reduce another's oxide or vice versa, as we show later. There is a simple rationale for all oxidation reactions having nearly the same slope on an Ellingham diagram. Let's consider the reaction of calcium to calcium oxide as in equation 24.5:

$$2\mathrm{Ca}(s) + \mathrm{O_2}\,(g) \rightleftharpoons 2\mathrm{CaO}(s) \tag{24.5}$$

The overall standard free energy of reaction can be determined by equation 24.6:

$$\Delta G^o = \Delta H^o - T\Delta S^o \approx A + BT \tag{24.6}$$

What we are interested in is the Ellingham slope (B) in equation 24.6, which amounts to the change in standard entropy as given by equation 24.7:

$$\Delta S^o = 2S^o_{CaO(s)} - 2S^o_{Ca(s)} - S^o_{\mathrm{O_2(g)}} \tag{24.7}$$

If we consult thermodynamic data for the three terms on the right side of equation, we obtain equation 24.8:

$$\Delta S^o = 2\,(38.1 - 41.6) - 205.1 = -212.1 \frac{\text{J}}{\text{mole} \cdot \text{K}} \tag{24.8}$$

It can be seen that the first two terms, the standard entropy terms of the two solids (calcium oxide, calcium), roughly cancel and that the overall value is dominated by the standard entropy of the oxygen gas (the third term). Hence, the slopes of all oxidation reactions on Ellingham diagrams involving solid metals and oxides will be very similar, owing to the fact that $B = -\Delta S^o \approx S^o_{O_2(g)}$ in equation 24.6. An actual Ellingham diagram is shown in Figure 24.2.

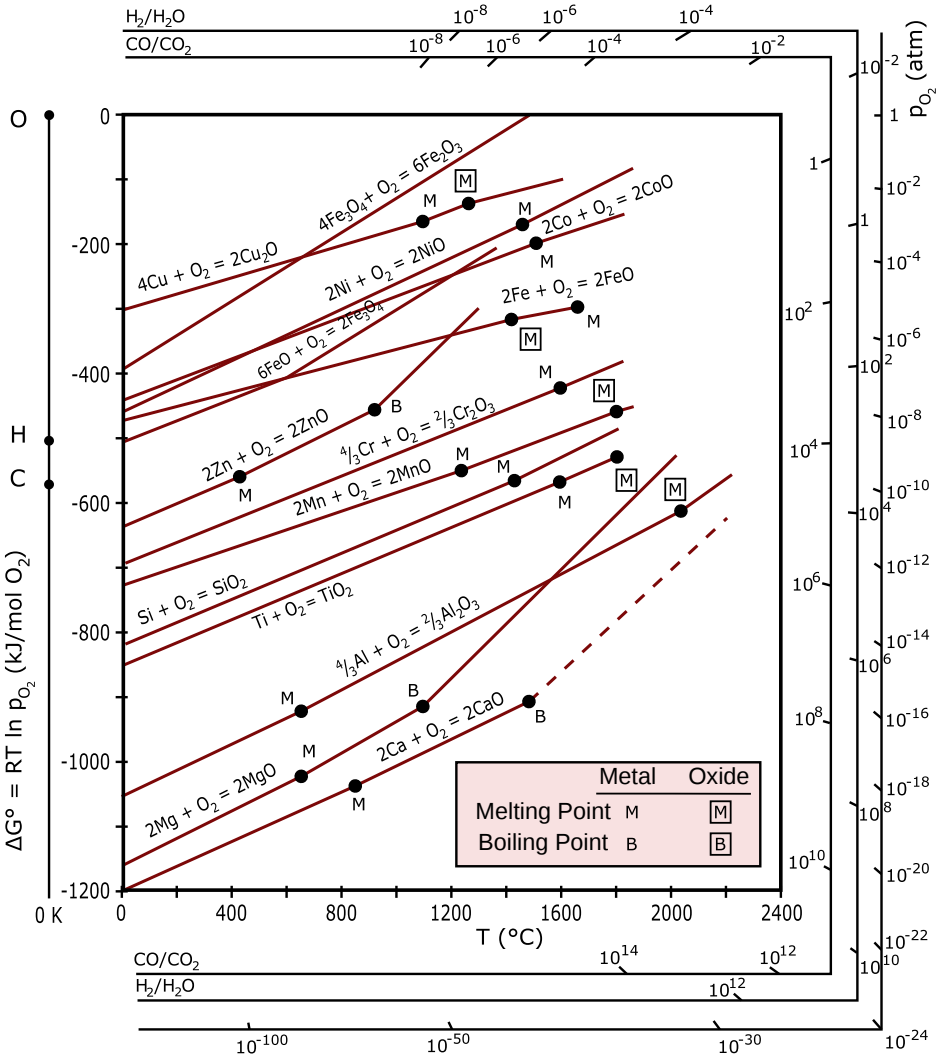


Figure 24.2: The Ellingham diagram (adapted from [27]).

At first, this may seem like a complicated diagram. However, the following discussion and "case studies" should help to simplify it and demonstrate its usefulness. As can be seen there are three different nomographic scales on the sides of the diagram. These were added by Richardson; hence, the diagram is often referred to the Richardson-Ellingham diagram. We will highlight each of these scales as we come to them. First let's consider three ways to arrive at a specific x, y or $\Delta G^o, T$ coordinate on the diagram. Consider the reaction of Ti with $O_2(g)$ to yield TiO_2 at 1000^oC. From the line and the diagram legend we know that both the metal and the oxide are solids at this temperature, because we do not encounter an "M" symbol (where the metal melts) until much higher temperature ($T \sim 1650^oC$), and there is no boxed "M" symbol, which stands for the melting point of the oxide. This means that the oxide melts at a temperature above the melting point of the metal, however no thermodynamic data are provided for higher temperatures. Note: a "B" symbol is fairly rare and corresponds to the boiling point of the metal, as in the case of Mg and Ca, the bottommost lines on the diagram. The large increase in slope at such a

"B" point is due to the fact that both reactants on the left side of the Ellingham equation 24.1 are in gaseous form (oxygen gas plus metal vapor) and therefore contribute to the entropy or slope of the line.

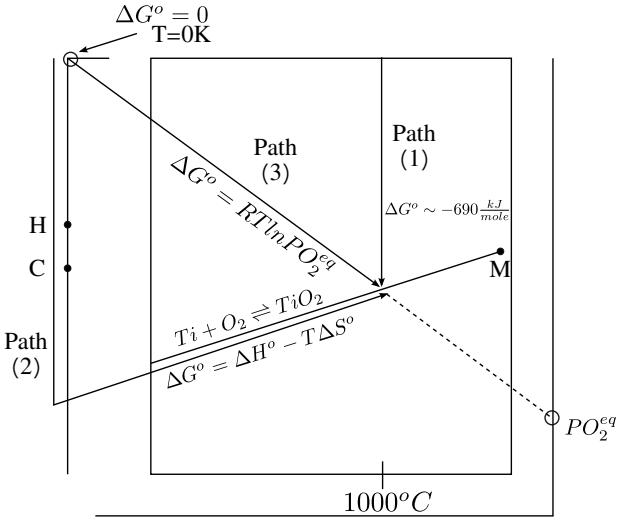


Figure 24.3: Three paths to reach a given coordinate on an Ellingham diagram.

The three ways to reach a specific coordinate are illustrated in Figure 24.3. The first way or path (1) to reach the coordinates in question is by what I call "direct read." The topmost horizontal line of the Ellingham diagram, directly below the p_{H_2}/p_{H_2O} or p_{CO}/p_{CO_2} nomographic scales, is the line of zero ΔG^o. At 1000^oC we draw an arrow down from this line until we hit the line representing the Ti/TiO_2 equilibrium at a value of $\Delta G^o \sim -690 kJ/mole$. This is illustrated in Figure 24.3. The second path (2) to reach the same coordinates is to use the Ellingham relation in equation 24.6. This is also illustrated in Figure 24.3. It is important to realize where the coordinates (0,0) occur on the diagram. On the very left of the Ellingham diagram in Figure 24.2 is a vertical line with "0 K" indicated. Since the x-axis is in degrees C, the absolute zero in degrees Kelvin is to the left another -273.15^oC. So the actual (0,0) point of the Ellingham diagram is at the top left corner where the line which passes through the "C" and "H" points makes an angle with the horizontal line of zero ΔG^o. This is a very important point on the Ellingham diagram, which I tend to call the "O" point (O for oxygen) and later the O-fulcrum. Starting at the O-point, we can draw the $A + BT$ line from the origin as shown in Figure 24.3. One can crudely think of this in terms of $\Delta H^o - T\Delta S^o$ for the Ti/TiO_2 equilibrium.

By the way, there is perfectly good thermodynamic reason why Ellingham did not extend the lines on the diagram below 0^oC. You may recall from your basic thermodynamics background that the heat capacity of a solid begins to vary dramatically below its "Debye" temperature approaching absolute zero. This would make for large deviations from linearity of the lines on the Ellingham diagram below 0^oC approaching $0K$; hence, the lines terminate at 0^oC.

The third way or path to reach the same coordinates of $\Delta G^o \sim -690 kJ/mole$ at 1000^oC requires some explanation. As found in basic chemistry textbooks we know from equation 24.9:

$$\Delta G = \Delta G^o + RT \ln Q \tag{24.9}$$

that the ΔG of a reaction is related to the standard free energy of that reaction plus a second term that depends upon the so-called "activity quotient" or Q. In the case of Ti/TiO_2 this becomes equation 24.10:

$$\Delta G = \Delta G^o + RT \ln \frac{a_{TiO_2}}{a_{Ti} p_{O_2}} \tag{24.10}$$

where the activities of the solid phases can be assumed to be unity (assuming pure metal and oxide) and the activity of oxygen is given by its partial pressure. However, if the metal and oxide are in equilibrium we know that $\Delta G = 0$, yielding the following equation 24.11:

$$\Delta G^o = -RT \ln K_{eq} = -RT \ln \frac{1}{p_{O_2}^{eq}} = RT \ln p_{O_2}^{eq} \tag{24.11}$$

where $p_{O_2}^{eq}$ is the oxygen partial pressure where $Ti(s)$ and $TiO_2(s)$ are in equilibrium. In effect, path (3) is a line with zero intercept and a slope of $R \ln p_{O_2}^{eq}$ vs. temperature, as shown in Figure 24.3. We can solve mathematically for $p_{O_2}^{eq}$ by plugging $-690 kJ/mole$ for ΔG^o and 1000^oC or rather 1273K into equation 24.11 to arrive at a value of $4.9 \times 10^{-29} atm$, for which the $\log p_{O_2}^{eq}$ (base 10) is -28.3. This is where the nomographic scale comes in handy. If we draw a line from the origin or "O-point" through the coordinates in question (the Ti/TiO_2 line at 1000^oC) to the p_{O_2} nomographic scale, we get approximately the same value. Keep in mind that this is really a log scale, so we must interpolate the "logs," for example one quarter of the way from 10^{-28} to 10^{-30} is $10^{-28.5}$ and definitely not 5×10^{-29} or 5×10^{-28}. So we have a short cut or "easy button" for finding the $\log p_{O_2}$ value for any set of coordinates on the Ellingham diagram. Simply take a ruler and connect it from the O-point through the coordinates in question and read the $\log p_{O_2}$ value off the nomographic scale. Since all lines radiate from the "O-point," I tend to refer to this point as the "O-fulcrum. You will note on the Ellingham diagram that all the "tick" marks on the O-nomographic scale point back to the O-fulcrum.

Of course, achieving such a low oxygen partial pressure is impossible with even the best available vacuum systems. That is where the outer two nomographic scales come in. These involve so-called "buffer gas systems." Consider the reaction of equation 24.12:

$$2CO(g) + O_2(g) \rightleftharpoons 2CO_2(g) \tag{24.12}$$

Let's flow an arbitrary mixture of $CO(g)$ and $CO_2(g)$ through a furnace at 1 atm total pressure. The equilibrium constant would be given by equation 24.13:

$$K_{eq} = \exp\left(\frac{-\Delta G^o}{RT}\right) = \frac{p_{CO_2}^2}{p_{CO}^2 p_{O_2}} = \frac{X_{CO_2}^2 P^2}{X_{CO}^2 P^2 X_{O_2} P} \tag{24.13}$$

where partial pressures are now expressed in terms of mole fractions and total pressure. If we let the total pressure be 1 atm and assume that the amount of oxygen produced is negligible compared to the moles of CO and CO_2 and let $r_c = X_{CO}/X_{CO_2}$ we arrive at a simplified equation 24.14:

$$K_{eq} = \frac{1}{r_c^2 p_{O_2}} \tag{24.14}$$

Let's go back to the situation we considered above, namely the Ti/TiO_2 equilibrium at 1000^oC with an equilibrium p_{O_2} of 4.9×10^{-29}. Given the ΔG^o for reaction 24.12 is $-564,800 + 173.62T$ J/mole and plugging this into equations 24.13 and 24.14, we can solve for an R value of 1.26×10^7. On a base 10 log scale this corresponds to 7.1. Now let's use the second nomographic scale and its corresponding C-fulcrum (this is the letter "C" on the line to the left side of the Ellingham diagram 24.2) to solve the same problem. Note that all the tick marks on the p_{CO}/p_{CO_2} radiate from the C-fulcrum. As illustrated in Figure 24.4, using a ruler to draw a line from the C-fulcrum through the Ti/TiO_2 line where it crosses 1000^oC all the way to the second nomographic scale and we obtain $10^{7.1}$, in excellent agreement with the calculations. Even though the p_{CO_2} would be quite small (on the order of 10^{-7} atm) this is still way larger than the value of p_{O_2}, so our assumption that the amount of oxygen can be neglected is mathematically justified. In reality, however, just as with aqueous buffers, there are limits to buffer reliability. For example, buffer gases become unreliable if the R value is too large or too small, owing to the potential for oxygen "leaks" in the lines feeding gases into a commercial furnace. Therefore, buffer gases are usually limited to values of $10^{-5} \leq r_c \leq 10^5$. Nevertheless, we have a valuable short cut to obtain the r_c value for any coordinates on the Ellingham diagram.

You will notice that there is still another nomographic scale on the Ellingham diagram of Figure 24.2. This nomographic scale involves a different buffer gas system of reaction 24.15:

$$2H_2(g) + O_2(g) \rightleftharpoons 2H_2O(g) \tag{24.15}$$

Here we are mixing hydrogen gas and water vapor, whose ratios are given along the outermost nomographic scale. Again, note that all tick marks radiate to the H-point or H-fulcrum on the line to the very left of the diagram. If we want to know a mixture of hydrogen gas and water vapor that would correspond to a set of coordinates on the Ellingham diagram, we would use a ruler to draw a line between the H-fulcrum through those coordinates to the nomographic scale, once again being careful to interpolate the log values.

Another use of the Ellingham diagram is to find a driving force for a given reaction. Consider the reaction of Mn metal with oxygen to form MnO by the reaction 24.16:

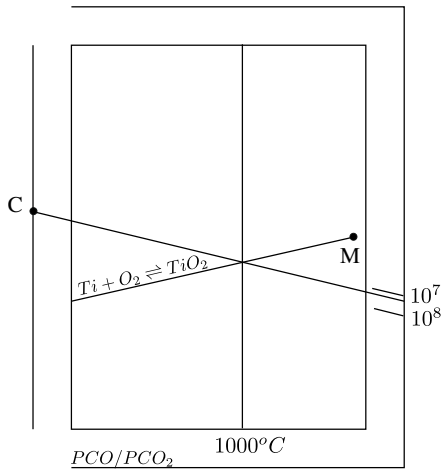


Figure 24.4: Illustrating the CO/CO_2nomographic scale on the Ellingham diagram.

$$2Mn(s) + \mathrm{O_2(g)} \rightleftharpoons 2MnO(s) \tag{24.16}$$

If we subject a mixture of Mn/MnO to an "applied" oxygen pressure of $p_{O_2}^{app} = 10^{-20}$ atm (for example, by using a buffer gas mixture) at 1000^oC, what is the driving force for the reaction to take place? There are several ways to solve for this. They each derive from equation 24.9. In this case, assuming Mn and MnO to be pure solids (activity=1) we would obtain equation 24.17:

$$\Delta G = \Delta G^o + RT \ln \frac{1}{p_{O_2}^{app}} \tag{24.17}$$

From the Ellingham diagram of Figure 24.2 we can find that the ΔG^o of the reaction is approximately -580 kJ/mole at 1000^oC. Plugging 1273K and $p_{O_2}^{app} = 10^{-20}$ into equation 24.17, we obtain $\Delta G = -92.6$ kJ/mole. But we also know from equation 24.11 that $\Delta G^o = RT \ln p_{O_2}^{eq}$. Plugging this into equation 24.17 we obtain equation 24.18:

$$\Delta G = RT \ln p_{O_2}^{eq} + RT \ln \frac{1}{p_{O_2}^{app}} = RT \ln \frac{p_{O_2}^{eq}}{p_{O_2}^{app}} \tag{24.18}$$

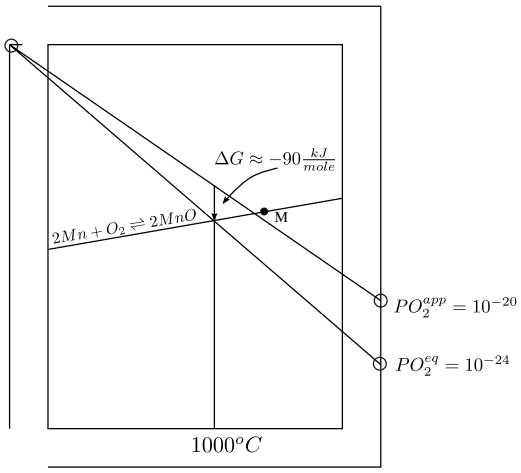


Figure 24.5: Using the Ellingham equation to obtain driving forces.

Using the O-nomographic scale on the Ellingham diagram of Figure 24.2 we can find the $p_{O_2}^{eq}$ to be very close to 10^{-24} atm at 1000^oC. Plugging this into equation 24.18 we obtain $\Delta G = -97.5 kJ/mole$. The third method is what I refer to as "direct read." This is illustrated on Figure 24.5. We always draw the arrow at constant temperature from

the *applied* condition to the *equilibrium* condition, which falls on the ΔG^o line for the reaction in question. What we obtain is approximately -90 kJ/mole. All three values are in agreement with one another, with a relatively small error determined by our ability to accurately extract values from the Ellingham diagram.
To summarize, Ellingham diagrams are characterized by the following features:

1. Curves in Ellingham diagrams for the formation of metallic oxides are basically straight lines with a positive slope. The slope is proportional to ΔS, which is fairly constant with temperature.

2. The lower position of a metal's line in the Ellingham diagram, the greater is the stability of its oxide. For example, the line for Al (oxidation of aluminum) is found to be below that for Fe (formation of Fe_2O_3).

3. Stability of metallic oxides decreases with increasing temperature. Highly unstable oxides like Ag_2O and HgO easily undergo thermal decomposition.

4. A reduced substance (such as a metal), whose Gibbs free energy of formation is lower on the diagram at a given temperature, will reduce an oxide whose free energy of formation is higher on the diagram. For example, metallic aluminum can reduce iron oxide to metallic iron, with the aluminum itself being oxidized to aluminum oxide.

5. The greater the gap between any two lines, the greater the effectiveness of the reducing agent corresponding to the lower line.

24.2 Two More Type I Phase Diagrams

So regardless of the method used to obtain the driving force, it is obvious that that driving force is negative; reaction 24.16 will proceed to the right and Mn metal will be oxidized to its oxide. This also allows us to see that we can make a Type I phase diagram out of each "line" (or metal/oxide pair) on the Ellingham diagram. If we consider the Mn/MnO "line" in Figure 24.6, it follows that a "direct read" arrow from any set of coordinates above the line (corresponding oxygen pressures larger than $p_{O_2}^{eq}$) to the equilibrium Mn/MnO line will be negative, i.e., the ΔG will be negative so the oxide will be favored. On the other hand, a "direct read" arrow from any set of coordinates below the line (corresponding to oxygen pressures smaller than $p_{O_2}^{eq}$) to the equilibrium Mn/MnO line will be positive, i.e., ΔG will be positive so the reaction of does not proceed; the metal will be stable. Therefore manganese oxide exists everywhere above the "line" and manganese metal exists everywhere below the "line." But there are two different forms of manganese metal and therefore two Type I phase boundaries: one between the solid oxide and the solid metal, below $T_m(Mn)$, and another between the solid oxide and the liquid metal, above $T_m(Mn)$. So at $T_m(Mn)$ we get a Type I triple point. The vertical phase boundary at $T_m(Mn)$ corresponds to the melting of Mn; solid Mn is stable to the left and liquid Mn is stable to the right. This vertical line intersects the other two at the triple point, where both solid and liquid Mn exist in equilibrium with solid MnO.

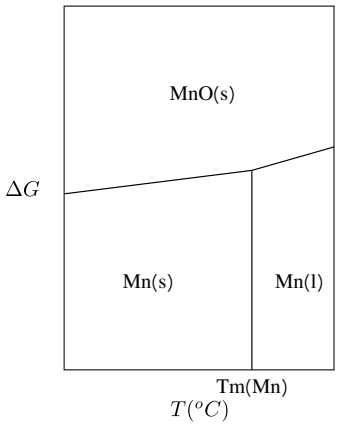


Figure 24.6: An Ellingham line turned into a phase diagram. (Note: P=1 atm.)

A more useful Type I phase diagram for the laboratory, however, is a plot of T vs. $\log p_{O_2}$. For example, we can convert the Ellingham-type phase diagram of Figure 24.6 into such a diagram for the Mn-MnO system by either 1) using the O-nomographic scale over and over to estimate the values of $\log p_{O_2}^{eq}$ for Mn/MnO equilibrium each

temperature, or 2) solving $\Delta G^o = RT \ln p_{O_2}^{eq}$ for each temperature, given the thermodynamic data for reaction 24.16. A schematic of this diagram is shown in Figure 24.7. There are many applications of such diagrams. For example, in Mn metal heat treating, we want to keep the $p_{O_2}^{app}$ below the pressure of the phase boundary with MnO. For ceramists dealing the MnO the opposite would hold true: we would want to maintain the $p_{O_2}^{app}$ above that of the oxygen partial pressure of the phase boundary with either solid or liquid Mn.

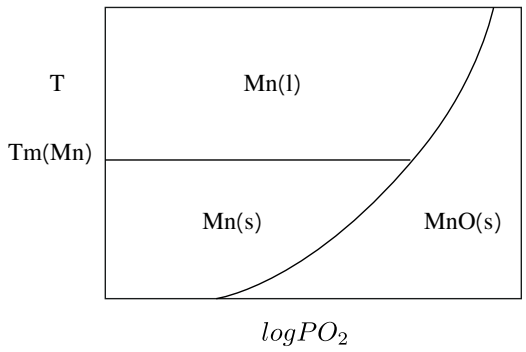


Figure 24.7: Schematic T vs. log p_{O_2} Type I phase diagram for the Mn-O system. (Note: P=1 atm.)

We can apply Gibbs' phase rule to both kinds of phase diagrams. We know that F=C+2-$\underline{P}$, however, the overall pressure is understood to be 1 atm for both diagrams, so the phase rule reduces to F=C+1-$\underline{P}$. As opposed to the water P-T phase diagram where the H:O ratio was everywhere 2 on the diagram, here the O:Mn ratio differs from phase field to phase field (e.g., it is 1:1 for MnO but 0:1 for Mn); hence, C=2 (for Mn and O). This yields a phase rule of F=3-$\underline{P}$, which means that in both the Ellingham-like Type I phase diagram of Figure 24.6 and in the T vs. log p_{O_2} phase diagram of Figure 24.7, we have the same features as we had for the H_2O P-T diagram: single-phase areas have 2 degrees of freedom, i.e., both T and log p_{O_2} must be specified, two-phase situations are phase boundaries/lines of F=1, i.e., if we fix one variable (say T), we immediately know the other (log p_{O_2}) or vice versa, and three-phase situations have zero degrees of freedom at three-line junctions or "triple points." As with the triple point of water, we have no control over the Mn(s)/Mn(l)/MnO(s) triple point (unless we increase or decrease the total pressure from 1 atm).

25 Type II (Binary) Phase Diagrams

As illustrated in Figure 23.3 Type II phase diagrams are really quite different from Type I phase diagrams. Instead of two thermodynamic potential axes, one of the axes is a ratio of conjugate extensive variables. Later, when dealing with free energy vs. composition diagrams, we will return to answering the question of how a potential axis can be replaced by a "ratio" axis to establish thermodynamic equilibrium. For now, suffice it to say that if we eliminate the -VdP term in the Gibbs-Duhem equation (by holding pressure fixed) we obtain for a two-component system:

$$0 = SdT + n_A d\mu_A + n_B d\mu_B \tag{25.1}$$

One can chose to fix T and one of the chemical potentials ($\mu_A or \mu_B$), which is relatively difficult to do, or one can chose to fix T and the ratio of the moles of B to the moles of A. Fixing n_B/n_A is inconvenient, however, since at one extreme ($n_A \rightarrow 0$) the ratio goes to infinity. Instead, as you well know, we fix the mole fraction of B, as seen in the x-axis of the common binary eutectic phase diagram sketched schematically in Figure 25.1. The following sections deal with how we can estimate each type of phase boundary (liquidus, solidus, solvus) in Type II phase diagrams from the simple solution thermodynamic models you already know (Raoultian, Henrian, Regular), as long as we make some simplifying assumptions. More complicated situations are better handled by software dedicated to predicting phase diagrams, taking into account more sophisticated models and behavior of individual solutions. Such software programs are discussed briefly at the end of this text. But for now, the following sections will build confidence in linking phase diagrams with their underlying solution thermodynamic origins, and will hopefully cause you to think about and question the specific models that lie behind the "black boxes" of modern phase diagram algorithms.

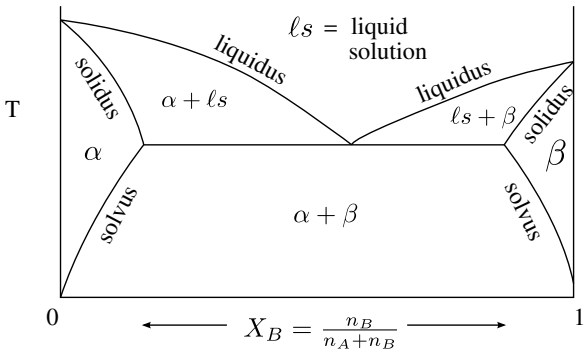


Figure 25.1: Binary eutectic with lines labeled.

25.1 Estimating Liquidus Lines on a Binary Eutectic with Negligible Solid Solution

To simplify our prediction of liquidus behavior, let's assume there to be negligible solid solubility and that the liquid is ideal or Raoultian. The former assumption is reflected in Figure 25.2 by the notations, "A" and "B," denoting nearly pure solid A and B, and ℓs denoting a liquid solution. Of course, we know from solution thermodynamics that there is no such thing as a perfectly pure solid. The assumption of liquid ideality just means that the activity of each component in the liquid is approximately equal to its mole fraction ($a_i = X_i$). Given these two assumptions, it is fairly straightforward to estimate the liquidus line for a "negligible solid solubility" system. Consider the situation in Figure 25.2:

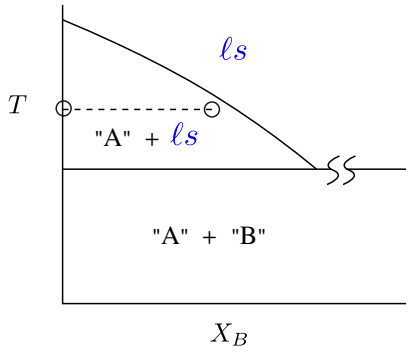


Figure 25.2: An ideal liquid in equilibrium with "pure" solid A.

Thermodynamically, the equilibrium of essentially pure solid A and the ideal liquid solution at the temperature shown in Figure 25.2 can be expressed by the equality of their chemical potentials, as shown in equation 25.2:

$$\mu_A^o(s) = \mu_A(\ell) \tag{25.2}$$

where $\mu_A^o(s)$ is the standard state chemical potential of pure solid A. From solution thermodynamics we know that the chemical potential in a solution (in this case, the liquid solution) is related to the chemical potential in the pure state by equation 25.3:

$$\mu_i = \mu_i^o + RT \ln a_i \tag{25.3}$$

We also know that the activity of the liquid can be replaced by its mole fraction (ideal solution). This leads to equation 25.4:

$$\mu_A^o(s) \simeq \mu_A^o(\ell) + RT \ln X_A(\ell) \tag{25.4}$$

Rearranging, we obtain equation :

$$-RT \ln X_A(\ell) \simeq \mu_A^o(\ell) - \mu_A^o(s) \tag{25.5}$$

But the right side of equation 25.5 is simply the ΔG_m of melting of component A per mole, which we know to be $\Delta H_m - T\Delta S_m$. At the melting point (T_m) we know that $\Delta G_m = 0$, such that $\Delta H_m = T_m \Delta S_m$ or $\Delta S_m = \Delta H_m / T_m$.

However, the temperature of equilibrium in Figure 25.2 is not the melting temperature. Modern software packages can account for changes in enthalpy and entropy for pure liquid and solid A at different temperatures. But we can at least make an estimate of what might happen by employing a further simplification, and assume that the enthalpy of melting is approximately constant and does not change significantly with temperature. This is often referred to as the "$\Delta c_p \approx 0$" approximation, namely that the difference in heat capacities between pure solid A and pure liquid A is negligible such that the enthalpy of melting is approximately temperature-independent. This gives the following equation 25.6:

$$\Delta G_m = \Delta H_m - T\Delta S_m \simeq \Delta H_m - T\left(\frac{\Delta H_m}{T_m}\right) \simeq \Delta H_m \left[1 - \frac{T}{T_m}\right] \tag{25.6}$$

By plugging this into equation 25.5 we arrive at equation 25.7:

$$-RT \ln X_A(\ell) \simeq \Delta H_m(A) \left[1 - \frac{T}{T_m(A)}\right] \tag{25.7}$$

which is a rough estimate of the point on the liquidus curve in Figure 25.2. In fact, we can solve the same equation for any temperature, starting with the melting temperature $T_m(A)$, for which the right side of equation 25.7 is zero. This requires that the mole fraction of A in the liquid be unity, i.e., $X_A(\ell) = 1$ or pure A, corresponding to the top of the liquidus in diagram 25.2. As we decrease the input temperature, the right side of the equation becomes increasingly positive, corresponding to smaller and smaller fractional values of $X(\ell)$ decreasing from unity to go along with the steady reduction in the liquidus temperature, as shown schematically in Fig. 25.2. By writing the corresponding equation for the B-liquidus at the other side of the phase diagram, again assuming negligible solid solution and ideal liquid behavior, we obtain equation 25.8:

$$-RT \ln X_B(\ell) \simeq \Delta H_m(B) \left[1 - \frac{T}{T_m(B)}\right] \tag{25.8}$$

Both liquidus lines (actually curves) are captured schematically in Figure 25.3. They fall away from the pure end-members, A and B, and can be extended far beyond the horizontal line shown on the figure. The horizontal line is where the two liquidus curves intersect. At this point, the same liquid solution is in equilibrium with both solids A and B, and equations 25.7 and 25.8 are simultaneously satisfied. According to the phase rule, there are three phases in equilibrium and there are no degrees of freedom. We have arrived at an "invariant point," which you know well as a binary eutectic.

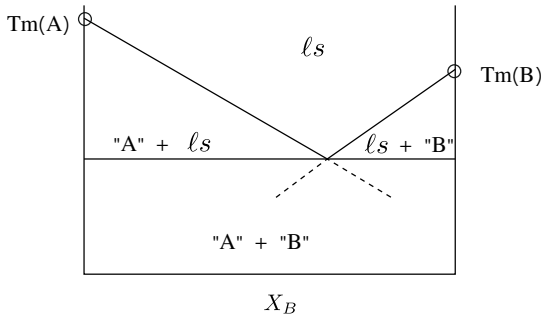


Figure 25.3: Negligible solid solubility binary eutectic diagram.

Let's pause for a moment to consider a couple of things about the behavior we have just described. First of all, the falling liquidus lines are examples of "freezing point lowering." This important phenomenon is used to advantage on icy streets by applying salt, which lowers the freezing point of ice. Of course, this requires that the ice melt and the liquid dissolve some salt as component "B" in Figure 25.2. Upon refreezing, however, the liquidus (first occurrence of solid water) is not reached until a significantly lower temperature. Furthermore, when we drop the temperature still lower, we enter a two-phase region where ice is in equilibrium with liquid salt solution, a mixture we commonly refer to as "slush."

Another good example of "freezing point lowering" is in the manufacture of Portland cement by the process of "klinkering." There are two compounds produced by this process, $(CaO)_2SiO_2$ or "C_2S" in "cement speak" and $(CaO)_3SiO_2$ or "C_3S." These compounds, when pulverized to powders and mixed with water, react to form the so-called C-S-H gel ("H" for H_2O or OH), the "glue" that upon "hardening" holds everything together in mortar (Portland cement plus sand) and concrete (Portland cement plus sand plus "aggregate"/crushed rock).

What does this have to do with "freezing point lowering? Well, with a rare exception in the dessicated regions of Israel, C_2S and C_3S do not occur in nature. If they did, they would spontaneously react with any water to form C-S-H gel. By the way, if you ever have a concrete sidewalk or driveway poured, don't let it "dry," which is a common (and disastrous!) misconception. Cement "hydration"/hardening actually consumes water, so once initial "set" has taken place, gently hose is down (or cover it with plastic) to keep it from drying out, which can lead to ruinous surface cracking.

The point here is that C_2S and C_3S are man-made compounds, and their manufacture depends upon "freezing point lowering." Both compounds have melting points in excess of 2000^oC, higher than just about any low-cost "refractories" (the high-melting ceramics used to line furnaces and kilns). But with certain "fluxing agents," for example, Fe_2O_3 and Al_2O_3, the liquidus drops dramatically to the 1350^oC to 1450^oC range. Since this is now at least a ternary system (CaO, SiO_2, Al_2O_3), we will come back to the phase diagram when considering Type III phase diagrams. For now it is enough to know that in enormous, gradually-sloped rotary kilns (pronounced "kills") the length of football fields, C_2S and C_3S "balls" (referred to by the German word, "klinker") tumble in their quasi-equilibrium liquid. Imagine an overall composition midway between "A" (C_2S/C_3S) and the eutectic composition in Fig. 25.2, but at the temperature indicated by the horizontal dashed line (a roughly 50:50 combination of solid and liquid would result). The klinker "balls" that emerge from the lower end of a cement kiln (both in terms of height and temperature), when cooled, pulverized, and ground to the consistency of fine flour become what we refer to as "Portland cement." This process is only possible owing to "freezing point lowering." By the way, Portland cement is a very important man-made material. Every year, approximately one ton of concrete is poured per capita in the developed and developing countries of the world!

One last point can be made regarding the the origin of the name "eutectic," whose Greek origins refer to "easy melting." In the binary eutectic diagram of Fig. 25.3, the liquidus lines fall away from the A and B end-members to the eutectic point, which is therefore the lowest-melting composition in the entire A-B system. We refer to the composition as the "eutectic" or easy-melting composition.

25.2 Estimating Liquidus and Solidus Lines for a Binary Isomorphous System

In the previous example we were dealing with a continuous liquid solution, but negligible solid solution. But what happens if both liquid and solid solutions are continuous across the A-B phase diagram? An example of such a phase diagram is shown schematically in Figure 25.4, where the upper line is the liquidus and the lower line is the solidus. This is a very unique situation that only happens if the end-member solids obey certain requirements, as put forth in the well-known Hume-Rothery rules: 1) the two solids must have the same crystal structure, 2) the two species should have similar electronegativities, and 3) their atomic radii must not differ by more than 15 percent. Thermodynamically speaking, the enthalpies of mixing should be nearly the same for the liquid solution as for the solid solution, or $\Delta H_{ls}^M \simeq \Delta H_{ss}^M$. For our estimation of liquidus and solidus lines we will assume both enthalpies of solution to be zero, or that both solutions are ideal or Raoultian.

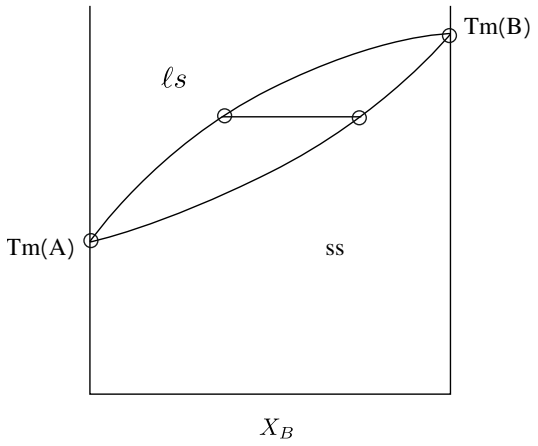


Figure 25.4: Binary isomorphous phase diagram.

Let's start with the two-phase equilibrium between liquid solution and solid solution as shown in Figure 25.4 and as expressed by equation 25.9:

$$\mu_A(\ell) = \mu_A(s) \tag{25.9}$$

However, as per equation 25.3 we can replace each side of equation 25.9 with the appropriate $\mu_A^o + RT \ln a_A$ term, giving us equation :

$$\mu_A^o(\ell) + RT \ln a_A(\ell) = \mu_A^o(s) + RT \ln a_A(s) \tag{25.10}$$

Rearranging this equation and substituting mole fractions in place of activities (we are assuming both solutions to be ideal), we obtain equation 25.11:

$$RT \ln \frac{X_A(s)}{X_A(\ell)} = \mu_A^o(\ell) - \mu_A^o(s) \tag{25.11}$$

The right side of this equation is the free energy of melting per mole of pure A, which we previously approximated by equation 25.6. Making the same simplifying approximation and rearranging, we obtain equation 25.12:

$$RT \ln \frac{X_A(s)}{X_A(\ell)} \simeq \Delta H_m(A) \left[1 - \frac{T}{T_m(A)}\right] \tag{25.12}$$

This equation only gives us the ratio of mole fractions of component A at a given equilibrium temperature. Fortunately, an analogous equation can be derived for component B:

$$RT \ln \frac{X_B(s)}{X_B(\ell)} \simeq \Delta H_m(B) \left[1 - \frac{T}{T_m(B)}\right] \tag{25.13}$$

For each given temperature, we can solve for each ratio in equations 25.12 and 25.13. If we let the two ratios be $\gamma = X_A(s)/X_A(\ell)$ and $\delta = X_B(s)/X_B(\ell)$, and remind ourselves of the fact that the mole fractions must sum to unity for each solution ($X_A(s) + X_B(s) = 1; X_A(\ell) + X_B(\ell) = 1$), we can show that:

$$\delta = \frac{X_B(s)}{X_B(\ell)} = \frac{1 - X_A(s)}{1 - X_A(\ell)} = \frac{1 - \gamma X_A(\ell)}{1 - X_A(\ell)} \tag{25.14}$$

from which the mole fraction of A on the liquidus can be obtained, namely $X_A(\ell) = (1 - \delta) / (\gamma - \delta)$. The mole fraction of A on the solidus can then be obtained from the ratio γ . Of course, to plot rather in terms of the mole fraction of B, one need only employ the $X_A + X_B = 1$ relations for each solution.

This example shows how the binary isomorphous diagram can be estimated for an A-B system, given only the two melting points and the enthalpies of melting. The process above need only be repeated systematically between the melting points of the two end-members to arrive at a diagram like the schematic in Figure 25.4.

Before the development of chemical vapor deposition, currently employed to purify silicon to transistor-grade levels (impurities at parts per billion!), a method called "zone-refining" was used to clean up crystalline ingots of silicon. The idea can be understood from the binary isomorphous diagram of Figure 25.4. Let component "B" be silicon. If we melt silicon with an impurity content at the left end of the equilibrium line in the figure, it can be seen that the solid crystallizing from the melt is significantly cleaner. Suppose we could isolate this cleaner solid. If we could repeat the process with this solid, melting it and crystallizing it, the composition of the solid would move progressively to the right, toward higher purity silicon. In fact, this was accomplished by repeated passes of a localized heater along the length of a cylindrical silicon crystal held at nearly its melting point. The "molten zone" was held in place by surface tension and it concentrated impurities and took them along for the ride. This happened to impurities with a positive "distribution coefficient," or $k = X(s)/X(\ell) > 1$ as in Figure 25.4. Impurities with a negative distribution coefficient went the opposite direction and were left behind. Either way, with each pass of the "molten zone" the central portion of the crystalline ingot became more and more pure, as the impurities were dragged/left behind at the ends, which were cut off and discarded.

25.3 Estimating Solvus Lines

It is rare when a system satisfies the conditions for a continuous solid solution. Instead, we see phase separation into two solids at low temperatures as on the schematic binary eutectic diagram of Figure 25.1. These can be entirely different crystal structures as in Figure 25.1, described as phase α and phase β. Or they can be the same crystal structure, but phase-separation into phases α_1 and α_2 occurs at low temperature. The latter behavior can be described by the regular solution model you learned about in solution thermodynamics. In the following development, let's assume that the entropy of mixing is solely "configurational," namely that it consists of only the ideal entropy of mixing in equation 25.15:

$$\Delta S^M = \Delta S^{M,id} = -R\left[X_A \ln X_A + X_B \ln X_B\right] \tag{25.15}$$

Now, for the excess free energy of mixing, let's assume the symmetrical enthalpy of mixing for a Regular solution as in equation 25.16:

$$G^{xs} = \Delta H^M = \Omega X_A X_B \tag{25.16}$$

where Ω is known as the "interaction parameter." It describes how A and B interact upon dissolving in one another. For example, the type of phase separation we will describe requires significantly large positive values of Ω, which raises the free energy of mixing, especially in the middle of the solution ($X_A \approx X_B$). The overall free energy of mixing is the sum of equation 25.15 and 25.16:

$$\Delta G^M = \Delta H^M - T\Delta S^{M,id} = X_A X_B + RT\left[X_A \ln X_A + X_B \ln X_B\right] \tag{25.17}$$

Let's take the first derivative of this equation with respect to X_B:

$$\left(\frac{\partial \Delta G^M}{\partial X_B}\right) = \Omega(X_A - X_B) + RT\left(\ln X_B - \ln X_A\right) \tag{25.18}$$

Remember that $dX_A = -dX_B$ when differentiating. It turns out that this is the equation describing the phase boundaries of the dome-shaped solvus at any temperature in Figure 25.5. We can additional useful information by taking the second and third derivatives of equation 25.17. The second derivative is:

$$\left(\frac{\partial^2 \Delta G^M}{\partial X_B^2}\right) = -2\Omega + RT\left[\frac{1}{X_A} + \frac{1}{X_B}\right] \tag{25.19}$$

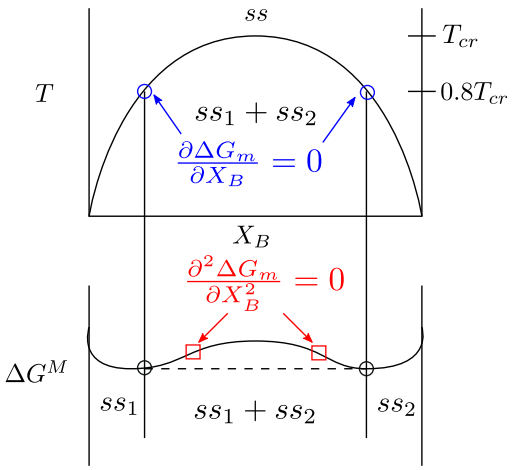


Figure 25.5: Simple solvus and free energy of mixing curve for the Regular solution model.

This equation also has special significance, namely the second derivative marks the inflection points in the ΔG^M vs composition curve of Figure 25.5. Outside of each inflection point, marked with a square, the second derivative is positive (the curve is concave up) whereas in the middle the the second derivative is negative (the curve is concave down). You will discuss in later materials science and engineering coursework the importance of these "spinodes," especially with respect to a process known as "spinodal decomposition."

To arrive at useful forms of both derivative functions, we need to take yet another derivative. It turns out that at the very top of the solvus in Figure 25.5 all three derivatives are zero. The third derivative of equation 25.17 is:

$$\left(\frac{\partial^3 \Delta G^M}{\partial X_B^3}\right) = RT\left[\frac{1}{X_A^2} - \frac{1}{X_B^2}\right] \tag{25.20}$$

Now we can begin to put all the derivatives to good use. For the third derivative to be zero requires that $X_A = X_B = 0.5$. This tells us that the top of the solvus is at the equimolar composition as shown in Figure 25.5. But the second derivative is also zero at the top of the solvus. For equation 25.19 to be zero requires that $2\Omega/RT =$

$1/X_A + 1/X_B = 1/0.5 + 1/0.5 = 4$ or $\Omega = 2RT$. We refer to this as the "critical" temperature, T_{cr}, such that $T_{cr} = \Omega/2R$. This is an important relationship. Given the critical temperature or top of the solvus in Figure 25.5, we can estimate the interaction parameter. Or given the interaction parameter, we can estimate the top of the solvus. Now let's plug these results into the first derivative equation 25.18, also setting it equal to zero:

$$-\ln X_A + \ln X_B = \ln \frac{X_B}{X_A} = \frac{\Omega}{RT}(X_B - X_A) = \frac{2RT_{cr}}{RT}(X_B - X_A) \tag{25.21}$$

or more simply:

$$\ln \frac{X_B}{X_A} = \frac{2T_{cr}}{T}(X_B - X_A) \tag{25.22}$$

There are two solutions (phase boundaries) to this equation. For example, if we chose a temperature 80% of the critical temperature or $T/T_{cr} = 0.8$ the two solutions are at $X_B = 0.145$ and $X_B = 0.855$, the latter being the symmetrical solution ($X_A = 0.145$). In the lower diagram of Figure 25.5 the lowest free energy situation between the two phase boundaries is to strike out along the dashed line, meaning an assemblage of two separate phases rather than a continuous solid solution, which is at higher free energies. It is actually easier to isolate T in equation 25.22 and solve for it by plugging in a composition. To obtain the following equation, remember that $X_A = 1 - X_B$ such that $X_B - X_A = X_B - (1 - X_B) = 2X_B - 1$.

$$T = \frac{2T_{cr}(2X_B - 1)}{\ln\left(\frac{X_B}{1-X_B}\right)} \tag{25.23}$$

Again, don't forget the two solutions at each temperature; the second X_B solution is the value of X_A for the first solution. To find the "spinodes," we need to set the second derivative equal to zero. The quantity $(1/X_A + 1/X_B)$ on the right side of equation 25.19 can be replaced by $(X_B + X_A)/X_A X_B$, which is simply $(1/X_A X_B)$. For the second derivative to be zero requires that $2\Omega/RT = (1/X_A + 1/X_B)$ or:

$$2\frac{\Omega}{RT} = 2\frac{2RT_{cr}}{RT} = \frac{4T_{cr}}{T} = \frac{1}{X_A X_B} \tag{25.24}$$

Inverting both sides of this equation yields:

$$\frac{T}{4T_{cr}} = X_A X_B = (1 - X_B)X_B = X_B - X_B^2 \tag{25.25}$$

This turns out to be a simple quadratic function, which can be readily solved for composition:

$$X_B = \frac{1 \pm \sqrt{1 - 4(\frac{T}{4T_{cr}})}}{2} = \frac{1 \pm \sqrt{1 - 4\left(\frac{T}{T_{cr}}\right)}}{2} \tag{25.26}$$

Again, there are two solutions at every temperature. For example, at $0.8T_{cr}$ the two solutions are X_B =0.276 and X_B =0.724. There is much more to spinodes and spinodal decomposition than these very simplified equations, as you will discover in higher level materials coursework.

25.4 Activity vs. Composition Plots

Although computer programs are able to do a far better job predicting the liquidus, solidus and solvus lines on Type II or "binary" phase diagrams, the previous three sections illustrate how far we can get with some very simple models and assumptions. In that same vein, we now turn to how thermodynamic activity varies with composition in the very phase diagrams considered thus far.

25.4.1 Binary Isomorphous Systems

Consider the binary isomorphous phase diagram in Figure 25.6. As above, we will consider both liquid and solid solutions to behave "ideally," meaning that they each follow Raoult's Law ($a_i = X_i$). At the melting point of B, $T_m(B)$, the activity vs. composition plot is very simple, as shown to the right. The line follows Raoult's Law, hence the "RL" label. Since pure liquid B and pure solid B are in equilibrium at $T_m(B)$, the plot will be the same regardless of which we chose to be the "standard state" at that temperature.

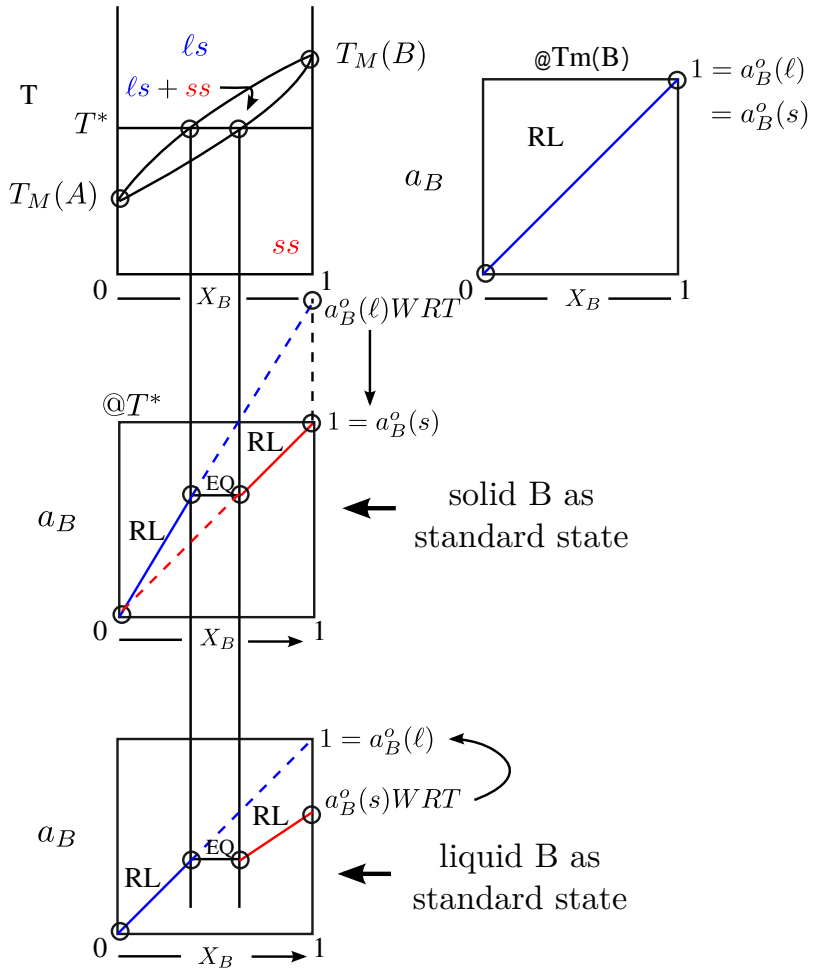


Figure 25.6: Activity vs. composition plots for an ideal binary isomorphous system.

However, consider a temperature midway between the two melting temperatures (T*), as depicted in Figure 25.6. It makes sense to chose pure solid B as the standard state, since we are well below the melting point of B. The plot immediately below the phase diagram shows how the activity changes with composition. We always start with the phase that is in the same state as the standard state. In this case we begin at the $a_B^o(s) = 1$ point (top right) and begin working backwards down the dashed Raoult's Law line until we reach the two-phase equilibrium between liquid solution and solid solution. In a two-component system at fixed temperature and pressure the degrees of freedom are 2- $\underline{P}$, or zero in the two-phase region. This requires that both chemical potential and thermodynamic activity be constant in this region, marked "EQ" for equilibrium. The leftmost regime is seemingly straightforward, since we know that thermodynamic activity must go to zero at zero composition. But this region is also labeled "RL," yet the line drawn is far from the dashed line for Raoult's Law. The reason is that we are now dealing with a liquid solution on a plot for which the standard state is pure B solid. In fact, if we extrapolate Raoult's Law in the liquid solution all the way to the right side of the activity plot, we obtain the activity of pure liquid B (if it could be obtained at this temperature) with respect to ("WRT") pure solid B being the standard state and having unit activity.

In the bottommost activity plot, we have chosen pure liquid B to be the standard state, having unit activity. But here we must start our activity plot in the regime where liquid exists, which is on the left side of the phase diagram in Figure 25.6. We know that the activity must be zero at $X_B = 0$, so (0,0) is our starting point. Since we have assumed a Raoultian liquid, the activity follows Raoult's Law ("RL") until the two-phase equilibrium ("EQ") between liquid and solid solutions, where the activity is constant. The rightmost region is also labeled "RL" for Raoult's Law, but is again far from the dashed Raoult's Law line. Again, the reason is that we are now dealing with a solid solution on a plot for which the standard state is pure B liquid (if it could be obtained at this temperature). By fitting a line from the (0,0) point through the rightmost circled point of the two-phase equilibrium and continuing it to the right side of the diagram, we obtain the activity of pure solid B with respect to ("WRT") pure liquid B being the

standard state at T*. It can be shown that the ratio of $a_B^o(\ell)/a_B^o$(s) is the same for the two diagrams. Furthermore, it is determined by the free energy of melting or fusion of component B, as we will now show.
At the two-phase equilibrium between liquid solution and solid solution, we can write the following relationships between chemical potential and activity. For the solid we get equation 25.27:

$$\mu_B(s) = \mu_B^o(s) + RT \ln \frac{a_B(s)}{a_B^o(s)} \tag{25.27}$$

and for the liquid we get equation 25.28:

$$\mu_B(\ell) = \mu_B^o(\ell) + RT \ln \frac{a_B(\ell)}{a_B^o(\ell)} \tag{25.28}$$

Until now, we have always assumed the standard state activity to be unity, namely $a_i^o = 1$ such that $a_i/a_i^o = a_i$. However, now we can make a choice as to which standard state we set equal to unity. If we consider the two-phase equilibrium in Figure 25.6, where $\mu_B(s) = \mu_B(\ell)$ and $a_B(s) = a_B(\ell)$, we can subtract equation 25.28 from equation 25.27 to obtain equation 25.29:

$$0 = \mu_B^o(s) - \mu_B^o(\ell) + RT \ln \frac{a_B^o(\ell)}{a_B^o(s)} \tag{25.29}$$

Recognizing that $\mu_B^o(\ell) - \mu_B^o(s)$ is the free energy of melting per mole of pure B or $\Delta G_m(B)$, we can use the same approximation as in Equation 25.6, namely that $\Delta c_p \approx 0$ or $\Delta H_m(B)$ is not a function of temperature to obtain equation 25.30:

$$RT \ln \frac{a_B^o(\ell)}{a_B^o(s)} \simeq \Delta H_m(B) \left[1 - \frac{T}{T_m(B)}\right] \tag{25.30}$$

This provides the explanation for how the activity vs. composition plots behave in Figure 25.6. At $T = T_m(B)$ the right side of equation 25.30 is zero; the two standard state activities are the same ($a_B^o(s) = a_B^o(\ell) = 1$). However, at other temperatures, we have the choice of which standard state activity we set equal to unity, hence the two diagrams for the temperature T*. If we divide both sides of equation 25.30 by RT and exponentiate, we find that at a fixed temperature less than $T_m(B)$ (e.g., T*) the activity ratio $a_B^o(\ell)/a_B^o(s)$ is a constant and is greater than unity, given that the right side of equation 25.30 is now positive. If we set $a_B^o(s) = 1$, it follows that the projected activity of pure liquid B (if it could exist at T*) would have an activity greater than unity. On the other hand, if we set $a_B^o(\ell) = 1$ it follows that the projected activity of pure solid B will have a value less than unity, as on the lower diagram of Figure 25.6. It should be stressed, however, that Raoult's Law is really $a_B = X_B a_B^o$. When dealing with the same "phase" (liquid or solid solution) as the standard state, it is understood that $a_B^o = 1$ and the "RL" line will fall on the dashed line in Figure 25.6. However, when dealing with the opposite "phase" (liquid or solid solution) from the standard state, Raoult's Law will still be a line, but its slope will be governed by the activity ratio $a_B^o(\ell)/a_B^o(s)$. In the lower two activity plots of Figure 25.6, we draw a line from the origin to or through the point at which we know the activity relative to the opposing standard state scale. In the bottom diagram, this linear projection ends at the activity of pure B solid with respect to pure B liquid having unit activity.

25.4.2 Binary Eutectic Systems with Dilute Solid Solutions

We can also draw schematic activity vs. composition diagrams for many binary eutectic systems by making the simplifying assumption that the solid solutions are "dilute" solutions. A dilute solution is one for which the solute (the minor component) behaves in a Henry's Law fashion and the solvent (the majority or "host" component) behaves in a Raoult's Law fashion. In fact, it can be proven that if the solute is Henrian, the solvent must be Raoultian. In a dilute solution, the solute (B) atoms are only surrounded by A atoms. It makes sense that the activity coefficient (γ_B) in the general equation 25.31:

$$a_B = \gamma_B X_B \tag{25.31}$$

will not vary with composition over the "dilute" regime until it begins to encounter other B atoms in its surroundings, giving us Henry's Law (equation 25.32):

$$a_B = \gamma_B X_B; \gamma_B \neq f(X_B) \tag{25.32}$$

Let's find out what this requires of the solvent, component A. Since $\gamma_B^{'}$ is a constant, it follows that $d\ln a_B = d\ln X_B$. But consider a version of the Gibbs-Duhem equation 25.33:

$$X_A d\mu_A + X_B d\mu_B = X_A RTd\ln a_A + X_B RTd\ln a_B = 0 \tag{25.33}$$

Dividing out the "RT" term and rearranging gives us equation 25.34, where we employ our knowledge of the solute behavior ($d\ln a_B = d\ln X_B$) to replace $d\ln a_B$ on the right by $d\ln X_B$:

$$d\ln a_A = -\frac{X_B}{X_A} d\ln a_B = -\frac{X_B}{X_A} d\ln X_B \tag{25.34}$$

But, of course, $d\ln X_B = dX_B/X_B$, $dX_B = -dX_A$ and $dX_A/X_A = d\ln X_A$ such that equation 25.34 becomes equation 25.35:

$$d\ln a_A = d\ln X_A \tag{25.35}$$

If we integrate both sides of this equation, we obtain equation 25.36:

$$a_A = const X_A \tag{25.36}$$

However, by definition, the activity of A must be unity when $X_A = 1$, requiring that the integration constant be unity, or $a_A = X_A$, which is Raoult's Law. This means that over the composition range that the solute behaves in a Henrian fashion, the solvent behaves in a Raoultian fashion. This will greatly aid us in sketching activity vs. composition plots. For the sake of simplicity, we will also make the simplifying assumption that the liquid is an ideal solution.

Figure 25.7 is a binary eutectic system with limited solid solubility; we will assume "dilute" behavior. As with the binary isomorphous phase diagram, at the melting point of B it follows that both standard state activities are unity so the activity-composition plot on the right is quite straightforward. The activity of B follows Raoult's Law across the diagram. Similarly, at the temperature T_1 the behavior is quite similar to what we saw with the binary isomorphous phase diagram. We have the choice of B two standard states: solid and liquid. In each diagram we have a two-phase equilibrium ("EQ") separating two Raoult's Law lines (one for the liquid and one for the solvent of the "dilute" solution on the right) and the ratio of $a_B^o(\ell)/a_B^o(s)$ is a constant. If we assume further that the enthalpy of melting of pure B is roughly temperature-independent, we can invoke equation 25.30 to estimate the ratio, given the melting temperature and melting enthalpy of B.

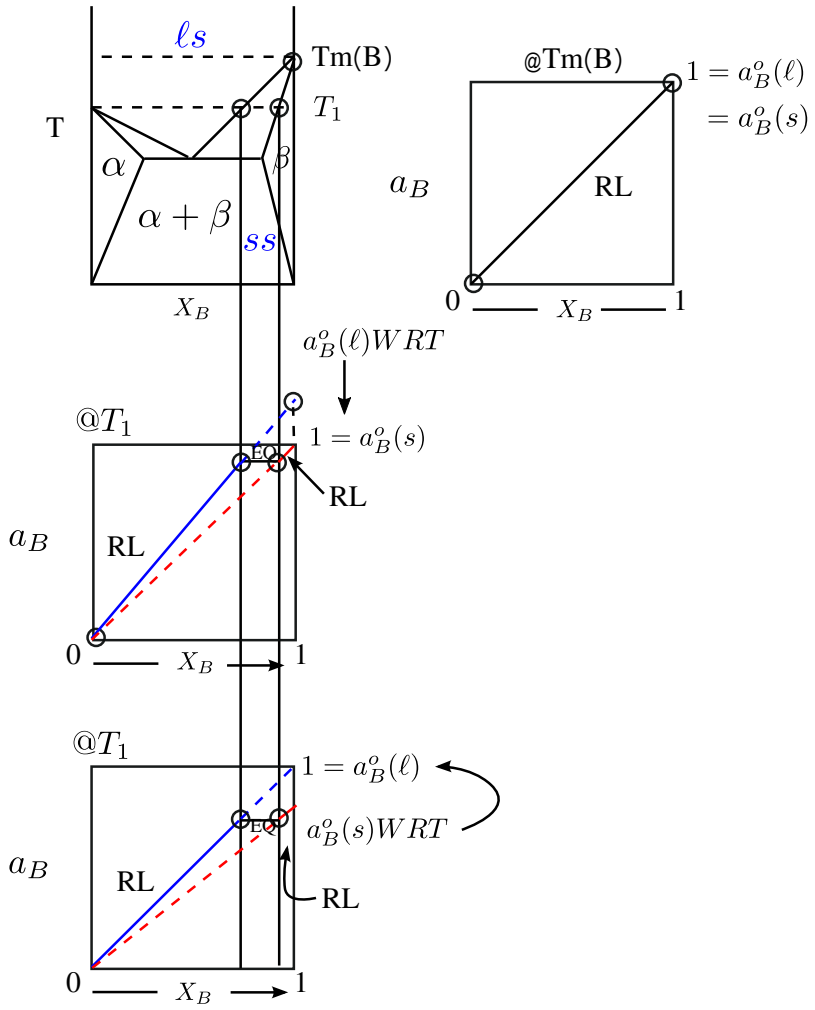


Figure 25.7: Binary eutectic with limited solid solution. Activity vs. composition plots.

At temperature T_2 the situation is a bit more complicated, as depicted in Figure 25.8. If we chose solid B as the standard state, we would begin on the right of the diagram and follow the dashed line for Raoult's Law (B being the ideal solvent) down to the phase boundary of the two-phase regime, where the activity would be constant owing to equilibrium "EQ" between liquid solution and the β solid solution. The third line is marked "RL" for Raoult's Law, since we have assumed that the liquid solution behaves this way. The reason that this "RL" line lies above the dashed line is that this is Raoult's Law in the liquid solution with respect to pure solid B having unit activity. To obtain this line, simply draw a line from the origin through the leftmost point of the previous "EQ" situation. The extrapolation of this line to the right side of the diagram would give the activity of pure liquid B (if it could be obtained at this temperature) with respect to pure solid B. Continuing to the left on the plot, we have yet another two-phase equilibrium (between liquid solution and the α solid solution) and a horizontal "EQ" situation. The final regime is strictly in the α phase, for which the B-component is the dilute solute, hence Henry's Law ("HL") is obeyed.

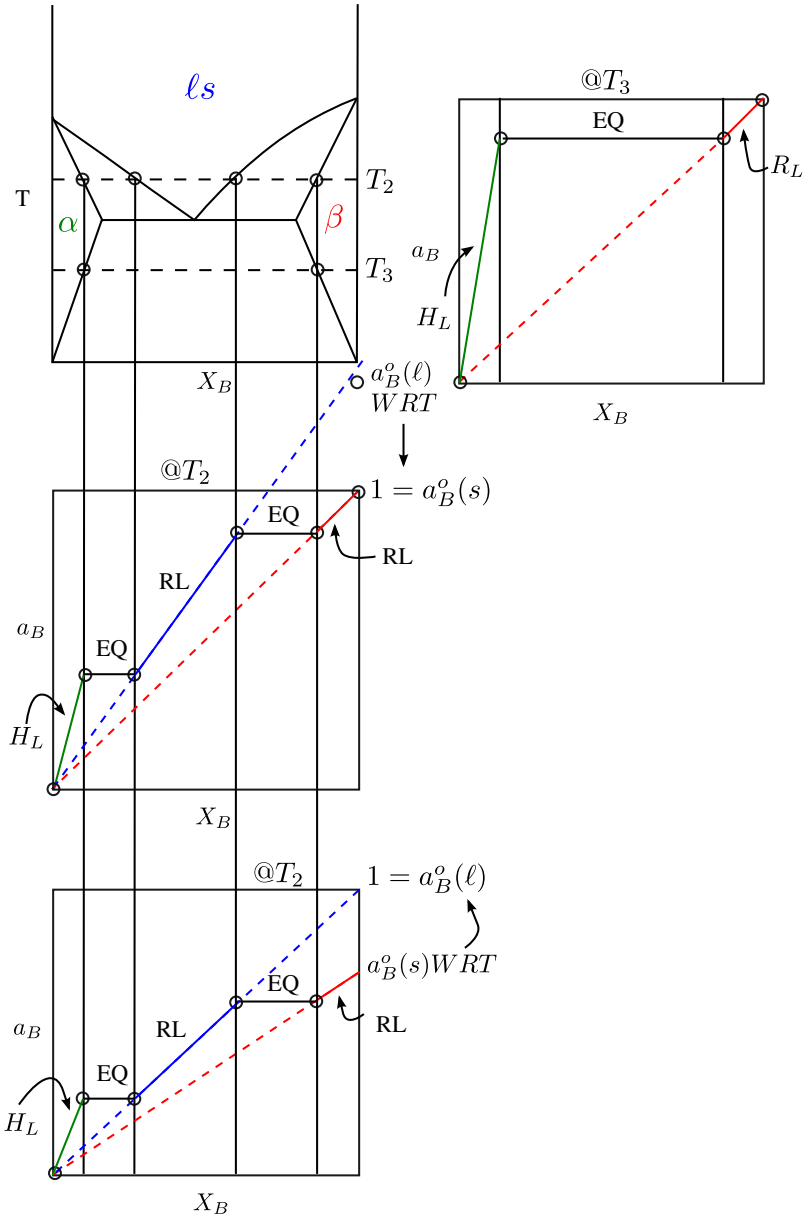


Figure 25.8: Binary eutectic with limited solid solution. More activity vs. composition plots.

But we have another choice of standard state that can be made, namely pure B liquid, shown in the bottom diagram of Figure 25.8. As before, we must begin where the phase in question is in the same state as the standard state (liquid). So we start in the very middle of the diagram where liquid solution exists by itself. The activity of B must fall on the standard Raoult's law $a_B = X_B$ dashed line as shown. At each end of this "RL" line segment we encounter two-phase equilibria between liquid solution and α solid solution on the left, and between liquid solution and β solid solution on the right. These are the two horizontal segments marked "EQ" on the plot. The final segment on the left brings the activity to zero at the origin of the plot. This segment is marked "HL" for Henry's Law, since B is the "dilute" solute in the α solid solution. For the final segment on the right, we know that B is the solvent in the β solid solution, and should behave according to Raoult's Law as marked on the diagram. However, this is Raoult's law in the solid solution with respect to pure liquid B being the standard state. Therefore the segment falls on a lower line extrapolated from the origin through the activity at the rightmost point of the two-phase liquid solution-β equilibrium ("EQ"). Where it strikes the right axis of the plot corresponds to the activity of pure solid B with respect to pure liquid B having unit activity. As with the binary isomorphous example, the ratio of $a_B^o(\ell)/a_B^o(s)$ in the two lower plots must be the same. Again, if we assume that the enthalpy of melting of pure B is roughly temperature-independent, we can invoke equation 25.30 to estimate the ratio, given the melting temperature and melting enthalpy of component B.

Below the eutectic temperature we are only dealing with solid solutions, so the situation is quite straightforward. The $a_B - X_B$ plot to the right of the phase diagram in Figure 25.8 shows how the activity of B varies with composition. We start by defining pure solid B as the standard state. Since B is the solvent in the β solid solution, the plot follows Raoult's Law until the two-phase equilibrium between the α and β solutions, the long horizontal segment marked "EQ" in the diagram. The third and final segment takes the activity of B to zero at the origin. Since B is the "dilute" solute in the α phase, this segment is marked "HL" for Henry's Law.

25.5 Schematic Free Energy vs. Composition Diagrams

In order to compare the free energies of two or more phases in Type II phase diagrams, we need to plot absolute free energies rather than the free energy change upon mixing (ΔG^M) as in Figure 25.5. Before working with such diagrams, however, we need to have two more thermodynamic "tools" in our toolbox. The first deals with determining the chemical potentials at a given composition from free energy vs. composition curves, and the second has to do with how free energy curves approach the y-axes (G-axes) at either end of such diagrams.

25.5.1 Method of Tangential Intercepts

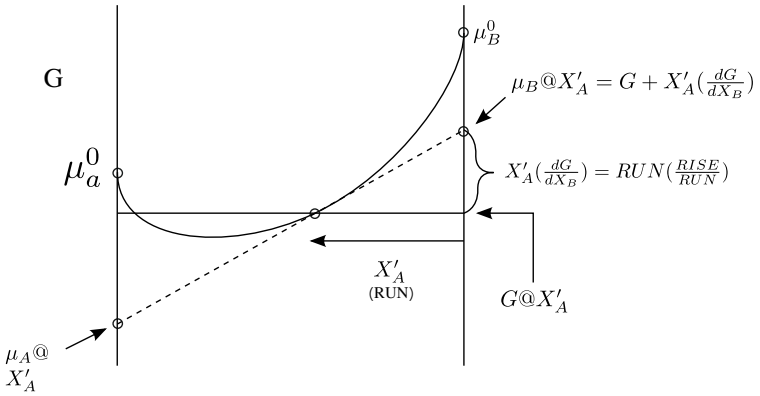


Figure 25.9: Schematic diagram of free energy vs. composition of a continuous solution.

In Figure 25.9 is a schematic diagram of absolute free energy vs. composition of a continuous solution in the A-B system. This could be a liquid solution or a solid solution. The total free energy is given by equation 25.37:

$$G' = n_A\mu_A + n_B\mu_B \tag{25.37}$$

If we divide both sides by $(n_A + n_B)$ we arrive at the free energy per mole of solution or equation 25.38:

$$G = X_A\mu_A + X_B\mu_B \tag{25.38}$$

Taking the total differential, we obtain equation 25.39:

$$dG = X_A d\mu_A + \mu_A dX_A + X_B d\mu_B + \mu_B dX_B \tag{25.39}$$

However, from the Gibbs-Duhem equation (per mole of solution) at constant temperature and pressure in equation 25.33 we have equation 25.40:

$$0 = X_A d\mu_A + X_B d\mu_B \tag{25.40}$$

Now let's subtract equation 25.40 from equation 25.39 to obtain equation 25.41:

$$dG = \mu_A dX_A + \mu_B dX_B = (\mu_A - \mu_B)dX_A \tag{25.41}$$

since $dX_B = -dX_A$. If we multiply both sides of equation 25.41 by X_B/dX_A we obtain equation 25.42:

$$X_B\frac{dG}{dX_A} = X_B\mu_A - X_B\mu_B \tag{25.42}$$

Now let's add equations 25.38 and 25.42 to obtain equation 25.43:

$$G + X_B \frac{dG}{dX_A} = \mu_A (X_A + X_B) = \mu_A \tag{25.43}$$

By analogy, we can derive a similar equation for the chemical potential of B as equation 25.44:

$$G + X_A \frac{dG}{dX_B} = \mu_B (X_A + X_B) = \mu_B \tag{25.44}$$

These are very important equations. For example, equation 25.44 tells us that if we take the slope of the G vs. X_B curve at any point (e.g., $X_A^{'}$ in Figure 25.9) and place the corresponding line on the diagram, the right intercept will be the chemical potential of B at that particular composition (μ_B at $X_A^{'}$). This follows from the left side of equation 25.44. The sum of the absolute value of G at that composition plus the product of the mole fraction of A (marked as "run" on the diagram) and the slope dG/dX_B (marked "rise/run" on the diagram) yields μ_B at that composition. Similarly, at the left end of this tangent line will be the chemical potential of component A at that same composition.

This procedure is often referred to as the "method of tangential intercepts." It is important for several reasons: 1) It conclusively proves that we can fix two chemical potentials by fixing the ratio of two conjugate extensive variables, in this case the ratio of n_B/n_A or rather the mole fraction, $n_B/(n_A + n_B)$. 2) Given a specific G vs. X_B curve, we can determine the chemical potentials at any composition. 3) Since we know the chemical potential of the pure end-members A (μ_A^o) and B (μ_B^o), namely the extreme ends of each G vs. X_B curve, we can also know the activity at any composition using the $\mu_i = \mu_i^o + RT \ln a_i$ relationship. 4) Finally, if we have two phases in equilibrium, they must share the same common tangent and intercepts so that the chemical potentials and activities are the same in both phases. We will illustrate this when we draw schematic free energy vs. composition curves for actual phase diagrams (below).

25.5.2 Terminal Slopes on Free Energy vs. Composition Plots

It can be shown that the terminal slopes on a free energy vs. composition plot should be infinitely negative as $X_B \to 0$ and infinitely positive as $X_B \to 1$. Let's begin with the regular solution model of equation 25.17:

$$\Delta G^M = \Delta H^M - T\Delta S^{M,id} = \Omega X_A X_B + RT\left[X_A \ln X_A + X_B \ln X_B\right] \tag{25.45}$$

The first derivative of this equation with respect to X_B gives us equation 25.46:

$$\frac{\partial G^M}{\partial X_B} = \Omega(X_A - X_B) + RT\left[-\ln X_A + \ln X_B\right] \tag{25.46}$$

In the limit that $X_B \to 0$ ($X_A \to 1$) the derivative becomes negative infinity and in the opposite limit where $X_B \to 1$ ($X_A \to 0$) the derivative becomes positive infinity. These hold true regardless of the size of the interaction parameter Ω. What this means is that in any free energy vs. composition plots we sketch, the terminal slopes must be $-\infty$ at the left and $+\infty$ at the right.

25.5.3 Schematic Free Energy vs. Composition Curves for a Binary Isomorphous System

Given these tools, we can sketch schematic free energy vs. composition curves for specific phase diagrams. In Figure 25.10 are three such plots for three different temperatures in the A-B binary isomorphous system. Temperature T_1 is above the melting point of component B and the liquid solution has the lowest free energy at all compositions compared to the solid solution. The situation at the melting point of B would be essentially the same, with the curve for the liquid solution being everywhere below that of the solid solution with the exception of the composition, $X_B = 1$, where the two curves would meet. The situation is reversed at temperature T_3. This temperature is below the melting point of component A and the solid solution now has the lowest free energy at all compositions compared to the liquid solution. The situation at the melting point of A would be essentially the same, with the curve for the solid solution being everywhere below that of the liquid solution with the exception of the composition, $X_B = 0$, where the two curves would meet.

At temperature T_2, however, the two curves overlap in such a way that the liquid solution has the lowest free energy on the left side of the phase diagram and the solid solution has the lowest free energy on the right side of the phase diagram. Between the two phase boundaries, however, the lowest free energy situation can be found on the "common tangent" or dashed line between the two curves. In other words, the equilibrium situation is an assemblage of the two solutions, liquid and solid. In fact, any combination of composition and temperature

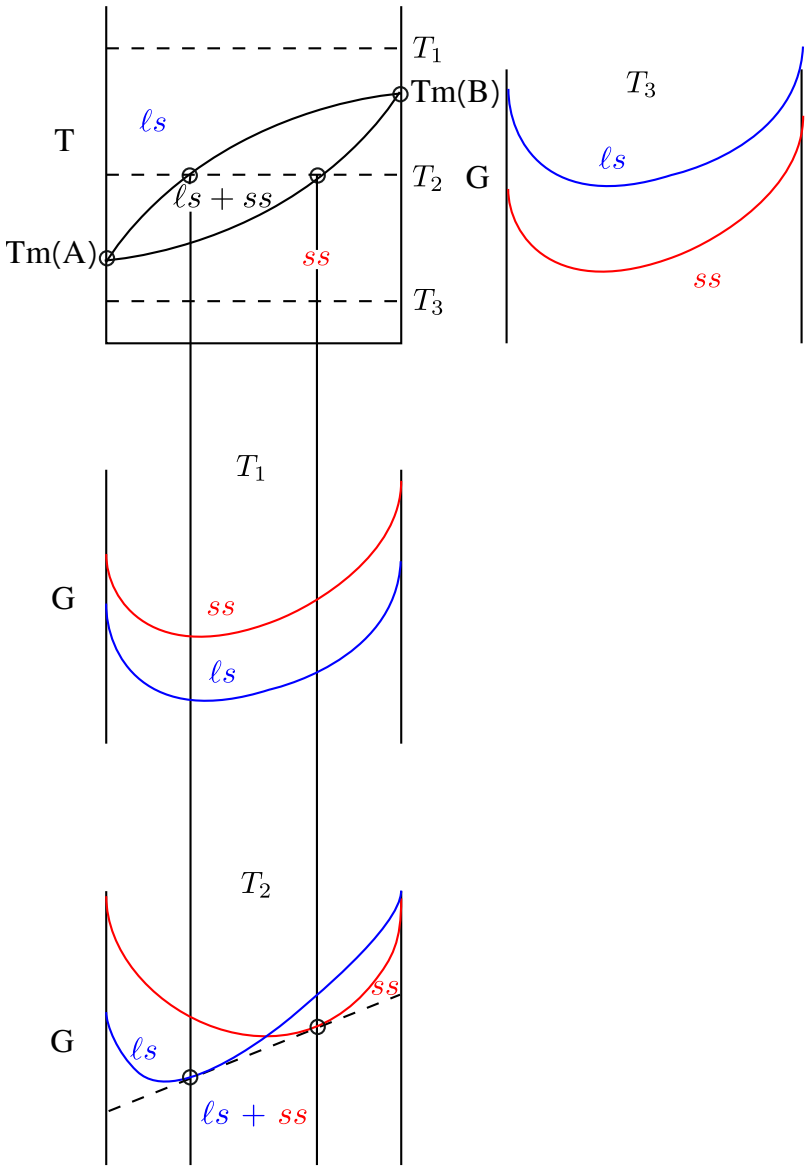


Figure 25.10: Schematic free energy vs. composition curves for a binary isomorphous system.

inside the "lens" of the phase diagram will have a microstructure consisting of a combination of liquid solution and solid solution, whose compositions are determined by the ends of the horizontal "tie line" for the temperature in question, as shown for T_2 in the phase diagram.

25.5.4 Schematic Free Energy vs. Composition Plots for Binary Eutectic Diagrams

When dealing with a binary eutectic phase diagram, we have to consider the thermodynamic interaction of three phases: the liquid solution (ls) and two solid solutions. When the α solid solution on the A side of the diagram and the β solid solution on the B side of the diagram have different crystal structures, as in Figure 25.11, there will be three distinct free energy vs. composition curves at each temperature.
The temperature T_1 is above the eutectic temperature. In the middle of the phase diagram the liquid solution (ls) has the lowest free energy. But beyond the liquidus curves to left and right we have two-phase equilibria between liquid solution and one of the solid solutions: α solid solution on the left and β solid solution on the right. The lowest free energy situation in each case is along the respective line of common tangency, as shown. In these two-phase regions, the chemical potentials can be found at the extremes of the associated common tangent line. Beyond the solidus lines to the left and right, we have either α solid solution by itself ($X_A \to 1$) or β solid solution by itself ($X_B \to 1$). In each case, the lowest free energy situation is the curve for that phase by itself (α or β).

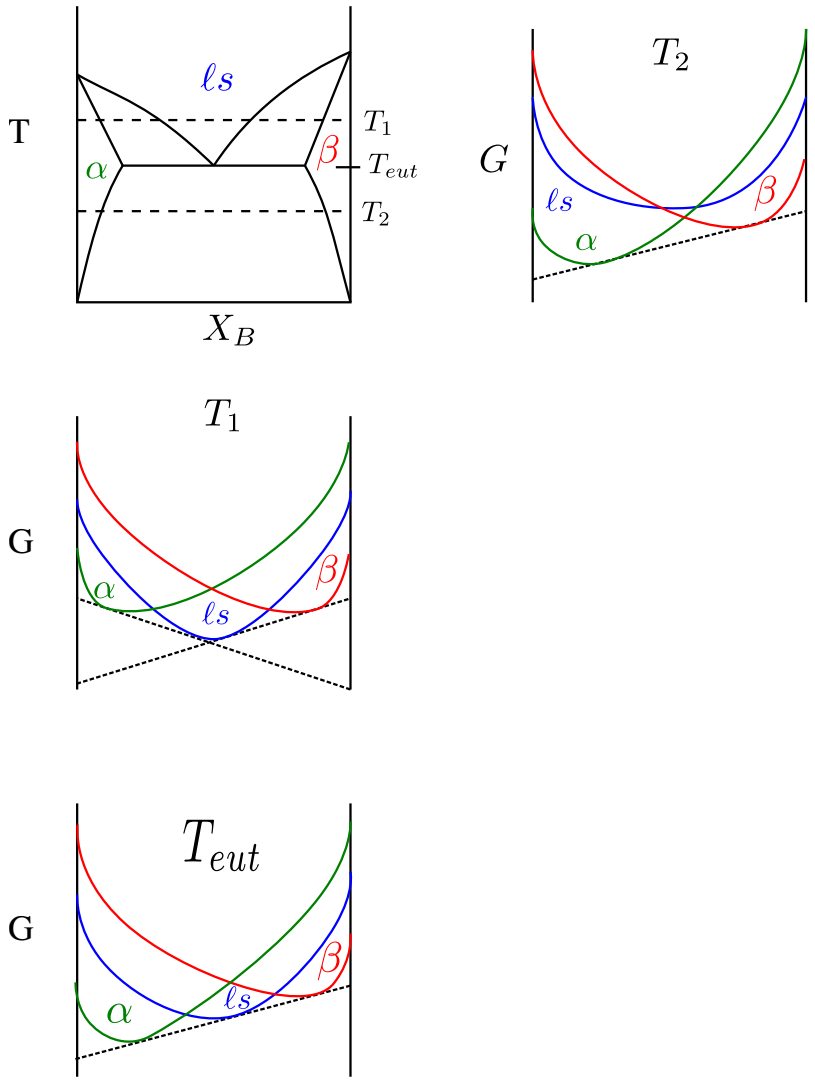


Figure 25.11: Schematic free energy vs. composition curves for a binary eutectic system with distinct phases α and β.

On the other hand, the temperature T_2 is below the eutectic temperature. The free energy vs. compostion curve for the liquid solution lies everywhere above those of the two solid solutions. The lowest free energy situation is either a single solid solution by itself, namely α solid solution or β solid solution on the left and right ends of the phase diagram, respectively, or a microstructure incorporating both α and β solid solutions in the middle of the phase diagram. Their equilibrium is reflected in the common tangent line between their free energy vs. composition curves in the diagram to the right of the phase diagram.

At the eutectic temperature all three phases (ls, α, β) are in equilibrium. This is reflected in the free energy vs. composition curves for all three phases sharing the same common tangent line at T_{eut}, hence the chemical potentials and thermodynamic activities are the same in all three phases. Only at the left and right extremes of the phase diagram do we have α solid solution by itself ($X_A \rightarrow 1$) or β solid solution by itself ($X_B \rightarrow 1$). The solidus and solvus lines intersect the horizontal eutectic line at the maximum solubilities (of B in α and of A in β), at least for this phase diagram. Any composition between these points will exhibit a three-phase microstructure involving differing amounts of α solid solution, liquid solution, and β solid solution, depending upon overall composition.

There is another way, however, to realize a binary eutectic phase diagram similar to that in Figure 25.11. Such a diagram and its associated free energy vs. composition curves are shown in Figure 25.12.

In this case, the solid solution on either side of the phase diagram is essentially the same solid solution with the same crystal structure. However, owing to interaction between the A and B atoms, the solid solution is prone to phase-separation. We described this previously, making use of the Regular solution model with a positive interaction parameter Ω. Temperature T_2 is below the eutectic temperature. As shown in the diagram to the right,

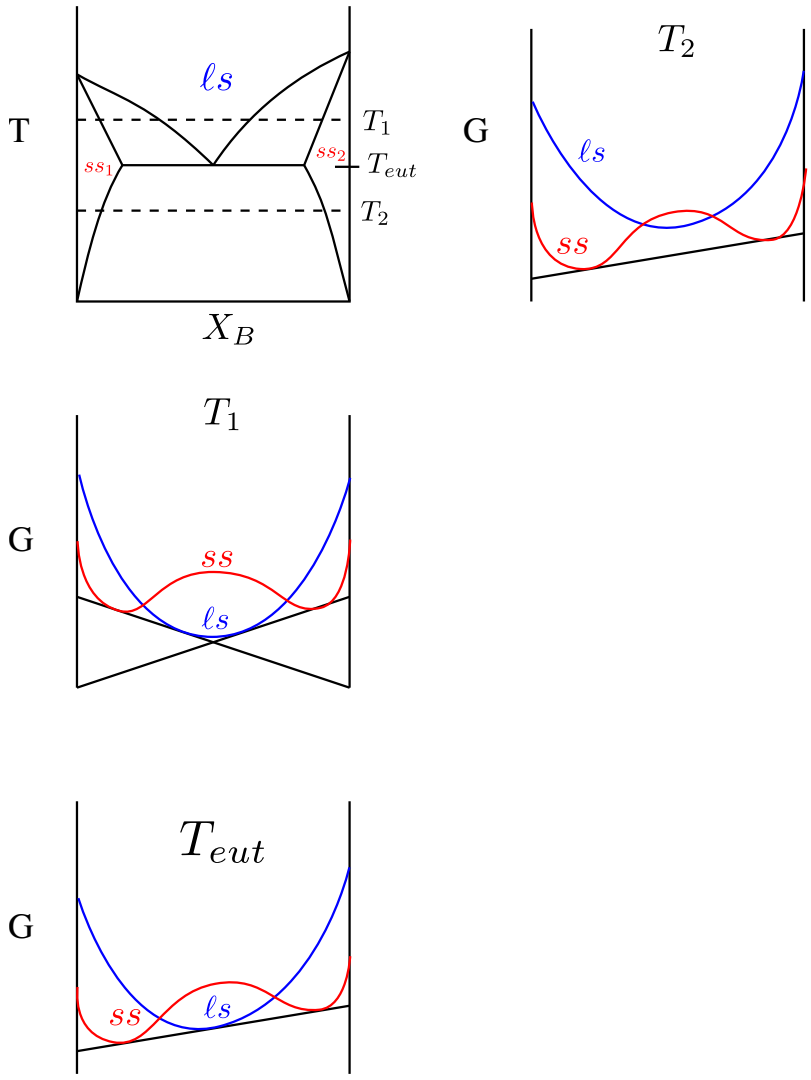


Figure 25.12: Schematic free energy vs. compostion curves for a binary eutectic system with two solid solutions of the same crystal structure.

the free energy vs. composition curve for the liquid solution lies everywhere above that of the solid solution. However, the free energy vs. composition curve for the solid solution is consistent with phase separation into two solid solutions, ss_1 and ss_2. At the eutectic temperature, all three "phases" share the same common tangent line. The solidus and solvus lines intersect the horizontal eutectic line at the maximum solubilities (of B in ss_1 and of A in ss_2). Any composition between these points will exhibit a three-phase microstructure involving differing amounts of ss_1 solid soution, liquid solution, and ss_2 solid solution, depending upon overall composition. The temperature T_1 lies above the eutectic temperature. In the middle of the phase diagram the liquid solution (ls) has the lowest free energy. But beyond the liquidus curves to left and right we have two-phase equilibria between liquid solution and either ss_1 or ss_2. The lowest free energy situation in each case is along the respective line of common tangency, as shown. In these two-phase regions, the chemical potentials can be found at the extremes of the associated common tangent line. Beyond the solidus lines to the left and right, the lowest energy situation is the curve for solid solution by itself, either ss_1 on the left or ss_2 on the right.

25.6 Regular Solution Predictions

Pelton and Thompson showed what could be done by just employing the Regular solution model for both liquid and solid solutions [29]. The resulting set of diagrams is given in Figure 25.13.
The phase diagram toward the middle of Figure 25.13 (row 3, column 3) can be easily recognized as corresponding to a binary isomorphous system, for which the interaction parameters for both solutions (liquid, solid) are equal.

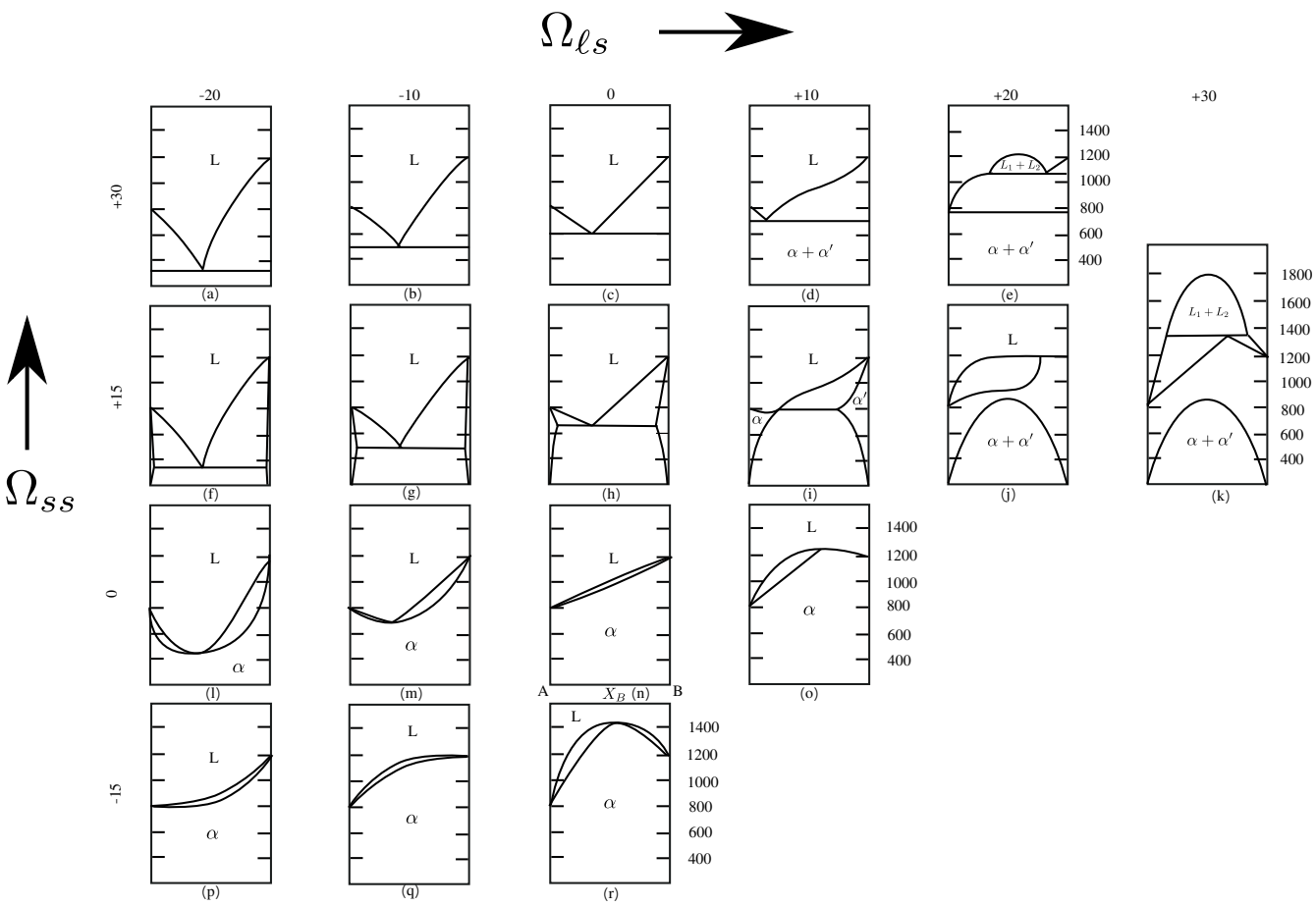


Figure 25.13: Predicted phase diagrams vs. regular solution interaction parameters (after Pelton and Thompson).

In this case, it is assumed that the interaction parameters are both zero, $\Omega_{ls} = \Omega_{ss} = 0$. In other words, both solutions are taken to be ideal or Raoultian. Note what happens as we stay in the same column ($\Omega_{ls} = 0$) but moving upward, the value of the interaction parameter for the solid solution is made progressively more and more positive. We obtain the conventional binary eutectic phase diagram, first with limited solid solubility ($\Omega_{ss} =$ 15 kJ/mole) and then with negligible solid solubility ($\Omega_{ss} = 30$ kJ/mole). The behavior at the top-right of Fig. 25.13 is particularly distinctive. This is referred to as "monotectic" behavior. If we dramatically increase the interaction parameter of the liquid solution (to +20 or +30 kJ/mole), we can produce phase-separation in the liquid, i.e., the coexistence of two liquid solutions. Such behavior is employed in the manufacture of a low-cost, high-silica glass called Vycor (by the Corning Company). Glass bodies are made out of borosilicate glass, which then phase-separates into two interconnected phases. The one with lower silica content is leached out by acid at room temperature, leaving the silica-rich phase, which is then densified at elevated temperature into an essentially dense silica body.

25.7 Another Category of Type II Phase Diagrams: Oxygen Partial Pressure vs. Composition

Just as we saw with Type I phase diagrams, there is more than one kind of phase diagram within the Type II category. In the case of Type I diagrams we became aware of several kinds, namely P vs. T, chemical potential vs. T (both Ellingham-type and log p_{O_2} vs. T), and chemical potential vs. chemical potential (stability or predominance area diagrams). For Type II diagrams we have thus far only treated T vs. mole fraction diagrams, but there are definitely other kinds. Figure 25.14 shows a schematic log p_{O_2} vs. mole fraction diagram. Note that both the total pressure (1 atm) and temperature (1000K) had to be fixed (C+2 becomes C+0 in the phase rule) in order to result in a 2-dimensional phase diagram for this three-component A-B-O system, namely, F=C+0-$\underline{P}$ or F=3-$\underline{P}$.
At first glance, this would seem to be a T vs. mole fraction phase diagram for a binary isomorphous system as in Fig. 25.4, which it closely resembles. However, this is not the case. In a T-X binary isomorphous phase diagram

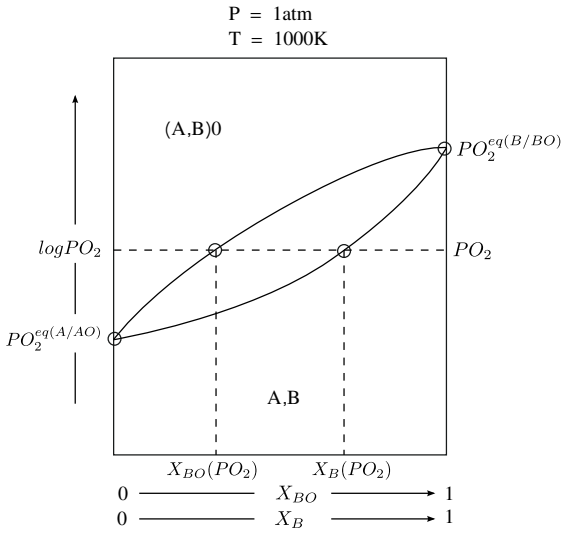


Figure 25.14: Schematic log p_{O_2} vs. mole fraction Type II phase diagram for the A-B-O system at fixed temperature and pressure.

the two phases in equilibrium at the "lens" are liquid solution (prevails at higher temperature) and solid solution (prevails at lower temperature). In Fig. 25.14 the two phases in equilibrium are an oxide AO-BO solid solution (prevails at higher p_{O_2}) and an alloy A-B solid solution (prevails at lower p_{O_2}). The x-axis is the mole fraction of B in either phase, namely $X_B/(X_A + X_B)$, which translates into $X_{BO}/(X_{AO} + X_{BO})$ in the oxide solid solution (see the dual x-axis in Fig. 25.14). Both solutions are 1) solid and 2) continuous. So this is an oxidation phase diagram. The $p_{O_2}^{eq}$ values at either end of the "lens" correspond the point of oxidation of pure metal to pure oxide, either A to AO (on the left) or B to BO (on the right). Since the interpretation of all phase diagrams within a given Type is the same, the "lens" corresponds to two-phase equilibrium between an oxide solid solution (ceramists never refer to solid solutions as "alloys!") and an alloy solid solution, as at the ends of the tie-line at the selected p_{O_2} in the phase diagram.
In fact, we can actually simulate the phase diagram in Fig. 25.14. Of course, this is an extremely unique situation, involving continuous solid solutions of both phases. This would mean that Hume-Rothery rules would need to be satisfied for both solid solutions. For our simulation, we will assume that both solid solutions behave ideally, i.e., they behave according to Raoult's law such that $a_A = X_A$ and $a_B = X_B$ in the alloy and $a_{AO} = X_{AO}$ and $a_{BO} = X_{BO}$ in the oxide solid solution. We begin by writing the oxidation reaction for component A as equation 25.47:

$$2A(s) + \mathrm{O_2(g)} \rightleftharpoons 2AO(s) \tag{25.47}$$

We can write the equilibrium relationship for this reaction as follows in equation 25.48:

$$K_1 = \exp(-\Delta G_1^o/RT) = \frac{a_{AO}^2}{a_A^2 p_{O_2}} \tag{25.48}$$

We can evaluate the equilibrium constant K_1 in one of two ways. First, we can find the value of ΔG_1^o at 1000K from the Ellingham diagram in Fig. 24.2. Second, we can use the p_{O_2} nomographic scale on the Ellingham diagram to solve the following equilibrium equation :

$$K_1 = \frac{a_{AO}^2}{a_A^2 p_{O_2}} = \frac{1^2}{1^2 p_{O_2}^{eq(A/AO)}} \tag{25.49}$$

We are dealing here exclusively with the point at the leftmost end of the "lens" in Fig. 25.14, where essentially pure A metal is in equilibrium with pure AO oxide. The value of $p_{O_2}^{eq(A/AO)}$ can be established by drawing a line from the O-fulcrum through the 1000K intersection of the A/AO line to the O-nomographic scale on the Ellingham diagram of Fig. 24.2. A similar process can be carried out at the right side of the "lens" in Fig. 25.14 involving the oxidation reaction for component B as per equation 25.50:

$$2B(s) + O_2(g) \rightleftharpoons 2BO(s) \tag{25.50}$$

yielding an analogous equilibrium relationship as described by equation 25.51:

$$K_2 = \exp(-\Delta G_2^o/RT) = \frac{a_{BO}^2}{a_B^2 p_{O_2}} = \frac{1^2}{1^2 p_{O_2}^{eq(B/BO)}} \tag{25.51}$$

As for the A/AO equilibrium, we can establish the value of K_2 from either the ΔG_2^o value at 1000K from the Ellingham diagram, or alternatively the value of $p_{O_2}^{eq(B/BO)}$ by drawing a line from the O-fulcrum through the 1000K intersection of the B/BO line to the O-nomographic scale on the Ellingham diagram of Fig. 24.2.
This gives us the two endpoints of the "lens," but how can we establish the phase boundaries of the "lens" for, say, a given p_{O_2} value as illustrated by the tie line in Fig. 25.14? This can be accomplished be recognizing that both equilibrium relationships of equations 25.49 and 25.51 must also hold for every p_{O_2} value between the extrema ($p_{O_2}^{eq(A/AO)}$ and $p_{O_2}^{eq(B/BO)}$). Furthermore, the situation is simplified by our assumption of Raoultian behavior for each solid solution. We can therefore write two equilibrium expressions 25.52 and 25.53:

$$K_1 = \frac{a_{AO}^2}{a_A^2 p_{O_2}} = \frac{X_{AO}^2}{X_A^2 p_{O_2}} \tag{25.52}$$

$$K_2 = \frac{a_{BO}^2}{a_B^2 p_{O_2}} = \frac{X_{BO}^2}{X_B^2 p_{O_2}} \tag{25.53}$$

Taking into account that $X_A + X_B = 1$ and $X_{AO} + X_{BO} = 1$,we can divide equation 25.53 by equation 25.52 to eliminate the unknown p_{O_2} to arrive at equation 25.54:

$$\frac{K_2}{K_1} = \frac{X_{BO}^2 X_A^2}{X_B^2 X_{AO}^2} = \frac{(1-X_{AO})^2 X_A^2}{(1-X_A)^2 X_{AO}^2} \tag{25.54}$$

Taking the square root of both sides and rearranging, we obtain equation 25.55:

$$\frac{(1-X_{AO})}{X_{AO}} = \left(\frac{K_2}{K_1}\right)^{1/2} \frac{(1-X_A)}{X_A} \tag{25.55}$$

Since at each temperature, including 1000K, the ratio K_2/K_1 is a constant, we can solve equation 25.55 for a unique value of X_{AO} for each inputted value of X_A. We can then plug the values of X_A and X_{AO} into equation 25.52 to solve for the corresponding value of p_{O_2}. By varying the value of X_A, we can solve for individual tie-lines ($\log p_{O_2}, X_A, X_{AO}$). Once again, $X_B = 1 - X_A$ and $X_{BO} = 1 - X_{AO}$. Intrepreting the phase diagram in Fig. 25.14 is straightforward. In each of the single-phase regions there are two degrees of freedom, i.e., to establish thermodynamic equilibrium (in addition to fixing overall pressure (1 atm) and temperature (1000K), one would have to fix the overall composition (B/A ratio) and the oxygen partial pressure. Inside the "lens" however, there is a single degree of freedom. If we fix the oxygen partial pressure, both phases' compositions are fixed. Alternatively, we need only fix one composition of the two phases in equilibrium, say X_B, and the composition of the other phase, X_{BO}, plus the oxygen partial pressure are thereby also fixed.

26 Type III (Ternary) Phase Diagrams

As we begin to introduce Type III phase diagrams, it is useful to compare and contrast their appearance with that of Type I and Type II phase diagrams as schematically represented in Fig. 26.1.
Type I diagrams are distinctive in their absence of tie-lines. Areas are always single-phase regions, and lines are always phase boundaries between two adjacent phases. Type II diagrams are distinctive because of their having tie-lines perpendicular to the "thermodynamic potential" (ϕ_i) axis. The dashed lines are usually not drawn as in the Type II diagram of Fig. 26.1; but they are understood. In fact, since we can always draw another tie-line between any adjacent pair of tie-lines, there is an infinite array of tie-lines within the three two-phase regions on the diagram. However, when these tie-lines converge to unique three-phase (F=O) invariancies, these are represented as solid lines perpendicular to the ϕ_i-axis. For example, we are quite familiar with the horizontal invariant lines in T vs. mole fraction eutectic (as in Fig. 25.1) or eutectoid diagrams (see Fig. 22.2). So there are two different

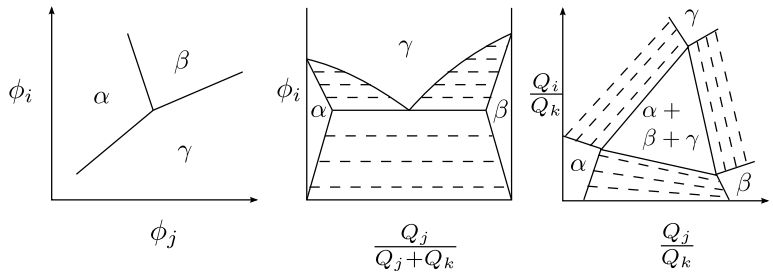


Figure 26.1: Schematic comparison of Type I, Type II and Type III phase diagrams.

"areas" in Type II diagrams: single-phase regions (α, β, γ) and two-phase regions consisting of an infinite array of parallel tie-lines perpendicular to the ϕ-axis ($\alpha+\beta, \alpha+\gamma, \gamma+\beta$). Type III diagrams are distinctive owing to the presence of tie-triangles involving three phases in equilibrium (α, β, and γ in Fig. 26.1). They are also distinctive in that tie-lines need neither be parallel nor perpendicular to any axis, as shown in Fig. 26.1. In contrast with Type II diagrams, tie-lines for two-phase equilibria are typically drawn. However, once again since we can always find another tie-line between any adjacent pair of tie-lines, so there is an infinite array of tie-lines within the two-phase regions on the diagram. Furthermore, as these converge to the sides of unique three-phase equilibrium triangles, these solid lines become very important, representing the terminal compositions of the two phases in equilibrium with that of the third. So lines are either phase boundaries of single-phase regions (can be curved) or the straight line sides of tie-triangles. As can be seen, open "areas" can either be single-phase regions (α, β, γ) or tie-triangles. These tie-triangles are usually referred to as "compatibility" triangles, meaning that the terminal compositions of the three phases are "compatible" and do not react with one another.
Various axis schemes have been employed for Type III phase diagrams, as illustrated in Fig. 26.2.

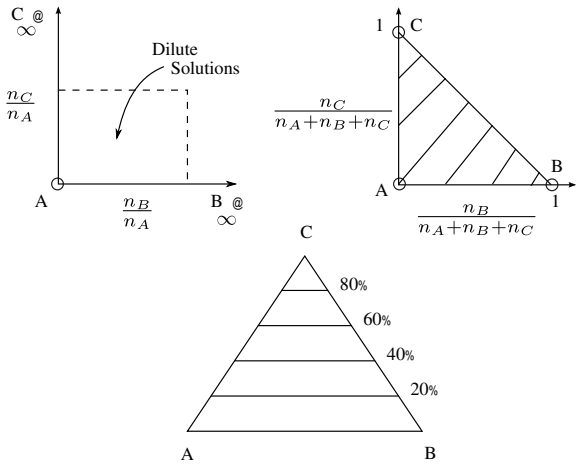


Figure 26.2: Axis schemes for Type III phase diagrams.

The strictly Q_i/Q_k vs. Q_j/Q_k scheme at the top-left finds limited use in the literature, but can be useful for representing phase equilibria in dilute solutions. The reason that this representation is not practicable for full-scale diagrams has to do with the loci of the end-members B and C on the diagram. Component A lies at the origin, however pure B and pure C exist at n_B/n_A and n_C/n_A values of inifinity on their respective axes. The mole fraction C vs. mole fraction B diagram at the top-right corrects this problem. Solutions to the overall mole fraction equation, $X_A + X_B + X_C = 1$, exist only within the dashed region. Again, this axis scheme has found limited use in the literature. The near-universal axis scheme used by materials scientists and engineers involves collapsing the diagram at the upper right of Fig. 26.2 into an equilateral triangle as shown in the bottom diagram of the Figure. Since this axis scheme was developed by J. Willard Gibbs of Gibbs free energy fame, it is usually referred to as a Gibbs phase triangle. The mole fraction of a given component is read from equally-spaced lines drawn parallel to the opposing side of the diagram, as shown for component C in Fig. 26.2. Comparable lines can be drawn parallel to the other two sides of the diagram, to establish the mole fractions of components A and B for a specific composition.

26.1 The Ternary Lever Rule

The Gibbs phase triangle axis scheme actually is a specific case of the ternary lever rule. This is represented for a more general three-phase (α, β, γ) equilibrium in Fig. 26.3.

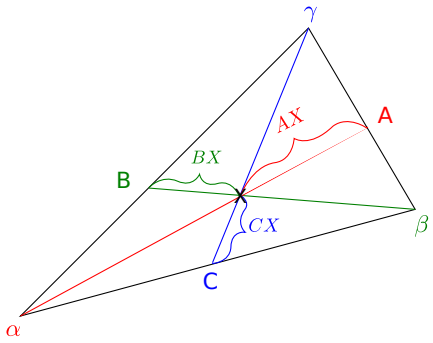


Figure 26.3: The ternary lever rule.

The ternary lever rule is applicable to all tie-triangles, and not just equilateral ones. The best approach is to employ a ruler to measure the length of line segments in the specific triangle. For example, take the composition marked by the "X" in Fig. 26.3. The fraction of the α phase, f^{α} is the length of the line segment marked AX divided by the entire length of the line drawn from the $\beta - \gamma$ side of the tie triangle through the overall composition "X" to the α vertex. Similarly, the fractions of the β and γ phases can be determined from their relative line lengths, so we have:

$$\begin{aligned} f^{\alpha} &= \frac{AX}{A\alpha} \\ f^{\beta} &= \frac{BX}{B\beta} \\ f^{\gamma} &= \frac{CX}{C\gamma} \end{aligned} \tag{26.1}$$

The three fractions sum to unity. It can be readily seen that when this scheme is applied to an equilateral triangle, we obtain the axis scheme associated with Gibbs' phase triangle. We will employ the ternary lever rule quite frequently when dealing with so-called "horizontal sections" and crystallization sequences in so-called "liquidus projection" diagrams.

26.2 Dealing with an Additional Degree of Freedom

It can be easily demonstrated that we need 3-dimensions to do justice to Type III or "ternary" phase diagrams. Figure 26.4 shows such a schematic three-dimensional phase diagram for the system A-B-C with negligible solid solubility.
Since there are three components (C=3), even if pressure is fixed at 1 atm such that (C+2) becomes (C+1), the phase rule will be F=4-$\underline{P}$. This means that we need three axes to fully represent the phase equilibria: temperature (T), and two of the three mole fractions in the equation, $X_A + X_B + X_C = 1$, e.g., the mole fraction of component A (X_A) plus the mole fraction of B (X_B), hence the 3-D diagram in Fig. 26.4 with temperature (T) as the vertical axis. The major features of this diagram are the three "mountain peaks" known as "primary phase fields" in phase diagram parlance, which meet at three "valleys" or liquidus phase boundaries that descend from the "passes" (think of "Donner Pass") at the binary eutectics on the bounding binary eutectic phase diagrams into the interior of the ternary diagram until all three valleys converge at the ternary eutectic (think of "Jackson Hole").

26.3 Liquidus Projection Diagrams

This is all very visual, but we tend to be "flat-landers," preferring 2-dimensional representations. One way to achieve 3-D perspective on a 2-D diagram is by projection, as in the bottom of Fig. 26.4. This is referred to as a "liquidus projection diagram." To capture all the 3-D information, however, the projection needs to include liquidus "isotherms" or lines of constant liquidus temperature, as shown in Fig. 26.5 for the NaCl-NaF-NaI system. One can think of such diagrams as being like topographic maps, which contain "contour lines" of constant altitude. We will return to the "mountain/valley/hole" analogy and and use these isotherms to advantage when we discuss

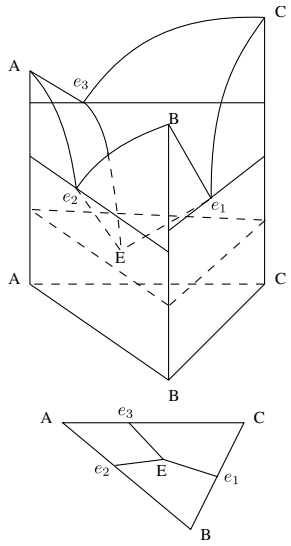


Figure 26.4: Schematic 3-D phase diagram for the A-B-C system with negligible solid solution.

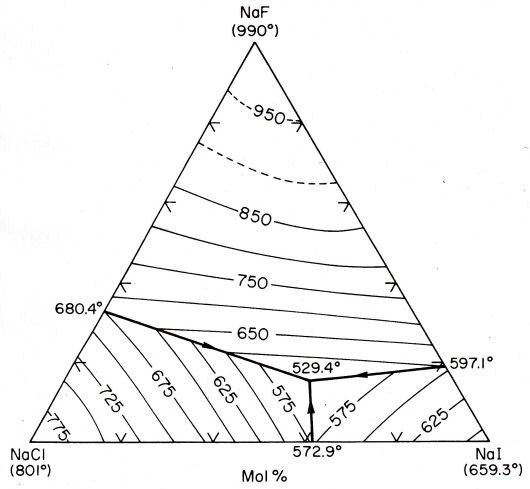


Figure 26.5: Liquidus projection diagram for the NaCl-NaF-NaI system (diagram 5908 from Phase Diagrams for Ceramists).

"isothermal sections" (below). Note that C=3 (NaCl, NaF, NaI) rather than C=4, owing to the fact that, as with the H_2O Type I phase diagram (C=1), the ratios of Na:Cl, NaF, and Na:I remain fixed.

We can actually produce a schematic liquidus projection diagram based upon some thermodynamic data (melting points and enthalpies of melting for the end members) and some simple assumptions, namely that there is negligible solid solubility and that the liquid solution behaves in a Raoultian fashion. These assumptions are not far from reality for many ceramic systems, for which fulfillment of the Hume Rothery rules (same crystal structure, same cation valence, similar electronegativities, similar cation radii) is fairly rare. We previously found an equation for the liquidus line in a binary system with negligible solid solubility and an ideal liquid (equation 25.7). We employed two such equations to establish the two liquidus curves and their intersection (the binary eutectic) in Fig. 25.3. By combining each of the three possible pairs of equations, we can produce the bounding binary eutectic phase diagrams in Fig. 26.4. However, if we assume the ternary liquid solution to be Raoultian, namely $a_i = X_i$ for each component, there is no reason that the same procedure cannot be extended into the ternary diagram. First, we write the three equations for each of the three components:

$$-RT \ln X_A(\ell) \simeq \Delta H_m(A) \left[1 - \frac{T}{T_m(A)}\right] \tag{26.2}$$

$$-RT\ln X_B(\ell) \simeq \Delta H_m(B)\left[1-\frac{T}{T_m(B)}\right] \tag{26.3}$$

$$-RT\ln X_C(\ell) \simeq \Delta H_m(C)\left[1-\frac{T}{T_m(C)}\right] \tag{26.4}$$

The procedure for predicting the liquidus projection diagram using these three equations is outlined in Fig. 26.6.

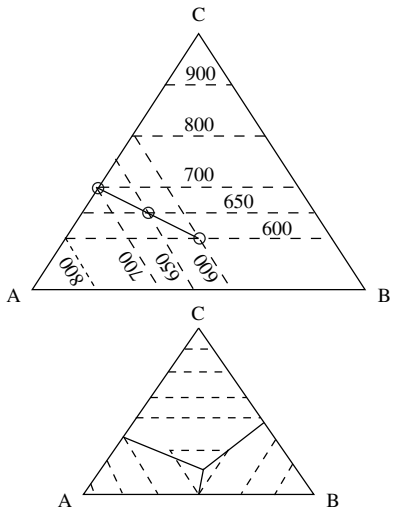


Figure 26.6: Procedure for calculating a liquidus projection diagram for a negligible solid solution ternary.

Solutions for each equation are lines parallel to the opposite boundary of the phase diagram, i.e., at constant mole fraction of that particular component. For example, at 900K the solution to equation 26.4 is a constant mole fraction of C as shown ($X_C = const$). Similarly, at 800K the solution to equation 26.2 is a constant mole fraction of A as shown ($X_A = const$). Note that if we keep ramping down the temperature the two solutions meet at the temperature of 700K, the binary eutectic temperature for the A-C system. But the solutions can extend to smaller mole fractions as well. This is the genesis of the solid line extending into the interior of the phase diagram from the binary eutectic (see solutions for 650K and 600K). This would require smaller value of X_A and X_C than at the binary eutectic, where $X_A + X_C = 1$. However, this is not a problem since we are now dealing with ternary compositions, such that $X_A + X_B + X_C = 1$. We are now following the A plus C plus liquid phase boundary ("valley") descending from the "pass" of the A-C binary eutectic into the interior of the phase diagram as in Fig. 26.6. If we do the same procedure for the A plus B plus liquid phase boundary and the B plus C plus liquid phase boundary, the three valleys will meet at the ternary eutectic point ("E" in Fig. 26.4).
The resulting "liquidus projection diagram" should look something like the bottom diagram in Fig. 26.6. There are obvious differences between our predicted phase diagram (Fig. 26.6) and an actual liquidus projection diagram (see Fig. 26.5). The main difference is that predicted isotherms are straight and parallel to one another and parallel to the opposing side of the triangle. This is an automatic result of the ideal liquid solution assumption underlyling equations 26.2 through 26.4. In the real phase diagram, although the isotherms are approximately parallel, they are curved and deviate substatially from being parallel to the opposing side of the triangle. This simply makes us challenge our assumptions underlying our predictions. First, the liquid solution may be far from Raoultian and second, there may actually be appreciable solid solubility.

26.4 Hummel's Rules for Ternary Systems with Negligible Solid Solubility

We can relax the Raoultian liquid specification in our considerations, but if a given ternary phase diagram involves negligible solid solubility it can be shown that very specific rules hold for the interpretation of liquidus projection diagrams. Once we have laid the foundation for interpretation of such Type III phase diagrams, we can turn to more sophisticated simulation engines/algorithms to help us interpret ternary diagrams with appreciable solid solubility. For examples, many metal alloy ternaries fall into the latter category.
In order to introduce Hummel's Rules (adapted from F. A. Hummel's class notes and his textbook, *Phase Equilibria in Ceramic Systems*, Marcel Dekker, 1984), we need to develop a binary analogy. We need to think carefully

about how to generate both eutectic and peritectic behavior in the absence of solid solubility. Conventional binary eutectic and peritectic phase diagrams are displayed in Fig. 26.7. We have already described the "easy melting" character of eutectic systems, but the melting/heating reactions are quite different between eutectic and peritectic diagrams. We can describe them as follows in equations 26.5 and 26.6:

$$\alpha + \beta \rightleftharpoons \text{liquid solution} \tag{26.5}$$

$$\alpha \rightleftharpoons \beta + \text{liquid solution} \tag{26.6}$$

In these reactions, the forward arrow stands for heating and the reverse arrow stands for cooling. From the Greek, "peritectic" means "covered melting." This can be seen in that the "primary phase field" of β (the two-phase area where only β is in equilibrium with liquid) overhangs or covers the α phase. Another way to describe peritectic melting is that the α phase melts to a liquid of a different composition and another solid (β). This is referred to as "incongruent melting," as described below.

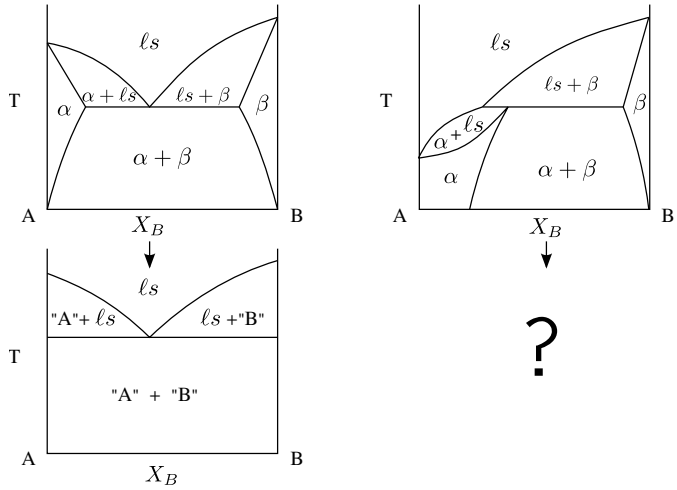


Figure 26.7: Schematic binary eutectic and peritectic phase diagrams, and what happens as solid solubility goes to zero.

But now let's imagine what happens on both diagrams as the solid solubilities, namely the solubility of B in α and the solubility of A in β, go to zero. This works for the eutectic system (as in Fig. 25.3), where the quotation marks around "A" and "B" remind us that thermodynamically speaking, there is no such thing as an absolutely "pure" solid. In contrast, note that it is impossible to preserve peritectic behavior when solid solubilities go to zero in a simple binary peritectic diagram of Fig. 26.7. However, we can introduce peritectic behavior in a negligible solid solubility system by introducing an intermediate compound (AB), as shown in Fig. 26.8. The diagram on the left has an intermediate compound (AB), which melts "congruently," namely it melts to a liquid of identical composition, AB. In fact, we can recognize that this diagram can be thought of as consisting of two side-by-side eutectic diagrams: A-AB and AB-B. However, this cannot be said of the diagram on the right. It similarly has an intermediate compound (AB), however in this case the compound (AB) melts "incongruently" to a liquid of a quite different composition (the "peritectic" composition at p_1) and a different solid ("B"). It obeys the peritectic melting behavior described by equation 26.6. The apparatus beneath the two phase diagrams will be discussed shortly.

As we introduce each of Hummel's Rules we will, 1) make reference to the binary analogues in Fig. 26.8, and 2) try to place each rule on a firm solution thermodynamic footing. Before we proceed, however, we need to introduce and define common phase diagram terminology. The first term is "primary phase," referring to the first phase to crystallize from the melt on cooling, thus the major or "primary" solid phase in the developing microstructure. For example, in the dual-eutectic diagram on the left side of Fig. 26.8 between the pure A end-member and the binary eutectic, e_1, A is the "primary phase." The second term is closely related. A "primary phase field" is the area in a liquidus projection diagram where a single (primary) phase is in equilibrium with liquid. For example, in the dual-eutectic diagram on the left side of Fig. 26.8 between the two eutectic points, e_1 and e_2, exists a quite large "primary phase field" for compound AB. In contrast, the primary phase field of AB is significantly compressed in the eutectic-peritectic diagram on the right side of Fig. 26.8, existing only between the eutectic (e_1) and peritectic (p_1) points. The third term, "subsolidus compatibility," requires additional explanation. First of all, we need to

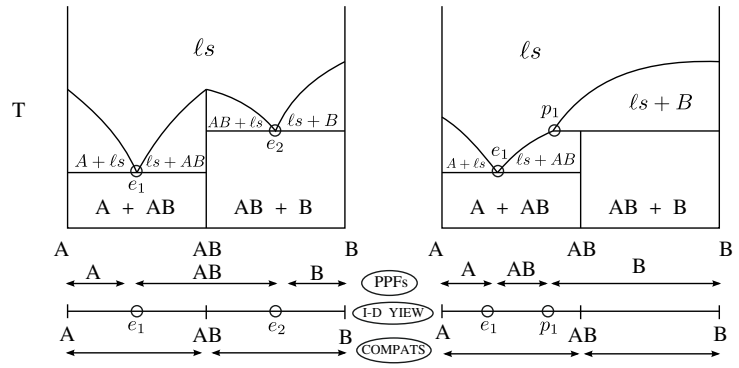


Figure 26.8: Binary analogues for eutectic and peritectic melting involving a congruetly-melting (left) and incongruently-melting (right) intermediate compound AB.

understand what "subsolidus" means. On the left side of Fig. 26.9 we see a conventional binary eutectic phase diagram with limited solid solubility. The "solidus" lines are circled on the left and right and correspond to the first appearance of liquid upon heating, or the last occurrence of liquid upon cooling; hence "subsolidus," meaning "below the solidus." However, you will see that the horizontal eutectic line is also circled. For all compositions between the ends of this line (the extrema of solid solubility) the first liquid appears on heating upon crossing this line. Conversely, upon cooling the last liquid disappears upon cooling through this line. Therefore, the horizontal eutectic line can be thought of as part of the "solidus." If we shrink the solid solubilities to essentially zero, we obtain the phase diagram on the right side of Fig. 26.9. Here, "subsolidus" applies to all compositions when below the eutectic temperature. We can now introduce two additional terms that are valid for negligible solid solubility systems: "subsolidus compatibility join" and "subsolidus compatibility triangle." A "subsolidus compatibility join" is any line connecting two phases that are compatible, meaning that they do not react to form other phases in the subsolidus. In Type III ternary phase diagrams we have both "subsolidus compatibility joins" and "subsolidus compatibility triangles," the latter referring to composition triangles connecting three compatible solid phases in equilibrium in the subsolidus. As mentioned previously with regard to the distinctive characteristics of the various phase diagram Types in Fig. 26.1, compatibility triangles are the hallmark of Type III diagrams.

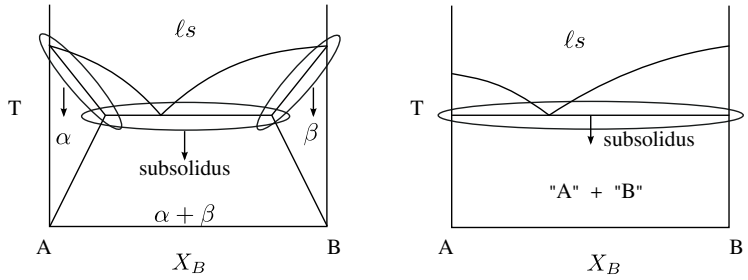


Figure 26.9: Identifying subsolidus regimes in binary eutectics with and without appreciable solid solubility.

We can now put the analogues and the apparatus at the bottom of Fig. 26.8 to good use. Hummel's first rule states that "liquidus surfaces always fall away (in temperature) from the corresponding primary phase." Thermodynamically, we can understand this from the familiar equation 26.7, developed for situations where the "primary phase" exhibits negligible solid solubility and the liquid solution behaves in a Raoultian fashion.

$$-RT \ln X_i(\ell) \simeq \Delta H_m(i) \left[1 - \frac{T}{T_m(i)}\right] \tag{26.7}$$

where "i" stands for the Raoultian liquid host (A, B, AB). Starting with the melting temperature ($T_m(i)$), for which the right side of equation 26.7 is zero, requiring the mole fraction of component "i" in the liquid to be 100%, we can see that as the temperature decreases the right side of this equation becomes increasingly positive, requiring increasingly fractional mole fractions of "i." In other words, the liquidus falls away (in temperature and in composition) from the corresponding primary phase. This can be readily seen in Fig. 26.8; in both diagrams the liquidus in a given primary phase field always falls away from the corresponding primary phase. This holds true even for the incongruently melting intermediate compound, AB, in the eutectic-peritectic diagram of Fig. 26.8. Even

though the primary phase field is at some distance from the AB composition, it can be seen that the liquidus still falls away from the primary phase (AB).

Hummel's second rule state that "there are as many primary phase fields as there are primary phases that melt." To understand this rule, let's take an imaginary helicopter trip up above "mount A," "mount B" and "mount AB" in each of the diagrams in Fig. 26.8. Looking down on each diagram, it is easy to identify three "mountains" or primary phase fields, one for A, one for B, and one for AB. It is true that we only see a "shoulder" of mount AB in the eutectic-peritectic diagram on the right of Fig. 26.8. Nevertheless, the primary phase field is visible for the AB primary phase. In each diagram, there are three primary phase fields (PPFs) and three primary phases (A, AB, B). This may seem overly obvious on these binary diagrams, but this rule will be a big help when dealing with ternary liquidus projection diagrams, where primary phase fields are areas rather than lines. The rule refers to "primary phases that melt" for good reason. Figure 26.10 shows an intermediate compound with "an upper temperature limit of stability," namely the compound AB decomposes into solid A and solid B well before it has the chance to melt. Above the decomposition temperature, it is as if the system is a straightforward A-B eutectic system. Believe it or not, it is possible for the compound AB to have a primary phase field in a ternary A-B-C system, but this complication is beyond the scope of the present treatment; hence the disclaimer referring to "primary phases that melt."

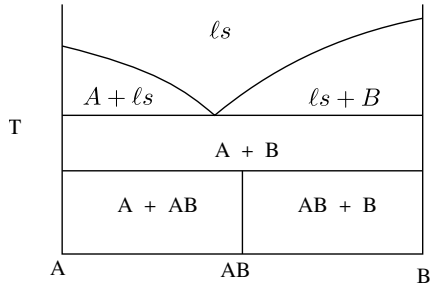


Figure 26.10: Binary system with an intermediate compound having an upper temperature limit of stability.

Hummel's third rule states that "there is a one-to-one correspondence between liquidus invariant points and the corresponding subsolidus compatibilities." If we go down in temperature into the subsolidus, there are two subsolidus compatibilities in each diagram of Fig. 26.8, namely A-AB and AB-B. In other words, for any composition between A and the intermediate compound, AB, in the subsolidus there will be two phases in equilibrium: A and AB. Similarly, for any composition between AB and B in the subsolidus, there will be two solid phases in equilibrium: AB and B. Since there are precisely two subsolidus compatibilities in each diagram, from Hummel's third rule we can anticipate two invariant points. On the dual-eutectic diagram on the left side of Fig. 26.8, there is an invariant point where the primary phase fields of A and AB meet (e_1) and one where the primary phase fields of AB and B meet (e_2). On the eutectic-peritectic diagram to the right side of Fig. 26.8, there is an invariant point where the primary phase fields of A and AB meet (e_1) and one where the primary phase fields of AB and B meet (p_1). This rule may seem simplistic when dealing with the uncomplicated binary diagrams in Fig. 26.8, but will be quite powerful when dealing with a complicated ternary liquidus projection diagram. If there are 10 subsolidus compatibility triangles, there will be precisely 10 ternary invariant points, with a one-to-one correspondence between them.

Hummel's fourth rule states that "invariant points inside their respective subsolidus compatibilities are eutectics, whereas those outside their respective subsolidus compatibilities are peritectics." Looking in 2-D at the two phase diagrams in Fig. 26.8, it is quite obvious which points are eutectics and which are peritectics. But now imagine that we are back in the helicopter hovering above the 1-D phase diagrams (see the lines labeled "1-D view" in the apparatus beneath each diagram). It is clear on the left diagram that the point where the primary phase fields of A and AB (PPFs) come together lies within the corresponding A-AB subsolidus compatibility (see "compats" in the apparatus). This is therefore a eutectic. The same is true for the eutectic-peritectic diagram on the right side of Fig. 26.8. In the case of the invariant point involving AB and B, the situation is different between the two phase diagrams. From the helicopter, we can see in the apparatus for the diagram on the left side of Fig. 26.8 that the point where the primary phase fields of AB and B come together lies within the corresponding AB-B subsolidus compatibility; hence, this is a eutectic. In contrast, the point where the primary phase fields of AB and B come together in the diagram on the right side of Fig. 26.8 lies outside the corresponding AB-B subsolidus compatibility; hence, this is a peritectic. Again, these relationships may seem quite simplistic when dealing with binary phase diagrams, but rule no. 4 will be quite powerful when dealing with complex liquidus projection

diagrams. Ternary invariant points inside their corresponding subsolidus compatibility triangles will be eutectics; those falling outside their subsolidus compatibility triangles with be peritectics.
Let's begin to illustrate the first four Hummel's rules with regard to the liquidus projection diagram in Fig. 26.11.

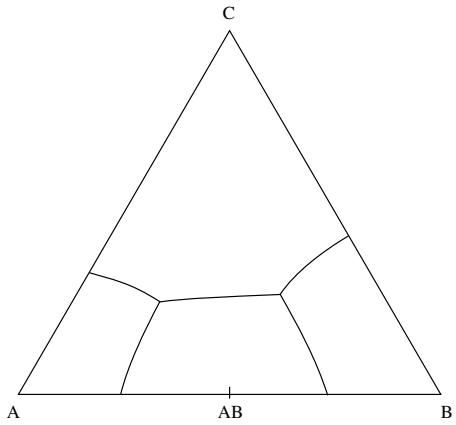


Figure 26.11: Liquidus projection diagram for the A-B-C system with negligible solid solubility and an intermediate compound AB.

It may seem that this diagram is useless in the absence of primary phase field labels and liquidus projection isotherms, but application of Hummel's rules will show this not to be the case. First, let's apply rules no. 2 and 3: "there are as many primary phase fields as there are primary phases that melt" and "there is a one-to-one correspondence between liquidus invariant points and the corresponding subsolidus compatibilities." If we go around the diagram we can count four "primary phases that melt," namely the end-members A, B, C, and the binary compound AB. In ternary phase space, primary phase fields are areas. We can identify four primary phase fields in ternary space. We can make assignments beginning from the vertices (end-members) and then dealing with any binary and ternary compounds. This is done in Fig. 26.12.

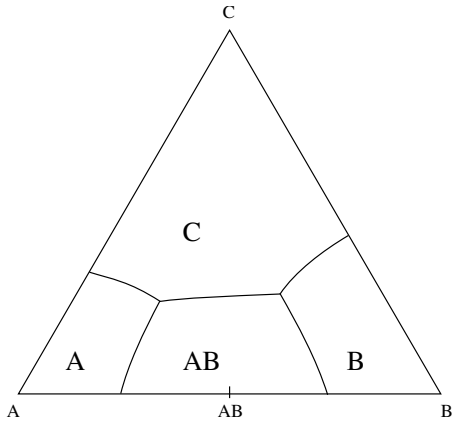


Figure 26.12: A-B-C liquidus projection diagram with primary phase fields labeled.

Using the third rule ("there is a one-to-one correspondence between liquidus invariant points and the corresponding subsolidus compatibilities") we can now sketch in the subsolidus compatibilities. There are two ternary invariant points (intersections of three phase boundaries or "valleys" as well as intersections of three primary phase fields or "mountains," namely where the primary phase fields of A, C, and AB come together and where the primary phase fields of AB, C, and B come together). The one-to-one correspondence of Hummel's third rule requires that there be two subsolidus compatibility triangles: A-C-AB and AB-C-B, as shown in Fig. 26.13.
Furthermore, Hummel's rule no. 4 allows us to identify the nature of the ternary invariant points: "invariant points inside their respective subsolidus compatibilities are eutectics, whereas those outside their respective subsolidus compatibilities are peritectics." In this case, both invariant points are inside their respective subsolidus compatibility triangles; hence, they are both ternary eutectics, as labeled in Fig. 26.13. In other words, the point where the primary phase fields of A, C, and AB converge is inside the corresponding A-C-AB compatibility triangle, and similarly for the PPFs of AB, C, and B.

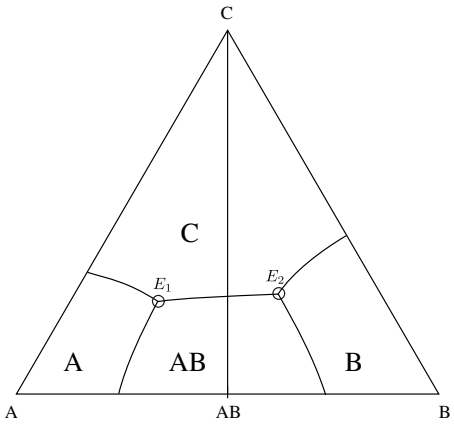


Figure 26.13: A-B-C liquidus projection diagram with subsolidus compatibility triangles.

However, there is much more that can be done with a schematic liquidus projection diagram, as captured in Hummel's fifth and sixth rules. The fifth rule, in combination with rule no. 1, allows us to draw in "directions of falling temperature" on all the liquidus phase boundaries or "valleys." It states, "the direction of falling temperature on a liquid-solid(1)-solid(2) phase boundary is always away from the corresponding solid(1)-solid(2) join." This follows from what we did in predicting liquidus phase boundaries ("valleys") in the simple ternary eutectic phase diagram of Fig. 26.6 (assuming ideal liquid solution and negligible solid solubility). By the governing equations for two of the components, e.g., A and C:

$$-RT \ln X_A(\ell) \simeq \Delta H_m(A) \left[1 - \frac{T}{T_m(A)}\right] \tag{26.8}$$

$$-RT \ln X_C(\ell) \simeq \Delta H_m(C) \left[1 - \frac{T}{T_m(C)}\right] \tag{26.9}$$

we found the phase boundary or "valley" descending from the "pass" or binary eutectic on the A-C binary into the interior of the ternary diagram and headed for the ternary eutectic. Figure 26.14 shows most of the directions of falling temperature (DFTs) for the current ternary phase diagram. For example, beginning at the A-C binary eutectic (e_1) the liquid-solid(A)-solid(C) phase boundary "falls away" from the A-C join, which happens to be the bounding binary. The same argument goes for all the "valleys" descending into the interior of the phase diagram from the bounding binaries. In the middle of the diagram we find a special case, but rule 5 still holds: the liquid-solid(C)-solid(AB) phase boundary falls to the left of the AB-C join toward one ternary eutectic (E_1), just as the liquid-solid(C)-solid(AB) phase boundary falls to the right of the AB-C join toward the other ternary eutectic (E_2).

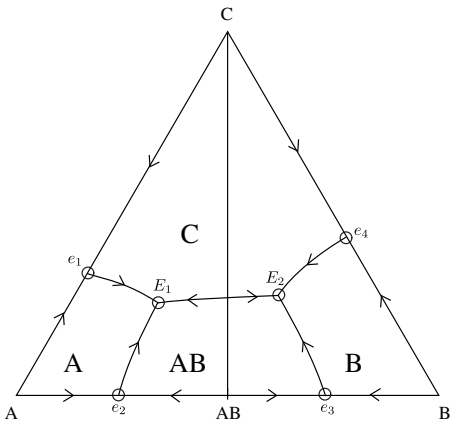


Figure 26.14: A-B-C liquidus projection diagram with most of the directions of falling temperature.

We are nearly finished with the phase diagram but we need to introduce the final Hummel's rule: No. 6. This may seem complicated at first, but it is nicely illustrated in Fig. 26.15. Rule no. 6, also referred to as "the Alkemade

Theorem" (pronounced "awl-keh-mah-deh") states that "a compatibility join solid(1)-solid(2) intersected solely by its own liquid-solid(1)-solid(2) boundary is a true binary eutectic in its own right; temperature falls to this point (a true binary eutectic) along this join, but falls away to either side in the ternary diagram." Hence, the crossing point labeled as e_5 in the diagram is a "saddle point," with DFTs descending to it from the end points (pure A, pure AB) but descending into the ternary on either side (also fulfilling rule no. 5).

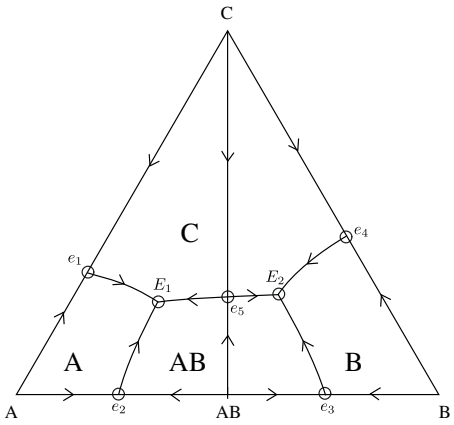


Figure 26.15: Completed A-B-C liquidus projection diagram and demonstration of the Alkemade Theorem.

You will note that directions of falling temperature on all the bounding binaries have also been drawn in. This is in accordance with Hummel rule no. 1, "liquidus surfaces always fall away (in temperature) from the corresponding primary phase." This completes the ternary liquidus projection diagram, but we can also sketch all the binary phase diagrams, at least schematically. As long as we are provided no isotherms or melting temperatures of the end members (A,B,C) or the intermediate compound (AB), only schematic diagrams can be sketched. So if you sketch a diagram with the melting point of C lower that that of AB, I cannot argue with you. The resulting schematics are shown in Fig. 26.16. Note that the AB-C eutectic diagram, determined by the Alkemade Theorem, is also represented in schematic fashion.

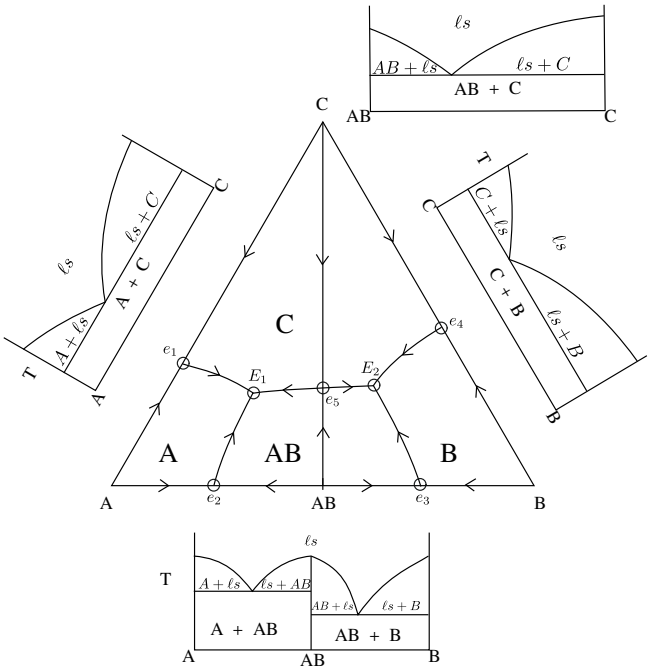


Figure 26.16: Completed A-B-C liquidus projection diagram with schematic bounding binaries and Alkemade eutectic AB-C.

Since we are dealing with schematic phase diagrams, just as with the choice of relative melting temperatures for the end-members (and compounds) the choice of relative eutectic temperatures is also arbitrary. For example, in

Fig. 26.16 the A-AB eutectic temperature is shown as being higher than for the AB-B eutectic. Were you to sketch it in the reverse fashion I could not argue with you, given the absence of more precise input information. However, there is one situation that is forbidden. This is shown in Fig. 26.17, where both A-AB and AB-B eutectics seem to have the same temperature. This is a violation of Gibbs phase rule. For any 2-D phase diagram, Gibbs phase rule is F=3-$\underline{P}$. In Fig. 26.17 there are actually five phases in equilibrium: three solid phases (A,AB,B) and two liquid solutions (the A-AB eutectic liquid and the AB-B eutectic liquid). Therefore, such a representation is to be avoided.

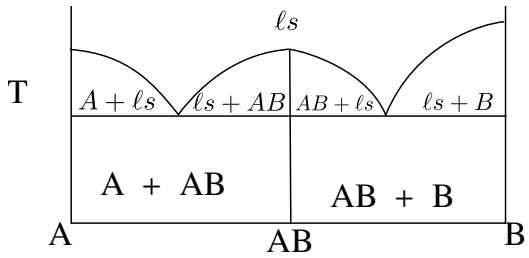


Figure 26.17: Dual-eutectic A-B system with a violation of Gibbs phase rule.

A slightly more complicated liquidus projection diagram is given in Fig. 26.18, which involves an incongruently melting binary compound, C_2B. You are more than welcome to take a crack at the diagram, using Hummel's rules to 1) label all primary phase fields, 2) determine the four subsolidus compatibility triangles (there are four ternary invariant points), 3) label all invariant points, both ternary and binary, 4) label directions of falling temperature on all ternary liquid-solid(1)-solid(2) boundaries and on the bounding binaries, and 5) sketch schematics of the bounding binary phase diagrams and any other true binary diagrams (satisfying the Alkemade Theorem) within the ternary. There are some "quirks" to this phase diagram, whose solution is given in Fig. 26.19.

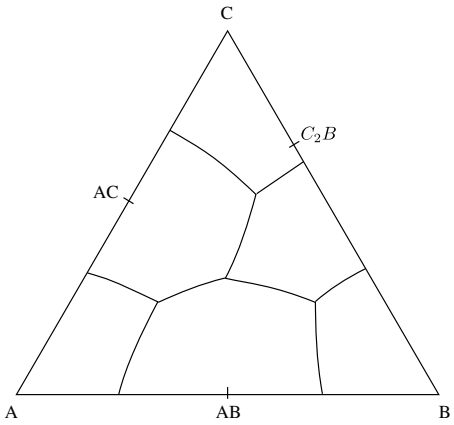


Figure 26.18: Raw A-B-C liquidus projection diagram involving incongruently melting C_2B compound.

The AC-AB binary is a clear solution to Hummel's rule no. 6 (the Alkemade Theorem). Only the liquid-solid(AC)-solid(AB) liquidus boundary crosses the AC-AB join. Hence, the intersection point is a true binary eutectic (e_6), which is a saddle-point in the liquidus; temperature falls to this point along the AC-AB binary (see the figure in the upper right), but falls to either side toward the ternary eutectics (E_1, E_2). But what about the seemingly corresponding point on the $AB - C_2B$ join? According to Hummel's rule no. 1, temperature falls away from the end members on this join, as shown. Furthermore, according to Hummel's rule no. 5, temperature on the liquid-solid(AB)-solid(C_2B) phase phase boundary falls away from the AB-C_2B join, as shown. However, the AB-C_2B join is NOT a true binary eutectic, owing to the fact that the primary phase field of C overlaps the AB-C_2B join at the upper right. There is no way on an AB-C_2B binary eutectic to combine AB and C_2B and arrive at pure C. So the AB-C_2B join is NOT a true binary eutectic phase diagram.

All this results from the peritectic behavior in the C-B binary. As shown in the C-B phase diagram, the primary phase field of solid C overhangs the C_2B composition, requiring that C_2B melt incongruently to a different solid (C) and a liquid of a different composition (p_1). This behavior persists into the ternary, with the primary phase field of C overhanging the liquid-solid(C)-solid(C_2B) phase boundary. The two circled arrows in Fig. 26.20 showing temperature descending from the C-B binary peritectic (p_1) to the ternary peritectic (P_1) follow from Hummel's rule no. 5: "the direction of falling temperature on a liquid-solid(1)-solid(2) phase boundary is always away from

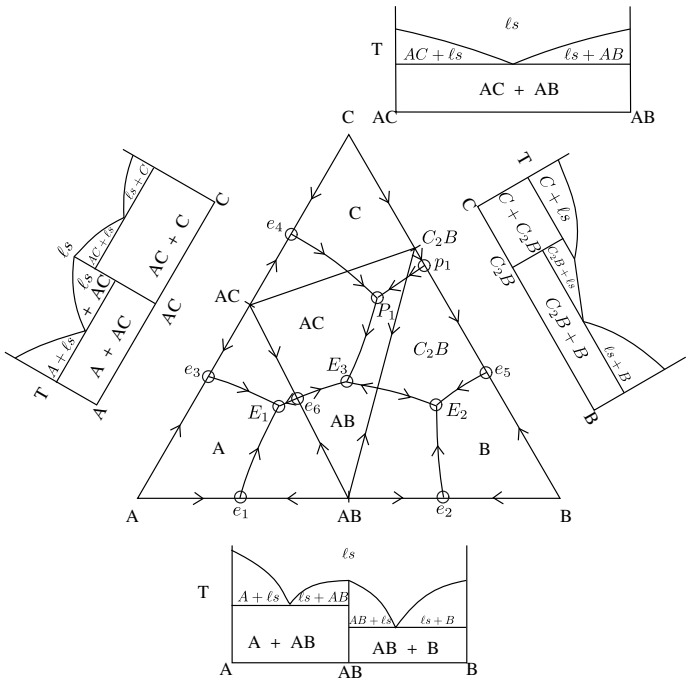


Figure 26.19: Completed A-B-C liquidus projection diagram of Fig. 26.18.

the corresponding solid(1)-solid(2) join." In this case we have to extend the C-C_2B subsolidus compatibility join, as shown by the dotted line in Fig. 26.20, away from which the liquid-solid(C)-solid(C_2B) boundary descends from the binary peritectic to the ternary peritectic. A final clarification has to do with the single arrow along the liquid-solid(AC)-solid(C_2B) liquidus phase boundary, between the ternary peritectic (P_1) and the ternary eutectic (E_3). This satisfies Hummel's fifth rule, with the direction of falling temperature being away from the corresponding AC-C_2B join, which is well above it on the diagram.

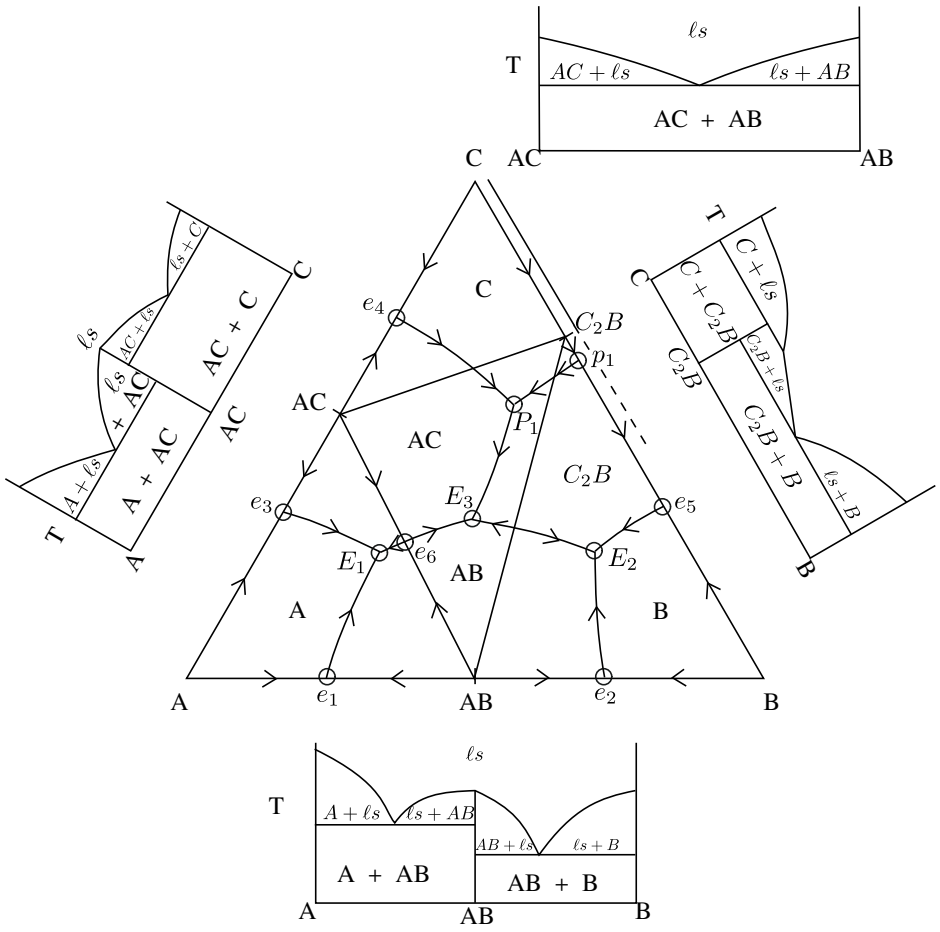


Figure 26.20: Specifics of the A-B-C liquidus projection diagram.

26.5 Isothermal Sections

Believe it or not, the ternary phase diagrams considered thus far are actually NOT valid Type III phase diagrams. This is owing to their 3-D character, or at least the projection of 3-D character onto liquidus projection diagrams. We now turn to so-called "isothermal sections," which are true Type III phase diagrams. In Fig. 26.21 we have reproduced the limited solid solubility liquidus projection diagram for the NaCl-NaF-NaI system.

We will now derive a series of isothermal sections from this phase diagram. It should be emphasized that 1) these are NOT schematic, but rather real phase diagrams which, 2) are true Type III phase diagrams. The first diagram at 991^{o}C is particularly straightforward in Fig. 26.22, being one degree C above the melting point of the most refractory end-member (refractory is from the French, meaning "high-melting;" NaF melts at 990°C). The entire isothermal section consists of a continuous liquid solution. We can return to the "mountain/valley/hole" analogies and the great Noahic flood; water is above the tops of all the mountains.

But now let's begin to drain the water from "Jackson Hole," bringing its level down to 850 meters, or rather the temperature to 850^{o}C as in Fig. 26.23. In doing such isothermal sections, it is helpful to overlay tracing or other "see-through" paper to trace the boundaries of the ternary and the specific "contours" (isotherms). At 850 meters the water is still above the tops of "mount NaCl" and "mount NaI." But the top of "mount NaF is exposed. Actually, for the diagram in Fig. 26.23 a "strip mining" analogy is probably more appropriate. Imagine removing the top of the mountain down to the 850 meter (or 850^{o}C) level (plus one centimeter as a "levee" to keep the water out). Water would be everywhere to the south on the diagram; to the north would be the flattened top of mount NaF. In the diagram you see dashed tie lines connecting various liquidus compositions around the perimeter back to the NaF vertex. This is a result of negligible solid solubility, namely that little NaCl or NaI tends to dissolve in solid NaF. All tie-lines radiate from the vertex. Imagine a friend standing and holding a rope at the composition of "pure" NaF, with you holding the opposite end and walking the "perimeter" or the shoreline of the water (the 850^{o}C isotherm) all the while keeping the rope taut. I refer to these constructions as "waffle cone" features. But we still don't see the tie-triangles characteristic of Type III diagrams. This happens at lower temperatures, however.

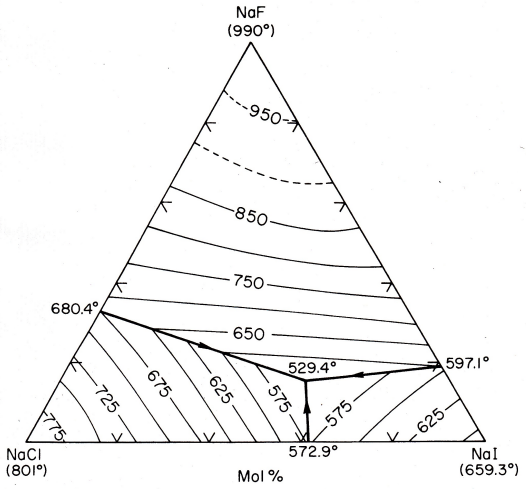


Figure 26.21: The liquidus projection diagram for the NaCl-NaF-NaI system showing isotherms (diagram 5908 from *Phase Diagrams for Ceramists*).

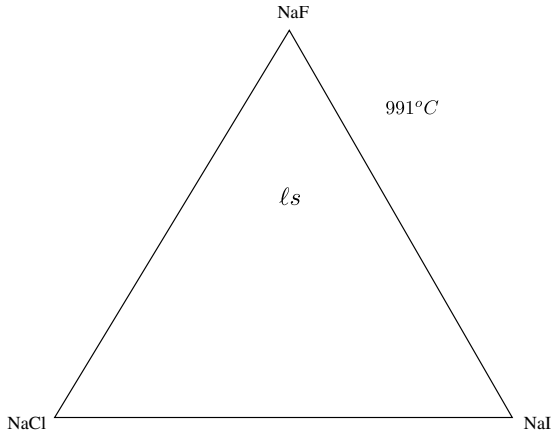


Figure 26.22: Isothermal section of the NaCl-NaF-NaI system at 991^{o}C.

For example, at 650^{o}C in Fig. 26.24 we are well below the melting points of NaF and NaCl, and even the very top of "mount" NaI (melts at 659.3^{o}C) is exposed. More importantly, the dropping water has exposed the NaCl-NaF-liquid "valley" descending from the NaCl-NaF eutectic at 680.4^{o}C. The remaining liquid solution portion is significantly smaller. There are now three "waffle cone" features, involving each of the primary phases in equilibrium with a range of liquid solutions. Again, tie-lines radiate from the "pure" end-members. Remember, we can always draw an intermediate tie-line between any two adjacent tie-lines; the "waffle cone" regions consist of an infinite array of tie-lines connecting primary phases and liquidus. More significantly, we see our first "tie-triangle" involving solid NaCl, solid NaF, and liquid solution. As pointed out previously, such tie-triangles are the hallmark of Type III phase diagrams.

Let's go further down in altitude (or rather temperature). There are no contour lines at 550 meters (550^{o}C), but we can certainly envision/approximate them as being roughly midway between 575 meters and the 529.4 meter altitude of "Jackson Hole" (529.4^{o}C). The resulting phase diagram in Fig. 26.25 has a greatly shrunken liquid solution region. There are now three "waffle cone," or rather "ice cream cone," constructions involving each of the primary phases and different small ranges of liquid solution. Dominating the diagram, around the perimeter are three tie-triangles involving two solid phases and three unique compositions of liquid solution. Clockwise, these are solid NaCl plus solid NaF plus liquid solution(1), solid NaF plus solid NaI plus liquid solution(2), and finally solid NaCl plus solid NaI plus liquid solution(3).

Let's pause and consider the nature of these isothermal sections. Later we return to the question of microstructure, however these series of diagrams give a good representation of isothermal sections which, once again, are true Type III phase diagrams. The phase rule for a C=3 system at fixed overall pressure and temperature, where C+2

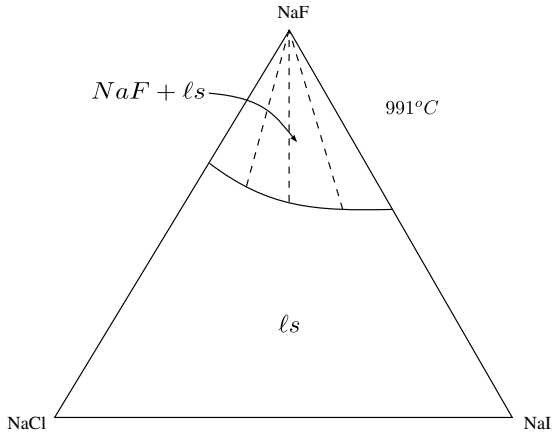


Figure 26.23: Isothermal section of the NaCl-NaF-NaI system at 850^{o}C.

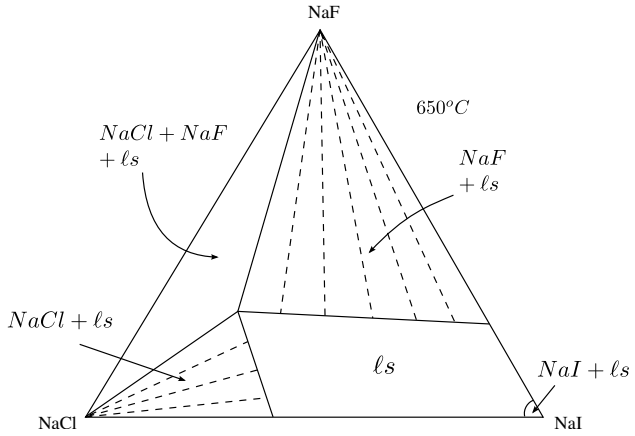


Figure 26.24: Isothermal section of the NaCl-NaF-NaI system at 650°C.

becomes C+0, is given as F=3-$\underline{P}$. Applying this to Fig. 26.25, we would find that F=2 in the small liquid solution region ($\underline{P}$=1); we would have to specify two of the three mole fractions to establish equilibrium. In the "ice cream cone" features $\underline{P}$=2, so we would have one degree of freedom. Specifying overall composition determines which tie line we are on, and the particular liquid solution in equilibrium with the associated solid phase. Or given the liquid solution composition, we would know on which tie-line the overall composition falls. Finally, in each of the tie-triangles $\underline{P}$=3 making F=0; all of the compositions are fixed, including that of the associated liquid solution.

The situation is a bit different at the ternary eutectic temperature. As we continue to go down in temperature it is apparent that the liquid solution region in Fig. 26.25 is shrinking and approaching the "drain" ("Jackson Hole" in our mountain/valley/hole analogy). What happens at precisely the point where the altitude of the "hole" is reached is displayed in Fig. 26.26. At first, this would seem to be a violation of the phase rule (F=3-$\underline{P}$), with four phases in equilibrium: solid(NaCl), solid(NaF), solid(NaI), and ternary eutectic liquid. However, we do not have the luxury of selecting the temperature, in the same way that we have no control over the altitude of Jackson Hole. The eutectic temperature is fixed for us, meaning that we can only look it up in a handbook (or on the published phase diagram of Fig. 26.21). We have no ability to change it. So (C+2) does not reduce to (C+0) as for the other isothermal sections, for which both temperature and pressure is fixed. Instead, temperature remains a free variable such that (C+2) becomes (C+1), the "1" standing for temperature, and the phase rule becomes F=4-$\underline{P}$. So under the unique conditions of a ternary invariant point, we can have four phases in equilibrium.

Let's complete our isothermal section "journey" by reducing the temperature to below the "hole" or the ternary invariant point. We are now in the subsolidus, and the entire phase triangle becomes one big tie-triangle involving the three "pure" end-members in equilibrium. The temperature of 529^{o}C is 4/10ths of a degree below the ternary eutectic temperature. There is no remaining liquid solution; everything has solidified. So the resulting isothermal section in Fig. 26.27 is nearly as straightforward as the one with which we started.

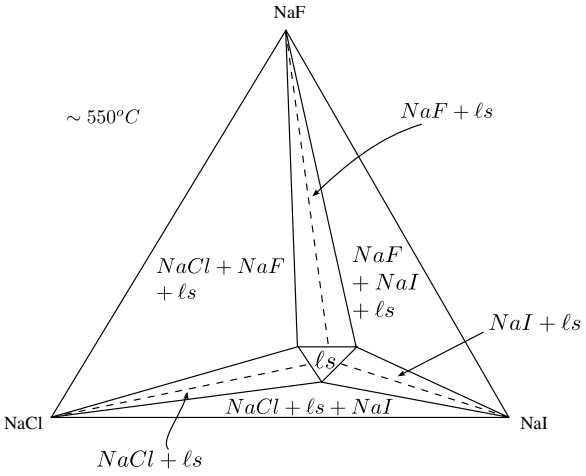


Figure 26.25: Isothermal section of the NaCl-MaF-NaI system at ~550C.

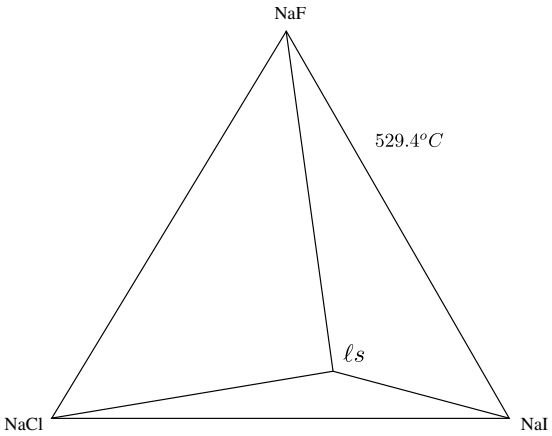


Figure 26.26: Isothermal section of the NaCl-NaF-NaI system at the ternary eutectic temperature of 529.4^{o}C.

26.6 Crystallization Paths and Microstructure Evolution

Just as with binary phase diagrams, crystallization sequences/paths that take place during cooling in Type III (ternary) liquidus-projection diagrams play a major role in establishing the the ultimate microstructures obtained. This is a prime example of the processing$\Longleftrightarrow$ microstructure chain link in the materials science and engineering paradigm in Fig 22.1. It must be stressed that we will be considering so-called "equilibrium" cooling sequences/liquidus paths. In reality, microstructural evolution also depends upon cooling rates, for example, the occurrence of dendrites (dendritic growth) treated in later materials science and engineering coursework. Let's begin by considering some crystallization sequences in the simple binary eutectic of Fig. 26.28. This will be a quick review of what you should have already learned in your "Intro to Materials Science and Engineering" coursework. The sequence is particularly simple for the eutectic composition (X_{eut}). Above the liquidus/eutectic we have 100% liquid solution of eutectic composition which, by binary lever rule is 46% B and 54% A. Upon cooling through the eutectic reaction, all the eutectic liquid is converted to the lamellar eutectic microstructural constituent indicated by the banded regions in the bottom diagram on the left of Fig. 26.28. This layered microstructural constituent results from the growth of layers being the most efficient way to separate A and B (by diffusion at the growth front) from a solution into separate phases, as shown schematically in Fig. 26.29.
In contrast, the crystallization sequence and liquidus "path" for the composition X' is quite different. Above the liquidus, we have 100% liquid of the composition X'. As we cool below the liquidus the "primary phase" to crystallize from solution is B. As B is removed from the liquid solution, the liquid becomes progressively richer in A, following the liquidus line as shown, until it reaches the eutectic point. Since the composition (X') is approximately half way between the end-member B and the eutectic composition, at just above the eutectic temperature we would have a microstructure with 50% primary B grains and 50% liquid of eutectic composition. Now, if we somehow sieved out the primary crystals, what would remain would be a liquid of eutectic composition, which would go

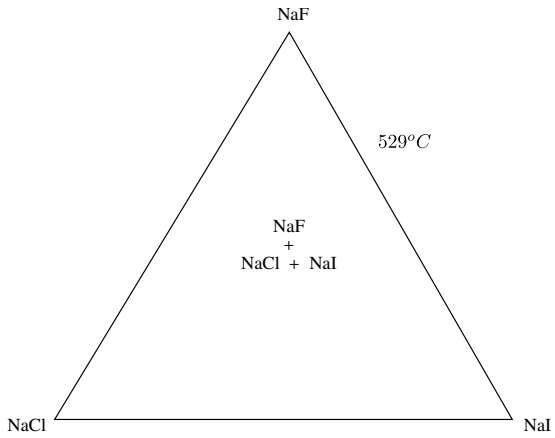


Figure 26.27: Isothermal section of the NaCl-NaF-NaI system at 529°C.

through the same crystallization to lamellar eutectic microstructural constituent as happened for X_{eut}. The ultimate microstructure would consist of 50% primary B grains suspended in 50% lamellar eutectic microstructural constituent, as shown on the bottom-right of Fig. 26.28.

Before considering crystallization sequences/liquidus paths in ternary liquidus-projection diagrams, we need to learn a framework for their interpretation. I refer to these as the "Crystallization Path Dicta" or "CPD," as spelled out in Figure 26.31. Note the reversal of "CPD" to "DPC" in the diagram. The C stands for "compatibility" and gives us the "D" or "destination," namely the ending point of the liquidus path: the ternary invariant point where the final liquid disappears in the crystallization sequence. Since there is a one-to-one correspondence of subsolidus compatibilities and ternary invariant points, the "destination" will always be the ternary eutectic or peritectic corresponding to the subsolidus compatibility in which the overall composition lies. In the simple ternary eutectic phase diagram of Fig. 26.30, there is only one subsolidus compatibility and therefore only one invariant point. "All paths lead to Rome," they say. In this system, for all compositions within the A-B-C compatibility triangle, all paths end up the the ternary eutectic. The path taken to get there will differ according to the overall composition.

The second letter in CPD, "P" stands for "primary phase field" and tells us the initial "P" for "path." If we return to the "mountain/valley/hole" analogy, our helicopter has just dropped a skier or snowboarder on the liquidus surface at the point marked "X." Since essentially pure C is crystallizing from the liquid solution, our skier's/snowboarder's "path" (or rather the "path" of the liquid) is directly down the mountainside away from the peak, namely directly away from "mount C." The composition "X" was chosen so that its crystallization path would intersect the ternary eutectic. If we pause just at/above the ternary eutectic, what would be the microstructure? We would employ the binary lever rule, since there are only two phases. On the line from C through the original composition "X" to the eutectic point, the length marked "f_{PriC}" divided by the entire length of the line would give us the fraction of primary C grains in the microstructure (~50%), with the remainder being eutectic liquid. Once again, if we somehow sieved out the primary crystals, what would remain would be a liquid of eutectic composition, which would go through crystallization to ~50% lamellar eutectic microstructural constituent, this time consisting of three layers: A plus B plus C. If we wanted to know the fractions of phases in the lamellar eutectic, we could do the triangular lever rule at the eutectic composition, using the entire Gibbs triangle for the calculation.

Now let's consider the composition marked X' on the ternary liquidus-projection diagram of Figure 26.30. Again, "C" for "compatibility" tells us that we have to end up somehow at the ternary eutectic ("D"=destination). "P" for "primary phase field" gives us the initial "P" or path, namely directly away from "mount C" as shown in the diagram. Again, our skier/snowboarder is headed full-tilt down the mountain directly away from its peak. This time, however, his/her path intersects the solid(C)-solid(A)-liquid phase boundary or "valley." Just at this point, labeled "$T \approx T_{A-C-liqboundary}$" in the second microstructure on the right side of Figure 26.30, we can use the binary lever rule to establish the microstructure. The point X' is roughly half way between end-member C and the solid(C)-solid(A)-liquid boundary. This means we have ~50% primary C grains and ~50% liquid of the composition at the boundary. Here we employ the final dictum of our "CPD" dicta. The letter "D" stands for "directions of falling temperature" and informs us of "changes in path" or the "C" in DPC. Quite naturally, the skier/snowboarder carves to the left to follow the valley down to the hole. What this actually means is that both C and A are crystallizing simultaneously from the liquid, the latter as a secondary crystallization product; its grains will end up smaller in size than primary phase C grains, owing to the fact that they have had less time to nucleate

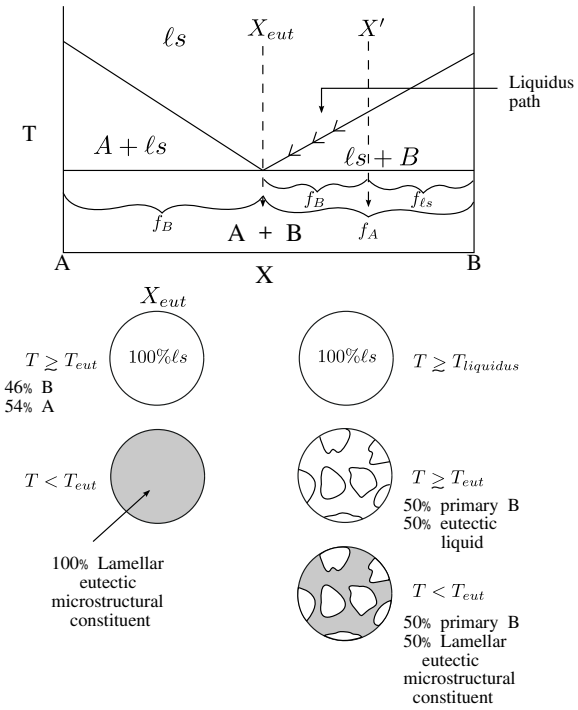


Figure 26.28: Crystallization sequences/"paths" in a simply binary eutectic phase diagram.

and grow. Now let's analyze the situation at a temperature just above the ternary eutectic. Since we now have three phases in equilibrium, we need to employ the ternary lever rule of Fig. 26.3. The analysis is shown in Fig. 26.32. A dashed triangle is drawn from end-member C to end-member A and to the eutectic composition, on which we carry out the ternary lever rule to establish phase fractions. Lines from each vertex of this triangle are drawn through the overall composition X'. The relative line lengths (divided by the total line lengths) give us the fractions of primary C (~65%), secondary A(~16.5%), and eutectic liquid (~18.5%). As before, if we somehow sieved out both primary and secondary crystals, what would remain would be a liquid of eutectic composition, which would go through crystallization to lamellar eutectic microstructural constituent. The resulting microstructure (~65% primary C, ~16.5% secondary A, ~18.5% lamellar eutectic) is shown at the bottom right of Fig. 26.30.

26.7 Isothermal Sections of "Real" A-B-C Systems

Of course, very few actual Type III/ternary systems satisfy the assumptions we have made thus far (for simplification), namely that the liquid solution is everywhere ideal and that there are essentially no solid solutions. It should be *strongly* emphasized that Hummel's rules strictly apply only to such narrowly-constrained systems. "All bets are off!" when dealing with "real" ternary systems, in which non-ideality of the liquid solution and significant, even substantial, solid solutions can be present. However, as we will see, all of the same characteristic features of Type III diagrams will be present, namely 1) the presence of tie-triangles, 2) the occurrence of tie-lines that are neither parallel nor perpendicular to any axis, and 3) open areas that can either be tie-triangles or single-phase solutions. There are, however, important differences. In "real" Type III isothermal sections, open areas that are not tie-triangles can also be solid solution regions in addition to liquid solution regions, as we will show. More importantly, "waffle cone" and "ice cream cone" regions, in which solid-liquid tie-lines radiate from a point (pure solid), will morph into quite a variety of quadrilateral constructions, with tie-lines connecting compositions on one side (solid solution) to compatible compositions on the other side (liquid solutions).

As you can imagine, experimental determination of even one isothermal section can be highly time-consuming. This has led to the development of and usage of powerful phase diagram calculation programs. Two of these are FactSage (C.R.C.T., Ecole Polytechnique de Montreal) and Thermo-Calc (Thermo-Calc Software, McMurray, PA; Thermo-Calc Software AB, Stockholm, Sweden), but there are many more. We will employ Thermo-Calc to calculate both the liquidus and one isotherm for the Sn-Pb-Bi phase diagram. But the user should be aware that there is a lot that does not "meet the eye" when using such high-powered "black box" programs. Behind the scenes, each such program employs a massive database of thermodynamic parameters, including solution

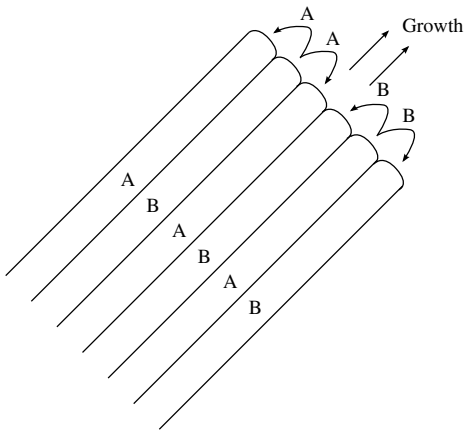


Figure 26.29: Schematic growth mechanism of binary eutectic microstructural constituent.

thermodynamic parameters which, in certain cases (not all!) can be far more sophisticated than the dilute solution and regular solution models employed in this text. Truly inquisitive scientists and engineers are encouraged to check what is "under the hood" rather than just "kick the tires."

The liquidus-projection phase diagram calculated by Thermo-Calc for the Sn-Pb-Bi system is shown in Fig. 26.33. On the "surface" (no pun intended!) this would look like just any other liquidus-projection diagram for an A-B-C Type III system involving ideal liquid solution and negligible solid solution. But there are *big* problems, which may have already occurred to you. First of all, there are four primary phase fields, but only three end-members. Second, there are two ternary invariant points, but seemingly only one compatibility triangle (involving the three end-members) in the subsolidus!

Things get even more interesting if we do an isothermal section. In Fig. 26.34 we see the isothermal section predicted by Thermo-Calc at 423.15K (150^oC). Note that the low temperature has to do with the relatively low melting points (for metals) of Sn (231.9^oC), Pb (327.5^oC), and Bi (271.4^oC). Again, at first glance, things seem to be familiar. There is a "waffle cone" structure at the Bi-vertex, where liquid solutions of a wide range of compositions are compatible with/have tie-lines connecting with essentially "pure" Bi. (But the "ice cream" seems to have melted, i.e., the liquidus boundary is convex.) Also, there is one tie-triangle involving Sn-solid solution, Pb-solid solution, and liquid solution (ls). But elsewhere this isothermal section is quite different from what we have seen thus far. There are significant solubilities in both solid Sn (Bi is much more soluble than Pb) and in solid Pb (both Sn and Bi are quite soluble, with Bi being roughly twice as soluble as Sn). We also see a small solid solubility range for an intermediate compound in the Pb-Bi system. If we look at the Pb-Bi phase diagram (not shown) we find out that the intermediate compound is called "ϵ-Pb." It is apparent in the isothermal section of Fig. 26.34 that ϵ-Pb dissolves a considerable amount of Sn, but to determine the amount of excess Bi or Pb it dissolves, we would have to consult the binary diagram to ascertain the 150^oC values in comparison with its nominal Bi/Pb stoichiometry. Anyway, this explains the four primary phase fields and the two ternary invariant points in the liquidus diagram of Fig. 26.33. However, the biggest differences in the isothermal section of Fig. 26.34 are the regions of tie-lines connecting various phases in equilibrium. These turn out to be strangely-shaped quadrilateral constructions bearing no resemblance to the "waffle cone" and "ice cream cone" constructions with which we are familiar.

The most important distinction between "real" Type III systems (non-ideal liquid solution, non-negligible solid solutions) and the constrained Type III systems (ideal liquid solution, negligible solid solutions) we have studied thus far has to do with crystallization sequences/liquidus paths and microstructure evolution. For example, in a two-phase equilibrium region, the compositions of the solutions at the end of tie-lines through a fixed overall composition change with temperature. Think of the overall composition as "anchoring" all the tie-lines, but the end-points (solution compositions) change progressively and in opposite directions as temperature is lowered. A schematic representation of this phenomenon is given in Fig. 26.35. From above, the resulting two paths, e.g., of liquid solution and of solid solution, would trace out a "butterfly-shaped" or "hour glass-shaped' construction, as in the diagram on the right. The important "take away" point is that crystallization becomes much more complicated. For example, liquidus paths are seldom straight lines in "real" Type III/ternary systems, but are rather curved trajectories.

Micro structurally speaking, since the solid solution ranges change with temperature, the likelihood of a process called "coring" increases. A very simplified perspective on coring can be taken from the simple binary eutectic

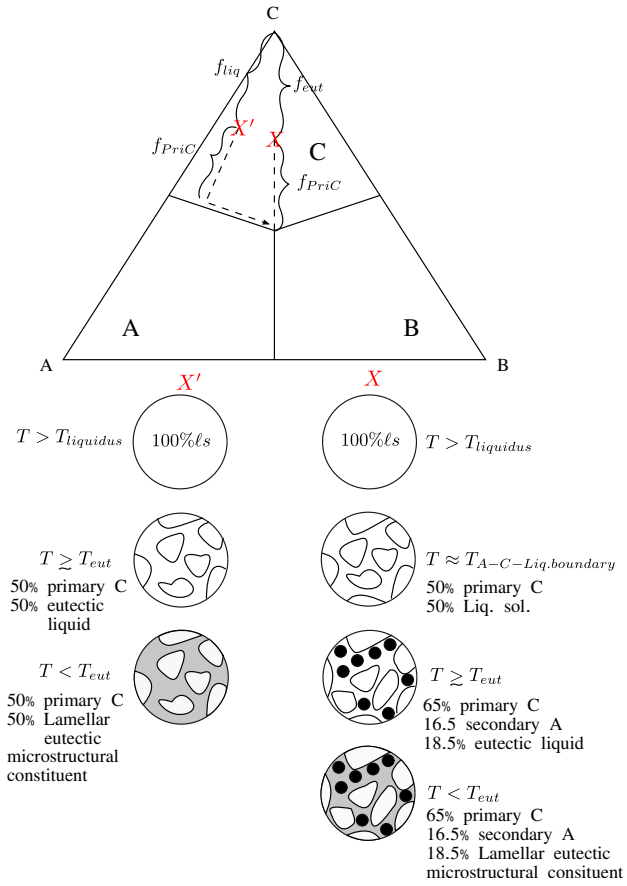


Figure 26.30: Crystallization sequences/liquidus paths in the A-B-C simple ternary eutectic diagram.

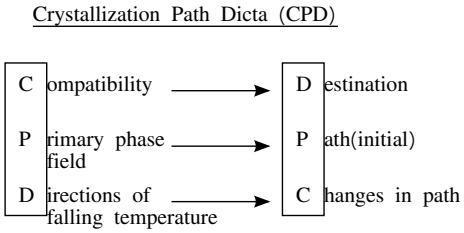


Figure 26.31: Crystallization Path Dicta (CPD) for ternary liquidus projection diagrams with negligible solid solubility.

in Fig. 26.36. This A-B-system shows significant solid solution at both ends of the diagram. The diagram on the right shows a blow-up of the liquidus and α-solidus. When a liquid of composition "X" is cooled to the liquidus, the first solid α to precipitate from solution has a fairly high A-content, as shown. However, at a much lower temperature, the equilibrium solid α phase will have a significantly smaller A-content, if equilibrium were maintained. The problem is that alloys are seldom cooled at rates slow enough to maintain equilibrium. That is because, once precipitated out, a solid phase "locks up" its constituents, since diffusion in solids is so much slower than in liquids. So what happens is that the new α' precipitating out will do so at a more B-rich composition. In the schematic microstructure beneath the two phase diagrams this is represented as a distinct layer of α' around a core of α. In reality, there will be a gradual shift in composition from the center outward and, in fact, this will occur by dendritic growth. You will learn about dendrites and dendritic growth in later materials science and engineering coursework. For now, we need only point out that the "average" solid composition (averaging the composition of α, α', α'', etc.) deviates from the solidus line. Since less A is being removed from the liquid, the liquidus trajectory must also therefore deviate from the equilibrium liquidus, as shown. All this is beyond the scope of the present text, but gives you a good idea of what you can look forward to in upcoming microstructure-evolution coursework. The conclusion of this discussion is that crystallization sequences/liquidus paths in "real" Type III/ternary systems can be quite complicated, along with microstructure evolution, which will also be cooling

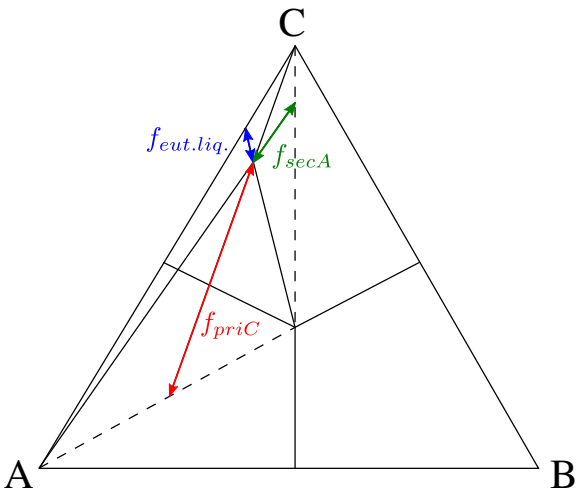


Figure 26.32: Application of the ternary phase rule to a point in the crystallization sequence/liquidus path of composition X' in Fig. 26.30.

rate-dependent. This is a good place to close the present text, by pointing out once again the synergistic influences of both thermodynamics and kinetics in the processing ⇒ microstructure "chain link" of the materials science and engineering processing ⇒ microstructure ⇒ properties ⇒ performance paradigm. Students interested in pursuing further the topic of the interpretation of "real" Type III/ternary diagrams are directed to old treatises like Georg Masing's *Ternary Systems: Introduction to the Theory of Three Component Systems* [24].

26.8 Some Technologically Important Type III Phase Diagrams

We mentioned "freezing point lowering" when discussing binary (Type II) phase diagrams. There we mentioned the klinkering of cement as one example of major technological importance. A portion of the $CaO - SiO_2 - Al_2O_3$ liquidus projection diagram is reproduced in Fig. 26.37. On the $CaO - SiO_2$ binary diagram to the left, the incongruent melting point of C_3S is given as $\sim 2050^oC$ and the congruent melting point of C_2S is given as $\sim 2130^oC$, which are way too high for economical manufacturing. The dashed circle shows the composition ranges over which so-called "Portland cement" is made by klinkering. I will leave it to you to apply your ternary phase diagram prowess and decipher the lower lying ternary invariant points involving the C_3S and C_2S phases.
Another technologically important application is that of "solders." A solder is a fusible alloy of two or more metals used to join metal parts together and having a melting point lower than those of the constituents or of the parts being joined. Solders are commonly used in plumbing, sheet metal joining, and electronics. The application in electronics is of particular importance, since solder is employed for making permanent mechanical and electrical connection of parts to the printed circuit boards on which they are mounted. Eutectic compositions are of special interest owing to the low melting points and absence of any primary phase that could disrupt electrical contact, for example if the joint solidifies in the so-called "pasty" (eutectic liquid plus primary phase) state. For decades, the alloy of choice for electrical soldering was 60/40 Sn/Pb. This is because the eutectic composition in the Sn-Pb system is 63% Sn and 37% Pb, which melts at the lowest possible (eutectic) temperature of 188^oC. On the other hand, the Sn/Pb ratio is closer to 50/50 in plumbing solder, so chosen because this alloy solidifies more slowly and manageably.
In recent years, however, environmental safety concerns have motivated a steady move away from Pb-based solders owing to the recognized toxicity of Pb in both manufacturing and recycling. In the European Union, directives issued in 2006 prohibited the inclusion of significant Pb contents in consumer electronics. In the U.S., there are also tax incentives for the reduction of Pb content in consumer electronics. There are many binary and ternary (largely eutectic) systems that have been investigated and employed as Pb-free solders. One of the most commercially successful solders to date comes from the Sn-Ag-Cu ternary system shown in Fig. 26.38. Unfortunately, the eutectic in question is where the primary phase fields of Sn, Ag_3Sn and Cu_3Sn come together, which is very near the Sn-vertex and therefore extremely difficult to see on the full ternary. Figure 26.39 is a blow-up of the Sn-corner showing the low-melting eutectic composition of approximately 4 mass% silver and 1 mass% copper (~95 mass% Sn) at 218^oC. It seems appropriate to conclude our discussion of ternary phase diagrams with a true Type III or Q_i/Q_k vs. Q_j/Q_k phase diagram, in this case the mass fraction of Ag/Sn plotted vs. the mass fraction of Cu/Sn

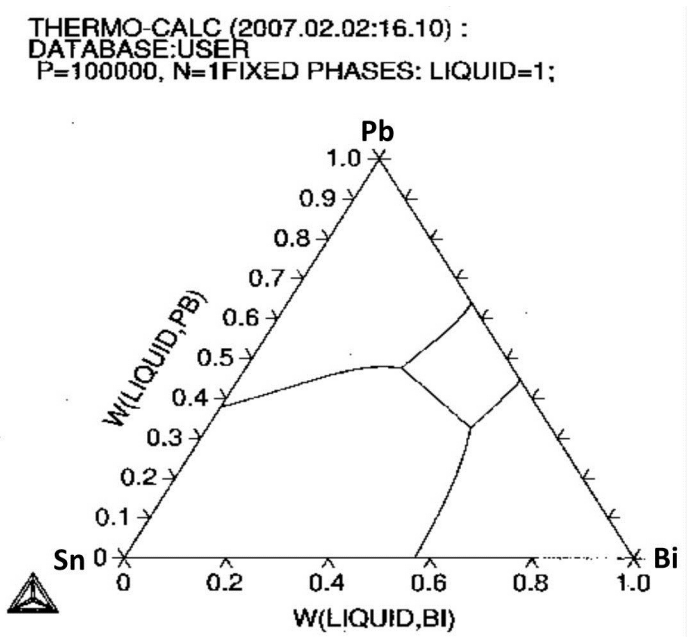


Figure 26.33: Calculated liquidus projection diagram for the Sn-Pb-Bi Type III phase diagram (by Thermo-Calc).

plotted in proper Type III rectilinear fashion.

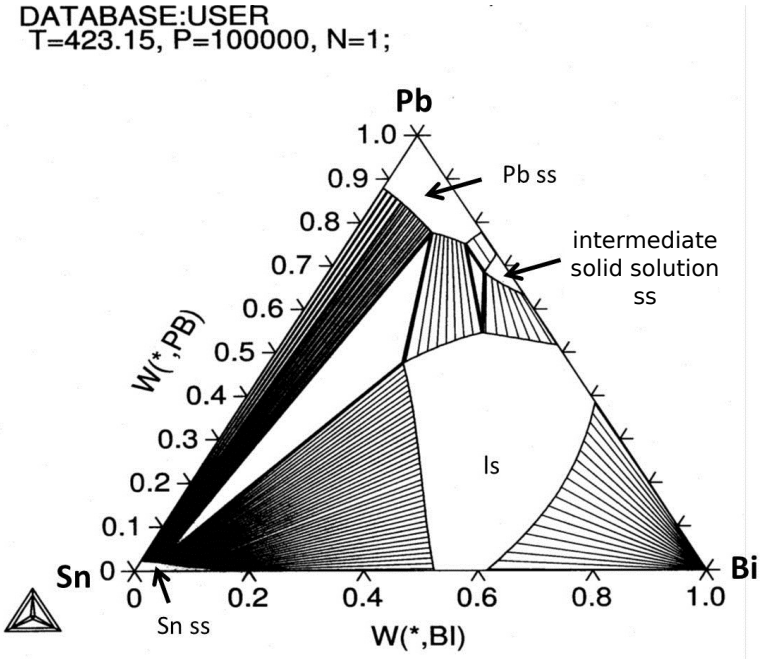


Figure 26.34: Calculated 150^{o}C isothermal section for the Sn-Pb-Bi Type III system (by Thermo-Calc.

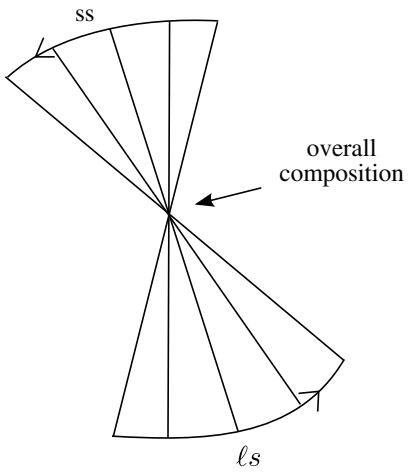


Figure 26.35: Schematic representations of crystallization "paths" in a two-phase region of a "real" Type III ternary with solid and liquid solubilities that change with temperature.

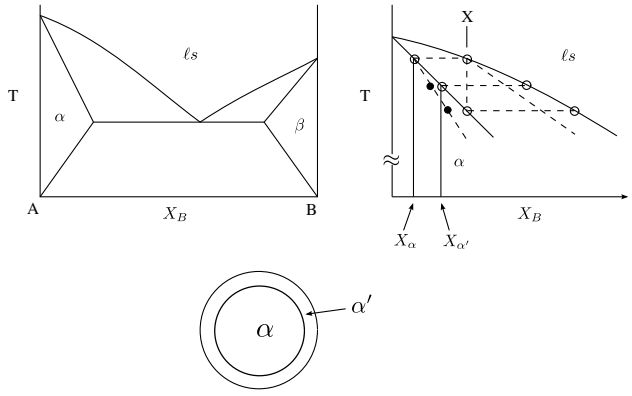


Figure 26.36: Schematic representation of the phenomenon of "coring" in a binary eutectic system.

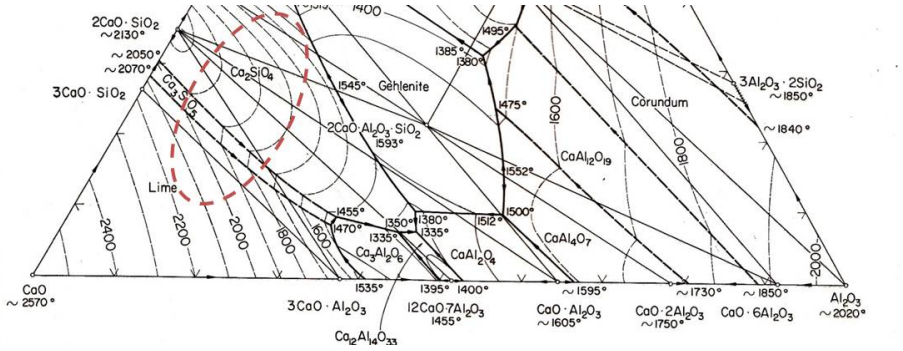


Figure 26.37: Lower portion of the $CaO - SiO_2 - Al_2O_3$ liquidus projection phase diagram (from Phase Diagrams for Ceramists, The American Ceramic Society).

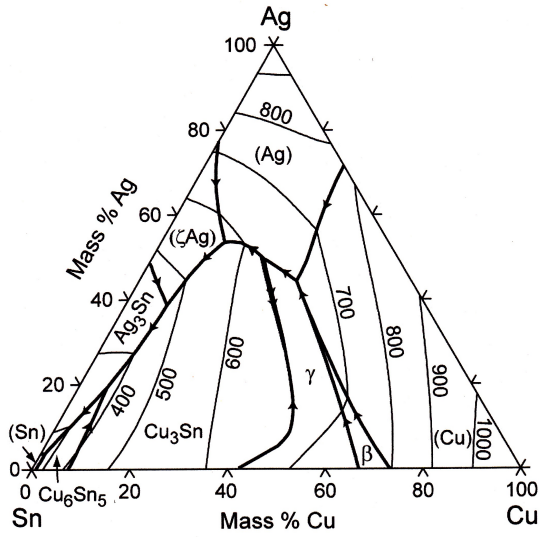


Figure 26.38: The calculated Sn-Ag-Cu liquidus projection phase diagram (from the NIST database).

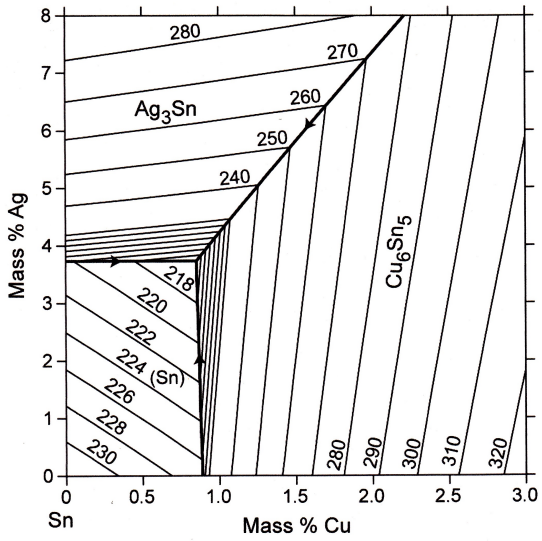


Figure 26.39: The Sn corner of the calculated Sn-Ag-Cu liquidus projection diagram plotted in rectilinear fashion (from the NIST database).

27 Fick's Laws of Diffusion

Just as there are fundamental laws governing the thermodynamics of materials, so there are basic laws governing diffusion or how atoms/ions move around in materials. Consider the two adjacent planes in an isotropic solid as shown in Figure 27.1a:

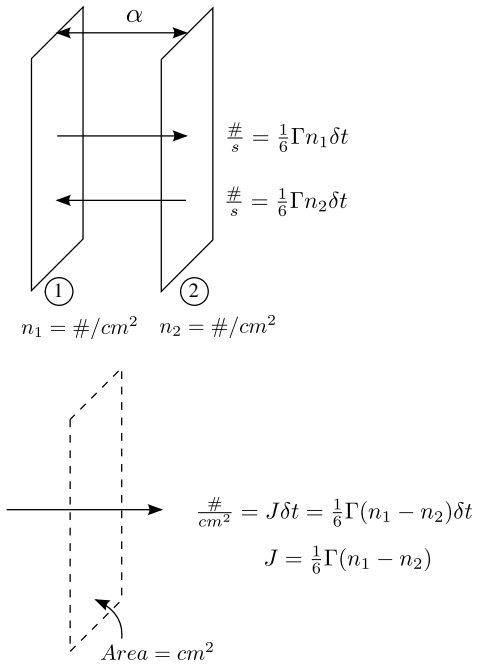


Figure 27.1: Schematic of flux between adjacent planes in an isotropic solid.

By isotropic, we mean that the diffusion rate of an impurity atom is independent of direction. The planes are separated by the distance, α, which we will refer to as the jump distance. Each plane has a different areal density of impurity atoms (# per cm^2), n_1 on plane 1 and n_2 on plane 2. Things are simplified by assuming an isotropic solid, since we can consider jumps from plane 1 to plane 2 (along the $+x$ axis) to be only one of six possible jump directions ($\pm x, \pm y, \pm z$). Therefore, the number of atoms jumping from plane 1 to plane 2 during a time period δt will be equation 27.1:

$$\#_{1\to 2} = \frac{1}{6}\Gamma n_1 \delta t \tag{27.1}$$

where Γ is the jump frequency of an atom (#/s). Similarly, the counter jumps from plane 2 to plane 1 during the same time period is equation 27.2:

$$\#_{2\to 1} = \frac{1}{6}\Gamma n_2 \delta t \tag{27.2}$$

The net flow to the right through an imaginary plane midway between the two planes, as in Figure 27.1b, is given by subtracting equation 27.2 from equation 27.1 or equation 27.3:

$$net\#_{1\to 2} = \frac{1}{6}\Gamma(n_1 - n_2)\delta t \tag{27.3}$$

If we define the net flux J as the number of atoms passing through a unit cm^2 area per unit time (s), we arrive at equation 27.4:

$$net\#_{1\to 2} = \frac{1}{6}\Gamma(n_1 - n_2)\delta t = J\delta t \tag{27.4}$$

such that (see Figure 27.1b) the flux can be written as equation 27.5:

$$J = \frac{1}{6}\Gamma(n_1 - n_2) \tag{27.5}$$

Two immediate observations can be made from Eq. 27.5. The units of flux are $(\#/cm^2)/s$. Furthermore, if the area concentration of the impurity is the same on the two planes, the net flux will be zero. However, this does not indicate a static situation; it only means that the flux to the right ($J_1 = \frac{1}{6}\Gamma n_1$) is the same as the flux to the left ($J_2 = \frac{1}{6}\Gamma n_2$). In fact, each of these fluxes can be enormous. Net zero flux only requires that their magnitudes be the same, namely $J = J_1 - J_2 = 0$.

27.1 Fick's First Law

Until now we have only considered area concentrations of impurities ($\#/cm^2$), however Fick's Laws are given in terms of volume concentrations ($\#/cm^3$). We can convert to volume concentrations in Figure 27.1 by shifting the frame of reference by one-half a jump distance to the right, as shown in Figure 27.6:

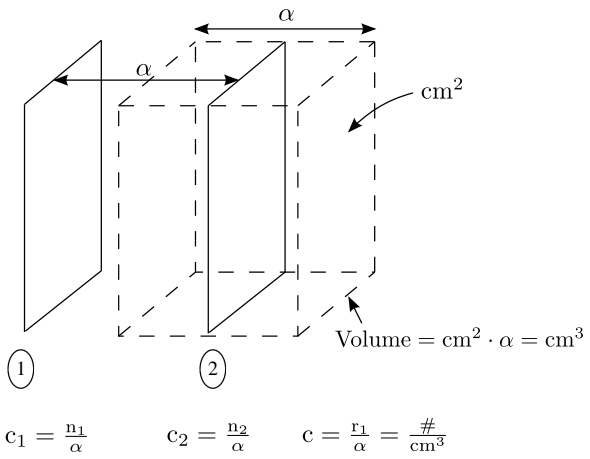


Figure 27.2: Converting area concentration to volume concentration.

If we consider the jump distance in centimeters, it follows that the area inside the dashed box will be $cm^2 \cdot cm = cm^3$. The volume concentration in this block (c_2) is the area concentration of impurities, ($n_2 = \#/cm^2$), divided by the jump distance, $\alpha = cm$, or $\#/cm^3$. The net flux from plane 1 to plane 2 can therefore be written as equation 27.6:

$$J = \frac{1}{6}\Gamma(n_1 - n_2) = \frac{1}{6}\Gamma\alpha(c_1 - c_2) \tag{27.6}$$

Of course, concentration is seldom treated at a perfectly atomistic, plane-by-plane level. Rather it is treated as continuous, as displayed schematically by the dashed line in Figure 27.3.

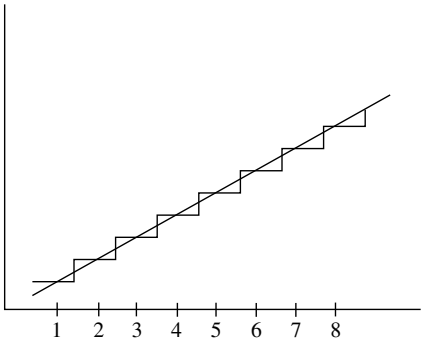


Figure 27.3: Converting the concentration profile to a continuous function.

If we do a Taylor series expansion about c_1 and ignore higher order terms, we obtain equation 27.7:

$$c_2 = c_1 + \alpha\frac{\partial c}{\partial x} \tag{27.7}$$

However, since the "gradient" between planes 1 and 2 is given by equation 27.8:

$$\frac{\partial c}{\partial x} = \left(\frac{c_2 - c_1}{\alpha}\right) = -\left(\frac{c_1 - c_2}{\alpha}\right) \tag{27.8}$$

the concentration difference between the two planes can be expressed as equation 27.9:

$$(c_1 - c_2) = -\alpha \frac{\partial c}{\partial x} \tag{27.9}$$

Substituting for $(c_1 - c_2)$ in equation 27.6, we find that equation 27.10 holds:

$$J = \frac{1}{6}\Gamma\alpha(c_1 - c_2) = -\frac{1}{6}\Gamma\alpha^2 \frac{\partial c}{\partial x} \tag{27.10}$$

This important equation tells us that the flux is proportional to the negative of the concentration gradient at time t, or $\nabla_c = (\partial c/\partial x)_t$, and the proportionality coefficient is given by equation 27.11:

$$D = \frac{1}{6}\Gamma\alpha^2 \tag{27.11}$$

where D is the isotropic diffusion coefficient. Since the jump frequency (Γ) has units of $\#/s$ and the jump distance (α) has units of cm, it follows that the diffusion coefficient will have units of cm^2/s. Many times we will see diffusion coefficients in other units, such as m^2/s. Regardless, we can rewrite equation 27.10 as equation 27.12, which is the most common form of Fick's First Law of diffusion:

$$J = -D\frac{\partial c}{\partial x} \tag{27.12}$$

It should be stressed that the diffusion coefficient of equation 27.11 holds strictly true only for isotropic solids. These include amorphous solids (glasses) and cubic crystal structures, such as face-centered cubic (fcc), body-centered cubic (bcc), sodium chloride (NaCl), fluorite (CaF_2), etc. More complex equations describe diffusion in non-cubic systems, e.g., hexagonal close-packed (hcp), which are beyond the scope of the present treatment.

27.2 Random Walk Diffusion

Above we mentioned that lots of jumps are taking place in both directions between planes 1 and 2 in Figure 27.1, even under a condition of zero net-flux. Consider the diffusion coefficient of interstitial carbon in fcc $\gamma - Fe$, which has been found experimentally to be $2.5x10^{-11}m^2/s$ at 1000^oC. Figure 27.4 shows the nearest-neighbor jump distance of an interstitial on a (100) plane in the fcc structure.

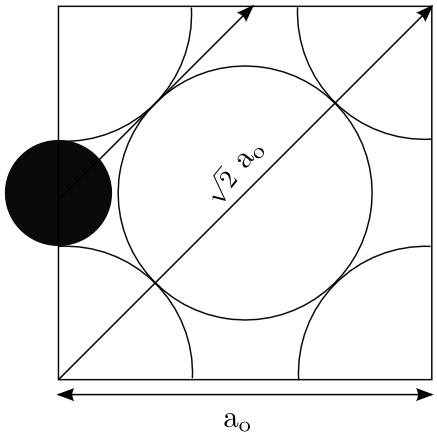


Figure 27.4: Schematic showing one interstitial jump distance in the (100) plane of fcc $\gamma - Fe$.

The lattice parameter of $\gamma - Fe$ (a_o) is known to be 0.37 nm, such that the face-diagonal is $\sqrt{2}a_o$. The jump distance shown is half of that distance, or $\sqrt{2}a_o/2$, which would be 0.26 nm. Rearranging equation 27.11 to yield the jump frequency, we obtain equation 27.13:

$$\Gamma = \frac{6D}{\alpha^2} = \frac{6(2.5x10^{-11}m^2)}{(0.26x10^{-9}m)^2} = 2.2x10^9/s \tag{27.13}$$

Wow! An average interstitial carbon atom makes approximately 2.2 billion jumps in just one second at this temperature! Yet the lattice vibrational frequency (recall the Debye frequency from your thermodynamics background, ν_D) is on the order of $10^{13}/s$. We can look at this frequency as the "attempt" frequency, or how many times per second the interstitial is attempting to jump from the original position to the new one. So only approximately 2 out of 10,000 attempts is successful at this temperature.
It is also of interest to compare how far an average interstitial atom moves from its initial position in one second vs. the total distance of back-and-forth motions it makes in that same second. Consider the sketch in Figure 27.5, where the sphere represents the initial carbon atom position, and arrows represent individual jumps in each of the $\pm x$, $\pm y$, and $\pm z$ directions.

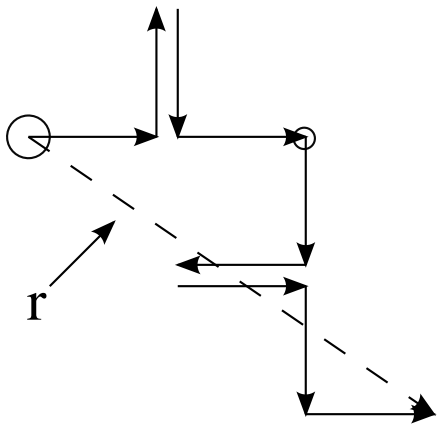


Figure 27.5: Schematic of random walk distance after a series of individual jumps in an isotropic solid.

The circle with the dot represents a jump out of the plane of the diagram and then back. It can be seen that many of the jumps are ineffective for long-range diffusion, since they are simply reversed. Nevertheless, the interstitial atom makes steady progress away from its initial position, and the dashed line connects the initial position with the final position after all the jumps are considered. This is referred to as the "random walk distance" (r). Computer simulations can keep track of both the random walk distance and the "total" distance traveled by the impurity, taking into account the sum of all back-and-forth motions. The "total" distance traveled is given by equation 27.14:

$$x_{tot} = \alpha \Gamma t = (0.26x10^{-9}m)(2.2x10^{9}/s)(1s) = 0.57m \tag{27.14}$$

On the other hand, the random walk distance is given by equation 27.15:

$$r = \alpha\sqrt{\Gamma t} = (0.26x10^{-9}m)[(2.2x10^{9}/s)(1s)]^{1/2} = 1.2x10^{-5}m \tag{27.15}$$

This reinforces the facts 1) that lots of jumps are taking place each second, and 2) that lots of the jumps are ineffective insofar as long-distance diffusion is concerned. An average interstitial finds itself 12 micrometers away from its starting point, but having traveled a staggering half a meter plus to get there!
You may see another form of the random walk equation 27.15. Plugging $\sqrt{6D/\Gamma}$ for α (from equation 27.11) into equation 27.15, we obtain equation 27.16:

$$r = \sqrt{6Dt} = [(6(2.5x10^{-11}m^2/s)1s)]^{1/2} = 1.2x10^{-5}m \tag{27.16}$$

in agreement with what we obtained previously.

27.3 Steady State Diffusion

A "steady state" situation is one in which the local concentration of impurities does not change with time, such that $(\partial c/\partial t)_x = 0$ at each value of x. A good example involves the gradual diffusion of hydrogen through the steel wall of a pressurized gas tank, as represented schematically in Figure 27.6.
It has been found that hydrogen dissolves to some extent in steel, maintaining a constant surface composition on the inner surface of the tank that depends upon the tank pressure. Since air surrounds the tank and the hydrogen content in air at ground level is less than 1 part per million, we can assume the hydrogen content on the outside of the tank to be zero. Let the equilibrium hydrogen content on the inner surface be c_H and the wall thickness be

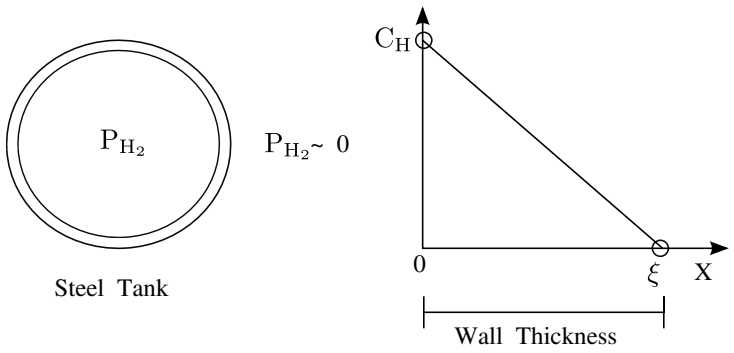


Figure 27.6: Steady state diffusion of hydrogen through a steel tank wall.

ξ. A steady state situation means that the hydrogen content decreases linearly with distance from c_H at the inner wall to zero at the outer wall ($x = \xi$). This can be expressed mathematically as equation 27.17:

$$\left(\frac{\partial c}{\partial x}\right) = \frac{c(x = \xi) - c(x = 0)}{\xi - 0} = \frac{0 - c_H}{\xi} \tag{27.17}$$

Plugging this into Fick's First Law of equation 27.12, we obtain equation 27.18:

$$J = \frac{Dc_H}{\xi} \tag{27.18}$$

The product $D \cdot c_H$ is often called the "gas permeability," since to be permeable through the solid wall the gas has to be both soluble (c_H) and mobile (D). Gas permeability is given in units of $g/cm \cdot s$ or $kg/m \cdot s$, which reflects the product of the diffusion coefficient (cm^2/s or m^2/s) and the appropriate mass concentration (g/cm^3 or kg/m^3). Fortunately, the gas permeability of hydrogen through steel is relatively small at room temperature. Nevertheless, given sufficient time and sufficient diffusion the internal gas pressure (and c_H) will gradually decline to yet another steady state situation. A compelling demonstration of this effect is the loss of helium from mylar balloons. Helium rapidly diffuses through typical rubber balloons. However, mylar balloons have a thin aluminum metal coating that helps to slow down helium diffusion. Nevertheless, helium eventually escapes causing the balloon to slowly lose buoyancy. The situation at any point can be well described by equation 27.18.

27.4 Fick's Second Law

The steady state example just described is an exception rather than the rule. In virtually all other instances of diffusion, the impurity concentration at a given point is changing with time. For such situations we will need another diffusion law, referred to as Fick's Second Law of diffusion. Consider the two concentration profiles in Figure 27.7 and what is happening at imaginary planes 1 and 2, in each case separated by the distance, Δx.

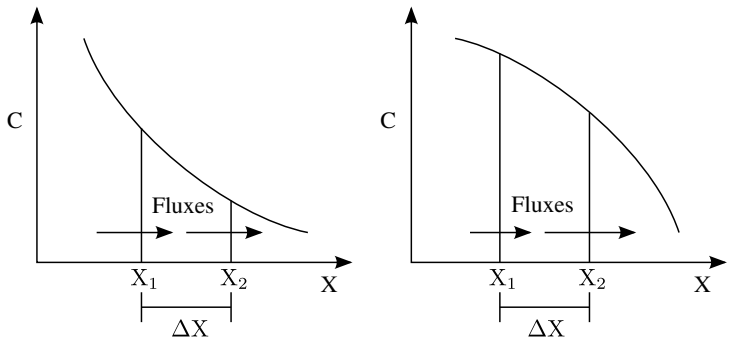


Figure 27.7: Two contrasting concentration profiles and the relative fluxes at two planes separated by a distance, Δx.

Fick's First Law in equation 27.12 tells us that the flux is proportional to the negative of the concentration gradient at each point. Based upon the gradient (slope) at each plane, it can be seen that the flux across plane 1 is greater

than the flux across plane 2 in the concave-up c vs. x profile on the left, and the opposite is true in the concave-down c vs. x profile on the right. We might therefore expect the local concentration between the planes to increase with time in the first situation, but decrease with time in the second. To simplify our discussion, let's assume that the flux is a linear function of distance, as shown in Figure 27.8.

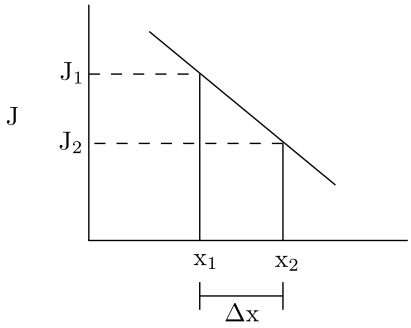


Figure 27.8: Hypothetical linear flux vs. distance relationship that might correspond to the concave-up concentration profile in Figure 27.7.

This might correspond to the first situation in Figure 27.7. As can be seen, the flux into the volume between the planes is greater than the flux out, so we would expect the impurity concentration in that volume to increase with time. If we specify an identical area, A, on each plane, we can calculate the number of impurity atoms that are added to the volume between the two planes during an increment of time, δt, with the following equation 27.19:

$$(J_1 - J_2)A\delta t = A\Delta x \delta c \tag{27.19}$$

Since the units of flux are $(\#/cm^2)/s$, when we multiply $(J_1 - J_2)$ by the product of area (A in cm^2) and time increment (δt in s), we arrive at the number of impurity atoms that are added to the volume between the planes. Since the volume in question is just the product of A and Δx, this amounts to an increment in the impurity concentration (δc in $\#/cm^3$). We can use equation 27.19 to help us derive Fick's Second Law. Let's begin by doing a Taylor series expansion of flux around J_1 in equation 27.20:

$$J_2 = J_1 + \left(\frac{\partial J}{\partial x}\right)\Delta x \tag{27.20}$$

again ignoring higher order terms. Rearranging equation 27.20 we obtain equation 27.21:

$$J_1 - J_2 = -\left(\frac{\partial J}{\partial x}\right)\Delta x \tag{27.21}$$

Inserting the right side of this equation in place of $(J_1 - J_2)$ in equation 27.19 and canceling the A and Δx terms on both sides of the resulting equation yields equation :

$$-\left(\frac{\partial J}{\partial x}\right)\delta t = \delta c \tag{27.22}$$

In the limit that the increments (δt, δc) go to zero, it follows that equation 27.22 can be expressed as equation 27.23:

$$\left(\frac{\partial c}{\partial t}\right) = -\left(\frac{\partial J}{\partial x}\right) \tag{27.23}$$

However, we know from Fick's First Law that $J = -D(\partial c/\partial x)$. Substituting $-D(\partial c/\partial x)$ for J in equation 27.23, we obtain equation 27.24:

$$\left(\frac{\partial c}{\partial t}\right) = \frac{\partial}{\partial x}\left(D\frac{\partial c}{\partial x}\right) \tag{27.24}$$

This is the most general form of Fick's Second Law, which holds for all situations, including those in which the diffusion coefficient is a function of composition (and therefore position). Fortunately, in many situations the diffusion coefficient does not vary significantly over the concentration range involved and therefore does not

change significantly with position. This simplifies things considerably. Fick's Second Law of diffusion becomes equation 27.25:

$$\left(\frac{\partial c}{\partial t}\right) = D\left(\frac{\partial^2 c}{\partial x^2}\right) \tag{27.25}$$

This equation tells us that the rate of change of concentration at a given point is proportional (by the diffusion coefficient) to the second derivative of concentration with respect to distance at that point. We can use this equation to 1) ascertain whether the concentration of an impurity is increasing or decreasing with time at a chosen point, given a specific concentration profile, and 2) to test whether specific diffusion equations are valid, namely that all such equations must obey Fick's Second Law as expressed in equation 27.25 (see the following section). As an illustration of the first application, consider the three concentration profiles in Figure 27.9:

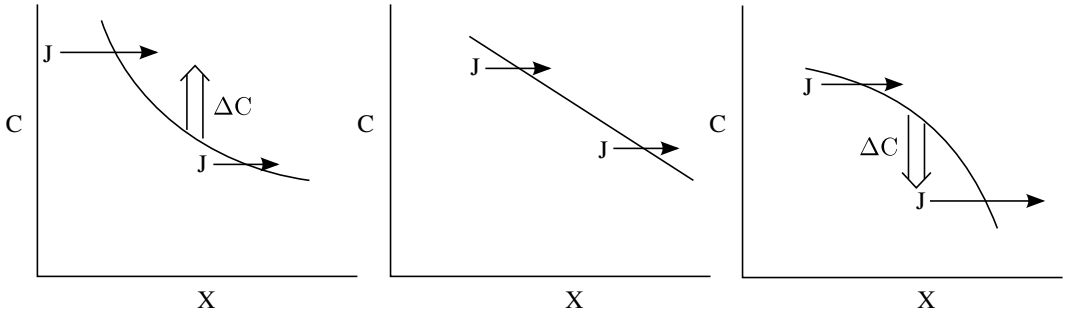


Figure 27.9: Analysis of three concentration profiles by Fick's Second Law.

The concave-up concentration profile on the left has a second derivative of concentration with respect to distance that is everywhere positive. This means that $(\partial c/\partial t)$ must be positive; concentration is increasing with time as shown by the vertical double-arrow. The linear concentration profile in the middle plot has a second derivative with respect to distance that is zero. This means that $(\partial c/\partial t)$ is zero; local concentration is not changing with time. This is the "steady-state" situation discussed in the previous section. The concave-down concentration profile on the right has a second derivative of concentration with respect to distance that is everywhere negative. This means the $(\partial c/\partial t)$ is everywhere negative; concentration is decreasing with time as shown by the vertical double-arrow.

28 Applications of Fick's Laws

28.1 Homogenization and Point Defect Relaxation

Imagine a sinusoidal concentration profile as shown in Figure 28.1. Using Fick's Second Law, it follows that in the concave-down portions of the profile where the local concentration is greater than the average composition, $\bar{c}$, the concentration will be decreasing with time. At the same time, in the concave-up portions of the profile where the local concentration is less than the average composition, the concentration will be increasing with time.

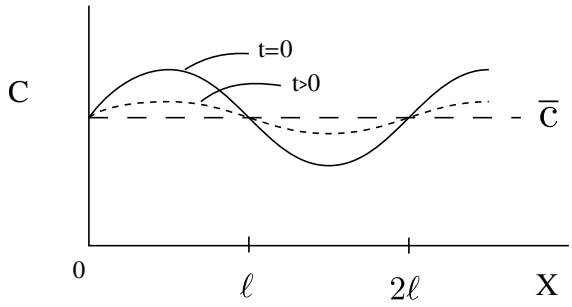


Figure 28.1: Sinusoidal composition profile.

At time zero we can express the concentration profile as equation 28.1:

$$c(x,0) = \bar{c} + \beta(0) sin\left(\frac{\pi x}{l}\right) \tag{28.1}$$

where $\beta(0)$ is the amplitude of the sine wave at time zero. If we allow diffusion to proceed, the concentration profile will relax as shown in Figure 28.1. It turns out that the solution to Fick's Second Law for these boundary conditions is given by equation :

$$c(x,t) = \bar{c} + \beta(0) sin\left(\frac{\pi x}{l}\right) \exp\left(-\frac{Dt\pi^2}{l^2}\right) \tag{28.2}$$

You can prove to yourself that this equation is a solution to Fick's Second Law by taking the first derivative with respect to time (t) and comparing the result with the second derivative with respect to distance (x). The two derivatives should be identical. Let's test this equation by setting $t = 0$. This simply reverts to the time zero function in equation 28.1. Now let's consider the variation of composition at the point, $x = l/2$. Since the sine of $(\pi(l/2)/l)$ or $\pi/2$ is unity, equation 28.2 becomes equation 28.3:

$$c(\frac{l}{2},t) - \bar{c} = \beta(0) \exp\left(-\frac{Dt\pi^2}{l^2}\right) = \beta(0) \exp\left(-\frac{t}{\tau}\right) \tag{28.3}$$

The parameter, τ, is referred to as the "relaxation time," and is given by equation :

$$\tau = \frac{l^2}{\pi^2 D} \tag{28.4}$$

When time is equal to the relaxation time $(t = \tau)$ the amplitude above the average composition at the point, $x = l/2$, according to equation 28.3 will be $\beta(0)/e$ or $0.368\beta(0)$. This will, in fact, be true of every point along the profile; each composition will be 36.8% of its value at time equal to zero, as shown in Figure 28.2.

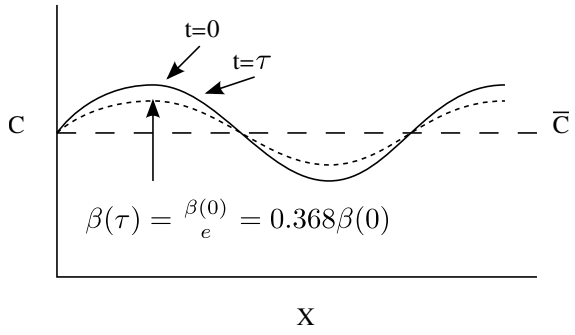


Figure 28.2: Initial profile and concentration profile at one relaxation time.

One very important application of the sine function concentration profile of equation 28.2 has to do with point defect relaxation. In metals and in some oxides, we can quench in a high vacancy population by quenching from high temperature. If we then take the specimen to an intermediate temperature, excess vacancies can be annihilated, but only by migrating to a surface. In polycrystalline materials, grain boundaries can act as internal "surfaces." In either single crystal or polycrystalline materials, dislocation cores can also act as internal "surfaces." Figure 28.3 shows schematic diagrams of how grain boundaries and dislocations can act as "sinks" for excess vacancies.

In the case of grain boundaries we can equate the grain size to the value of l in the relaxation time equation 28.4. If instead dislocations dominate, we can equate the average dislocation spacing to the value of l in equation 28.4. In both cases, the diffusivity of interest will be the vacancy diffusion coefficient at the temperature of interest. Polishing and etching can be employed to determine the dislocation density (etch pits) or $\#/cm^2$. Inverting the value obtained gives an area per dislocation or cm^2 per dislocation, as shown in Figure 28.4. If we take the square root of this value, we obtain the average separation distance of dislocations to use as l in the relaxation time equation 28.4.

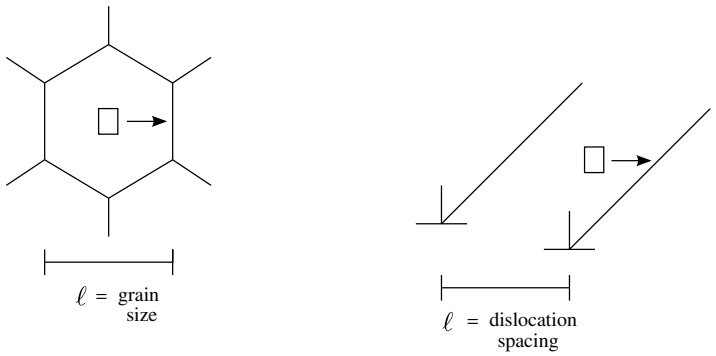


Figure 28.3: Schematics of grain boundaries and/or dislocation cores acting as "sinks" for vacancies, indicates by the squares.

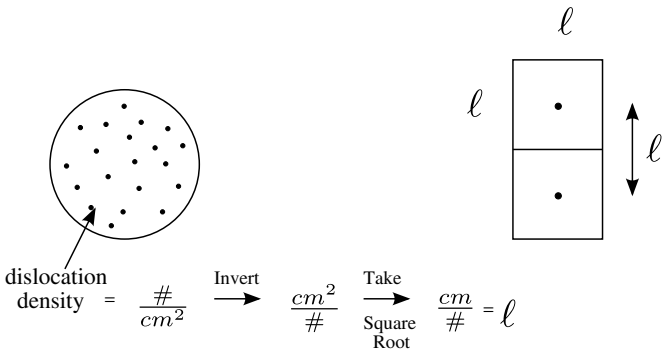


Figure 28.4: Acquiring the average dislocation spacing from a measured dislocation density.

28.2 Non-Infinite Systems

"Non-infinite systems" are those where concentration profiles span the entire specimen, from one end to the other. Or if diffusion is taking place from both sides of a plate or slab, the two concentration profiles overlap in the middle of the specimen. A classic example is diffusion out of a slab. For example, in Figure 28.5 a slab of thickness, h, with initial impurity concentration, c_o, is held in an environment that takes the surface concentration, c_s, to zero.

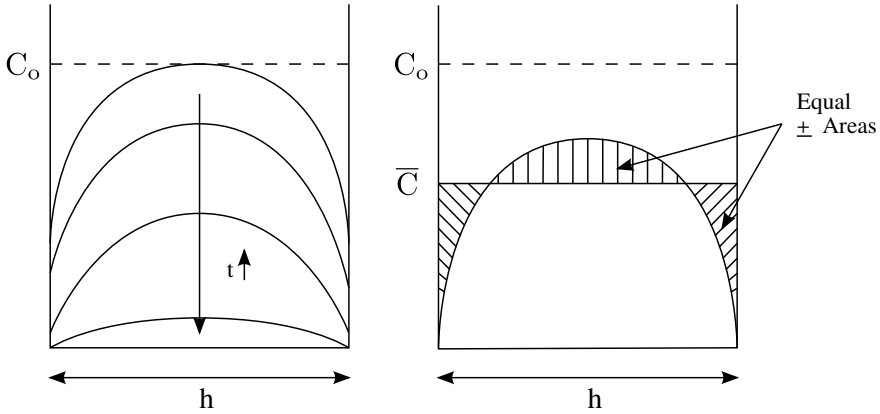


Figure 28.5: Diffusion of an impurity out of a slab.

At short times the diffusion profiles do not overlap. They are described by "semi-infinite" solutions to Fick's Second Law, as will be described in the following section. At longer times, however, the diffusion profiles do overlap. The solution to Fick's Second Law for this situation is complicated, as shown in equation 28.5.

$$c(x,t) = \frac{4c_o}{\pi}\sum_{j=0}^{\infty}\left(\frac{1}{2j+1}\right) sin\frac{(2j+1)\pi x}{h}\exp\left[-\left(\frac{(2j+1)\pi}{h}\right)^2 Dt\right] \tag{28.5}$$

Fortunately, the first term dominates, giving us equation 28.6:

$$c(x,t) = \frac{4c_o}{\pi} sin(\frac{\pi x}{h})\exp\left(-\frac{Dt\pi^2}{h^2}\right) \tag{28.6}$$

This equation would be used if one wanted to estimate a composition at a specific point and time. More often, however, we are interested in the average composition in the overall slab, as shown in the diagram on the right side of Figure 28.5. This is found by integrating equation 28.5 to obtain equation 28.7:

$$\bar{c}(t) = \frac{1}{h}\int_0^h c(x,t)dx = \frac{8c_o}{\pi^2}\sum_{j=0}^{\infty}\frac{1}{(2j+1)^2}\exp\left[-\left(\frac{(2j+1)\pi}{h}\right)^2 Dt\right] \tag{28.7}$$

Again, fortunately, the first term dominates, and for average compositions less than 80% of the initial composition ($\bar{c} \leq 0.8c_o$) the first term is an excellent approximation to the solution. The results in equation 28.8:

$$\frac{\bar{c}}{c_o} = \frac{8}{\pi^2}\exp(-\frac{t}{\tau}) \tag{28.8}$$

where $\tau = h^2/\pi^2 D$ is called the relaxation time.
Equation 28.8 is very useful for describing the degassing of metals, including decarburization of steels. In this case the diffusion coefficient in the relaxation time would be that of the particular gas species or of carbon in the particular metal. Another application involves the relaxation of vacancies in a supersaturated metal, for example a metal quenched from high temperature has a large vacancy population. If the metal is taken to an intermediate temperature where the equilibrium vacancy concentration is small (effectively zero compared to the quenched-in concentration, c_o) and vacancies are sufficiently mobile that they can annihilate at sinks such as dislocation cores or grain boundaries, relaxation will occur. The situation involving dislocation cores is represented schematically in Figure 28.6, where the concentration of vacancies is shown during relaxation from the initial, quenched-in concentration, c_o, to a point where the average vacancy concentration is $\bar{c}$. In this case, h would be the average dislocation spacing and the diffusion coefficient would be that of vacancies in the metal of interest.

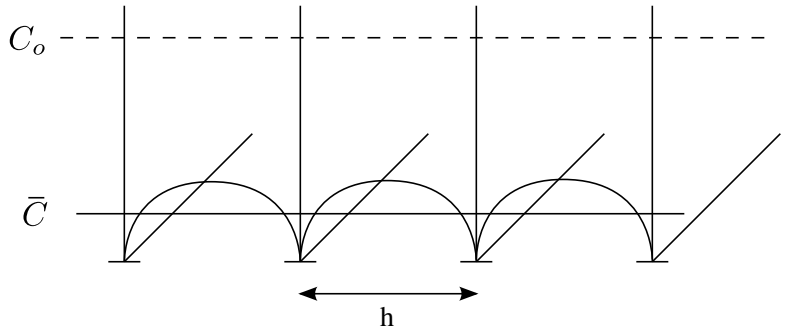


Figure 28.6: Relaxation of vacancies to dislocation cores in a metal.

28.3 Semi-Infinite Systems

By "semi-infinite" systems we are referring to situations where the specimen can be treated as essentially "infinite" in size compared to the extent of the impurity concentration profile at the surface. Another way of describing this is that the concentration profile resulting from diffusion never reaches the other end of the specimen. The specimen, although not infinite in extent, is effectively "infinite" insofar as diffusion is concerned. In terms of diffusion it is "semi-infinite."

28.3.1 Thin Film Tracer Diffusion

Thin film "tracer" diffusion is of special importance to materials science and engineering. It is how diffusion coefficients are often measured. Imagine a thin film containing an area concentration ($\#/cm^2$) of a "tracer" that is deposited on a highly polished, flat surface of a single crystal (crystal #1 in Figure 28.7). To prevent any loss of "tracer" (for example by evaporation) and also to provide a second crystal in which to study diffusion, the highly polished, flat surface of crystal #2 in Figure 28.7 is butted up against crystal #1, sandwiching the tracer thin film between them. For self-diffusion studies, the tracer is the same chemical species as constitutes both crystals, however with a different atomic mass (namely, another isotope than that of the host atoms making up the two crystals). In many instances, a radioactive tracer is employed at low concentrations (and low radioactive emission levels to protect lab workers). In this case, the relative concentration of tracer at a specific diffusion depth is determined by the number of radioactive "counts" registered at that depth, usually by "serial sectioning" (see below). In other cases, the amount of a non-radioactive tracer diffused to a given depth can be established by the use of a mass spectrometer that can differentiate and quantify the relative amounts of tracer species vs. naturally occurring (host crystal) species.

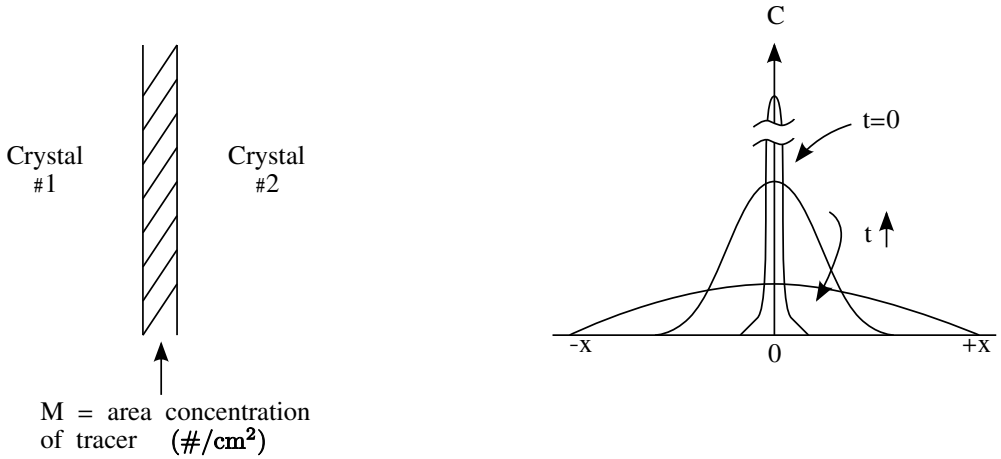


Figure 28.7: Schematic of thin film tracer diffusion.

For impurity diffusion studies, the tracer is a different chemical species from the host. Again, either radioactive or non-radioactive tracers can be employed, with mass spectrometry being used to in the latter case to quantify the amount of a diffused species found at a given depth from the surface by serial sectioning, as described below.
The unique aspect of thin film tracer diffusion is that the overall amount of tracer remains constant, once applied to the surface and sandwiched between the two crystals. On the right side of Figure 28.7 are the concentration profiles at early, intermediate, and long times of diffusion at a specified temperature. Assuming no loss of tracer, the area under each curve (the total amount of tracer) must remain the same at all times.
The actual diffusion process is carried out by heating the pair of crystals to a predetermined temperature as rapidly as possible, holding for a set time, and "quenching" to room temperature as quickly as possible. For ceramic samples, heating and cooling rates are limited by what the crystals can sustain without fracture caused by thermal shock. Once a diffusion heat treatment is completed, the two crystals are cleaved at the thin film interface and each crystal is subjected to the process called "serial sectioning," as represented in Figure 28.8. Roughly equivalent "sections" are carefully removed, beginning with the surface to which tracer was applied, and being careful to maintain a flat/planar surface throughout the process.

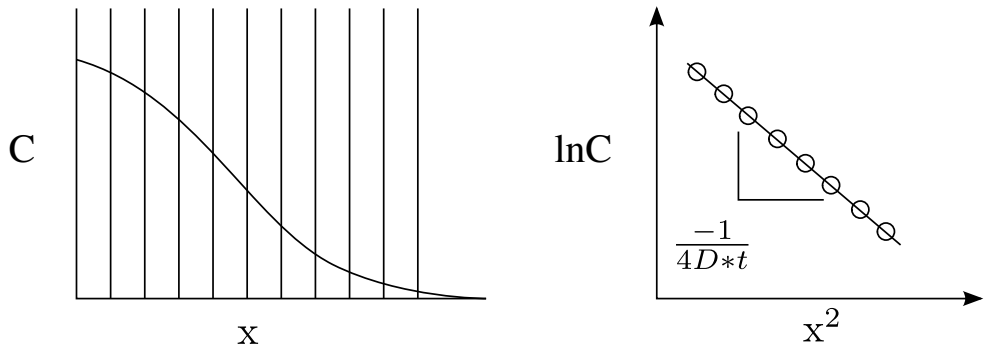


Figure 28.8: Schematic of the serial sectioning process, and how the tracer diffusion coefficient is determined.

Sectioning is typically accomplished using the same grinding media (abrasive cloths, powders) used for metallographic sample preparation. It is very important, however, to quantify both the "depth" removed (each Δx in Figure 28.8) and the tracer concentration (C in Figure 28.8) for each section. Alternatively, the count rate (counts/-time) can be measured by a radiation detector. The count rate will be directly proportional to the concentration in each section. The increment of depth can be measured directly by a high precision micro-caliper. Alternatively, it can be indirectly calculated from the mass removed, which can be measured quite accurately. In the case of non-radioactive tracers, mass spectrometry is employed to calculate the tracer content in a given section. With a radioactive tracer, the abrasive cloth holding the powder of the "section" removed can be placed in a radiation detector and counted for a predetermined time period. The solution to Fick's Second Law that reflects the boundary conditions for thin film tracer diffusion is given in equation 28.9:

$$C = \frac{M}{2\sqrt{\pi D t}} \exp\left(\frac{-x^2}{4Dt}\right) \tag{28.9}$$

where C is either the tracer concentration or the tracer's radioactive count rate, and M is the area concentration of tracer deposited on the surface of crystal #1 in Figure 28.7. Taking the natural logarithm of both sides of equation 28.9 gives equation 28.10:

$$\ln C = \ln\left(\frac{M}{2\sqrt{\pi D t}}\right) - \frac{x^2}{4Dt} \tag{28.10}$$

The slope of a plot of the natural logarithm of C vs. x^2 yields $-1/4D^*t$, as depicted on the right side of Figure 28.8. Since the diffusion time is known, the diffusion coefficient can be calculated. This process is repeated at other temperatures in order to establish the pre-exponential factor (D_o) and activation energy (Q) of the tracer diffusion coefficient (D^*) as in equation 28.11:

$$D^* = D_o \exp\left(-\frac{Q}{RT}\right) \tag{28.11}$$

where D_o is called the "pre-exponential factor" and Q is the activation energy of diffusion. We discuss the origin(s) of this characteristic of diffusion vs. temperature behavior (so-called Arrhenius behavior) in a later section.

28.3.2 Constant Surface Composition Situations

As opposed to thin film tracer diffusion, where the "surface" composition diminishes with time, there are a number of situations in materials science and engineering where the "surface" composition is maintained constant with time. These include "doping" situations, where a solid is held in an atmosphere that keeps the composition of an impurity at a constant level on the exposed surface. Similarly, the carburization of iron or steel can be controlled by holding the surface in an atmosphere with a fixed ratio of carbon monoxide and carbon dioxide. The amount of carbon in solution at the surface is thereby fixed according to equation 28.12:

$$2CO(gas) \Leftrightarrow CO_2(gas) + \bar{C}(insolution) \tag{28.12}$$

As carbon diffuses into the interior of the solid, the above reaction guarantees that additional carbon is added to the surface to maintain the surface carbon composition. The overall impurity (or carbon) content of the solid is given by equation 28.13:

$$c_{tot} = A\int_0^{\infty} c(x)dx \tag{28.13}$$

where A is the area of the surface into which solute is diffusing. This means, of course, that as opposed to thin film tracer diffusion, the overall solute content steadily increases with time. This is shown schematically on the left side of Figure 28.9, where c_o is the initial carbon content and c_s is the surface composition. The concentration profile moves progressively to the right with time, as the interior of the solid is progressively enriched with solute. We later consider how thick the sample must be to be "semi-infinite" compared to the diffusion profile, but from Figure 28.9 it follows that the sample dimension should be large relative to the diffusion profile at the longest time applied.

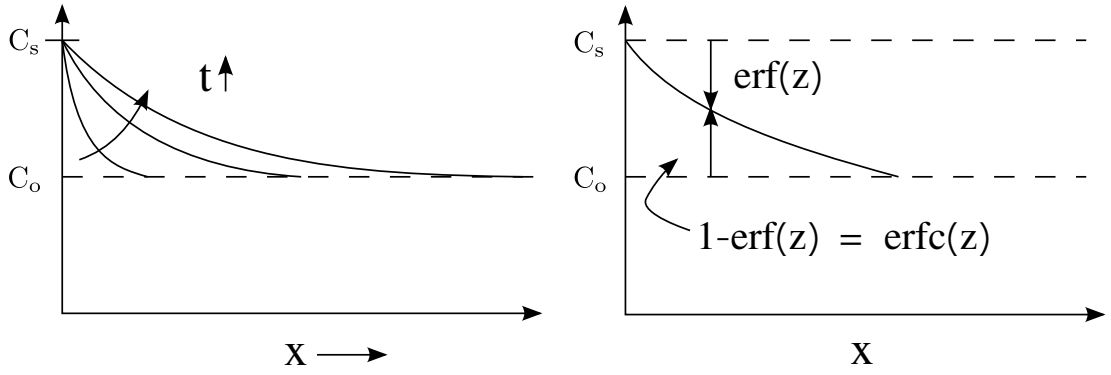


Figure 28.9: Schematic of impurity diffusion from a fixed surface concentration into a solid with a background concentration and, on the right the situation cast in terms of the error function and the complementary error function.

Before we introduce the solution to Fick's Second Law that meets the boundary conditions of Figure 28.9, we will make the simplifying assumption that the diffusion coefficient does not vary with impurity content, and therefore distance from the surface (x). This is valid for the doping of semiconductors from gas phase precursors, where relatively small dopant levels are involved. However, in the case of carburization, relatively large changes in solute (carbon) content occur and the diffusion coefficient is known to vary somewhat with carbon composition (and therefore position) along the diffusion profile. Fortunately, relatively good predictions of carbon diffusion profiles can be obtained by employing an "average" carbon diffusion coefficient (discussed later) that can be assumed invariant with carbon content (and position).
The solution to Fick's Second Law corresponding to the boundary conditions in Figure 28.9 is given by equation 28.14:

$$c(x,t) = c_s - (c_s - c_o)\text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.14}$$

which is usually written as equation 28.15:

$$\frac{c_s - c(x,t)}{c_s - c_o} = \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.15}$$

where x is the distance from the surface, D is the diffusion coefficient (a constant diffusion temperature is assumed), and t is the time period over which diffusion takes place. You should have encountered the "error function" (erf) in prior math courses. It is an indefinite integral of the form in equation 28.16:

$$erf(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-\eta^2}d\eta \tag{28.16}$$

where η is known as a "dummy variable." In our case, $z = x/(2\sqrt{Dt})$. In days past, with no closed-form solution to equation 28.16, one resorted to tables of the error function. Fortunately, many modern calculators now include the error function. Alternatively, there have been many mathematical approximations. The most straightforward amongst the collection in the Handbook of Mathematical Functions [12], having a reported accuracy of $\pm 5x10^{-4}$, is equation 28.17:

$$1 - erf(z) = \text{erfc}(z) \approx (1 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4)^{-4} \tag{28.17}$$

where $a_1 = 0.278393$, $a_2 = 0.230389$, $a_3 = 0.000972$, and $a_4 = 0.078108$. This approximation can be used in a either a computer program or in a mathematical spreadsheet to estimate erf(z). In equation 28.17 we have also introduced the "complementary error function," denoted as erf$c(z)$. The simple relationship between the two functions in equation 28.18:

$$\text{erfc}(z) = 1 - \text{erf}(z) \tag{28.18}$$

is illustrated in Figure 28.10. On the right side of Figure 28.9 we see these two functions as applied to the above mentioned instance of doping/carburization of a sample with an initial dopant concentration. An abbreviated table of error function values is given in Table 1.

Table 28.1: Examples of Error Function Values

z	erf(z)	z	erf(z)	z	erf(z)
0	0	0.8	0.742	1.6	0.976
0.2	0.223	1.0	0.842	1.8	0.989
0.4	0.428	1.2	0.910	2.0	0.995
0.6	0.604	1.4	0.952	2.2	0.998

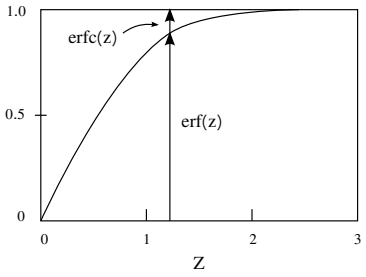


Figure 28.10: The error function and the complementary function.

We can generate a useful "metric" to describe diffusion in cases of constant surface composition (error function solutions to Fick's Second Law). In the case just described, consider the diffusion profile in Figure 28.11.

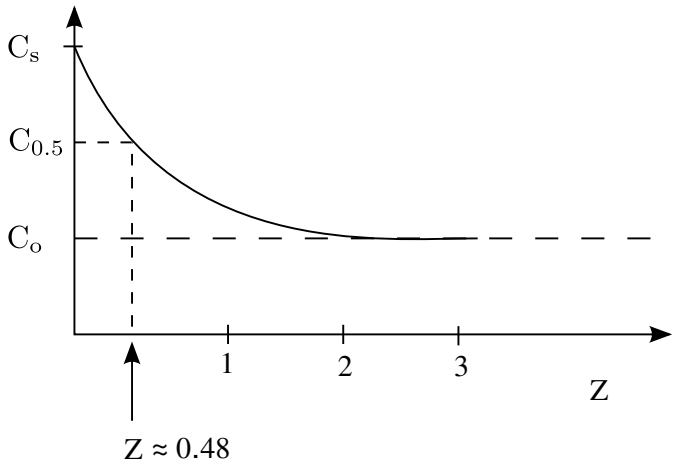


Figure 28.11: Schematic representation of the "diffusion depth" for the diffusion process in Figure 28.9.

You will see one special composition highlighted, labeled $c_{0.5}$. This corresponds to the point where the composition is half way between the surface composition (c_s) and the original sample composition (c_o), which modifies equation 28.15 to equation 28.19:

$$\frac{c_s - c_{0.5}}{c_s - c_o} = 0.5 = \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.19}$$

It turns out that erf(z)=0.5 for a value of z that is very close to 0.5 (see the error function Table or Figure 28.10), or more precisely $z = x/2\sqrt{Dt} = 0.477$. If we let $z \approx 0.5$, it follows that $x_{0.5} \approx \sqrt{Dt}$. This distance of "root Dt" is commonly referred to as the "diffusion depth." It is a useful "metric" for diffusion in situations involving constant surface composition. For example, one might ask, "How much longer must I diffuse at the same temperature to double the diffusion depth?" The answer follows from the definition of $x_{0.5}$ in equation 28.20:

$$\frac{x'_{0.5}}{x_{0.5}} = 2 = \frac{\sqrt{Dt'}}{\sqrt{Dt}} = \sqrt{\frac{t'}{t}}; \frac{t'}{t} = 4 \tag{28.20}$$

or four times the length of time. We can also establish the thickness of sample required to be "semi-infinite" with respect to a given combination of diffusion coefficient and time. From the previous Table of error function values, we can observe that at $z = x/2\sqrt{Dt} = 2$ the diffusion profile has diminished by 99.5% toward the original composition. In other words, very little (0.5%) impurity has been delivered at this point in the diffusion profile. For diffusion from one side, a specimen must therefore be greater than $2\sqrt{Dt}$, or twice the "diffusion depth." For diffusion from both sides of a plate we might define "semi-infinite" as being $> 4\sqrt{Dt}$ or roughly four times the "diffusion depth." Of course, this is arbitrary and more stringent definitions of "semi-infinite" can be made (e.g., $z = 2.5$, for which erf$(z) = 0.9996$). For diffusion from both sides of a plate, this would amount to 0.08% added impurity at the middle of the specimen.

Of course, we should not forget our non-infinite solutions to Fick's Second Law, which are required in place of error function (semi-infinite) solutions when there is significant overlap of diffusion profiles from the two sides of a plate. Equations 28.5 and 28.6 would be used to determine the concentration at a given position, whereas equations 28.7 and 28.8 would be employed to determine the average composition (above the initial composition) for the plate. We revisit such overlap when we consider decarburization (below).

All the solutions that follow are simply permutations of equation 28.15. For instance, let's consider carburization or doping of an initially "pure" host. Of course, from thermodynamics we know that there is no such thing as complete purity. Impurities, including the one we are interested in, are always present to some degree. However, their level can be considered so inconsequential as to be effectively zero, namely c_o can be taken to be zero in equation 28.15. By making this assumption and rearranging, we obtain equation 28.21:

$$\frac{c(x,t)}{c_s} = 1 - \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) = \text{erfc}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.21}$$

A representative diffusion profile is displayed in Figure 28.12, showing how the complementary error function comes into the picture.

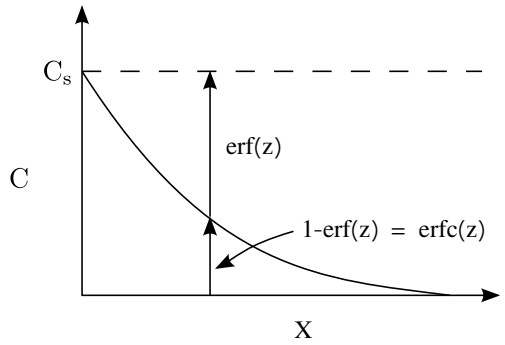


Figure 28.12: Carburization or doping of a solid with an essentially zero initial impurity content.

Now let's consider decarburizing, which is the opposite of carburizing (adding carbon to iron or steel). On the left side of Figure C, we are holding the surface carbon content at a lower value than the initial carbon content of the specimen. This can be accomplished by employing appropriate carbon monoxide/carbon dioxide gas ratios, as per equation 28.12. As time proceeds, carbon is removed by diffusion from progressively larger depths. Keep in mind, however, that the specimen is "semi-infinite," namely that its width is at least 2 diffusion depths (or 4 diffusion depths if decarburized from both sides). The solution to Fick's Second Law satisfying the boundary conditions in Figure C is given in equation 28.22:

$$\frac{c(x,t) - c_s}{c_o - c_s} = \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.22}$$

In the event that complete decarburization is desired, the surface carbon content is taken to zero, simplifying equation 28.22 to equation 28.23:

$$\frac{c(x,t)}{c_o} = \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.23}$$

A schematic illustrating this situation is given on the right side of Figure 28.13, where it can be seen that the straightforward error function is the simple solution for the boundary conditions applied. It should once again be stressed that the sample must be "semi-infinite," namely if being decarburized from both sides it should be several (>4) "diffusion lengths" in width, such that there is no overlap of error function solutions at the middle of the sample. If there is considerable overlap, we once again turn to our non-infinite solutions. These take into account both the carbon contents vs. position (equation 28.6) and the average carbon content of the slab (for values below 80% of the original, equation 28.8). It might be a good time to refer back to Figure 28.5 for a schematic of decarburization in the "non-infinite" case.

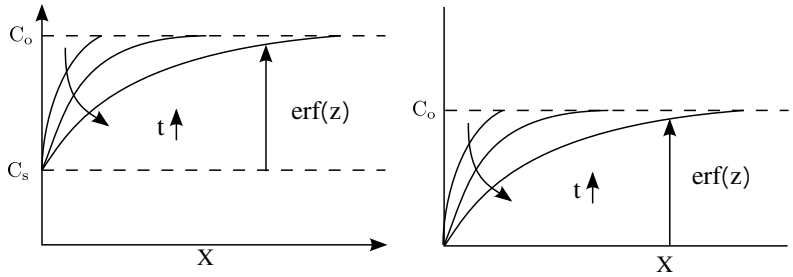


Figure 28.13: Decarburization of a sample with an initial carbon content with the surface content taken to a finite, but lower value (left) or to effectively zero (on the right).

Before considering the special case of interdiffusion, let's introduce some practical uses for all the error function solutions considered thus far. For many years, the n-type and p-type regions of microelectronic circuits were made by exposing "semiconductor-grade" (ultra-high purity) silicon wafers to gases containing n-type dopants or p-type dopants through removable masks in a process called "photolithography." Although diffusion is sometimes still used, for instance to set the background impurity level of a silicon layer, nowadays doping is usually accomplished by ion-implantation of dopants owing to improved controllability and speed.

Carburization and decarburization are fundamental to the processing of steels and for the preparation of their surfaces. For example, carburization is well known to increase the hardness of steels. In the case of a machine gear, for instance, the materials engineer may be interested in maintaining different properties of the gear core (high strength with toughness) and the gear surface (high strength with hardness). This calls for a different carbon content in the core vs. on its surface. In particular, carburization is employed to boost the carbon content at the surface, making it hard and wear-resistant. This can be done by a process of "pack carburization," namely packing the gear in a high temperature bed of charcoal, often referred to as "case hardening." However, carburization is more often accomplished by exposing steel surfaces to carbon-containing gases (see equation 28.12) or to carbon-containing plasmas.

The final instance of constant surface composition that we consider involves the interdiffusion of solute across an interface, as depicted in Figure 28.14.

Here a solid with impurity concentration, c_1, is butted up against a solid having a lower impurity concentration, c_2. It should be stressed that for the error function solution to be employed, the differences in solute content should be small and/or the diffusion coefficient must not vary much with solute content (and therefore distance in the diffusion "couple"). With these qualifications, the mid-point of the diffusion profile, both in terms of composition and in terms of distance, remains stationary with time as shown in Figure 28.14. It can be seen that the composition at this point remains the average of the two initial solute contents, or $(c_1 + c_2)/2$. Looking to the right from the interface ($x = 0$), the situation looks identical to the carburization situation in Figure 28.9, and we can insert $(c_1 + c_2)/2$ in place of c_s in equation 28.15 as is done in equation 28.24:

$$\left[\frac{(\frac{c_1+c_2}{2}) - c(x,t)}{(\frac{c_1+c_2}{2}) - c_2}\right] = \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.24}$$

The denominator of the left side of equation can be simplified to $(c_1 - c_2)/2$, and the resulting solution to Fick's Second Law for interdiffusion with a constant diffusion coefficient is equation 28.25:

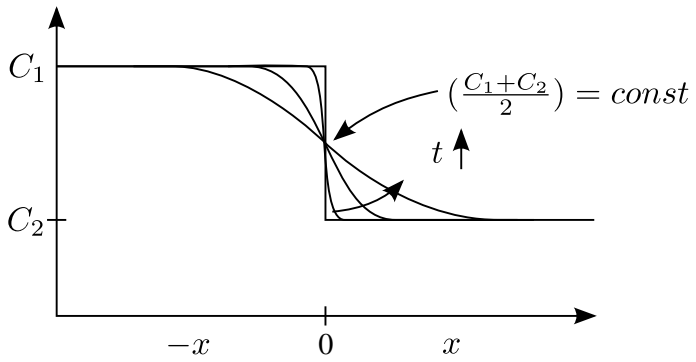


Figure 28.14: An interdiffusion "couple" assuming a constant value of diffusion coefficient.

$$c(x,t) = \left(\frac{c_1 + c_2}{2}\right) - \left(\frac{c_1 - c_2}{2}\right) \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.25}$$

This equation also holds for the left side of the interdiffusion couple in Figure , since $\text{erf}(-z) = -\text{erf}(z)$ as shown in Figure 28.15.

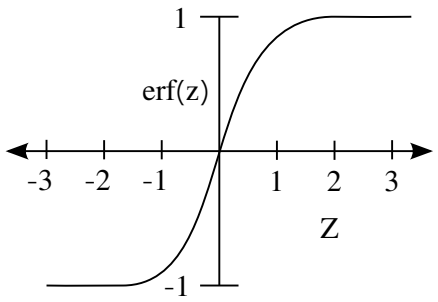


Figure 28.15: The error function for both positive and negative values of z.

Some textbooks have equation 28.25 as equation 28.26:

$$c(x,t) = \left(\frac{c_1 + c_2}{2}\right) + \left(\frac{c_1 - c_2}{2}\right) \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{28.26}$$

This arises from reversing the sense of "x" (and therefore "z"), namely that the solid with the higher solute concentration is now on the right, so be careful to use the solution that matches your boundary conditions. Note how the overall error function in Figure 28.15 actually looks like an interdiffusion profile with the more highly doped specimen on the right.

Before moving on, it should be stressed that interdiffusion with an essentially composition-independent and therefore position-independent diffusion coefficient is rather the exception than the rule. Most interdiffusion problems involve "interdiffusion" coefficients that are clearly functions of concentration and therefore of position. In an A:B couple, where A and B are different metals, the usual case is for A to diffuse faster into B or vice versa. In fact, the point of average composition, $(c_1 + c_2)/2$, moves with diffusion time. Fortunately, procedures exist to solve interdiffusion profiles for composition-dependent diffusion coefficients at each point. You will learn about these methods in lab projects and in upper level materials science and engineering coursework.

29 Atomistics of Diffusion

We now turn to the "basics" of diffusion, namely the "nuts and bolts" of how diffusion takes place in solids. We want to "unpack" all the contributions, the so-called "atomistics" that contribute to the pre-exponential factor (D_o) and the activation energy (Q) of the general equation for diffusion in solids, as per equation 29.1:

$$D = D_o \exp\left(\frac{-Q}{RT}\right) \tag{29.1}$$

It turns out that we can do a pretty good job estimating ranges of values for the pre-exponential factor for each of the primary diffusion mechanisms. In what follows, we take a modular "plug-in" approach to the problem, as outlined in Figure 29.1:

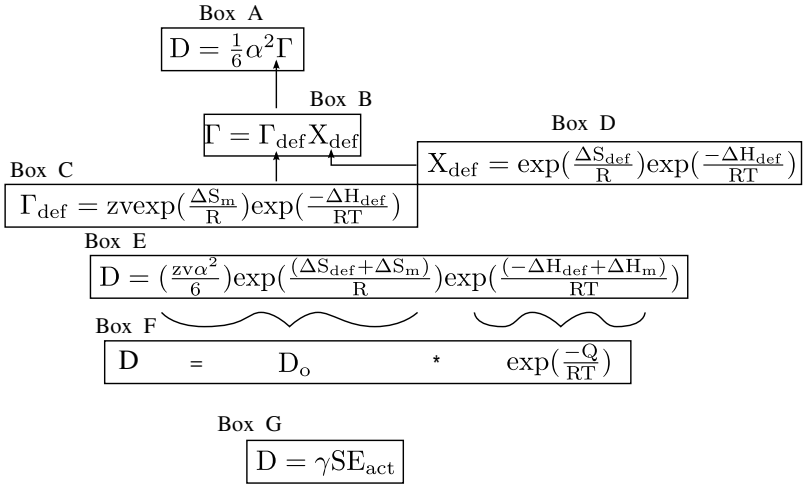


Figure 29.1: A modular plug-in approach for understanding the atomistics of diffusion in cubic crystals.

We begin with our basic diffusion equation 27.11 repeated and reorganized here as equation 29.2:

$$D = \frac{1}{6}\Gamma\alpha^2 \tag{29.2}$$

To review, α is the atomic jump distance (determined by the crystal structure) and Γ is the atomic jump frequency. This equation (29.2) is found in Box A of Figure 29.1 and will be the "jumping off point" for each of the diffusion mechanisms we consider. The first "plug-in" is equation 29.3:

$$\Gamma = \Gamma_{def} X_{def} \tag{29.3}$$

Here the atomistic jump frequency is related to the product of a "defect" jump frequency (Γ_{def}) and a defect availability factor (X_{def}). The availability factor will be further described as we consider each of the individual diffusion mechanisms. The rest of Figure 29.1 may seem quite daunting, however not every diffusion mechanism involves all of the "plug-ins" or boxes in the figure. We will proceed from simplest (interstitial diffusion) to most complex (vacancy and interstitialcy diffusion). Along the way, the "plug-ins"/boxes should come into focus, so be patient. The object is to develop for each mechanism an atomistic-based description of the diffusion coefficient. Right now Box E in Figure 29.1 seems quite complicated, however this massive equation can be reduced to the standard equation 29.1 in Box F or to the even simpler equation 29.4:

$$D = \gamma S E_{act} \tag{29.4}$$

in Box G, which tells us that three groupings of factors govern each individual diffusion mechanism: a geometric factor (γ), and entropic factor (S), and a thermally-activated term (E_{act}). Let's get started by considering the most straightforward of diffusion mechanisms, interstitial diffusion.

29.1 Interstitial Diffusion

Although self-diffusion by interstitials is possible in certain ceramic materials, we will restrict our consideration to impurity diffusion where the impurity spends all of its time in the interstices of the crystal lattice. A good example is carbon diffusion in iron. This type of impurity diffusion has special requirements, namely that 1) the atomic radius of the impurity is significantly less than that of the host (e.g., $r_C < r_{Fe}$), and 2) the atomic radius of the impurity is comparable to that of the interstices. It should be stressed that self-diffusion in close-packed metals never takes place by interstitial diffusion. The radius of a host atom is far too great for it to squeeze into the much smaller interstices in between host atoms. To place a host atom into such an interstice would be highly unfavorable energetically; self-interstitials are therefore highly unlikely.

Let's begin to analyze interstitial diffusion using the modular plug-in chart in Figure 29.1. Since the interstitial is always an interstitial, there is no thermal activation process associated with its formation as in Box D. Furthermore, the availability factor for interstitial jumps can be assumed to be essentially unity ($X_{int} \approx 1$), namely adjacent interstitial sites into which the interstitial can jump can be assumed to be empty. Of course, at the highest carbon contents there is a finite possibility that an adjacent interstitial is occupied. We could express the availability factor as $(1 - X_i)$, where X_i is the site fraction of carbon interstitials. But for simplicity we will assume that $X_i \ll 1$ and assume an availability factor (X_{int}) of unity.
It is time to add important functions to our diffusion "toolkit." These have to do with the defect jump frequency in Box C of Figure 29.1. In moving from one interstitial site to another, a certain amount of lattice dilation must occur, as illustrated in Figure 29.2.

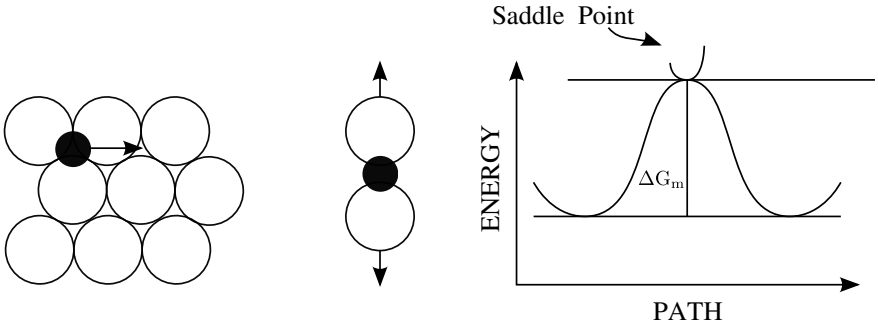


Figure 29.2: Schematic showing interstitial motion, local lattice dilation and the resulting energy barrier/saddle point.

In other words, adjacent atoms must move apart to enable the interstitial to pass. Energetically, this amounts to surpassing an energy barrier that separates the two sites. Figure 29.2 illustrates the special configuration with the interstitial at the so-called "saddle-point". The energy landscape is shaped like a saddle. Alternately, you can think about a mountain pass. You go up and over the pass, but when you are at the top of the pass you find the land rising to either side. From statistical mechanics, the probability that an interstitial has sufficient energy to reach the saddle point is small. It is governed by the height of the pass (ΔG_m) and the temperature. Higher temperature means more energy being available to the interstitials, and an increasing fraction of them having the probability to make it to/past the saddle point. The relationship is given by equation 29.5:

$$probability \propto \exp\left(\frac{-\Delta G_m}{RT}\right) \tag{29.5}$$

We refer to ΔG_m as the free energy of motion which, in turn, can be separated into enthalpy of motion (ΔH_m) and entropy of motion (ΔS_m) components as in equation 29.6:

$$\Delta G_m = \Delta H_m - T\Delta S_m \tag{29.6}$$

It should be apparent that an interstitial at the saddle point has a different entropy. Certainly, its vibrational frequency, and thus its vibrational entropy, should be different at the dilated/congested saddle than in either interstice, i.e., the energy wells on either side of the energy barrier in Figure 29.2. We can use the information in equation 29.6 to arrive at the equation in Box C of Figure 29.1, or equation 29.7:

$$\Gamma_i = z\nu \exp\left(\frac{\Delta S_{mi}}{R}\right) \exp\left(\frac{-\Delta H_{mi}}{RT}\right) \tag{29.7}$$

Here we have replaced Γ_{def} with Γ_i (i for interstitial), and ΔS_{mi} and ΔH_{mi} for ΔS_{def} and ΔH_{def}, respectively. The other constants are ν, the attempt frequency, and z, the coordination number of how many empty interstitial sites surround an occupied interstitial site. The attempt frequency can be assumed to be the lattice vibrational frequency, on the order of 10^{13} per second. You probably encountered the "Debye frequency" in your prior materials thermodynamics work.
Plugging Γ_i from equation for Γ_{def} in Boxes C and B and $X_i = 1$ for X_{def} in Boxes D and B of Figure 29.1, we obtain the following equation :

$$D = \left[\frac{1}{6}\alpha^2 z\nu \exp\left(\frac{\Delta S_{mi}}{R}\right)\right] \exp\left(\frac{-\Delta H_{mi}}{RT}\right) \tag{29.8}$$

If we consider A to be the host lattice and B to be the species moving via interstices, we can simplify equation 29.8 to equation 29.9:

$$D_B = D_{Bo} \exp\left(\frac{-Q_I}{R}\right) \tag{29.9}$$

where D_{Bo} is the pre-exponential factor for impurity diffusion of species B by interstitial mechanism, which includes all the entities within the brackets of equation 29.8, and Q_I is the activation energy for interstitial motion. Since there is no defect formation energy, Q_I is the same as the enthalpy of motion or ΔH_{mi}. Before we take a try at predicting pre-exponential factors for interstitial diffusion, let's rearrange equation 29.9 to separate out geometric (γ) and entropic (S) factors, as in equation 29.10:

$$D = \gamma S E_{act} = \left(\frac{\alpha^2 z \nu}{6}\right)\left(\exp\frac{\Delta S_{mi}}{R}\right)\exp\left(\frac{-\Delta H_{mi}}{RT}\right) \tag{29.10}$$

Let's first unpack the underlying parameters in the geometric factor (γ). We already know that the attempt frequency (ν) is on the order of $10^{13}/s$. But the jump distance (α) and coordination number (z) depend upon not only the crystal structure, but upon the actual element/metal under consideration. Let's focus our concentration on iron in both fcc (austenite) and bcc (ferrite) forms, and consider the case of the interstitial diffusion of carbon. The two crystal structures are displayed in Figure 29.3. More importantly, the interstitial positions are marked with "X's." In addition, a typical jump distance is shown by an arrow for each structure.

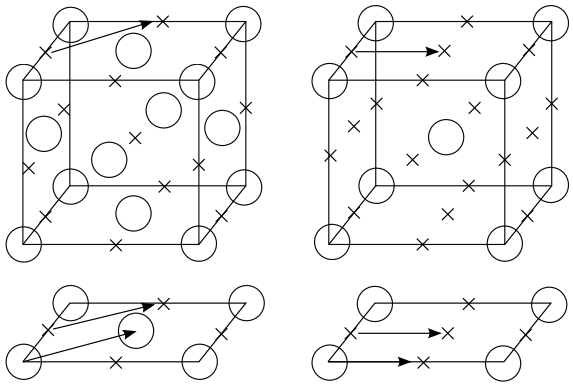


Figure 29.3: Interstitial positions in fcc (austenite) on left and in bcc (ferrite) on right.

Let's first consider carbon diffusion in fcc-Fe (austenite), whose structure is shown on the left side of Figure 29.3. It is obvious that the two sets of lattice positions (host atoms, interstices) form interpenetrating fcc lattices. In fact, we will encounter interpenetrating lattices when we consider the "rocksalt" structure of sodium chloride or table salt (NaCl), where the electropositive Na^+ "cations" sit on one set of sites and the electronegative Cl^- "anions" sit on the other. In austenite, most of the second interpenetrating (interstitial) lattice is empty. However, the fact that it (the interstitial lattice) is also fcc is quite important. It means that the coordination number of empty interstitials sites around any given occupied interstitial site will be the same as the coordination number of neighboring iron atoms around any given iron host atom. This gives us the value we need for the coordination number in γ of equation 29.10, namely 12.

On the other hand, the jump distance (or the arrow length in Figure 29.3) is the same as the distance between close-packed atoms (see the lower diagram in Figure 29.3), or one-half of the face-diagonal $\left(\sqrt{2}a_o/2\right)$. Since the lattice parameter (a_o) of fcc iron is ~ 0.37 nm. the jump distance would be $8\sim 0.26$ nm. Diffusion coefficients are rarely given in terms of nm^2/s, however. So let's convert 0.26 nm to centimeters, or $2.6x10^{-8}$ cm. This yields the following for the geometric factor for diffusion:

$$\gamma = z\blacksquare ff^2/6 = 12\left(10^{13}\,s^{-1}\right)\left(2.6x10^{-8}\,cm\right)^2/6 = 0.014\,cm^2/s \tag{29.11}$$

There have been several estimates of the entropic term (S) in equation 29.10. One of the most straightforward calculations took into account the elastic strain associated with the saddle point (C. Zener in Imperfections in Nearly Perfect Crystals, 1952, p.289) and also the change in vibrational entropy associated with the saddle point [19], arriving at the result that the entropy of interstitial motion was $\Delta S_{mi} \approx R$, such that the entropic term ($S = \exp(\Delta S_{mi}/RT)$)

was approximately 2.7. The product of the geometric factor (γ) and the entropic factor (S) in equation 29.10 is the pre-exponential coefficient of equation 29.9. Based upon our "back of the envelope" calculations, we predict a value for D_{Bo} of $\sim 0.38\,\mathrm{cm}^2/\mathrm{s}$.
Before we make a comparison with experimental values, it must be stressed that owing to the larger interstice size in fcc iron vs. bcc iron, the former (austenite) can incorporate a much larger carbon content than the latter (ferrite). We mentioned previously that the diffusion of carbon in fcc iron was not a constant, but rather varied measurably with carbon content. Fortunately, the variation is not extreme and a practicable composition-independent approximation appears in most diffusion textbooks. This approximation is given by the following equation 29.12:

$$\bar{D}_C(\text{austenite}) \approx 0.2\frac{\mathrm{cm}^2}{\mathrm{s}} \exp\left(\frac{-138\mathrm{kJ/mole}}{\mathrm{RT}}\right) \tag{29.12}$$

Given our rather simplistic assumptions and the fact that the experimental equation is itself an approximation, the agreement of pre-exponential values (estimated: $0.38\,\mathrm{cm}^2/\mathrm{s}$ vs. experimental: $0.2\,\mathrm{cm}^2/\mathrm{s}$) is remarkable.
Let's see how we do estimating the geometric factor and pre-exponent for carbon diffusion in the bcc (ferrite) form of iron. Interestingly, the interstices in bcc iron exist only on the faces of the unit cell as shown on the right side of Figure 29.3. These interstices are found along the edges and at the face centers of the unit cell. As a result, the jump distance is half of the cube edge or lattice parameter, such that $\alpha = a_o/2 = 0.29nm/2 = 0.14nm$. Since there are no bcc interstices within the unit cell, the coordination number can be determined by looking at the face-centered interstice on any face in Figure 29.3. If occupied, it is surrounded by 4 empty interstices at the 4 cube edge positions, so $z = 4$. Assuming once again an attempt frequency of $\sim 10^{13}/s$, we arrive at a geometric factor of $\approx 0.0013cm^2/s$.
Once again, there have been several estimates of the entropic term (S) in equation 29.10. If we assume, as with the fcc structure, that the entropy of interstitial motion is $\Delta S_{mi} \approx R$ such that the entropic term of $S = \exp(\Delta S_{mi}/RT)$ is approximately 2.7, we arrive at a "back of the envelope" estimate for D_{Bo} of $\sim 3.5x10^{-3}\,\mathrm{cm}^2/\mathrm{s}$. The experimental equation for carbon diffusion in bcc iron is [36]:

$$D_C(\text{ferrite}) = 2.0x10^{-2}\mathrm{cm}^2/\mathrm{s}\exp\left(\frac{-84.1\mathrm{kJ/mole}}{RT}\right) \tag{29.13}$$

This time there is only agreement of pre-exponent within an order of magnitude, however this can be considered quite good considering all the approximations we have made. As additional support for our atomistic approach, the corresponding D_{Bo} values are $3.0x10^{-3}\,\mathrm{cm}^2/\mathrm{s}$ for nitrogen [37] and $1.0x10^{-3}\,\mathrm{cm}^2/\mathrm{s}$ for hydrogen [21], in quite satisfactory agreement with our prediction. The corresponding enthalpies of motion (Q_I) are 76.1 kJ/mole and 13.4 kJ/mole for N [37] and H [21], respectively.

29.2 Vacancy Diffusion

Self-diffusion or impurity-diffusion by vacancy mechanism can be likened to the "tile" puzzles we have handled physically or solved as a computer game. Consider the tile puzzle in Figure 29.4. The configuration on the top left is insoluble owing to the absence of a vacancy, whereas the configuration on the top right is soluble by virtue of the vacancy's presence. The analogy to solids is that without such defects, there can be no diffusion and thus no reactivity. This is further illustrated in the bottom diagrams of Figure 29.4, each illustrating a close-packed plane in either hcp or fcc solids. With the addition of the vacancy in the bottom-right configuration, it is obvious that self diffusion of the atom to the left of the vacancy takes place by exchange with the vacancy. By moving the vacancy all around the lattice, each atom will eventually have its turn to exchange with the vacancy. From the outset, we might guess that the vacancy diffusivity will be much larger than that of the host or "self" atoms; the self-diffusion coefficient will be much smaller than the vacancy diffusion coefficient.

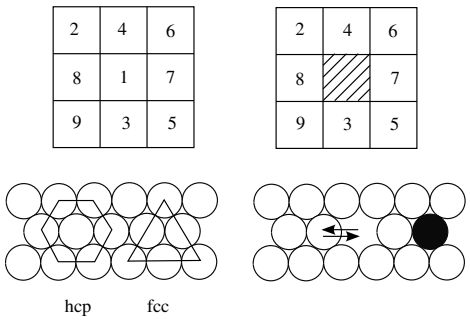


Figure 29.4: Tile puzzle analogy for both self-diffusion and substitutional impurity diffusion.

Substitutional impurities, namely those that sit on host atom sites rather than in its interstices, also depend upon vacancy motion for diffusion to take place. The shaded atom in the bottom-right configuration in Figure Z represents such a substitutional impurity. It is obvious that for this to take place, the atomic radius of the impurity must be comparable to that of the host. Impurity atoms will similarly bide their time waiting to exchange with a host vacancy that comes their way. They may exchange more readily or less so than the host atoms, but both diffuse by so-called "vacancy mechanism."

Let's consider vacancy diffusion within the rubric of the modular "plug-in" diagram of Figure 29.1. Box B of Figure 29.1 will become equation 29.14:

$$\Gamma = \Gamma_v X_v \tag{29.14}$$

where Γ_v is the vacancy jump frequency and X_v is the "availability" factor or site fraction of vacancies. The latter amounts to the probability that each adjacent site is unoccupied. The first of these two terms can be derived by analogy to equation 29.7 for interstitial motion, or equation 29.15:

$$\Gamma_v = z\nu \exp\left(\frac{\Delta S_{mv}}{R}\right) \exp\left(\frac{-\Delta H_{mv}}{RT}\right) \tag{29.15}$$

where z is the coordination number, ν is the attempt frequency, and ΔS_{mv} and ΔH_{mv} are the entropy and enthalpy of vacancy motion, respectively.

We will address each of these parameters later, but for now let's concentrate on the site fraction of vacancies or X_v. If temperature is sufficiently high, the vacancy concentration will be in thermodynamic equilibrium. We can approach this problem as was done for the regular solution model in your prior thermodynamics coursework. Expressions were written for both the enthalpy of mixing in equation 29.16:

$$\Delta H^M = \Omega X_A X_B \tag{29.16}$$

and the entropy of mixing in equation 29.17:

$$-T\Delta S^M = -T\left[-R(X_A \ln X_A + X_B \ln X_B)\right] \tag{29.17}$$

where Ω is the "interaction parameter" and X_A and X_B represent the mole fractions of the two components. The phase boundaries of the regular solution solvus were found by plugging these two equations into the overall expression for the free energy of mixing, equation 29.18:

$$\Delta G^M = \Delta H^M - T\Delta S^M \tag{29.18}$$

taking the derivative with respect to composition (X_B), and setting the result equal to zero.

In the case of the formation of vacancies in a solid metal, the corresponding enthalpy term is given by equation 29.19:

$$\Delta H \simeq X_v \Delta H_v \tag{29.19}$$

where ΔH_v represents the enthalpy increase per mole of vacancies. Similarly, the entropy of forming a mole of vacancies is given by equation 29.20:

$$\Delta S = \Delta S_{nonconfig} + \Delta S_{config} \tag{29.20}$$

where the non-configurational entropy, incorporating the thermal/vibrational entropy per mole of vacancies (ΔS_v), is given by equation 29.21:

$$\Delta S_{nonconfig} \simeq X_v \Delta S_v \tag{29.21}$$

and the configurational entropy looks a lot like the entropy of mixing (in equation 29.17) or equation 29.22:

$$\Delta S_{config} = -R[X \ln X_v + (1 - X_v) \ln(1 - X_v)] \tag{29.22}$$

If we begin with the standard free energy of the "perfect" crystal (G_A^o), and insert all the enthalpic and entropic terms for vacancy formation, the Gibbs free energy of the defective crystal compared to that of the perfect crystal is given by equation 29.23:

$$G_A - G_A^o = \Delta G_v = X_v \Delta H_v - T X_v \Delta S_v + RT \left[X_v \ln X_v + (1 - X_v) \ln(1 - X_v)\right] \tag{29.23}$$

Figure 29.5 shows how the enthalpic and entropic factors contribute to the overall free energy of vacancy formation.

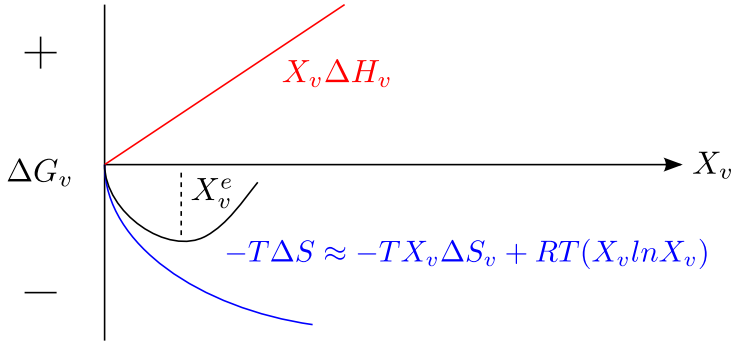


Figure 29.5: Competition of the enthalpic and entropic terms in vacancy formation, resulting in an equilibrium concentration at X_v^e.

At first, ΔG_v decreases with increasing vacancy concentration, being dominated by the entropy terms. Eventually, however, the enthalpy term turns things around, dominating at higher vacancy concentrations. A minimum in the free energy of vacancy formation occurs at the so-called equilibrium value (X_v^e). This value can be derived by taking the first derivative of equation 29.23 and setting it equation to zero as in equation 29.24:

$$\left(\frac{\partial G}{\partial X_v}\right)_{X_v = X_v^e} = 0 = \Delta H_v - T\Delta S_v + RT \ln\left(\frac{X_v^e}{1 - X_v^e}\right) \tag{29.24}$$

Rearranging equation 29.24, exponentiating, and making the assumption that $X_v^e \ll 1$, we obtain equation 29.25, which relates the equilibrium vacancy concentration (site fraction of vacancies) to temperature:

$$X_v^e = \exp\left(\frac{\Delta S_v}{R}\right) \exp\left(\frac{-\Delta H_v}{RT}\right) \tag{29.25}$$

For now, let's leave the testing of our assumption ($X_v^e \ll 1$) for later, and plug this equation (29.25) into Box D of Figure 29.1. We previously derived equation 29.15, which is Box C of Figure 29.1. Plugging these equations into the jump frequency expression of Box B, we obtain an overall equation for self-diffusion (Box A) by vacancy mechanism, or equation 29.26:

$$D_A = \left[\frac{1}{6}\alpha^2 z \nu \exp\left(\frac{\Delta S_v + \Delta S_{mv}}{R}\right)\right] \exp\left[\frac{-(\Delta H_v + \Delta H_{mv})}{RT}\right] \tag{29.26}$$

which can be simplified to the basic equation 29.27:

$$D_A = D_{Ao} \exp\left(\frac{-Q_{SD}}{RT}\right) \tag{29.27}$$

where Q_{SD}, the activation energy of self-diffusion, is the sum of formation (ΔH_v) and motion (ΔH_{mv}) enthalpies, and D_{Ao} is the pre-exponential factor that incorporates all the factors within the first bracket of equation 29.26. As

with interstitial diffusion, we can reorganize equation 29.26 into geometric (γ), entropic (S), and activation energy (E_{act}) factors, as in equation 29.28:

$$D_A = \gamma S E_{act} = \left(\frac{\alpha^2 z\nu}{6}\right)\left[\exp\left(\frac{\Delta S_v + \Delta S_{mv}}{R}\right)\right]\left[\exp\left(\frac{-(\Delta H_v + \Delta H_{mv})}{RT}\right)\right] \tag{29.28}$$

Before we examine the factors in the pre-exponent and estimate its value, it is useful to calculate the diffusion coefficient of the vacancies themselves. Plugging equation 29.15 into Box C of Figure 29.1 and inserting this for the Γ term in Box A, we obtain equation 29.29:

$$D_v = \frac{1}{6}\alpha^2\Gamma_v = \left[\frac{1}{6}z\nu\alpha^2\exp\left(\frac{\Delta S_{mv}}{R}\right)\right]\exp\left(\frac{-\Delta H_{mv}}{RT}\right) \tag{29.29}$$

Comparing equations 29.26 and 29.29, two important relationships result, equations 29.30 and 29.31:

$$\Gamma_A = \Gamma_v X_v \tag{29.30}$$

$$D_A = D_v X_v \tag{29.31}$$

As we said above, the vacancies are moving around much more frequently than individual host atoms. In the following section, we mention that the vacancy concentration in fcc metals can be as high as $X_v \sim 0.001$ at just below their melting points. Simple math shows us that the ratio of jump frequencies (Γ_v/Γ_A) or diffusivities (D_v/D_A), that is of vacancies vs. host atoms, will therefore be 1000x! Actually, you can prove this to yourself by solving a "tile" puzzle and counting the number of jumps made by each numbered tile compared to the number of jumps the "vacancy" makes to solve the puzzle.

Now let's estimate a range of pre-exponential factors for self-diffusion in fcc metals based upon our best guess of atomistic factors. We will assume a representative value for the jump distance in equation 29.26 of approximately $0.3nm$ or $0.3x10^{-7}cm$. For example, recall that we obtained $\alpha = 0.26nm$ for fcc iron. We also established that the coordination number (z) of adjacent sites around any given host atom, on which a vacancy might sit, is 12. Again, we will assume the familiar value for the attempt frequency (ν) of $\sim 10^{13}/s$. Theorists find values for both the entropy of formation and motion to be on the order of R, or $\exp(\Delta S/R) \sim 2.7$, however a range of values is found in the literature. We will employ a range of values of $1 \lesssim \exp(\Delta S_v) \lesssim 10$ and $1 \lesssim \exp(\Delta S_{mv}) \lesssim 10$. Plugging all these values into equation 29.26, we obtain a range of pre-exponent values of $0.018cm^2/s \lesssim D_{Ao} \lesssim 1.8cm^2/s$.

So how did we do? The pre-exponential factors for self diffusion are $1.7cm^2/s$, $1.9cm^2/s$, and $0.30cm^2/s$ for fcc aluminum, nickel, and copper, respectively [15]. Given all the assumptions we made, the agreement is quite good.

What about self-diffusion in bcc metals? Again, let's assume a representative value for the jump distance in equation of approximately $0.3nm$ or $0.3x10^{-7}cm$. As opposed to fcc metals, the jump distance is different for site-to-site jumps vs. interstitial-to-interstitial jumps. This is demonstrated in the bcc structure of Figure 29.6, where the near-neighbor jump is one-half the body diagonal, or $(\sqrt{3}a_o)/2$. In bcc iron, this jump distance would be $0.25nm$ or $.25x10^{-7}cm$. From the same Figure 29.6, the coordination number is obviously 8 as opposed to 12 for fcc metals. Again, assuming a typical attempt frequency (ν) of $10^{13}/s$ and a range of entropic factors, such that $1 \lesssim \exp(\Delta S_v) \lesssim 10$ and $1 \lesssim \exp(\Delta S_{mv}) \lesssim 10$, and inserting these values into the first bracket of equation 29.26, we obtain as an estimate, $0.012cm^2/s \lesssim D_{Ao} \lesssim 1.2cm^2/s$.

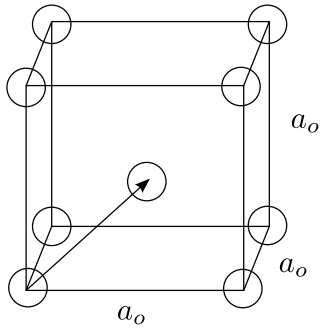


Figure 29.6: Jump distance for self-diffusion in bcc metals.

Again, how did we do? The pre-exponential factors for self-diffusion are $2.0 cm^2/s$, $1.8 cm^2/s$, and $0.2 cm^2/s$ for bcc iron, molybdenum, and chromium, respectively[15]. Again, the experimental values are in good agreement with our simple-minded predictions.

29.3 Interstitialcy Diffusion

Of the three diffusion mechanisms we have thus far considered, interstitialcy diffusion is certainly the most unusual. The German word for this mechanism is "zwischengitterstossmechanismus," which does a far better job conveying the concept. The words that make up this moniker will help us understand its meaning. The word "gitter" refers to the normal lattice sites. The word "zwischen" means "between." And the word "stoss" means "shove." This mechanism is exclusive to compounds with more than one sublattice. This is because the interstice size in metals is never big enough to accommodate a self-interstitial. However, in certain ceramic materials one species, for example the electropositive cations, can be much smaller than the electronegative anions. And in this case the cation interstices can be large enough to accommodate cation self-interstitials. The classic example is AgBr, which readily forms Ag self-interstitials. But rather than moving around by interstitial mechanism, as carbon does in iron, a coordinated two-cation motion mechanism takes place, in which an interstitial cation bumps another cation off its normal site and takes its place. That's where the "shove" part of zwischengitterstossmechanismus comes in. The normal cation now becomes an interstitial cation, and the process continues. A schematic the first step of such a process in AgBr is shown in Figure 29.7 in terms of "before" and "after" positions of ions. Only the Ag^+ species are shown, since all the action involves silver cations.

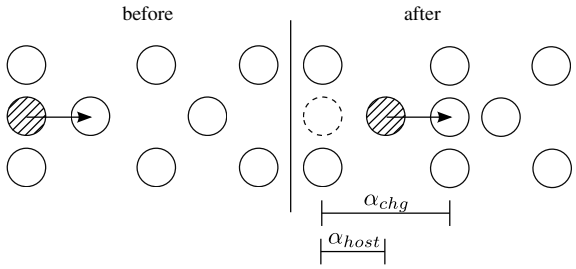


Figure 29.7: Schematic of the "before" and "after" positions of cations undergoing interstitialcy diffusion.

I like to think of one of two analogies for interstitialcy mechanism. The first involves "sending" an opponent's ball in croquet. Once I touch your ball with mine, I have the option of positioning my ball under my foot right next to yours and then clobbering my ball, the transferred momentum sending your ball way off course. The other analogy involves shuffleboard (or curling, if you have ever seen it), where the momentum of my puck/stone is transferred to your puck, which (hopefully) is removed from the scoring area while mine takes its place. Two points need to be made about the interstitialcy event. As shown in Figure 29.7, the process is "collinear," meaning that the pushed atom moves along the same line as the pushing atom. This need not be the case, as we well know from the shuffleboard/curling analogy. Non-collinear interstitialcy mechanisms are definitely known to occur. Second, note that the interstitial has moved twice as far as the host atom that got pushed. This may not seem significant, but it is actually one way to confirm that the collinear interstitialcy mechanism is occurring. Later we will find out how to use electrical conductivity to establish the diffusivity by charged ions, so-called ionic diffusivity/conductivity. It turns out that when we measure the diffusivity of silver ions by tracer methods in AgBr, we get an answer that is approximately half the diffusivity of silver ions measured by electrical conductivity. Why is this? Look again at Figure 29.7. The charge (the interstitial) is moving twice as far as the shaded host ion (which could be a tracer ion). Given the complexity of the interstitialcy mechanism and the relative dearth of representative pre-exponential factors, we will not subject this mechanism to the intense atomistic scrutiny as we did for interstitial and vacancy mechanisms. However, we can make some remarks about "jump frequency" and "defect availability" in Box B of Figure 29.1. It makes sense that the interstitialcy defects are jumping with a frequency, Γ_{iy} ("iy" for interstitialcy). But if I am a silver cation on a host site, instead of waiting for "pullers" to come along (vacancies), I am now biding my time awaiting the arrival of a "pusher" to come along (interstitialcy defects) and knock me on my way. Hence the defect availability factor will be the fraction of occupied interstices, or X_{iy}. By analogy with the vacancy mechanism, the corresponding Box E interstitialcy diffusion coefficient (with direct analogy to equation 29.26) is equation 29.32:

$$D_A = \left[\frac{1}{6}\alpha^2 z\nu \exp\left(\frac{\Delta S_{iy} + \Delta S_{miy}}{R}\right)\right] \exp\left[\frac{-(\Delta H_{iy} + \Delta H_{miy})}{RT}\right] \tag{29.32}$$

As with the vacancy mechanism, this equation has geometric (γ), entropic (S), and thermally-activated components (E_{act}), as in equation 29.33:

$$D_A = \gamma S E_{act} = \left(\frac{\alpha^2 z\nu}{6}\right)\left[\exp\left(\frac{\Delta S_{iy} + \Delta S_{miy}}{R}\right)\right]\left[\exp\left(\frac{-(\Delta H_{iy} + \Delta H_{miyv})}{RT}\right)\right] \tag{29.33}$$

The motion-related components are relatively straightforward to understand, however the formation components (ΔS_{iy}) and (ΔH_{iy}) will have to wait until we have considered point defects in ionic solids, including the introduction of a special shorthand representation for ionic defects called "Kröger-Vink" notation.

29.4 The Arrhenius Behavior of Diffusion

We have shown that for all diffusion mechanisms considered thus far, Arrhenius behavior results, namely that the diffusion coefficient is thermally activated as per the equation in Box F of Figure 29.1. We already saw this behavior for tracer diffusion in equation 28.7 and for general diffusion in equation 29.1, which is repeated here:

$$D = D_o \exp\left(\frac{-Q}{RT}\right) \tag{29.34}$$

As materials scientists and engineers, we need to be confident in using and manipulating diffusion data that behave in this Arrhenius, thermally-activated fashion. Here are some things we need to be able to do: 1) plot experimental data for diffusion on a $\ln D$ vs inverse temperature (T^{-1}) plot to extract pre-exponent and activation energy, 2) given the pre-exponent and activation energy for diffusion, calculate the diffusion coefficient for a specified temperature, or make the corresponding Arrhenius plot of $\ln D$ vs. T^{-1}, 3) given the diffusion coefficient at one temperature and the activation energy, predict the diffusion coefficient at another temperature, or 4) given diffusion coefficients at two temperatures, be able to extract pre-exponential factor and activation energy, plus predict the diffusion coefficient at an intermediate temperature. All of these tasks are quite straightforward once we recognize that taking the logarithm of both sides of equation 29.34 results in the formula of a line, as in equation 29.35:

$$\ln D = \ln D_o + (\frac{-Q}{R})T^{-1} = b + mx \tag{29.35}$$

where the axes are $y = \ln D$ and $x = T^{-1}$, the slope is $m = -Q/R$ and the intercept is $b = \ln D_o$. Alternatively, the slope is given by equation 29.36:

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\ln D(T_2) - \ln D(T_1)}{T_2^{-1} - T_1^{-1}} \tag{29.36}$$

These relationships are represented schematically in Figure A. It should be noted that diffusion data are more commonly plotted on base-10 $\log D$ vs. T^{-1} plots, but it is easy enough to modify equation 29.35 to equation 29.37:

$$\log D = \log D_o + (\frac{-Q}{2.303R})T^{-1} = b + mx \tag{29.37}$$

Be certain not to make the mistake of plotting temperature in ^oC! All diffusion plots employ temperature in degrees Kelvin.

30 How We Know That Point Defects Exist and Move, Plus Some Useful Relationships

We are going to "turn back the hands of time" to answer these questions. Of course, we can now image individual atoms on very sharp tips in machines called (what else?) atom probes. But long before the existence of such atom probes or high resolution electron microscopes, we knew about atoms and their absence (vacancies) and that such defects moved around. At the times in question (the 1950s and 1960s) the following methods were "de rigueur." Along the way we have learned quite a bit about diffusion in crystalline solids (metals, ceramics) and distilled some important relationships concerning their transport behavior.

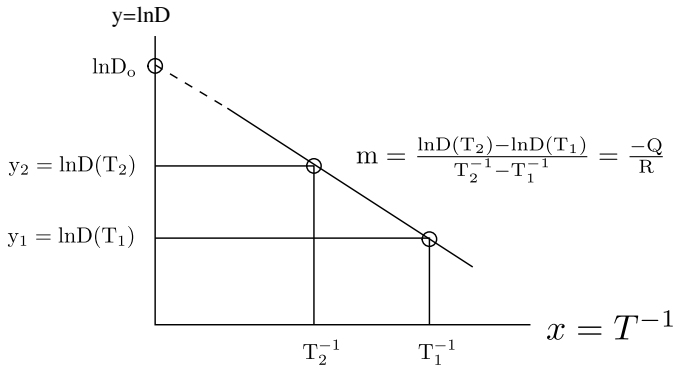


Figure 29.8: Illustration of the Arrhenius behavior of diffusion.

30.1 The Classic Simmons-Balluffi Experiment

You can read about the original experiment for yourself in the reference [31], but this classic experiment employed simultaneous measurement of macroscopic expansion (using a high-accuracy dilatometer) and sub-microscopic/lattice spacing changes (using in situ X-ray diffraction) on a single crystal specimen of an fcc metal during controlled heating. This may seem like a circular argument, but we already anticipate that the vacancy concentration will increase with temperature (see equation 29.25) and should be highest near the melting point. So the procedure involved intrepid experimenters willing to go to high temperatures, even approaching the melting temperature. The general concept is given schematically in Figure 30.1.

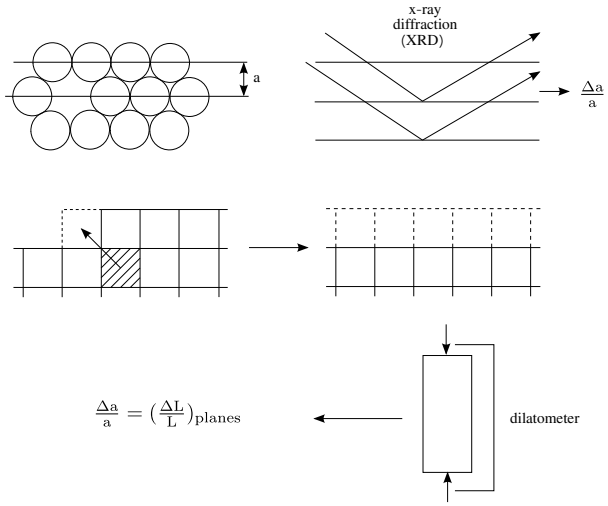


Figure 30.1: Basic concepts of the Simmons-Balluffi method to determine the vacancy concentrations in an fcc metal vs. temperature.

First we see a vacancy added to one of the close-packed planes in the fcc structure. We anticipate that the average lattice spacing would change (possibly decreasing), but in the end this influence ends up being factored out. The second effect has to do with the fact that vacancies can only be added at surfaces, as shown. Given enough vacancies, however, this will amount to extra planes being added to the crystal. This should be reflected in the dilatometer measurements, if we are able to measure very small strains. So whereas the change in lattice parameter, $\Delta a/a$, will be reflected in both X-ray diffraction or XRD (see the schematic of X-rays diffracting from the planes in Figure 30.1) and in dilatometry, the increase in the number of planes owing to vacancy formation ($(\Delta L/L)_{planes}$) will only be reflected in dilatometry. What we need to do is relate a uniaxial (1D) length change to the overall volume change (3D). Turning to Figure 30.2, if the volume of a cube of material at room temperature is given by $V = L^3$, its volume once expanded (and at temperature) will be given by $V' = (L + \Delta L)^3$.
Doing the math for $\Delta V/V = \{[(L + \Delta L)^3 - L^3]/L^3\}$ and dropping all terms involving $(\Delta L)^2$ or $(\Delta L)^3$, since ΔL itself is quite small, we find equation 30.1:

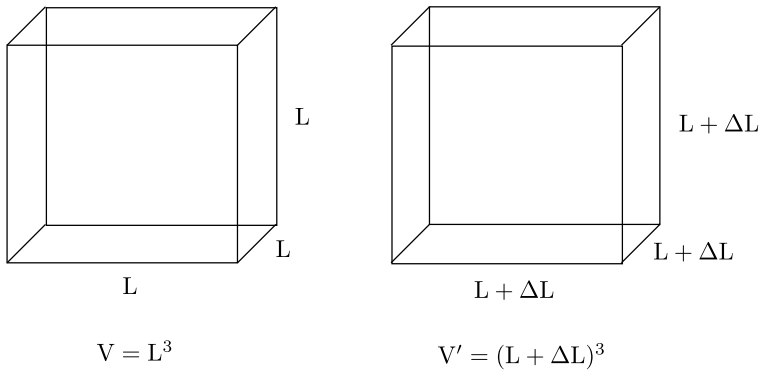


Figure 30.2: Changing from 1D expansion to 3D expansion.

$$\frac{\Delta V}{V} \simeq \frac{3\Delta L}{L} \simeq X_v + \beta X_v \tag{30.1}$$

where the first term on the right reflects the addition of planes owing to vacancy formation, and the second term takes into account the lattice contraction and/or expansion in proportion to the site fraction of vacancies being formed. But we can also argue equation 30.2:

$$\frac{3\Delta a}{a} = \beta X_v \tag{30.2}$$

that the contraction and/or expansion of lattice planes by X-ray diffraction is in proportion to the site fraction of vacancies. Subtracting equation 30.2 from equation 30.1 eliminates the βX_vterm (so we needn't concern ourselves with the value of β) and yields the desired equation 30.3:

$$X_v = 3\left(\frac{\Delta L}{L} - \frac{\Delta a}{a}\right) \tag{30.3}$$

This equation may seem simple enough, but careful examination reveals just how difficult (how de rigueur) this experiment was at the time. Simmons and Balluffi had to measure extremely small differences between two very small quantities as they changed with temperature. You can pull up the original references to examine the original data and plots, but a schematic representation is given in Figure 30.3.

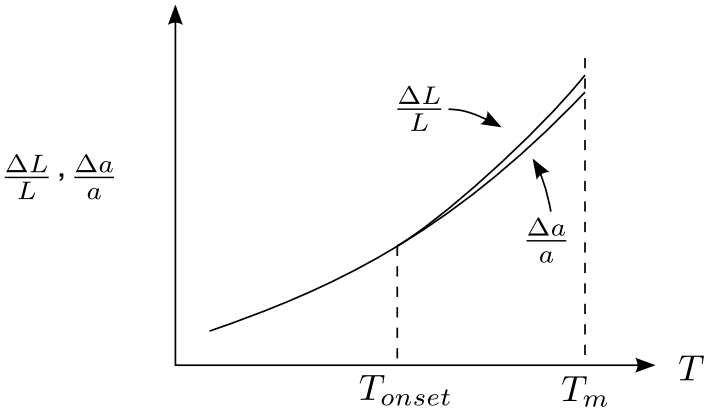


Figure 30.3: Schematic representation of Simmons-Balluffi data for an fcc metal.

A first observation is that both quantities ($\Delta L/L$) and ($\Delta a/a$) increase significantly with temperature, much more than can be explained by the addition of vacancies, and that both are concave upward with temperature. In fact, both quantities are reflecting the fact that thermal expansion is occurring with increasing temperature. This might seem to be a "deal-breaker" at first, until you realize that thermal expansion should be reflected identically in both quantities (dilatometry, XRD). Therefore, as per equation 30.3, any thermal expansion-induced changes will cancel out. A second observation is that up to an onset temperature (T_{onset}) there is no apparent difference in the two

quantities. Actually there should be a difference at every temperature (as long as the vacancy concentration is in thermal equilibrium) however, below T_{onset} the vacancy concentration is simply too small to be registered, meaning that any difference will be undetectable within the experimental limits of the two measurement methods. Above T_{onset}, however, there is a noticeable difference between $(\Delta L/L)$ and $(\Delta a/a)$, whose value is precisely one-third the vacancy concentration. It is obvious that the vacancy concentration increase monotonically with temperature up to the melting point.
Taking the natural logarithm of both sides of equation 29.25 we obtain equation 30.4:

$$\ln X_v = \frac{\Delta S_v}{R} + (\frac{-\Delta H_v}{R})T^{-1} = b + mx \tag{30.4}$$

This tells us that by plotting the Simmons-Balluffi-derived X_v vs. temperature data on an Arrhenius plot of $\ln X_v$ vs. inverse temperature, as shown schematically in Figure 30.4, we can extract both the entropy of vacancy formation (from the y-axis intercept) and the enthalpy of formation (from the slope).

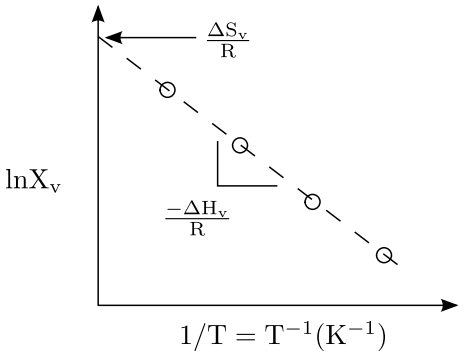


Figure 30.4: Schematic Arrhenius plot of Simmons-Balluffi data for an fcc metal.

In the original work on aluminum, the values obtained for ΔS_v and ΔH_v were $2.2R$ and $72.4kJ/mole$ (or 0.75 eV), respectively.

30.2 Some Important Numbers and a Useful Relationship Regarding Vacancy Formation in fcc Metals

Now that we have the ability to establish equilibrium vacancy concentrations in metals, we can derive useful relationships, for example between the enthalpies of vacancy formation and the melting points of fcc metals. It turns out experimentally that the equilibrium vacancy concentrations at the melting points of fcc metals are almost always in the range, $10^{-4} \lesssim X_V(T_m) \lesssim 10^{-3}$. Let's plug this range of values into equation 29.25 and assume that $\Delta S_v \approx 2.3R$ or $\exp(2.3R/R) \approx 10$. This yields equation 30.5:

$$10^{-4} - 10^{-3} \approx 10 \exp\left(\frac{-\Delta H_v}{RT_m}\right) \tag{30.5}$$

from which it follows that $9.2RT_m \lesssim \Delta H_v \lesssim 11.5RT_m$, or approximately equation 30.6:

$$\Delta H_v(fcc) \approx 10RT_m \tag{30.6}$$

This rough estimate can come in quite handy, for example when you know the melting point of an fcc metal, but lack a readily available measured value for the enthalpy of vacancy formation. In addition, knowing that vacancy concentrations of fcc metals are on the order of $10^{-4}$to 10^{-3}at their melting points is also good to put to memory.

30.3 The Classic Bauerle and Kohler Experiment

So far we have good evidence for the existence of point defects like vacancies, but is there evidence for their motion? We can certainly argue "yes" on the basis of tracer diffusion experiments. However, this question was answered in another way by Bauerle and Kohler [14]. This experiment is a bit more difficult to describe, but following the development will pay dividends. Since the experiment involved the use of electrical resistivity measurements, we need to introduce what is referred to as Matthiessen's rule, given by equation 30.7:

$$\rho = \rho_t + \rho_v + \rho_d \tag{30.7}$$

where the total resistivity of a metal is shown to be governed by several electron-scattering mechanisms, including thermal vibrations (ρ_t), scattering by vacancies (ρ_v), and scattering by other defects like dislocations (ρ_d). Since the resistivity of metals tends to be quite low, Bauerle and Kohler employed wires (to enhance the measurable resistance, given by $\rho l/A$, where l is wire length and A is cross-sectional area). The experimental campaign is represented schematically in Figure 30.5.

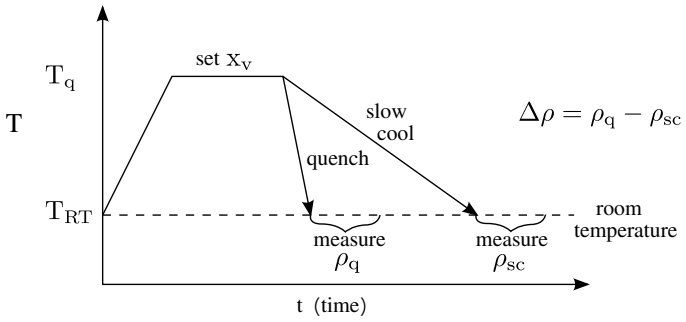


Figure 30.5: Schematic of the Bauerle and Kohler experimental campaign.

A wire of fcc metal was heated to what we will refer to as a "quench temperature" (T_q) in order to set the equilibrium vacancy concentration for that temperature. In two separate experiments, the wire was heated and held at T_q and then quenched to room temperature, freezing in the equilibrium vacancy concentration of the quench temperature, or $X_v(T_q)$. After electrical measurement at room temperature, the same wire was heated and held at T_q and then slow-cooled to room temperature. This second process allows for vacancy annihilation to take place, such that the resulting vacancy concentration is negligibly small. By subtracting the wire's resistivity slow-cooled (ρ_{sc}) from that quenched (ρ_q), we find equation :

$$\Delta\rho = \rho_q - \rho_{sc} = (\rho_t + \rho_v + \rho_d) - (\rho_t + \rho_d) = \rho_v = \alpha X_v(T_q) \tag{30.8}$$

We are assuming that there are no changes in the other defects (including dislocations) during quenching vs. slow cooling, and that the thermal contribution to resistivity will be the same, since both measurements are made at room temperature. The thermal and other defect contributions therefore cancel, leaving only the vacancy scattering contribution, which we can set equal to a constant (α) times the vacancy concentration frozen-in from the quench temperature [$X_v(T_q)$], as in equation 30.8. This gives us something proportional to the equilibrium vacancy concentration at only one quench temperature, so Bauerle and Kohler repeated the procedure in Figure at several other quench temperatures. They then plotted the natural logarithm of $\Delta\rho(T_q)$ vs. the inverse of quench temperature in Arrhenius fashion, as shown schematically in Figure 30.6. Since α is a constant that does not change with quench temperature, the slope of the plot should reflect only the activation energy for vacancy formation, ΔH_v, as shown in the Figure.

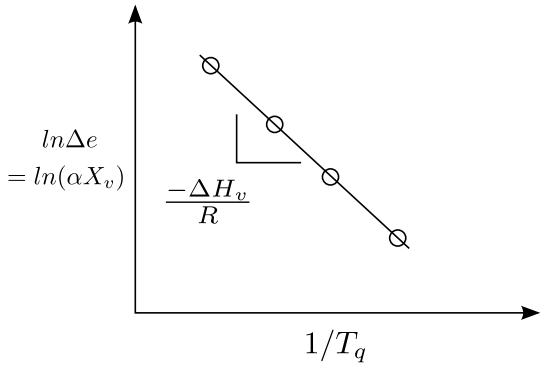


Figure 30.6: Schematic Arrhenius plot of quench-resistivity data from Bauerle and Kohler.

The result obtained for fcc gold by Bauerle and Kohler was 94.6$kJ/mole$ or 0.98 eV, in excellent agreement with later work of Simmons and Balluffi by dilatometry/XRD (90.7$kJ/mole$ or 0.94 eV). We should stress that since the

quench-resistivity method only yields a quantity that is proportional to the vacancy concentration as opposed to the vacancy concentration itself (as in the dilatometry/XRD work), it is impossible to extract information regarding the pre-exponential factor and underlying entropic term by quench-resistivity method.
The real advantage of the Bauerle-Kohler method, however, is that in addition to the enthalpy of vacancy formation, the enthalpy of vacancy motion can also be determined. To appreciate how this is done, we have to resurrect an old equation for the relaxation of vacancies, namely equation 28.8 repeated here:

$$\frac{\bar{c}}{c_o} = \frac{8}{\pi^2}\exp(-\frac{t}{\tau}) \tag{30.9}$$

You will recall this as the "diffusion out of a slab" equation that we also employed to describe vacancy relaxation to sinks such as dislocations or grain boundaries. It is valid when the average composition relaxes to below 80% of the initial composition. In the present situation (quench-resistivity measurements) we can modify this equation to equation 30.10:

$$\bar{X_v} = X_{vo}\frac{8}{\pi^2}\exp(-\frac{t}{\tau}) \tag{30.10}$$

where $\bar{X_v}$is the average vacancy concentration and X_{vo} is the quenched-in starting vacancy concentration. This, in turn, can be related to the quench-resistivity measurements by equation 30.11:

$$\Delta\rho = \alpha X_{vo}\frac{8}{\pi^2}\exp(-\frac{t}{\tau}) \tag{30.11}$$

having inserted the relationship from equation 30.8, namely that $\Delta\rho = \alpha\bar{X_v}$.
Here is where the process of obtaining the motion enthalpy of vacancies gets a little complicated, so stay with me. Figure 30.7 shows a schematic of the temperature-time history of the series of experiments.

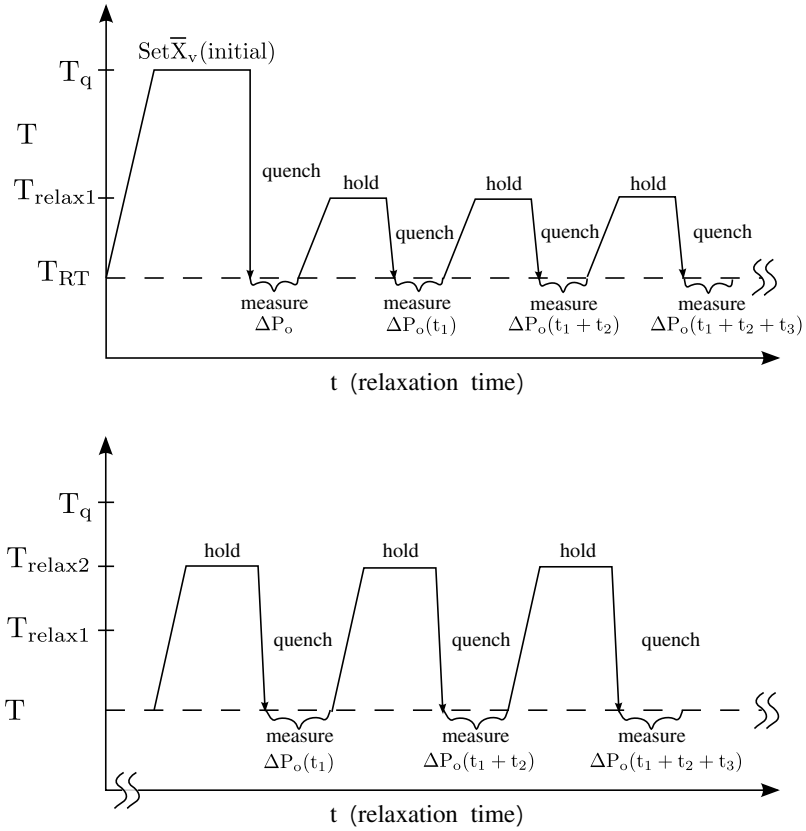


Figure 30.7: Schematic of the Bauerle-Kohler process for obtaining the enthalpy of vacancy motion.

Bauerle and Kohler began by setting a high initial concentration of quenched-in vacancies, by quenching from T_q to room temperature, where the initial value of $\Delta\rho_o$was measured. Remember that $\Delta\rho$ is always the as-quenched resistivity minus the slow-cooled resistivity. They then proceeded to take the wire to a relatively low intermediate

temperature, T_{relax1}, high enough to facilitate vacancy relaxation to sinks, but not so high that a significant fraction of the residual vacancies are lost. This involved a series of sub-experiments: 1) heat at T_{relax1} for a set time, 2) quench to room temperature, and 3) measure $\Delta\rho(t_1)$, and then repeat. As shown in Figure 30.7, the measured $\Delta\rho_t$ is for the cumulative time, e.g., $t = t_1 + t_2 + t_3 + etc.$, at the relaxation temperature. They then went to a higher relaxation temperature, T_{relax2}, and repeated the same series of heat-quench-measure procedures. By taking the natural logarithm of both sides of equation 30.11, noting that α, X_{vo}, and $8/\pi^2$ are all constants, we obtain equation 30.12:

$$\ln \Delta\rho = \ln(const) - \frac{t}{\tau} \tag{30.12}$$

We can normalize each $\Delta\rho$ by the initial $\Delta\rho_o$, and plot the natural logarithm of $\Delta\rho/\Delta\rho_o$ vs. cumulative relaxation time, as shown schematically in Figure 30.8.

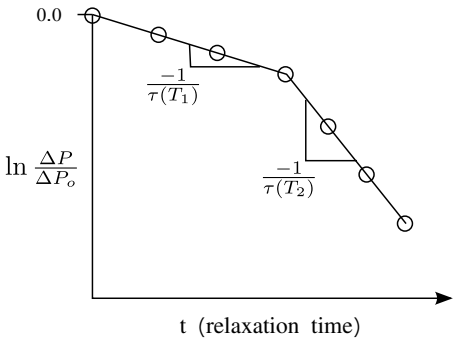


Figure 30.8: Schematic of Bauerle-Kohler vacancy relaxation at two temperatures.

The plot starts at time zero, where $\Delta\rho = \Delta\rho_o$and $\ln(\Delta\rho/\Delta\rho_o)$ vs. time follows a straight line of slope, $-1/\tau(T_1)$, up to the point where the change was made to relax at the higher temperature,T_{relax2} , after which the slope increases, becoming $-1/\tau(T_2)$. You might be asking, "Are we there yet?" And the answer is, "Almost!" Remember the definition of the relaxation time in equation 28.4, repeated here in terms of vacancies:

$$\tau = \frac{l^2}{\pi^2 D_v} \tag{30.13}$$

where l is the spacing between sinks, which we assume is not changing throughout thermal history, and $D_v$is the vacancy diffusivity. Using our knowledge of vacancy diffusivity, we can write the following equation 30.14:

$$D_v = D_{vo} \exp\left(\frac{-\Delta H_{mv}}{RT}\right) \tag{30.14}$$

Since l and $\pi^2$are constants, we can relate the relaxation times at the two temperatures in Figure 30.8 to the corresponding vacancy diffusivities, and the enthalpy of motion in equation 30.15:

$$\frac{\tau(T_{relax1})}{\tau(T_{relax2})} = \frac{D_v(T_{relax2})}{D_v(T_{relax1})} = \exp\left[\frac{\Delta H_{mv}}{R}\left(\frac{1}{T_{relax2}} - \frac{1}{T_{relax1}}\right)\right] \tag{30.15}$$

Using this procedure Bauerle and Kohler found the enthalpy of vacancy motion in solid gold to be 79.1 kJ/mole or 0.82 eV. Adding this to their enthalpy of vacancy formation (94.6 kJ/mole or 0.98 eV), we obtain a value of 173.7 kJ/mole or 1.80 eV. This happens to be virtually identical to the activation energy for gold self-diffusion (by tracer method), which is a nice double-check.

30.4 Some Useful Relationships Regarding Self-Diffusion and Vacancy Motion in fcc Metals

Another highly useful relationship has been found for fcc metals, as expressed in equation 30.16:

$$Q_{SD}(fcc) \approx 18RT_m \tag{30.16}$$

where the activation energy for self-diffusion (typically measured by tracer diffusion) is approximately 18 times the product of R and the melting point (in degrees Kelvin). Since $Q_{SD} = \Delta H_v + \Delta H_{mv}$, subtracting equation 30.6 from equation 30.16 gives the approximation of equation 30.17:

$$\Delta H_{mv}(fcc) \approx 8RT_m \tag{30.17}$$

For example, the melting point of gold is 961.8^oC or $1234.8K$, for which we would predict a vacancy motion enthalpy of 82.1 kJ/mole or 0.85 eV, in good agreement with the Bauerle and Kohler value (79.1 kJ/mole or 0.82 eV).

It turns out that the normalized self-diffusion activation energy, Q_{SD}/RT_m, is roughly constant for each type of crystal structure (including ceramics in addition to metals), however there is a different normalization factor for each structure type. It would pay to do a literature search for the appropriate relationship when exploring a solid with a crystal structure new to you.

This is also a good chance to introduce the concept of "homologous temperature," which is defined as the fraction of an element's melting temperature on an absolute temperature scale. You will encounter this a lot when dealing with kinetic phenomena in materials. For example, the two "intermediate" temperatures employed for vacancy relaxation in quenched gold, represented schematically as $T_1$and T_2in Figures 30.7 and 30.8 were 40^oC and 60^oC, or $313K$ and $333K$ on the absolute temperature scale. These temperatures are 25.4% and 27.0% of the melting temperature of gold. Were we to make similar measurements on a different fcc metal, a good guess for relaxation temperatures would be to start with the same homologous temperatures of $0.254 \lesssim (T/T_m) \lesssim 0.270$.

One last useful number to keep in mind is the melting point diffusivity of solid fcc metals. From the fact we were just given, namely that $Q_{SD}/RT_m \approx 18$, we can exponentiate as in equation 30.18:

$$\exp\left(\frac{-Q_{SD}}{RT_m}\right) \approx \exp(-18) = 1.8x10^{-8} \tag{30.18}$$

You may recall that we employed a range of pre-exponential values for fcc metals, namely $0.018cm^2/s \lesssim D_{Ao} \lesssim 1.8\,\mathrm{cm}^2/\mathrm{s}$. If we arbitrarily choose a value of $1.0\,\mathrm{cm}^2/\mathrm{s}$, this means that the melting point diffusivity of fcc metals should be on the order of $10^{-8}\,\mathrm{cm}^2/\mathrm{s}$, which is in fact quite common experimentally. This number, $D_A(T_m, fcc) \approx 10^{-8}\,\mathrm{cm}^2/\mathrm{s}$, is therefore a good one to put to memory.

31 Point Defects and Transport in Ionic Solids (Ceramics)

Many properties/functions of ionic solids/ceramics are governed by their point defect structures, including ionic diffusivity and ionic charge transport/conductivity. However, in order to understand defect-related structure-property relationships we need first to master the notation now near universally employed to describe point defects, their formation reactions, and their point defect equilibria–otherwise known as "point defect chemistry." The particular notation employed is referred to as "Kröger-Vink" notation, which we introduce in the following section.

31.1 Kröger-Vink Notation

A helpful shorthand scheme for Kröger-Vink (K-V) notation is given in Figure 31.1. An easy way to recall this shorthand scheme is by remembering it as the M, S and C in "Materials SCience." The "M" stands for matter, or the lack thereof (for example a vacancy). The "S" stands for the site on which the defect sits. And the "C" stands for charge. However, in K-V notation the "C" is not real charge (as on the individual ions), but rather effective charge, namely the charge of a defect species relative to the perfect crystal. How these all work can best be illustrated by considering a number of examples.

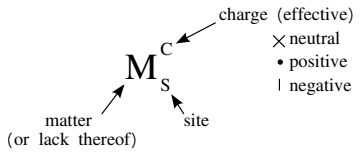


Figure 31.1: Schematic representation of Kröger-Vink notation.

Let's begin by considering host species and fully-charged point defects in zirconia, ZrO_2. Before we start, it is useful to consider the real charges on the ions with which we are dealing. We can write the formula unit of ZrO_2 as $Zr^{4+}O_2^{2-}$. Hence, we can write the host cation species as $(Zr^{4+}_{Zr^{4+}})^x$, where the superscript "x" represents a neutral species insofar as effective charge is concerned. That is, the perfect (non-defective) crystal is charge-neutral. What we have written is that a Zr^{4+} cation on a host Zr^{4+}site has an effective charge of zero (represented by the superscript "x" in K-V notation. Similarly, we can write the host anion species as $(O^{2-}_{O^{2-}})^x$, representing an O^{2-} host anion on an O^{2-} site, with neutral effective charge. For a while we will continue to write host and defect species with real charge inside the parentheses, however this practice not part of the K-V system, but only a temporary "crutch" to help get us comfortable with the notation scheme. In K-V notation, the host species we have thus far considered would be, Zr^x_{Zr} and O^x_O.

But what about point defects? Let's first consider oxygen defects, both vacancies and interstitials. Oxygen vacancies could be written as, $(V^{0-}_{O^{2-}})^{\bullet\bullet}$, representing a vacancy with zero real charge on an oxygen 2- site. The effective charge is thus +2, represented by the double-dots. Another way to think about this is to begin with a charge-neutral occupied oxygen site and remove from it (and the perfect crystal) an O^{2-}species. Since the crystal began electrically neutral, what remains must be doubly-charged positive. How about oxygen interstitials? Remember that, as opposed to close-packed metals where self-interstitials are energetically unfavorable, self-interstitials of cations or anions in ionic crystals are quite possible so long as the ionic radius of the interstitial species is close to that of the interstice on which it resides. We can write an oxygen anion interstitial as, $(O^{2-}_{i^{0-}})''$, which represents a doubly-charged oxygen anion on an originally uncharged interstice. The effective charge is doubly-negative, as represented by the double-strike marks. Another way to think about this is to insert a doubly-charged negative ion in an empty interstice in an otherwise neutral crystal. The effective charge should be doubly-negative. In K-V notation, these two species would be represented as $V_O^{\bullet\bullet}$and $O_i^{''}$ for vacancy and interstitial, respectively. Are you beginning to get the overall idea?

Let's test out our understanding by considering cation defects. A cation vacancy would be $(V^{0+}_{Zr^{4+}})''''$ or $V^{''''}_{Zr}$ in K-V notation. Again, you can think of this as removing a a Zr^{4+}cation from an otherwise neutral crystal. What is left behind must have a quadruply-charged positive effective charge. Note that instead of a superscript of (4'), K-V writes the effective charge in terms of quadruple-strike marks. Similarly, a cation interstitial would be $(Zr^{4+}_{i^{0+}})$ or $Zr_i^{\bullet\bullet\bullet\bullet}$ in K-V-notation. Again, you can think of this as inserting a Zr^{4+} cation in an empty interstice in an otherwise neutral crystal.

Now let's consider donor and acceptor doping. For example, let's put a 5-valent tungsten cation on a Zr^{4+}site. We could represent this as $(W^{5+}_{Zr^{4+}})^{\bullet}$ or $W^{\bullet}_{Zr}$ in K-V notation. The fact that this is a donor is obvious from its positive effective charge. If we compensated for this doping by electronic species, we would have to chose electrons over electron holes. We will consider their K-V representation shortly. Now, consider putting a 2-valent calcium ion on a Zr^{4+} site. We could represent this as $(Ca^{2+}_{Zr^{4+}})^{''}$ or $Ca^{''}_{Zr}$ in K-V notation. The fact that this is an acceptor is obvious from its negative effective charge. If we compensated by electronic species, we would have to chose electron holes over electrons.

Finally, how do we represent those electronic species–electrons and holes. Electrons have the following notation in K-V: $e^{'}$. And electron holes have the following notation in K-V: $h^{\bullet}$. Note that in neither symbol is there a subscript. This is because electrons and holes do not belong to specific sites, but rather to the crystal at large.

So we are now equipped to begin writing balanced point defect reaction equations and mass-action relationships. But there is one caveat that must be made. In most cases the concentration of a specific species is indicated by putting a set of brackets around the particular species. For example, the concentration of oxygen vacancies and zirconium cation interstitials would be represented as $[V_O^{\bullet\bullet}]$ and $[Zr_i^{\bullet\bullet\bullet\bullet}]$, respectively. However, there is a special practice for the electronic species, namely that the concentrations of electrons and holes are represented as n and p, respectively, rather than $[e^{'}]$ and $[h^{\bullet}]$. ·But this is easy to remember, since you are probably aware of n-type vs. p-type behavior in donor-doped vs. acceptor-doped semiconductors.

31.2 Rules for Balancing Point Defect Reactions

The Kröger-Vink rules for writing balanced point defect reactions can also be made to play off the M, S and C in "Materials SCience, as shown in Figure KV2. Here "M" refers to mass-balance, namely that the mass on the left of a reaction must equal the mass on the right. The "C" stands for charge-balance, meaning that the charge on the left side of a reaction must equal the charge on the right. It should be stressed that the charge need not be zero for this to hold; both sides may have net effective charge, as long as that charge is the same both before and after the reaction proceeds. Finally, the "S" stands for site ratio and not site-balance. What this means is that sites need

not be balanced. In fact, we can create sites with our reaction, as long as we preserve the site ratio of the host. For example, if we create two cation sites in a reaction involving Al_2O_3 as the host, as long as we simultaneously create three oxygen sites, preserving the 2:3 ratio of cation:anion sites, we will satisfy this rule. It is best to learn and apply these rules by writing some actual balanced point defect reactions, which we do in the following sections.

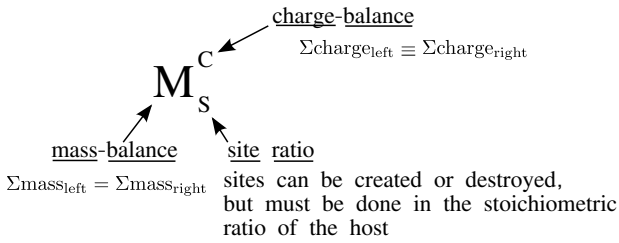


Figure 31.2: Schematic representation of the "M_S^C" rules for writing balanced Kröger-Vink reactions.

31.2.1 Stoichiometric Point Defect Reactions

In what follows three important distinctions must be made: 1) between "intrinsic" and "extrinsic" defects, 2) between stoichiometric and non-stoichiometric point defect reactions, and 3) between homogeneous and inhomogeneous point defect reactions. By "intrinsic" we mean that the reactions involve only species that are part of the undoped host crystal. You can think of the defects as being generated from "within." On the other hand, "extrinsic" defects come from "without," for example dopant species. The following two sections on stoichiometric and non-stoichiometric point defect reactions involve "intrinsic" point defects exclusively. In the third section, we consider doping and are therefore dealing with "extrinsic" point defects. Stoichiometric point defect reactions preserve the stoichiometry of the host, namely the O:M ratio in an oxide, M_xO_y, will be preserved at fixed y/x throughout the reaction. In non-stoichiometric point defect reactions (dealt with in the following section), the O:M ratio will be shifted from y/x(stoichiometric). The third distinction involves homogeneous vs. inhomogeneous point defect reactions. Homogeneous point defect reactions can take place anywhere throughout the crystal, and do not require surfaces to act as a source/sink for defects. No point defect gradients are generated between surface and bulk. On the other hand, inhomogeneous point defect reactions require a surface to act as a source/sink. Furthermore, as they occur a gradient of point defects is generated between the surface and the bulk. They also tend to be relatively slow vis-a-vis homogeneous reactions, which can take place at any point in the crystal, and no migration to/from a surface source/sink is required. This may seem confusing right now, but the distinction will be made clear as we consider specific cases.

The first stoichiometric reaction involves the formation of vacancy/interstitial pairs in what are known as Frenkel point defect reactions. These can be either cation Frenkel or anion Frenkel pairs. For sake of simplicity, we will consider all the point defect reactions of the present section as taking place in the M-monoxide host, MO or $M^{2+}O^{2-}$. I strongly suggest putting the host above the double arrows representing the reaction equilibrium in every case to remind yourself of the host you are dealing with. The first reaction involves cation Frenkel pair formation, as per equation 31.1:

$$V_i^x + M_M^x \overset{MO}{\rightleftarrows} M_i^{\bullet\bullet} + V_M'' \tag{31.1}$$

The species within the parenthesis stands for an unoccupied interstitial site. Since the interstitial population is usually quite small in ionic solids, the site fraction of unoccupied interstices is usually considered to be unity, and this part of the equation is usually not written, since it is understood to exist as in the revised equation 31.2:

$$M_M^x \overset{MO}{\rightleftarrows} M_i^{\bullet\bullet} + V_M'' \tag{31.2}$$

This is the form of the reaction that you will see in textbooks. Let's do two things with this reaction:

1. Confirm that it satisfies all the requirements for a balanced K-V reaction as per the M_S^C rules in Figure 31.2:

   (a) **Mass Balance**: As we consider reaction 31.2, we can see one M species on each side of the equation; this makes it mass-balanced.

(b) **Charge Balance**: The left side of the reaction has zero effective charge and the right side has two positive effective charges and two negative effective charges, which sum to zero; the equation is therefore charge-balanced.

(c) **Site Ratio**: Since there is one cation site on the left and one cation site on the right, there is no change in site ratio; the reaction satisfies the site-ratio rule. You might question the appearance of the interstice on the right, whereas there is no interstice on the left. However, site ratio pertains strictly to the normal host sites, and the availability of unoccupied interstices (as expressed in the equation) is understood. So this is a balanced K-V point defect reaction.

2. Consider whether it is homogeneous or non-homogeneous.

   As shown in Figure 31.3, the reaction simply involves a metal cation jumping from a normal host site into an ever-present empty interstice. This can occur everywhere homogeneously throughout the crystal, as shown, and does not require the presence of a surface to serve as source/sink. As a result, there will be no gradient of defect population between the surface and bulk of the crystal.

The corresponding oxygen interstitial reaction can be written as equation 31.3:

$$O_O^x \overset{MO}{\rightleftarrows} O_i'' + V_O^{\bullet\bullet} \tag{31.3}$$

Once again, we can test this equation against the M_S^C rules. There is one oxygen species on either side, so mass-balance is obeyed. There is zero effective charge on either side, so charge-balance in achieved. Finally, there is one anion site on either side, so site-balance is maintained. This is once again a balanced K-V point defect reaction. Of course, the empty interstitial on the left side of the reaction is understood. As with the cation Frenkel reaction, the anion Frenkel reaction can occur homogeneously throughout the crystal, as represented schematically in Figure 31.3. So oxygen Frenkel pair formation is a homogeneous point defect reaction. One final very important point should be made. Whereas the cation Frenkel reaction will vary with host crystal stoichiometry, namely the aluminum interstitial in Al_2O_3will be trivalent ($Al_i^{\bullet\bullet\bullet}$) as will the vacancy left behind ($V_{Al}^{'''}$), the anion Frenkel reaction is universal to all oxides. In other words, the identical reaction as in equation 31.3 can be written for any oxide, including Al_2O_3.

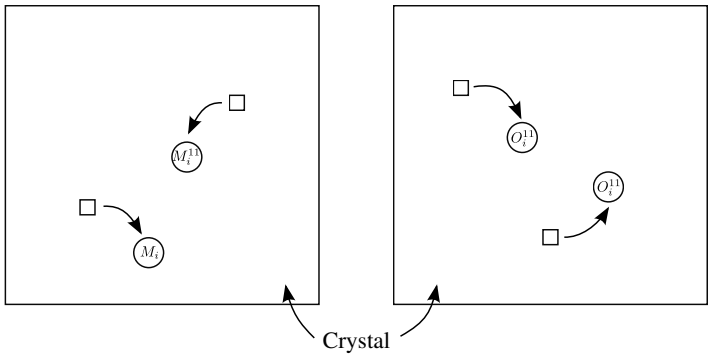


Figure 31.3: Schematic of the homogeneous nature of Frenkel reactions, whether cationic (left) or anionic (right).

The second stoichiometric point defect reaction is the formation of Schottky defects. This will provide an important example of an inhomogeneous point defect reaction. The balanced point defect reaction is given by equation 31.4:

$$null \overset{MO}{\rightleftarrows} V_M'' + V_O^{\bullet\bullet} \tag{31.4}$$

At first glance this seems like an incomplete reaction. There is no mass present, and what the heck is "null?" Well, believe it or not this is, in fact, a balanced K-V point defect reaction. There is no mass on either side, so mass balance is maintained. Both sides are charge-neutral (the 2+ and 2- on the right side cancel), so charge balance is preserved. But what about the fact that there are no sites on the left, whereas there are two on the right? The site-ratio rule does not prohibit the formation or annihilation of sites, but rather specifies that if formed or annihilated, this must be done in the stoichiometric ratio of the host. In this case we create anion and cation sites in the requisite 1:1 ratio of the host MO crystal (O:M=1). The "null" simply represents the perfect, non-defective crystal prior to the formation of the Schottky pair.

But now let's consider what makes this an inhomogeneous point defect reaction. Consider the schematic of a surface of MO crystal in Figure 31.4, where the subscript "S" represents a surface species.

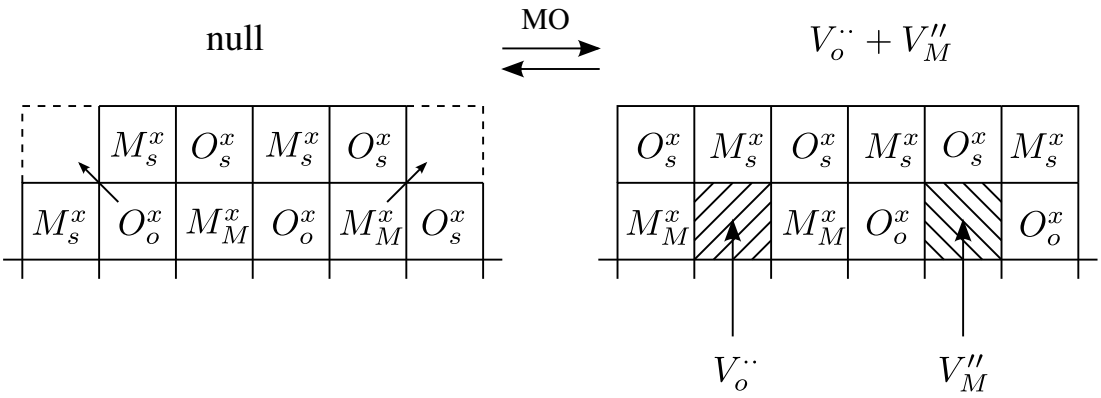


Figure 31.4: Schematic of Schottky pair formation at the surface of the oxide MO.

Let's move a bulk oxygen species to the surface, as shown, making it a surface anion and turning the underlying M surface species into a bulk cation. Next move a bulk metal species to the surface, as shown, making it a surface cation and turning the underlying O surface species into a bulk O species. By these two sub-steps, two vacancies are formed, one cationic and one anionic. The net point defect reaction can be written as equation :

$$M_S^x + O_S^x + M_M^x + O_O^x \overset{MO}{\rightleftarrows} M_S^x + O_S^x + M_M^x + O_O^x + V_M'' + V_O^{\bullet\bullet} \tag{31.5}$$

It can be seen that all species involving mass (an M cation or an O anion) cancel, leaving the much simpler equation 31.4. However, the exercise has been an important one, since pairs of vacancies cannot arbitrarily form in the bulk of a crystal. Instead, Schottky pairs must be formed at surfaces and then migrate into the interior of the crystal, as shown in Figure 31.5, thereby creating a point defect concentration gradient, hence the inhomogeneous nature of Schottky defect formation.

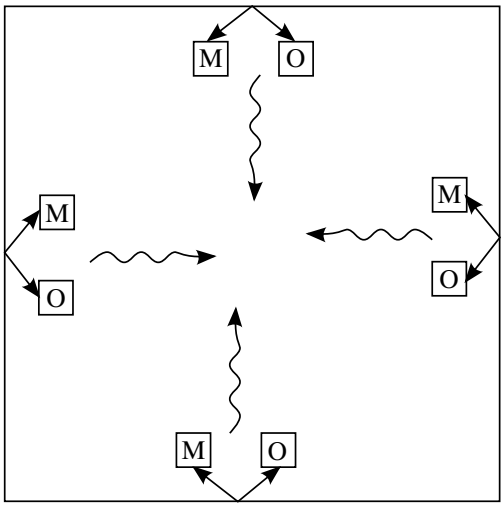


Figure 31.5: Schematic showing the inhomogeneous nature of Schottky defect formation.

Two additional caveats must be made, however. Grain boundaries and dislocation cores can also serve as internal surfaces in ionic solids, where Schottky reactions can take place in either forward direction (they act as sources) or reverse direction (they act as sinks). The second caveat is to be careful to refer to "Schottky pairs" only when the O:M ratio is 1:1, as in MO. In Al_2O_3, for example, the Schottky formation reaction would produce five point defects, as in equation 31.6:

$$null \overset{Al_2O_3}{\rightleftarrows} 2V_M''' + 3V_O^{\bullet\bullet} \tag{31.6}$$

I will leave it to you to check and see whether or not this equation satisfies the M_S^C rules.
In addition to Frenkel and Schottky defect reactions, there are two additional "stoichiometric" point defect reactions. The first involves site-exchange. For example, in the spinel $MgAl_2O_4$ there are two different coordination environments for cations. The Mg cations reside on tetrahedral sites, meaning that they are bonded to four oxygen anions. The Al cations reside on octahedral sites, meaning that they are bonded to six oxygen anions. However, at elevated temperature the two cations can switch sites, as represented by the following equation 31.7:

$$Mg_{Mg}^{x} + Al_{Al}^{x} \overset{MgAl_2O_4}{\rightleftarrows} Mg_{Al}^{'} + Al_{Mg}^{\bullet} \tag{31.7}$$

As with Frenkel reactions, this exchange need not take place only at surfaces, since the two cations are only changing places. Hence, such cation exchange reactions are homogeneous in nature. Once again, we can test whether or not this equation is balanced in terms of the M_S^C rules. There is one Mg species and one Al species on each side, so the equation is mass-balanced. There is one tetrahedral site and one octahedral site on each side, so the equation preserves site balance. In this case, the site ratio is O:Al:Mg equal to 4:2:1. Finally, both sides are charge-neutral. This is a balanced K-V reaction.
The final homogeneous point defect reaction involves electrons and electron holes. Figure 31.6 shows a schematic representation of the top of the filled valence band (CBM or conduction band maximum) and bottom of the unfilled conduction band (CBM or conduction band minimum) of a semiconducting ceramic, and the thermal promotion of an electron from the top of the VBM to the bottom of the CBM.

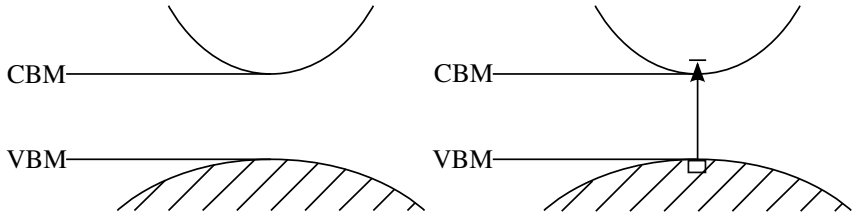


Figure 31.6: Schematic of the valence band maximum (VBM) and conduction band minimum (CBM) and electron promotion across the band gap of a semiconducting oxide.

This process in Kröger-Vink notation is given in the intrinsic electronic reaction of equation 31.8:

$$null \overset{MO}{\rightleftarrows} e^{'} + h^{\bullet} \tag{31.8}$$

So we have considered four different cases of stoichiometric point defect reactions, namely those that result in no change of the overall stoichiometry of the host. The ratio of O:M is maintained by each such reaction.

31.2.2 Non-stoichiometric Point Defect Reactions

We can divide non-stoichiometric point defect reactions into two categories, namely 1) those that decrease the O:M ratio, otherwise known as reduction reactions, and 2) those that increase the O:M ratio, otherwise known as oxidation reactions. As we will show, the former always produce electrons for charge-compensation and are n-type in nature, whereas the latter always produce electron hole for charge-compensation and are p-type in nature. Let's first consider reduction/n-type point defect reactions. Since oxygen gas either is produced (reduction reactions) or consumed (oxidation reactions) and this can only take place at surfaces, all non-stoichiometric point defect reactions are inhomogeneous.
We can decrease the O:M ratio by either decreasing the oxygen content or by increasing the cation content. Both are accomplished by the removal of oxygen, resulting in oxygen gas being released, plus the simultaneous production of electrons. The oxygen-decreasing reaction, otherwise referred to as an "oxygen deficit" reaction, results in the formation of oxygen vacancies, as per equation 31.9:

$$O_O^x \overset{MO}{\rightleftarrows} \frac{1}{2}O_2(g) + V_O^{\bullet\bullet} + 2e^{'} \tag{31.9}$$

This is probably unnecessary at this point, but to keep you in the practice of checking off the M_S^C rules, it can be seen that 1) there is one oxygen species on each side of the equation, 2) both sides are charge-neutral, and 3) we started with one anion site and end up with one anion site such that site-ratio is preserved. This is a balanced K-V reaction.

We can also decrease the O:M ratio by increasing the cation content. The cation-increasing reaction, otherwise referred to as a "metal-excess" reaction results in the formation of metal interstitials, as per equation 31.10:

$$M_M^x + O_O^x \overset{MO}{\rightleftarrows} \frac{1}{2}O_2(g) + M_i^{\bullet\bullet} + 2e^{'} \tag{31.10}$$

I will leave it to you to check and see whether or not this equation satisfies the M_S^C rules.
We can increase the O:M ratio by either increasing the oxygen content or by decreasing the cation content. Both are accomplished by the addition of oxygen and the simultaneous production of electron holes. The oxygen-increasing reaction, otherwise referred to as an "oxygen excess" reaction, results in the formation of oxygen interstitials, as per equation 31.11:

$$\frac{1}{2}O_2(g) \overset{MO}{\rightleftarrows} O_i^{''} + 2h^{\bullet} \tag{31.11}$$

It should be stressed that this equation is universal in oxides, meaning that the O:M ratio of the host (above the arrow in the equation) does not matter; the same reaction is as true in Li_2O as it is in V_2O_5. The cation-decreasing reaction, known otherwise as a "cation deficit" reaction, results in the formation of cation vacancies, as per equation 31.12:

$$\frac{1}{2}O_2(g) \overset{MO}{\rightleftarrows} O_O^x + V_M^{''} + 2h^{\bullet} \tag{31.12}$$

We must be careful, however, since this reaction is not universal but rather depends upon the O:M ratio of the host. For example, the corresponding cation deficit reaction for Al_2O_3would be given by equation 31.13:

$$\frac{3}{2}O_2(g) \overset{Al_2O_3}{\rightleftarrows} 3O_O^x + 2V_{Al}^{'''} + 6h^{\bullet} \tag{31.13}$$

Note that we are producing both anion sites and cation sites, but in the requisite 3:2 ratio for Al_2O_3.

31.2.3 Aliovalent Doping Reactions

We now consider how the oxide, MO, might be aliovalently-doped. Isovalent doping would be doping MO with oxide NO, where the valence state of the species N is 2+, as it is in the MO host. Aliovalent doping means doping with oxides having cations of different prevailing valence states than the host. We can write reactions for donor-doping, where the doping cation has a higher valence state than M^{2+}, and for acceptor-doping, where the doping cation has a lower valence state than M^{2+}. In each case we can charge-compensate with ionic point defects or with electronic defects. For example, let's consider donor-doping of MO with the trivalent cation N, or N_2O_3. The ionic compensation reaction involves the formation of cation vacancies, as per equation 31.14:

$$N_2O_3 \xrightarrow{MO} 3O_O^x + 2N_M^{\bullet} + V_M^{''} \tag{31.14}$$

This is a good reaction on which to test your K-V balancing skills. The reaction is obviously mass-balanced and charge-balanced, but what about site-ratio? Well, the reaction yields 3 anion sites and 3 cation sites, which are created in precisely the O:M ratio of the host (3:3 equals 1:1).
The alternative to ionic compensation is electronic compensation. Note that in so doing, we need to release some of the oxygen as gas to maintain the M:O ratio of the host. The resulting reaction is given by equation 31.15:

$$N_2O_3 \xrightarrow{MO} 2O_O^x + 2N_M^{\bullet} + 2e^{'} + \frac{1}{2}O_2(g) \tag{31.15}$$

Now let's consider acceptor-doping of MO with the monovalent cation N, or N_2O for which the valence states are $N_2^+O^{2-}$. The ionic compensation reaction involves the formation of oxygen vacancies, as per equation 31.16:

$$N_2O \xrightarrow{MO} O_O^x + 2N_M^{'} + V_o^{\bullet\bullet} \tag{31.16}$$

The electronic compensation reaction requires the addition of some oxygen on the left side of the equation 31.17:

$$N_2O + \frac{1}{2}O_2(g) \xrightarrow{MO} 2O_O^x + 2N_M^{'} + 2h^{\bullet} \tag{31.17}$$

An obvious question to ask is which compensation mechanism (ionic, electronic) will take place in a certain dopant oxide/host oxide combination? The answer is that it depends upon the defect formation energetics, which we consider in the following section. You may also have missed a critical changeover from the previous two sections on stoichiometric and non-stoichiometric reactions to the present one on aliovalent doping reactions. You will notice forward and reverse arrows in the reactions of the prior two sections, but only a forward arrow in the reactions of the present section. This is owing to the fact that the first two sections dealt with equilibrium point defect reactions, meaning that the forward and reverse arrows were balanced. With aliovalent doping of ceramics, this is not the case, unless one has excess dopant oxide present at high temperature, namely a two-phase mixture. Instead, ceramics are donor-doped and acceptor-doped by incorporating a fixed amount of aliovalent dopant in the initial processing of the doped host ceramic. It is expected that the forward reaction goes to completion during processing at elevated temperature. Therefore, there is only a forward arrow. In other words, in equation 31.14 for example, we can set the vacancy concentration equal to one-half the trivalent donor concentration, or $[V_M^{''}] = \frac{1}{2}[N_M^{\bullet}]$. It should be quite clear that all doping reactions, whether isovalent or aliovalent are "extrinsic" reactions, since the dopant species comes from "outside" the host crystal. It should also be clear that all doping reactions are inhomogeneous, reflecting the fact that dopants react with the surfaces of host oxide crystals/particles and undergo what is referred to as "solid state reaction" (and long range diffusion) to achieve doping uniformity throughout the doped ceramic.

31.3 Point Defect Thermodynamics in Ionic Solids/Ceramics

We can approach the thermodynamics of point defect reactions in ionic solids/ceramics in the same way we dealt with vacancy formation in metals. For example, for the formation of Schottky pairs in the oxide MO the reaction is repeated in equation 31.18:

$$null \overset{MO}{\rightleftarrows} V_M^{''} + V_O^{\bullet\bullet} \tag{31.18}$$

The only difference is that we cannot consider the formation of individual cation vacancies or oxygen vacancies. Given their charge, their formation would be highly unfavorable energetically. Instead, we form pairs consisting of a cation vacancy and an anion vacancy to maintain electroneutrality of the crystal, as per equation 31.18. The enthalpy of this reaction can be written as equation 31.19:

$$\Delta H = X_S \Delta H_S \tag{31.19}$$

where X_S is the site fraction of Schottky pairs, namely the site fraction of each kind of vacancy on its sublattice, and ΔH_S is the enthalpy per mole of Schottky pairs. The entropy of Schottky pair formation can be written as in equation 31.20:

$$\Delta S = X_S \Delta S_S - R[(X_M \ln X_M + (1 - X_M)\ln(1 - X_M)) + (X_O \ln X_O + (1 - X_O)\ln(1 - X_O))] \tag{31.20}$$

where ΔS_S represents the non-configurational entropy of Schottky pair formation (per mole of Schottky pairs) and the term inside the brackets represents the configurational entropy. If we plot the two functions in equations 31.19 and 31.20 vs. the site fraction of Schottky pairs, X_S, and recall that $\Delta G = \Delta H - T\Delta S$, we obtain the curves in Figure 31.7.

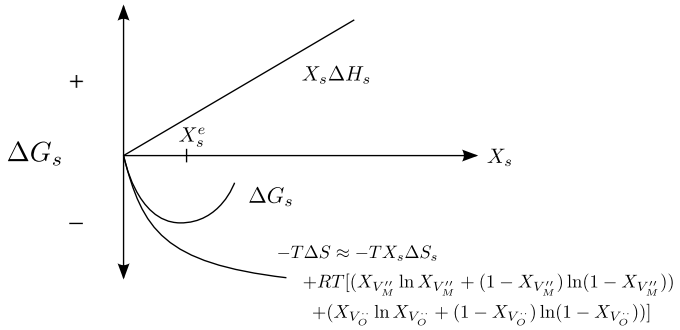


Figure 31.7: The enthalpic and entropic contributions to the free energy of Schottky pair formation.

The behavior closely resembles the same plot for vacancies in metals seen in Figure 29.5. Initially, the entropy term dominates, but with increasing Schottky pair production the enthalpy term begins to dominate. The result is a minimum in the free energy vs. X_S curve corresponding to the equilibrium Schottky pair concentration (X_S^e). We can write a mass-action expression (an equilibrium constant expression) for the Schottky pair formation, as in equation 31.21:

$$K_S^{'} = a_{V_M^{''}} a_{V_O^{\bullet\bullet}} = \gamma[V_M^{''}]\gamma[V_O^{\bullet\bullet}] \tag{31.21}$$

where $a_{V_M^{''}}$ and $a_{V_O^{\bullet\bullet}}$ represent the individual defect activities. We have assumed dilute solution behavior so the two γ factors are Henry's Law constants. These have to be the same for both cation and anion vacancies, such that they are formed in equal populations. By dividing both sides of equation by γ^2 we arrive at equation 31.22:

$$K_S = \frac{K_S^{'}}{\gamma^2} = [V_M^{''}][V_O^{\bullet\bullet}] = \exp\left(\frac{\Delta S_S}{R}\right)\exp\left(\frac{-\Delta H_S}{RT}\right) \tag{31.22}$$

Similarly, we can write an equilibrium expression for oxygen Frenkel formation, repeated as equation 31.23:

$$O_O^x \overset{MO}{\rightleftarrows} O_i^{''} + V_O^{\bullet\bullet} \tag{31.23}$$

for which the equilibrium expression is given by equation 31.24:

$$K_F = \frac{[O_i^{''}][V_O^{\bullet\bullet}]}{[O_O^x]} = [O_i^{''}][V_O^{\bullet\bullet}] = \exp\left(\frac{\Delta S_F}{R}\right)\exp\left(\frac{-\Delta H_F}{RT}\right) \tag{31.24}$$

In this expression, the site fraction of oxygen on oxygen sites is assumed to be unity, since the site fraction of oxygen vacancies is negligibly small. Hence, their activity can be assumed to be unity. We will employ these reactions and equilibrium expressions when generating and interpreting so-called Brouwer diagrams in the following sections.
Similarly, we can write point defect equilibrium expressions for each of the reduction and oxidation reactions introduced previously. For the oxygen-deficit reaction, repeated in equation 31.25:

$$O_O^x \overset{MO}{\rightleftarrows} \frac{1}{2}\mathrm{O_2(g)} + V_O^{\bullet\bullet} + 2e^{'} \tag{31.25}$$

the equilibrium expression would be equation 31.26:

$$K_{red} = p_{O_2}^{1/2}[V_O^{\bullet\bullet}]n^2 \tag{31.26}$$

For the metal-excess reaction, repeated in equation 31.27:

$$M_M^x + O_O^x \overset{MO}{\rightleftarrows} \frac{1}{2}\mathrm{O_2(g)} + M_i^{\bullet\bullet} + 2e^{'} \tag{31.27}$$

the equilibrium expression would be equation 31.28:

$$K_{red} = p_{O_2}^{1/2}[M_i^{\bullet\bullet}]n^2 \tag{31.28}$$

It should be pointed out that both the oxygen-deficit reaction (equation 31.25) and the metal-excess reaction (equation 31.27) have point defect concentrations decreasing as the oxygen partial pressure increases (at constant temperature). This can be seen by applying Le Chatelier's principle, namely increasing oxygen partial pressure causes a response in the direction (arrow) opposite to the side on which the p_{O_2} term occurs. It also follows from the point defect equilibrium expressions of equations 31.26 and 31.28, namely that to maintain the equilibrium constant, increasing P_{O_2} must be accompanied by a decrease in point defect concentration.
Similarly, we can now consider oxidation reactions. For the oxygen-excess reaction, repeated in equation 31.29:

$$\frac{1}{2}\mathrm{O_2(g)} \overset{MO}{\rightleftarrows} O_i^{''} + 2h^{\bullet} \tag{31.29}$$

the equilibrium expression would be equation 31.30:

$$K_{ox} = \frac{[O_i^{''}]p^2}{p_{O_2}^{1/2}} \tag{31.30}$$

And for the metal-deficit reaction, repeated in equation 31.31:

$$\frac{1}{2}O_2(g) \overset{MO}{\rightleftarrows} O_O^x + V_M^{''} + 2h^{\bullet} \tag{31.31}$$

the equilibrium expression would be equation 31.32:

$$K_{ox} = \frac{[V_M^{''}]p^2}{p_{O_2}^{1/2}} \tag{31.32}$$

It should be pointed out that both the oxygen-excess reaction (equation 31.29) and the metal-excess reaction (equation 31.31) have point defect concentrations increasing as the oxygen partial pressure increases (at constant temperature). Again, this can be seen by applying Le Chatelier's principle, namely increasing oxygen partial pressure causes a response in the direction (arrow) opposite to the side on which the p_{O_2} term occurs. It also follows from the point defect equilibrium expressions of equations 31.30 and 31.32, namely that to maintain the equilibrium constant, increasing P_{O_2} must be accompanied by an increase in point defect concentration. The reactions and equilibria of the present section will form the foundation upon which schematic Brouwer diagrams can be constructed and interpreted.
We must consider one last equilibrium expression, namely that between electrons and electron holes as described by equation 31.8. The equilibrium expression for this reaction is :

$$K_i = np = const \exp\left(\frac{-E_g}{RT}\right) \tag{31.33}$$

where E_g is the "band gap," corresponding to the difference in energy between the conduction band minimum and the valence band minimum as shown in Figure 31.6.

31.4 Brouwer Diagrams

In 1954, G. Brouwer introduced what are now referred to as "Brouwer diagrams" (Phillips Res. Rep., 9, 366). Brouwer diagrams are logarithmic plots of defect populations 1) vs. temperature (or rather inverse temperature), 2) vs. oxygen partial pressure, or 3) vs. dopant concentration, again on a logarithmic scale. We consider schematic versions for each of these in the following sections.

31.4.1 Brouwer Diagrams of Defect Concentration vs. Inverse Temperature

This type of diagram is useful for determining the energetics of point defect formation and motion in ionic solids/ceramics. For example, consider the extrinsic doping of NaCl by $CaCl_2$, where Ca is divalent ($Ca^{2+}Cl_2^-$), as given by equation 31.34:

$$CaCl_2 \xrightarrow{NaCl} Ca_{Na}^{\bullet} + 2Cl_{Cl}^x + V_{Na}^{'} \tag{31.34}$$

Note the use of the forward arrow only, connoting that this reaction goes to completion during the initial processing of the Ca-doped NaCl ceramic. Now consider the Schottky pair formation in NaCl, according to equation 31.35:

$$null \overset{NaCl}{\rightleftarrows} V_{Na}^{'} + V_{Cl}^{\bullet} \tag{31.35}$$

for which the equilibrium expression is given by equation 31.36:

$$K_S = [V_{Na}^{'}][V_{Cl}^{\bullet}] = \exp\left(\frac{\Delta S_S}{R}\right)\exp\left(\frac{-\Delta H_S}{RT}\right) \tag{31.36}$$

The next step is to write what is commonly called the "electroneutrality condition," which we will henceforth refer to as the "ENC." This condition is required to maintain the overall electroneutrality of the host crystal. The ENC is a straightforward expression setting the sum of all the negatively-charged point defects equal to the sum of all the positively-charged point defects, as per equation 31.37:

$$[V_{Na}^{'}] = [Ca_{Na}^{\bullet}] + [V_{Cl}^{\bullet}] \tag{31.37}$$

The next step is to make so-called "Brouwer approximations." These are regimes in which two species–one positively charged and one negatively charged–predominate. As we consider equation 31.37, it is clear that there should be two such regimes: 1) one where the Na-vacancy concentration is set by the Ca-donor doping, or $[V_{Na}^{'}] = [Ca_{Na}^{\bullet}]$, and 2) one where the Na-vacancy concentration is determined by the Schottky equilibrium with Cl-vacancies, or $[V_{Na}^{'}] = [V_{Cl}^{\bullet}]$. But which regime will be on the left and which will be on the right on a log [defect] vs. inverse temperature plot? Since temperature increases from right to left on plot with an inverse temperature (1/T) x-axis, it makes sense that the extrinsic, donor-doped regime would dominate at low temperatures, or to the right, whereas intrinsic, Schottky pair defects would dominate at high temperatures, or to the left. The resulting schematic Brouwer diagram is given in Figure T, with the two Brouwer approximations/regimes clearly marked.

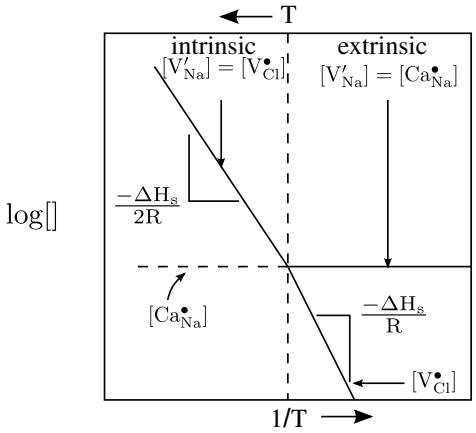


Figure 31.8: Brouwer diagram vs. inverse temperature for $CaCl_2$-doped NaCl.

The slope of the Na-vacancy concentration (important for later) is marked as $-\Delta H_S/2R$. This follows from inserting the ENC for that regime, namely $[V_{Na}^{'}] = [V_{Cl}^{\bullet}]$, into the equilibrium expression of equation 31.36 such that $K_S = [V_{Na}^{'}]^2$. This means that $[V_{Na}^{'}] = K_S^{1/2}$and that the activation energy for this regime will be $-\Delta H_S/2R$. However, the diagram isn't complete until we show a line for each species in both regimes. The donor-dopant, , is easy. It is just a horizontal line extending from the extrinsic regime into the intrinsic regime as shown. The Ca-dopant is a minority species in this regime. So we have lines for each of the three species in the ENC of equation 31.37 in the intrinsic regime. But what about the behavior of $[V_{Cl}^{\bullet}]$ in the extrinsic regime? Well, we know its concentration at the dashed line demarking the transition between electroneutrality regimes. In the extrinsic regime, the concentration of Na-vacancies is fixed by the aliovalent dopant concentration, namely $[V_{Na}^{'}] = [Ca_{Na}^{\bullet}]$ =constant (not a function of temperature). This means from the equilibrium expression of equation that Cl-vacancy concentration will be directly proportional to the Schottky formation constant, or $[V_{Cl}^{\bullet}] = K_S/[V_{Na}^{'}] = K_S/[Ca_{Na}^{\bullet}] = K_S/constant$. Therefore, the slope on the Cl-vacancy concentration in the extrinsic regime will be $-\Delta H_S/R$, as shown.
This Brouwer diagram is especially helpful in determining the formation and motion enthalpies of Na-vacancies in NaCl. A schematic representation of the tracer-diffusion data for Na in NaCl is shown in Figure 31.9.
Based upon the Brouwer diagram of Figure 31.8, we know that there is no energy of formation of Na-vacancies in the extrinsic regime; the slope is attributable solely to the enthalpy of motion, or $-\Delta H_{mv}/R$. However, in the extrinsic regime there are contributions from both the enthalpy of motion and the enthalpy of formation. From the Brouwer diagram of Figure 31.8, we know that the contribution from formation is $-\Delta H_S/2R$, so the combined slope is $-(\Delta H_{mv} + \Delta H_S/2)/R$. Since we know the enthalpy of motion from the slope of the diffusion plot in the extrinsic regime, this means that we can now calculate the enthalpy of Schottky pair formation from the slope of the diffusion plot in the intrinsic regime. This demonstrates the power of Brouwer diagrams for the study of point defect formation and motion in ionic solids/ceramics.

31.4.2 Brouwer Diagrams of Defect Concentration vs. Oxygen Partial Pressure

The second type of Brouwer diagram plots the defect concentrations vs. the oxygen partial pressure on a log-log plot made at constant temperature. The most commonly portrayed diagrams of this type are often referred to as "butterfly" diagrams, owing to their appearance. Although we will not construct such diagrams, we will display a couple of examples and discuss their construction and interpretation. For example, let's consider an oxide MO

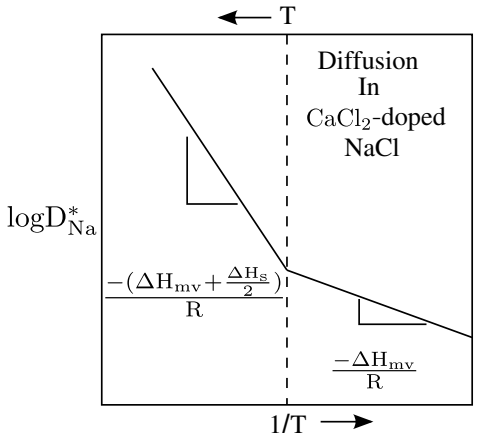


Figure 31.9: Na tracer diffusion in $CaCl_2$-doped NaCl.

that has prevailing Schottky disorder. We have already written the stoichiometric point defect reactions that can occur in such a material, repeated here in equations 31.38 and 31.39:

$$null \overset{MO}{\rightleftarrows} V_M^{''} + V_O^{\bullet\bullet} \tag{31.38}$$

$$null \overset{MO}{\rightleftarrows} e^{'} + h^{\bullet} \tag{31.39}$$

for which the equilibrium expressions are equations 31.40 and 31.41:

$$K_S = [V_M^{''}][V_O^{\bullet\bullet}] = \exp\left(\frac{\Delta S_S}{R}\right)\exp\left(\frac{-\Delta H_S}{RT}\right) \tag{31.40}$$

$$K_i = np = const \exp\left(\frac{-E_g}{RT}\right) \tag{31.41}$$

respectively. In addition, we write the appropriate reduction and oxidation reactions, repeated here in equations 31.42 and 31.43:

$$O_O^x \overset{MO}{\rightleftarrows} \frac{1}{2}O_2(g) + V_O^{\bullet\bullet} + 2e^{'} \tag{31.42}$$

$$\frac{1}{2}O_2(g) \overset{MO}{\rightleftarrows} O_O^x + V_M^{''} + 2h^{\bullet} \tag{31.43}$$

respectively. The equilibrium expressions for these reactions are repeated here as equations 31.44 and 31.45:

$$K_{red} = p_{O_2}^{1/2}[V_O^{\bullet\bullet}]n^2 \tag{31.44}$$

$$K_{ox} = \frac{[V_M^{''}]p^2}{p_{O_2}^{1/2}} \tag{31.45}$$

respectively. Now let's write the ENC for this system, as in equation :

$$n + 2[V_M^{''}] = p + 2[V_O^{\bullet\bullet}] \tag{31.46}$$

At first glance it would seem like the factors of "2" are in front of the wrong species, until we consider that according to the reduction reaction of equation 31.42, two electrons are formed for each oxygen vacancy formed. Hence, the electron population should be twice the concentration of oxygen vacancies. Were these the prevailing species, the Brouwer approximation would be $n = 2[V_O^{\bullet\bullet}]$.
Based upon the overall ENC, it would seem that there should be four Brouwer approximations, namely $n = p$, $n = 2[V_O^{\bullet\bullet}]$, $2[V_M^{''}] = p$, and $2[V_M^{''}] = 2[V_O^{\bullet\bullet}]$, however this is not the case. Instead, the oxide in question will either

have ionic species dominating at stoichiometry, such that the electrons and electron holes are minority species, or electronic species dominating at stoichiometry, such that cation vacancies and anion vacancies are minority species. This makes for two slightly different "butterfly" diagrams, as shown in Figure 31.10. We can understand and interpret such diagrams by considering their point defect equilibria.

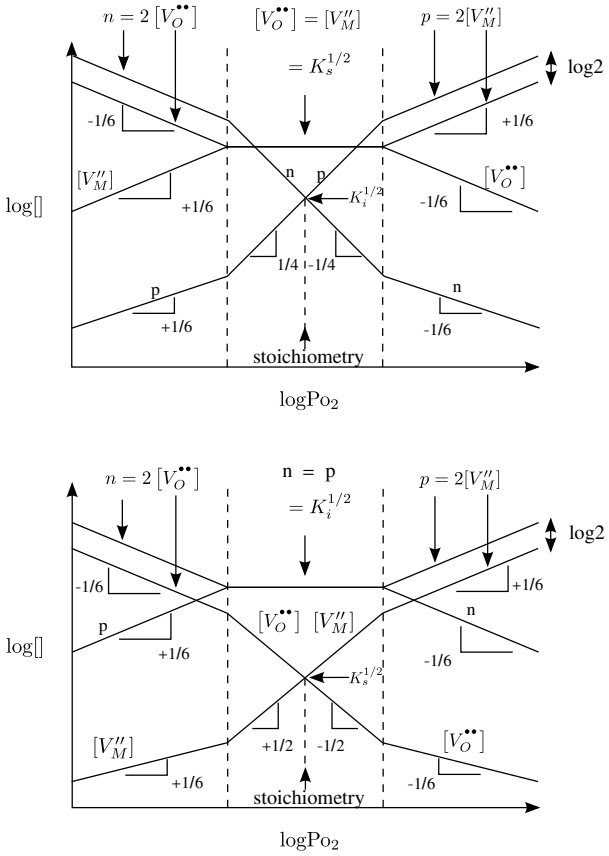


Figure 31.10: Butterfly log[defect] vs. log p_{O_2}Brouwer diagrams for prevailing ionic disorder (top) and for prevailing electronic disorder (bottom.).

The diagram on the top is for ionic species dominating at stoichiometry, such that $[V_M''] = [V_O^{\bullet\bullet}] = K_S^{1/2}$. With increasing oxygen partial pressure, the oxidation reaction takes over, such that the electroneutrality condition transitions to $p = 2[V_M'']$. Inserting this Brouwer approximation into the equilibrium expression of equation 31.43, we obtain equation 31.47:

$$K_{ox} = \frac{[V_M'']p^2}{p_{O_2}^{1/2}} = \frac{4[V_M'']^3}{p_{O_2}^{1/2}} \tag{31.47}$$

Taking the logarithm of this expression and rearranging a bit, we obtain equation 31.48:

$$3\log[V_M''] = \log\frac{K_{ox}}{4} + \frac{1}{2}\log p_{O_2} \tag{31.48}$$

Dividing both sides by 3, it follows that $\partial \log[V_M'']/\partial \log p_{O_2}$ will be +1/6 in this regime, which appears on the right side of the upper diagram in Figure 31.10. Note that the electron hole line lies above the cation vacancy line, owing to the factor of 2 in the Brouwer approximation, $p = 2[V_M'']$. The shift upwards of the electron hole line vs. the cation vacancy line will be precisely the logarithm of 2, as shown.

Now let's consider what happens under reduction, when the reduction reaction takes over, such that the electroneutrality condition transitions to $n = 2[V_O^{\bullet\bullet}]$. Inserting this Brouwer approximation into the equilibrium expression of equation 31.42, we obtain equation 31.49:

$$K_{red} = p_{O_2}^{1/2}[V_O^{\bullet\bullet}]n^2 = p_{O_2}^{1/2}4[V_O^{\bullet\bullet}]^3 \tag{31.49}$$

Again, taking the logarithm of this expression and rearranging a bit, we obtain equation 31.50:

$$3\log[V_O^{\bullet\bullet}] = \log\frac{K_{red}}{4} - \frac{1}{2}\log p_{O_2} \tag{31.50}$$

Dividing both sides by 3, it follows that $\partial\log[V_O^{\bullet\bullet}]/\partial\log p_{O_2}$ will be -1/6 in this regime, which appears on the left side of upper diagram in Figure 31.10. The slopes of the electrons and electron holes in the central regime follows from consideration of oxidation and reduction reactions. For example, reorganizing the reduction reaction of equation 31.49 we obtain equation 31.51:

$$K_{red}/[V_O^{\bullet\bullet}] = p_{O_2}^{1/2}n^2 \tag{31.51}$$

Taking the logarithm of this expression and rearranging a bit, we obtain equation :

$$2\log n = \log\frac{K_{red}}{[V_O^{\bullet\bullet}]} - \frac{1}{2}\log p_{O_2} \tag{31.52}$$

from which it follows that, since $[V_O^{\bullet\bullet}]$ is constant in the middle regime, $\partial\log n/\partial\log p_{O_2}$will be -1/4. Hence the -1/4 slope in the middle regime. For all other species, one can play off the corresponding majority defect to establish individual slopes. For example, from the intrinsic electronic disorder reaction of equation 31.39, for $K_i = np$ to remain constant in a regime where the electron population is found to vary with $p_{O_2}$to the -1/6 power, the electron hole population must vary with p_{O_2} to the +1/6 power. Similarly, from the Schottky pair formation reaction of 31.38, for $K_S = [V_M^{''}][V_O^{\bullet\bullet}]$ to remain constant in a regime where the metal vacancy population is found to vary with p_{O_2} to the +1/6 power, the oxygen vacancy population must vary with p_{O_2} to the -1/6 power.
The second diagram in Figure 31.10 is for the situation where electronic disorder predominates in the stoichiometric (middle) regime, such that $K_i = np$. It can be seen that in this regime the ionic defects (cation vacancies and anion vacancies) are minority species. Note, however, that the Brouwer approximations for the oxidation and reduction regimes are the same as for the upper diagram in Figure 31.10.
There are six such Brouwer diagrams like the two represented in Figure 31.10. This is communicated by Figure 31.11, where the six situations are delineated. There are two each for the three main types of stoichiometric disorder: Schottky, Cation Frenkel, and Anion Frenkel. In each case there will be a diagram where ionic disorder predominates ($K_{ion} >> K_i$, where K_{ion}stands for either the Schottky or Frenkel equilibrium constant and K_i stands for the intrinsic electronic disorder equilibrium constant) and a diagram where electronic disorder predominates ($K_i >> K_{ion}$).

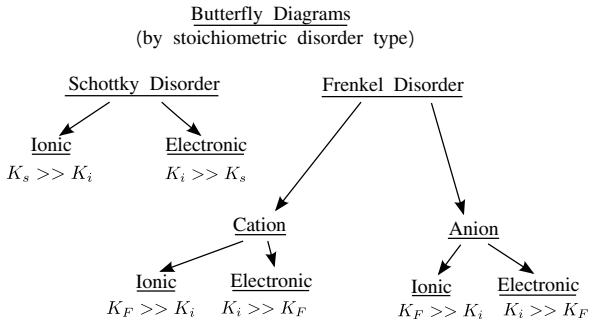


Figure 31.11: Six possible butterfly diagrams for Schottky, Cation Frenkel, and Anion Frenkel disorder.

The power of Brouwer diagrams like those in Figure 31.10 can be demonstrated by their predictive capability. For example, if the MO oxide were a Schottky pair-former, with $K_S >> K_i$ as in the top diagram of Figure 31.10, with cation motion governed by vacancy mechanism, this could be confirmed by tracer diffusion studies of M-cations vs. oxygen partial pressure. According to the top diagram of Figure 31.10, we would predict +1/6 slopes vs. oxygen partial pressure at extreme reducing and extreme oxidizing conditions, with a plateau where the diffusion coefficient should not vary with p_{O_2} in the middle regime. Similar predictions could be made concerning the O-anion diffusivity.
One must be careful in drawing "butterfly" diagrams in oxides for which the O:M stoichiometry is other than 1:1. For example, the oxidation reaction for the oxide M_2O (assumed to be a Schottky former) would be given by equation 31.53:

$$\frac{1}{2}\mathrm{O}_2(\mathrm{g}) \overset{M_2O}{\rightleftarrows} O_O^x + 2V_M^{'} + 2h^{\bullet} \tag{31.53}$$

whereas the reduction reaction would be equation 31.54:

$$O_O^x \overset{M_2O}{\rightleftarrows} \frac{1}{2}\mathrm{O}_2(\mathrm{g}) + V_O^{\bullet\bullet} + 2e^{\prime} \tag{31.54}$$

I'll give you a moment to establish the slopes of the majority species on either side of this "butterfly" plot. They will prove NOT to be the same, which tells us that asymmetrical "butterfly" Brouwer diagrams (warped butterflies) can be anticipated for oxides with stoichiometries other than O:M of 1:1. (Spoiler alert! I am about to tell you the slopes in the "wings" of the M_2O butterfly diagram. If you want to test your Kröger-Vink/Brouwer prowess, look away and see if you can solve for the slopes.)
The slopes on the M_2O butterfly would be -1/6 on the left (this is essentially unchanged from the MO butterfly situation), and +1/8 on the right. This latter slope can be obtained if we write the equilibrium reaction for equation 31.53 as equation 31.55:

$$K_{ox} = \frac{[V_M^{'}]^2 p^2}{p_{O_2}^{1/2}} \tag{31.55}$$

Reminding ourselves that the Brouwer approximation for the ENC in the right "wing" will be $p = [V_m^{'}]$, and taking the logarithm of equation 31.55 and rearranging, we obtain equation 31.56:

$$4\log p = \log K_{ox} + \frac{1}{2}\log p_{O_2} \tag{31.56}$$

It follows that $\partial \log p / \partial \log p_{O_2}$ will be +1/8.
We end this section by pointing out that "butterfly" Brouwer diagrams have limited practical application, owing to two facts. First, few oxides are amphoteric, meaning that they exhibit both n-type and p-type behavior. Most go in one direction (oxidation or reduction, p-type or n-type), but seldom both. Second, few oxides are employed in an undoped intrinsic state. We will discuss additional log[defect] vs. $\log p_{O_2}$ diagrams as we consider two-regime extrinsic/intrinsic Brouwer diagrams in the following section.

31.4.3 Brouwer Diagrams of Defect Concentration vs. Dopant Concentration

A very useful type of Brouwer diagram involves plots of defect populations vs. dopant concentration (at fixed T and p_{O_2}) on a log-log plot. As with the log[defect] vs. inverse temperature plots already discussed, these have two Brouwer approximations, namely two ENC regimes: one intrinsic and one extrinsic. The procedures for creating such Brouwer diagrams are as follows:

1. Write the appropriate intrinsic point defect reaction and the corresponding mass-action (equilibrium constant) equation.
2. Write the extrinsic doping reaction.
3. Write the overall electroneutrality condition, from which the two Brouwer approximations will be made clear.
4. Decide which regime goes where, for example on which side the extrinsic regime should fall.
5. Solve the mass-action (equilibrium constant) equation for all point defect species in each Brouwer regime.

Let's start by considering cation-deficit oxide MO, which is known to be a p-type semiconductor under oxidizing conditions. It is known that it can be acceptor-doped by the oxide, N_2O, where the dopant is monovalent, namely $N_2^{+}O^{2-}$, resulting in electronic compensation by electron holes. The intrinsic point defect reaction is the oxidation reaction to form cation vacancies, as in equation 31.57:

$$\frac{1}{2}\mathrm{O}_2(\mathrm{g}) \overset{MO}{\rightleftarrows} O_O^x + V_M^{''} + 2h^{\bullet} \tag{31.57}$$

for which the mass-action (or equilibrium constant) equation is equation 31.58:

$$K_{ox} = \frac{[V_M^{''}]p^2}{p_{O_2}^{1/2}} \tag{31.58}$$

Next let's write the acceptor-doping reaction, repeated as equation 31.59:

$$N_2O + \frac{1}{2}O_2(g) \xrightarrow{MO} 2O_O^x + 2N_M^{'} + 2h^{\bullet} \tag{31.59}$$

As we discussed previously, the forward arrow (with no reverse arrow) indicates that the reaction goes to completion during the processing of N-doped MO. The overall electroneutrality condition will be equation 31.60:

$$p = 2[V_M^{''}] + [N_M^{'}] \tag{31.60}$$

The next task is to decide on which side of the Brouwer diagram to place the extrinsic regime. This is pretty straightforward in the present case, owing to the fact that with increasing donor-doping, extrinsic behavior should "win out" over intrinsic behavior. We can begin by drawing a dashed line of slope +1 on the Brouwer plot, representing the concentration of acceptor species, or $[N_M^{'}]$, as shown in Figure 31.12.

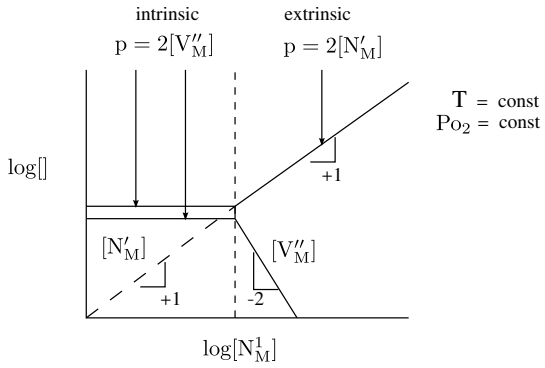


Figure 31.12: Schematic Brouwer diagram of defect population in oxide MO vs. N_2O level.

The vertical dashed line represents the transition between intrinsic behavior (at low doping levels, $p = 2[V_M^{''}]$) and extrinsic behavior (at high doping levels, $p = [N_M^{'}]$). In the intrinsic regime, the Brouwer approximation is given by $p = 2[V_M^{''}]$. Rearranging equation 31.58 we obtain equation 31.61:

$$[V_M^{''}]p^2 = 4[V_M^{''}]^3 = K_{ox}p_{O_2}^{1/2} \tag{31.61}$$

However, since temperature (and therefore K_{ox}) and p_{O_2} are both constants, the metal vacancy concentration and hole content will also be fixed ($p = 2[V_M^{''}]$) and independent of doping level. It can be seen that all three defect species ($p, [V_M^{''}], [N_M^{'}]$) are represented in the intrinsic regime, but it remains to be established how the metal vacancy concentration varies with doping level in the extrinsic regime. Rearranging equation 31.61 and substituting $[N_M^{'}]$ for p we obtain equation :

$$[V_M^{''}] = K_{ox}p_{O_2}^{1/2}/p^2 = K_{ox}p_{O_2}^{1/2}/[N_M^{'}]^2 \tag{31.62}$$

Since temperature (and K_{ox}) and p_{O_2} are constants, it follows that $[V_M^{''}]$ will be proportional to the dopant concentration to the -2 power, as shown in Figure 31.12.
We can also create a log[defect] vs. $\log p_{O_2}$ diagram (at fixed temperature and fixed acceptor dopant level) for this system. Again, we need to decide where the extrinsic regime belongs, on the left or on the right. This is not as straightforward as with the log[defect] vs. log[dopant] diagram. However, examination of the intrinsic defect reaction of equation 31.57 indicates that as the oxygen partial pressure increases, the metal vacancy concentration will increase. This suggests that the intrinsic regime should be on the right with the extrinsic regime on the left. (Hint: if you try doing it the other way around, the resulting diagram will not make sense.) We begin by drawing a horizontal dashed line in Figure 31.13 indicating the fixed acceptor dopant level.

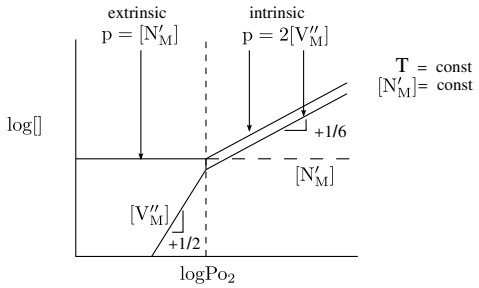


Figure 31.13: Schematic Brouwer diagram of dopant concentration vs. oxygen partial pressure for oxide MO at fixed temperature and N_2O doping level.

Again, the vertical dashed line indicates the transition between the two Brouwer approximations or ENC regimes. The intrinsic regime is solved the same way as for the butterfly diagrams of the previous section. Inserting the ENC, namely that $p = 2[V_M^{''}]$, and rearranging equation 31.58 we obtain equation 31.63:

$$[V_M^{''}]p^2 = 4[V_M^{''}]^3 = K_{ox}p_{O_2}^{1/2} \tag{31.63}$$

If we take the logarithm of both sides and rearrange, we obtain equation :

$$3\log[V_M^{''}] = \log\frac{K_{ox}}{4} + \frac{1}{2}\log p_{O_2} \tag{31.64}$$

from which it follows that $\partial \log[V_M^{''}]/\partial \log p_{O_2}$ will be +1/6, as shown in Figure 31.13. All three species (p, $[V_M^{''}]$, $[N_M^{'}]$) are represented in the intrinsic regime, but we need to decide how the metal vacancy concentration behaves in the extrinsic regime. Once again we can employ equation 31.58, which rearranged gives us equation 31.65:

$$[V_M^{''}] = (K_{ox}p_{O_2}^{1/2})/p^2 = (K_{ox}p_{O_2}^{1/2})/[N_M^{'}]^2 \tag{31.65}$$

Since both K_{ox} (fixed by temperature) and $[N_M^{'}]$ are constants, it follows that the metal vacancy concentration will vary with p_{O_2} to the +1/2 power, as shown in Figure 31.13.
Now let's consider a cation-excess oxide MO that becomes an n-type semiconductor under reducing conditions. It can also be donor-doped with N_2O_3 additions during initial processing, again with electronic compensation. As with the prior example, we will generate dual-regime isothermal Brouwer diagrams vs. log p_{O_2} and also vs. dopant concentration. We begin by writing the metal-excess (metal interstitial) equilibrium, reproduced in equation 31.66:

$$M_M^x + O_O^x \overset{MO}{\rightleftarrows} \frac{1}{2}\mathrm{O_2(g)} + M_i^{\bullet\bullet} + 2e^{'} \tag{31.66}$$

for which the mass-action/equilibrium relationship is given by equation 31.67:

$$K_{red} = p_{O_2}^{1/2}[M_i^{\bullet\bullet}]n^2 \tag{31.67}$$

The N_2O_3 donor-doping reaction is given by equation 31.68:

$$N_2O_3 \xrightarrow{MO} 2O_O^x + 2N_M^{\bullet} + 2e^{'} + \frac{1}{2}\mathrm{O_2(g)} \tag{31.68}$$

Note that the one-way arrow indicates that this reaction goes to completion during the firing of N_2O_3-doped MO. The overall ENC can be written as equation 31.69:

$$n = 2[M_i^{\bullet\bullet}] + [N_M^{\bullet}] \tag{31.69}$$

This time, let's begin with the log[defect] vs. log p_{O_2} Brouwer diagram (at fixed dopant level). But will the intrinsic regime be on the left or on the right? One way of resolving this issue is to examine the intrinsic defect reaction of equation 31.66 and the corresponding mass-action/equilibrium relationship of equation 31.67. It can be seen that the metal interstitial population will increase (and overwhelm the extrinsic dopant level) as the oxygen partial pressure is reduced. This tells us to expect the intrinsic regime on the left (low p_{O_2}). In Figure 31.14 we begin by drawing a horizontal dashed line to represent the constant N_2O_3 or $[N_M^{\bullet}]$ level.

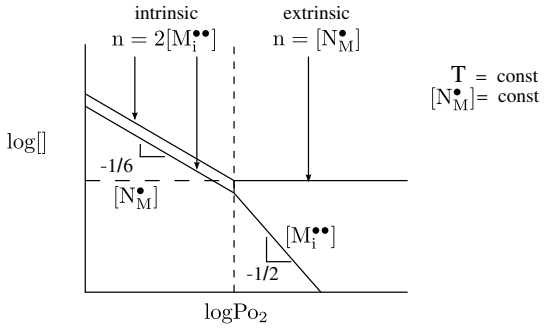


Figure 31.14: Schematic Brouwer diagram vs. log p_{O_2} of N_2O_3-doped MO at fixed temperature and doping level.

The vertical dashed line represents that transition from intrinsic behavior on the left to extrinsic behavior on the right. Assuming the Brouwer approximation in the intrinsic regime to be $n = 2[M_i^{\bullet\bullet}]$ and rearranging equation 31.67, we obtain equation 31.70:

$$[M_i^{\bullet\bullet}]n^2 = 4[M_i^{\bullet\bullet}]^3 = K_{red}p_{O_2}^{-1/2} \tag{31.70}$$

Rearranging once again and taking the logarithm of both sides, we obtain equation 31.71:

$$3\log[M_i^{\bullet\bullet}] = \log\frac{K_{red}}{4} - \frac{1}{2}\log p_{O_2} \tag{31.71}$$

from which it follows that $\partial \log[M_i^{\bullet\bullet}]/\partial \log p_{O_2}$ will be -1/6, as shown in Figure 31.14. We have lines for each of the three species in the ENC for the intrinsic regime, but how does the metal interstitial concentration behave in the extrinsic regime? Rearranging equation 31.67 and substituting $n = [N_M^{\bullet}] = const$, we obtain equation 31.72:

$$[M_i^{\bullet\bullet}] = (K_{red}/n^2)p_{O_2}^{-1/2} = (K_{red}/[N_M^{\bullet}]^2)p_{O_2}^{-1/2} \tag{31.72}$$

from which it follows that $\partial \log[M_i^{\bullet\bullet}]/\partial \log p_{O_2}$ will be -1/2, as shown in Figure 31.14.
Now let's consider the corresponding log[defect] vs. log[dopant] Brouwer diagram (at constant p_{O_2}). The increasing donor-dopant level is represented by the dashed line of slope +1 in Figure 31.15.

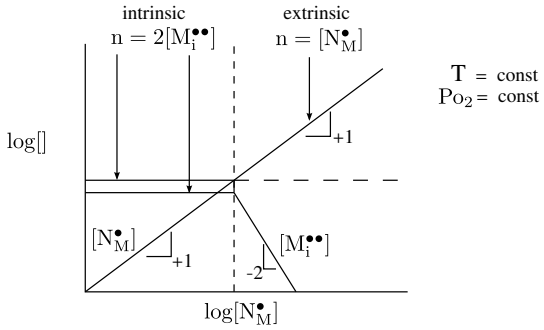


Figure 31.15: Schematic Brouwer diagram vs. doping level in N_2O_3-doped oxide MO at constant temperature and constant p_{O_2}.

The vertical dashed line represents the transition from intrinsic behavior to extrinsic behavior, the latter of which must fall on the right side of the diagram (high dopant level). In the intrinsic regime, since temperature (and therefore K_{red}) and p_{O_2} are both constant, if we rearrange equation 31.67 and insert the ENC ($n = 2[M_i^{\bullet\bullet}]$), we obtain equation 31.73:

$$[M_i^{\bullet\bullet}]n^2 = 4[M_i^{\bullet\bullet}]^3 = K_{red}p_{O_2}^{-1/2} = const \tag{31.73}$$

from which it follows that $n = 2[M_i^{\bullet\bullet}] = const$, as shown in Figure 31.15. We have lines for all three species in the overall ENC in the intrinsic regime, but how does the metal interstitial population behave in the extrinsic regime? Again, rearranging equation and inserting the ENC ($n = [N_M^{\bullet}]$) we obtain equation :

$$[M_i^{\bullet\bullet}] = (K_{red} p_{O_2}^{-1/2})/n^2 = (K_{red} p_{O_2}^{-1/2})/[N_M^{\bullet}]^2 \tag{31.74}$$

Again, since $K_{red}$is a constant (because T is constant) and p_{O_2} is also constant, the metal interstitial population will vary with the dopant concentration to the -2 power, as shown in Figure 31.15. We will revisit this diagram when we consider electrical conductivity in the following section.
As a transition to our consideration of ionic conductivity, we will construct one last log[defect] vs. log p_{O_2} Brouwer diagram. Calcia-doped zirconia is an oxygen vacancy ionic conductor except at very reducing oxygen partial pressures, where electrons can begin to contribute electronic conductivity. The intrinsic equilibrium between oxygen vacancies and electrons is given by equation 31.75:

$$O_O^x \overset{ZrO_2}{\rightleftarrows} \frac{1}{2}O_2(g) + V_O^{\bullet\bullet} + 2e^{'} \tag{31.75}$$

for which the mass-action/equilibrium relationship is given by equation 31.76:

$$K_{red} = p_{O_2}^{1/2}[V_O^{\bullet\bullet}]n^2 \tag{31.76}$$

The calcia (CaO) doping reaction is given by equation 31.77:

$$CaO \xrightarrow{MO} Ca_{Zr}^{''} + O_O^x + V_O^{\bullet\bullet} \tag{31.77}$$

Note here that the dopant cations are divalent ($Ca^{2+}O^{2-}$) whereas the host cations are tetravalent ($Zr^{4+}O_2^{2-}$). Note also the forward arrow, which indicates that this reaction goes to completion during the processing of CaO-doped ZrO_2. The overall electroneutrality condition is given by equation 31.78:

$$n + 2[Ca_{Zr}^{''}] = 2[V_O^{\bullet\bullet}] \tag{31.78}$$

This time we will focus exclusively on the log[defect] vs. log p_{O_2} Brouwer diagram, which is presented in Figure 31.16.

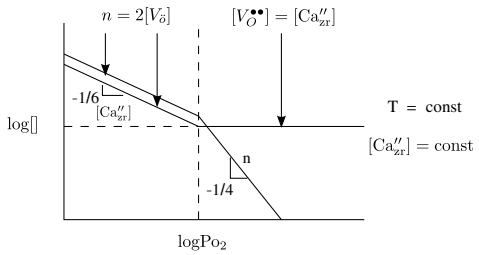


Figure 31.16: Schematic Brouwer diagram vs. p_{O_2} for CaO-doped ZrO_2 at constant temperature and constant doping level.

You should be able to rationalize: 1) that the intrinsic regime belongs on the left, and 2) that the slope in the intrinsic regime should be -1/6. It may not be obvious why the electrons have a -1/4 slope in the extrinsic regime. Isolating the electron population in equation 31.76, we obtain equation 31.79:

$$n^2 = (K_{red} p_{O_2}^{-1/2})/[V_O^{\bullet\bullet}] = (K_{red} p_{O_2}^{-1/2})/[Ca_{Zr}^{''}] \tag{31.79}$$

Taking the logarithm of both sides, and remembering that K_{red} is constant (because temperature is constant) and $[Ca_{Zr}^{''}]$ is also constant, it follows from equation 31.80:

$$2\log n = \log \frac{K_{red}}{[Ca_{Zr}^{''}]} - \frac{1}{2}\log p_{O_2} \tag{31.80}$$

that $\partial \log n / \partial \log p_{O_2}$ will be -1/4. We will return to this Brouwer diagram (Figure) in the following section.

32 Electrical Conductivity

The usefulness and predictive power of schematic Brouwer diagrams can best be illustrated by considering electrical conductivity, which is controlled by the point defect species with the highest product of concentration and mobility. It turns out that there can be contributions from first, electronic species (electrons and electron holes), and from second, mobile ionic species. The former determines the electronic conductivity of a solid and the second determines the ionic conductivity of that solid. Together, these two contributions determine the overall electrical conductivity.

32.1 Electronic Conductivity

From introductory materials science and engineering courses, you should be familiar with the following equation for the electronic conductivity:

$$\sigma = ne\mu_e + pe\mu_h \tag{32.1}$$

where e is the unit of electron charge, μ_e and μ_h are the mobilities of electrons and holes, respectively, and n and p are their concentrations, as previously defined. Usually, however, one species will dominate (as a majority species) whereas the other will have negligible concentration (a minority species). For example, electrons are majority species in the n-type semiconducting oxide MO, as reflected in the Brouwer diagram of Figure 31.14. In this case, the electronic conductivity reduces to equation :

$$\sigma = ne\mu_e \tag{32.2}$$

It is useful to do a unit analysis on the parameters in this equation to understand common usage in the materials community. The basic unit of conductivity is a Sieman per centimeter, or S/cm. The Sieman is the same as a reciprocal ohm ($1/\Omega$), and according to Ohm's law an ohm is voltage/current (V/I) or a Joule per coulomb (J/C) divided by a coulomb per second (C/s), such that an ohm is given by $(\mathrm{J}{\cdot}s)/C^2$. Placing these values into equation 32.2, we obtain equation 32.3:

$$] \frac{\mathrm{C}^2}{(\mathrm{J}{\cdot}\mathrm{s}{\cdot}\mathrm{cm})} = \frac{\#}{\mathrm{cm}^3}\mathrm{C}^{-}_{\mathrm{e}} \tag{32.3}$$

It follows that the basic units for mobility are cm^2 per V·s ($\mathrm{cm}^2/\mathrm{V}{\cdot}\mathrm{s}$). This may make more sense written as a centimeter per second (carrier velocity) per volt per centimeter (driving force), or $(\mathrm{cm/s})/(\mathrm{V/cm})$.
Here is where the predictive power of schematic Brouwer diagrams comes in handy. If the electronic mobility is independent of carrier concentration, then the conductivity in equation 32.2 will depend solely upon the carrier concentration, n. Let's consider the case of M-excess semiconductor MO, whose Brouwer diagram is represented in Figure 31.14. The corresponding predicted behavior of conductivity vs. oxygen partial pressure is given in Figure 32.1. Note that if the donor-dopant level is reduced to a minimum, the intrinsic regime ($n = 2[M_i^{\bullet\bullet}]$) can be preserved to high p_{O_2} values, as reflected by the "undoped" behavior in the Figure.

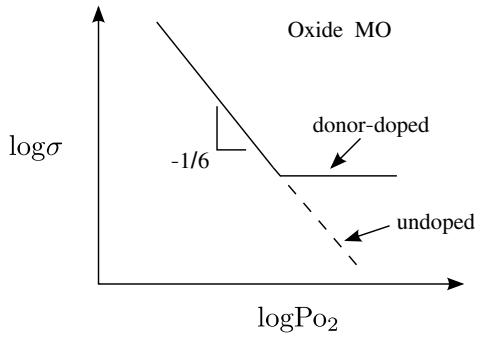


Figure 32.1: Predicted conductivity vs. oxygen partial pressure behavior of n-type cation-excess MO based upon the Brouwer diagram in Figure 31.14.

In the 1970s, the Ford Motor company developed a resistance-based oxygen sensor based upon such a $\log\sigma$ vs. $\log p_{O_2}$behavior as shown in Figure 32.1. Employing a highly pure titanium metal filament, which was subsequently oxidized to TiO_2, they achieved a reproducible conductance (the inverse of resistance, or G=1/R) vs.

p_{O_2}behavior that could be employed to monitor the oxygen content in the exhaust gases of the internal combustion engine. This enabled engineers to control the air-to-fuel ratio, otherwise known at the lambda-ratio. Hence, automotive oxygen sensors are typically referred to as lambda-sensors. It should be noted that given the different cation valence (4+ in TiO_2 vs. 2+ for MO), and therefore the different charge on an interstitial cation, the slope on a log-log plot of G vs. p_{O_2} was not the -1/6 for titania. Nevertheless, the invention of a resistance-based oxygen sensor was an important development. However, resistance-based oxygen sensors lost out to electrochemical sensors as described in the following section.

32.2 Ionic Conductivity

It makes sense that Ionic conductivity will likewise be determined by the product of ion concentration, ion charge, and ion mobility. A development of the so-called "Nernst-Einstein" relationship is beyond the scope of the present text, but one form of it describes the ionic mobility of a charged species, as given in equation 32.4:

$$\mu_i = \frac{D_i q_i}{kT} \tag{32.4}$$

where D_i is the diffusivity of an ionic species of charge q_i, T is absolute temperature and k is Boltzmann's constant. Since the product kT is in units of Joules, the product of cm^2/s and C divided by Joules yields the appropriate units for mobility of cm^2 per $V \cdot s$. The carrier concentration will be the concentration of mobile ions or c_i in units of $\#/cm^3$. The charge of the mobile species is given by q_i, which in turn is equal to the product of z_i (the number of charges on a mobile ion, e.g., -2 in the case of oxygen ions) and e, the unit of electron charge, or $q_i = z_i e$. Putting all this together, the product of carrier content, carrier charge, and carrier mobility is given by equation 32.5:

$$\sigma_i = c_i q_i \left(\frac{D_i q_i}{kT}\right) = \left(\frac{c_i q_i^2}{kT}\right) D_i \tag{32.5}$$

which is yet another version of the Nernst-Einstein relationship. Let's apply this equation to the situation of calcia-doped zirconia, whose doping reaction is repeated here as equation 32.6:

$$CaO \xrightarrow{MO} Ca''_{Zr} + O_O^x + V_O^{\bullet\bullet} \tag{32.6}$$

As always, the forward arrow indicates a reaction that goes to completion during the firing of calcia-doped zirconia. The resulting electroneutrality condition is $[V_O^{\bullet\bullet}] = [Ca''_{Zr}]$, as shown on the right side of Figure 31.16. The formula unit for CaO-doped ZrO_2 can be written as in equation 32.7:

$$(ZrO_2)_{1-x}(CaO\square)_x = (Zr_{1-x}Ca_x)O_{2-x}\square_x \tag{32.7}$$

where the box represents oxygen vacancies. In fact, we can write the ionic conductivity in terms of either oxygen diffusivity or oxygen vacancy diffusivity, as in equation 32.8:

$$\sigma_i = \left(\frac{c_{O^{2-}} q_{O^{2-}}^2}{kT}\right) D_O = \left(\frac{c_{V_O^{\bullet\bullet}} q_{V_O^{\bullet\bullet}}^2}{kT}\right) D_{V_O^{\bullet\bullet}} \tag{32.8}$$

where the $V_O^{\bullet\bullet}$ subscript refers to oxygen vacancies. Since $q_{O^{2-}}^2 = (-2e)^2 = q_{V_O^{\bullet\bullet}}^2 = (+2e)^2$, we can cancel these terms and also the kT product from both sides to obtain equation 32.9:

$$c_{O^{2-}} D_O = c_{V_O^{\bullet\bullet}} D_{V_O^{\bullet\bullet}} \tag{32.9}$$

If we divide both sides by the concentration of overall oxygen sites (c_{OS}), and replace $c_{O^{2-}}/c_{OS}$ and $c_{V_O^{\bullet\bullet}}/c_{OS}$ by the respective site fractions, X_O and $X_{V_O^{\bullet\bullet}}$, respectively, we obtain equation 32.10:

$$X_O D_O = X_{V_O^{\bullet\bullet}} D_{V_O^{\bullet\bullet}} \tag{32.10}$$

This may not look familiar, but we have previously dealt with a version of this equation when dealing with self diffusion by vacancy mechanism in metals. You can find it back in equation 29.31, which can be rewritten as $D_A = X_v D_v$. In reality, this expression should be $X_A D_A = (1 - X_v) D_A = X_v D_v$. In the case of metals, however, the site fraction of vacancies is negligibly small such that $X_A = (1 - X_v) \approx 1$. However, in the case of CaO-doped ZrO_2 the sizable vacancy content cannot be ignored.

Let's consider zirconia doped by 15 mole percent of calcia. The formula unit can be written as in equation 32.11:

$$(ZrO_2)_{0.85}(CaO\square)_{0.15} = (Zr_{0.85}Ca_{0.15})O_{1.85}\square_{0.15} \tag{32.11}$$

such that the site fractions will be $X_O = 1.85/2 = 0.925$ and $X_{V_O^{\bullet\bullet}} = 0.15/2 = 0.075$. Note that the ratio of site fractions $(X_O/X_{V_O^{\bullet\bullet}})$ is 12.33. From equation 32.10 it follows that the vacancy diffusivity at every temperature will be 12.33 times the oxygen ion diffusivity.
What we need to be able to do is to go back and forth between ionic conductivity and either of the diffusivities (oxygen, vacancy). If we know the concentration of overall oxygen sites, we can calculate the concentration of either occupied oxygen sites $(X_O c_{OS})$ or of oxygen vacancies $(X_{V_O^{\bullet\bullet}} c_{OS})$. The concentration of overall oxygen sites can be calculated in a couple of ways. The cubic fluorite structure of ZrO_2 is given in Figure 32.2.

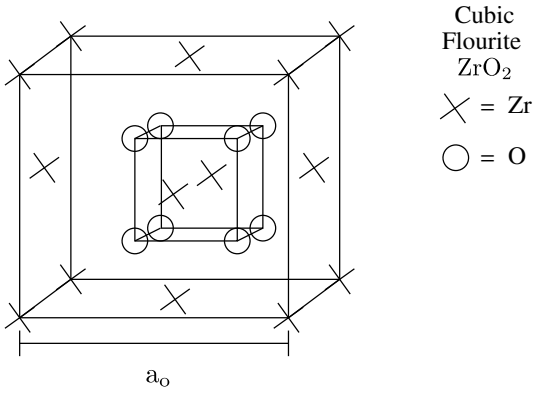


Figure 32.2: Schematic representation of the cubic fluorite structure of calcia-doped zirconia.

There are 4 formula units or 8 oxygen sites per unit cell. If given the lattice parameter (a_o), the concentration of oxygen sites will be $c_{OS} = 8/a_o^3$. Alternatively, we might be given the density of the the 15 mole percent-doped zirconia in g/cm^3. The mass per formula unit (FU) in equation 32.11 can be calculated as in equation 32.12:

$$\frac{\text{mass}}{\text{FU}} = [0.85AW_{Zr} + 0.15AW_{Ca}] + 1.85AW_O \tag{32.12}$$

Where "AW" stands for atomic weight. The density (g/cm^3) divided by the mass per FU (g/FU) gives the number of formula units per cm^3. The number of oxygen sites per cm^3 (c_{OS}) is just twice this value, since there are 2 oxygen sites per formula unit. Either way, we can now go back and forth between diffusivity (oxygen ion or oxygen vacancy) and ionic conductivity using the Nernst-Einstein relationship of equation 32.8 as illustrated in Figure 32.3.
Given that the oxygen vacancy concentration is fixed by the CaO doping, it follows that the oxygen diffusivity will be given by equation 32.13:

$$D_O = (D_O)_o \exp\left(\frac{-\Delta H_{mv}}{RT}\right) \tag{32.13}$$

Incorporating this expression into equation 32.8 we obtain equation :

$$\sigma_i = \left[\frac{c_O q_{O^{2-}}^2 (D_O)_o}{kT}\right] \exp\left(\frac{-\Delta H_{mv}}{RT}\right) \tag{32.14}$$

Except for temperature, all the other parameters within the brackets are constant or nearly so. This allows us to rearrange equation 32.14 into equation 32.15:

$$\sigma_i = \frac{\sigma_{io}}{T} \exp\left(\frac{-\Delta H_{mv}}{RT}\right) \tag{32.15}$$

By plotting the natural logarithm of the product of ionic conductivity and temperature $(\sigma_i T)$ vs. inverse temperature, we can determine the enthalpy of oxygen/vacancy motion, as shown in Figure 32.3. This figure shows how to calculate individual diffusivities from the ionic conductivity, and vice versa. It should be pointed out that once one diffusivity is determined, the other diffusivity can be readily calculated from the $X_O D_O = X_{V_O^{\bullet\bullet}} D_{V_O^{\bullet\bullet}}$ relationship.

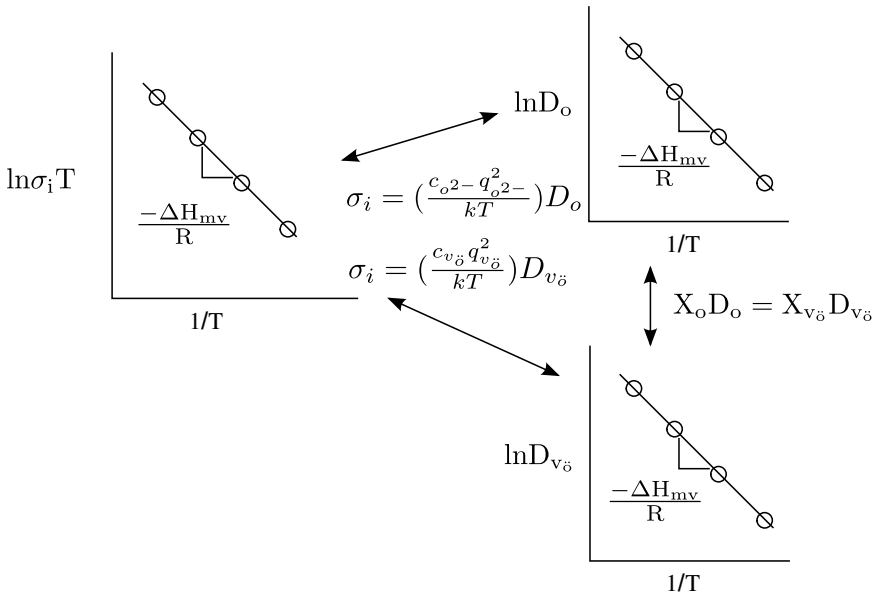


Figure 32.3: Using the Nernst-Einstein equation to go back and forth between ionic conductivity and diffusivity (ion or vacancy).

The oxygen sensors in the internal combustion engines of gas-powered automobiles are zirconia-based. However, instead of registering conductance (resistance) vs. oxygen partial pressure, the zirconia sensor serves as a solid electrolyte membrane between two environments. One side is exposed to air ($p_{O_2} = 0.21$) as a reference, while the other is exposed to the hot combustion gases coming from the engine. High temperature is required to keep the ionic conductivity high enough for the zirconia sensor to function. An electrochemical voltage is produced between the two sides of the zirconia electrolyte that is proportional to the difference in p_{O_2} between the two sides. When the air-to-fuel ratio is high, excess oxygen decreases the difference (and the voltage). But when the air-to-fuel ratio is low, the reducing products of combustion increase the difference (and the voltage). A microprocessor uses the output of the oxygen sensor to adjust the air-to-fuel ratio by controlling the fuel injectors in the engine. You might suppose that the purpose of the oxygen sensor would be to keep the engine operating at optimal efficiency, but this is not the case. The purpose of the oxygen sensor is to maintain the air-to-fuel ratio in a narrow range where all three catalysts in the catalytic converter function optimally to react unwanted pollutants (reduction of NO_x, oxidation of CO, and oxidation of unburnt hydrocarbons).

33 315 Problems

1. Use the Ellingham Diagram (reproduced here as Figure 33.1) to answer the following.
   (a) Find the temperature and partial pressure of O_2 where Ni(s), Ni(l), and NiO(s) are in equilibrium.
   (b) Can the same equilibrium be achieved with H_2 and H_2O instead of oxygen? If so, what is the ratio of H_2/H_2O?

2. Use the Ellingham Diagram to answer the following question. At 1245°C, H_2, and H_2O with a ratio of partial pressures of 10:1 is flowed through a tube furnace containing a crucible filled with MnO powder. Determine the driving force for the reaction

3. Using the Ellingham diagram, can you safely melt aluminum in a magnesia (MgO) container? Why or why not? What is the resulting reaction and its driving force?

4. Establish the T-log P_{O_2} phase diagram between 1000°C and 1500°C fort he Mn-O system at 1 atm total pressure.

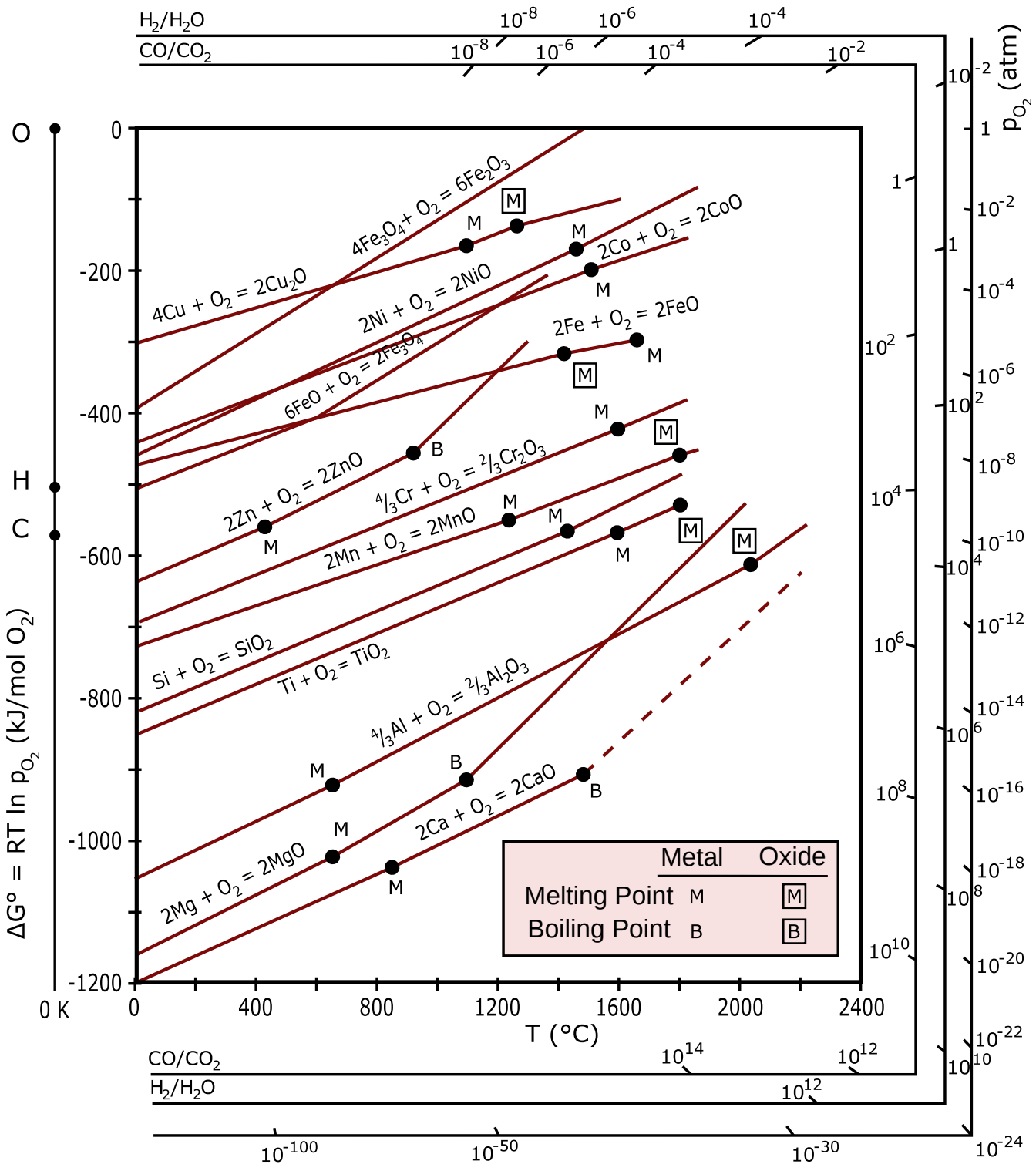


Figure 33.1: Ellingham diagram.

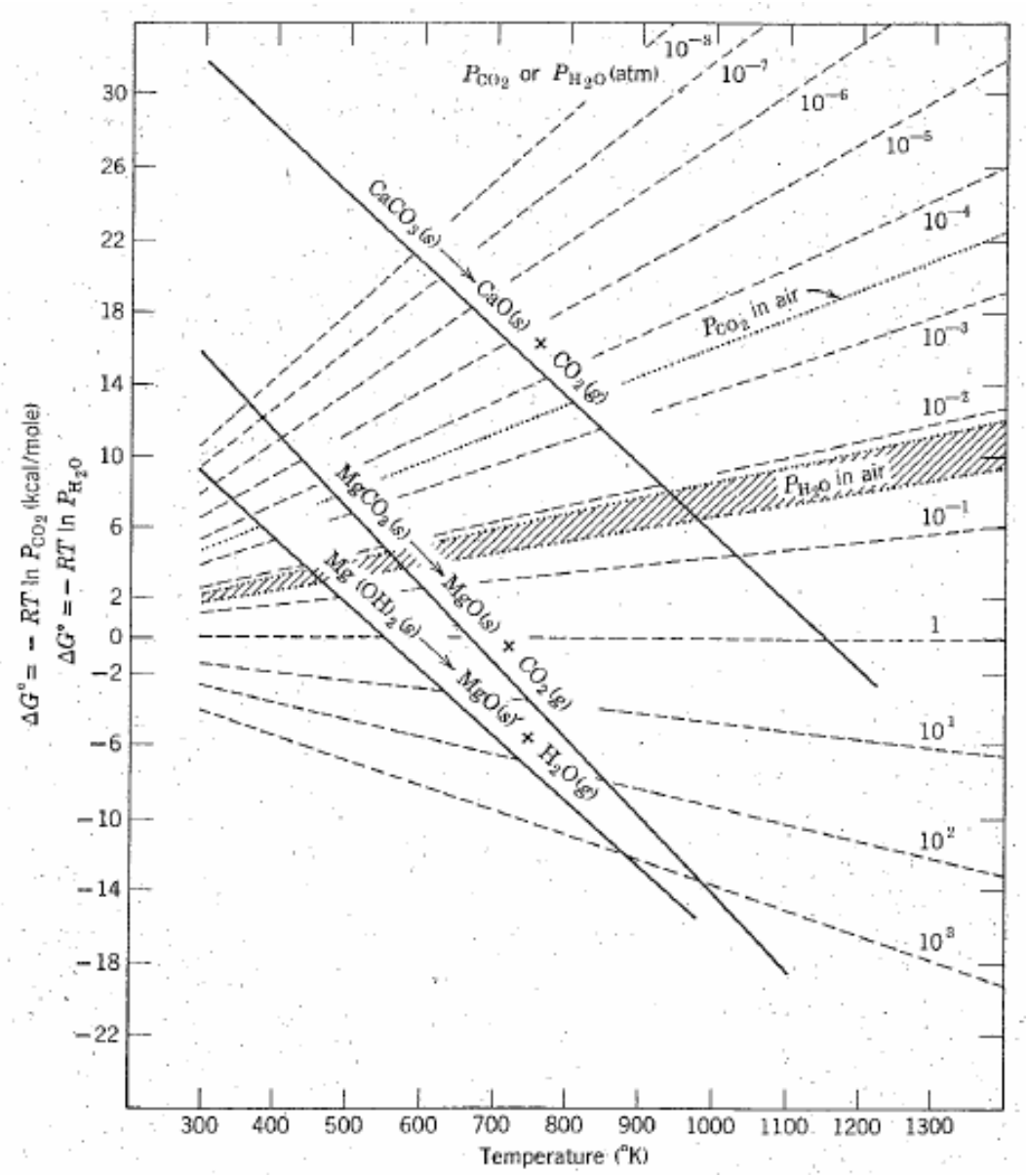


Figure 33.2: Thermodynamic data for Mg and Ca oxides and carbonates.

5. In the days before the industrial revolution the P_{CO_2} in the earth's atmosphere was 275 ppm. Use the data in Figure 33.2 to calculate how high one would need to heat $CaCO_3$ to decompose it at a P_{CO_2} for the preindustrial concentration of 275 ppm and for present day P_{CO_2} (You will have to look this up, please give your source). Also calculate how high one would need to heat $CaCO_3$ to decompose it if the CO_2 level in the atmosphere reaches 500 ppm.

6. Based on Raoultian liquid solution behavior, calculate the Sn-Bi eutectic phase diagram (using Excel, Mathematica, MATLAB, etc). Assume that there is negligible solid solubility of both Sn and Bi in the other component, and that $\Delta Cp \approx 0$ for both end members. Use the following melting points and enthalpies of fusion:

Material	Tm(K)	$\Delta H(s \rightarrow l)(J/mol)$
Sn	505.12	7030
Bi	544.59	11300

7. Use MATLAB or a spreadsheet to calculate liquidus and solidus lines for a "lens-type" T-X diagram for the A-B system, using the data below. You may assume both the liquid and solid solutions behave ideally.

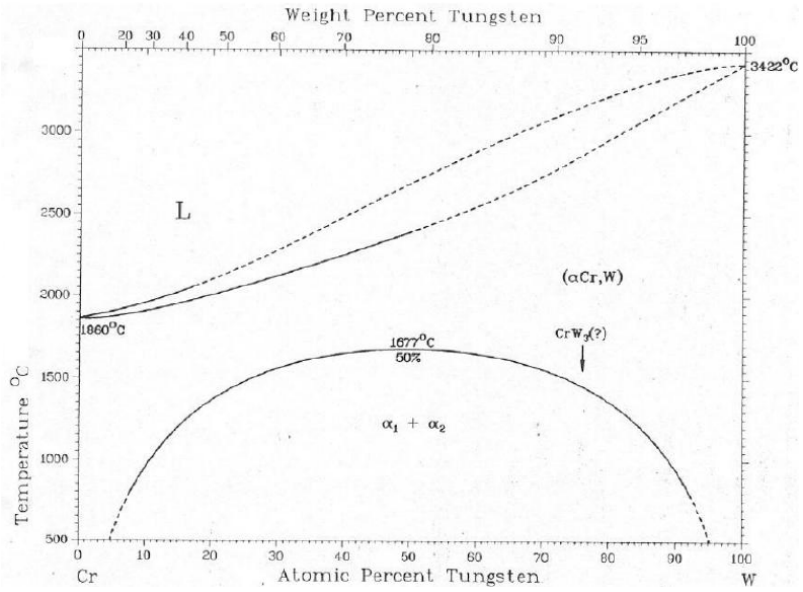


Figure 33.3: Cr-W Phase diagram.

Type	Tm(°C)	ΔHm(J/mol)
A	910	34700
B	1300	49800

(a) Plot the T vs. X phase diagram. Label each region on the diagram with the phases present and the degrees of freedom.

(b) For the temperatures 800°C, 1100°C, and 1500°C, draw plots of the activity of component A vs. composition. Include two plots for each: One with respect to liquid as the reference state and one with respect to solid as the reference state.

8. Based upon the temperature at the top of the miscibility gap in the Cr-W system (see Figure 33.3), do the following:

(a) Predict the miscibility gap (solvus) and spinodals based upon the regular solution model. Use the spreadsheet and plot the results.

(b) Compare your miscibility gap with the experimental one in the attached figure. Speculate about why there might be differences.

9. Consider the Pb-Sn phase diagram (see Figure 33.4).

(a) Label each region on the diagram with the degrees of freedom.

(b) Sketch free energy vs. composition curves for all phases at 150°C, 200°C, 250°C, and the eutectic temperature.

(c) For each temperature from part (b), draw plots of the activity of Sn vs. composition. Include two plots for each: One with respect to liquid as the reference state and one with respect to solid as the reference state. You may assume the liquid solution to be Raoultia. At 150°C, only plot activity of Sn vs. composition with respect to the solid reference state. At 250°C, only plot activity of Sn vs. composition with respect to the liquid reference state.

10. Calculate and plot the liquidus projection of the ternary phase diagram for the NaF-NaCl-NaI system. The melting temperatures and heats of fusion are as follows: NaF (990°C, 29,300 J/mol), NaCl (801°C, 30,200 J/mol) and NaI (659.3°C, 22,300 J/mol). Assume and ideal liquid solution and negligible solid solubility. Compare your result with the experimental diagram shown in Figure 33.5. Why might they be different?

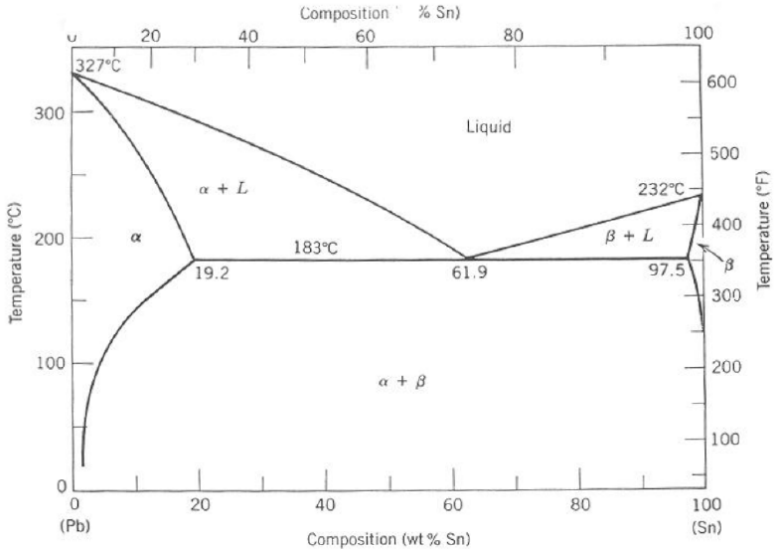


Figure 33.4: Pb-Sn phase diagram.

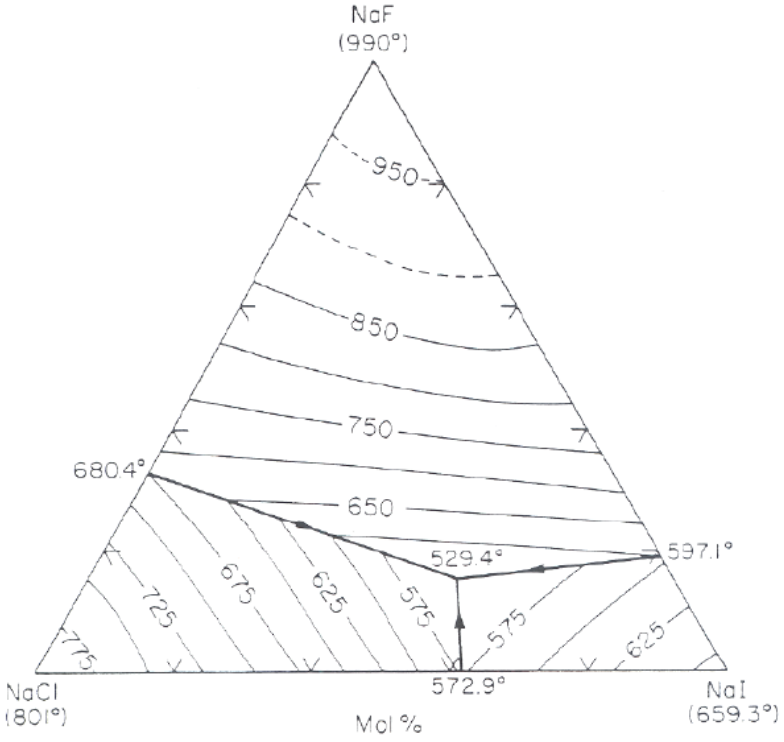


Figure 33.5: NaF-NaCl-NaI phase diagram. From ref. [20].

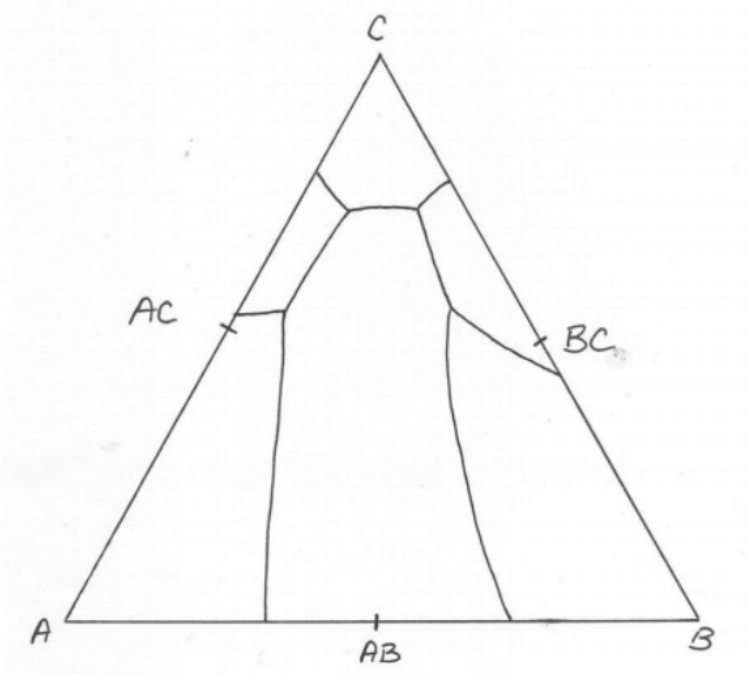


Figure 33.6: Liquidus projection diagram.

11. On the liquidus projection diagram for the hypothetical system A-B-C shown in Figure 33.6), complete the following:

    (a) Label primary phase fields
    (b) Draw the subsolidus compatibility joins.
    (c) Label all the binary and ternary invariant points.
    (d) Indicate the directions of falling temperature (binaries and ternary).
    (e) Sketch all the binary phase diagrams (including those formed by subsolidus compatibility joins).

12. Using the attached liquidus projection diagram for the hypothetical system A-B-C shown in Figure 33.7, complete the following:

    (a) Determine the equilibrium crystallization path for the composition marked with the star.
    (b) Determine the microstructural constituents:
        i. Just prior tot he liquid striking the phase boundary (liquid + solid 1 + solid 2).
        ii. At the eutectic but just prior to eutectic crystallization.
        iii. After crystallization is complete.

13. On the $(LiCl)_2 - CaCl_2 - (KCl)_2$ phase diagram shown in Figure 33.8, draw isothermal sections at the following temperatures: (note- Ternary eutectic E_1 is at 332°C and ternary eutectic E_2 is at 412°C)

    (a) 600°C
    (b) 450°C
    (c) 400°C
    (d) 300°C
    (e) Also determine the precise (not schematic!) $(LiCl)_2 - KCaCl_3$ phase diagram.

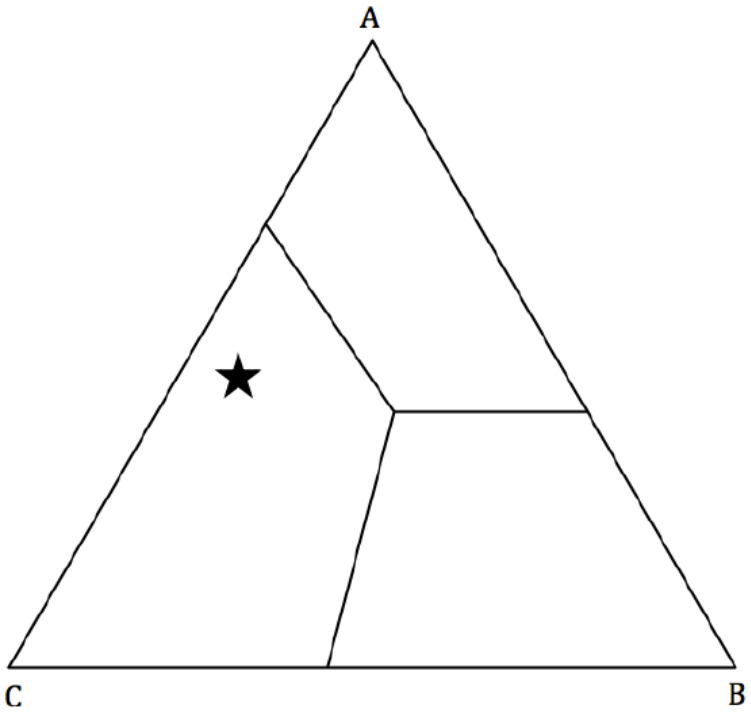
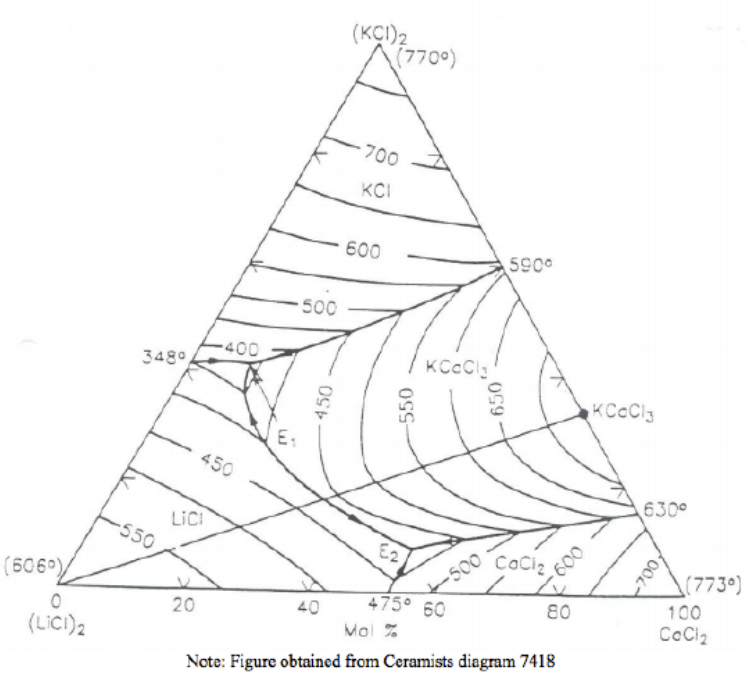


Figure 33.7: Liquidus projection diagram.



Note: Figure obtained from Ceramists diagram 7418

Figure 33.8: $(LiCl)_2 - CaCl_2 - (KCl)_2$ phase diagram.

14. (Bonus question - 10% of problem set value) Starting with the regular solution model, prove that regardless of how positive the interaction parameter (or heat of mixing) might be, the inital slope on any free energy vs. composition curve must be infinitely negative on the left side ($X_B \rightarrow 0$) and infinitely positive on the right side ($X_B \rightarrow 1$).

15. A steel tank contains hydrogen at 15 atm pressure. If the solubility of hydrogen in steel is 1x10^{-2} g/cm^3 under 15 atm pressure, the diffusion coefficient is 8x10^{-5}cm^2/s at room temperature and the tank is placed in a vacuum, calculate the flux of hydrogen through a 3.5 mm thick wall.

16. Austenite ($\gamma-$Fe) with .85 wt% carbon has a diffusion coefficient of 1.9x10^{-11}m^2/s at 900^{o}C.

    (a) Determine the jump distance in terms of the lattice parameter a_o and the coordination number for carbon diffusion in this structure.

    (b) How many jumps does a carbon interstitial make each second? Assuming a lattice vibration frequency of 10^{13}s^{-1}, what fraction of jumps is successful?

    (c) Calculate and compare the random walk distance with the total distance (back and forth) traveled by an interstitial carbon atom in one second.

17. Ferrite ($\alpha-$Fe) (BCC structure) dissolves carbon to a lesser extent than austenite (FCC structure).

    (a) Determine the jump distance in terms of the lattice parameter a_o and the coordination number for carbon diffusion in this structure.

    (b) Given the data in Table 2.1 of Porter, Easterling & Sherif, make an Arrhenius plot of diffusion coefficients of carbon and nitrogen from room temperature to 800^{o}C.

    (c) A different interstitial solute diffuses at a rate of 4.1x10^{-2}mm^2/s at 300^{o}C and 7.3x10^{-2}mm^2/s at 600^{o}C. Determine its activation energy and pre-exponential factor.

18. Write a MATLAB code to evaluate the composition as a function of distance for the draining plate problem.

    (a) For t/τ = 0.05 how many terms in the series is necessary to obtain a composition that is converged to within 1% of the exact answer. The percent error is the maximum value of $|c(x) - c_{exact}(x)|/c_{exact}(c)$x100. To determine the exact answer evaluate the summation to $j = 200$. L=100um, $C_o = 0.1$at. %

    (b) Plot the converged solution as a function of x for t/τ = 0.05, 0.5, 1.0, 2.0.

    (c) For what approximate value of t/τ does a single term in the summation with $j = 0$ provide an approximation to the exact solution within 10%?

19. (After Shewmon 2-13) We wish to consider the rate at which the vacancy concentration increases in a specimen after an increase in temperature. We assume that the vacancy concentration in the lattice near the free surface, grain boundaries and edge dislocations will rise to the new equilibrium value of the new temperature as soon as the temperature is raised. The vacancy concentration far from these vacancy sources rises only as fast as vacancies can diffuse to the region from the source.

    (a) Assume that vacancies come only from grain boundaries, and the grain diameter is approximately 1 mm. Calculate the relaxation time in two regimes, at high temperatures where the diffusion coefficient Dv is 10^{-5}cm^2/s.

    (b) Calculate the relaxation time (τ) given a dislocation line length (dislocation density) of 10^7cm/cm^3. (Hint: First, calculate the distance between dislocations, i.e, the vacancy sources.)

20. The diffusion coefficient of carbon in austenite can be approximated as:

$$D_c = 0.2 \exp\left(\frac{-136,000\,\text{J/mol}}{RT}\right) \text{cm}^2/\text{s}$$

    (a) How long does it take for the composition $c_{0.5}$ during carburization to penetrate .45 mm at 900^{o}C? How long for 5 mm?

    (b) What annealing temperature is required to double the penetration in a given time?

21. Consider two blocks initially one pure A and the other pure B that are welded together and annealed at 1100°C. Plot the diffusion profile as a function of distance after half an hour. Assume that the diffusion coefficient of both specie is $D = 4.5\text{x}10^{-11}\text{m}^2/\text{s}$ and that D is not a function of concentration.

22. Calculate the enthalpy and entropy of vacancy formation ($\Delta H_v, \Delta S_v$) for a system given the equilibrium concentration of vacancies (X_v^β) is $1.7\text{x}10^{-8}$ at 440 K and $1.5\text{x}10^{-5}$ at 650 K.

23. Given that $D = \frac{1}{6}\Gamma v a^2$, consider the diffusion of vacancies in an FCC lattice:

    (a) Let $\frac{\Delta S_m}{R} = 2$ and $v = 10^{13}\text{s}^{-1}$. Calculate the pre-exponential factor D_o for vacancies (assume $a_o = 0.4$ nm.)

    (b) If $\Delta H_m = 6.5$ kJ/mol, calculate D_v for vacancies at 750°C

24. Below are the linear thermal expansion ($\Delta L/L_o$) and X-ray lattice parameter ($\Delta a/a_o$) results at different temperatures for aluminum. Calculate and plot $\ln X_v$ from this data versus T^{-1} and determine the enthalpy and entropy of vacancy formation in aluminum. Show all equations used.

25. Below is a table of linear thermal expansion (ΔL/L) and lattice parameter expansion (Δa/a) vs. temperature for aluminum. Calculate and plot $\ln X_v$ form this data versus 1/T and determine the enthalpy and entropy of vacancy formation in this material. Show all equations used.

26. A gold specimen is quenched from 700°C to room temperature (25°C). An identical specimen is air cooled from 700°C to room temperature. The difference in their resistances is $\Delta\rho_o$. The quenched specimen is annealed at 40°C for 120 hours and then annealed at 60°C. Resistivity measurements were taken periodically by quenching the sample to room temperature. From the two slopes shown in attached Figure 2-16, find ΔH_{motion}.

27. Write balanced Kröger-Vink reactions for the following reactions, assuming full ionic charge for all ionic species.

    (a) Schottky defect formation in Li_2O

    (b) Anion Frenkel defect formation in Nb_2O_5

    (c) Oxidation of CdO to yield Cd/O < 1/1 (write both possible reactions)

    (d) Doping Al_2O_3 with ZnO to produce oxygen vacancies.

28. Pure ZnO is an n-type semiconductor dominated by oxygen vacancies, but it can be further donor-doped by substituting Al^{3+} for Zn^{2+} sites.

    (a) Draw a schematic Brouwer diagram as a function of pO_2 at fixed Al content.

    (b) Draw a schematic Brouwer diagram at fixed pO_2 as a function of Al content. Include all relevant point defect reactions and mass-action relationships.

29. Given the following oxygen ion diffusivities for calcia-stabilized zirconia (CSZ) of composition $(ZrO_2)_{0.85}(CaO)_{0.15}$ and density 5.5 g/cm^3, calculate a) the ionic conductivity at each temperature, and b) the enthalpy of motion. Assume oxygen vacancies are the dominant defect.

33.1 301 Computational Exercises

(Nothing yet to include)

Part IV
316-1: Microstructural Dynamics I

34 Diffusion

34.1 Review of the Basic Equations

The diffusion equation describes the evolution of the composition profile with time as the individual components diffuse within a sample. These components can be either atoms or molecules, but for our purposes we'll assume that the diffusing species are atoms (as in a metallic sample). For a binary system of A and B components, we can use either C_A or C_B (the respective concentrations of A and B species in atoms/volume) to describe the composition. These compositions sum to the total atomic concentration, C_0:

$$C_0 = C_A + C_B \tag{34.1}$$

If the molar volumes of A and B are equal to one another, then C_0 is fixed, so that the following conditions hold:

$$\frac{\partial C_A}{\partial z} = -\frac{\partial C_B}{\partial z} \tag{34.2}$$

$$\frac{\partial C_A}{\partial t} = -\frac{\partial C_B}{\partial t} \tag{34.3}$$

Note that we are using z as our spatial variable. For a binary A/B alloy we can use either C_A or C_B to describe the overall composition of the alloy. The flux of atoms is given by Fick's first law:

$$J_i = -D_i \frac{\partial C_i}{\partial z} \tag{34.4}$$

Here D_i is the intrinsic diffusion coefficient for component i and J_i is the diffusive flux of component i referenced to a given lattice plane in the material. In a binary system there are two intrinsic diffusion coefficients, D_A and D_B, and two diffusive fluxes, J_A and J_B. The time evolution of the composition is given by the continuity condition that relates a change in local concentration must be related to the spatial derivative of the flux:

$$\frac{\partial C_i}{\partial t} = -\frac{\partial J_i}{\partial z} \tag{34.5}$$

The diffusion equation is obtained by combining Eqs. 34.4 and 34.5:

$$\frac{\partial C_i}{\partial t} = \frac{\partial}{\partial z} D_i \left(\frac{\partial C_i}{\partial z} \right) \tag{34.6}$$

If the diffusion coefficient is independent of concentration (and hence, independent of x as well) then the diffusion equation can be written as follows:

$$\frac{\partial C_i}{\partial t} = D_i \frac{\partial^2 C_i}{\partial z^2} \tag{34.7}$$

The diffusion equation involves a first derivative with respect to time an a second derivative with respect to distance, so in general we need an initial condition and two boundary conditions. Consider for example the following situation:

- Boundary conditions: $C_A = C_1$ for $z \to -\infty$ and $C_A = C_2$ for $z \to \infty$
- Initial condition: The concentration jumps discontinuously from C_1 to C_2 at $x = 0$

With these initial and boundary conditions, the solution to Eq. 34.7 is:

$$C_A(x,t) = \frac{C_1 + C_2}{2} + \frac{C_2 - C_1}{2} \text{erf}\left(\frac{z}{w(t)} \right) \tag{34.8}$$

Here w is the following diffusion length, which enters into all diffusion problems:

$$w(t) = 2\sqrt{D_A t} \tag{34.9}$$

The solution to Eq. 34.8 is shown in Figure 34.1. To show how the concentration profile evolves with time, we have included values of w in the plot.

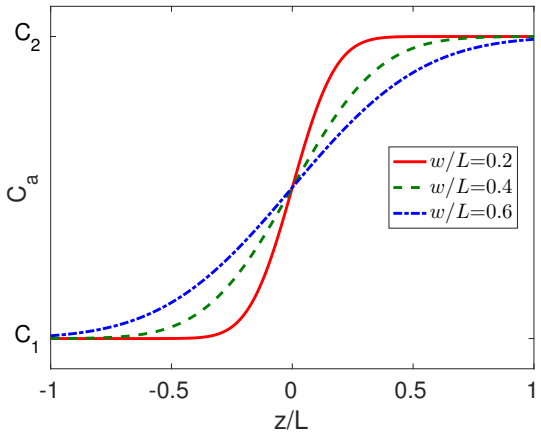


Figure 34.1: Representations of Eq. 34.8 for different values of $w = 2\sqrt{Dt}$.

This program was used to generate Figure 34.1:

```
figure
figformat % set some defaults so the figures look pretty
z=linspace(-1,1,200);  % These are the z points
w=[0.2,0.4,0.6]; % these are the three values of the normalized diffusion length that we will
  include in our calculations
c=@(z,w) erf(z/w);  % define a function of two variables, z and w
col={[1,0,0],[0,0.5,0],[0,0,1]};  % these are the three colors (rgb format)
linetype={'-','--','-.'};  % these are the three line types we well used (plain, dashed and
  dash-dot)
axes
hold on
for i=1:3
    plot(z,c(z,w(i)),'color',col{i},'linestyle',linetype{i})
    legendtext{i}=['$w/L$=' num2str(w(i))];
end
legend(legendtext,'location','best','interpreter','latex')
ylabel('C_{a}')
xlabel('z/L')
ylim([-1.2 1.2])
set(gca,'ytick',[-1,1])
set(gca,'yticklabel',[])  % turn off the y axis tick labls by making 'yticklable' an empty
  vector
text(-1.15, -1, 'C_{1}','fontsize',16)
text(-1.15, 1, 'C_{2}','fontsize',16)
print(gcf,'../figures/erfsolution.eps','-depsc2') % save as an eps file
```

We can also consider the situation where we have layer of material at $z = 0$, which diffuses in the positive and negative directions into the bulk material. In this case the initial and boundary conditions are as follows:

- Boundary conditions: $C_A = 0$ for $z = \pm\infty$
- Initial condition: All of the A component is confined to a very layer at $z = 0$, with a surface coverage (atoms/area) of C_s
- Normalization condition: The total amount of material in the sample must be conserved, so if we integrate the concentration profile we must end up with C_s:

$$\int_{-\infty}^{\infty} C_A(z)\, dz = C_s \tag{34.10}$$

In this case the following solution to the diffusion equation is obtained:

$$C_A(z,t) = \frac{C_s}{w(t)\sqrt{\pi}} \exp\left(-(z/w)^2\right) \tag{34.11}$$

Eq. 34.11 is plotted in Figure 34.2 for three different time points.

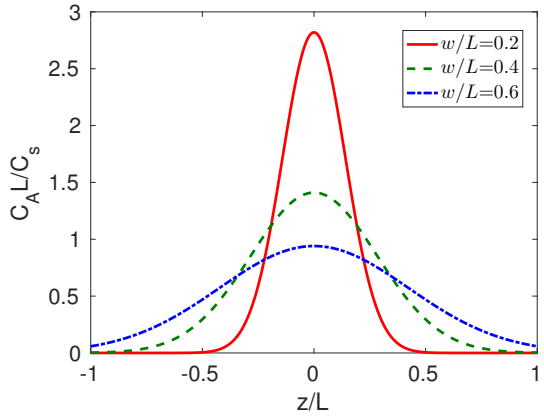


Figure 34.2: Representations of Eq. 34.8 for different values of the diffusion length, ℓ_A^D.

In many cases all we need to know is the diffusion length, ℓ_A^D, in order to understand what is going on at a pretty high level of detail. For example, ℓ_A^D describes both the width of interfacial mixing for two materials that are brought into contact with one another (Figure 34.1) and the diffusive broadening of a thin interfacial layer (Figure 34.2). The quantitative interpretation of the diffusion length in these two circumstances is illustrated in Figure 34.3. In Figure 34.3a we plot the interfacial broadening for a thin layer that is diffusing in the positive and negative z directions. The width of the diffusion profile can be characterized by the half-width of the peak, w', evaluated at half the total peak height. In Figure 34.3b we plot the concentration after bars with bulk concentrations of C_1 and C_2 are brought into contact with one another. In this case w' is obtained by drawing a tangent to the concentration profile at the midpoint between C_1 and C_2, and taking $2w'$ as the horizontal distance between the points where this tangent line reaches concentrations of C_1 and C_2. The value of w' in both cases is quite close to the diffusion length, w.

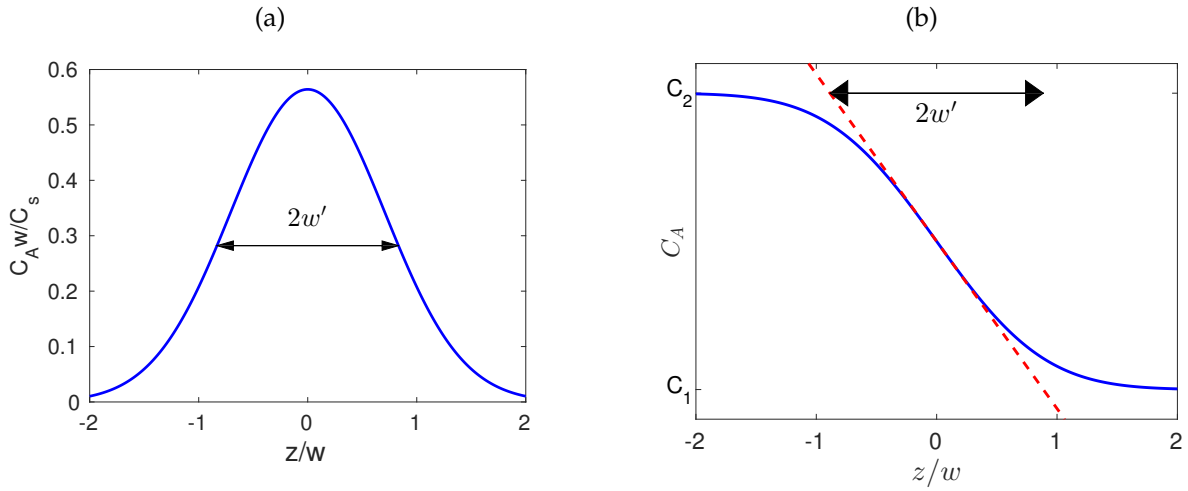


Figure 34.3: Illustration of the interfacial width for the thin-film solution (a) and error function solution (b) to the diffusion equation. In part a, $w' = 0.83w$ and in part b $w' = 0.89w$.

34.2 Mole Fractions and Volume Fractions

An assumption that we make throughout this text is that the atomic volumes of different chemical species are all identical, equal to V_0. In reality, this is almost never exactly true. Fortunately, it doesn't really matter when thinking

about diffusion because we can always work with volume fractions instead of mole fractions. In a generalized formulation the molecular volumes of the A and B molecules are given by multiplying the reference volume by a factor of N, which is not necessarily the same for each molecule:

$$\begin{aligned} V_A &= N_A V_0 \\ V_B &= N_B V_0 \end{aligned} \tag{34.12}$$

We can relate concentrations to mole fractions and volume fractions by considering a binary A/B system with n total atoms. Of these, nX_A are A atoms and nX_B are B atoms. Multiplying by the the atomic volume gives the total volume of each component. The total volume of A atoms is $nX_A N_A V_0$ and the total volume of B atoms is $nX_B N_B V_0$. From these expressions we obtain the following for ϕ_A, the volume fraction of A atoms in the system:

$$\begin{aligned} \phi_A &= \frac{nX_A N_A V_0}{V_{tot}} = C_A N_A V_0 \\ \phi_B &= \frac{nX_A N_A V_0}{V_{tot}} = C_B N_B V_0 \end{aligned} \tag{34.13}$$

where V_{tot} is the total volume of the system. Note that we have used $C_A = nX_A/V_{tot}$ and $C_B = nX_B/V_{tot}$. Throughout the rest of this text we generally assume that $N_A = N_B = 1$. In the case where N_A and/or N_B are not equal to one we can define renormalized concentrations, C_A^0 and C_B^0 that describe the concentration of subunits of volume V_0. These fluxes are related to the atomic fluxes, C_A and C_B by multiplying by the appropriate value of N:

$$\begin{aligned} C_A^0 &= C_A N_A = \phi_A / V_0 \\ C_B^0 &= C_B N_B = \phi_B / V_0 \end{aligned} \tag{34.14}$$

The renormalized fluxes, J_A^0 and J_B^0 are obtained by a similar normalization:

$$\begin{aligned} J_A^0 &= J_A N_A \\ J_B^0 &= J_B N_B \end{aligned} \tag{34.15}$$

Fick's first law still holds for these renormalized fluxes and concentrations, since we are just multiplying each side of Eq. 34.4 by N_i. Fick's second law applies for a similar reason. We can also use Eq. 34.14 to substitute ϕ_i for C_i^0:

$$\frac{\partial \phi_i}{\partial t} = \frac{\partial}{\partial x} D_i \left(\frac{\partial \phi_i}{\partial z} \right) \tag{34.16}$$

The bottom line of all this is that Fick's second law still applies, with same diffusion coefficient used for the case where the atomic volumes are equal, provided that we simply replace concentrations with volume fractions.

34.3 Vacancy Diffusion Mechanism

Figure 34.4 shows the output of a vacancy diffusion simulation of the interdiffusion between two materials. Vacancies move when an atom from an adjacent site moves into the vacancy. The resulting net motion of the atoms provides a means for diffusive mixing across an interface, and this is the process being illustrated in Figure 34.4 If the probability of hopping into a vacancy is different for A and B atoms, then $|J_A| \neq |J_B|$, and we need to consider additional effects. These are described below in our discussion of the Kirkendall effect.

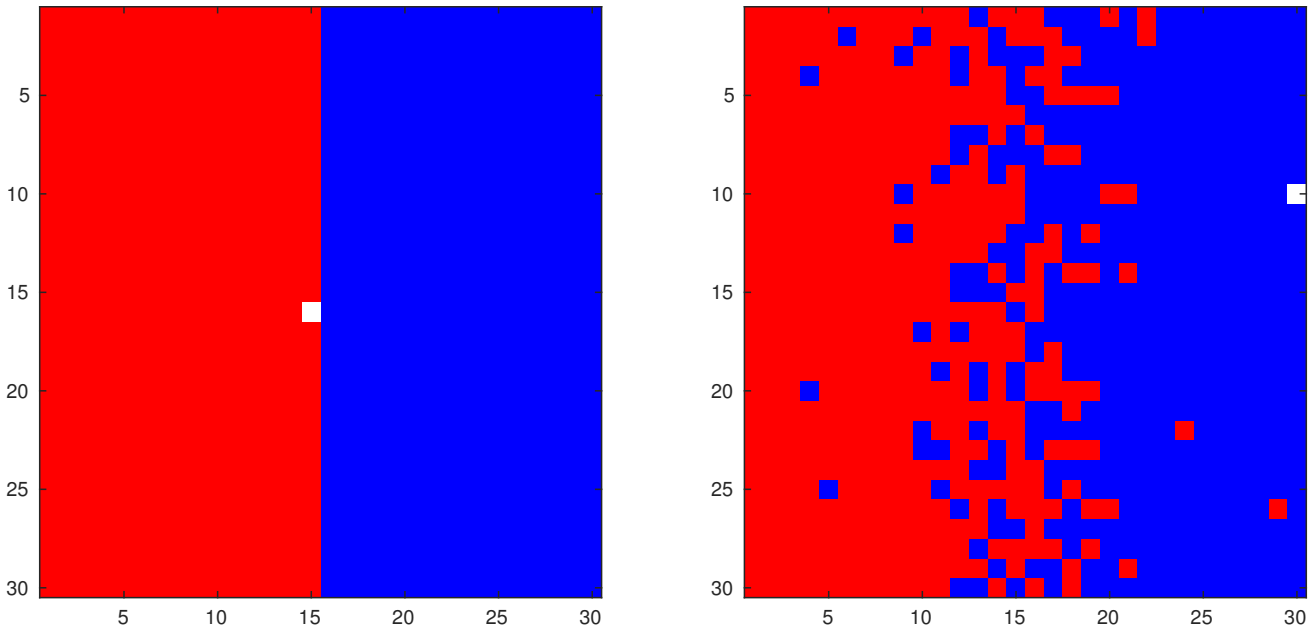


Figure 34.4: Diffusion couple in its initial state (a) and after 100,000 hops of the included vacancy (b). The size of the 2d lattice is 30×30..

The following program was used to generate the images shown in Figure 34.4.

```
tic % start a time so that we can see how long the program takes to run
n=30;  % set the number of boxes across the square grid
vfrac=0.01;  % vacancy fraction
matrix=ones(n);
map=[1,1,1;1,0,0;0,0,1];  % define 3 colors:  white, red, blue
figure
colormap(map)  % set the mapping of values in 'matrix' to a specific color
caxis([0 2]) % range of values in matrix goes from 0 (vacancy) to 2
% the previous three commands set things up so a 0 will be white, a 1 will
% be red and a 2 sill be blue
matrix(:,n/2+1:n)=2;  % set the right half of the matrix to 'blue'
i=round(n/2);   % put one vacancy in the middle
j=round(n/2);
matrix(i,j)=0;
imagesc(matrix); % this is the command that takes the matrix and turns it into a plot
t=0;
times=[1e4,2e4,5e4,1e5];
showallimages=1; % set to zero if you want to speed things up by not showing images,  set to 1
  if you want to show all the images during the simulation

%% now we start to move things around
vacancydiff.matrices={};  % makea blank cell array
while t<max(times);
    t=t+1;
    dir=round(4*rand+0.5);
    if dir==1
        in=i+1;
        jn=j;
        if in==n+1; in=1; end
    elseif dir==2
        in=i-1;
        jn=j;
        if in==0; in=n; end
    elseif dir==3
        in=i;
        jn=j+1;
        if jn>n; jn=n; end
    elseif dir==4
        in=i;
        jn=j-1;
```

```
        if jn==0; jn=1; end
    end
    % now we need to make switch
    neighborix=sub2ind([n n],in,jn);
    vacix=sub2ind([n n],i,j);
    matrix([vacix neighborix])=matrix([neighborix vacix]);
    if showallimages
        imagesc(matrix);
        drawnow
    end
    if ismember(t,times)
        vacancydiff.matrices=[vacancydiff.matrices {matrix}]; % append matrix to output file
        imagesc(matrix);
        set(gcf,'paperposition',[0 0 5 5])
        set(gcf,'papersize',[5 5])
        print(gcf,['vacdiff' num2str(t) '.eps'],'-depsc2')
    end
    i=in;
    j=jn;
end
vacancydiff.times=times;
vacancydiff.n=n;
save('vacancydiff.mat','vacancydiff') % writes the vacancydiff structure to a .mat file that we
    can read in later
toc
```

34.4 Kirkendall Effect

The geometry of the Kirkendall experiment (1947) is shown in Figure 34.5 [32]. In the experiment a small block of brass (70% copper, 30% Zn) was surrounded by inert, Molybdenum (Mo) wires. The sample was then coated with copper, and heated to a high temperature to allow atoms within the material to diffuse. In the measurement, the distance, w, between the Mo markers decreased as a function of time. This result implies that the flux of Zn out of the brass portion of the sample is larger than the copper flux back into the brass from the outside.

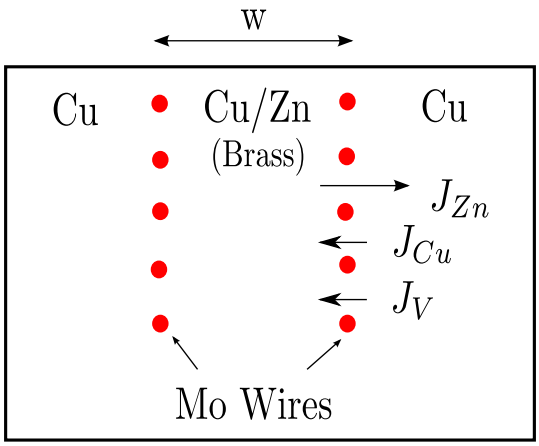


Figure 34.5: Experimental geometry of the original Kirkendall experiment

Diffusion does not need to occur by a vacancy motion in order for the Kirkendall effect to be observed, all that is needed is an asymmetry in the diffusion coefficients of the individual components in the material. However, for our purposes we will assume for now that diffusion occurs by a vacancy hopping mechanism. This assumption is valid for the original Kirkendall experiment, and it also enables us to make a connection to the relevant microscopic diffusion mechanisms. It is an excellent example of the structure/property relationships that define the field of materials science.

Our starting point is to assume that the vacancy concentration remains at equilibrium, so that the total number of lattice sites (including vacant sites) remains constant. A consequence of this assumption is that the fluxes of of A atoms, B atoms and vacancies must sum to zero:

$$J_A + J_B + J_v = 0 \tag{34.17}$$

Rearrangement of this equation, in combination with Fick's first law (Eq. 34.4) and the requirement that $\frac{\partial C_A}{\partial z} = -\frac{\partial C_B}{\partial z}$ leads to the following:

$$J_v = (D_A - D_B)\frac{\partial C_A}{\partial z} \tag{34.18}$$

The situation for $D_A - D_B > 0$ is illustrated in Figure 34.6. In this case the net vacancy flux is negative (to the left), and has a maximum magnitude at the point where the concentration gradient is the largest. Because the vacancy flux varies with position, there will be a time dependent increase or decrease in the local vacancy concentration that can be obtained from a site conservation equation similar to Eq. 34.5:

$$\frac{\partial C_v}{\partial t} = -\frac{\partial J_v}{\partial z} \tag{34.19}$$

This results in a net depletion of vacancies in some regions of the sample (the right in Figure 34.6c) and a net supersaturation of the vacancy concentration in other regions of the sample (the left in Figure 34.6c). In most cases processes exist that enable these concentration variations to be eliminated, by the creation of vacancies at the right portion of the sample and the destruction of vacancies at the left portion of the sample. Typically, these processes involve the addition or removal of vacancies to the core of a dislocation.

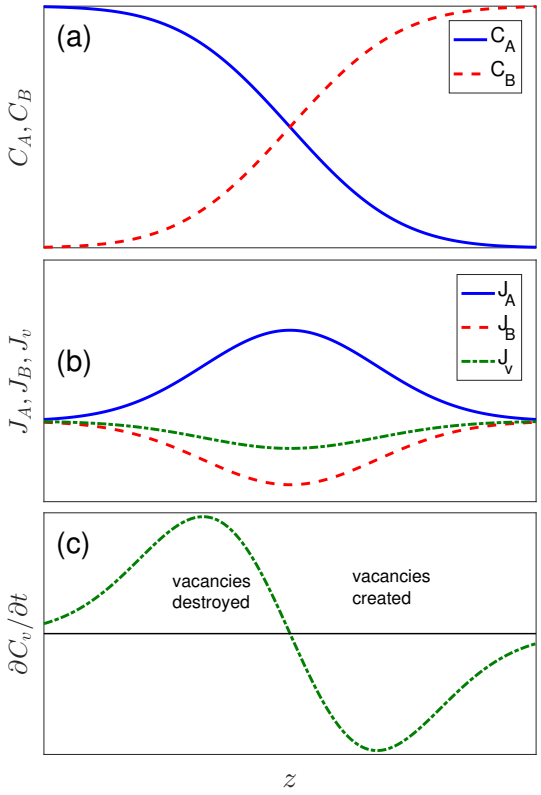


Figure 34.6: Representative concentration profiles during diffusion in a binary system (a), along with the fluxes for the case where $D_A > D_B$, and the change in local vacancy concentration due to the diffusive fluxes.

34.5 The Interdiffusion Coefficient

In general, a material flux, J, within a material can be related to a velocity, v. This velocity is obtained simply by multiplying J by the reference volume, V_0, that is used to define the diffusive flux: ($V_0 = 1/C_0$):

$$v\left(\frac{\mathrm{m}}{\mathrm{s}}\right) = J\left(\frac{1}{\mathrm{m}^2 \cdot \mathrm{s}}\right) \cdot V_0\left(\mathrm{m}^3\right) \tag{34.20}$$

The relevant velocity for us is a net, material velocity with respect to a set of inert markers, corresponding, for example to markers shown in the schematic representation of the Kirkendall experiment (Figure 34.5). It is easy to see from this picture, that if there is a net material flux to the right (in the positive direction), the result will be

a net motion of the markers to the left. The value of v_ℓ, the net velocity of the markers with respect to the ends of the samples, is determined by using -$(J_A + J_B)$ as the relevant flux in Eq. 34.20:

$$v_\ell = -\left(J_A + J_B\right) V_0 \tag{34.21}$$

We will often find it useful to use mole fractions instead of concentrations in our expressions, so we need to keep the following relationship in mind:

$$C_i = X_i C_0 \tag{34.22}$$

with i=A we can differentiate Eq. 34.22 to obtain:

$$\frac{\partial X_A}{\partial z} = \frac{1}{C_0}\frac{\partial C_A}{\partial z} \tag{34.23}$$

We can now combine Fick's first law (Eq. 34.4) with Eqs. 34.21 and 34.22 to obtain:

$$v_\ell = \left(D_A - D_B\right)\frac{\partial X_A}{\partial z} \tag{34.24}$$

This is the velocity that individual planes are moving with respect to a fixed position in the sample that is far from the interface (the ends of the sample, for example). The fluxes obtained from Fick's first law are defined in terms of a reference plane that is moving with a velocity v_ℓ. We can also define fluxes of A and B atoms across stationary planes, and we refer to these fluxes as J'_A and J'_B. We can get J'_A by adding $v_\ell C_A$ to J_A, where $v_\ell C_A$ is the net flux of A atoms across a fixed plane in space due to the lattice plane velocity:

$$J'_A = -D_A\frac{\partial C_A}{\partial z} + v_\ell C_A \tag{34.25}$$

We can combine this expression with Eq. for v_ℓ to get:

$$J'_A = -D_A\frac{\partial C_A}{\partial z} + \left(D_A - D_B\right)C_A\frac{\partial X_A}{\partial z} \tag{34.26}$$

With $C_A/C_0 = X_A$, we can combine Eqs. 34.23 and 34.26 to obtain:

$$J'_A = -D_A\frac{\partial C_A}{\partial z} + \left(D_A - D_B\right)X_A\frac{\partial C_A}{\partial z} \tag{34.27}$$

After a bit of algebra, keeping in mind that $X_A + X_B = 1$, we obtain the following:

$$J'_A = -\left[D_A X_B + X_A D_B\right]\frac{\partial C_A}{\partial z} \tag{34.28}$$

Now we can define an interdiffusion coefficient, $\tilde{D}$:

$$\tilde{D} = \left[X_B D_A + X_A D_B\right] \tag{34.29}$$

with this definition we have:

$$J'_A = -\tilde{D}\frac{\partial C_A}{\partial z} \tag{34.30}$$

A similar approach can be used to show that $J'_B = -J'_B$ and that the same value of $\tilde{D}$ can be used to relate J'_B to $\frac{\partial C_B}{\partial z}$. In addition, this same value of $\tilde{D}$ now appears in Fick's second law, where we can see from Eq. 34.29 that the value of $\tilde{D}$ is generally going to be composition dependent. The concentration profile therefore evolves according to the time-dependent solution of the following form of Fick's second law:

$$\frac{\partial C_i}{\partial t} = \frac{\partial}{\partial z}\tilde{D}\left(\frac{\partial C_i}{\partial z}\right) \tag{34.31}$$

34.6 Connection to thermodynamics

At equilibrium the chemical potential of component i, μ_i is a constant. If μ_i is not constant, then we must have diffusive fluxes as the system move towards equilibrium. It's not a gradient in concentration that generates the flux, it's really the gradient in the chemical potential, μ. A simple example illustrating this point is the abrupt change in concentration that exists at an equilibrated interface between two coexisting phases, shown as the α and β phases in Figure 34.7. Even though there is a large composition gradient at the interface, there is no diffusion for an equilibrated system because the chemical potential is spatially uniform.

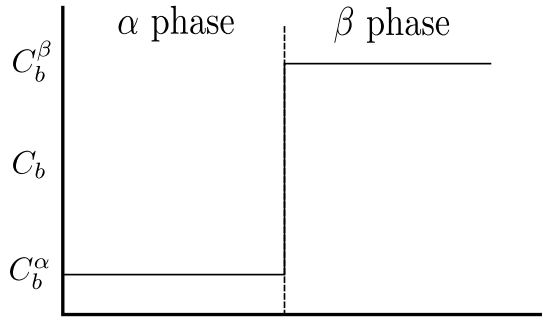


Figure 34.7: Schematic representation of the interface between α and β phases.

This example illustrates the fact that diffusion involves more than just concentration gradients, but involves thermodynamic factors as well. In order to account for these we need to revisit Fick's first law, but write things in terms of the chemical potentials. The flux of B atoms is can be written as the product of C_B and a diffusive velocity v_B where v_B is the average velocity at which the B atoms are moving. This velocity is related to the concentration gradient by a mobility coefficient, M_B:

$$v_B = -M_B \frac{\partial \mu_B}{\partial z} \tag{34.32}$$

Note that M_B must be positive. Atoms always move down a chemical potential potential gradient, although in some cases diffusion may take place up a concentration gradient (more on this in 316-2). We can use the previous expression for v_B to obtain the following expression for the diffusive flux of B atoms:

$$J_B = C_B v_B = -M_B C_B \frac{\partial \mu_B}{\partial X_B} \frac{\partial X_B}{\partial z} \tag{34.33}$$

We can use Eq. 34.23 to write this in terms of a concentration gradient:

$$J_B = -M_B X_B \frac{\partial \mu_B}{\partial X_B} \frac{\partial C_B}{\partial z} \tag{34.34}$$

Comparing to Fick's first law (Eq. 34.4), we obtain the following for D_B:

$$D_B = M_B X_B \frac{\partial \mu_B}{\partial X_B} \tag{34.35}$$

We see the this intrinsic diffusion coefficients involves purely kinetic parameter (the mobility, M_B), and a thermodynamic parameter (the derivative of μ_B with concentration). As discussed in more detail in 315 (see the sectional on Type II (binary) phase diagrams), chemical potentials are most commonly expressed in terms of activity coefficients in the following way:

$$\mu_B = \mu_B^0 + RT \ln a_B \tag{34.36}$$

Here μ_b^0 is the standard state, which is generally defined to be zero for a pure material at thermodynamic equilibrium. Tracer diffusion coefficients are defined in the dilute limit, where the activity of B atoms increases linearly with concentration in a way that is given by the Henry's law coefficient, H_B:

$$a_B = H_B X_B \tag{34.37}$$

In the Henry's law regime we see that $\partial\mu_B/\partial X_B = RT/X_B$, so the tracer diffusion coefficient for the B atoms is defined as follows:

$$D_B^* = RTM_B \tag{34.38}$$

The general relationship between D_B^* and D_B, valid for all compositions, and not just in the Henry's law regime, is obtained by comparing Eqs. 34.35 and 34.38:

$$D_B = \frac{D_B^* X_B}{RT}\frac{\partial\mu_B}{\partial X_B} \tag{34.39}$$

Composition dependent tracer diffusion coefficients can also be defined, so that D_B^* does not always correspond to the diffusion coefficient of B atoms in an alloy that is almost entirely made up of B. We can also imagine an alloy of any composition, but where we label a small fraction of B atoms (by isotopic substitution, for example). In this case the diffusing species is itself dilute, but the the alloy in which it is diffusing can have any composition.

34.7 Summary of Diffusion in a Binary System

We have defined three interrelated types diffusion coefficients: $\tilde{D}$, D_i and D_i^*. Here we provide a brief summary of these different diffusion coefficients and the relationships between them.

1. $\tilde{D}$: The interdiffusion coefficient (often referred to as the mutual diffusion coefficient). If you are interested in the time-dependent evolution of the composition profile, this is the diffusion coefficient that you use when you are solving the diffusion equation:

$$\frac{\partial C_B}{\partial t} = \frac{\partial}{\partial x}\tilde{D}\left(C_B\right)\left(\frac{\partial C_B}{\partial z}\right)$$

   note that for a binary system, we only need to specify one of the compositions, since $C_A = C_0 - C_B$. Also note that in general, $\tilde{D}$ depends on the composition, so cannot be treated as a constant.

2. D_A and D_B: The intrinsic diffusion coefficients for the individual components. These are important for two reasons. First, they are needed if you want to describe motions of atomic planes relative to the external boundaries of the sample (the Kirkendall effect). This motion was determined from the the atomic fluxes relative to atomic planes, as opposed to fixed points in space. These fluxes are determined by the appropriate intrinsic diffusion coefficient. For example, for the B component, we have:

$$J_B = -D_B\frac{\partial C_B}{\partial z}$$

   Also, predictive models of interdiffusion are generally based on the relationship between these intrinsic diffusion coefficients and the interdiffusion coefficient through the following expression:

$$\tilde{D} = [X_B D_A + X_A D_B]$$

3. D_A^* and D_B^*: The tracer diffusion coefficients for the individual components. Imagine a single atom in a homogeneous material. The tracer diffusion coefficient describes the probability that the atom has diffused a certain distance in a given period of time. These diffusion coefficients are purely kinetic parameters, and can be expressed in terms of a mobilty:

$$D_B^* = RTM_B$$

   Unlike the interdiffusion and intrinsic diffusion coefficients, they are not affected by the thermodynamics of the system. In general, values of the tracer diffusion coefficients will depend on the concentration of the material in which the tracer atoms are diffusing. The special cases of D_A^* at $X_A = 1$ and D_B^* at $X_B = 1$ are self diffusion coefficients. The intrinsic diffusion coefficients are related to the tracer diffusion coefficients through the following relationship:

$$D_B = \frac{D_B^* X_B}{RT}\frac{\partial\mu_B}{\partial X_B}$$

34.8 Diffusion in Ternary Systems

Atomic diffusion in ternary systems is driven by chemical potential gradients, just as it is does in binary systems. In systems with more than two components, however, the composition is no longer specified by a single composition variable. Some interesting effects can be observed in this case, as exemplified by carbon diffusion in Fe-Si-C ternary alloys . The carbon chemical potential is now a function of the concentration of both the silicon and carbon in the alloy:

$$\mu_c = f\left(X_{Si}, X_C\right) \tag{34.40}$$

The diffusion coefficient of carbon is much larger than the diffusion coefficient for silicon ($D_c \gg D_{Si}$), so we can assume that the silicon remains stationary during a diffusion experiment, as shown in Figure 34.8. Silicon and carbon have a unfavorable thermodynamic interaction within the alloy, so μ_c increases with increasing silicon content, X_{Si}. In order for the carbon chemical potential to remain constant across and interface between two regions of differing Si content, the carbon concentration in the region with low Si content needs to be smaller than the carbon concentration in the region with high Si content. This chemical potential discontinuity at the interface is eliminated by the jump of carbon atoms from the left (high Si side) to the right (low Si side) of the interface. Diffusion then continues from left to right, down the carbon potential gradient that has been established.

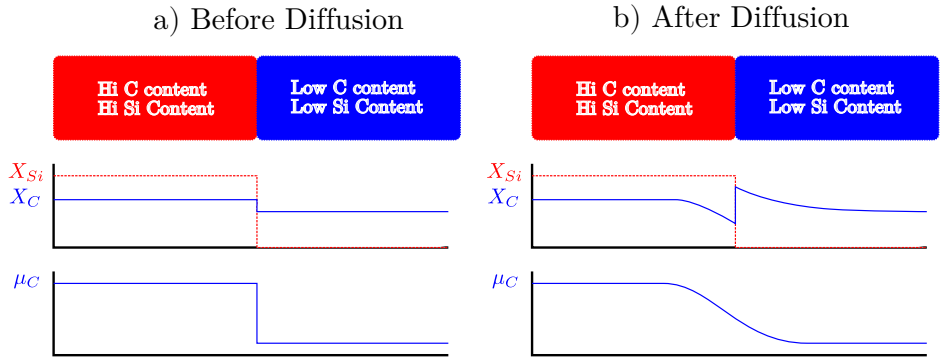


Figure 34.8: Chemical composition and carbon chemical potential across a ternary Si-C-Fe diffusion couple, showing the situation before diffusion (a) and after diffusion (b).

34.9 Crystal Defects and High Diffusivity Paths

> "Crystals are like people. It is the defects in them which tend to make them interesting." - Colin Humphreys

Real crystals are never perfect, and they always contain some sort of defects. These defects can be classified into four categories, based on their dimension:

- 0-dimensional (point) defects: These include missing atoms (vacancies), or atoms in location where they would not be in a perfect crystal structure (interstitials or substitutional impurities. From purely thermodynamic considerations we know that point defects must exist at some finite concentration for temperatures above 0K.
- 1-dimensional (line) defects: These are dislocations.
- 2-dimensional (planar) defects: These include grain boundaries, which are internal interfaces between regions of different crystalline orientation, and the external surfaces of a material.
- 3-dimensional (volume) defects: These are geometric imperfections in a material, like pores and cracks. We don't consider these types of defects in this class, but they become very important when we discuss the fracture properties of bulk, brittle materials in subsequent courses.

As illustrated in Figure 34.9, dislocation, grain boundaries and surfaces are associated with a more open structure. As a result diffusion along these defects is much faster than in the bulk of the material.

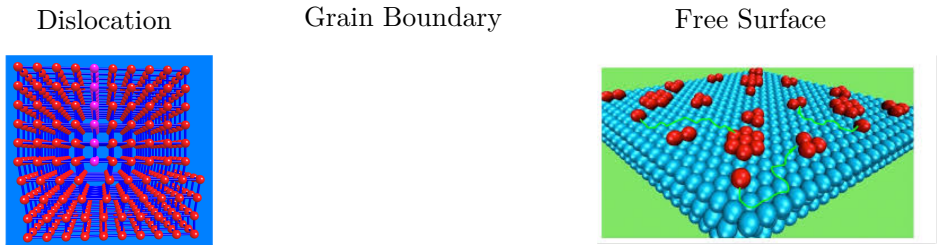


Figure 34.9: Examples of 1 and 2-dimensional defects that act as high-diffusivity paths.

35 Dislocations

Plastic deformation of a crystalline solid occurs by the motion of dislocations, which are one dimensional defects in the crystal structure. In general, deformation of a material occurs by shear along specified planes called slip planes. An illustration of this effect in single crystal aluminum is shown in Figure 35.1. The material in this image is being deformed in tension, but the slip occurs along suitably oriented planes that are experiencing a high degree of shear.

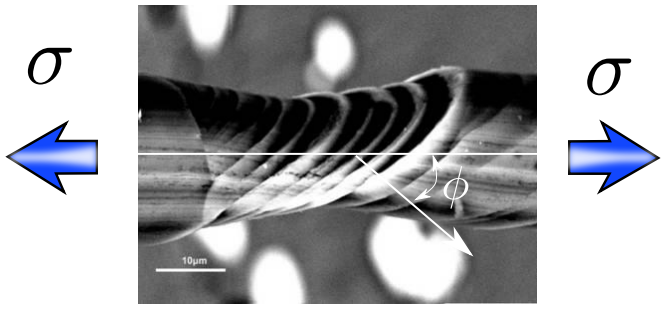


Figure 35.1: Slip bands in single crystal aluminum undergoing tensile deformation.

When a stress is applied to a single crystal, deformation takes place when the resolved shear stress on an appropriately aligned shear plane exceeds a critical value, referred to as the critical resolved shear stress, τ_{crss}. The relationship between the tensile stress, σ and the resolved shear stress, τ_{rss} is illustrated in Figure 35.2. In mathematical terms we have:

$$\tau_{rss} = \sigma \cos\phi \cos\lambda \tag{35.1}$$

where ϕ is the angle between the tensile axis and the slip plane normal, $\vec{n}$, and λ is the angle between the tensile axis and the slip direction, $\vec{d}$.

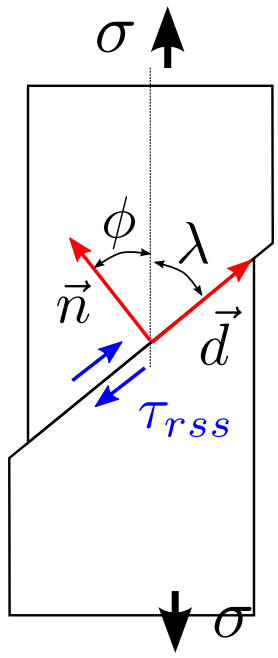


Figure 35.2: Relationship between an applied tensile stress, σ and the resolved shear stress, τ_{rss}.

Values of this quantity for different single crystals are shown in Table 35.1. For the materials with close packed crystals structures on this list (fcc and hcp), the value of τ_{crss} is about four orders of magnitude less than the shear modulus, G.

Table 35.1: Critical resolved shear stress for single crystals (Read-Hill, "Physics of Metals Principles", chap. 4 (1964).

Metal	Structure	G (psi)	G (Pa)	τ_{crss} (psi)	τ_{crss} (Pa)
Al	fcc	$3.9\text{x}10^6$	$27\text{x}10^9$	148	$1.0\text{x}10^6$
Cu	fcc	$7.0\text{x}10^6$	$48\text{x}10^9$	92	$0.64\text{x}10^6$
Mg	hcp	$2.4\text{x}10^6$	$17\text{x}10^9$	63	$0.44\text{x}10^6$
Zn	hcp	$5.6\text{x}10^6$	$38\text{x}10^9$	26	$0.18\text{x}10^6$
α-Fe	bcc	$9\text{x}10^6$	$27\text{x}10^9$	4000	$28\text{x}10^6$

.

A note about units of stress:
The SI unit of stress is a pascal (Pa), or N/m^2. We generally use SI units in this text, but English units (pounds per square inch, or psi), are still often used in engineering fields. One useful number to remember is that atmospheric pressure is $\approx 10^5$ Pa, or 14.7 psi. The exact conversion is that 1 psi = 6895 Pa = 6.895 kPa.

Exercise: From the critical resolved shear stress for single crystal aluminum shown in Table 35.1, calculate the minimum force (in pounds) that must be applied to a one half inch diameter rod of single crystal Al to deform plastically.
Solution: The critical resolved shear stress for pure, single crystal Al is 148 psi, so we need to figure out what tensile stress on the sample will produce this value for the resolved shear stress, τ_{rss}. The smallest value of σ for which τ_{rss} is equal to the critical value of 148 occurs for the slip system with $\phi = \lambda = 45°$, so from Eq. 68.1 we get $\sigma = 2\tau_{rss} = 296$. Multiplying by the cross sectional area of the rod gives:

$$F = (296\,\text{psi}) \cdot ß \cdot (0.25\,\text{in})^2 = 58\,\text{pounds}$$

This is a pretty small force, and is much less than the force required to deform a stock piece of aluminum that I would find in the machine shop.

Why is the force to deform a single crystal so low? We'll start by considering what we would expect for the critical resolved shear stress if the shear deformation were to occur by the sliding of atomic planes over one another, as shown conceptually in Figure 35.3. We refer to the stress required to slide these planes over one another as the dislocation-free critical resolved shear stress, τ^0_{crss}.

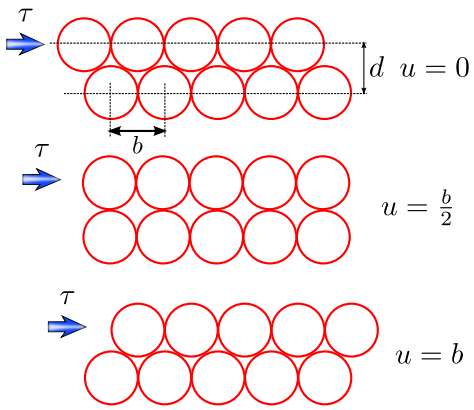


Figure 35.3: Sliding of close packed planes on top of one another.

We'll start by reminding ourselves of the definition of a shear strain, illustrated in Figure 35.4. In shear deformation, two parallel surfaces separated by a distance, d, are translated by an amount u with respect to one another. If the deformation occurs in the x-y plane, we refer to the shear strain as e_{xy}, which is given by:

$$e_{xy} = \frac{u}{d} \tag{35.2}$$

For a linearly elastic material, the shear stress, τ is proportional to e_{xy}, with the shear modulus G defined as the ratio of shear stress over shear strain:

$$\tau = G\frac{u}{d} \tag{35.3}$$

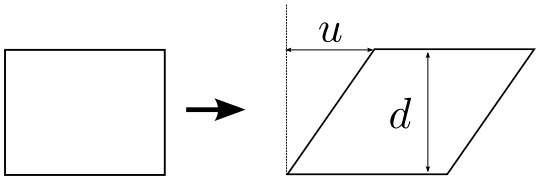


Figure 35.4: Application of a shear strain to a material.

In Figure 35.5 show a schematic representation of the stress as a function of displacement for the atomic planes shown in Figure 35.3. The stress function has the following features:

1. The stress is a periodic function, with the stress repeating every time the displacement is increased by an amount equal to b, the distance between atoms along the slip direction.
2. The stress is equal to zero at the stable equilibrium positions at $u = 0$, b, $2b$, etc.
3. For $u < b/2$ the stress is positive because we need to apply a stress to move the atoms out of their stable equilibrium positions.
4. At $u = b/2$ the system is at an unstable equilibrium. The stress is also equal to zero at this position, but the equilibrium is unstable because any slight perturbation in the displacement will cause the atomic plane to fall back into an equilibrium position at $u = 0$ or $u = b$.
5. The maximum stress is at $u = b/4$. The stress actually reverses sign for $u > b/2$, since a stress must be applied to avoid having the atoms fall into the equilibrium position at $u = b$.

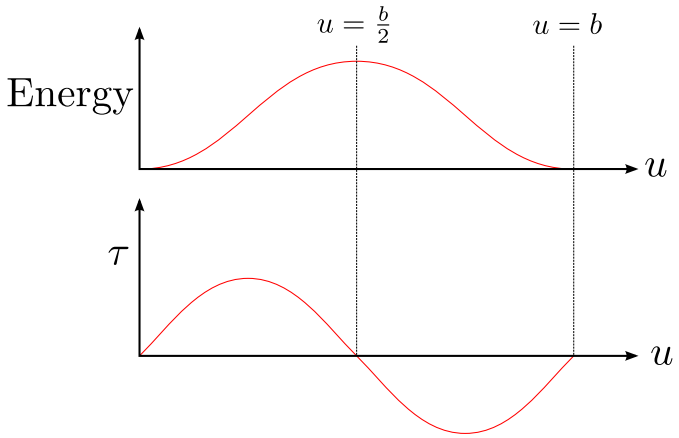


Figure 35.5: Schematic representation of the stress vs. displacement as the atomic planes in Figure 35.3 slide over one another.

The simplest mathematical expression for the shear stress that has the right periodicity is a sinusoidal function:

$$\tau = a \sin\left(\frac{2\pi u}{b}\right) \tag{35.4}$$

Now we need to figure out what the constant a is in terms of actual material properties. For small displacements the material is in the linear regime, and we can use the definition of the shear modulus (Eq. 35.3) to obtain the following:

$$\left.\frac{d\tau}{du}\right|_{u=0} = G/d \tag{35.5}$$

Comparison of Eqs. 35.4 and 35.5 gives $a = bG/2\pi d$, so the shear stress becomes:

$$\tau = \frac{bG}{2\pi d} \sin\left(\frac{2\pi u}{b}\right) \tag{35.6}$$

The critical resolved shear stress in this picture corresponds to the maximum value of τ, equal to $bG/2\pi d$. The interplanar spacing, d is comparable to b. (We're not going to worry about the exact numerical factor here, since we're just aiming to get an approximate expression for τ_{crss}). We take $b \approx d$ and $2\pi \approx 6$ to end up with the following expression for the ideal critical resolved shear stress, τ^0_{crss}, which is the value of the critical resolved shear stress we would expect to have if dis:

$$\tau^0_{crss} \approx G/6. \tag{35.7}$$

In reality, $\tau_{crss} \approx G/10^4$, so this picture of atomic planes sliding over one another can't be correct. What is really going on here? The answer is that slip occurs by the motion of dislocations, not by the concerted motion of entire planes of atoms across one another. The concept of slip by dislocation motion can be illustrated conceptually by the force required to slide a carpet across a floor. If the friction between the rug and the floor is very high, it's going to be very difficult to move the rug along the floor simply by grabbing it from one end and pulling. This situation is analogous to sliding atomic planes across one another as illustrated in Figure 35.3. If the rug just needs to be moved a small distance it is much easier to create a wrinkle at one end of the rug and move it to the other end of the carpet. At the end of the process, the carpet has moved by a length equal to the length of extra carpet stored in the wrinkle. Dislocations are line defects in crystalline materials that are analogous to these wrinkles.

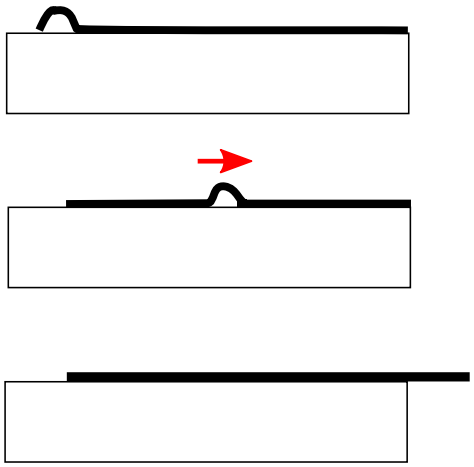


Figure 35.6: Moving a carpet by propagating a defect along its length.

35.1 Edge Dislocations

The easiest type of dislocation to visualize is an edge dislocation. A dislocation is formed by slipping part of the top half of a crystal relative to the bottom half by the application of a shear stress, τ, as illustrated schematically in Figure 35.7. The slip plane corresponds to the interface between the slipped and unslipped regions of the sample. An edge dislocation can be viewed as the termination of an extra half plane of atoms, and is illustrated for a simple cubic lattice in Figure 35.8.

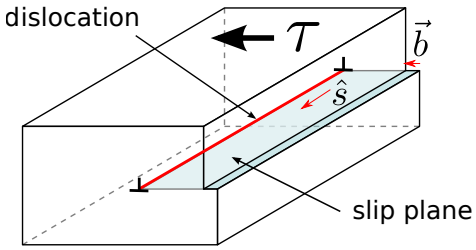


Figure 35.7: Illustration of the boundary between regions that have slipped from the application of a shear stress. The dislocation in this example refers to the boundary between the slipped and unslipped regions.

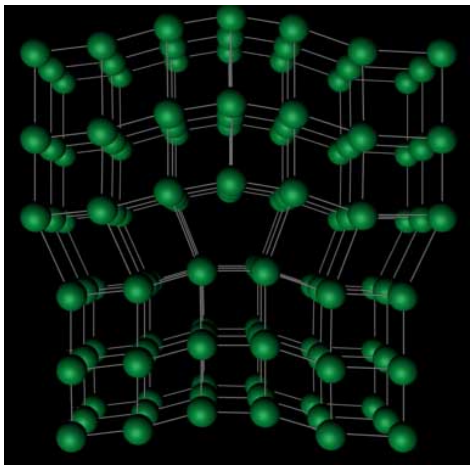


Figure 35.8: Edge dislocation in a simple cubic lattice.

Motion of an edge dislocation is illustrated in response to an applied shear stress is illustrated in Figure 35.9. Note that for every atom moving away from its equilibrium on one side of the dislocation core, there is an equivalent

atom moving toward an equilibrium position on the other side of the dislocation core. In energetic terms, for every atom that must be forced out of its lowest energy position, there is atom moving toward its lowest energy position. As a result the energy changes cancel (or very nearly so), and the energy barrier to moving a dislocation is much less than the barrier to slide surfaces across one another. As a result the net force to move a dislocation is very small. The stress needed to move a dislocation is generally much less than $G/6$, and is as low or lower than the observed critical resolved shear stress for single crystals.

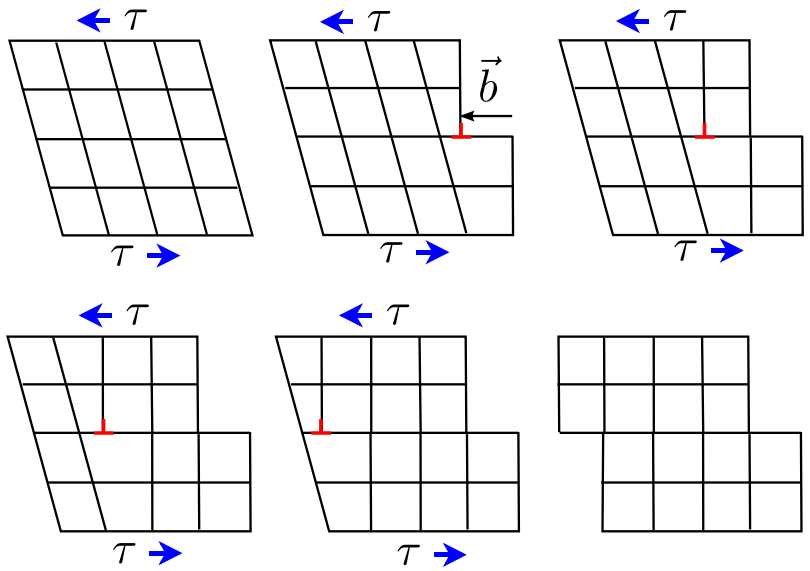


Figure 35.9: Schematic representation edge dislocation motion in the slip plane, illustrating the Burgers vector, $\vec{b}$ for an edge dislocation.

The relative displacement of the two halves of the crystal caused by the motion of a single dislocation through it is the Burgers vector, $\vec{b}$, which is the single most important characteristic of the dislocation. For an edge dislocation $\vec{b}$ is perpendicular to the dislocation line, which we represent by the unit vector $\hat{s}$ (the) *i.e.* $\vec{b} \cdot \hat{s} = 0$. Note that dislocations of opposite sign moving in opposite directions give the same final shear. This is illustrated by comparing Figures 35.9 and 35.10, which both result in the final deformed state of the material. Finally, when two edge dislocations with opposite Burgers vectors ($\vec{b}$ and $-\vec{b}$) meet on the same glide plane, they annihilate each other (see Figure 35.11).

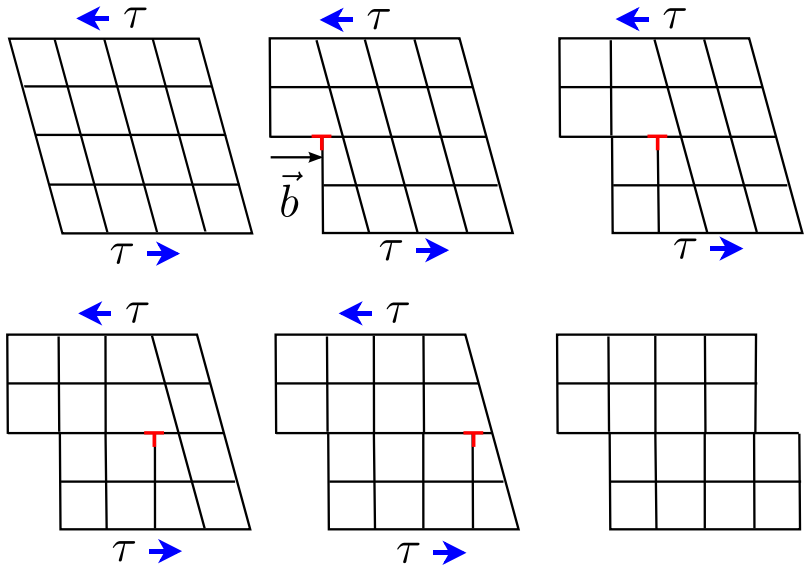


Figure 35.10: Deformation from Fig. 35.9, but resulting from an edge dislocation of opposite sign moving in the opposite direction.

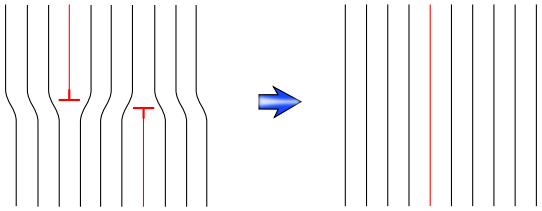


Figure 35.11: Annihilation of two dislocations of opposite sign that are moving in the same glide plane.

35.2 Screw Dislocations

As with edge dislocations, a screw dislocation line marks the boundary between 'slipped' and 'unslipped' regions of the sample, but for a screw dislocation the displacement described by $\vec{b}$ is parallel to the dislocation line, *i.e.* $\vec{b} \cdot \hat{s} = \left|\vec{b}\right|$. (Note that in order to simplify our notation, we'll refer to $\left|\vec{b}\right|$, the magnitude of the Burgers vector, simply as b in this text. A schematic representation of a the displacements associated with a screw dislocation is shown in Figure 35.12.

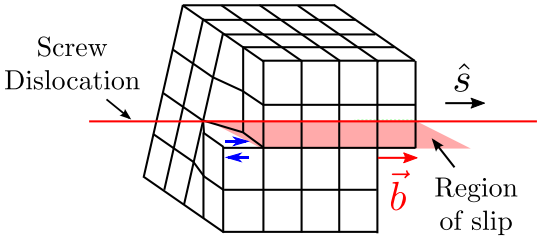


Figure 35.12: Schematic representation of a screw dislocation.

Figure 35.13 illustrates the the motion of a screw dislocation through a crystal. In this case the dislocation moves from the front of the crystal to the back of the crystal. The net effect of this motion is for the top and bottom halves of the crystal to be displaced to the right, by an amount and in the direction given by the Burgers vector. This figure illustrates the following:

1. When a dislocation line travels through a material, the motion of the line traces out a plane.

2. The relative displacement between the material on either side of this plane is given by the Burgers vector $\vec{b}$.

Note that this is true for ANY dislocation (edge, screw, or mixed).

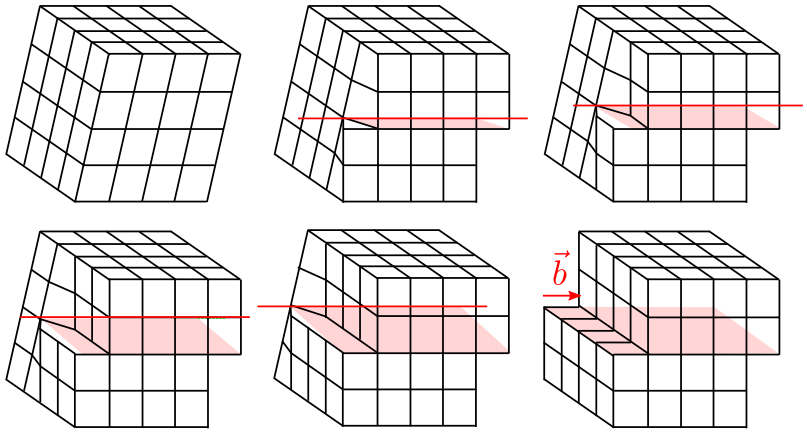


Figure 35.13: Motion of a screw dislocation The dislocation moves from the front of the crystal to the back. The net result is the production of a step edge with magnitude and direction equal tot he Burgers vector, $\vec{b}$.

35.3 The Burgers Circuit

In the previous section we have described some of the basic features of edge and dislocations, and have shown that they differ in the relationship between the orientation of the Burgers vector with respect to the dislocation line. Now we introduce a formal procedure that can be used to determine the value of $\vec{b}$ for any dislocation. The procedure is based on the use of a Burgers circuit, as described here:

1. Draw a circuit around the dislocation line that starts end ends at the same point. A 'right handed' convention is typically used to describe the direction that we take the circuit. (Clockwise looking along the direction of $\hat{s}$, counterclockwise if $\hat{s}$ is pointed at you).
2. Repeat the procedure, using the same numbers of atomic steps in each direction in a perfect crystal.
3. The Burgers vector is the vector connecting the start and end positions for the circuit drawn in the perfect crystal.

Use of the procedure is illustrated in Figure 35.14 for an edge dislocation with an extra half plane in the top half of the crystal. The circuit around the dislocation begins and ends at point a and proceeds as follows:

1. Move four steps down (a to b)
2. Move three steps to the right (b to c)
3. Move four steps up (c to d)
4. Move four steps to the left (d back to a)

When this same procedure is repeated in the perfect crystal we end up at point e, which is one step to the left of our starting point at point a. Our convention is to define the $\vec{b}$ as the vector starting at point a and ending at point b. When the procedure is repeated for a dislocation where the half plane is in the bottom half of the crystal we end up with the Burgers vector pointing in the opposite direction, as shown in Figure 35.15.

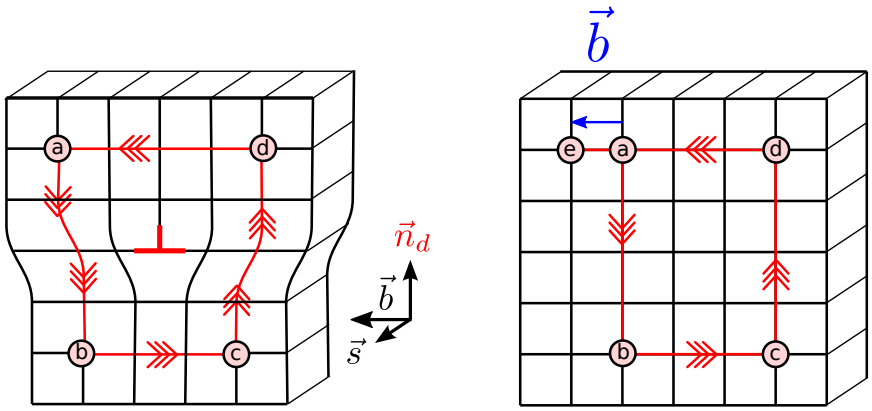


Figure 35.14: Determination of $\vec{b}$ for an edge dislocation. In this case $\hat{s}$ is defined so that it is pointing into the plane of the figure. The vector $\vec{n}_d$ is defined in Eq. 35.8.

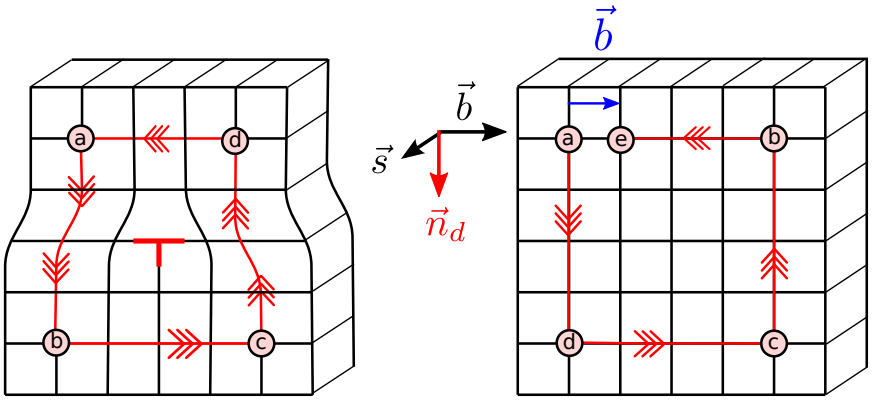


Figure 35.15: Determination of $\vec{b}$ for an edge dislocation with an opposite sign to the dislocation from Figure 35.14.

In Figure 35.16 we repeat the same process for a screw dislocation. In this example we have defined the direction of $\hat{s}$ so that the dislocation is pointed toward the bottom of the figure. The procedure for determining $\vec{b}$ is as follows:

1. Draw a circuit in the clockwise direction (viewed from the top, so we are looking in the direction of $\hat{s}$) around the dislocation line. The circuit begins and ends at point a.
2. Repeat the circuit in a perfect part of the crystal. The circuit begins at point s and ends at point f.
3. The Burgers vector is obtained as the vector that starts at s and ends at f.

Note that $\vec{b}$ is parallel to $\hat{s}$, as it must be for a screw dislocation, but that $\vec{b}$ and $\hat{s}$ are pointed in opposite directions, *i.e.*, they are anti-parallel. With our convention of drawing the $\vec{b}$ from the starting point to the ending point of the Burgers circuit in the perfect crystal, right handed screw dislocations have negative Burgers vectors and left handed screw dislocations have positive Burgers vectors. The left handed version of the dislocation shown in Figure 35.16 is shown in Figure 35.17.

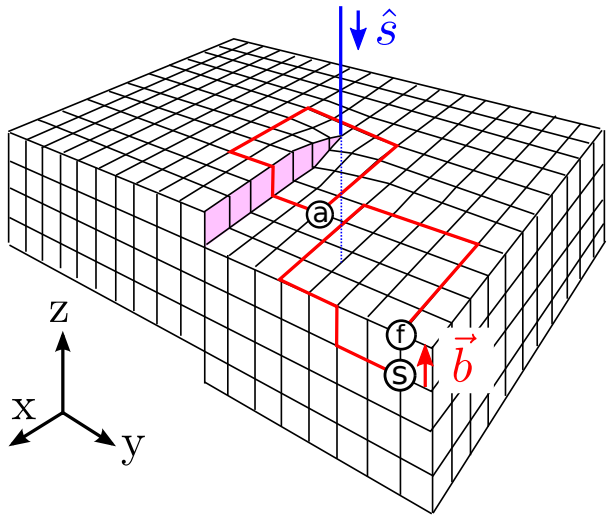


Figure 35.16: Burgers circuit for a right-handed screw dislocation, with $\vec{s}$ defined so that the the positive direction of the dislocation line is toward the bottom of the crystal (along the negative z direction).

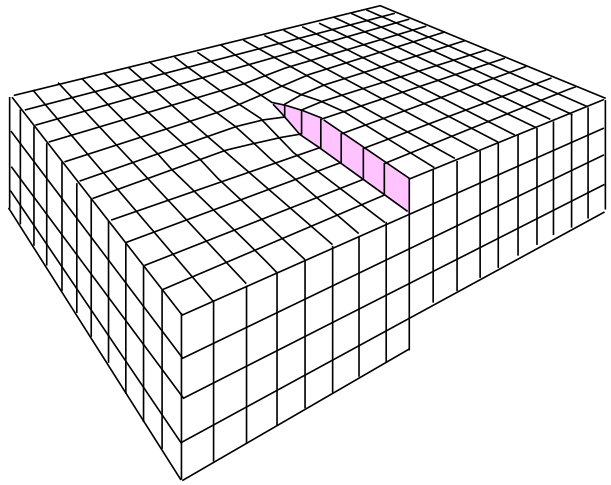


Figure 35.17: Left-handed version of the dislocation from Figure 35.16.

Exercise: Does the handedness of a screw dislocation (right handed or left handed) depend on the way you define the direction of $\hat{s}$?
Solution: No! If you you can see this by taking your right thumb and directing it along the dislocation line in Figure 35.16. In the direction that your figures are pointing, the planes spiral upward toward your thumb. It doesn't matter which way you orient your thumb when you do this. For the dislocation shown in Figure 35.17 you need to use your left hand to get this to work.

35.4 The $\vec{b} \times \hat{s}$ cross product

The concept of the Burgers circuit is a useful formalism that can always be used to specify the Burgers vector for a given dislocation. The confusing part about the procedure is that the sign of the Burgers vector depends on some arbitrary conventions that are not used the same way by everyone. For example, our convention is to define $\vec{b}$ as the vector linking the start to the finish of the Burgers circuit in the perfect crystal (linking points s to f in Figure 35.16), but you can find plenty of other people who draw the vector the other way around (drawing $\vec{b}$ from point f to point s). Nevertheless, we remove any ambiguity by always using this 'start-to-finish' definition for the Burgers vector. Similarly, we remove ambiguity regarding the direction in which we take the Burgers circuit by always

doing it the same way. In our case we use the right hand rule, directing our thumb along $\hat{s}$ and drawing the circuit in the direction in which our fingers are pointing.
Unfortunately, the ambiguity introduced by our definition of the direction of $\hat{s}$ along the dislocation line is impossible to remove. In figure 35.16 we defined $\hat{s}$ so that it points along the negative z direction, but there's no reason that we couldn't have defined $\vec{s}$ so that it is directed in the positive z direction instead. We end up with a Burgers vector that points in one of two opposite directions, depending on how we define $\hat{s}$ in the first place. The good news is that $\vec{n}_d$, the vector cross product of $\hat{s}$ and $\vec{b}$ is independent of our convention for defining the direction of $\vec{s}$. As a reminder, the vector cross product between vectors $\hat{s}$ and $\vec{b}$ is defined as follows, as illustrated in Figure 35.18:[7]

$$\vec{n}_d = \vec{b} \times \hat{s} = |\vec{b}|\,|\hat{s}| \sin\theta \hat{n} = b \sin\theta \hat{n}_d \tag{35.8}$$

Here $\hat{n}_d$ is a unit vector in the direction perpendicular to the plane containing $\hat{s}$ and $\vec{b}$. It's orientation is defined using the right hand rule: We place our right hand along $\hat{s}$, with our fingers oriented in the positive θ direction. Our right thumb is then pointed along $\hat{n}_d$.

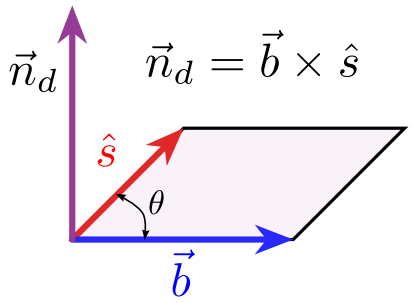


Figure 35.18: Definition of $\vec{n}_d$.

When defined in this way, $\vec{n}_d$ has the following properties:

- Because redefining $\vec{s}$ to have the opposite orientation also changes the orientation of $\vec{b}$, the negative signs cancel and we end up with a value for $\vec{n}_d$ that is independent of the way that we choose to define $\vec{s}$.
- For a pure screw dislocation, $\theta = 0$ or $\theta = 180^\circ$. In either case, $\vec{n}_d = 0$.
- For an edge dislocation, the magnitude of $\vec{n}_d$ is equal to the b, the magnitude of Burgers vector. In addition, $\vec{n}_d$ points toward the extra half plane.

This last point is perhaps the most important one, because it provides an easy way to figure out how the extra half plane is oriented in an edge dislocation, once we specify the orientations of $\vec{b}$ and $\hat{s}$. We just use the right hand rule, cross $\vec{b}$ into $\hat{s}$, and our thumb will be pointed along the direction of the extra half plane. To convince yourself that this actually works, you can try it with the edge dislocations pictured in Figures 35.14 and 35.15.
With our convention for using the Burgers circuit to obtain $\vec{b}$ (Right-hand-rule, start to finish), we have the following relationships between $\hat{s}$ and $\vec{b}$:

- Right-handed screw dislocation: $\hat{s}$ and $\vec{b}$ point in opposite directions.
- Left-handed screw dislocation: $\hat{s}$ and $\vec{b}$ point in the same direction.
- Edge dislocation: $\hat{s}$ perpendicular to $\vec{b}$, $\vec{b} \times \hat{s}$ points to the extra half plane.

35.5 Connection to the Crystal Structure

The Burgers vector must correspond to an atomic repeat distance in the crystal structure. As we show below, the energy of a dislocation is proportional to the square of the magnitude of the Burgers vector. For this reason the Burgers vector will correspond to closest atomic distance in crystal structure. As shown in Figure 35.19, the Burgers vector is half the unit cell diagonal for the BCC structure, and half the face diagonal of the unit cell in the FCC structure.

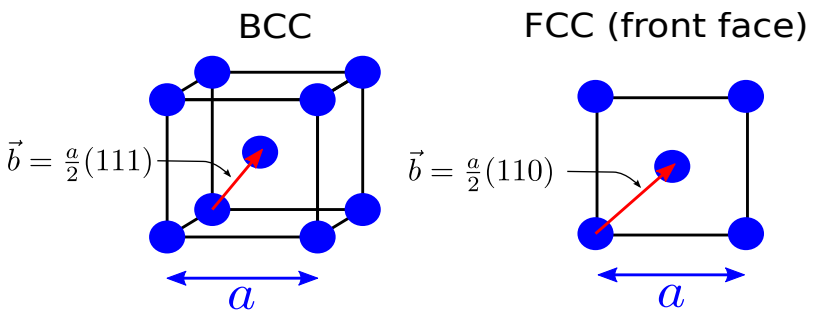


Figure 35.19: Burgers vectors for the BCC and FCC crystal structures.

35.6 Dislocation loops

A dislocation cannot terminate within a crystal, although it can terminate at a grain boundary or crystal surface. Also, while the Burgers vector along a given dislocation is constant, the dislocation itself is not necessarily a straight line. In other words, $\vec{b}$ is fixed, but $\hat{s}$ can change as the direction of the dislocation changes. Consider for example the dislocation loop shown in Figure 35.20.

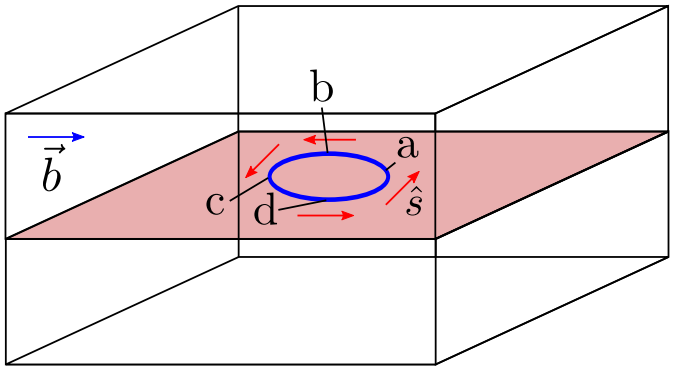
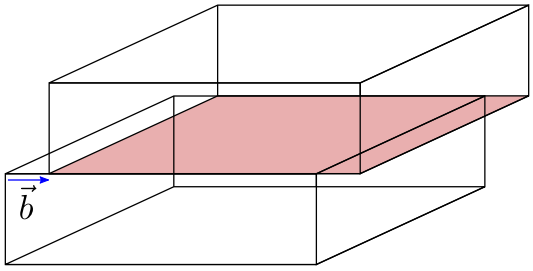


Figure 35.20: Dislocation loop with screw, edge and mixed character at different points along the loop. The direction of the dislocation line, $\vec{s}$, is indicated by the arrows surrounding the dislocation loop.

Exercise: Describe the orientation of the extra half planes for the portions of the dislocation loop in Figure 35.20 that have an edge character.

Solution: For an edge we have $\hat{s} \perp \vec{b}$, which occurs at points a and c. We just need to figure out if the extra half plane is in the top half of the figure or in the bottom half of the figure in each case. The easiest way to do this is to use the fact that $\vec{b} \times \hat{s}$ points in the direction of the extra half plane. At point a $\vec{b} \times \hat{s}$ points up, so the extra half plane is in the top half of the figure. At point c, $\hat{s}$ has reversed, $\vec{b} \times \hat{s}$ points down, and the extra half plane is in the bottom of the figure.

Exercise: What happens to the shape of the crystal in Figure 35.20 if the loop expands and exits the crystal on all sides?
Solution: When a dislocation line moves, it introduces a net displacement of $\vec{b}$ between parts of the crystal on either side of the plane defined by the motion of the dislocation line. So if the dislocation exits the crystal, it will result net translation of $\vec{b}$ of the two parts of the crystal. The trick here is to figure out if the top half moves by an amount $\vec{b}$ or if the bottom half moves by this amount. In this case the top half must be shifted by $\vec{b}$ because the extra half plane at the top of the crystal ends up at the right edge of the crystal, and the extra half plane in the bottom part of the crystal ends up at the left edge. The final situation is as follows:



Exercise: What happens to the shape of the crystal in Figure 35.20 if the loop contracts to nothing and disappears?
Solution: The dislocation just disappears, and a perfect crystal (at least in this region) is recovered.

35.7 Dislocation Density

The following two definitions of the dislocation density are often used:

- Total line length of dislocations per volume.
- The number of intersections that the dislocations make with a plane of unit area.

Both definitions give dislocation densities with units of 1/area, and are equivalent if the dislocations are straight. Typical dislocation densities are as follows:

- A well annealed metal: $10^6 - 10^8/\text{cm}^2$.
- Plastically deformed metal: can be as high as $5x10^{11}/\text{cm}^2$.
- Ceramics: Much lower, typically $10/\text{cm}^2$.
- Si used in microelectronics: dislocation density of zero! Macroscopic single crystals are typically grown without a single dislocation. The down side of this is that Si is very brittle, since there is no plastic deformation mechanism.

35.8 Dislocation Motion

35.8.1 Dislocation Glide

Dislocation glide (which is sometimes referred to simply as slip) corresponds to dislocation motion within a glide plane that contains along the plane that contains both the Burgers vector, $\vec{b}$ and the sense, $\vec{s}$, of the dislocation. For an edge dislocation or a dislocation with mixed edge and screw character, a single slip plane exists that is perpendicular to the vector $\vec{n}_d$, given by the cross product of $\hat{s}$ and $\vec{b}$ (see Figure 35.18). Slip does not require atomic diffusion, and so is not strongly temperature dependent. For an edge dislocation it occurs when the extra half plane of atoms reattaches to a new atomic plane, moving the half plane by a distance equal to $\vec{b}$. The process is illustrated schematically in Figure 35.21.

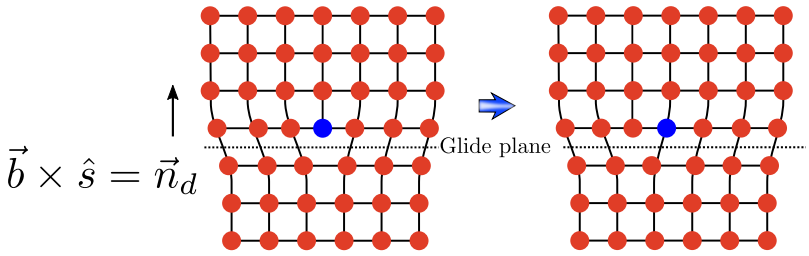


Figure 35.21: Glide of an edge dislocation.

For a pure screw dislocation, because $\hat{s}$ and $\vec{b}$ are collinear, a variety of glide planes are available. As a result, screw dislocations can more easily navigate their way around obstacles (like a precipitate particle) by changing the slip plane on which they are moving. The process is called cross slip and is illustrated schematically in Figure 35.22. This illustration could correspond, for example, to the motion of a screw dislocation with $\hat{s}$ oriented along the $[1\bar{1}0]$ direction that moves along the (111) plane initially, switches to the $(11\bar{1})$ plane and then begins moving again in the (111) plane. (Note - if you forget the Miller index notation for planes and directions, the Wikipedia page [9] is a useful refresher).

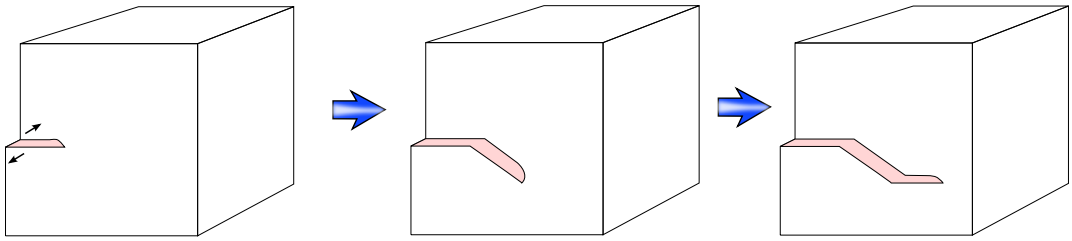


Figure 35.22: Cross slip of a screw dislocation.

35.8.2 Dislocation Climb

Edge dislocations can climb out of the glide plane by the addition or subtraction of vacancies to the dislocation core. The process is illustrated in Figure 35.23 for a situation where $\vec{n}_d$ is directed toward the top of the figure (*i.e.* the extra half plane is above the glide plane). In this example an atom at the end of the extra half plane jumps into a vacancy. The net result is that the vacancy is destroyed, and the dislocation climbs up, away from the initial glide plane. Because the process requires the diffusive hopping of atoms from one site to another, climb is a thermally activated process that becomes more important at elevated temperatures.

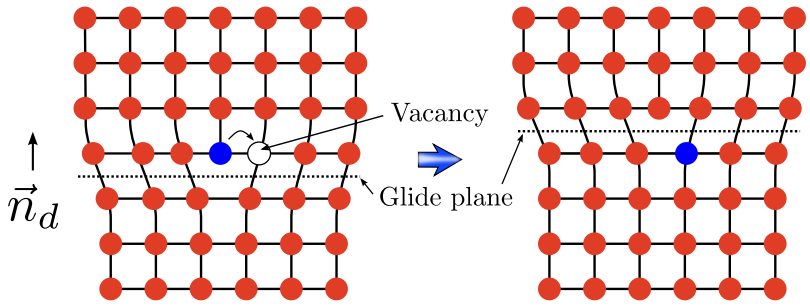


Figure 35.23: Schematic representation of dislocation climb.

If dislocations climb in the direction of $\vec{n}_d$ (in the direction of the extra half plane) as illustrated in Figure 35.23, vacancies are destroyed. If they climb in the other direction (adding atoms to the extra half plane instead of removing them), the opposite occurs and vacancies are created. Dislocation climb therefore provides an mechanisms for equilibrating the vacancy concentration. For metals it is the process that allows us to assume that the vacancy concentration remains at equilibrium, and important assumption of our analysis of the Kirkendall experiment in Section 34.4.

35.9 Dislocation Energy

Dislocation disrupts the regularity of the lattice, and introduces strain into the sample. The strain field that results from the presence of a dislocation has a very long range, and can easily be more than 100 times the unit cell dimension. As a result the total strain energy is very large as well. This strain field and the energy associated with it is important because it provides a mechanism for dislocations to interact with one another over long distances. In essence, dislocations 'talk' to each other through these strain fields.

35.9.1 Screw Dislocations

For a screw dislocation we can use some simple concepts to calculate this strain energy, so we'll start with this example. Our starting point is that the material surrounding a screw dislocation is in a state of pure shear, with shear deformation as defined in Figure 35.4. We see this by considering a cylindrical portion of the material around a screw dislocation, using the illustration in Figure 35.24. The displacement applied across the dislocation is given by the Burgers vector, $\vec{b}$ (Figure 35.24a). (When referring to the magnitude of the Burgers vector we'll drop the vector symbol and just use the b). If we unwrap the circumference of the cylinder at a distance r from the dislocation line (Figure 35.24b) we see that the shear displacement of b is applied over a distance of $2\pi r$. The cylinder has been unwrapped in the circumferential direction, *i.e.* the θ direction, and the displacement is along the z direction so we have the following for th shear strain, $e_{z\theta}$:

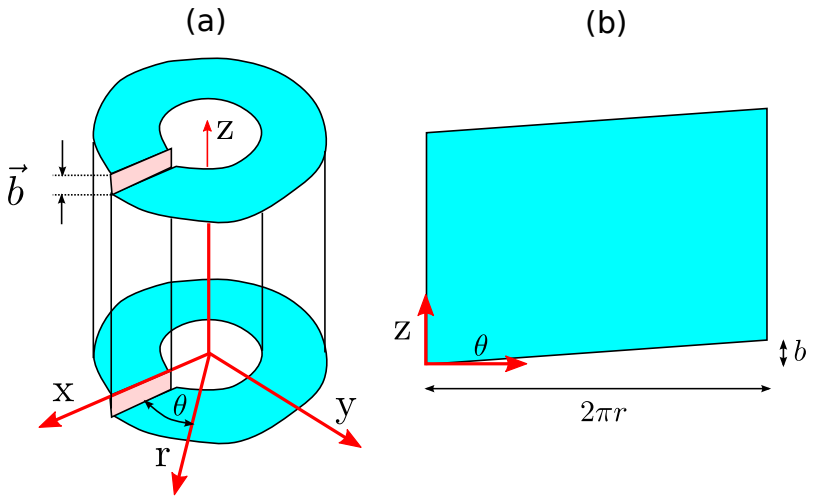


Figure 35.24: Strain field around a screw dislocation that is centered along the z axis.

The distortion is pure shear, with a shear strain at a radius of r given by the following:

$$e_{z\theta} = \frac{b}{2\pi r} \tag{35.9}$$

Note that as $r \to 0$, $e_{z\theta} \to \infty$. The strain can't really go to infinity, so we have a problem here that we're going to have to deal with eventually. The elastic stress is obtained by multiplying by the shear modulus:

$$\tau_{z\theta} = \frac{Gb}{2\pi r} \tag{35.10}$$

The elastic strain energy per unit volume, E_v is obtained from the following expression:

$$E_v = \int_0^{e_{r\theta}} \tau_{z\theta} de_{z\theta} = \frac{G}{2} e_{z\theta}^2 \tag{35.11}$$

Dimensionally this makes sense, since G has units of force/area, or energy/volume.
We can combine Eqs. 35.9 and 35.11 to obtain the following:

$$E_v = \frac{G}{2}\left(\frac{b}{2\pi r}\right)^2 \tag{35.12}$$

Because the strain energy is radially symmetric, we can get the energy per unit length of the dislocation (E_s/h) by integrating all values of r. Because the area element in radial coordinates is $2\pi r dr$, we have:

$$E_s/h = 2\pi \int_0^{r_{max}} E_v\,(r)\,dr \tag{35.13}$$

Substituting E_v from Eq. 35.12 into this equation gives:

$$E_s = \frac{Gb^2h}{4\pi}\int_0^{r_{max}} \frac{1}{r} dr \tag{35.14}$$

After integration we obtain:

$$E_s = \frac{Gb^2h}{4\pi}[\ln r_{max} - \ln 0] \tag{35.15}$$

We have a problem here because $\ln 0 = -\infty$. This is because for $r \to 0$, the shear strain goes to infinity and we can no longer use the simple, continuum picture of linear elasticity to describe what is going on. Instead what we typically do is separate out a core energy, E_s^{core} that corresponds to the strain energy inside some small core radius, r_0. We do this simply be adding E_s^{core} to E_s and replacing the lower bound on the integration from 0 to r_o.

$$E_s = E_s^{core} + \frac{Gb^2h}{4\pi}\ln\left(\frac{r_{max}}{r_0}\right) \tag{35.16}$$

We can generally choose r_0 so that it is large enough so that our assumption of linear elasticity holds for $r > r_0$, yet it is small enough so that E_s^{core} is a relatively small fraction of the overall dislocation energy. In this case we can ignore the core energy and approximate the dislocation energy as follows:

$$E_s \approx \frac{Gb^2h}{4\pi}\ln\left(\frac{r_{max}}{r_0}\right) \tag{35.17}$$

35.9.2 Edge Dislocations

The stress field for an this case is much more complicated, as illustrated in Figure 35.25. The distinctive features of the strain field are as follows:

- In the slip plane itself the material is in a state of pure shear.
- Above the slip plane there is compressive component to the strain field.
- Below the slip plane there is a tensile component to the strain field.

A more detailed calculation shows that the strain still decays as $1/r$, with an expression for the edge dislocation energy per length that is similar to the expression obtained for a screw dislocation:

$$E_s \approx \frac{Gb^2h}{4\pi\,(1-\nu)}\ln\left(\frac{r_{max}}{r_0}\right) \tag{35.18}$$

Here, ν is Poisson's ratio. Typically ≈ 0.3 for most metals.

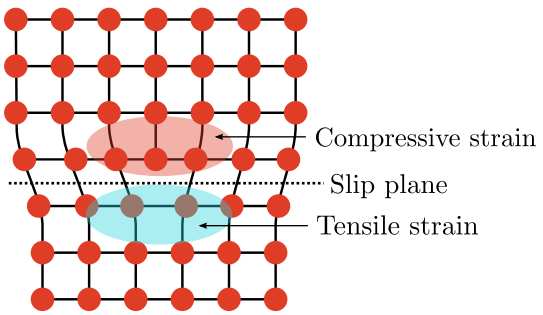


Figure 35.25: Schematic representation of an edge dislocation, showing the regions of tensile and compressive strain.

The conceptual picture shown in Figure 35.25 is useful, but we can do a little bit better by reminding ourselves of some definitions pertaining to a stress state. A two-dimensional stress state in the x-y plane has three independent components of the stress: the shear stress, τ, and two normal stresses, σ_{xx} and σ_{xy}, as shown in Figure 91.1. In Figure 35.27a the regions around an edge dislocation where stress components have different signs are illustrated. Figure 35.27b has similar information, but in this case we plot contours of equal stress for each of the three stress components, σ_{xx}, σ_{yy} and σ_{xy}.

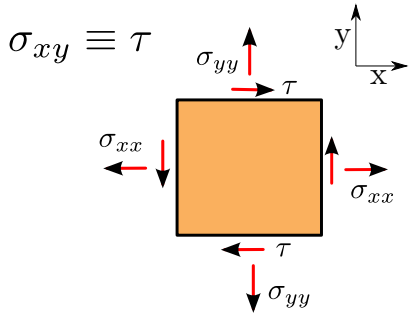


Figure 35.26: 2-dimensional stress tensor

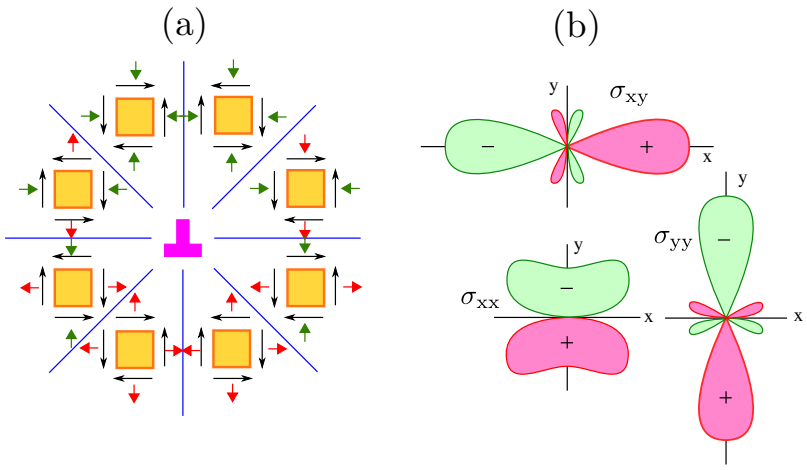


Figure 35.27: Stress distribution around an edge dislocation.

35.9.3 General Comments

The following general comments are valid for edge, screw and mixed dislocations:

1. Elastic strain energy scales with $\ln r$ so it has a very long range.
2. The boundary conditions matter, so the energy depends on the shape of the sample. A small crystal with low value of r_{max} will have a lower dislocation energy than a large crystal with a very large value of r_{max}.
3. Energy scales as b^2. Dislocations with small values of b are therefore preferred, which is why the Burgers vector in a material corresponds to the smallest interatomic spacing in the material.
4. Energy scales as h. Energy is proportional to the length of the dislocation. This means the strain energy will decreases as the line length decreases.

This last point seems trivial at first, but it has some important consequences. Consider for example a dislocation loop. If the radius of a circular loop decreases, the energy associated with the loop will decrease as well. There's a line tension acting on the loop causing it to contract. This tension is like the tension in a rubber band that once to squeeze things inward, and can be viewed as a driving force for the dislocation loop to shrink in size. An applied stress can cause a dislocation loop to grow instead of shrink, and this will be considered later.

Let's compare some numbers to see how the dislocation energy compares to the energy of other defects, like vacancies, for example. To do this we'll estimate the energy per atomic length along a screw dislocation line by taking $h = b$ in Eq. 35.17. We'll also assume typical values for the other parameters in the expression for E_s, with b=0.3 nm, $r_0 = 1$nm. $r_{max} = 1\,\mu m$, and $G = 3x10^{10}$ Pa:

$$E_s \approx \frac{Gb^3}{4\pi}\ln\left(\frac{r_{max}}{r_0}\right) = \frac{(3x10^{10}\,\text{Pa})\,(3x10^{-10}\,\text{m})^3}{4ß}\ln\left(\frac{1000\,\text{nm}}{1\,\text{nm}}\right) = 4.5x10^{-19}\,\text{J} \tag{35.19}$$

A more convenient energy scale on an atomic basis is the electron volt, which we obtain from the energy in Joules by dividing by the electron charge ($1.6x10^{-19}$ C). In these units the energy per atom along the dislocation line is 2.8 eV. This energy of comparable magnitude to typical vacancy formation energies of $\approx$ 1eV, but is actually larger because of the nature of the long-range strain field that is produced around a dislocation.

35.10 Dislocation Line Tension

An energy per unit length has dimensions of a force. The dislocation energy per unit length is therefore equivalent to a force, or tension, exerted by the dislocation. We refer to this line tension as T_s:

$$T_s = E_s/h \tag{35.20}$$

This line tension is a one dimensional analog of the interfacial free energy, with units of energy per length instead of energy per area. The comparison is summarized in Table 35.2.

Table 35.2: Comparison of line tension and the interfacial free energy.

Quantity	Energy units	Force Units
T_s	J/m	N
γ	J/m^2	N/m

The line tension itself is a force, and it gives rise to a force per unit length acting perpendicular to a curved dislocation line, in the same way that the interfacial free energy results in a pressure difference (force per unit area) across a curved surface. The work done against this one dimensional pressure, which we refer to as F_s^r is equal to the increase in free energy associated with the increased length of the dislocation line. By considering the graphical construction shown in Figure 35.28:

$$F_s^r 2\pi r dr = T_s 2\pi dr \tag{35.21}$$

Rearrangement gives the following expression for F_s^r:

$$F_s^r = \frac{T_s}{r} \tag{35.22}$$

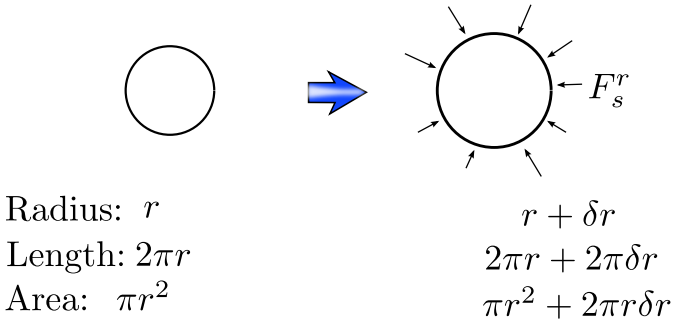


Figure 35.28: Relationship between the curvature and the the one-dimensional pressure acting on a dislocation line.

35.11 Effect of Applied Stress

The line tension acts to decrease the area of a dislocation loop, but we need loops to expand in order for a material to plastically deform. So how does an applied stress induce a force on a dislocation? An applied shear stress, τ results in an additional force per unit length, F_s^τ. It's easiest to visualize the relationship between τ and F_s^τ for an edge dislocation, as we illustrate in Figure 35.30. To do this we use an energy balance. When the dislocation has propagated across the entire sample a total applied shear force, P, results in a net translation of the material above the slip plane by an amount given by the Burgers vector, $\vec{b}$. The total work put into the system is simply Pb (force times displacement). With $\tau = P/w\ell$ this total work is:

$$work = w\ell\tau b \tag{35.23}$$

This work goes into moving the dislocation, and must be equal to the force applied to the dislocation multiplied by the distance the dislocation moves as it translates across the sample. In our notation this distance is the sample width, w, so we have:

$$work = F_s^\tau \ell w \tag{35.24}$$

Equating these two expressions for the work gives:

$$F_s^\tau = \tau b \tag{35.25}$$

So the force per unit length acting on the dislocation is simply the shear stress multiplied by the magnitude of the Burgers vector.

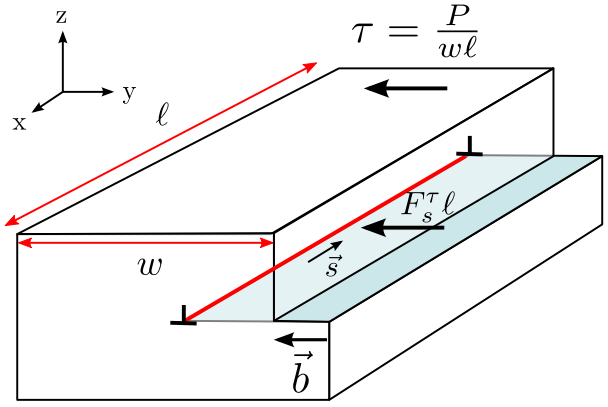


Figure 35.29: Force force acting on a dislocation.

The only real assumption in Eq. 35.25 is that τ is the component of the shear stress oriented along the direction of the Burgers vector. The resulting force is perpendicular to the dislocation line itself, regardless of the specific orientation of the dislocation line. This point as an important one that is not completely obvious, so we illustrate it for a screw dislocation in Figure 35.30. The orientation of the Burgers vector is identical to that of the Burgers vector for the edge dislocation in Figure 35.29, but the dislocation is now a screw dislocation oriented along the y direction that propagates in the negative x direction as the dislocation moves through the crystal. Because the final state of the crystal is the same as for the edge dislocation in Figure 35.29, the work done by the applied stress is still given by $w\ell\tau b$, *i.e.* Eq. 35.23 still applies. The dislocation moves a distance ℓ in this case. Because the length of the dislocation is w, the total force applied to the dislocation is $F_s^\tau w$, and the energy required to translate it by a distance ℓ is $F_s^\tau w\ell$. So we see that Eq. 35.24 still applies as well. The net result is that the F_s^τ is still given by τb, just as it was for an edge dislocation. It can be shown that the same must be true for a mixed dislocation as well.

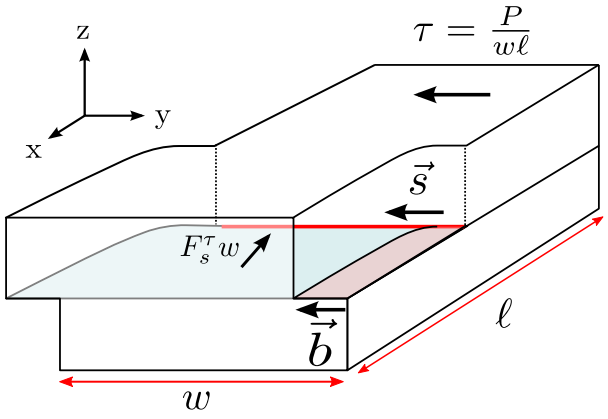


Figure 35.30: Force force acting on a dislocation.

Now we can look at the stress required to expand a circular dislocation loop. We'll assume that the energy/length of the dislocation is a constant. In other words, we are neglecting the factor of $1-\nu$ in Eq. 35.18 that gives a small energy difference between edge and screw dislocations. The total force per unit length acting on a circular dislocation loop is the sum of F_s^r, which acts toward the center of the loop and therefore negative, and F_s^τ, which for an appropriately aligned shear stress is positive:

$$F_s = F_s^\tau - F_s^r = \tau b - \frac{T_s}{r} \tag{35.26}$$

At equilibrium the net force acting on the dislocation is zero ($F_s = 0$). This occurs when the applied stress is equal to a critical value that we refer to as τ^*:

$$\tau^* = \frac{T_s}{rb} \tag{35.27}$$

If $\tau > \tau^*$ the dislocation loop expands, and if $\tau < \tau^*$ the dislocation shrinks and disappears altogether.
So why do precipitates strengthen a material? The answer is connected to Eq. 35.27. Consider a dislocation that is moving toward two precipitates. The applied stress results in a force per unit length, F_s^τ that moves the dislocation. The pinning of the dislocation between the precipitates results in a curvature, r, with an associated stress τ^* that must be applied in order for the dislocation to move. The maximum value of τ^* corresponds to the minimum value of the dislocation curvature r, which is equal to half the interparticle spacing. In this example, τ^* is the critical resolved shear stress for the material. For optimum strengthening, what we want very small precipitates with a correspondingly small interparticle spacing.

35.12 Dislocation Multiplication

Where do these dislocations come from in the first place? Shape change associated with the emergence of dislocation to the exterior of the crystal must be decreasing their density. A typical dislocation density of $\approx 10^7/\text{cm}^2$ is way too small to give the experimentally measured plastic strain observed in a typical metal. So there must be some mechanism of creating new dislocations. One possibility we can consider is that the applied stress is itself sufficient to nucleate a dislocation loop. To figure out if this makes sense, we can calculate the shear stress required to expand a relatively small dislocation loop with a radius, r, of $10b$. We'll assume $r_0 = b$ and $r_{max} = 10b$ and estimate the dislocation line tension From Eqs. 35.17 and 35.20:

$$T_s \approx \frac{Gb^2}{4\pi}\ln(10) \approx \frac{Gb^2}{8} \tag{35.28}$$

The resolved shear stress, τ^*, required to expand the loop is given by Eq. 35.27:

$$\tau^* = \frac{T_s}{rb} = \frac{G}{80} \tag{35.29}$$

Actual values of the critical resolved shear stress are $\approx G/10^4$ (see Table 35.1), so there must be some other mechanism operating at a lower stress that enables new dislocations to be created. This mechanism is the Frank-Read source.

The process by which new dislocations are produced by a Frank-Read source is illustrated in Figure 35.31. It is based on the behavior of a dislocation segment that is pinned between two points (precipitate particles for example), labeled A and B in Figure 35.31. In the absence of an applied shear stress, this dislocation is a straight line between points A and B, (line 1 in Figure 35.31). As a shear stress is applied to the material the dislocation expands outward in a series of arcs, labeled as 2, 3, 4 and 5 in Figure 35.31. Because F_s^τ is always acting normal to the dislocation line, pushing it outward, the dislocation bends around the pinning points. Eventually, two segments of the dislocation with opposite $\hat{s}$ are in close proximity to each other (arc 4 in Figure 35.31). These segments of the dislocation annihilate each other, and the dislocation breaks into two separate arcs, both of which are now labeled as 5. The larger of these arcs is a dislocation loop that continues to expand, and the smaller of the arcs repeats the process as it expands in response to the stress. In this way an unlimited number of dislocation loops can be created by the original segment of the dislocation.

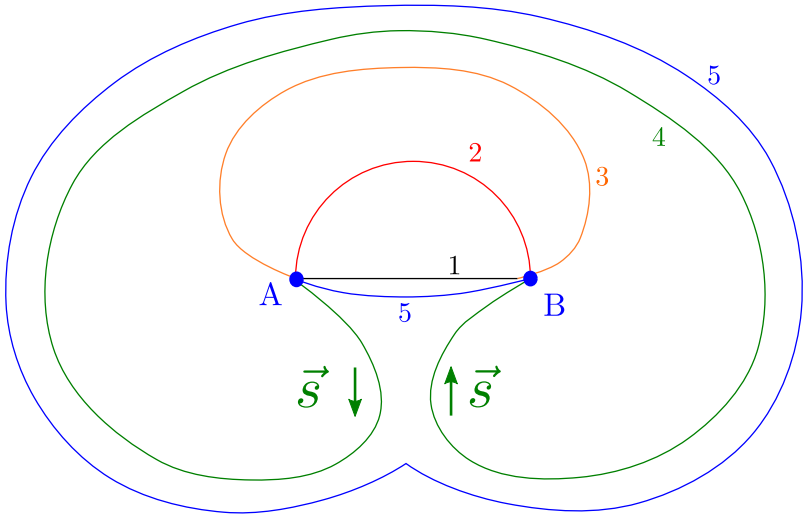


Figure 35.31: Schematic representation of a Frank-Read dislocation source. Note that all sections of all of the dislocations have the same Burgers vector.

We can also use this argument to obtain values for the critical resolved shear stress in the system. Because the shear stress needed to expand the dislocation is inversely proportional to the dislocation radius of curvature, r (from Eq. 35.27), the largest stress corresponds to the smallest radius of curvature for the dislocation line. In its original unstressed configuration (line 1 in Figure 35.31), the dislocation is a straight line, with $r = \infty$. Then the radius of curvature decreases as the dislocation begins to grown in response to the applied stress. The minimum radius of curvature is $d/2$, where d is the distance between the pinning points of the dislocation. This corresponds to line 1 in Figure 35.31. This corresponds to the maximum applied stress, which for $r = d/2$ is $2T_s/db$. This is the critical resolved shear stress, τ_{CRSS}, for the system, if dislocation pinning is the strengthening mechanism in the material. If we estimate T_s as $Gb^2/8$ as we did above, we obtain:

$$\tau_{crss} \approx \frac{Gb}{4d} \tag{35.30}$$

Precipitation strengthening of a material is based on the introduction of very closely spaced nano-scale precipitates, giving the smallest possible value of d, and hence the maximum τ_{CRSS}.

35.13 Interactions Between Dislocations

Dislocations interact through the long range strain fields induced by the dislocations. Consider two screw dislocations with Burgers vectors $\vec{b}_1$ and $\vec{b}_2$ that separated by d. The stress at dislocation 2 due to the presence of dislocation 1 is given by Eq.35.10:

$$\vec{\tau} = \frac{G\vec{b}}{2\pi d} \tag{35.31}$$

We use a vector notation for the stress here to remind us that τ the force associated with the shear stress is directed along $\vec{b}$.

This stress induces a force on the dislocation, given by Eq. 35.25:

$$F_s^\tau = \frac{G\vec{b}_2 \cdot \vec{b}_1}{2\pi d} \tag{35.32}$$

If $\vec{b}_1$ and $\vec{b}_2$ are pointing in the same direction, the force is positive and the interaction is repulsive. If they are pointing in the opposite direction, the force is negative (attractive). Because the force scales as $1/d$ it has a very long range.
We can also make a qualitative argument based on the energetics to see what is going on. From Eq. 35.32 we see that $E \propto b^2$. If the dislocations have opposite signs, the dislocation come together and the net b is zero. If they have the same sign, then they form a total Burgers vector with a magnitude of $2b$, and an energy of $4b^2$. So the energy is twice as large as it was originally. The energy as a function of separation must look the plot shown in Figure 35.32. If the energy of each individual dislocation is E_1 for $d \to \infty$, then the total dislocation energy at $d = 0$ is zero for dislocations of opposite sign and $4E_1$ for dislocations of the same sign.

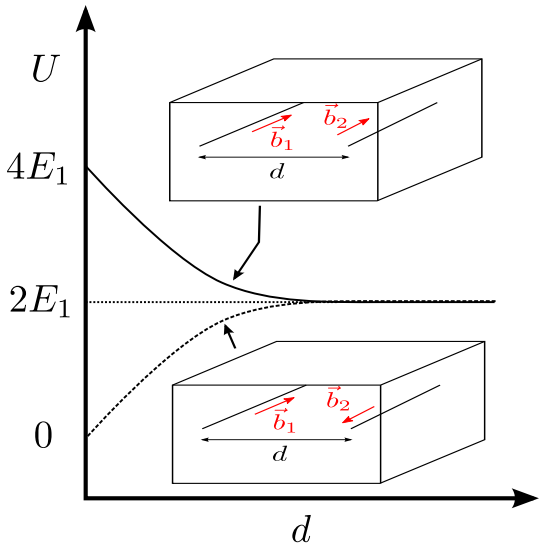


Figure 35.32: Interactions between screw dislocations of the same sign (solid line) and opposite sign(dashed line).

Screw dislocations are easy to think about because the stress field is axially symmetric about the dislocation line, and the stress field is always pure shear. We already know from our previous discussion of stress fields that edge dislocations are more complicated. A simple limiting case involves two edge dislocations on the same slip plane, since within the slip plane we are in a state of pure shear. In this case the discussion from the previous section is still valid, and we get the following expression for the interaction force:

$$F_s^\tau = \frac{G\vec{b}_2 \cdot \vec{b}_1}{2\pi\,(1-\nu)\,d} \tag{35.33}$$

The expression is the same as the expression for the screw dislocations, with the extra factor of $1 - \nu$ that also appears in the expression for the dislocation energy for an edge dislocation (see Eq. 35.18).
If the two edge dislocations do not lie on the same slip plane the situation is more complicated, but we can still make some qualitative arguments based on the nature of the strain fields around the dislocations. Consider two edge dislocations on glide planes that are separated by a distance h, as illustrated in Figure 35.33. In this case the Dislocations begin to line up due to cancellation of the strains in the regions where these strain field overlap. The tensile strain induced below the upper dislocation is partially compensated for by the compressive strain above the lower dislocation. As a result the dislocations move along their respective glide planes so that they lie on top of one another.

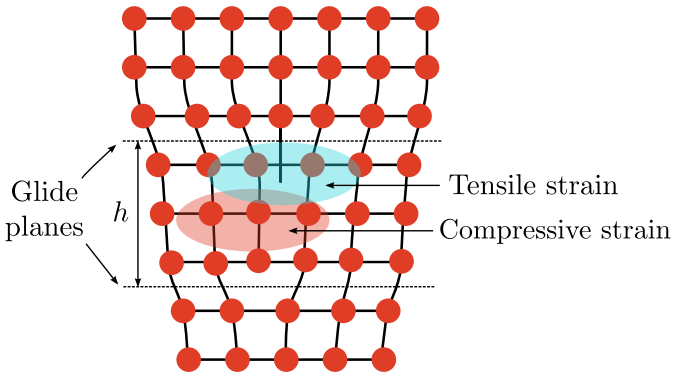


Figure 35.33: Overlapping stress fields for two adjacent dislocations of the same sign.

Dislocations with the same sign repel each other when they are in the same glide plane (see Figure 35.32), but they move toward each other so that they line up on top of one another when they are in different glide planes (see Figure 35.33). The overall situation is summarized in Figure 35.34, which shows the regions of attractive and repulsive interactions for the dislocations.

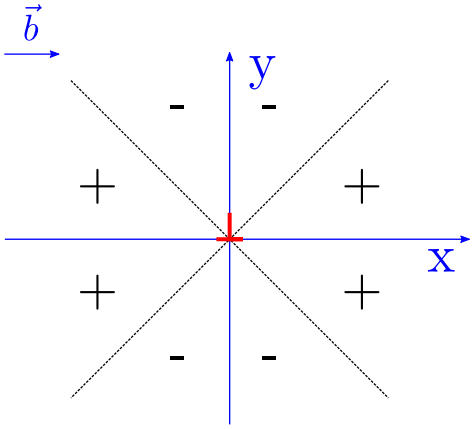


Figure 35.34: Map showing the regimes where two edge dislocations move toward one another with attractive interaction (-) or apart from one another with a repulsive interaction (+). One of the dislocations is at the origin, with $\hat{s}$ pointing into the plane of the figure. The other dislocation is assumed to have the same value for $\hat{s}$ and $\vec{b}$. Both dislocations move in glide planes that are perpendicular to the y axis. These dislocations move toward each other if the second dislocation is located within a region that is attractive (-), and apart from one another if the second dislocation is located in a region that is repulsive (+).

Dislocations can also interact with point defects or other atoms. The reason for this is that point defects also distort the lattice, giving strain fields that overlap with the strain field of a dislocation. This effect is easiest to visualize for edge dislocations, as illustrated in Figure 35.35. Substitutional impurities with different sizes than the majority component atoms can reduce the strain fields surrounding the dislocation by moving to the appropriate region near the dislocation. Large atoms will tend to substitute for atoms in regions where the local strain is tensile, and small atoms will tend to substitute for atoms in regions where the local strain is compressive. Interstitial impurities will segregate to regions of tensile stress as well.
As a result of this impurity segregation, dislocations collect 'clouds' of impurity atoms. For a dislocation to move, it must either break away from this atmosphere or move the impurity atoms along with it. The net result in either case is an increase in the critical resolved shear stress, a process referred to as solid solution strengthening.

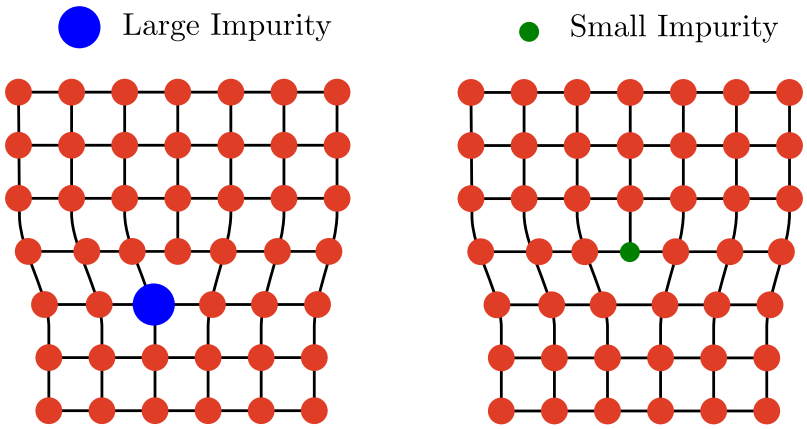


Figure 35.35: Favored locations the segregation of substitutional impurities to the core of an edge dislocation.

36 Thermodynamics of interfaces

36.1 A Brief Review of the Thermodynamic Potentials

A statement of the combined first and second laws of thermodynamics is that the internal energy, U, is minimized at equilibrium under conditions of fixed temperature and entropy. We more commonly fix the temperature instead of the entropy. For this reason we o closely related thermodynamic potentials: the Helmhtoltz free energy, F, and the Gibbs free energy, G.

36.1.1 Helmholtz Free Energy

The Helmholtz free energy, F, is defined in the following way:

$$F \equiv U - TS \tag{36.1}$$

The differential of F is:

$$\delta F = \delta U - T\delta S - S\delta T \tag{36.2}$$

We already an expression for δU (see Eq. 36.14):

$$\delta U = T\delta S - PdV + \sum_i \mu_i \delta n_i \tag{36.3}$$

Substituting this value for δU into Eq. 36.2 for δF gives:

$$\delta F = -PdV - S\delta T + \sum_i \mu_i \delta n_i \tag{36.4}$$

So that $\delta F = 0$ for fixed V, T, n_i. This means that at equilibrium under conditions of fixed temperature and volume, the Helmholtz free energy is minimized.

36.1.2 Gibbs Free Energy

The Gibbs free energy is defined in a very similar manner, but in this case we replace the internal energy, U, with the enthalpy, H.

$$G \equiv H - TS \tag{36.5}$$

The enthalpy is defined as:

$$H = U + PV \tag{36.6}$$

By comparing these equations to Eq. 36.1, we see that all we've really done is add a PV term to the Helmholtz free energy:

$$G = F + PV \tag{36.7}$$

In differential form

$$\delta G = \delta F + P\delta V + V\delta P \tag{36.8}$$

Using Eq. 36.4 for δF gives:

$$\delta G = V\delta P - S\delta T + \sum_i \mu_i \delta n_i \tag{36.9}$$

So that $\delta G = 0$ for fixed P, T, n_i. This means that at equilibrium under conditions of fixed temperature and volume, the Gibbs free energy is minimized.

36.2 Temperature and Chemical Potential Equilibrium

Interfaces have an energy associated with them. This is easiest to see in the case where there is a big structural change across the interface (a solid-vapor interface, for example). In the simple example illustrated in Figure 36.1 the atoms at the surface have fewer bonds than the atoms in the bulk of the material. The lower number of bonds implies that there is an excess energy associated with atoms near the surface. In the simple nearest neighbor picture only those atoms at the surface are affected. In most cases, however, many atoms near the surface are affected, especially in cases where the density and/or structure of the phases are very similar. For example, in a liquid/liquid system like the interface between oil and water, structural changes across the interface are more subtle, and the interface can be very wide on an atomic scale. If we plot the concentration of one of the components across the interface between α and β phases as shown schematically in Figure 36.2, we see that it transitions from C^α to C^β over an interfacial region that can in some cases be many atomic dimensions wide.

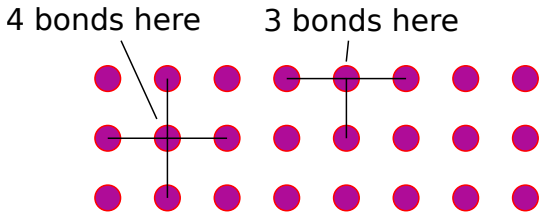


Figure 36.1: Reduced bonding at the free surface of a crystalline material.

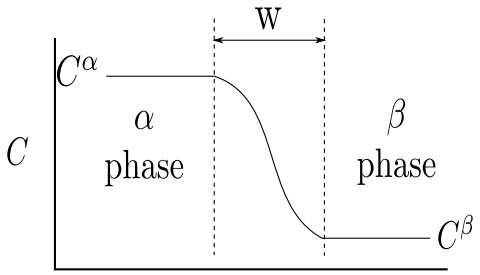


Figure 36.2: Interfacial profile between two phases (in this case a solid and liquid), illustrating the existence of an interfacial region of width w.

The change in density across the interface means that the energy in the transition zone is different than the energy in either of the bulk phases. Even if the structure is the same, for example a coherent interface between alloys of the same crystal structure, the change in composition across the interface will lead to a region of the material with a different energy.

How can we develop a generalized description of interfacial thermodynamics that is valid for all types of interface (crystalline, amorphous, narrow, broad, etc.)? Fortunately, this was done by Gibbs even before atoms were discovered (c. 1880)! Our basic assumption is that all quantities vary across the interface in a continuous manner, like

the density plot shown in Figure 36.2. We need to develop the corresponding condition for the inhomogeneous interfacial region of finite width. We begin by considering the interface between two phases, α and β. As shown in Figure 36.3 we can separate the system in to three regions: α and β bulk phase regions where the properties are completely uniform, and an interfacial region I, where the properties (S, n_i, etc. are non-uniform).

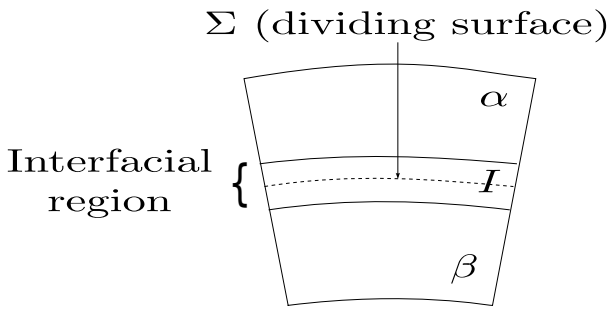


Figure 36.3: Formal representation of the interfacial region between α and β phases.

We begin by recognizing that the total energy is the sum of the energies of the three different regions:

$$U = U^{\alpha} + U^{\beta} + U^{I} \tag{36.10}$$

This same summation holds for the total entropy and for the number of atoms of each component in the system:

$$S = S^{\alpha} + S^{\beta} + S^{I} \tag{36.11}$$

$$n_i = n_i^{\alpha} + n_i^{\beta} + n_i^{I} \tag{36.12}$$

So far we haven't invoked any thermodynamic principles. We just divided the entropy, energy into portions that reside in each of the three zones of Figure 36.3 (α, β and I). Now we invoke the second law of thermodynamics which for us we word in the following way:
At equilibrium we have $\delta U = 0$ (energy is minimized) at fixed S, n_i, V:

$$\delta U = \delta U^{\alpha} + \delta U^{\beta} + \delta U^{I} = 0 \tag{36.13}$$

Each of the terms in δU can be expressed in a sum involving partial derivatives with respect to S and n_i and V. For example, for δU^{α} we have:

$$\delta U^{\alpha} = \frac{\partial U^{\alpha}}{\partial S^{\alpha}} \delta S^{\alpha} + \frac{\partial U^{\alpha}}{\partial V^{\alpha}} \delta V^{\alpha} + \sum_i \frac{\partial U^{\alpha}}{\partial n_i} \delta n_i \tag{36.14}$$

We can also use the following expressions for the chemical potentials and the temperature:

$$T^{\alpha} \equiv \frac{\partial U^{\alpha}}{\partial S^{\alpha}} \tag{36.15}$$

$$\mu_i^{\alpha} \equiv \frac{\partial U^{\alpha}}{\partial n_i^{\alpha}} \tag{36.16}$$

$$P^{\alpha} \equiv -\frac{\partial U^{\alpha}}{\partial V} \tag{36.17}$$

We'll fix the boundaries between the α, β and I region in Figure 36.3 so that $\delta V^{\alpha} = \delta V^{\beta} = \delta V^{I} = 0$ and the volume terms in Eq. 36.14 disappear. With this assumption, and the definitions of the temperature (Eq. 36.15) and chemical potential (Eq. 36.16), we can rewrite Eq. 36.14 as follows:

$$\delta U^{\alpha} = T^{\alpha} \delta S^{\alpha} + \sum_i \mu_i^{\alpha} \delta n_i \tag{36.18}$$

Exactly analogous expressions can be written down for δU^{β} and δU^{I}:

$$\delta U^{\beta} = T^{\beta}\delta S^{\beta} + \sum_i \mu_i^{\beta}\delta n_i \tag{36.19}$$

$$\delta U^{I} = T^{I}\delta S^{I} + \sum_i \mu_i^{I}\delta n_i \tag{36.20}$$

We can now substitute these expressions for δU^{α}, δU^{β} and δU^{I} into Eq. 36.13 to obtain the following:

$$T^{\alpha}\delta S^{\alpha} + T^{\beta}\delta S^{\beta} + T^{I}\delta S^{I} + \sum_i \mu_i^{\alpha}\delta n_i + \sum_i \mu_i^{\beta}\delta n_i^{\beta} + \sum_i \mu_i^{I}\delta n_i^{\beta} = 0 \tag{36.21}$$

Now we can consider the constraints on the system: $\delta S = 0$, $\delta n_i = 0$ (a closed thermodynamic system at equilibrium). In our case these constraints can be written as follows:

$$\delta S = \delta S^{\alpha} + \delta S^{\beta} + \delta S^{I} = 0 \tag{36.22}$$

$$\delta n_i = \delta n_i^{\alpha} + \delta n_i^{\beta} + \delta n_i^{I} = 0 \tag{36.23}$$

We can use these two equations to eliminate δn_i^{α} and δS^{α} from Eq. 36.21 by substituting $-\delta S^{I} - \delta S^{\beta}$ for δS^{α} and $-\delta n_i^{I} - \delta n_i^{\beta}$ for δn_i^{α}:

$$T^{\alpha}\left(-\delta S^{I} - \delta S^{\beta}\right) + T^{\beta}\delta S^{\beta} + T^{I}\delta S^{I} + \sum_i \mu_i^{\alpha}\left(-\delta n_i^{I} - \delta n_i^{\beta}\right) + \sum_i \mu_i^{\beta}\delta n_i + \sum_i \mu_i^{I}\delta n_i = 0 \tag{36.24}$$

This expression can be rearranged to the following:

$$\delta S^{\beta}\left(T^{\beta} - T^{\alpha}\right) + \delta S^{I}\left(T^{I} - T^{\alpha}\right) + \sum_i \left(\mu_i^{\beta} - \mu_i^{\alpha}\right)\delta n_i^{\beta} + \sum_i \left(\mu_i^{I} - \mu_i^{\alpha}\right)\delta n_i^{I} = 0 \tag{36.25}$$

All the terms in brackets must be zero, since this must hold true for any variation in S^{I}, S^{β}, n_i^{I}, n_i^{β}. This implies $T^{\alpha} = T^{\beta} = T^{I}$ (uniform temperature) and that the chemical potentials are uniform everywhere.

36.3 The Dividing Surface

At a planar interface, we still get the usual thermodynamic condition that the temperature and chemical potentials are uniform everywhere at equilibrium. But what about pressure effects? What if $\delta V^{\alpha} \neq 0$? To answer these questions we will make a relatively large conceptual leap and replace the real system where the interfacial has some finite width with an equivalent model system where the the interface is a true surface with no volume. This model system is obtained by extending bulk phase properties all the way up to the fictitious location of the dividing surface, Σ, where we have the two phases α and β in our example directly in contact with one another.

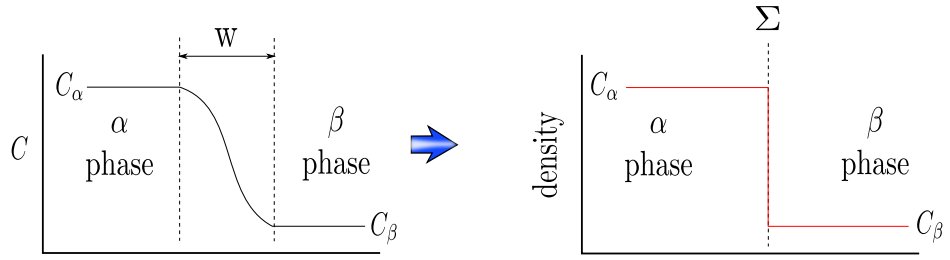


Figure 36.4: Conceptual representation of the replacement of a continuous concentration gradient a single dividing surface, Σ, that has no width.

Once we specify the precise location of the dividing surface we can determine the number of atoms that are associated with the th interface. Once we know where the dividing surface is, we also know the volumes of each phase, V^{α} and V^{β}. Multiplying by the bulk phase concentration gives the total number of atoms in each phase:

$$\begin{aligned} n_i^{\alpha} &= C_i^{\alpha} V^{\alpha} \\ n_i^{\beta} &= C_i^{\beta} V^{\beta} \end{aligned} \tag{36.26}$$

In general, the total number of atoms of component i, n_i, is not equal to $n_i^\alpha + n_i^\beta$. The excess is associated with the interfaces, and is referred to as n_i^Σ:

$$n_i^\Sigma = n_i - n_i^\alpha - n_i^\beta \tag{36.27}$$

We commonly divide by the interfacial area, A, to get an interfacial excess of component i per area, which we define as Γ_i:

$$\Gamma_i \equiv n_i^\Sigma / A \tag{36.28}$$

We can also define an interfacial energy (U^Σ) and an interfacial entropy (S^Σ) in a similar way:

$$U^\Sigma = U - U^\alpha - U^\beta \tag{36.29}$$

$$S^\Sigma = S - S^\alpha - S^\beta \tag{36.30}$$

Note that is is generally fairly obvious where we should put the dividing surface (at least approximately), although the detailed location we choose will affect the exact values of Γ_i, U^Σ and S^Σ.
We now proceed in a fashion that is similar to what we did in the previous section. The expressions for δU^α and δU^β are the same as Eqs. 36.18 and 36.19, but we will no longer insist that δV^α and δV^β are zero:

$$\delta U^\alpha = T^\alpha \delta S^\alpha - P^\alpha dV^\alpha + \sum \mu_i^\alpha \delta n_i^\alpha \tag{36.31}$$

$$\delta U^\beta = T^\beta \delta S^\beta - P^\beta dV^\beta + \sum \mu_i^\beta \delta n_i^\beta \tag{36.32}$$

The interface is different from the α and β phases in that it has no volume, but it has an area, A. In the expression for δU^Σ, the differential of the energy associated with the interface, we replace the partial derivative term involving δV with one involving δA. The version of Eq. 36.14 that holds for the interface can therefore be written as follows:

$$\delta U^\Sigma = \frac{\partial U^\Sigma}{\partial S^\Sigma}\delta S^\Sigma + \sum \frac{\partial U^\Sigma}{\partial n_i^\Sigma}\delta n_i^\Sigma + \frac{\partial U^\Sigma}{\partial A^\Sigma}\delta A^\Sigma \tag{36.33}$$

The partial derivative of the energy with respect to area that appears in Eq. 36.33 is defined as $U_a^{\alpha\beta}$ the interfacial energy between the α and β phases:

$$U_a^{\alpha\beta} = \frac{\partial U^\Sigma}{\partial A^\Sigma} \tag{36.34}$$

We can also define an interfacial entropy, $S_a^{\alpha\beta}$ in a similar way:

$$S_a^{\alpha\beta} = \frac{\partial S^\Sigma}{\partial A^\Sigma} \tag{36.35}$$

We can also define an interfacial free energies, in a way that is analogous to the definitions given in Section 36.1. Defining Gibbs and Helmholtz free versions of the interfacial quantities gives the following:

$$F_a^{\alpha\beta} = U_a^{\alpha\beta} - TS_a^{\alpha\beta} \tag{36.36}$$

$$G_a^{\alpha\beta} = U_a^{\alpha\beta} + PV^\Sigma - TS_a^{\alpha\beta} \tag{36.37}$$

Because the dividing surface is defined so that it has no volume ($V^\Sigma = 0$), these two versions of the interfacial free energy are equal to one another. We define them as the interfacial free energy, $\gamma_{\alpha\beta}$:

$$\gamma_{\alpha\beta} = F_a^{\alpha\beta} = G_a^{\alpha\beta} = U_a^{\alpha\beta} - TS_a^{\alpha\beta} \tag{36.38}$$

The interfacial contribution to the interfacial energy obeys the following version of Eq. 36.9:

$$\delta G = -S\delta T + \sum_i \mu_i \delta n_i \tag{36.39}$$

$$dG^{\Sigma} = -SdT^{\Sigma} + \sum \mu_i \delta n_i^{\Sigma} + \gamma_{\alpha\beta}\delta A^{\Sigma} \tag{36.40}$$

From the equilibrium condition, $\delta G = \delta G^{\alpha} + \delta G^{\beta} + \delta G^{\Sigma} = 0$, we get the same constant temperature and constant chemical potential requirement that we had in the previous section:

$$T^{\Sigma} = T^{\alpha} = T^{\beta} \tag{36.41}$$

$$\mu_i^{\alpha} = \mu_i^{\beta} = \mu_i^{\Sigma} \tag{36.42}$$

For a fixed temperature ($dT = 0$), and specified system volume, we use the Helmoltz free energy, since this is the thermodynamic potential to use when when T and V are the independent variables. The differential Helmholtz energy for the α phase, the β phase and the interfacial region reduces to the following:

$$\delta F = -P^{\alpha}\delta V^{\alpha} - P^{\beta}\delta V^{\beta} + \gamma_{\alpha\beta}\delta A^{\Sigma} = 0 \tag{36.43}$$

We can now look at the implications of this expression for a flat interface and for a curved interface.

36.3.1 Case 1: flat interface

Suppose the interface remains flat, and simply translates by a distance, δx, as illustrated in Figure 36.5. In this case the area of the interface remains fixed, so $\delta A^{\Sigma} = 0$. The volume of the α phase increase by $A^{\Sigma}\delta x$ and the volume of the β phase decreases by $A\delta x$. From Eq. 36.43 we have:

$$\delta F = -P^{\alpha}A^{\Sigma}dx + P^{\beta}A^{\Sigma}dx = 0 \tag{36.44}$$

This equilibrium condition for a flat interface can only be true when $P^{\alpha} = P^{\beta}$, *i.e.*, there is no pressure change across the interface.

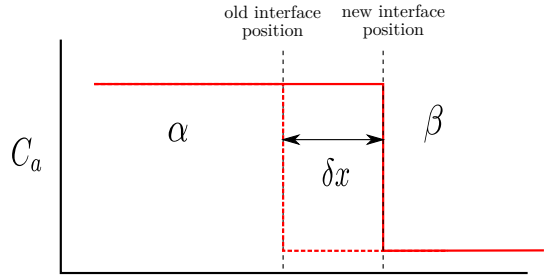


Figure 36.5: Displacement of a flat interface between α and β phases.

36.3.2 Case 2: curved interface

Assume a spherical particle of β in a matrix of α, as shown in Figure 36.6. We use the following relationships between the radius, interfacial area, and volume of the β precipitate:

$$V^{\beta} = \frac{4}{3}\pi R^3 \tag{36.45}$$

$$A^{\Sigma} = 4\pi R^2 \tag{36.46}$$

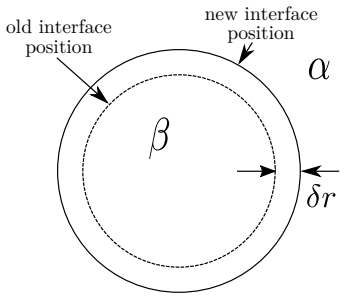


Figure 36.6: Displacement of a flat interface between α and β phases.

If R changes by a small amount δR, this leads to the following for δV^β and δA^Σ:

$$\delta V^\beta = 4\pi R^2 \delta R \tag{36.47}$$

$$\delta A^\Sigma = 8\pi R \delta R \tag{36.48}$$

We are working in a system with total fixed, *i.e.*, $\delta V = \delta V^\alpha + \delta V^\beta = 0$. From this we obtain:

$$\delta V^\alpha = -\delta V^\beta = -4\pi R^2 \delta R \tag{36.49}$$

Now we can now use these expressions for δV^β, δA^Σ and δV^α in Eq. 36.43 for δU in order to obtain the following:

$$P^\alpha 4\pi R^2 \delta R - P^\beta 4\pi R^2 \delta R + \gamma_{\alpha\beta} 8\pi R \delta R = 0 \tag{36.50}$$

After some rearrangement we obtain the Laplace pressure equation for a material with an isotropic surface energy:

$$P^\beta - P^\alpha = \frac{2\gamma_{\alpha\beta}}{R} \tag{36.51}$$

Note that this pressure equation is an additional equilibrium condition, in addition to those already obtained (constant temperature and constant chemical potential). Not that at equilibrium the chemical potentials are uniform everywhere, even in conditions where the pressure is non-uniform. For systems with curved interfaces we need to account of the effect of pressure (and hence, the interface curvature) on the chemical potential. These pressure-induced chemical potential differences drive a variety of important processes in microstructure development in materials, including coarsening and grain growth.

36.4 Chemical Potential Expressions

One useful thing that emerges from all of these expressions is that we can get some useful, equivalent expressions for the chemical potential. The chemical potential of component i is always given by the derivative of some thermodynamic potential with respect to n_i. The thermodynamic potential to use just depends on what we are holding constant during the differentiation: S, V if we use U; T, V if we use F; P, T if we use G. In mathematical terms we have:

$$\mu_i = \left.\frac{\partial U}{\partial n_i}\right|_{S,V} = \left.\frac{\partial F}{\partial n_i}\right|_{T,V} = \left.\frac{\partial G}{\partial n_i}\right|_{P,V} \tag{36.52}$$

The easiest way to see that this must be the case is to look at the corresponding expressions for δU (Eq. 36.3) , δF (Eq. 36.4) and δG (Eq. 36.9).
In this class we are going to be working primarily with the Gibbs free energy. A couple other statements about G and its relationship to the chemical potentials is useful here. The first is that the chemical potential is equivalent to the partial molar free energy. So writing the chemical potential as a derivative of the free energy, we can sum up the potentials to get the free energy:

$$G = \sum_i \mu_i n_i \tag{36.53}$$

36.5 A practical example

Semiconductor nanowires provide a useful illustration of the importance of thermodynamics in modern materials synthesis, and on effects that emerge when the relevant length scales become very small. A schematic representation of Si nanowires is shown in Figure 36.7, along with an illustration of the growth process. Growth occurs at the interface between a gold liquid phase (where the Si solubility is quite high) and a solid Si phase (which has a negligible solubility for the gold. The Si-Au phase diagram is obviously relevant to this problem, and is shown in Figure 36.8. The problem is that this phase diagram is for bulk materials, and will somehow be affected by the fact that the length scales are very small. Nanowire diameters are typically in the range of tens of nm, and we expect that things might behave differently at this length scale. This is is one of the issues addressed in this section.

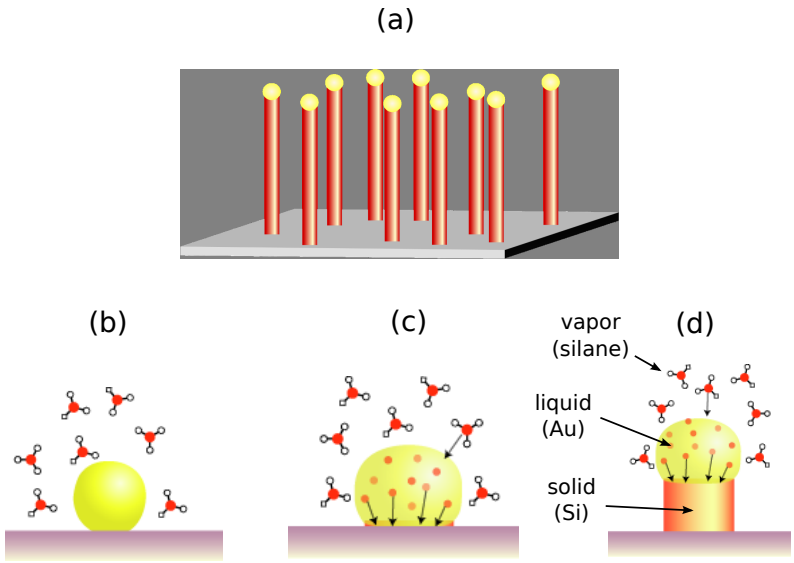


Figure 36.7: Schematic representation of Si nanowires (a) and their growth by the vapor-liquid-solid (VLS) mechanism. In this method growth occurs as Si-contaning precursor molecules react to form silicon that dissolves into a gold drop that is placed on a Si surface (b-d). The supersaturation of Si in the liquid phase causes Si to grow at the liquid/solid interface.

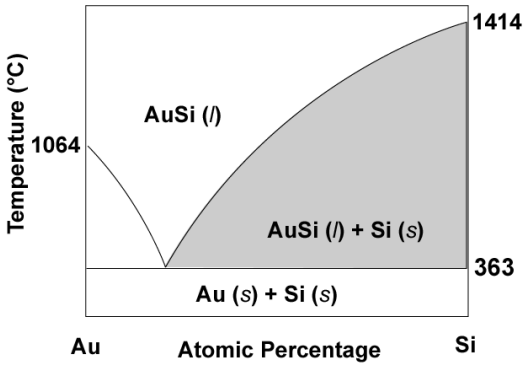


Figure 36.8: Si-Au phase diagram.

Example: Magnitude of the Laplace pressure How large is the Laplace pressure

36.6 Effects of Interfacial Curvature on the Melting Transition

How does the melting point of a pure material depend on its size? We'll assume that the solid material has a radius of R, and that the solid/liquid interfacial free energy is γ. As illustrated schematically in Figure 36.9

$P^s \neq P^\ell$ so $T \neq T_m$.

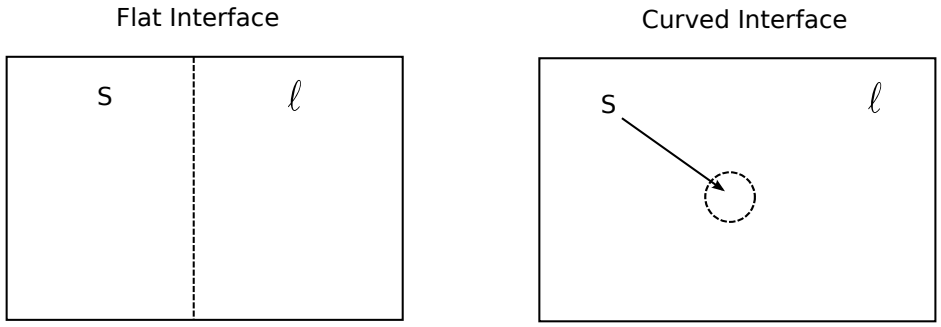


Figure 36.9: Solid/liquid equilibrium at flat and curved interfaces.

The equilibrium condition between the solid and liquid is obtained by equating the chemical potentials in the solid and liquid phase. For a single component system, the chemical potential on a molar basis (energy per mole of atoms as opposed to energy per atom) is equivalent to the molar free energy, G_m. What we need to do is find the temperature where the molar free energy of the material in the solid phase, G_m^s is equal to the molar free energy in the liquid phase (G_m^ℓ):

$$G_m^s(T, P^s) = G_m^\ell(T, P^\ell) \tag{36.54}$$

Note that we are interested in the case where the temperature is not necessarily equal to the equilibrium melting temperature ($T \neq T_m$). This temperature difference arises from the fact that the pressure in the solid and liquid phases differs by an amount given by the Laplace pressure difference, $P^s - P^\ell = 2\gamma_{s\ell}/R$. We assume that the differences between T and T_m and between P^s and P^ℓ are small, so we can get away with retaining only the first derivative terms in Taylor series expansions for G_m^s and G_m^ℓ:

$$\begin{cases} G_m^s(T,\, P^s) = G_m^s(T_m,\, P^\ell) + \left.\frac{\partial G_m^s}{\partial T}\right|_{P^\ell, T_m}(T - T_m) + \left.\frac{\partial G_m^s}{\partial P}\right|_{P^\ell, T_m}(P^S - P^\ell) \\ G_m^\ell(T, P^\ell) = G_m^\ell(T_m,\, P^\ell) + \left.\frac{\partial G_m^\ell}{\partial T}\right|_{P^L, T_m}(T - T_m) \end{cases} \tag{36.55}$$

Now we can use the following thermodynamic definitions:

$$\frac{\partial G_m^s}{\partial T} = -S_m^s,\ \frac{\partial G_m^\ell}{\partial T} = -S_m^\ell,\ \frac{\partial G_m^s}{\partial P} = V_m^s \tag{36.56}$$

Combination of Eqs. 45.20 and 45.21 gives the following:

$$\begin{cases} G_m^s(T, P^s) = G_m^s + (T_m, P^\ell) - S_m^s(T - T_m) + V_m^s(P^s - P^\ell) \\ G_m^\ell(T, P^\ell) = G_m^\ell(T_m, P^\ell) - S_m^\ell(T - T_m) \end{cases} \tag{36.57}$$

For a planar interface ($R = \infty$): $P^\ell = P^s$, and the liquid and solids are at equilibrium at $T = T_m$:

$$G_m^\ell(T_m, P^\ell) = G_m^s(T_m, P^\ell) \tag{36.58}$$

We define the differences, ΔH_m, ΔS_m, ΔT in the following way:

$$\Delta H_m \equiv H_m^\ell - H_m^S, \Delta S_m \equiv S_m^\ell - S_m^s,\ \Delta T \equiv T_m - T \tag{36.59}$$

Now we combine Eqs. 36.54, 45.22, 45.23, 45.24 to obtain the following for the temperature difference:

$$T_m - T = \frac{\left(P^s - P^\ell\right) V_m^s}{\Delta S_m} \tag{36.60}$$

The enthalpy of melting, ΔH_m, is a more intuitive quantity than the entropy of melting, and is more directly measured experimentally. At equilibrium for an interface between solid and liquid with a planar interface (so the pressure is the same in both phases), the free energies of the solid and liquid are equal to one another. This fact is generally used to write thermodynamic quantities in terms of ΔH_m instead of ΔS_m. We begin by recognizing that the free energy change between solid and liquid phases is zero at $T = T_m$:

$$\Delta G_m(T_m) = \Delta H_m - T_m \Delta S_m = 0 \tag{36.61}$$

We can use this equation to write the entropy of mixing in terms of the enthalpy of mixing and the equilibrium melting temperature:

$$\Delta S_m = \frac{\Delta H_m}{T_m} \tag{36.62}$$

Equation 36.60 for the melting point depression can therefore be rewritten in the following way:

$$T_m - T = \frac{\left(P^s - P^l\right) V_m^s T_m}{\Delta H_m} \tag{36.63}$$

Finally, with $P^s - P^\ell = 2\gamma_{s\ell}/r$ we have:

$$T_m - T = \frac{2\gamma_{s\ell} V_m^s T_m}{r \Delta H_m} \tag{36.64}$$

So how large is this effect? To understand this, we need to put in some real numbers. Let's consider the case for gold with a particle radius, R of 5 nm:

- γ (solid/liquid interfacial free energy for gold): 0.177 J/m^2.
- V_m^s (molar volume of solid gold): $\frac{197\,\text{g}}{\text{mole}} \cdot \frac{\text{cm}^3}{17.3\,\text{g}} \cdot \frac{10^{-6}\,\text{m}^3}{\text{cm}^3} = 1.14\text{x}10^{-5}\,\text{m}^3$
- ΔH_m (molar heat of fusion): 1.25x10^4 J/mole
- T_m (equilibrium melting temperature): 1064 °C (1337 K)
- R (droplet radius): 5x10^{-9} m

Putting all these numbers into Eq. 36.64 gives $T_m - T = 86$ K, which is certainly a significant effect.
We conclude this section with a useful graphical interpretation of the effect of pressure.

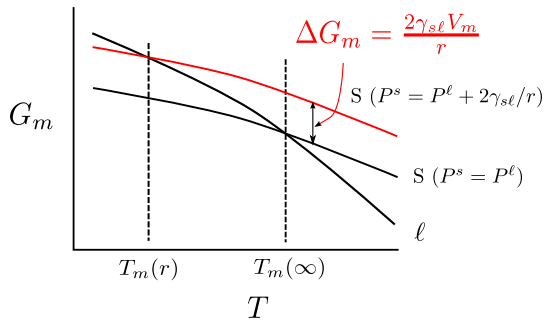


Figure 36.10: Graphical representation of the effect of the particle size on the melting temperature.

By taking only the first term in the Taylor expansion, we are assuming that the plots of G_m vs T are straight, *i.e.* we are neglecting any temperature dependence of the entropy. We are also assuming that $\partial G_m/\partial P$ is constant, which means we are saying that the molar volume is independent of the pressure (the system is assume to be incompressible). This is generally a reasonable approximation for most solid and liquid materials, but will fail miserably if one of the phases is a gas.

36.7 Size-dependent solubility

Another consequence of the increased pressure within a small precipitate is that small precipitates are more soluble in their surroundings than large precipitates.

36.7.1 General Concepts

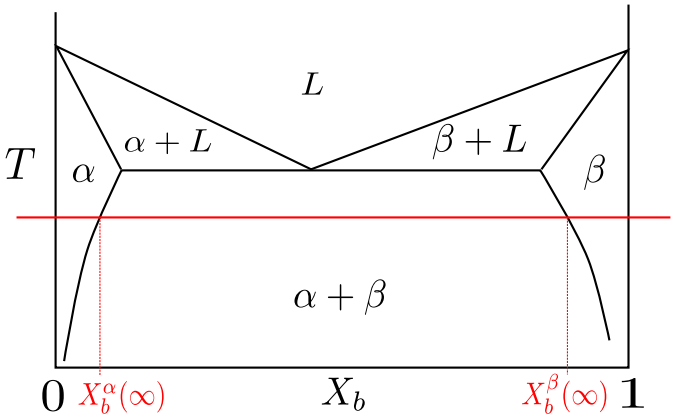


Figure 36.11: Eutectic phase diagram.

At temperatures below the eutectic temperature, solid α and solid β are in equilibrium with one another. For flat interfaces, ($R = \infty$) the phase compositions are given by the solvus lines, and are equal to $X_B^\alpha(\infty)$ and $X_B^\beta(\infty)$. How does this change if the interface is curved? Suppose we are in the A-rich portion of the phase diagram, where small β precipitates of radius r exist in a matrix of α.

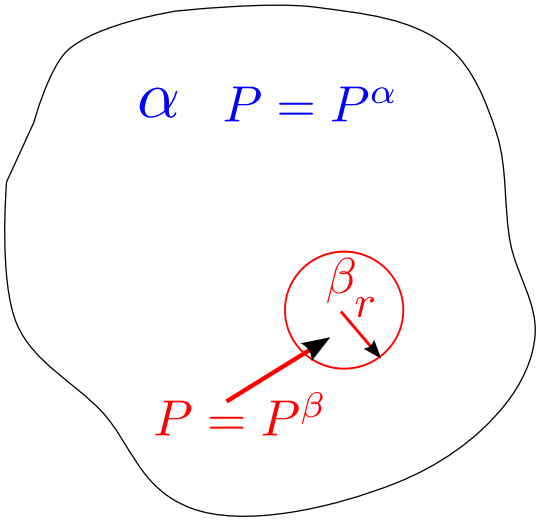


Figure 36.12: A β precipitate of radius R in a matrix of α.

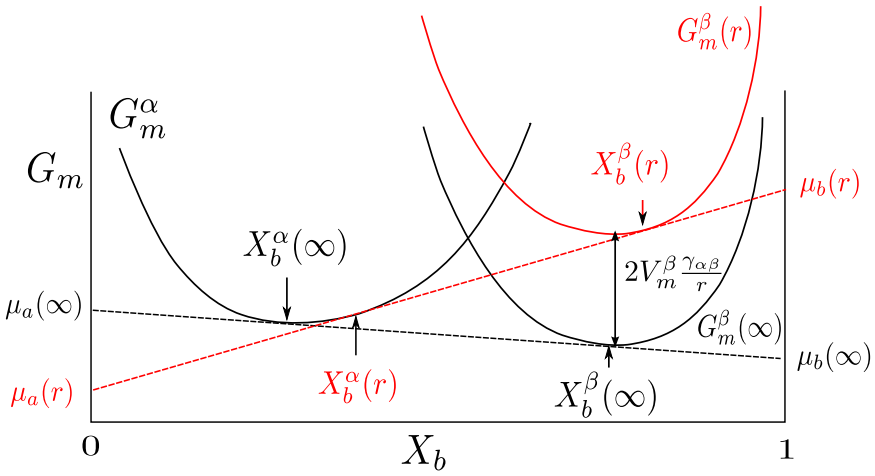


Figure 36.13: Shift of the free energy curves and equilibrium phase compositions due to curvature effects.

$$\frac{\partial G_m^\beta}{\partial P} = V_m^\beta \tag{36.65}$$

If we assume that β is incompressible, then V_m^β does not change with pressure and we have:

$$G_m^\beta\left(P^\beta\right) = G_m^\beta\left(P^\alpha\right) + V_m^\beta\left(P^\beta - P^\alpha\right) \tag{36.66}$$

With $P^\beta - P^\alpha = 2\gamma_{\alpha\beta}/R$, where $\gamma_{\alpha\beta}$ is the interfacial free energy for the interface between α and β phases, we have:

$$G_m^\beta\left(P^\beta\right) = G_m^\beta\left(P^\alpha\right) + \frac{2V_m^\beta\gamma_{\alpha\beta}}{R} \tag{36.67}$$

From the construction in Figure 36.13 we see that X_B^α and X_B^β are both functions of r. More specifically, we have the following inequalities:

$$\begin{aligned} X_B^\alpha(r) &> X_B^\alpha(\infty) \\ X_B^\beta(r) &> X_B^\beta(\infty) \end{aligned} \tag{36.68}$$

To develop expressions for $X_B^\alpha(r)$ and $X_B^\beta(r)$, we just need to equate the chemical potentials for A and B atoms in each phase:

$$\mu_i^\alpha\left(X_B^\alpha, P^\alpha\right) = \mu_i^\beta\left(X_B^\beta, P^\beta\right): \quad i = A, B \tag{36.69}$$

where $P^\beta = P^\alpha + 2\gamma_{\alpha\beta}/r$. In general Eq. 36.69 is a set of two, nonlinear equations (obtained by setting i to A or B) that must be solved numerically in order to obtain X_B^α and X_B^β, the compositions of the α and β phases that re in equilibrium with one another. Note that because $X_A = 1 - X_B$ we can use the single composition variable, X_B to describe the compositions.

36.7.2 Activity Coefficients

In general the chemical potential of species i is related to its activity coefficient, a_i:

$$\mu_i = \mu_i^0 + RT\ln(a_i) \tag{36.70}$$

Here R is the gas constant (8.314 J/mole·K) T is the absolute temperature and μ_i^0 is the chemical potential in it's standard state (which we'll take at a pressure of P^α). We'll define the standard state chemical potentials as zero, so the chemical potentials for $P = P^\alpha$ are simply:

$$\mu_i\left(P^\alpha\right) = RT\ln(a_i) \tag{36.71}$$

36.7.3 Effect of Pressure

In the beta phase, we need to account for the fact that pressure in the β phase, P^{β}, is no longer equal to the reference pressure, P^{α}:

$$\mu_i\left(P^{\beta}\right) = RT\ln a_i + \frac{\partial \mu_i^{\beta}}{\partial P}\left(P^{\beta} - P^{\alpha}\right) \tag{36.72}$$

The pressure derivatives appearing in these equations can be replaced by the partial molar volumes of the A and B components in the β phase, defined as follows:

$$\frac{\partial \mu_i}{\partial P} = \bar{V}_i \tag{36.73}$$

where $\bar{V}_i$ is the partial molar volume of component i. We can therefore right the chemical potential in the following generalized form, which accounts for its dependence on both composition (X_b) and pressure (P^{β}). and pressure concentration and pressure dependence:

$$\mu_i\left(X_B,\, P^{\beta}\right) = RT\ln a_i\left(X_B\right) + \bar{V}_i\left(P^{\beta} - P^{\alpha}\right) \tag{36.74}$$

36.7.4 Expression for $X_B^{\alpha}(r)$ in the dilute regime.

In order to illustrate how this is done, we'll consider the simplest possible case, where the α phase is nearly pure A and the β phase is nearly pure B. If component i is very dilute ($X_i \ll 1$), we are in the Henry's law regime where the activity coefficient increases linearly with the concentration:

$$a_i = H_i X_i \quad (\text{Raoult's Law}) \tag{36.75}$$

where H_i is the Henry's law coefficient. Similarly, if $X_i \approx 1$ (so that the phase of interest is nearly pure i), the activity coefficient is also proportional to the concentration, but with a slope of 1 (Raoult's law):

$$a_i = X_i \tag{36.76}$$

Flat Interface $R \to \infty$: For a flat interface the pressures in both phases are equal to the standard pressure of P^{α}. The chemical potential of B in the beta phase is given by combining Eq. 36.74 (with $P^{\alpha} = P^{\beta}$ and $X_b = X_b^{\beta}(\infty)$) and Eq. 36.76 (Raoult's law) to obtain:

$$\mu_B^{\beta} = RT\ln X_B^{\beta}(\infty) \approx 0 \tag{36.77}$$

We know that μ_B^{β} is close to zero because $X_B^{\beta}(\infty)$ is close to one. Component B is dilute in the the alpha phase, so we are in the Henry's law regime. The chemical potential is given as follows:

$$\mu_B^{\alpha} = RT\ln a_B^{\alpha} = RT\left[\ln H_B + \ln X_B^{\alpha}(\infty)\right] \approx 0 \tag{36.78}$$

We know that $\mu_B^{\alpha} = \mu_B^{\beta} \approx 0$, which is why we can set μ_B^{α} to zero in Eq. 36.78. With $\mu_B^{\alpha} = 0$ we have:

$$\ln H_B = -\ln X_B^{\alpha}(\infty) \tag{36.79}$$

Curved interface - finite R: The chemical potentials are now modified by the pressure contribution to the molar free energy of the beta phase, and are no longer zero. One consequence of this is that the equilibrium compositions are changed. Because the α phase pressure is taken as our reference pressure, the equations for the chemical potentials in this phase are unchanged. We just need to specify that r is no longer infinite:

$$\mu_B^{\alpha}(r) = RT\left[\ln H_B + \ln X_B^{\alpha}(r)\right]$$

Now we can develop some analytic expressions that are useful in the dilute limit. We'll Eq. 36.73, along with the fact that $P^{\beta} - P^{\alpha} = 2\gamma_{\alpha\beta}/r$ to simplify some of the expressions. We're specifically interested in the increase in the minority phase fraction when r becomes very small. In other words, how large is $X_B^{\alpha}(r)$ compared to $X_B^{\alpha}(\infty)$? Requiring that $\mu_B^{\beta}(r) = \mu_B^{\alpha}(r)$ gives:

$$RT\left[\ln H_B + \ln X_B^\alpha(r)\right] = RT\ln X_B^\beta(r) + \frac{2\gamma_{\alpha\beta}\bar{V}_B}{r} \tag{36.80}$$

Eq. 36.79 holds in the alpha phase, which is assumed to be nearly pure A, so we can replace $\ln H_B$ with $-\ln X_B^\alpha(\infty)$. Similarly, the β phase is assumed to be nearly pure B, soln $\left(X_B^\beta\right) \approx 0$ and we have:

$$RT\ln\left(\frac{X_B^\alpha(r)}{X_B^\alpha(\infty)}\right) = \frac{2\gamma_{\alpha\beta}\bar{V}_B^\beta}{r} \tag{36.81}$$

The molar volume of the β precipitate is related to the partial molar volumes in the following way:

$$V_m^\beta = X_A^\beta \bar{V}_A^\beta + X_B^\beta \bar{V}_B^\beta \tag{36.82}$$

we have $X_B^\beta \approx 1$ so we can replace $\bar{V}_B^\beta$ with V_m^β. Also, since the β is nearly pure B, V_m^β is really just the molar volume of B. Now we can rearrange Eq. 36.81 to obtain the following expression for $X_B^\alpha(r)$:

$$X_B^\alpha(r) = X_B^\alpha(\infty)\exp\left[\frac{2\gamma_{\alpha\beta}V_m^\beta}{RTr}\right] \tag{36.83}$$

For small x, $\exp(x) \approx 1 + x$. If r is not too small (generally the case) we can use this expansion for the exponential function to write $X_B^\alpha(r)$ as follows:

$$X_B^\alpha(r) = X_B^\alpha(\infty)\left[1 + \frac{2\gamma_{\alpha\beta}V_m^\beta}{RTr}\right] \tag{36.84}$$

We see that the surface free energy term makes small precipitates more soluble in the matrix than larger precipitates. This increased solubility drives the coarsening of the microstructure over time, giving larger precipitates over time. We're not going to do much more with this specific equation in 316-1, but it is very important when we start talking more about the evolution of microstructure in 316-2. It is given here largely as an illustration of the importance of the interfacial free energy.

37 Two Dimensional Defects in Crystals: Surfaces and grain boundaries.

Dislocations are one dimensional defects in a crystalline structure. We now consider crystal interfaces, which can be viewed as two-dimensional crystal defects. We'll consider three kinds of interfaces:

1. Free surfaces
2. Grain boundaries
3. Interphase interfaces

All of these interfaces have an interfacial energy, given by the energy required to create extra interfacial area:

$$\gamma = \frac{\partial U}{\partial A} \tag{37.1}$$

One of the unique features of crystalline materials is that γ is no longer isotropic. Certain crystal facets have a lower interfacial energy than other facets. This is why natural crystals like quartz (see Figure 37.1) have beautiful shapes and are not boring solid blobs. In the following sections we'll investigate some of the features that give rise to the anisotropy in γ, and will see how this anisotropy determines the equilibrium crystal shape.

Figure 37.1: Naturally-formed quartz crystals.

Crystal surfaces with the lowest surface energies tend to be ones with relatively high densities of atoms within the plane. For FCC crystal structures, these include the {111}, {200} and{220}, with hard sphere representations of these surfaces shown in Figure 37.2. Note that the {100} and {200} surfaces are identical, as are the {110} and {220} surfaces. We use the non-reduced notation so that we obtain the correct interplanar spacing for identical planes. For a cubic crystal structure d is given by the following expression:

$$d_{hkl} = \frac{a}{\left(h^2 + k^2 + \ell^2\right)^{1/2}} \tag{37.2}$$

here h, k and ℓ are the Miller indices [9] and a is the lattice parameter.

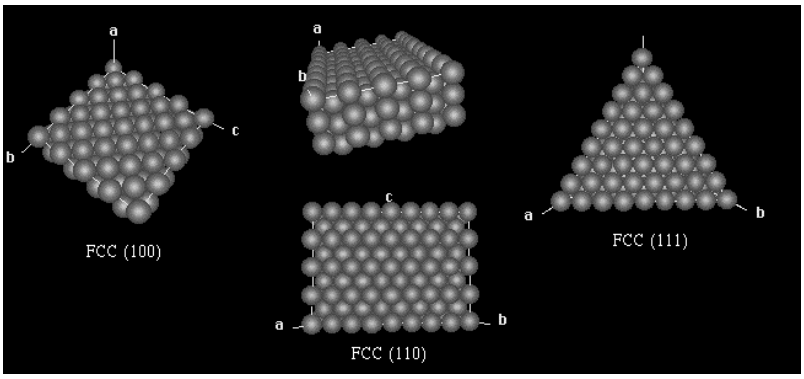


Figure 37.2: Representations of the the {100}, {110} and{111} crystal surfaces for an FCC material.

Miller indices: If you need a refresher on Miller indices, the Wikipedia page is very helpful. Here we include a reminder of the notation used to indicate planes and directions.

- **Specific planes:** We use round brackets - $(hk\ell)$
- **Class of planes:** To indicate all planes that are crystallographically identical, we use curly brackets - $\{hk\ell\}$
- **Specific direction:** We use square brackets - $[hk\ell]$
- **Class of directions:** To indicate all directions that are crystallographically identical, we use angle brackets - $\langle hk\ell \rangle$

37.1 Surface energy of a close packed plane

We can use the 'missing bond' picture of crystalline surfaces to estimate the surface energy for a close packed plane of atoms. Consider, for example a {111} surface in the FCC crystal structure. The FCC crystal structure consists of ABC stacking of these close-packed planes, as shown in Figure 37.3.

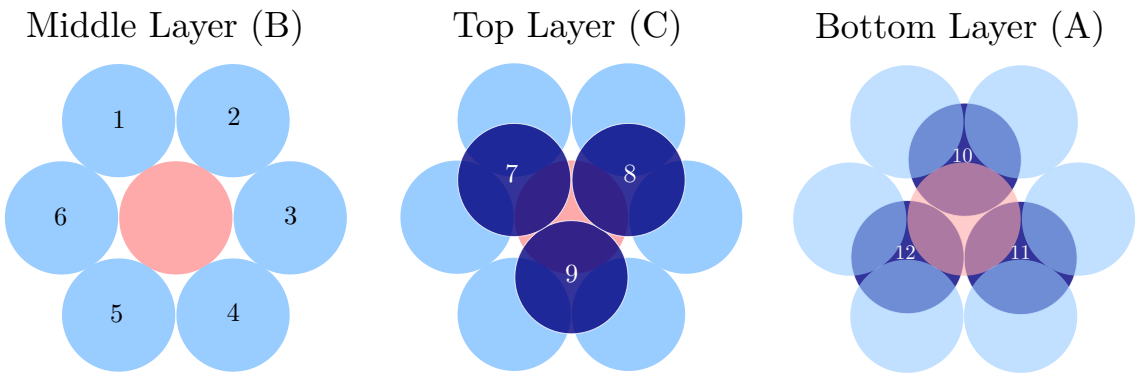


Figure 37.3: Packing of the close packed planes in an FCC crystal structure.

Consider an atom within one of the 'B' planes within the bulk crystal structure. Each atom within this close-packed plane has 12 nearest neighbors: 6 in same 'B' plane, 3 in the 'A' plane below, 3 in the 'C' plane above. Now suppose that this 'B' plane represents the crystal surface. We have removed the three bonds to the atoms in the 'C' layer, so that we have lost the energy associated with 3 of the 12 nearest neighbor bonds. Now suppose the energy per bond is ϵ. The bond energy/atom is $\epsilon/2$ (since the bond energy is shared between the two atoms). This means that every surface atom has an excess energy of $3\epsilon/2$ compared to the energy of an atom in the bulk of a material. All we need now is an estimate of ϵ. The easiest way to get this is to look at the molar heat of sublimation, L_m, which is the energy required to convert one mole of atoms from the solid to the vapor. (This is also referred to as the heat of vaporization, and is included in the Wikipedia entries for the different elements.) If one mole of atoms is vaporized, then $zN_{av}/2$ bonds are broken, where z is the number of nearest neighbors for a given atom (12 in our case) and N_{av} is Avogadro's number. For $z = 12$ we have:

$$L_m = 12N_{av}\frac{\epsilon}{2} = 6N_{av}\epsilon \tag{37.3}$$

Rearrangement of this equation gives:

$$\epsilon = L_m/6N_{av} \tag{37.4}$$

so the surface energy per atom is $L_m/4N_{av}$. There is also an excess entropy associated with the surface due to changes in the vibrations of the surface atoms, configuration entropy due to surface vacancies, but this is typically a small contribution to the overall surface free energy and is ignored in our treatment. The surface energy is obtained by multiplying the excess surface energy per atom by the surface density of atoms on the plane of interest, Σ_s:

$$\gamma_{sv} = \frac{3\epsilon\Sigma_s}{2} \tag{37.5}$$

When calculating Σ_s, it is easier to think in terms of its inverse, A_s, the area per surface atom in the plane of interest. The situation for a close-packed plane is shown in Figure 37.4, where we show the two dimensional unit cell for a hexagonal lattice. We obtain the following for Σ_s for a close-packed plane of atoms:

$$\Sigma_s = \frac{1}{A_s} = \frac{1}{2r^2\sqrt{3}} \tag{37.6}$$

where r is the atomic radius.

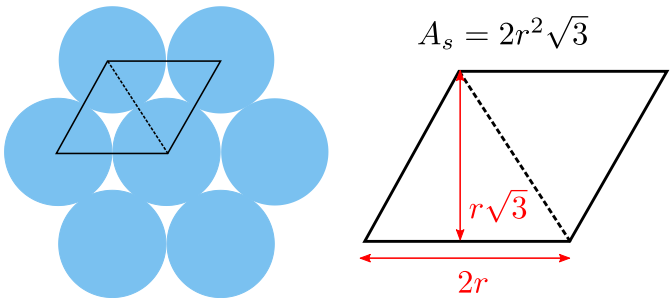


Figure 37.4: The two dimensional unit cell for a hexagonal lattice. The area per unit cell is the area per atom, A_s.

Combination of Eqs. 37.4, 37.5 and 37.6 gives:

$$\gamma_{sv} = \frac{L_m}{8r^2\sqrt{3}N_{av}} \tag{37.7}$$

In addition to having a higher value of γ_{sv}, we also expect that materials with a higher value of L_m will have higher melting points (T_m). Values for T_m and γ_{sv} are listed in Table 37.1.

Table 37.1: Melting temperatures and solid/vapor interfacial free energies.

Crystal	T_m (°C)	γ_{sv} (J/m^2)
Sn	232	0.68
Au	1063	1.39
W	3407	2.65

37.2 Orientation Dependence of the Surface Energy

In order to understand the faceting of single crystals we need to understand the anisotropy of the surface energy. The existence of this anisotropy is one of the key differences between a liquid and a solid. We can use simple bond counting arguments to understand where this anisotropy comes from, using drawing shown in Figure 37.5. This figure shows a square lattice of atoms with an exposed surface tilted by an angle θ with respect to one of the crystal axes. The only way to get this tilted surface is to add a series of atomic steps, each of which leaves a broken bond at the top and left surfaces. Consider surface of length ℓ along the surface of the material. The projection of this length onto the horizontal axis is $\ell\cos\theta$ and the projection onto the vertical axis is $\ell\sin\theta$. Along each of these directions the distance between broken bonds is a, so the total number of bonds is along the length ℓ is $(\ell/a)(\cos\theta + \sin\theta)$. The number of bonds along the width of the sample (the direction perpendicular to the plane of Figure 37.5) is simply w/a, where w is the sample width. The number of bonds per unit area, Σ_s is:

$$\Sigma_s = \frac{1}{a^2}(\cos\theta + \sin\theta) \tag{37.8}$$

The surface energy is obtained multiplying by the energy per bond, $\epsilon/2$:

$$\gamma_{sv} \approx \frac{\epsilon}{2a^2}(\cos\theta + \sin\theta) \tag{37.9}$$

The equation is approximate because we have not accounted for any entropic contributions to the surface free energy. Also, the model of just adding up the contributions from 'missing' bonds neglects the tendency for the atoms at the surface to reorganize into structures that lower the overall free energy. These surface reconstructions play an important role in the surface science of materials

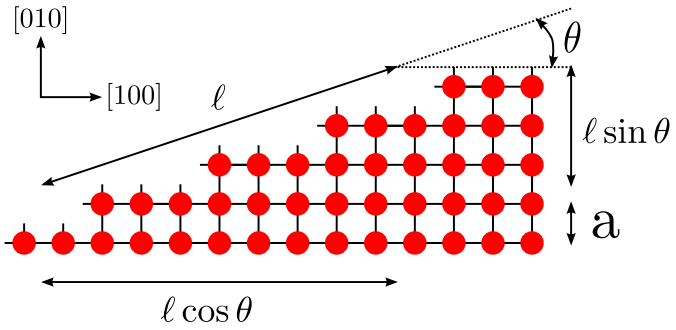


Figure 37.5: Schematic representation of broken bonds at a stepped surface.

Equation 37.9 is valid for values of θ between 0 and 90°, where $\sin\theta$ and $\cos\theta$ are both positive. These sin and cos terms came from the projected length of the tilted surface along the [100] and [010] directions. A

$$\gamma_{sv} \approx \frac{\epsilon}{2a^2}\left(|\cos\theta| + |\sin\theta|\right) \tag{37.10}$$

(Note that the equation here is approximate because we have neglected entropic contributions to surface free energy). There's no derivative at $\theta = 0$, so the free energy function must have a cusp. (show plot). How can we represent γ_{sv} as a function of θ? Describe in terms of θ_n (angle of normal of a plane with respect to the x axis). So we can plot γ_{sv}on a polar plot. In MATLAB we use the 'polar' command to do this. We'll give an example when we illustrate the Wulff construction in the following section.

37.3 Equilibrium Shape of Crystals

We know that γ is a function of the angle, but what are the implications on the equilibrium shape of the crystal? We need to minimize the total surface energy subject to volume conservation.

$$\int_A \gamma ds$$

If γ is a constant (independent of the angle), then we just need to minimize the overall surface area for a fixed volume. We get a sphere in this case. Now we have $\gamma = f(\theta)$. Suppose we have two facets with surface free energies of γ_1 and γ_2.

$$G_t = \gamma_1 A_1 + \gamma_2 A_2$$

So how do we minimize this? We use the Wulff construction to provide the shape with the lowest energy. The construction was proposed in 1901, but it was not proved mathematically until 1953. We won't attempt to show the proof here, but will instead focus on the use of the construction itself. The procedure is as follows:

1. Draw $\gamma = \gamma(\theta)$
2. Draw line from origin to γ curve for a given value of θ
3. Draw perpendicular to this line.
4. Repeat for all values of θ.
5. Inner envelope is equilibrium shape.

Here's a MATLAB script that does this for the surface energy expression given in Eq. 37.10:

```
close all
gamma=@(alpha) abs(cos(alpha))+abs(sin(alpha));
r=@(theta,alpha) gamma(alpha)/cos(theta-alpha);
alpha=linspace(0,2*pi,200);
polar(alpha,gamma(alpha),'r-');
title('\gamma=|cos\theta|+|sin\theta|', 'fontsize', 16)
hold on % plot all subsequent curves on existing axes
for alpha=linspace(0,2*pi,17)  % this  is the loop that draws all the lines
    theta(1)=alpha+2*pi/5;  % specify two angles on either side of alpha
```

```
    theta(2)=alpha-2*pi/5;
    rvals(1)=r(theta(1),alpha);  % use the equation provided to get r for each
    % of the specified angles
    rvals(2)=r(theta(2),alpha);
    polar(theta,rvals) % plot lines connecting the two points we just defined
end
set(gcf,'paperposition',[0 0 5 5],'papersize',[5 5])
print(gcf,'../figures/matlabwulffenergyexample.eps', '-depsc2') % save the eps file
```

The resulting construction is shown in Figure 37.6.

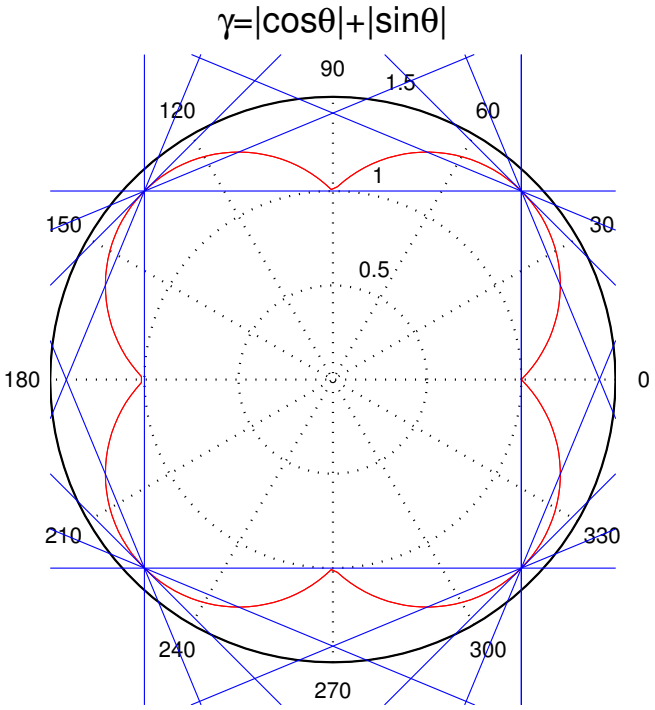


Figure 37.6: Wulff construction for the the surface free energy function given by Eq. 37.10.

38 Grain Boundaries

In the two dimensional Wulff construction the interface surface of interest is specified by a single variable, θ. For a true, three-dimensional crystal the Wulff circle becomes a Wulff sphere, and we need two different angles to specify the the orientation on this sphere. In other words, we have two degrees of freedom in specifying a specific surface of a three-dimensional crystal. We commonly specify a surface by using the normal vector, $\hat{n}$, that is perpendicular to that surface. This surface has three components, n_1, n_2 and n_3, in the x, y and z directions, respectively:

$$\hat{n} = (n_1\hat{x} + n_2\hat{y} + n_3\hat{z}) \tag{38.1}$$

Because $\hat{n}$ is a unit vector with $n_1^2 + n_2^2 + n_3^2 = 1$, only two of the three components of $\hat{n}$ are independent, so we again come to the conclusion that there are two degrees of freedom associated with the specification of a crystal surface.
In order to fully describe a grain boundary between two crystals we need to specify three additional degrees of freedom, so there are five degrees of freedom altogether. In order to illustrate these degrees of freedom, we can consider the following conceptual procedure for producing a grain boundary.

1. Cut the crystal along a plane specified by the unit normal to the plane, $\hat{n}$. Two degrees of freedom are associated with the specification of $\hat{n}$.

2. Rotate one of the two halves of the crystal by θ about an axis directed along the unit normal, $\hat{c}$.

Two additional degrees of freedom are used in the specification of $\hat{c}$, just as we use two degrees of freedom to specify $\hat{n}$. The fifth and final degree of freedom is the the rotation angle, θ.
The fact that 5 different parameters are needed to specify a grain boundary within a given crystal means that it is impossible for us to be exhaustive in our treatment of the different possibilities. Instead, we'll consider the following three cases:

1. Twist boundary: Twist boundaries correspond to rotation about an axis that is perpendicular to the plane. In terms of $\hat{c}$ and $\hat{n}$, they correspond to the case where these unit vectors are parallel to one another:

$$\hat{c} \parallel \hat{n}$$

2. Tilt boundary: Tilt boundaries correspond to the opposite limiting case, where $\hat{c}$ and $\hat{n}$ are perpendicular to one another:

$$\hat{c} \perp \hat{n}$$

3. Twin boundary. This is a special type of low energy tilt boundary, where lattice planes on either side of the boundary are in registry with one another.

Examples of pure twist and pure tilt boundaries are shown in Figure 38.1. In the following subsections we describe each of these boundaries in more detail.

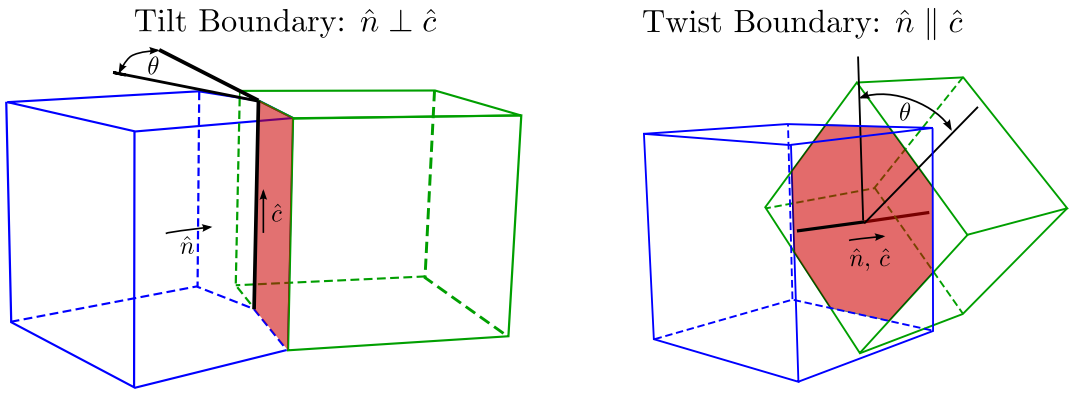


Figure 38.1: Schematic representation of tilt and twist boundaries.

38.1 Tilt Boundaries

A low-angle tilt boundary can be produced by introducing a series of edge dislocations along the grain boundary, as shown in Figure 38.2. The average distance, d, between dislocation in a low-energy tilt boundary is given by the following expression, which can be seen from Figure 38.2:

$$\tan(\theta/2) = \frac{b/2}{2\sin(\theta/2)} \tag{38.2}$$

For very small values of θ, we can assume $\tan(\theta/2) \approx \theta/2$, so we have:

$$d \approx b/\theta \tag{38.3}$$

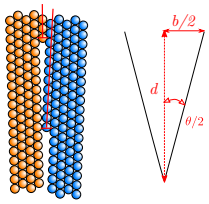


Figure 38.2: A low angle tilt boundary.

The energy per unit length of a dislocation is T_s. For a low angle grain boundary consisting of dislocations separated by a distance b, we expect the following for the grain boundary energy, γ_{gb}:

$$\gamma_{gb} = \frac{T_s}{d} \approx \frac{T_s \theta}{b} \tag{38.4}$$

where we have used Eq. 38.3 to approximate d. If we use Eq. 35.18 for the dislocation energy (with $T_s = E_s/h$), we have:

$$\gamma_{gb} = \frac{Gb\theta}{4\pi(1-\nu)} \ln(r_{max}/r_0) \tag{38.5}$$

The simplest thing to do here is to let the upper cutoff, r_{max}, correspond to the dislocation spacing d, which in our case is equal to b/θ. From this we get:

$$\gamma_{gb} = \gamma_0 \theta (\ln(b/r_0) - \ln\theta) \tag{38.6}$$

where we have defined γ_0 in the following way:

$$\gamma_0 = \frac{Gb}{4\pi(1-\nu)} \tag{38.7}$$

A more detailed treatment (the Read-Schockley model of low angle tilt boundaries) gives the following very similar form:

$$\gamma_{gb} = \gamma_0 \theta (B - \ln\theta) \tag{38.8}$$

where instead of $B = \ln(b/r_0)$ as in Eq. 38.6 we have:

$$B = 1 + \ln(b/2\pi r_0) \tag{38.9}$$

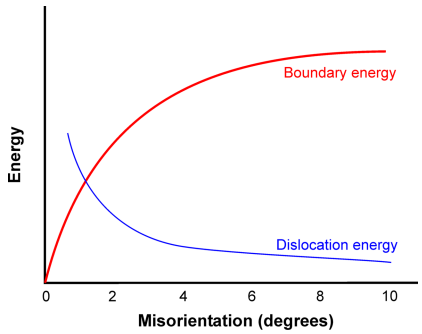


Figure 38.3: Grain boundary energy and energy per dislocation as a function of tilt angle according to the Read-Schockley model.

38.2 Twin Boundaries

Twin boundaries are special class of tilt boundaries with an exceptionally low energy. They are basically just a disruption in the stacking of the stacking of the layers in an FCC crystal structure, which is shown schematically in Figure 38.4. The point here is that if we look at one of the close packed {111} planes, and designate the location of the atom centers as 'A', we have two choices for the location of the centers of the next layer, labeled as 'B' and 'C' in Figure 38.4. Suppose we place the centers of the second layer of atoms at 'B'. We don't know if the structure is HCP or FCC until we put the third layer down. We have two choices:

1. The third layer goes above position 'A', so that the repeating structure of the stacking is ABABAB... This results in the FCC structure.

2. The third layer goes above position 'C', so that the repeating structure of the stacking is ABCABCABC... This stacking produces the FCC structure.

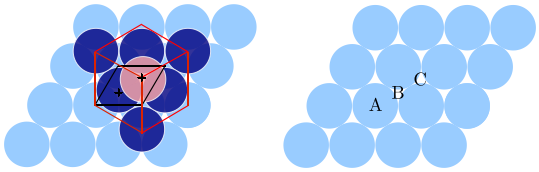


Figure 38.4: FCC crystal structure, illustrating the stacking of close-packed planes.

It is evident that there is actually a very small difference between the FCC and HCP crystal structures, with the difference depending on the way atoms interact with other atoms two layers away. In other words, there is not a big difference between energies of HCP and FCC crystals, since the first nearest neighbors are the same. We need to go to the second nearest neighbors to find a difference. A consequence of this small energy difference is that there can often be small regions of HCP-like structure in an FCC crystal. A twin boundary can be viewed as a case where three layers HCP stacking exist within an FCC structure. An example from a recent application involving a solar cell is shown in Figure 38.5.

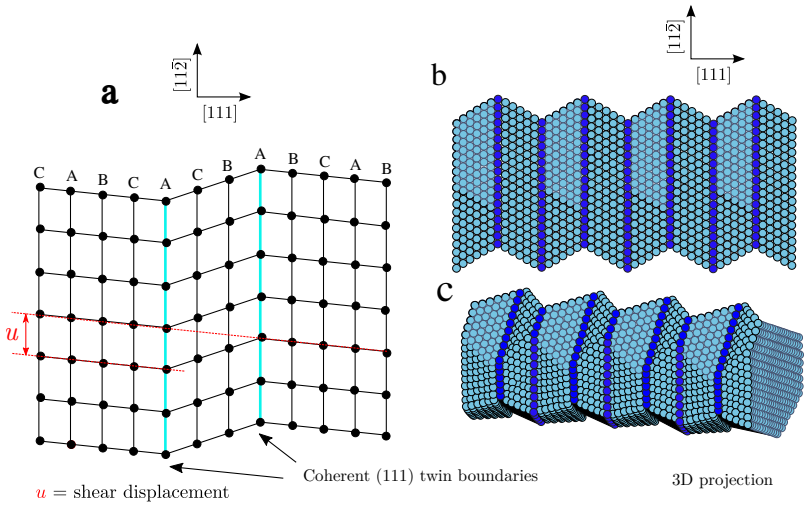


Figure 38.5: Example of a series of twin boundaries in an FCC crystal (from ref. [?], Reprinted by permission from Macmillan Publishers Ltd.).

Now we can talk about twin boundaries in an FCC structure. In an untwinned structure this is a regular, uninterupted of the stacking of the 'A', 'B' and 'C' close-packed planes:

- Untwinned: ABCABCABC

At a twin boundary this stacking gets interrupted, as in the following example.

- Twinned: ABCAB | C | BACBA (twin indicated)

The twin plane is the 'C' plane in the middle of this sequence. Note that the sequences on either side of the twin boundary are mirror images of one another. The sequence of planes working out from the twin plane is 'BACBACBAC...' in both cases. Twin boundaries have very low energies because there are no broken bonds, dislocations, step edges, etc. The energy only comes from the small unfavorable energy associted from second nearest neighbor interactions, as described above.

38.3 Twist Boundaries

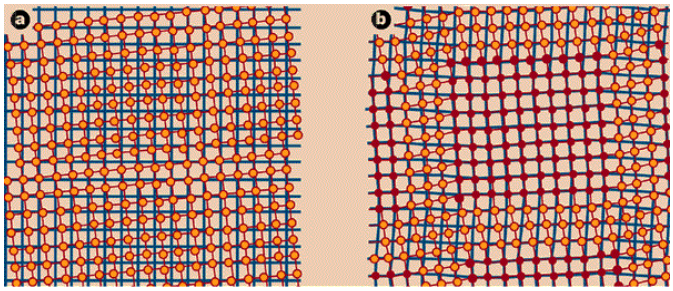


Figure 38.6: Twist boundary.

38.4 Interface Junctions

At this point it is useful to make some general statements about the junctions between different interfaces. We'll start by simplifying things a lot by ignoring the orientation dependence of the grain boundary energy and treating $\gamma_{\alpha\beta}$ as a constant.

38.5 Thermally activated migration of grain boundaries

(General treatment is much more general than just grain boundaries)
Assume isotropic grain boundary properties.
In general, grain boundaries are not flat. As a result the boundaries are subject to a force. Very similar to force exerted on a line by the line tension.
Recall the pressure dependence of the chemical potential:

$$\mu\left(P^2\right) = \mu\left(P^1\right) + \frac{2V_m\gamma}{r} \tag{38.10}$$

$$\frac{\Delta\mu}{V_m} = \Delta P = \frac{\text{force}}{\text{area}} \tag{38.11}$$

Boundaries move toward their center of curvature.
Now we know the driving force - this comes from thermodynamics
Need to study the kinetics to understand if the grain will actually move in response to this driving force.

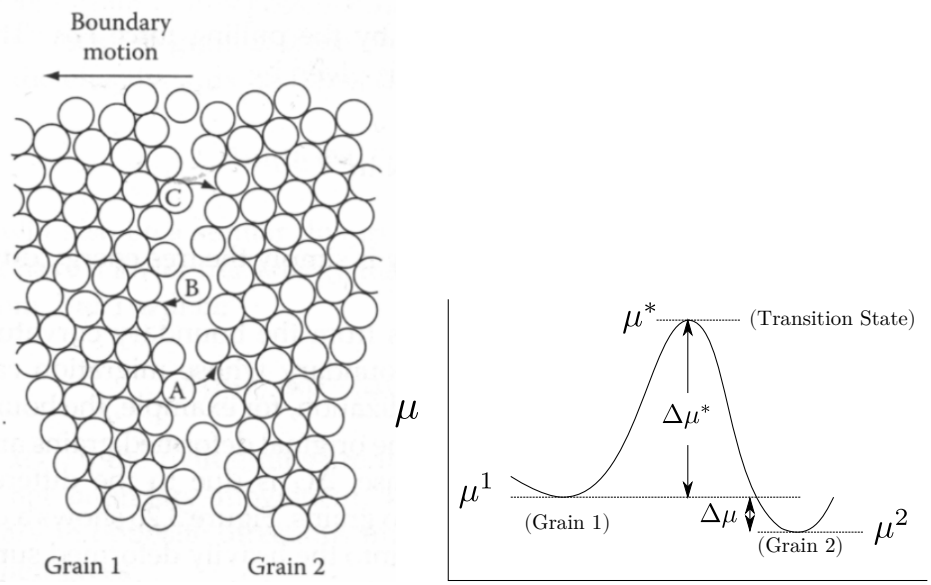


Figure 38.7

Now plot free energy as a function of position. Draw at equilibrium. An activation barrier exists that has a high of ΔG_a.
What if the interface is curved so that the curvature is toward grain 1 (grain 1 is smaller).
Redraw curve - now grain one has higher energy than grain 2.
This decreases μ a bit smaller than it was before. Also, have a negative μ for atoms going from grain 1 to grain 2.
Now the fluxes in the two directions are different. It's clear now that there is a net flux of atoms from grain 1 to grain 2. The actual flux from 1 to 2 is:

$$J_{1\to 2} = A_2 n_1 \nu_1 \exp\left(-\Delta\mu^*/RT\right) \tag{38.12}$$

A_2 is the probability that the atom is accommodated in grain 2.
n_1 the number of atoms that are able to make the jump (in molar units).
ν_1= vibrational frequency.
The flux in the backward direction (from 2 to 1) is given by:

$$J_{2\to 1} = A_1 n_2 \nu_2 \exp\left(-\left(\Delta\mu^* + \Delta\mu\right)/RT\right) = A_1 n_2 \nu_2 \exp\left(-\Delta\mu^*/RT\right)\exp\left(-\Delta\mu/RT\right) \tag{38.13}$$

If $\Delta\mu = 0$, we have equilibrium and $J_{1\to 2} = J_{2\to 1}$. A consequence of this is that the following $A_2 n_1 \nu_1 = A_1 n_2 v_2$.
If $\Delta\mu > 0$ the net flux is given by the difference:

$$J_{net} = J_{1\to 2} - J_{2\to 1} = A_2 n_1 v \exp\left(\frac{-\Delta\mu^*}{RT}\right)\left(1 - \exp\left(-\Delta\mu/RT\right)\right) \tag{38.14}$$

Assume that the $\Delta\mu \ll RT$, so we can use the approximation that $\exp(x) \approx 1 + x$ for small x. The expression for J_{net} then reduces to the following:

$$J_{net} = J_{1\to 2} - J_{2\to 1} = A_2 n_1 v \exp\left(\frac{-\Delta\mu^*}{RT}\right)\frac{\Delta\mu}{RT} \tag{38.15}$$

Now we can get an expression for the velocity of the grain boundary:

$$v = J_{net} V_m = A_2 n_1 V_m v \exp\left(\frac{-\Delta\mu^*}{RT}\right)\frac{\Delta\mu}{RT} \tag{38.16}$$

Can substitute this and get an expression for v. (expand exponential for small argument).
Now we define an interface mobility in the following way:

$$v = \frac{M\Delta\mu}{V_m} \tag{38.17}$$

Break things down into enthalpy and entropy:

$$\Delta\mu^* = \Delta H^* - T\Delta S^* \tag{38.18}$$

$$M = \frac{A_2 n_1 \nu_1 V_m^2}{RT}\exp\frac{\Delta S^*}{R}\exp\left[-\frac{\Delta H^*}{RT}\right] \tag{38.19}$$

What are the factors affecting M:

- Temperature
- Structure of the boundary (low angle vs. high angle boundary)
- Impurities (alloying elements)

Show Fig. 3.27
Impurities have a huge effect on grain boundary mobility. Grain boundaries are like garbage dumps for impurities.
Structure also plays a role. Tricks used in the heat treatment of high temperature superconductors.
Effect of impurities - Langmuir-Mclean model for grain boundary segregation.
X_{gB}= fraction of a monolayer adsorbed on the boundary

$$\left(\frac{X}{1-X}\right)_{gb} = \left(\frac{X}{1-X}\right)_{bulk}\exp\left[\frac{\Delta G_b}{RT}\right]$$

Here $\Delta G_B > 0$ for an element that adsorbs to the boundary.

- Strongly temperature dependent
- more dilute elements adsorb more (show figure)
- segregation affects the mobility of the boundaries

Show grain boundary composition vs. atomic solid solubility.

38.5.1 Transformation kinetics (crystallization, recrystallization):

Crystallization occurs by a nucleation and growth process, where crystalline regions nucleate from the parent and grow until they impinge on one another. A schematic example for a material that forms spherical crystals that impinge on one another to form individual grains is shown in Figure 38.8.

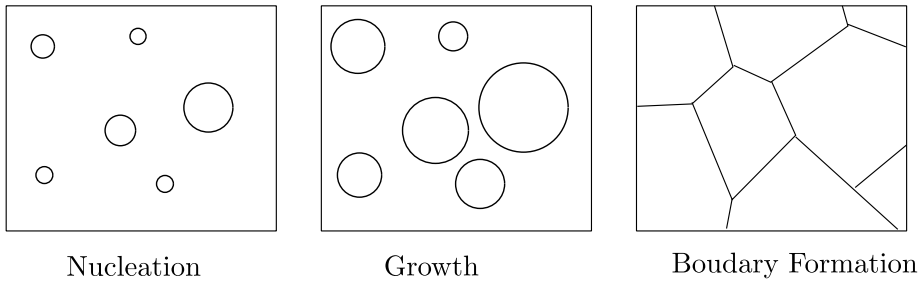


Figure 38.8: Schematic representation of a crystallization (or recrystallization) process.

What is time dependence of this type of transformation? We define the progress of the transition in terms of the transformed fraction, f, which has the sigmoidal time dependence illustrated in Figure 38.9.

$$f(t) = \frac{\text{fraction transformed at time t}}{\text{final fraction transformed}} \tag{38.20}$$

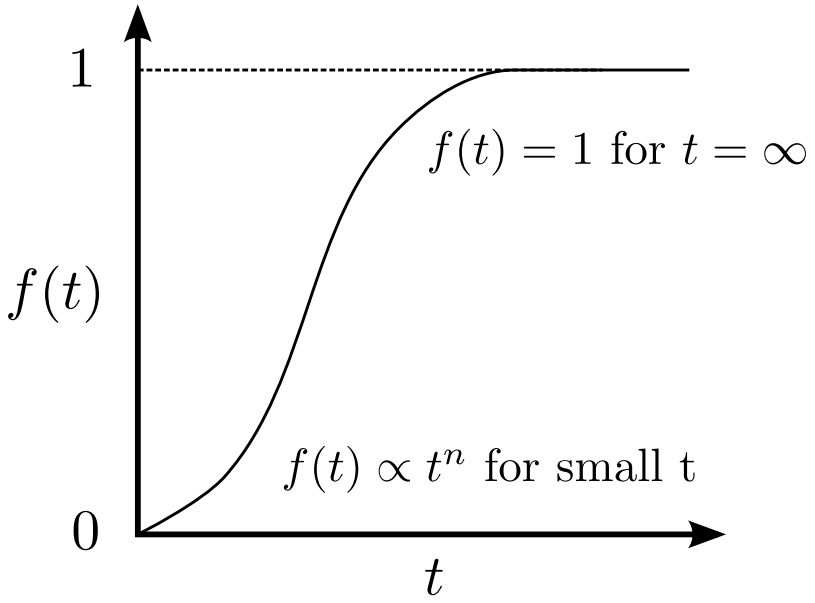


Figure 38.9: Time dependence of transformation.

We can derive an expression for $f\,(t)$ by assuming that nucleation occurs at a uniform rate, $\dot{N}_v$ (nuclei formed per volume per time) that does not depend on t. The volume of a single crystalline sphere of radius r is $4\pi r^3/3$. If the sphere forms at $t = 0$ and r increases linearly with a growth velocity of v, we have:

$$Vol(t) = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi(vt)^3 \tag{38.21}$$

If nucleation does not occur until $t = \tau$ we have:

$$Vol(t, \tau) = \frac{4}{3}\pi v^3 (t - \tau)^3 \tag{38.22}$$

The number of individual nuclei formed per unit unit volume during some time increment $d\tau$ is $\dot{N}_v d\tau$. Each of these have a volume given by Eq. 38.22. The total volume of crystallized material is given by integrating over all possible nucleus formation times:

$$f(t) = \dot{N}_v \int_0^t Vol(t, \tau) d\tau = \dot{N}_v \frac{4}{3}\pi v^3 \int_0^t (t - \tau)^3 d\tau = \dot{N}_v \frac{\pi}{3} v^3 t^4 \tag{38.23}$$

This equation is only valid for short times, since it neglects the fact that individual crystalline regions stop growing once they impinge on one another. In reality, $f(t)$ must reach an asymptotic value of 1 for $t = \infty$. A more detailed solution to the problem gives the following expression:

$$f(t) = 1 - \exp\left(-\frac{\pi}{3}\dot{N}_v v^3 t^4\right) \tag{38.24}$$

This is a specific example of the following more general expression, referred to as the Johnson-Mehl-Avrami-Kolmogorov (JMAK) equation:

$$f(t) = 1 - \exp(-kt^n) \tag{38.25}$$

where n is an empirical constant obtained from experimental data that is found to vary between 1 and 4. This is the simplest equation that has the basic behavior observed experimentally.

38.5.2 Relationship to Material Strength

Because grain boundaries impede dislocation motion, materials with a smaller grain size have a higher yield stress. Over a relatively large range of grain sizes, the relationship between the yield stress, σ_y, and the average grain size, d_g, is given by following relationship, referred to as the Hall-Petch relationship.

$$\sigma_y = \sigma_0 + \frac{k_y}{\sqrt{d}} \tag{38.26}$$

(See the Wikipedia page for more details)[8]

39 Three Phase Contact Lines

Two regions of space meet at a plane, and three regions of space meet at a line. These lines are important in a variety of problems in materials science. In Figure 39.1 we consider the most general case, where the 3 regions of space are labeled 1, 2 and 3. The junction between these three regions may correspond to three different material phases, or they may correspond to grain boundaries within a single phase region. In either case we refer to 1, 2 and 3 as 'phases', and refer to the line at which they meet as a '3-phase contact line'. The way in which the three phases meet at this contact line are specified by two angles. These two angles can be defined in a variety of ways, but we use the angles defined in Figure 39.1 as θ_1 and θ_2, which give the orientation of the 1/3 and 2/3 boundaries with respect to the 1/2 boundary.

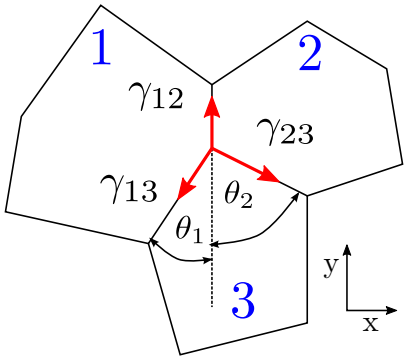


Figure 39.1: Force balance along a three-phase contact line.

At equilibrium θ_1 and θ_2 are related to the interfacial free energies of the 3 phases that meet at the contact line. Because interfaces have a contribution to the free energy that is associated with them, there is a thermodynamic driving force for any interface to shrink in area. As a result an interface exerts a force on the contact line along the direction of the interface, with a force per length of equal to the relevant interfacial free energy. At the three phase contact line three different forces, γ_{12}, γ_{23} and γ_{13}, are pulling on the contact line. At equilibrium the net force on the contact line is zero. We can obtain θ_1 and θ_2 by considering separate force balances in the directions parallel and perpendicular to the $1/2$ interface:

- Horizontal force balance (x direction):

$$\gamma_{13} \sin\theta_1 = \gamma_{23} \sin\theta_2 \tag{39.1}$$

- Vertical force balance (y direction):

$$\gamma_{12} = \gamma_{13} \cos\theta_1 + \gamma_{23} \cos\theta_2 \tag{39.2}$$

These are coupled, nonlinear equations that generally need to be solved numerically. An example procedure using MATLAB is given below in the section on wetting.

39.1 Wetting

Wetting refers to the case where one of the three phases is either air or vacuum. As an example, consider an oil droplet on the surface of water, as shown schematically in Figure 39.2. In order to determine the values of θ_1 and $\theta_2$in this case we need to know the following interfacial energies:

- γ_w: the surface free energy of water
- γ_o: the surface free energy of the oil
- γ_{ow}: the interfacial free energy between oil and water

Note that we refer to γ_w and γ_o as 'surface energies' and not 'interfacial energies' because one of the contact phases is air. We drop the subscript for air in this case, which is the convention that is commonly used. The horizontal force balance in this case can be written as follows:

$$\gamma_o \sin\theta_1 = \gamma_{ow} \sin\theta_2 \tag{39.3}$$

$$\gamma_w = \gamma_o \cos\theta_1 + \gamma_{ow} \cos\theta_2 \tag{39.4}$$

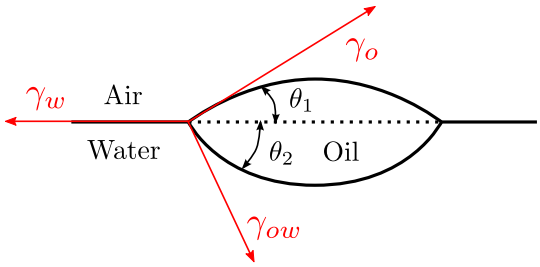


Figure 39.2: Force balance determining the shape of an oil droplet floating on the surface of water.

Here's a MATLAB script that solves the horizontal and vertical forces at the contact line for $\gamma_w = 72$, $\gamma_o = 30$ and $\gamma_{ow} = 50$. We give show the script here because it is an excellent example of the use of the MATLAB fsolve command to solve a series of coupled, nonlinear equations. (download link for script)

```
go=30; gow=50; gw=72;  % specify the different interfacial energies
verticalforce=@(theta) go*sind(theta(1))-gow*sind(theta(2)); % this is the net force in the
  vertical direction
horizontalforce=@(theta) gw-go*cosd(theta(1))-gow*cosd(theta(2));
ftosolve=@(theta) [verticalforce(theta), horizontalforce(theta)]; % write the function so that
  it returns the two components of the net force that both must be zero
thetaguess=[10,10]; % initial guess for theta1 and theta2
thetasolution=fsolve(ftosolve, thetaguess); % returns the solution as thetasolution
```

The values that we end up with are $\theta_1 = 33.9^\circ$ and $\theta_2 = 19.6^\circ$. This situation where the contact angles are greater than zero and the oil droplet forms a lens is referred to as partial wetting. What if the value of the oil surface energy (γ_o) is reduced so that the following inequality holds?

$$\gamma_o + \gamma_{ow} < \gamma_w \tag{39.5}$$

In this case there is not longer a solution to Eqs. 39.4 and 39.3, which means that the force on the contact line is never zero. Instead a force is directed outward so that the oil droplet spreads on the water surface, covering an enormous area and becoming exceptionally thin.

39.2 Grain Boundary Junctions

In this case the three phase contact line is actually a junction between grain boundaries, as opposed to a place where three distinct phases come into contact with one another. The important points are the following:

- If the three grain boundary energies for the boundaries meeting at the contact line are all equal to one another, $\theta_1 = \theta_2 = 60^\circ$. In other words, the interior angles between the different grains are all 120°.
- As a corollary to the point above, the boundaries of grains with fewer than 6 sides will be curved outward and the grain will tend to shrink, whereas grains with more than 6 sides will have boundaries that are curved and the grains will tend to grow.

40 Interphase Interfaces

Interfaces between two coexisting phases can have three types of interfaces: coherent, semicoherent and incoherent. Here we briefly describe these three types of interface.

40.1 Coherent Interfaces

Interfaces between different phases can be either coherent or incoherent. Coherent interfaces have atomic planes that are continuous across the interface as shown in Figure 40.1. As a result there are no broken bonds, and the interfacial energy, $\gamma_{\alpha\beta}$, is relatively low. It can be as low as 10^{-3}J/m^2, as in the case of an α/κ interface in the Cu-Si system. In general, $\approx 10 - 100\ \text{mJ/m}^2$ for coherent interfaces. Even for a fully coherent interface, there is still a

finite interfacial free energy though, because of unfavorable interactions between different atomic species that lead to phase separation in the first place. We refer to this interfacial energy as the coherent contribution, $\gamma_{coherent}$, so for coherent interfaces we have:

$$\gamma_{coherent} = \gamma_{ch} \tag{40.1}$$

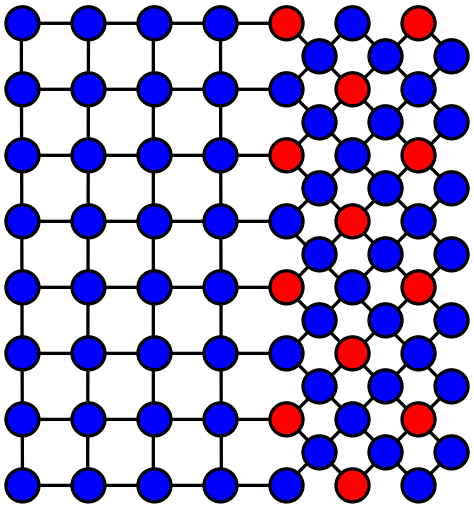


Figure 40.1: Schematic example of a coherent interface between two phases.

For interfaces between FCC and HCP crystal structures, only certain planes are coherent. For all planes to be coherent, both phases have to have the same crystal structure. However, they don't have to have the same lattice parameter. In this case elastic strains are generated.

40.2 Semicoherent Interfaces

Semicoherent interfaces are generally coherent interfaces with dislocations introduced at the interface to accommodate a small mismatch between the spacings of the atomic planes on either side of the interface as shown in Figure 40.2. These dislocations have an energy associated with them, which we refer to as γ_{st}, the structural component to the interfacial free energy. The total interfacial free energy for the semicoherent interface is given by adding this structural contribution to the chemical contribution.

$$\gamma_{semicoherent} = \gamma_{ch} + \gamma_{st} \tag{40.2}$$

Typical values for $\gamma_{semicoherent}$ are in the range of 0.1-0.5 $\mathrm{mJ/m^2}$. Note that

$$\gamma_{st} \propto \text{dislocation density}$$

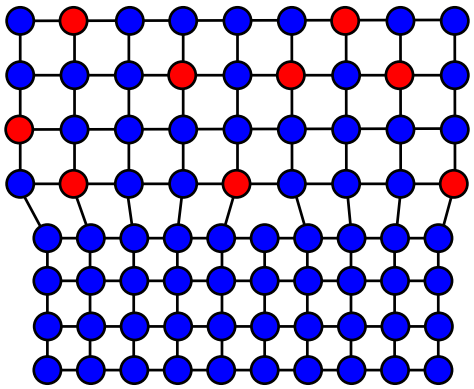


Figure 40.2: Example of a semicoherent interface.

40.3 Incoherent interfaces

If the lattice mismatch becomes too late, we the energy associated with all of the required dislocations to make the interface at least partially coherent is too high. Instead the interface becomes incoherent as shown in Figure the dislocation cores begin to overlap. The interface becomes incoherent, with $\gamma_{incoherent} \approx 500 - 1000\,\mathrm{mJ/m^2}$. $\gamma_{coherent}$ is relatively isotropic.

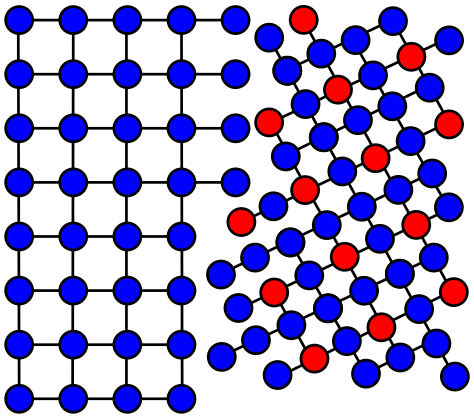


Figure 40.3: Example of an incoherent interface.

40.4 Case Study I: the Si-Ge system

A useful example of coherent and semicoherent interfaces is the silicon-germanium (Si-Ge) system. Both of these materials have the diamond cubic crystal structure, illustrated in Figure 40.4. The structure can be viewed as two interpenetrating FCC lattices. The Burgers vector for the most energetically favorable dislocation links atoms at the corner to the center of the face:

$$\vec{b} = \frac{a}{2}\langle 110 \rangle \tag{40.3}$$

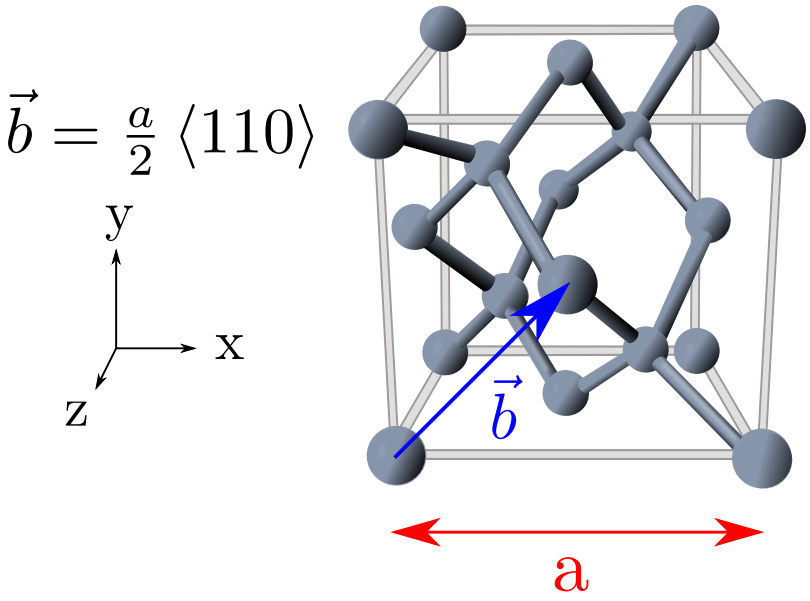


Figure 40.4: Diamond cubic crystal structure.

It makes sense that Si and Ge would have the same crystal structure, since the properties of these elements are very similar, with Ge residing just below Si on the periodic table. The lattice parameters are different however, and are given as follows:

Si: a_{Si}=5.431 Å

Ge: a_{Ge}=5.658 Å

This lattice parameter mismatch corresponds to a mismatch strain, δ_a, of 0.04 in this case, which we obtain from the relative difference between the lattice parameters:

$$\delta_a = \frac{a_{\text{Ge}} - a_{\text{Si}}}{a_{\text{Si}}} \tag{40.4}$$

Heteroepitaxy is an important thin film growth process where one material is deposited directly onto another material, with a coherent interface forming between the deposited film and the substrate. Suppose we deposit Ge onto a (010) surface of Si. The deposited Ge will have the same orientation, but will be strained by an amount δ because of the lattice parameter mismatch. Because Ge is larger than Si, we'll have some missing planes of atoms in the Ge film. The picture will look something like this:

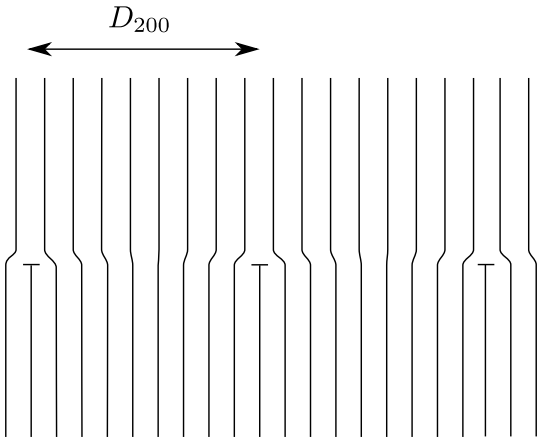


Figure 40.5: Dislocations in a semicoherent interface between materials with the same crystal structure but different lattice parameters.

These missing planes in the x direction are the (200) planes, since the spacing of these planes corresponds to the component of the Burgers vector that is in the x direction. From Eq. 37.2 we have the following for the spacing of the 200 planes in the two different phases:

$$d^{\text{Si}}_{200} = \frac{a_{\text{Si}}}{(h^2 + k^2 + \ell^2)} = \frac{a_{\text{Si}}}{2}$$

Similarly, we have $d^{\text{Ge}}_{200} = a_{\text{Ge}}/2 =$. The fractional difference in the interatomic spacings is the same as the fractional difference in the lattice parameters:
Define lattice misfit, δ:

$$\delta_{200} = \left(d^{\text{Ge}}_{200} - d^{\text{Si}}_{200}\right) / d^{\text{Si}}_{200} = \delta_a \tag{40.5}$$

We can rearrange this expression to get the following for d^{Ge}_{200}

$$d^{\text{Ge}}_{200} = d^{\text{Si}}_{200} + \delta_a d^{\text{Si}}_{200} \tag{40.6}$$

No we introduce a quantity D_{200} which is the average distance between dislocations in the x direction. Within this distance there are n (200) Ge planes but there are $n + 1$ (200) Si planes, so we have:

$$D_{200} = n d^{\text{Ge}}_{200} = (n+1)\, d^{\text{Si}}_{200} \tag{40.7}$$

From these two equations we obtain $n = 1/\delta_a$.
For $\delta_a = 0.04$ and $d^{Ge}_{200} = 2.82$ Å , we have $n = 25$ and $D_{200} = 70.5$ Å. We also have to account for the misfit in the other surface direction (the z direction in our case). The same argument holds in this direction as well, so we'll end up with dislocation spaced by D_{002} in this direction with $D_{002} = D_{200}$. Overall, we get a grid of dislocations at the interface, as shown in Figure 40.6.

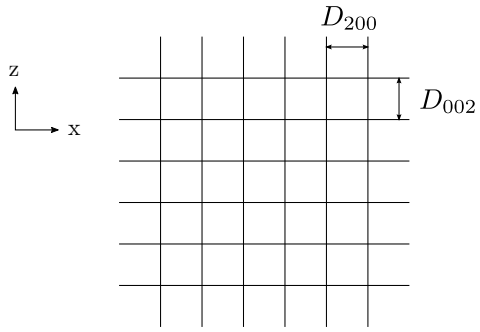


Figure 40.6: Grid of dislocations across a semicoherent interface.

40.5 Case Study II: The Cu/Al system

40.6 Second phase shape

The shape of a second phase particle is given by the Wulff construction. If the precipitate is fully coherent and the lattice parameters are very similar (no stress), then the interfacial energy is relatively isotropic and the precipitates are spherical. This situation is common in many precipitation hardened materials, like the Al-Cu system. In partially coherent precipitates the situation is much different, because the coherent interfaces have a much lower interfacial energy. In Figure 40.7 we show an example of the Wulff construction for a case where the interfacial free energy is radially symmetric with the exception of two deep cusps corresponding to the orientations for which coherent interfaces with the matrix can be formed. The coherent faces are flat, and the incoherent interfaces are curved. In addition, the aspect ratio of an equilibrium precipitate (length/width) is equal to the ratio of the incoherent interfacial free energy to the coherent interfacial free energy.

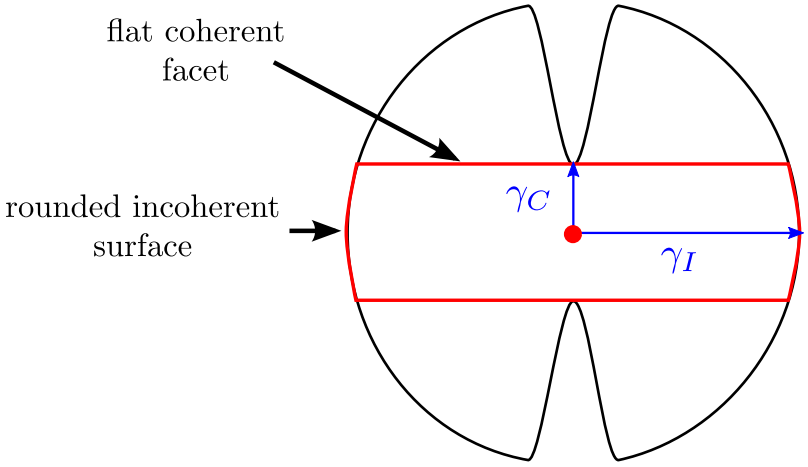


Figure 40.7: Wulff construction for a precipitate particle with an isotropic incoherent interfacial free energy, γ_I, and and a coherent interfacial free energy, γ_C.

40.7 Elastic Effects

Precipitates may transition from coherent to incoherent as they grow because of the elastic energy associated with the lattice distortions imposed by a lattice parameter mismatch across the interface. Consider a spherical precipitate of phase β in a matrix of phase α, as illustrated schematically Figure 40.8. We'll suppose for our case that α and β have the same crystal structure, and that for small values of precipitate radius r the interfacial free energy is isotropic. Because the interface is coherent it includes only a chemical component of the interfacial free energy, which we refer to here as γ_{ch}:

$$\gamma_{\alpha\beta} = \gamma_{ch} \tag{40.8}$$

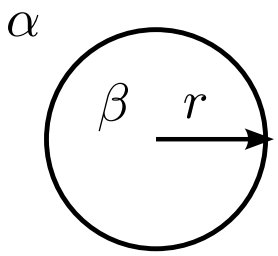


Figure 40.8: Beta precipitate in a matrix of alpha.

If the lattice parameters of the α and β phases do not match exactly, which will almost certainly be the case for any real system, there will be a positive elastic strain energy, W_{el} that we need to consider. For a spherical, coherent precipitate in an elastically isotropic medium, W_{el} is given by the following expression:

$$W_{el} = V^{\beta}\left[\frac{18\delta^2 G^{\alpha} K^{\beta}}{3K^{\beta} + 4G^{\alpha}}\right] \tag{40.9}$$

K^{β}= bulk modulus of β phase: $K^{\beta} = V\frac{\partial P}{\partial V^{\beta}}$
G^{α}= shear modulus of α phase: $G = \frac{\text{shear stress}}{\text{shear strain}}$
δ= misfit: $\delta = \frac{1}{3}\left[\frac{V_m^{\beta} - V_m^{\alpha}}{V_m^{\alpha}}\right]$
For cubic systems and for the small values of δ that are generally relevant here, δ is also equal to the fractional mismatch in the lattice parameter:

$$\delta = \left(\frac{a^{\beta} - a^{\alpha}}{a^{\alpha}}\right) \tag{40.10}$$

To provide some more insight into the behavior of Eq. 40.9 we consider the following three limiting cases:

1. compressible precipitate in a rigid matrix: $G^{\alpha} >> K^{\beta}$

$$\frac{W_{el}}{V^{\beta}} \approx \frac{9\delta^2 K^{\beta}}{2} \tag{40.11}$$

2. rigid precipitate in a deformable matrix: $K^{\beta} >> G^{\alpha}$

$$\frac{W_{el}}{V^{\beta}} \approx 6\delta^2 G^{\alpha} \tag{40.12}$$

3. Precipitate and matrix with the same elastic properties.

   An isotropic material is characterized by just two independent elastic constants. It's convenient to express K in terms of G and ν, using the following expression (note that the Wikipedia page has a very useful summary of the relationships between different elastic constants for an isotropic material)

$$K = \frac{2G\nu}{1 - 2\nu} \tag{40.13}$$

   As an example, we can take $\nu = 1/3$, in which case we get $K = 2G$, and W_{el} is given by the following expression.

$$\frac{W_{el}}{V^{\beta}} = 3.6\delta^2 G \tag{40.14}$$

In each of these cases the most important points to keep in mind are the following:

1. The elastic strain energy is proportional to the volume of the precipitate.
2. The elastic strain energy is proportional to the square of the lattice mismatch.

We are now in a position to compare the overall free energy of coherent and incoherent precipitates, and to see how each of these depend on the precipitate size. For a coherent precipitate we just need to add the elastic strain energy to the chemical contribution to the interfacial free energy:

$$W_{coh} = 4\pi r^2 \left(\gamma_{ch}\right) + \frac{4}{3}\pi r^3 \frac{W_{el}}{V^\beta} \tag{40.15}$$

Incoherent precipitates have a larger value of $\gamma_{\alpha\beta}$ because we also need to account for the structural component of the interfacial free energy that arises for the dislocations that are present at the interface between the α and β phases. However, the strain energy in the bulk is reduced to zero, so have the following for W_{inc}, the total excess free energy of an incoherent precipitate:

$$W_{inc} = 4\pi r^2 \left(\gamma_{ch} + \gamma_{st}\right) \tag{40.16}$$

For suitably small values of r, the contribution to the total free energy that scales with r^2 will always be more important than the contribution that scales with r^3. As a result W_{coh} will always be less than W_{inc} for sufficiently small values of r. As the precipitate grows and r increases, the contribution that scales with r^3 will become more important, and W_{coh} will exceed W_{inc}. The two free energies are equal to one another at a critical radius, $r_{crit,}$ which we illustrate schematically in Figure 40.9. Precipitates will remain coherent for sizes below r_{crit} and will become incoherent for sizes larger than r_{crit}. The value of r_{crit} is the value of r for which W_{coh} and W_{inc} are equal to one another. From Eqs. 40.15 and 40.16 we get:

$$r_{crit} = 3\gamma_{st}\frac{V^\beta}{W_{el}} \tag{40.17}$$

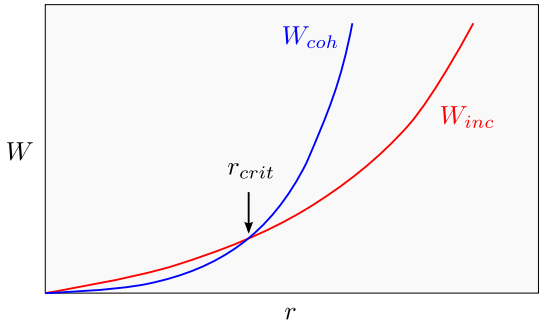


Figure 40.9: Size dependence of the overall excess free energies of spherical coherent and incoherent precipitates.

40.8 Effects of Elastic Anisotropy

in reality, no crystalline material is completely isotropic. An FCC crystal, for example, is generally stiffest along the [110] directions and softest along [100] directions. This is because the linear density of atoms is highest along the [110] direction, where in a hard sphere model of the crystal structure the atoms are in contact with one another. As a result coherent precipitaes end up with facets perpendicular to the 'soft' [100] directions. Faceting becomes more important as the precipitates grow (assuming they stay coherent), since the elastic contribution to the energy scales with the volume of the precipitate, whereas the total surface area scales with the 2/3 power of the volume.

41 Thin Film Growth

We can now consider wetting in solid systems, where elastic strain energy often plays a role. Here we consider the example of a thin, germanium film that is deposited on a silicon substrate. Both of these elements have the diamond cubic crystal structure, with a lattice parameter mismatch of about 4%. When an element like germanium is deposited from the vapor phase, the atoms land individually on the substrate as illustrated in Figure 41.1. If the atoms have sufficient mobility, the resulting film will be determined by the structure that minimizes the free energy. For the Ge/Si system, the equilibrium structure consists of a thin, continuous wetting layer below isolated Ge islands.

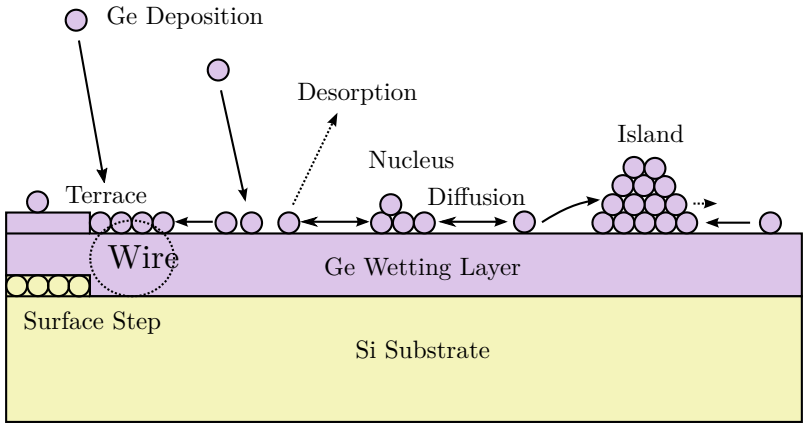


Figure 41.1: Schematic representation of Ge deposition on a single crystal Si substrate.

To understand the full behavior of the system, it is useful to develop a plot of the overall free energy per unit area of the system as a function of the Ge film thickness, t_{Ge}, which we show in Figure 41.2. The following contributions to the free energy need to be considered:

- γ_{Si}: The surface free energy of the Si substrate
- γ_{Ge}: The surface free energy of Ge
- γ_{Ch}: The chemical contribution to the free energy of the Si/Ge interface
- γ_{St}: The structural component to the free energy of the Si/Ge interface
- W_{el}/A: The strain energy per unit area within the Ge film.

The full thickness dependence of the free energy can be understood by investigating the way that these contributions contribute to the overall free energy as the Ge film thickness increases:

- For $t_{Ge} = 0$ the free energy is simply γ_{Si}, the surface free energy of the silicon substrate.
- For very thin Ge films the elastic strain energy within the Ge film does not contribute significantly to the strain energy, so the overall value of W/A is the sum of the energies of the Si/Ge and Ge/vapor interfaces. Because the Si/Ge interface is fully coherent for sufficiently thin Ge films, its interfacial free energy is just the chemical part, γ_{ch}, so the overall free energy for very thin Ge films is $\gamma_{Ge} + \gamma_{Ch}$. This free energy is less than γ_{Si}, so the system is in the wetting regime.
- As the film thickness increases the Ge film remains coherent, but W/A increases linearly with thickness, according to the thickness dependence of W_{el}.
- When the elastic energy exceeds the structural component of the Si/Ge interfacial free energy associated with the loss of full coherence, the Ge film becomes incoherent. The elastic energy W_{el} is now large enough so that it is energetically favorable for the Si/Ge interface to be less coherent.

For film thicknesses larger than the thickness for which W_{el} is a minimum, a thin Ge layer with a thickness corresponding to the thickness at the free energy minimum will coexist with Ge droplets that are much thicker - the 'islands' shown in Figure 41.1.

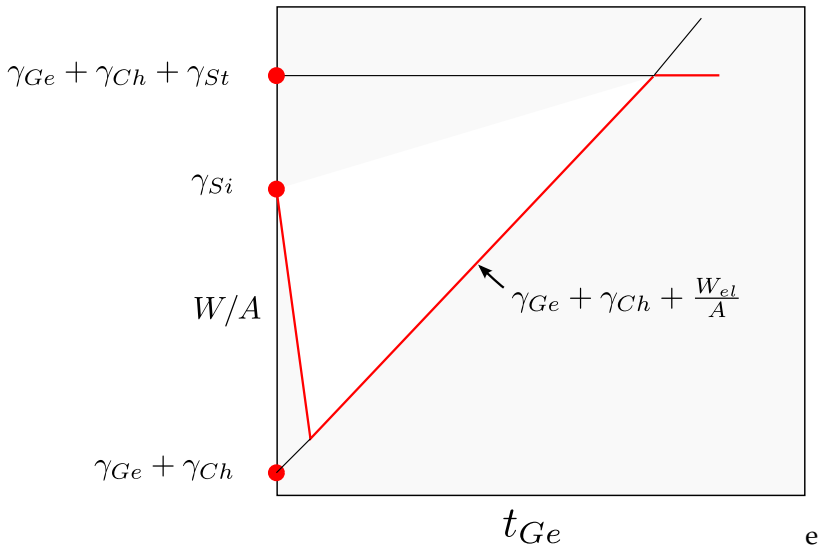


Figure 41.2: Free energy as a function of film thickness for an epitaxial Ge film on a single crystal Si substrate.

42 Review Questions

42.1 Diffusion

1. What are the tracer, interdiffusion and intrinsic diffusion coefficients? Which are purely kinetic quantities, and which involve thermodynamics? How are these diffusion coefficients related to one another?
2. What is the Kirkendall effect? How can you figure out the direction of vacancy motion?
3. What is the mechanism by which vacancies are either created or destroyed?
4. Under what conditions will voids form, and where will they form (Lab 1)?
5. Where must dislocations be created or destroyed to maintain an equilibrium vacancy concentration?

42.2 Dislocations

6. Explain the physical origin of shear bands observed on the surface of a plastically deformed metal.

1. What is the critical resolved shear stress? How is it calculated for a tensile experiment?
2. What is the value of the ratio of the theoretical critical resolved shear stress (in the absence of dislocations) to a typical experimental value of this same quantity.
3. Define an edge and a screw dislocation in terms of their Burgers vectors and the sense vectors.
4. What are the Burgers vectors of perfect dislocations in the simple cubic, face-centered cubic and body-centered cubic lattices?: use a drawing to illustrate the Burgers vectors. What are the magnitudes of these vectors.
5. Explain how to make pure edge or screw dislocations by cutting and slipping operations.
6. Explain how to make a curved dislocation by cutting and slipping operations. Demonstrate that its character varies from pure screw to pure edge as one moves along the curved dislocation line.
7. Demonstrate carefully how the Burgers vector of an edge or a screw dislocation is determined employing a Burgers circuit.
8. What is the difference between a right-hand and a left-hand screw dislocation?

9. Given $\vec{b}$ and $\vec{s}$, how do you know where the slip plane is, and what direction the dislocation will move in response to an applied shear stress.
10. Explain why a pure screw dislocation does not have a unique glide or slip plane.
11. Explain why a pure edge dislocation has a unique glide or slip plane.
12. Explain the main differences between glide motion of a dislocation and climb motion?
13. How does the temperature dependence of glide differ from the temperature dependence of climb? Which is more important at lower temperatures, and why?
14. Explain how climb of an edge dislocation can relieve a super- or subsaturation of vacancies or self-interstitial atoms.
15. Do pure screw dislocations cross-slip?
16. What are the physical origins of the energy of a dislocation line?
17. If no external stress is supplied to a dislocation loop, why does the loop shrink until it disappears from a crystal?
18. Describe qualitatively the state of stress associated with a pure edge dislocation and compare it with the state of stress associated with a pure screw dislocation.
19. What is the physical significance of F_s^τ and F_s^r?
20. Describe qualitatively the state of stress associated with a pure edge dislocation and compare it with the state of stress associated with a pure screw dislocation.
21. When do parallel edge dislocations move toward each other? Under what conditions do they move away from each other?
22. When do parallel screw dislocations move toward or away from each other?
23. How does a Frank-Read source work?
24. How does the stress need to be oriented to either expand or contract a dislocation loop?
25. How does precipitation hardening work? What is the role of the precipitate spacing and of Frank-Read sources? Why does the precipitate spacing matter?

42.3 Solid/Liquid and Solid/Vapor Interfaces

26. Why does the surface energy of a crystal depend on the orientation?
27. How does the interfacial energy affect the melting temperature for a small droplet?
28. How does the interfacial free energy affect precipitate solubility?
29. How can the surface energy be approximated from the crystal structure and thermodynamic data?
30. What is the Wulff construction and how is it used?

42.4 Grain Boundaries

31. What are the 5 parameters needed to fully characterize a grain boundary?
32. What special relationship exists between these parameters for twist and tilt boundaries?
33. How are dislocations arranged for low angle twist and tilt boundaries?
34. How does the force balance at the triple junctions of grains affect grain shape?
35. What happens to grains with different numbers of sides during grain growth?

36. What is the role of curvature in grain growth?
37. Derive the expected time dependence of the grain size for grain growth driven by curvature.
38. What is a twin boundary? For what crystal structures is it observed?

42.5 Interphase Interfaces

39. What is the general condition governing the equilibrium shape of a precipitate when there is no contribution from the elastic energy?
40. (a) What is the expression for the total elastic strain energy of a precipitate, if the matrix is elastically isotropic? (b) Explain the physical significance of each term in this equation.
41. How does the shape of a misfitting precipitate in an elastically anisotropic system vary with particle size? Why is this variation observed?
42. (a) Derive an approximate expression for the critical radius at which a spherical precipitate loses its coherency and become semi-coherent or incoherent. (b) Explain why precipitates often exceed this calculated critical value.
43. Describe the nature of a solid-liquid interface and how it differs from a solid-solid interface.
44. What is the significance of the chemical and structural components to the interfacial free energy between solids?
45. Why do Ge films on Si form islands on top of a thin wetting layer?

42.6 Crystallization or Recrystallization

46. Make a plot of the volume fraction, f, transformed, 0 to 1.0, as a function of time for a general phase transformation occurring by nucleation and growth.
47. Derive for $f(t) \ll 1$ the Johnson-Mehl-Avrami-Kolomogorov (JMAK) equation for the volume fraction transformed as a function of time, $f(t)$, under the assumption that a specimen contains a number n of effective point heterogeneities per unit volume and nucleation occurs at all of these points very quickly and that the nuclei have a spherical shape. State any and all assumptions made.
48. Derive a JMAK equation, for $f(t) \ll 1$, for the case where all the nuclei do not form at time $t = 0$, but rather form randomly throughout a specimen at a constant rate, which is N nuclei formed per unit volume per unit time of untransformed material. State any and all assumptions made.

42.7 MATLAB

49. How do I make plots suitable for publication when those plots generated by Excel just aren't good enough anymore?
50. How do I write arbitrary functions that can be plotted or compared with experimental data?
51. How do I fit a user-defined function to experimental data?
52. How do I use fsolve to solve a system of coupled equations?
53. How can I run a write and run a simple simulation in MATLAB (like the vacancy diffusion simulation)
54. What is a polar plot and how can I generate one?
55. How do I solve the Wulff construction numerically?

43 316-1 Problems

43.1 Introduction

Send an email to Prof. Shull (k-shull@northwestern.edu) and Nari (NariJeon2015@u.northwestern.edu) with the following information:

(a) Anything about yourself (why you are interested in MSE, previous work experience, etc., outside interests apart from MSE) that will help me get to know you a bit (feel free to be brief - any info here is fine).

(b) Your level of experience and comfort level with MATLAB. Be honest about your assessment (love it, hate it, don't understand it, etc.).

(c) Let us know if you have NOT taken 314 or 315 for some reason.

43.2 Diffusion

Consider a diffusion couple with composition C_1 as $z \to -\infty$ and C_2 as $z \to \infty$. The solution to the diffusion equation is:

$$C(z,t) = \frac{C_1 + C_2}{2} - \frac{C_1 - C_2}{2}\text{erf}\left(\frac{z}{2\sqrt{Dt}}\right)$$

where $\text{erf}(y) = \frac{2}{\pi}\int_0^y e^{-t^2}dt$. Note that in the definition of the error function t is a dummy variable of integration, thus the error function is a function of y. Also, erf(0)=0, and erf(∞)=1. You will determine if these boundary conditions are correct.

(a) Show that the boundary conditions at $z = \pm\infty$ are satisfied by the solution.

(b) Does the composition at $z = 0$ vary with time? If not, what is its value? Why do you think this is the case?

(c) Write the solution in terms of $\eta = z/t^{1/2}$.

(d) Show that the solution in terms of satisfies the following diffusion equation that is written in terms of η:

$$D\frac{d^2C}{d\eta^2} + \frac{\eta}{2}\frac{dC}{d\eta} = 0$$

You will needed to take a derivative of the error function. Leibniz's formula for the differentiation of integrals will be helpful:

$$\frac{d}{dz}\int_{h(z)}^{g(z)} f(t)\,dt = \frac{dg(z)}{dz}f(g(z)) - \frac{dh(z)}{dz}f(h(z))$$

A diffusion couple including inert wires was made by plating pure copper on to a block of α-brass with $X_{Zn} = 0.3$, as shown in Figure 43.1. After 56 days at 785 °C the marker velocity was 2.6x10^{-8} mm/s, with a composition at the markers of $X_{Zn} = 0.22$, and a composition gradient, $\partial X_{Zn}/\partial z$ of 0.089 mm^{-1}. A detailed analysis of the data gives $\tilde{D} = 4.5x10^{-13}$ m^2/s for $X_{Zn} = 0.22$. Use these data to calculate D_{Zn} and D_{Cu} for $X_{Zn} = 0.22$. How would you expect D_{Zn}, D_{Cu} and $\tilde{D}$ to vary as a function of composition?

In class we developed an expressions for J'_a. Show that $J'_a = -J'_b$. (Recall that these primed fluxes correspond to fluxes in the laboratory frame of reference).

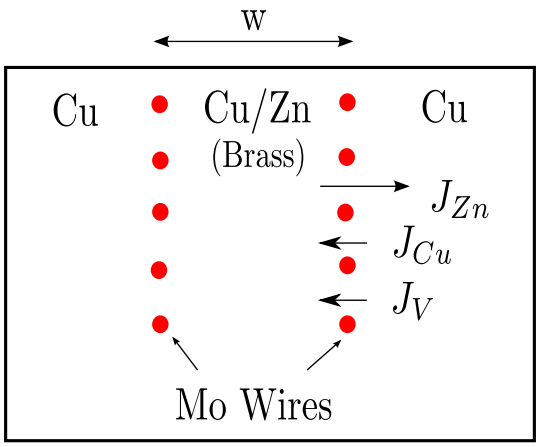


Figure 43.1: Experimental Geometry for the Kirkendall experiment.

Consider two binary alloys with compositions $X_b = X_1$ and $X_b = X_2$, shown in Figure 43.2 along with the free energy curves for α and β phases formed by this alloy. Draw the composition profile across the interface shortly after the two alloys are brought into contact with one another, assuming that the interface is in "local equilibrium", i.e. the interface compositions are given by the equilibrium phase diagram. Describe the direction in which you expect the B atoms to diffuse on each side of the interface.

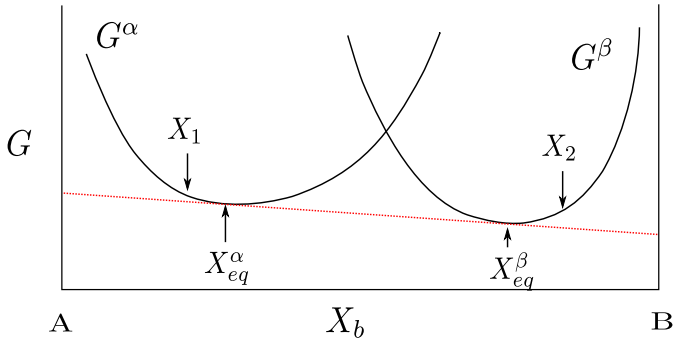


Figure 43.2: Free energy curves for a model A/B alloy.

The following MATLAB script runs the vacancy simulation shown in class. It saves the data into a 'structure' called output, which can be loaded into MATLAB later. The file can be downloaded from this link:
http://msecore.northwestern.edu/316-1/matlab/vacancydiffusion.m

```
tic % start a time so that we can see how long the program takes to run
n=30;  % set the number of boxes across the square grid
vfrac=0.01;  % vacancy fraction
matrix=ones(n);
map=[1,1,1;1,0,0;0,0,1];  % define 3 colors:  white, red, blue
figure
colormap(map)  % set the mapping of values in 'matrix' to a specific color
caxis([0 2]) % range of values in matrix goes from 0 (vacancy) to 2
% the previous three commands set things up so a 0 will be white, a 1 will
% be red and a 2 sill be blue
matrix(:,n/2+1:n)=2;  % set the right half of the matrix to 'blue'
i=round(n/2);   % put one vacancy in the middle
j=round(n/2);
matrix(i,j)=0;
imagesc(matrix); % this is the command that takes the matrix and turns it into a plot
t=0;
times=[1e4,2e4,5e4,1e5];
```

```
showallimages=1; % set to zero if you want to speed things up by not showing images,  set to 1
  if you want to show all the images during the simulation

%% now we start to move things around
vacancydiff.matrices={};  % makea blank cell array
while t<max(times);
    t=t+1;
    dir=round(4*rand+0.5);
    if dir==1
        in=i+1;
        jn=j;
        if in==n+1; in=1; end
    elseif dir==2
        in=i-1;
        jn=j;
        if in==0; in=n; end
    elseif dir==3
        in=i;
        jn=j+1;
        if jn>n; jn=n; end
    elseif dir==4
        in=i;
        jn=j-1;
        if jn==0; jn=1; end
    end
    % now we need to make switch
    neighborix=sub2ind([n n],in,jn);
    vacix=sub2ind([n n],i,j);
    matrix([vacix neighborix])=matrix([neighborix vacix]);
    if showallimages
        imagesc(matrix);
        drawnow
    end
    if ismember(t,times)
        vacancydiff.matrices=[vacancydiff.matrices {matrix}]; % append matrix to output file
        imagesc(matrix);
        set(gcf,'paperposition',[0 0 5 5])
        set(gcf,'papersize',[5 5])
        print(gcf,['vacdiff' num2str(t) '.eps'],'-depsc2')
    end
    i=in;
    j=jn;
end
vacancydiff.times=times;
vacancydiff.n=n;
save('vacancydiff.mat','vacancydiff') % writes the vacancydiff structure to a .mat file that we
   can read in later
toc
```

(a) Run the vacancy diffusion script, and include in your homework the .jpg files generated for time steps of 1e4, 2e4, 4e4 and 1e5.

(b) For the longest time step, develop a plot of average composition along the horizontal direction.

Here is the MATLAB script that I used to do this (available at http://msecore.northwestern.edu/316-1/matlab/vacancyplot.m):

```
load vacancydiff % load the previously saved output.mat file
figure
figformat % not necessary, this is the standard initialization script I use to standardize
   what my plots look like
n=vacancydiff.n;
matrix=vacancydiff.matrices{4};
matrixsum=sum(matrix,1); % sum of each column in the matrix
plot(1:n,matrixsum/n,'+b')
xlabel ('z')
ylabel ('C')
print(gcf,'../figures/vacancyplot.eps','-depsc2') % this creates an .eps file, which I use
   for the coursenotes but which may not be as useful for many of you as the jpg file
```

```
% saveas(gcf,'vacancyplot.jpg') % this is what to do if you just want to save a .jpg file
```

Note that 'figformat' is NOT a matlab command. This line calls another file called names figformat.m that includes a few commands to standardize plots that I am making for this class. Here's what it looks like:

```
set(0,'defaultaxesbox', 'on') % draw the axes box (including the top and right axes)
set(0,'defaultlinelinewidth',2)
set(0,'defaultaxesfontsize',16)
set(0,'defaultfigurepaperposition',[0,0,7,5])
set(0,'defaultfigurepapersize',[7,5]')
```

(c) In the previous problem set we obtained concentration profiles from the MATLAB. Now we'll take these concentration profiles and see if they are consistent with the solution to the diffusion equation.

i For each of the 4 time points used in the simulation, plot the concentration profile and fit it to the error function to the diffusion equation, using the interfacial width, w, ($w = 2\sqrt{Dt}$) as a fitting parameter:

$$C(z,t) = \frac{C_1 + C_2}{2} - \frac{C_1 - C_2}{2}\text{erf}\left(\frac{z}{w}\right)$$

Note: This problem is a curve fitting exercise in MATLAB. The most frustrating part is getting all the syntax right, but once you know the proper format for the MATLAB code, it's pretty straightforward. Take a look at the section entitled 'Fitting a Function to a Data Set' in the MSE MATLAB help file (http://msecore.northwestern.edu/matlab.pdf). This section includes a MATLAB script that you can download and modify as needed.

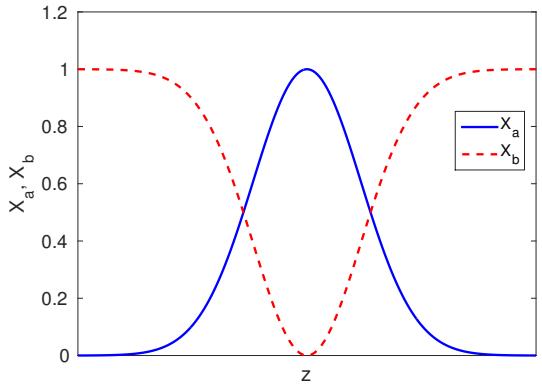
ii Plot w^2 as a function of the time (expressed here as the number of time steps in the simulation). Obtain the slope of a line drawn through the origin that best fits the data.

iii When diffusion occurs by a vacancy hopping mechanism in a 2-dimensional system like the one used in our simulation, the diffusion coefficient is given by the following expression:

$$D = KX_v\Gamma a^2$$

Here is the average hop frequency for any given vacancy and a is the hopping distance. From the the slope of the curve of w vs. the total number of jumps, extract an estimated value for K.

A region of material with a different composition is created in an infinitely long bar. The following plot shows the mole fraction of component A as a function of position. Assume that the intrinsic diffusion coefficient of the A atoms is twice as large as the intrinsic diffusion coefficient for the B atoms.



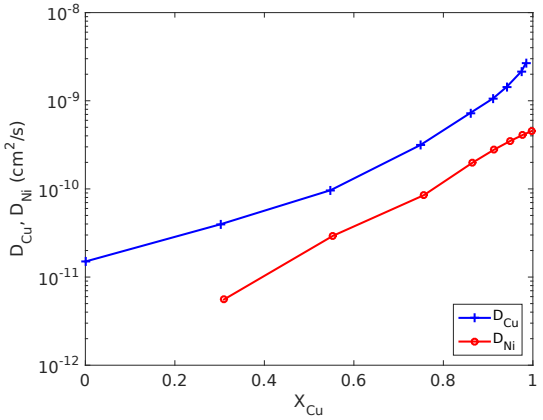
(a) Plot the flux of A and the flux of B relative to the lattice as a function of position in the graph above.

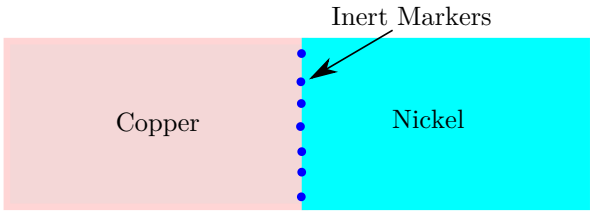
(b) Plot the vacancy creation rate as a function of position in the graph above.

(c) Plot the flux of A and B in the lab frame as a function of position in the graph above.

(d) Plot the lattice velocity as a function of position in the graph below. What are the physical implications of this plot?

The values for the intrinsic diffusion coefficients for Cu and Ni in a binary Cu/Ni alloy are shown below on the left (note that Cu and Ni are completely miscible in the solid state). A diffusion couple is made with the geometry shown below on the right.

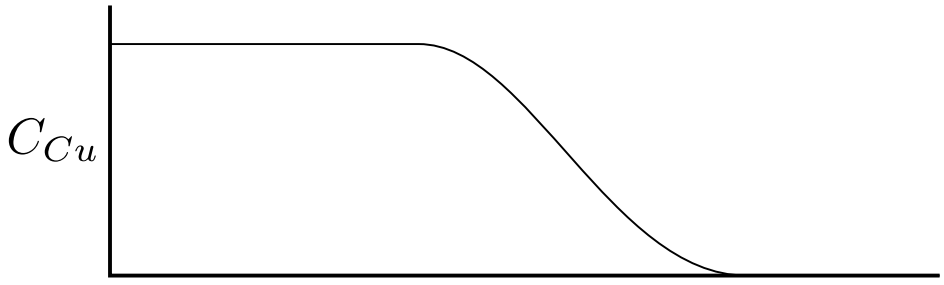




(a) What is the value of the interdiffusion coefficient $\tilde{D}$, for an alloy consisting of nearly pure Nickel?

(b) Will the markers placed initially at the Cu/Ni interface move toward the copper end of the sample, the nickel end of the sample, or stay at exactly the same location during the diffusion experiment.

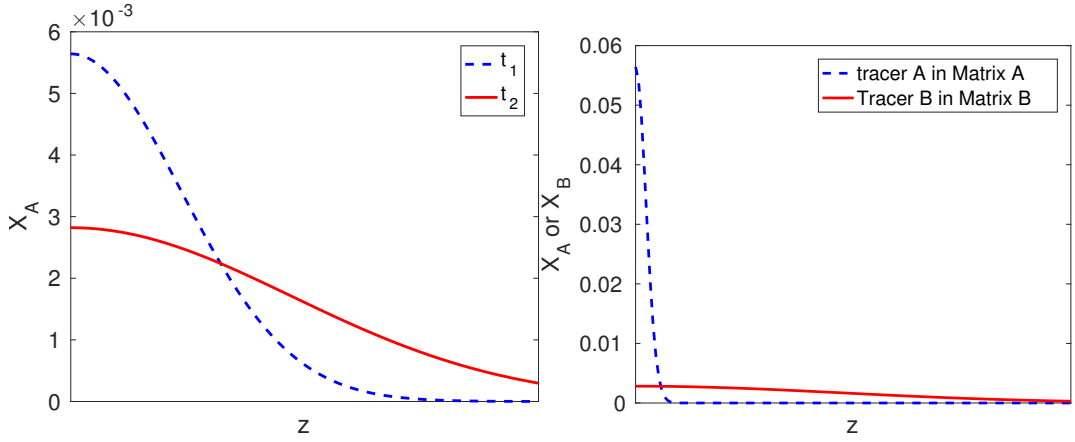
(c) The copper concentration across the sample is sketched below after diffusion has occurred for some time.



(d) Sketch the fluxes of Copper, Nickel and vacancies, defining positive fluxes as those moving to the right.

(e) Now sketch the rate at which vacancies are created or destroyed within the sample in order to maintain a constant overall vacancy concentration throughout.

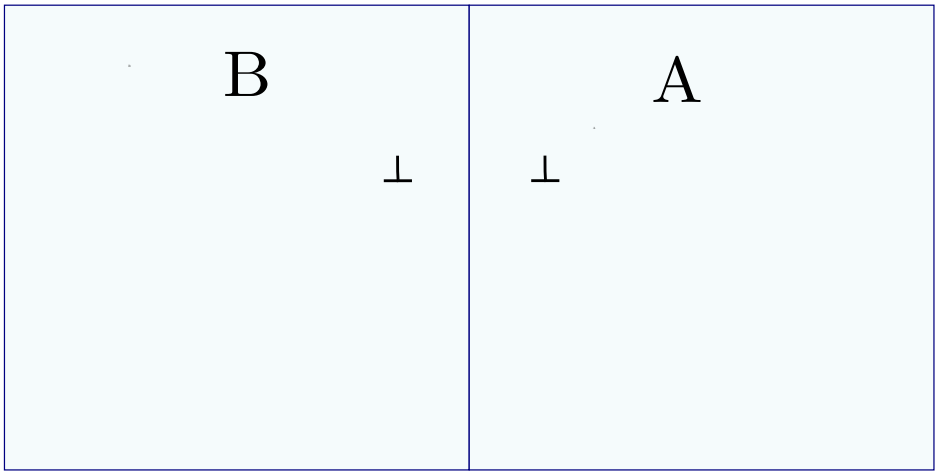
An experiment is performed to determine the tracer diffusion coefficient of metal A in a matrix of metal B. This is done by depositing a very thin film of metal A onto the surface of metal B and measuring the concentration profile of metal A into the depth of the material at different times. The concentration profiles in the left figure below are obtained at two times, t_1 and t_2:



(a) Estimate the ratio t_2/t_1

(b) Now suppose we measure the self diffusion coefficients of A and B. Performing measurements at the same time and temperature gives the concentration profiles shown in the figure above to the right. Which element (A or B) do you expect has the highest melting temperature, and why?

(c) Now we'll make a diffusion couple with element A on the right half and element B on the left half. Assume that A and B are miscible at the diffusion temperature, and form a one phase alloy. Mark up the following diagram as directed on the next page:



i Put an arrow labeled 'M' on the diagram indicating the direction that inert markers placed originally at the interface will move.

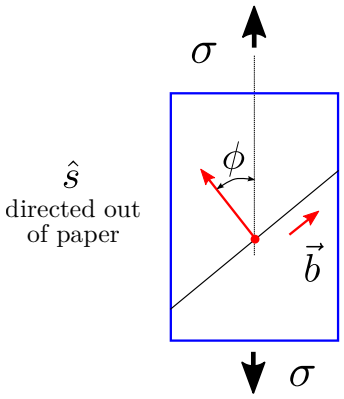
ii Put an arrow labeled 'V' on the diagram indicating the the net vacancy flux due to diffusion in the sample.

iii Put a 'C' on the region of the sample where you expect vacancies to be created, and a 'D' on the sample where you expect vacancies to be destroyed, assuming that the total vacancy concentration must remain at equilibrium.

iv Two edge dislocations are also indicated in the diagram. Place arrows on top of each dislocation to illustrate he directions you expect these dislocations to move in order to create or destroy the vacancies from part iii.

43.3 Stress and Strain

A tensile stress, σ, is applied to a single crystal of zinc, which has an HCP structure. The close packed planes of atoms (the slip plane for an HCP material) is oriented with its surface normal in the plane of the paper, inclined to the tensile axis by an angle ϕ as shown below, with $\phi = 30°$. Assume that the critical resolved shear stress for motion of the dislocation is 50 MPa (5x10^7 Pa). The shear modulus of Zn is 43 GPa (4.3x10^{10} Pa) and its atomic radius is 0.13 nm.



(a) Is this an edge dislocation, a screw dislocation, or a mixed dislocation, and how do you know?

(b) Put an arrow on the drawing above to indicate the direction in which the dislocation moves under an applied tensile stress.

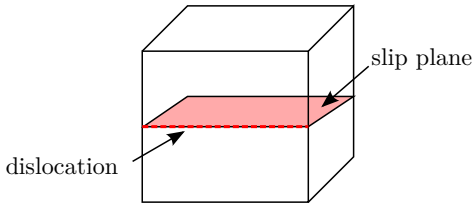
(c) Calculate the tensile yield stress for this sample.

(d) Suppose that the slip plane is oriented so that $\vec{b}$ is still in the plane of the paper, but that ϕ is increased to 60°. Will the yield stress increase, decrease or stay the same.

(e) Suppose that the dislocation is impeded by pinning points (precipitates, for example), that are uniformly spaced and separated by 1 μm (10^{-6} m). The resolved shear stress is determined by the stress required to move the dislocation around these pinning points. Use the information given in this problem to determine the energy per length of the dislocation. Compare this to the expressions given for the energies of edge and screw dislocations to see if it makes sense.

43.4 Dislocation Structure

A right handed screw dislocation initially located in the middle of the front face of the sample shown below moves toward the back of the sample in response to an applied shear stress on the sample.



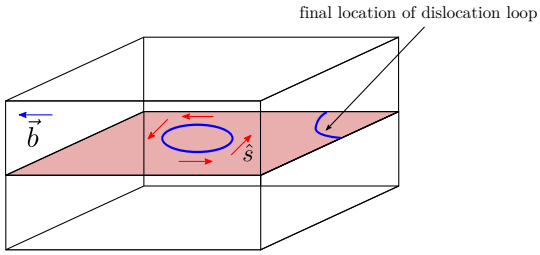
(a) Sketch the shape of the sample after the dislocation has propagated halfway through the sample, and again when it has propagated all the way through the sample. Use arrows to specify the shear force that is being applied.

(b) Repeat part a for a left-handed screw dislocation.

Draw an edge dislocation and on the same figure dot in the positions of the atoms after the dislocation has shifted by $\vec{b}$.

How can two edge dislocations with opposite Burgers vectors meet to form a row of vacancies? How can they meet to form a row of interstitials? Draw pictures of both situations.

Given a crystal containing a dislocation loop as shown in the following figure:



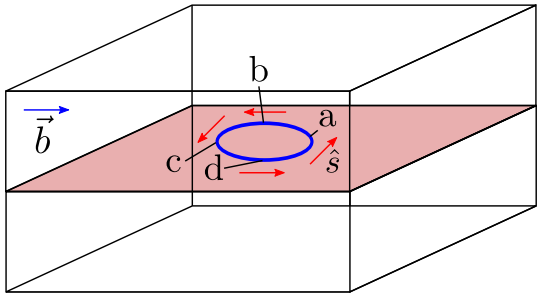
Let the loop be moved (at constant radius) toward a corner until three-fourths of the loop runs out of the crystal. This leaves a loop segment that goes in one face and comes out the orthogonal face. Sketch the resultant shape of the crystal, both above and below the slip plane.

Given a loop with a Burger's vector that is perpendicular everywhere to the dislocation line, determine the resulting surface morphology after the loop propagates out of the crystal. Assume that the loop moves only by glide.

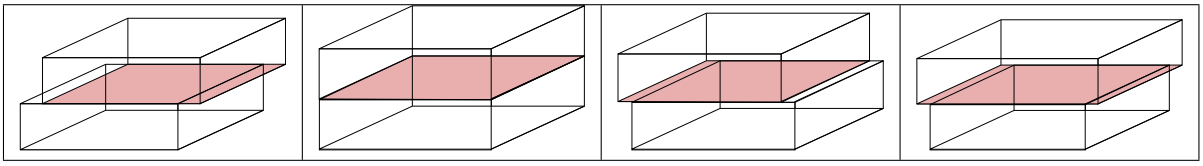
Show that it is impossible to make a dislocation loop all of whose segments are pure screw dislocations, but that it is possible with edge dislocations. For the case of the pure edge dislocation loop, describe the orientation of the extra half plane with respect to the dislocation loop.

Draw the compressive and tensile regions surrounding an edge dislocation.

Consider the dislocation loop shown below:



(a) Circle the drawing below that corresponds to the shape of the material after the dislocation has expanded and moved out outside the crystal.



(b) Indicate in the spaces below the locations (a, b, c, or d) where the dislocation has the following characteristics:

i It is a right handed screw dislocation:_____

ii It is a left handed screw dislocation:_____

iii It is an edge dislocation with the extra half plane above the plane of the loop:_____

iv It is an edge dislocation with the extra half plane below the plane of the loop:_____

(c) Add arrows to the illustration of the dislocation loop to show the orientation of the shear stress that will most efficiently cause the dislocation to loop to grow.

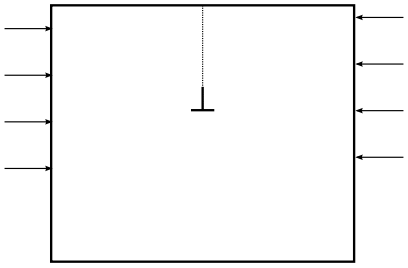
43.5 Dislocation Interactions

If edge dislocations with opposite signs of the Burger's vectors meet, does the energy of the crystal increase or decrease? Defend your answer.

A nanowire is grown such that it is free of dislocations. Why would the stress required to deform the nanowire be larger than a bulk material?

If an anisotropic alloy system has a nearly zero dislocation line tension, would you expect the precipitate spacing to have a large effect on the yield stress of the alloy? Explain your reasoning

Given an edge dislocation in a crystal, whose top two-thirds is under a compressive stress σ acting along the glide plane (see figure below):



(a) If diffusion occurs, which way will thee dislocation move? Explain why and tell where the atoms go that leave the dislocation.

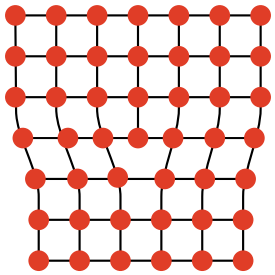
(b) Derive an equation relating the stress, σ to b and the force tending to make the dislocation move in the vertical plane.

(c) If the edge dislocation is replaced by a screw dislocation, which which way will the dislocation tend to move?

Construct a plot of the interaction energy vs. dislocation separation distance for two identical parallel edge dislocations that continue to lie one above the other as climb occurs. Justify your plot qualitatively by explaining how the strain energy changes with vertical separation.

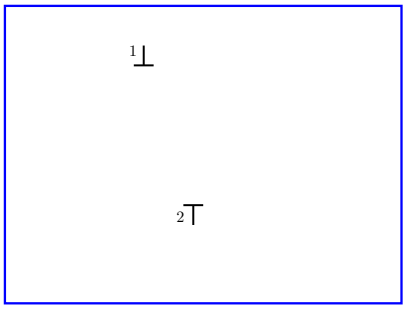
Repeat the previous problem for edge dislocations of opposite sign.

On the following sketch of a dislocation, indicate the direction that it must move in order for vacancies to be created.



Consider the two edge dislocations shown below. Suppose dislocation 1 remains fixed in place, but that dislocation 2 is able to move on its glide plane.

(a) Assume that the sense vector, $\vec{s}$, for each dislocation is defined so that $\vec{s}$ points into the page. Indicate the direction of $\vec{b}$ for each of the to dislocations.



(b) Indicate the glide plane for dislocation 2 with a dotted line.

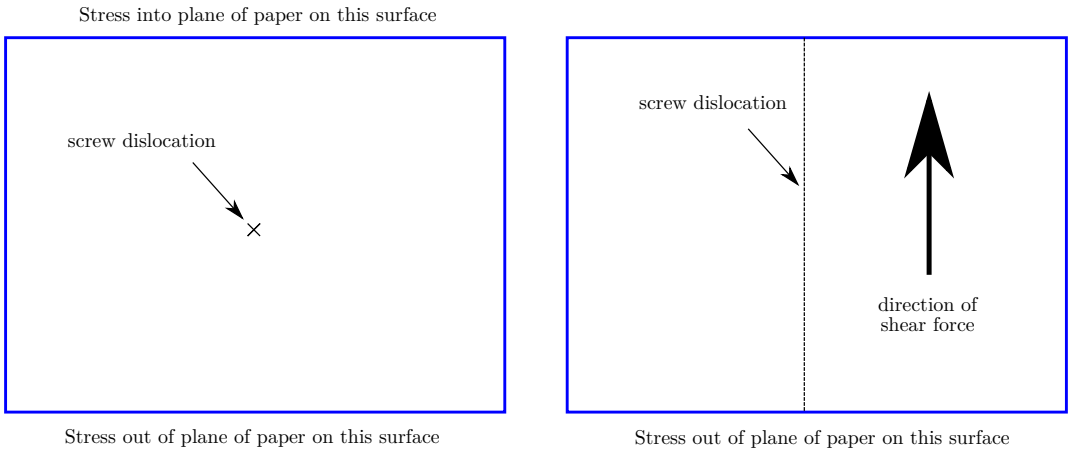
(c) Indicate with an X the location of dislocation 2 at the position within its glide plane that minimizes the total strain energy of the system.

(d) Now suppose that dislocation 1 is a fixed, left-handed screw dislocation and dislocation 2 is a mobile right-handed screw dislocation.

i Use a dotted line to indicate the plane on which you expect dislocation 2 to move in order to minimize the overall strain energy of the system.

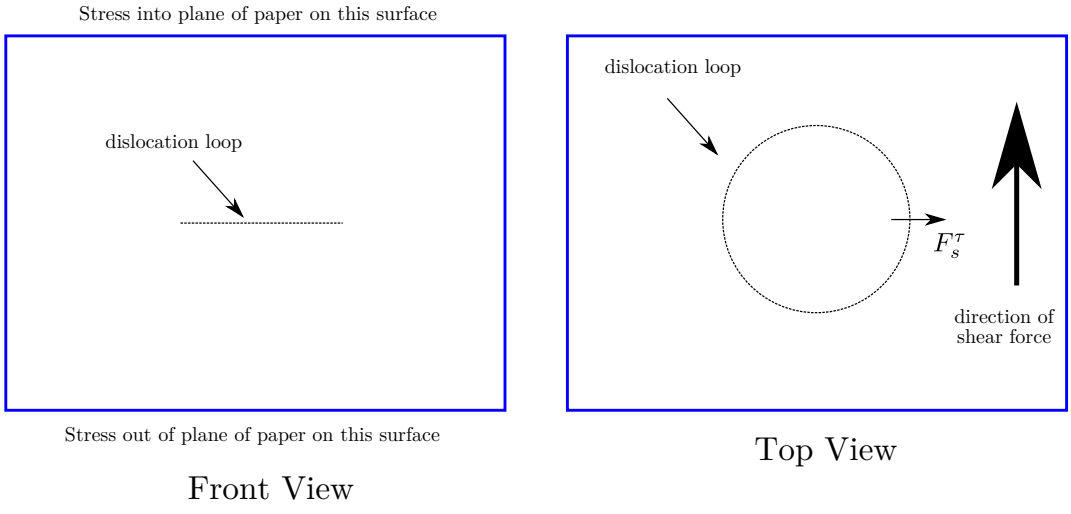
ii Plot the overall strain energy of the system as a function of the distance between the two screw dislocations.

Consider an isolated right-handed screw dislocation. Suppose a shear force is applied parallel to the dislocation line, as illustrated below.



(a) What is the direction of the force, F_s^τ, that is applied to the dislocation as a result of the applied stress.

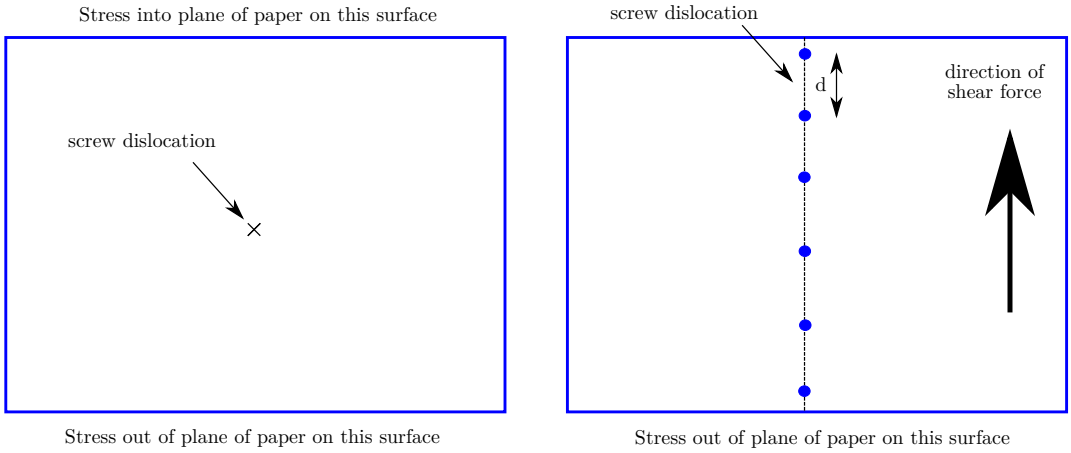
(b) Suppose the screw dislocation is replaced by a dislocation loop with the same Burgers vector as the dislocation from part a, as shown below. Use arrows to indicate the direction F_s^τ at different points along the dislocation loop. (The direction of F_s^τhas already been indicated at the right edge of the dislocation).



(c) Describe how the magnitude of F_s^τ changes (if at all) for different locations along the dislocation loop.

(d) What to you expect to happen to the dislocation loop if you remove the external applied stress (will the loop grow, shrink or stay the same size)?

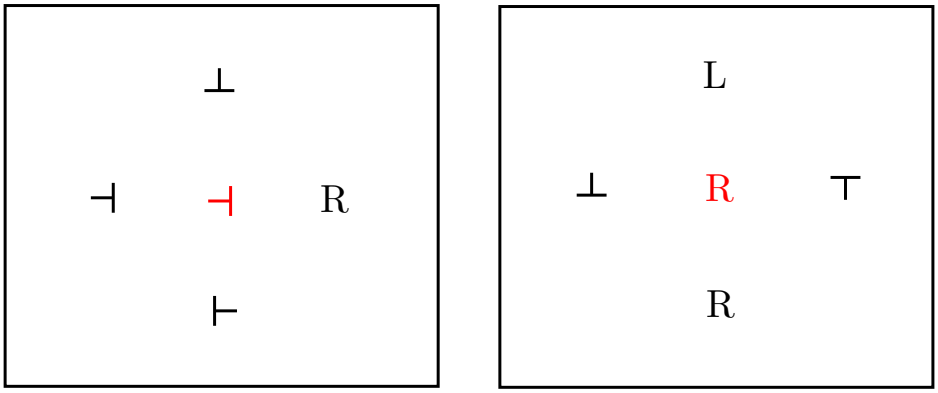
(e) Suppose the straight screw dislocation from is pinned by obstacles that are separated by a distance d, as illustrated in the following figure. Sketch the shape of the dislocation for an applied shear stress that is just large enough for dislocation to pass around the obstacles.



(f) What do you expect to happen to the critical resolved shear stress of the material if d is decreased by a factor of 2. (Will the critical resolved shear stress increase, decrease or stay the same).

Consider the two red dislocations at the center of the two diagrams shown below: (All of the dislocations are perpendicular to the plane of the paper.) We are interested in the effect that the central dislocation has on each case on the 4 surrounding black dislocations.

R = right-handed screw dislocation; L = left-handed screw dislocation



(a) For each of the 5 black edge dislocations, indicate the slip planes with a dashed line.

(b) Draw an arrow on each of the black edge dislocations, showing the direction of the force within its slip plane that is exerted by the red dislocation. If there is no force within the slip plane, circle the black dislocation instead.

(c) For the three black screw dislocations, draw an arrow on them to indicate the direction of the slip force exerted by the red dislocation. If the slip force is zero, circle the black screw dislocation instead.

43.6 Interfacial Thermodynamics

Consider the following:

(a) Is the molar latent heat positive or negative?

(b) Is the melting temperature, T , for a very small particle greater to or less than the equilibrium value of T_m for a bulk material?

(c) Must this always be the case?

(d) For metals, what is the typical value of r for which a change in melting temperature of 10K is observed. What about a change of 1K?

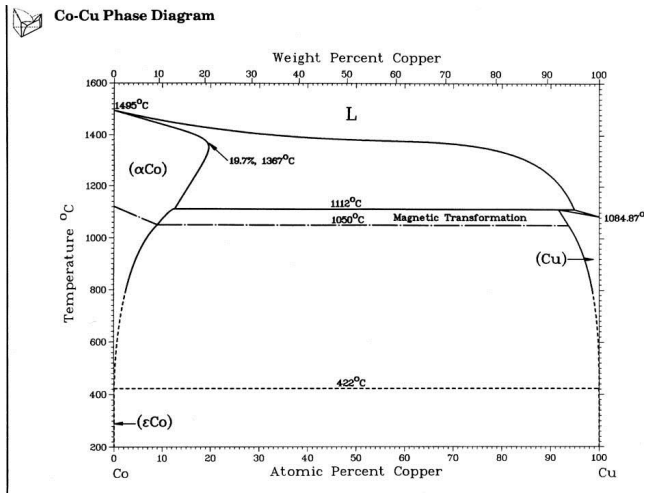
The molar enthalpy of a phase varies with temperature as

$$H_m(T) - H_m(T_0) + \int_{T_0}^{T} C_p(T)\, dT$$

where C_p is the molar heat capacity. Given this, at what temperature is the latent heat appearing in expression for the melting point reduction evaluated?

Consider the case of a pure liquid spherical droplet embedded in a pure solid. Create a graphical construction plotting the temperature dependence of the free energy of the solid and liquid phases for this case, and use it to determine if the melting point above or below the bulk melting temperature.

Consider the Co-Cu phase diagram shown below:



(a) Plot the equilibrium activity of Cobalt as a function of composition across the entire phase diagram at 900°C.

(b) From the phase diagram, estimate the solubility limit of Co in Cu at 900 °C. Suppose the interfacial free energy for the Cu/Co interface is 300 mJ/m^2. For what radius of a Co precipitate will this solubility limit be increased by 10%?

43.7 Surface and Interface Structure

Look up values for heats of sublimation for any of the materials in Table 8.1 that have close-packed crystal structures (FCC or HCP). Compare the estimated values of the surface free energy that you obtain from these heats of sublimation to the tabulated values in Table 8.1.

Determine the equilibrium shape of a crystal. This should be done using a computer and your favorite program or language (most likely MATLAB). The equation of a straight line in polar coordinates drawn from the origin of the polar coordinate system is $r\cos(\theta-\alpha)=d$, where (r,θ) locate the points on the line, d is the perpendicular distance from the origin to the line and α is the angle between the perpendicular to the line and the x-axis (see Figure 43.3).

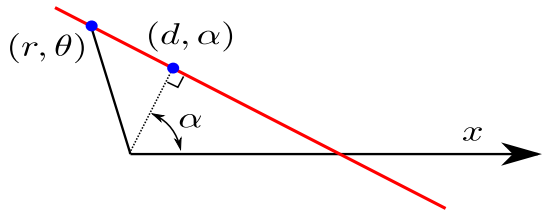


Figure 43.3: Representation of a line drawn a distance d from the origin.

(a) Determine the equilibrium shape of a crystal where the surface energy is given by $\gamma = 1\,\mathrm{J/m^2}$ (independent of α).

(b) Determine the equilibrium shape of a crystal where the surface energy is given by $\gamma = 1+0.05\cos(4\alpha)\,\mathrm{J/m^2}$ (α in radians). Are there any corners on the equilibrium shape?

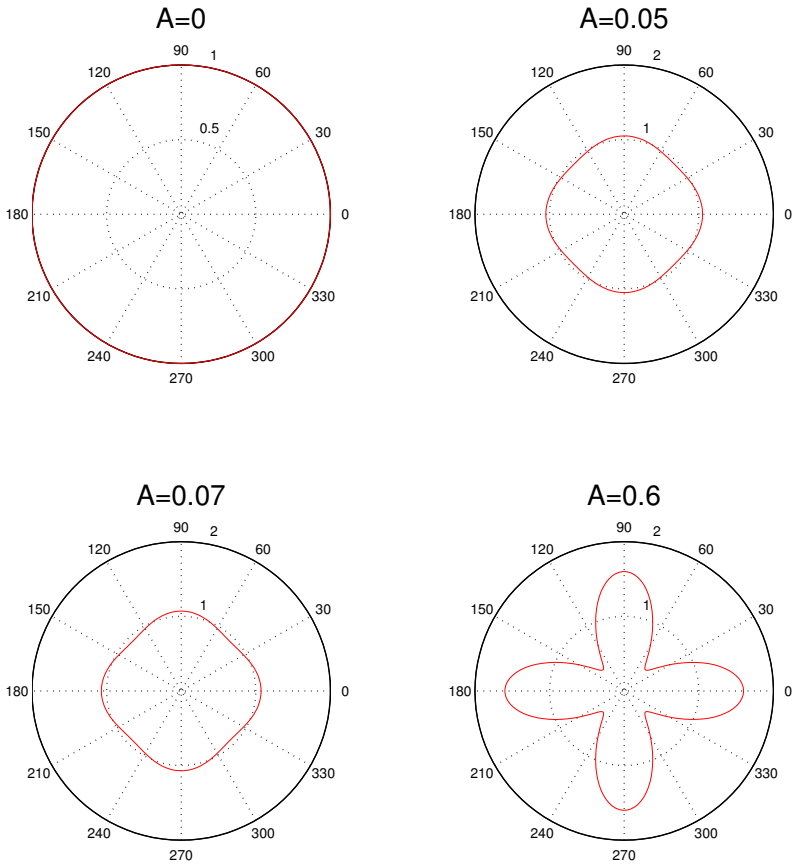
(c) Determine the equilibrium shape of a crystal where the surface energy is given by $\gamma = 1+0.07\cos(4\alpha)\,\mathrm{J/m^2}$. Are there any corners on the equilibrium shape?

(d) Determine the equilibrium shape of a crystal where the surface energy is given by $\gamma = 1+0.6\cos(4\alpha)\,\mathrm{J/m^2}$. Are there any corners on the equilibrium shape? How is the shape shown in (c) different from that in (d), and why (argue on the basis of the physics of the problem)?

As a headstart on this problem, here's a MATLAB script that generates polar plots of the γ as defined in the problem:

```
close all
A=[0,0.05,0.07,0.6];  % these are the 4 values of A defined in the problem
% define a function where the radius d is the surface energy and alpha
% is the angle
d=@(A,alpha) 1+A*cos(4*alpha);
figure
for k=1:4
    alpha=linspace(0,2*pi,200);
    subplot(2,2,k)  % this makes a 2 by 2 grid of plots
    polar(alpha,d(A(k),alpha),'r-');  % poloar is the command to make a polar plot
    title(['A=' num2str(A(k))],'fontsize',20)  % label each subplot
end
% adjust the print command as necessary to change the format, filename,
% etc.
print(gcf,'../figures/matlabwulffenergy.eps', '-depsc2') % save the eps file
```

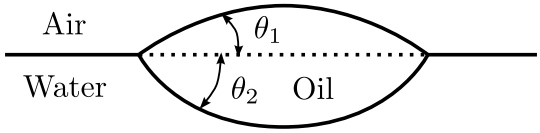
This generates the following polar plots for the four different functions that are given (with A defined so that $\gamma = 1 + A\cos(4\alpha)$).



Assume a simple cubic crystal structure with nearest neighbor interactions. Calculate the ratio of the surface energies for {110} and {100} surfaces.

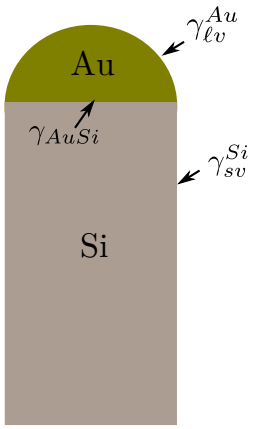
43.8 Wetting and Contact Angles

Consider the an oil droplet that forms on the surface of water, as shown schematically in the following Figure:



Determine θ_1 and θ_2 if the air/water interfacial free energy is 72 mJ/m^2, the air/oil interfacial free energy is 30 mJ/m^2 and the oil/water interfacial free energy is 50 mJ/m^2.

Suppose a, hemispherical liquid Au droplet with a radius of curvature of r is in contact with solid Si cylinder with the same radius as shown below. Derive a relationship between the three interfacial energies that must be valid in order for the equilibrium shape of the Au/Si interface to be flat, as drawn in the picture.



43.9 Grain Boundaries

The surface energy of the interface between nickel and its vapor is 1.580 J/m^2 at 1100K. The average dihedral angle measured for grain boundaries intersecting the free surface is 168°. Thoria dispersed nickel alloys are made by dispersing fine particles of ThO_2 in nickel powder and consolidating the aggregate. The particles are left at the grain boundaries in the nickel matrix. Prolonged heating at elevated temperatures gives the particles their equilibrium shape. The average dihedral angle measured inside the particle is 145°. Estimate the interfacial energy of the thoria-nickel interface. Assume the interfacial energies are isotropic.

Consider a gold line deposited on a silicon substrate. The film is 0.1 micron thick and all the grain boundaries are perpendicular to the gold-vapor interface. The grain boundaries run laterally completely across the line, giving a "bamboo" structure. The grain boundary energy of gold at 600K is 0.42 J/m^2 and the surface energy is 1.44 J/m^2. Assume all the interfacial energies are isotropic.

(a) Compute the dihedral angle where a grain boundary meets the external surface.

(b) Find the critical grain boundary spacing ℓ_c for which the equilibrium grain shape produces a hole in the film.

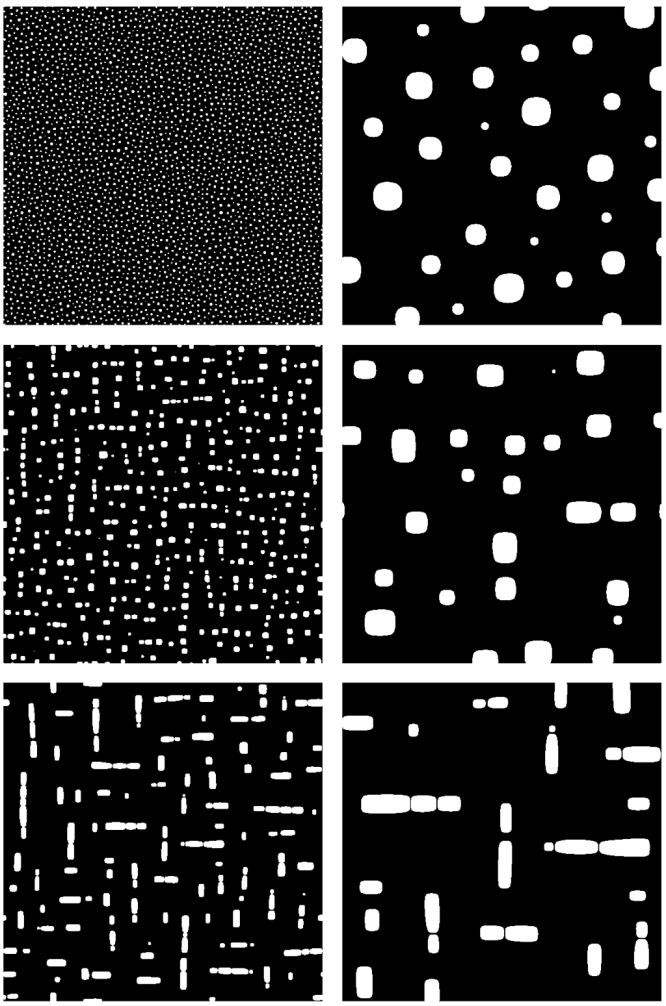
Why does the velocity of a grain boundary depend on temperature? Assume that the driving force for grain boundary motion is independent of temperature.

43.10 Transformation Kinetics

Does the time to 50% transformed increase or decrease with an increase in nucleation rate? Defend your answer without using any equations.

43.11 Interphase Interfaces

Consider the shapes of the particles in the simulations below of misfitting particles in an elastically anisotropic system. The left column is the entire system, whereas the right column is a magnification of a small region of the figure in the left column. These are snapshots taken as function of time while the particles are growing. Are these cuboidal shapes due to elastic stress, an anisotropic interfacial energy, or both?



Explain the structure and energies of coherent, semicoherent and incoherent interfaces, paying particular attention to the role of orientation relationships and misfit.

Explain why fully coherent precipitates tend to lose coherency as they grow.

The relationship between the the interfacial energy between α and β phases and the pressure difference across a curved interface is obtained from the following expression:

$$-P^{\alpha}\delta V^{\alpha} - P^{\beta}\delta V^{\beta} + \gamma_{\alpha\beta}\delta A^{\Sigma} = 0$$

(a) Use this expression to obtain the pressure difference between a cylinder of β phase with a radius r and a surrounding α phase.

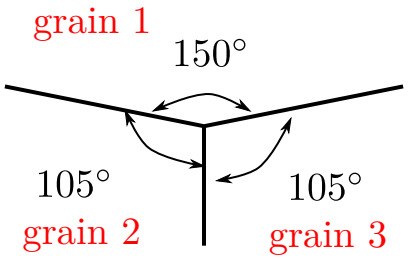
(b) Repeat the calculation for a cube where the length of each side is a. Assume that the surface energy of each of the cube faces is the same.

Why do very small precipitates tend to have coherent interfaces?

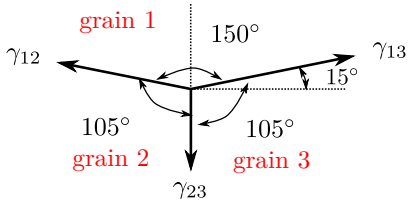
A thin film of Zn with an HCP crystal structure is deposited on a Ni FCC substrate with a {111} orientation. Which plane of the HCP crystal would you expect to contact the {111} Ni surface?

Given an example of an interface between two crystals that that displays a very large change in free energy with a change in the orientation of the interface.

Consider the following junction between three grains. Suppose that the grain boundary free energy between grains 1 and 2, and between 1 and 3, is 0.5 J/m^2. What is the grain boundary energy between grains 2 and 3?



Solution: The force balance is as shown below:

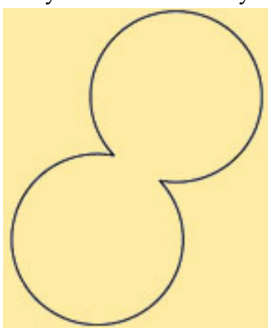


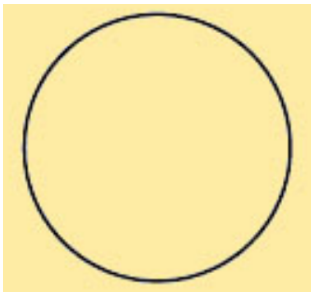
The force balance in the vertical direction in this case can be written in the following way:

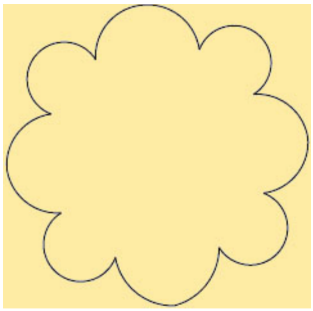
$$(\gamma_{12} + \gamma_{13}) \sin(15^\circ) = \gamma_{23}$$

with $\gamma_{12} = \gamma_{13} = 0.5\,\text{J/m}^2$ we get γ_{23} =0.26 J/m^2.

Consider a material with the orientational dependence of the surface energy shown in each of the 3 plots below. For each of these three materials, sketch the equilibrium shape that you would expect to obtain. On each drawing, indicate any interfaces that you expect to be coherent.







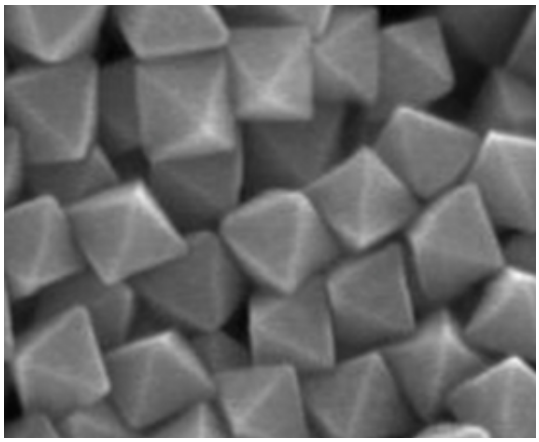
Consider an FCC metal (metal A) with a surface energy of 1 J/m^2. An HCP metal (metal B) with a surface energy of 0.7 J/m^2 is deposited onto the {111} surface of metal A. Assume that the atomic diameter of the HCP metal is 3% larger than the atomic diameter of the FCC metal, and that the chemical component of the interfacial energy between the two metals is 0.2 J/m^2.

(a) For B layers that are sufficiently thin, do you expect that a coherent interface will form between the A and B materials? Justify your answer.

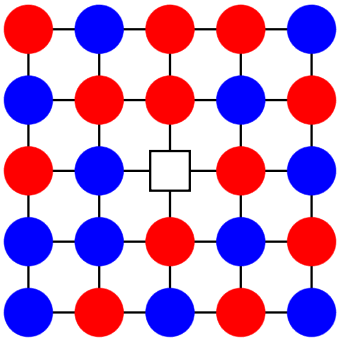
(b) How do you expect the interface between the A and B metals to change as the thickness of the B layer increases?

(c) Do you expect thick films to remain continuous, or will isolated drops of B be formed on the surface. Describe any assumptions that you make.

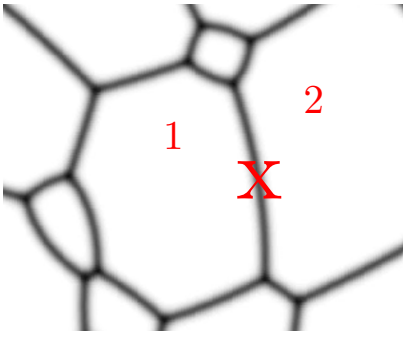
The octahedral particles of FCC gold shown below were created by controlling the growth rates of the different crystal facets. For these crystals, were the growth rates fastest in the ⟨100⟩ directions or in the ⟨111⟩ directions?



Consider the vacancy shown below, for a simulation of 'red' and 'blue' atoms that are undergoing phase separation. Is the vacancy more likely to move to the right or to the left? Justify your answer.



Consider the following image from the grain growth simulation:

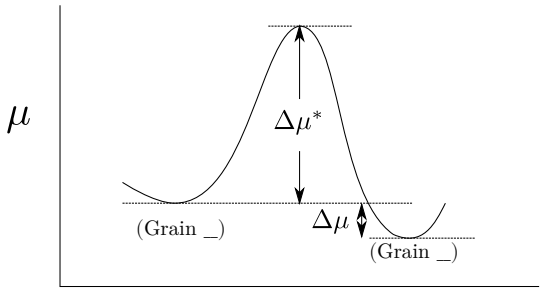


(a) The boundary marked with an 'X' separates grains 1 and 2. Do you expect this boundary to move toward grain 1 or grain 2 during the process of grain growth?

(b) Suppose that the interface marked above is the cross section through a grain boundary in aluminum, and that this section of the grain boundary has a spherical shape with a radius of curvature of 1μm. Assuming

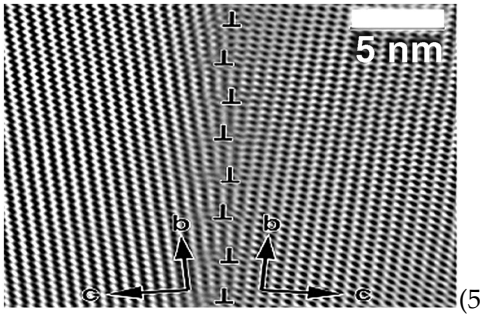
a grain boundary energy of 0.25 J/m^2, calculate the chemical potential difference, $\Delta\mu$ between Al atoms on the '1' and '2' sides of the grain boundary.

(c) On the schematic below, indicate which grain is grain 1 and which one is grain 2.

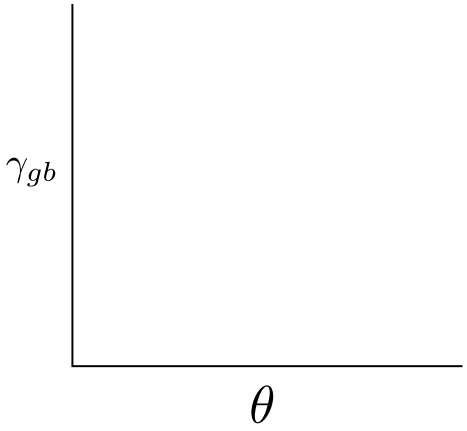


(d) Suppose $J_{1\to2}$ is the rate at which Al atoms hop from grain 1 to grain 2, and $J_{2\to1}$ is the rate at which

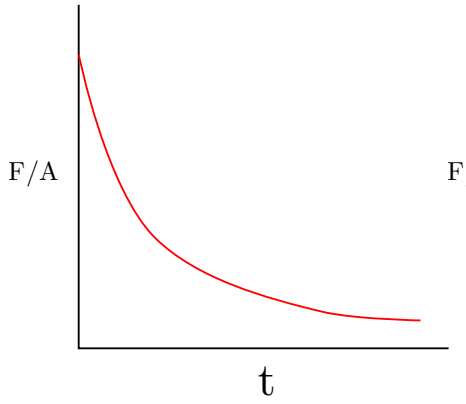
Consider the tilt boundary shown in the image to the left. On the axes on the right, sketch the relationship between the grain boundary free energy and the tilt angle that you expect to observe for values of theta between 0 and 10°.

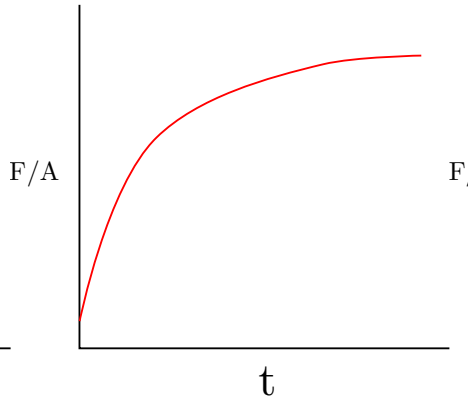


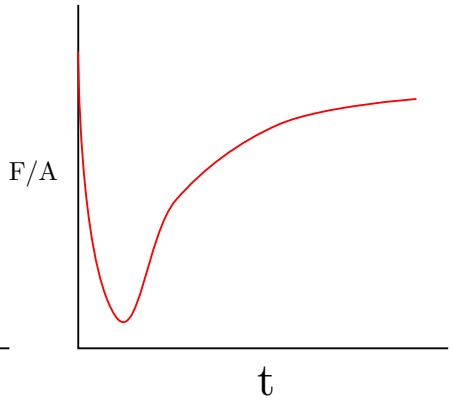
(5



Suppose you need to apply a coating to a surface, and you want the coating to spread as a smooth uniform film for all thicknesses. You have a choice of three different coatings, which have the thickness-dependent free energies shown below. Which material to you choose, and why?







44 316-1 Simulation Exercise: Monte Carlo Simulation of Decomposition in a Binary Alloy

44.1 Background

44.1.1 Scientific problem

We want to analyze the thermodynamic evolution of a A-B alloy by simulation. We assume that this system has the phase diagram presented in Figure 44.1. In this figure we see that for temperatures lower than T_C, the A-B alloy decomposes in two phases α and β with equilibrium concentrations X_B^α and X_B^β. The experiment that we want to model involves the following steps:

(1) We mix together the same number of moles of elements A and B to obtain a homogeneous alloy at some temperature above T_c.

(2) The temperature is reduced to T_0.

(3) The temperature is held fixed at T_0, and the system evolves to form two different phases, with compositions X_B^α and X_B^β.

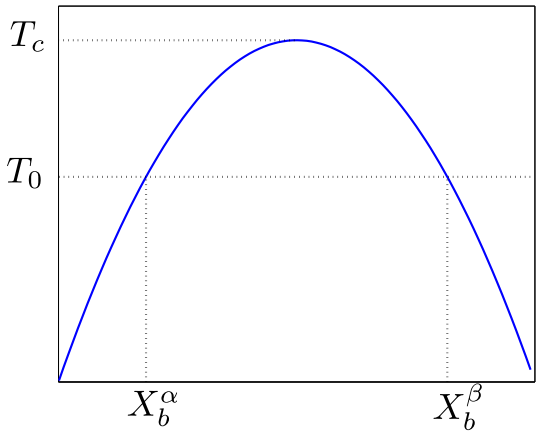


Figure 44.1: A-B alloy phase diagram

44.1.2 Atomistic Monte Carlo Model

In this section, we introduce the Atomistic Monte Carlo model that we will use to model the decomposition of the A-B alloy.

44.1.3 Atomistic model

In this model, we suppose that the two elements (A and B) have the same lattice structure. This lattice is represented by a matrix with periodic boundary conditions on its edges (see Figure 44.2). In 2 dimensions the left edge is connected to the right edge and the upper edge is connected to the lower edge. We reproduce the system evolution at the atomistic level: vacancies present in the lattice migrate from site to site by exchanging their position with their first nearest neighbors. The successive displacements of vacancies make the system evolve toward its equilibrium state.

44.1.4 Monte Carlo model

The thermodynamic evolution of the alloy is modeled with a Monte Carlo process. The principle of Monte Carlo simulations is to model the A-B alloy evolution in a statistic way. To understand this model we can consider individual jumps of a vacancy into one of the z nearest neighbor positions. Within a certain specified time step, Δt, these different possible jumps occur with a probability Γ_μ where μ is an index that indicates which direction

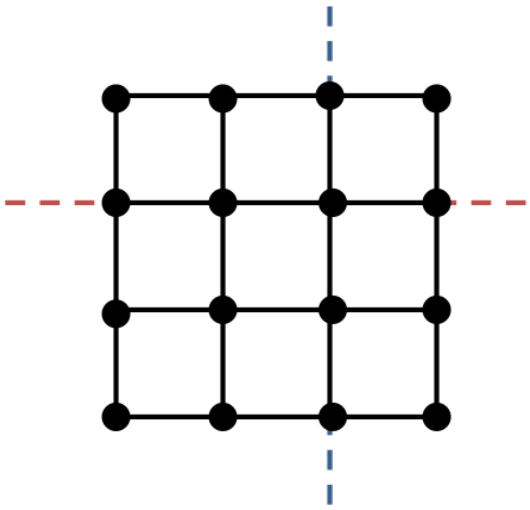


Figure 44.2: Lattice with periodic boundary conditions. The blue and red dashed lines represent bonds between sites induced by periodic boundary conditions.

the vacancy will move. In a simple cubic lattice, for example, $z = 6$, and the 6 values of μ correspond to jumps in the positive and negative x, y and z directions. The sum over all possible jump probabilities in the statistical time must sum to 1:

$$\sum_{\mu=1}^{z} \Gamma_\mu = 1 \tag{44.1}$$

To figure out which direction the vacancy moves, we draw a random number r_n between 0 and 1. The jump performed by the system during the time Δt is the k^{th} one such that the following condition holds:

$$\sum_{\mu=1}^{k-1} \Gamma_\mu < r_n \leq \sum_{\mu=1}^{k} \Gamma_\mu \tag{44.2}$$

Probabilities of transitions Γ_μ are related to the energetic barrier associated with vacancy motion, which we refer to as ΔE_μ. Because vacancy hopping is a thermally activated process, we can use an Arrhenius rate expression:

$$\Gamma_\mu = \Gamma_0 \exp\left(-\frac{\Delta E_\mu}{k_B T_0}\right) \tag{44.3}$$

where Γ_0 is a constant, k_B is Boltzmann's constant and T_0 is the temperature of the system.
The energy barrier is the difference between the maximum energy of the system during the jump (the position of the migrating atom at this maximum energy is called the saddle point) and the energy of the system before the jump.

$$\Delta E_\mu = E^{SP} - E^{ini} \tag{44.4}$$

Here the superscript SP refers to 'Saddle Point' and ini means 'initial', as shown in Figure 44.3.

44.1.5 Energetic model

To compute the energy barriers of the different possible jumps ΔE_μ, we have to use an energetic model. In Monte Carlo simulations, we usually use an Ising model or Broken bond model. In this energetic model, we assume that the total energy of the system is equal to the sum of interaction energies ε_{ij} between the different elements (atoms of type A and B and vacancies V) placed on the lattice sites.

$$E_\nu = \Sigma_{ij} \varepsilon_{ij} \tag{44.5}$$

With this energetic model, the migration barrier of an exchange between an element X and the vacancy V becomes:

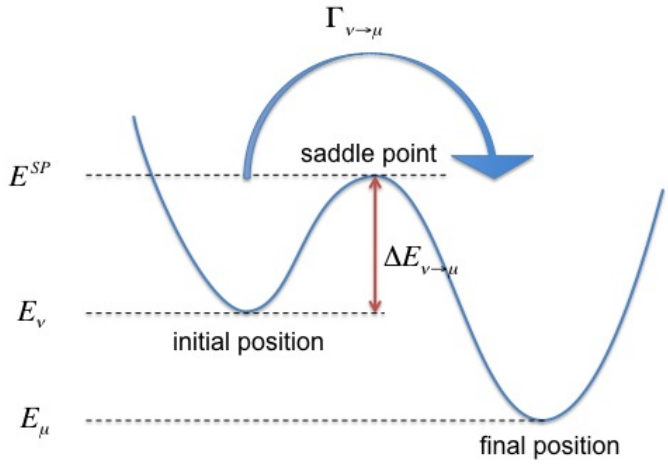


Figure 44.3: Schematic representation of the evolution of the system energy during a single atomic jump.

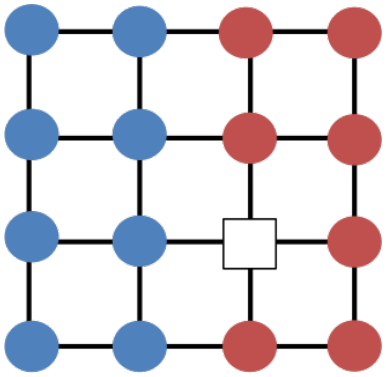


Figure 44.4: Example of configuration of atoms in the lattice. The red circles are A atoms and the blue circles are B atoms. The square is the vacancy

$$\Delta E_\nu = \sum_k \varepsilon^{SP}_{Xk} - \sum_i \varepsilon_{Xi} - \sum_j \varepsilon_{Vj} \tag{44.6}$$

where ε^{SP}_{Ak} are interaction energies between the atom migrating and its neighbors at the saddle point, ε_{Ai} are interaction energies between the atom migrating and its neighbors before the jump and ε_{Vj} are interaction energies between the vacancy and its neighbors before the jump. The indices i, j and k indicate the following neighbors:

Index	Meaning
i	nearest neighbors of the migrating atom before the jump
j	nearest neighbors of the vacancy before the jump
k	nearest neighbors of migrating atom at the saddle point

In theory, the range of interaction distances between elements are unlimited. In practice, we usually restrict these interactions to first and sometimes second nearest neighbors.
For example, the system presented in figure 44.4 has:

- 3 A-V interactions
- 1 B-V interaction
- 3 A-B interactions
- 9 A-A interactions
- 12 B-B interactions

Therefore, $E_v = 3\varepsilon_{AV} + 1\varepsilon_{BV} + 3\varepsilon_{AB} + 9\varepsilon_{AA} + 12\varepsilon_{BB}$. If we suppose that the vacancy exchange its position with the B atom on its left side, the configuration of the system at the saddle point is the one presented in figure 44.5.

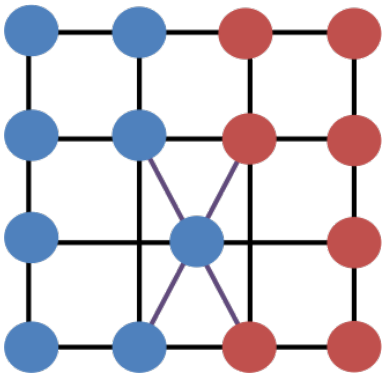


Figure 44.5: Configuration of the system at the saddle point if the vacancy echange its position with the B atom on its left

In this configuration, the system has:

- 2 B-B interactions at the saddle point
- 2 A-B ineractions at the saddle point
- 3 A-B interactions
- 9 A-A interactions
- 9 B-B interactions

so $E^{SP} = 2\varepsilon^{SP}_{BB} + 2\varepsilon^{SP}_{BA} + 3\varepsilon_{AB} + 9\varepsilon_{AA} + 9\varepsilon_{BB}$. The migration barrier of this jump is therefore:

$$\Delta E_v = 2\varepsilon^{SP}_{BB} + 2\varepsilon^{SP}_{BA} - 3\varepsilon_{AV} - 1\varepsilon_{BV} - 3\varepsilon_{BB}$$

44.1.6 Modeling of scientific problem

Here we assume that the two elements A and B have the same simple cubic lattice. We model the A-B alloy as a matrix in 2D with nx rows and ny columns and with periodic boundary conditions on its edges. To simplify the problem, we introduce only one vacancy in the lattice (so 1 vacancy for $nx \times ny$ sites), initially located in the middle of the matrix. As we only interest ourselves to the thermodynamic evolution of the system (and not to its kinetic evolution), we assume that the alloy evolves with normalized time steps of 1 until a maximum time t_{max}. At each time step, the vacancy exchanges its position with one of its neighbors.
To simplify the energetic model we suppose that the sum of interaction energies between the atom migrating and its neighbors at the saddle point $\sum_k \varepsilon^{SP}_{Xk}$ is a constant equal to $3\,eV$. In addition, we suppose that $\varepsilon_{AA} = \varepsilon_{BB} = \varepsilon_{AV} = \varepsilon_{BV} = 0\,eV$. The only interaction which can be different from zero is thus ε_{AB}.
The free enthalpy of the alloy is expressed by

$$\Delta G_{mix} = \Omega X_A X_B - T\Delta S_{mix} \tag{44.7}$$

with Ω the ordering energy of the alloy and ΔS_{mix} the configurational entropy of mixing of the alloy given by :

$$\Delta S_{mix} = -k_B \left[X_A \ln X_A + X_B \ln X_B\right] \tag{44.8}$$

For a symmetrical miscibility gap, the ordering energy is

$$\Omega = 2k_B T_C \tag{44.9}$$

where T_C is the critical temperature of the miscibility gap ($T_C = 1000\,K$ in this study). In broken bond models with only first nearest neighbors interactions we have:

$$\Omega = z\left(\epsilon_{AB} - \frac{1}{2}\left(\varepsilon_{AA} + \varepsilon_{BB}\right)\right) \tag{44.10}$$

where z is the number of first nearest neighbors for a given site.

44.1.7 Algorithmic scheme : Translation of problem in algorithm

In this section we translate the problem described previously in an algorithm scheme. As we are modeling an evolution according to time, our code will contain an initial state and an incremental loop on time which will start from the initial time (t_0) and finish at a final time (t_{max}). During the time loop (for example between time t_n and t_{n+1}), the code will repeat the same operations which will make the matrix go from the configuration at t_n to the one at t_{n+1}. In this code we suggest that the system evolves with the following steps in the time loop:

(1) Evolution of time from t_n and t_{n+1}

(2) Computation of jump frequencies of all possible jumps Γ_μ

(3) Drawing of a random number r_n and choice of a jump according to Eq. 44.2.

(4) Completion of chosen jump: exchange of position between vacancy and nearest neighbor chosen.

44.2 Exercise

Random walks
In this first work, we model the evolution of the system if the equilibrium configuration of the alloy is an homogenized state. As we only interest ourselves to the thermodynamic evolution of the system (and not to its kinetic evolution), we assume that the alloy evolves with normalized time steps of 1 until a maximum time t_{max}. At each time step, the vacancy exchanges its position with one of its neighbors. The vacancy can exchange its position with all its first nearest neighbors X (and only its first nearest neighbors). The difference is that in this section we suppose that all exchanges have the same jump frequency Γ_{XV}. This is called a" random walk".

44.2.1 Preliminary work

(1) Consider a vacancy located on the lattice site (xv, yv) as in Figure (44.6). In this figure, identify the first nearest neighbors of the vacancy by numbers and give the coordinates of these neighbors according to (xv, yv).

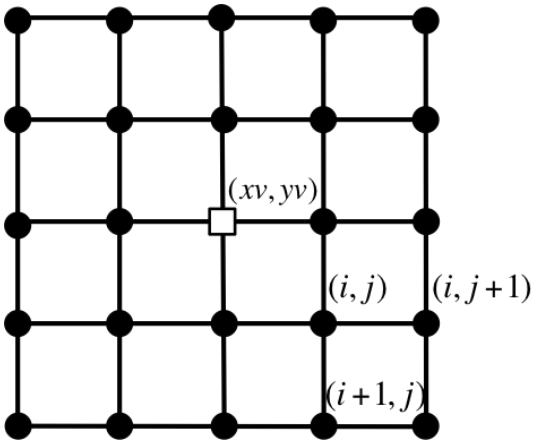


Figure 44.6: Identification of sites coordinates in the lattice and coordinates of a vacancy (represented by a square)

(2) Suppose that all exchanges of the vacancy with its first nearest neighbors have the same jump frequency. Using equation (44.1), give the probability of a given jump Γ_μ.

44.2.2 Simulation

(3) Create a folder for this MATLAB project. Open a new script in Matlab and save it in your folder as "part1.m".

(4) We first write the initial state of the system in the file part1.m. Save the matrix given in the file called 'matini.mat' available from the following link:

https://www.dropbox.com/s/8vhnam4lnburnit/matini.mat?dl=0

Load this matrix in part1.m as "matrix". Define nx and ny as the number of rows and column respectively of matrix. In this matrix, elements A are identified by a number 1 and elements B are identified by a number 2. Place a vacancy (identified by a 0) in the middle of the matrix:

- define in part1.m the coordinates (xv, yv) (where xv is the row and yv the column of the vacancy position) as the coordinates at the middle of the matrix.
- Place a 0 in the matrix at the appropriate coordinates (xv, yv) .

Initialize time t to 0.

(5) If the matrix has the configuration of figure 44.6, what does the matrix in Matlab look like (with the numbers)? (Include a printout of the matrix).

(6) Create a loop on time t where time evolves by steps of 1 as long as t remains lower than t_{max}. Place $t_{max} = 10$. We now have the part 1 in the algorithm (see section 44.1.7). Attach part1.m that includes all of the steps so far.

(7) We now have to create the next part of the algorithm: the computation of the jump frequency of all possible jumps. (Remark: in this random walk program, this part could be placed outside of the time loop since all jumps have the same frequency. However, we include it in the time loop to prepare the second part of the problem where we will have to compute the Γ_{XV} according to the environment). In the program, we call Gamma the vector such that Gamma(i) is the jump frequency of the exchange i. Use a "for" loop to compute the values of the different Gamma(i) components.

(8) We now have to choose a jump amount the different possibilities. For this, we suggest the MATLAB code shown below:

```
cumgam=cumsum(Gamma);
rn=rand;
njump=1;
for k=1:4,
    if cumgam(k)>=rn
       break
    end
    njump=k+1;
end
```

In this code, we draw a random number rn and translate the equation (44.2) in a Matlab program which uses the vector cumgam where $\text{cumgam(k)} = \sum_{i=1}^{k} \text{Gamma}(i)$. According to this algorithm, if $rn = 0.3$, what is the number of the jump chosen? What is 'njump' in this code?

(9) For the chosen jump, identify in your code by (xn, yn) the coordinates of the corresponding nearest neighbor according to (xv, yv). For this, we suggest you to define a matrix $(2 \times z)$ of the different possible evolutions (for example $\begin{pmatrix} +1 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$) and to write (xn, yn) according to (xv, yv) and the column of the matrix corresponding to the jump.

(10) We use periodic boundary conditions in this model (see part 44.1.3). For a site (x, y), verify that the following function enables to apply boundary conditions presented in figure (44.2)

$$x = \text{mod}(x - 1,\, nx) + 1$$
$$y = \text{mod}(y - 1,\, ny) + 1$$

For this, respond to the following questions: what is the value of x returned by this function if the x in input is between 1 and nx ? equal to 0? equal to $nx + 1$? Apply this function to xn and yn.

(11) Exchange types of elements corresponding to the vacancy an the neighbor migrating in the matrix.

(12) Update the vacancy coordinates to its new site.

(13) In this random walk model, what is the equilibrium state of the system? (Help: the fact that all the Gamma(i) are equal induces that the migration barriers for all possible jumps ΔE_μ are equal. From equation (44.6) it induces that all saddle point interactions ε_{Ak}^{SP} and ε_{Bk}^{SP} are equal, all atom-atom interations are equal and $\varepsilon_{AV} = \varepsilon_{BV}$. What is thus the value of the ordering energy Ω in equation (44.10)? And the value of T_C ? So at any temperature, what is the equilibrium state of the system?)

(14) Test: Replace the initial matrix by a matrix of same size with all A atoms on the half left side and all B elements on the half right side. Print an image of this initial matrix. Make the code run until $t_{max} = 10^6$. What do you observe? Print an image of the final matrix.

44.2.3 Introduction of alloy thermodynamic properties

We now have to introduce the alloy thermodynamic properties in the code. We thus have to compute the jump frequency of possible exchanges between the vacancy and its neighbors according to the alloy thermodynamic properties.

(15) We recall here that $\varepsilon_{AA} = \varepsilon_{BB} = 0\,eV$. Express ε_{AB} according to the ordering energy Ω and then to the critical temperature T_C. Give a numerical value of ε_{AB} in eV if $T_C = 1000\,K$.

(16) We analyze the migration barrier of an exchange between a vacancy V and one of its nearest neighbors X. We note NA the number of X first nearest neighbors of type A and NB the number of X first nearest neighbors of type B. How many first nearest neighbors does X have (we do not count the vacancy)? Express equation (44.6) according to NA, NB and $\varepsilon_{XA}$and ε_{XB} (look at the example given in section 44.1.5 for help). Using that $\sum_k \varepsilon_{Xk}^{SP} = 3\,eV$ and that $\varepsilon_{AA} = \varepsilon_{BB} = \varepsilon_{AV} = \varepsilon_{BV} = 0\,eV$, simplify the equation obtained if X is an element A. Same question if X is an element B. We observe from these calculations that, to compute the migration barrier of a jump, we need to know the type of the element of the exchange (so the type of X) and the type of all X first nearest neighbors (to compute NA and NB).

(17) For a given vacancy position, we want to compute the jump frequency of the jump i (so Gamma(i)). We note X the vacancy neighbor corresponding to this jump. We start by computing NA and NB (the number of X first nearest neighbors of type A and B). We note (xn, yn) the position of X and (xnk, ynk) the coordinates of X first nearest neighbor k (k goes from 1 to 3, the vacancy position is excluded from these nearest neighbors). We write

$$\begin{pmatrix} xnk \\ ynk \end{pmatrix} = \begin{pmatrix} xn \\ yn \end{pmatrix} + nveci(k)$$

where $nveci(k)$ is the column k of the 2×3 matrix of relative position of (xnk, ynk) compared to (xn, yn). Graph 44.7 gives the position of neighbors X compared to the vacancy.

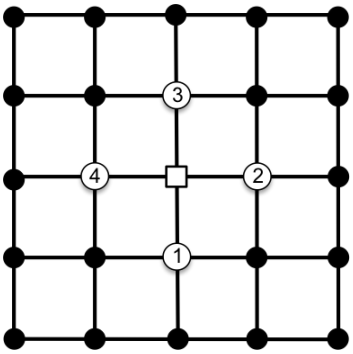


Figure 44.7: Position of X corresponding to the different possible jumps

For each of these jumps, associate the matrix nveci of the relative position of X first nearest neighbors.

$$(a)\, nveci = \begin{pmatrix} 0 & 0 & -1 \\ +1 & -1 & 0 \end{pmatrix} \quad (b)\, nveci = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$(c) nveci = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (d) nveci = \begin{pmatrix} 0 & 0 & +1 \\ +1 & -1 & 0 \end{pmatrix}$$

(18) Inside the loop to compute Gamma(i) coefficients write the following steps:

a define by (xn, yn) the vacancy neighbor corresponding to i (use the nvec matrix). Apply periodic conditions to (xn, yn).

b Initialize NA and NB to zero. Compute NA and NB of the exchange by analyzing the type of element on all (xnk, ynk) sites. To define the nveci matrix corresponding to the jump, you can distinguish the different cases with if-statements or you can use a structure with all the nveci matrices and load the one corresponding to the jump. Don't forget to apply boundary conditions to (xnk, ynk).

(19) Express the migration barrier of each jump depending on the type of the neighbor X (located on (xn, yn)) and NA and NB. Compute the jump frequency associated to this migration barrier (place the temperature to an arbitrary value-don't forget to define ε_{AB} in your code).

(20) Normalize the Gamma vector to 1 so that the sum of Gamma(i) is equal to 1.

(21) Analytic calculation: Suppose that for a given position, the vacancy can exchange it's position with either of 2 different A atoms. One on them is in a local configuration with NB=0 (the jump frequency of this exchange is noted $\Gamma_{NB=0}$) and the other one is in a local configuration with NB=3 (the jump frequency of this exchange is noted $\Gamma_{NB=3}$). Compute $\Gamma_{NB=3}/\Gamma_{NB=0}$ for T=100K and for T=2000K. Explain why these ratios are consistent with the alloy phase diagram.

(22) Place the temperature to 100K. Run the simulation until $t_{max} = 10^6$. What do you observe?

Part V
316-2: Microstructural Dynamics II

45 Nucleation

Much of this class is concerned with the appearance and growth of a new phase, referred to as β in Figure 45.1, from a matrix phase, α. In many cases the process can be divided into the following two phases:

(1) A nucleation phase, where individual very small (typically in the nanometer range) β precipitates are observed.

(2) A growth phase where the β precipitates grown in size.

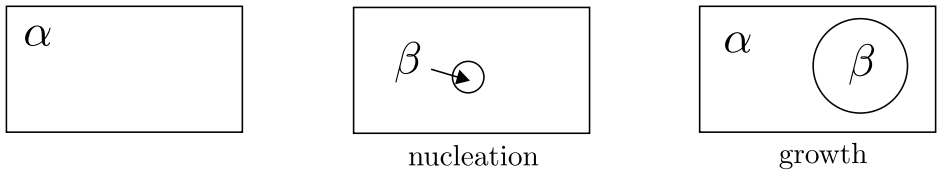


Figure 45.1: Schematic Representation of Nucleation and Growth.

The process is actually much more interesting than one might think by looking at the simple example shown in Figure 45.1. Consider, for example, the images of snowflakes shown in figure 45.2.

Figure 45.2: Snowflake images .

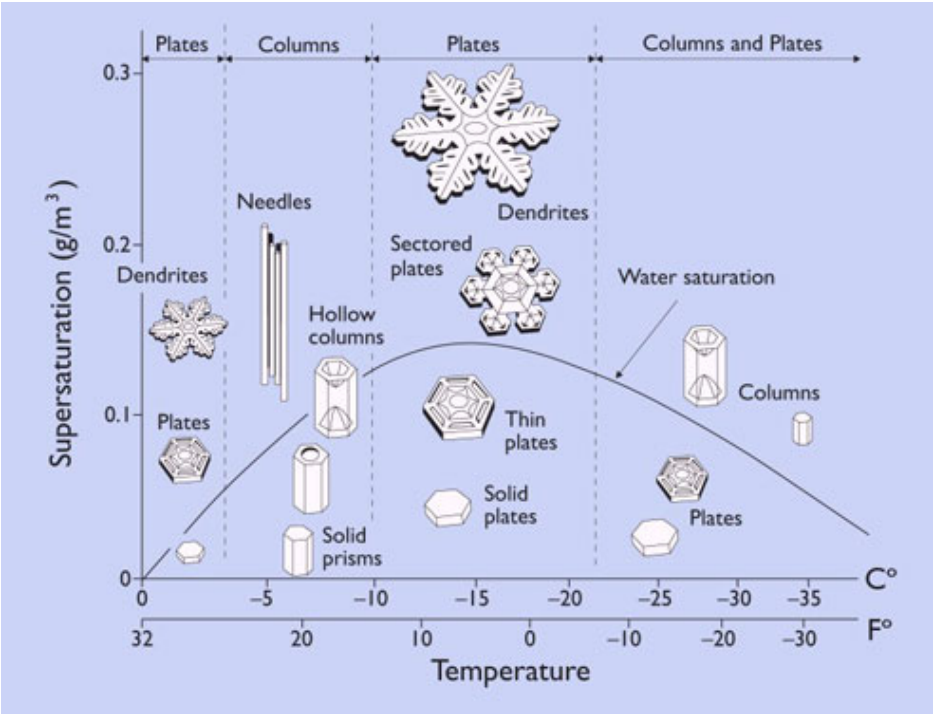


Figure 45.3: Morphology diagram for snowflakes .

If you are not motivated by snowflakes, plenty of modern technological examples exist that illustrate the concepts developed in this course. Silicon nanowires illustrated below are one excellent example.

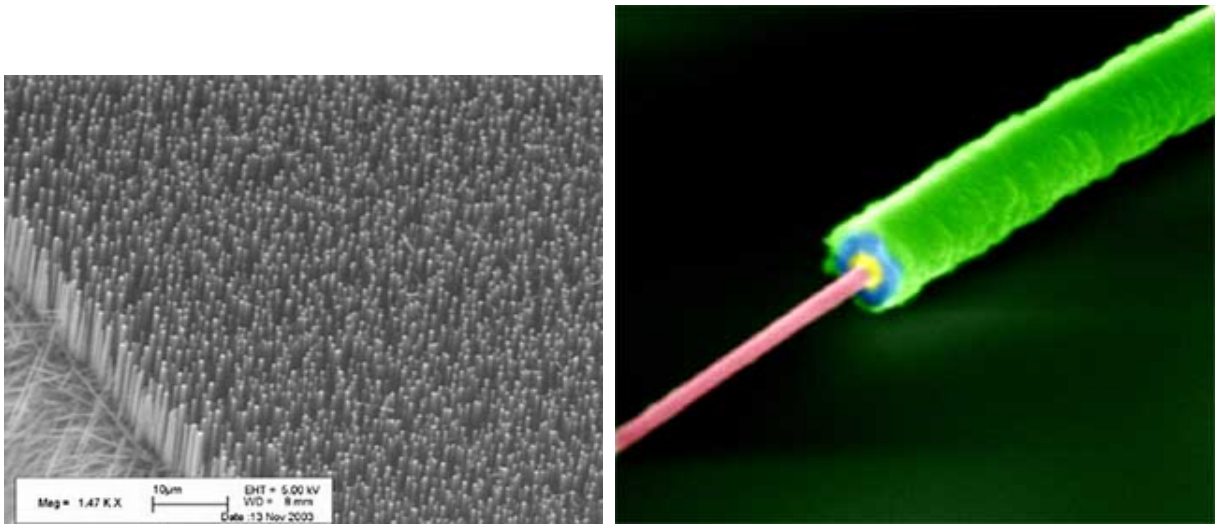


Figure 45.4: Two examples of silicon nanowires .

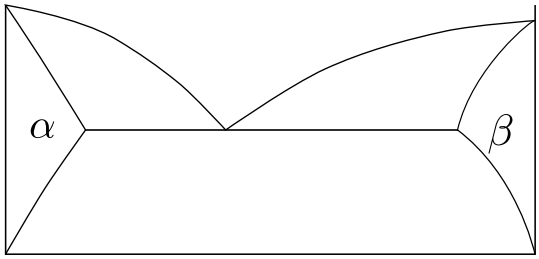


Figure 45.5: Generic eutectic phase diagram. A solid phase appears from the liquid as the temperature is cooled along the path indicated by the arrow.

The simplest example of these concepts is the freezing of pure water. The equilibrium phase at a given temperature is the one with the lowest molar free energy, G_m. A schematic representation of these free energies is shown in Figure 45.6. The free energies cross at 0 °C, with the liquid having a lower free energy at higher temperatures and the solid having a lower free energy at lower temperatures.

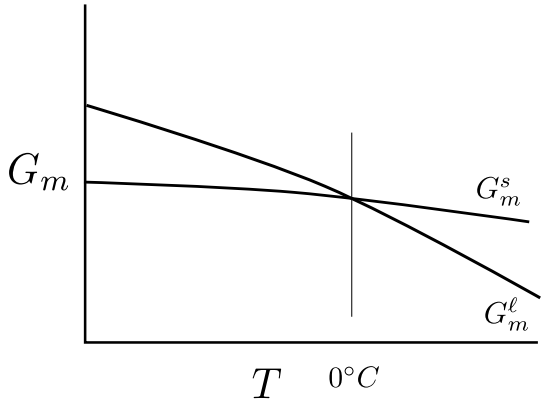


Figure 45.6: Schematic representation of the molar free energies of the solid and liquid phases of water.

Of course if we cool water below 0 °C, ice is not formed immediately - the solidification process takes time. Under appropriate conditions, it can be possible to maintain a liquid below its equilibrium melting point indefinitely, without ever seeing a solid. The ability to supercool a liquid is related to the fact that the interface between the solid and the liquid contributes a positive contribution to the overall free energy of the system.
Thermodynamics can be used to calculate the properties of equilibrium phases, including *unstable* equilibrium phases like the critical nucleus. The difference between a stable equilibrium (absolute minimum in free energy), a metastable equilibrium (local minimum in the free energy) and an unstable equilibrium (local maximum in the free energy) is illustrated in Figure 45.7.

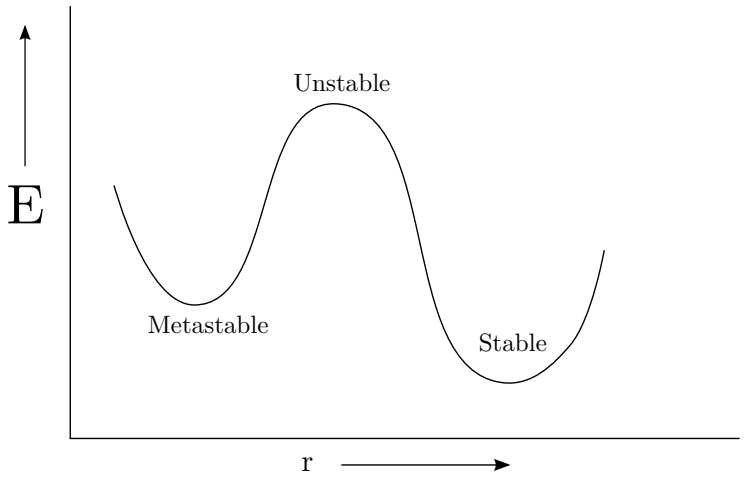


Figure 45.7: Energy as a function of nucleus size, illustrating stable, metastable and unstable states.

We need to compare the appropriate thermodynamic potentials for the cases where a critical nucleus is present and where it is absent. As illustrated schematically in Figure 45.8, we assume that the critical nucleus is spherical, with a radius of R^*.

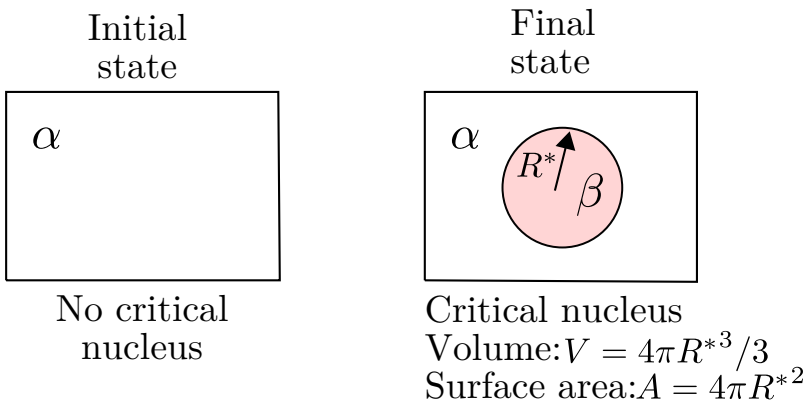


Figure 45.8: Initial and final states for the calculation of the work to form a critical nucleus.

Because of Laplace pressure, the pressure inside the critical nucleus is not equal to the pressure outside the critical nucleus. The correct thermodynamic potential to use is called the omega potential. The omega potential is relevant when a region of fixed volume is transformed from one phase to the other, in conditions where atoms are freely exchanged between the two phases. Here we give a brief review of thermodynamic potentials and introduce this new potential.

45.1 Thermodynamics Review

We begin by reminding ourselves of the differential forms of the internal energy (U), Helmholtz free energy (A) and Gibbs free energy (G). We begin with a closed system, where the number of particles is fixed, so chemical potentials are not relevant. We begin with expression for the differential of the internal energy, U:

$$dU = TdS - PdV \tag{45.1}$$

The natural variables for the internal energy are the volume and the entropy, since U remains unchanged ($dU = 0$) when these variables are constant ($dS = dV = 0$). Entropy is not a very useful natural variable, so we define the Helmholtz free energy, A, which has volume and temperature as the natural variables:

$$A = U–TS \tag{45.2}$$

$$dA = –PdV–SdT \tag{45.3}$$

The Gibbs free energy, G, has pressure and temperature as the natural variables:

$$G = A + PV \tag{45.4}$$

$$dG = VdP–SdT \tag{45.5}$$

For open systems where molecules or atoms are freely exchanges between different phases, we need to introduce the chemical potentials. The differential Helmholtz free energy becomes:

$$dA = -PdV - SdT + \sum \mu_i dn_i \tag{45.6}$$

which is zero at fixed V, T and n_i. Note that the index i in Eq. 45.6 includes all of the different components in the system. For example, for a binary system considering of two metallic elements, $i = 1$ corresponds to one of the elements and $i = 2$ corresponds to the other element. In an open system the chemical potentials are the relevant variables. We can define a new potential, Ω (the omega potential, also referred to as the Landau potential) in the following way:

$$\Omega = A - \sum u_i n_i \tag{45.7}$$

In differential form we have:

$$d\Omega = -PdV - SdT - \sum n_i d\mu_i \tag{45.8}$$

which is zero at equilibrium at fixed T, V, μ_i. A more useful expression for Ω is obtained by realizing that $G = \sum \mu_i n_i$, so we obtain:

$$\Omega = A–G = -PV \tag{45.9}$$

45.2 Thermodynamics of Pure Elements

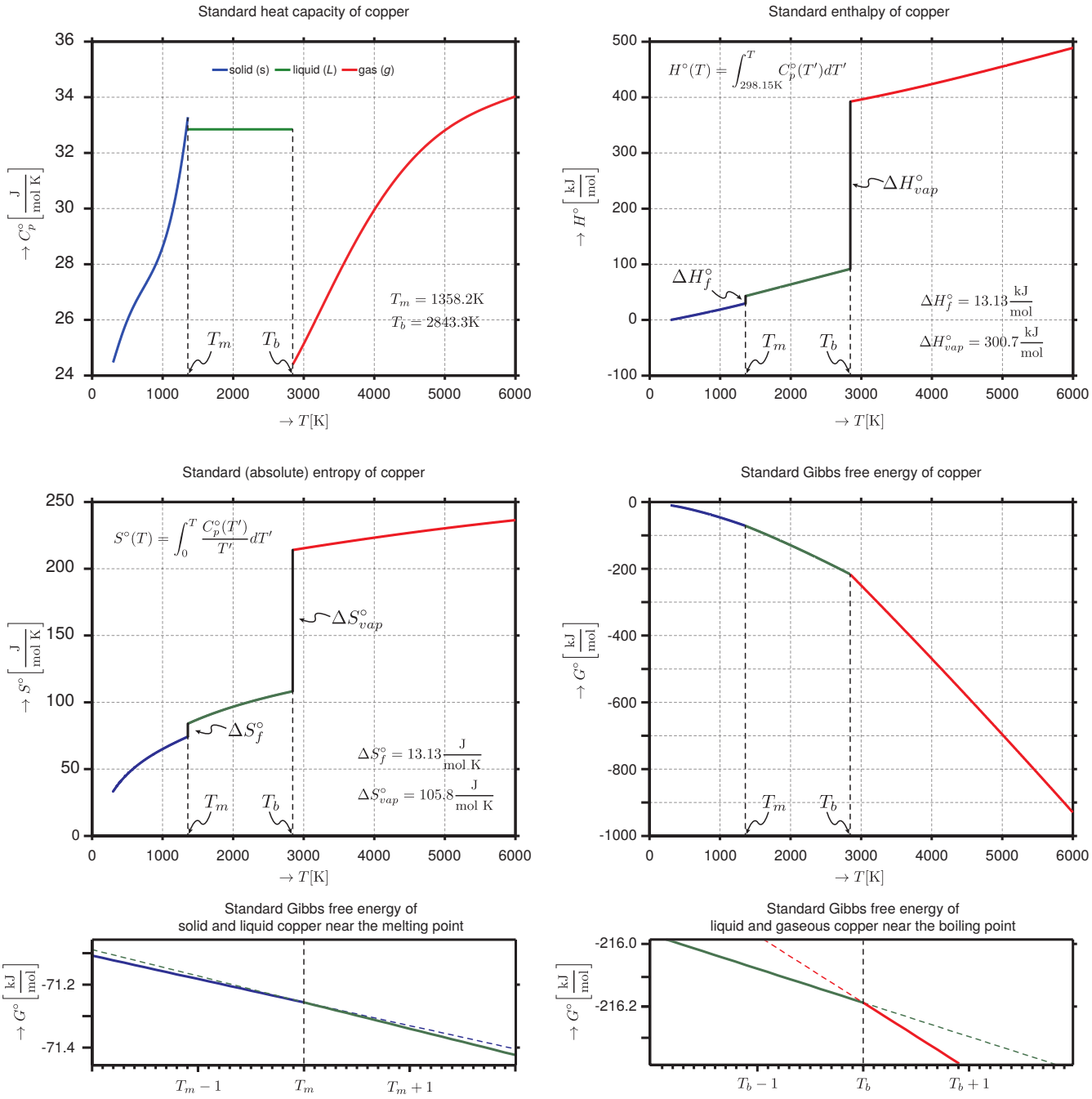


Figure 45.9: Thermodynamic data for copper. The curves were generated by the MATLAB script listed in Section 59.

45.3 The Critical Nucleus

We'll start by considering the transition of a pure one-component system between two phases, which we refer to as α (the phase present before the transition) and β (the phase present after the transition). The most important transition is solidification, in which case α is a liquid phase and β is a solid phase, but we'll keep things more generic for now so that the same formalism can be used to describe other phase transitions. The β phase is preferred because it is has lower Gibbs free energy at the specified temperature and pressure than the parent α phase. However, because of the interfacial free energy between the α and β phases, the relevant thermodynamic potential of the system is does decrease monotonically as the system transforms from α to β. Instead, this potential is maximized when the size of the transformed region of β corresponds to a critical nucleus, shown schematically in Figure 45.8.

The thermodynamic work to form a critical nucleus, W_R^*, is the difference in omega potentials between the final state (the existence of a critical nucleus of β) and the initial state (pure parent phase, α):
In addition to the PV term in the omega potential, we also need to include the interfacial term, given by the product of the interfacial free energy, γ, (an energy per unit area) and the interfacial area, A. Here we make the simplifying assumption that the critical nucleus is a sphere of radius R^*. Let's start with a review of the derivation of the Laplace equation for the pressure difference across a spherical interface. We start be equating the work done by inflating the sphere to the free energy increase associated with energy stored by the interfacial tension:

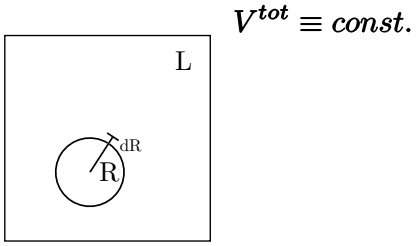


Figure 45.10: Change in

$$\Delta P dV = \gamma_{\alpha\beta} dA \tag{45.10}$$

We use the expressions for the volume and surface area of a sphere to get the relationship between dV and dA:

$$V = \frac{4}{3}\pi R^3, \; A = 4\pi R^2 \tag{45.11}$$

We differentiate the expressions in 45.11 to obtain the following:

$$dV = 4\pi R^2 dR, \; dA = 8\pi R dR \tag{45.12}$$

Combination of Eqs. 45.10-45.12 gives:

$$\Delta P = \frac{2\gamma_{\alpha\beta}}{R} \tag{45.13}$$

This is the well-known expression for the Laplace pressure difference between the inside and outside of a sphere of radius R.
Here ΔP is the difference in pressure between the inside of the sphere and the outside of the sphere. We are interested in applying thermodynamic principles to the critical nucleus with a radius that we define as R^*. The pressure difference, ΔP, is equal to $P^{\beta^*} - P^{\alpha}$. We can now use the concept of the omega potential to express W_R^* in terms of either $P^{\beta^*} - P^{\alpha}$ or R^*. We begin by defining W_R^*as the difference between the Ω potentials between the initial and final states in Figure 45.8:

$$W_R^* = \Omega_f \,(\text{final state}) – \Omega_i \,(\text{initial state}) \tag{45.14}$$

The relevant volume here is the volume, $V^* = 4\pi \left(R^*\right)^3/3$ of the critical nucleus. In the final state the pressure is t pressure, P^{S*} inside the critical nucleus. We also need to account for the interfacial energy, γ between the solid and liquid:

$$\Omega_f = -P^{S*} V^* + \gamma A^* \tag{45.15}$$

Here A^* is the surface area of the critical nucleus, *i.e.* $A^* = 4\pi \left(R^*\right)^2$. In the initial state is simply pure liquid, so there is no interface and the pressure is the liquid pressure, P^{ℓ}:

$$\Omega_i = -P^{\ell} V^* \tag{45.16}$$

The Laplace pressure expression (Eq. 45.17) can be used to relate the pressure difference to the interfacial energy and the critical radius, R^*:

$$R^* = \frac{2\gamma}{P^{s*} - P^{\ell}} \tag{45.17}$$

Combination of Eqs. 45.29-45.17 enable us to obtain the following expression for the W_R^* in terms of the pressure difference or R^*:

$$W_R^* = \frac{16\pi\gamma}{3\left(P^{s*} - P^{\alpha}\right)^2} = \frac{4}{3}\pi\gamma_{\alpha\beta}R^{*2} \tag{45.18}$$

45.4 Solidification

This expression assumes that $R = R^*$, but we still don't know what R^* (or equivalently, $P^{s*} - P^{\ell}$) is. Recognizing that we can use thermodynamic relationships to calculate the properties of the critical nucleus, we begin by writing the appropriate equilibrium condition between the solid, critical nucleus and the surrounding liquid:

$$G_m^S(T, P^{S*}) = G_m^L(T, P^{\ell}) \tag{45.19}$$

Note that we are interested in the case where the temperature is not necessarily equal to the equilibrium melting temperature ($T \neq T_m$). Also, because to the Laplace pressure difference, $P^{S*} \neq P^{\ell}$). We assume that the differences between T and T_m and between P^{S*} and P^{ℓ} are small, so we can get away with retaining only the first derivative terms in Taylor series expansions for G_m^{S*} and G_m^{ℓ}:

$$\begin{cases} G_m^S(T,\, P^{S*}) = G_m^{S*}(T_m,\, P^{\ell}) + \left.\frac{\partial G_m^S}{\partial T}\right|_{P^{\ell},T_m}(T - T_m) + \frac{\partial G_m^S}{\partial P}|_{P_l,T_m}(P^{S*} - P^{\ell}) \\ G_m^L(T, P^{\ell}) = G_m^L(T_m,\, P^{\ell}) + \left.\frac{\partial G_m^L}{\partial T}\right|_{P^{\ell},\,T_m}(T - T_m) \end{cases} \tag{45.20}$$

Now we can use the following thermodynamic definitions:

$$\frac{\partial G_m^S}{\partial T} = -S_m^S,\ \frac{\partial G_m^L}{\partial T} = -S_m^L,\ \frac{\partial G_m^S}{\partial P} = V_m^S \tag{45.21}$$

Combination of Eqs. 45.20 and 45.21 gives the following:

$$\begin{cases} G_m^S(T, P^{S*}) = G_m^{S*} + (T_m, P^{\ell}) - S_m^S(T - T_m) + V_m^S(P^{S*} - P^{\ell}) \\ G_m^L(T, P^{\ell}) = G_m^L(T_m, P^{\ell}) - S_m^L(T - T_m) \end{cases} \tag{45.22}$$

For a planar interface ($R = \infty$): $P^{\ell} = P^{S*}$, and the liquid and solids are at equilibrium at $T = T_m$:

$$G_m^L(T_m, P^{\ell}) = G_m^S(T_m, P^{\ell}) \tag{45.23}$$

We define the differences, ΔH_m, ΔS_m, ΔT in the following way:

$$\Delta H_m \equiv H_m^L - H_m^S, \Delta S_m \equiv S_m^L - S_m^S,\ \Delta T \equiv T_m - T \tag{45.24}$$

Now we combine Eqs. 45.19, 45.22, 45.23, 45.24 to obtain the following for the pressure difference:

$$P^s - P^{\ell} = \frac{\Delta T \Delta S_m}{V_m^s} \tag{45.25}$$

The enthalpy of melting, ΔH_m, is a more intuitive quantity than the entropy of melting, and is more directly measured experimentally. At equilibrium for an interface between solid and liquid with a planar interface (so the pressure is the same in both phases), the free energy of the solid and liquids are equal to one another. This fact is generally used to write thermodynamic quantities in terms of ΔH_m instead of ΔS_m. We begin by recognizing that the free energy change between solid and liquid phases is zero at $T = T_m$:

$$\Delta G_m(T_m) = \Delta H_m - T_m \Delta S_m = 0 \tag{45.26}$$

We can use this equation to write the entropy of mixing in terms of the enthalpy of mixing and the equilibrium melting temperature:

$$\Delta S_m = \frac{\Delta H_m}{T_m} \tag{45.27}$$

Equation 45.25 for the pressure difference can therefore be rewritten in the following way:

$$P^{s*} - P^{\ell} = \frac{\Delta T \Delta H_m}{T_m V_m^s} \tag{45.28}$$

The following points are useful to keep in mind:

(1) V_m^S, S_m^S, V_m^ℓ are all evaluated at the equilibrium melting temperature, *i.e.*, $T = T_m$.

(2) ΔH_m and ΔS_m are both > 0 for melting (increasing temperature through the transition). Our sign convention for these quantities is that they are measured during heating.

(3) $S_m^\ell > S_m^S$, so $P^{S*} - P^\ell = 2\gamma/R^* > 0$ If $T < T_m$. This means that a nucleus is able to form if $T < T_m$.

We can now combine Eq. 45.29 with Eq. 45.28 to obtain an expression for W_R^* in terms of measurable thermodynamic quantities:

$$W_R^* = \frac{16\pi\gamma^3 (V_m^S)^2 T_m^2}{3(\Delta T \Delta H_m)^2} \tag{45.29}$$

We can also obtain a comparable expression for the radius of the critical nucleus:

$$R^* = \frac{2\gamma}{P^{S*} - P^\ell} = \frac{2\gamma V_m^S T_m}{\Delta T \Delta H_m} \tag{45.30}$$

Note the following points:

- As $T \to T_m$, $R* \to \infty$, $W_R^* \to \infty$, difficult to nucleate near $T_m (\gamma > 0)$
- As ΔT (undercooling) increases, R^*decreases, W_R^*decreases
- $W_R^* \propto \gamma^3$, $W_R^* \propto \frac{1}{\Delta T^2}$

Note that the basic assumption here is that the nucleus is incompressible. In other words, $V_m = \frac{\partial G_m}{\partial P}$ is taken as a constant that is assumed to be unaffected by $P^s - P^\ell$. If the nucleus is compressible, higher order terms in the Taylor expansion (Eq. 45.20) must be included. In this course we make the assumption that the nucleus is incompressible. An important example where this assumption obviously fails is for the nucleation of gas bubbles in a liquid (carbon dioxide bubbles forming in a carbonated beverage, for example).

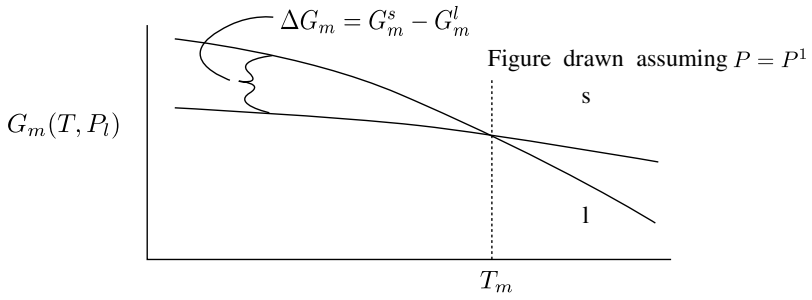


Figure 45.11: Temperature dependence of the molar free energies for the solid and liquid phases in the vicinity of the equilibrium melting temperature, T_m.

Below the melting point, the free energy change for solidification (excluding the surface energy) is negative, and above the equilibrium melting point, this free energy is positive:

$$\Delta G_m < 0 \, T < T_m \tag{45.31}$$

$$\Delta G_m > 0 \, T > T_m \tag{45.32}$$

A liquid cooled below $T_m$has a higher free energy, but work is still required to form a solid because of interfacial energy. The tradeoff between bulk free energy change (negative) and interfacial free energy (positive). It is most convenient to work in terms of a 'bulk' free energy per volume, ΔG_v, obtained by dividing ΔG_m by the molar volume, V_m:

$$\Delta G_v \equiv \Delta G_m / V_m \tag{45.33}$$

We can define an overall, size-dependent free energy for the nucleus with a term proportional to the volume, V^s of the solid nucleus, and a second term that is proportional to the surface area, A^s of the solid nucleus:

$$W_R = V^S \Delta G_v + \gamma A^S \tag{45.34}$$

Assume a sphere of radius R:

$$W_R = \frac{4}{3}\pi R^3 \Delta G_v + 4\pi R^2 \gamma \tag{45.35}$$

For $T < T_m$ ΔG_v is negative, and a plot of W_R vs. R looks like the plot shown in Figure 45.12. In this representation W_R^* is the maximum value of W_R, obtained where W_R is a maximum:

$$\left.\frac{dW_R}{dR}\right|_{R=R^*} = 0 \tag{45.36}$$

Solving Eqs. 45.35 and 45.36 gives the following for R^*:

$$R^* = \frac{2\gamma}{\Delta G_v}$$

which when substituted back into Eq. 45.35 gives:

$$W_R^* = \frac{16\pi\gamma^3}{3\left(\Delta G_v\right)^2} \tag{45.37}$$

Comparing to the previous expressions, we see that ΔG_v is given by the following:

$$\Delta G_v = P^\ell - P^s = \frac{\Delta T \Delta H_m}{T_m V_m^s} \tag{45.38}$$

An important assumption in this whole treatment based on Eq. 45.34 is that V_m^s the molar volume of the solid precipitate is constant, for all value of the pressure difference $P^s - P^\ell$. This is a reasonable assumption for many solid-phase precipitates but this assumption fails miserably if you are interested in gas nucleation from a supersaturated solution of gas in a solid or liquid (carbon dioxide in soda, for example).

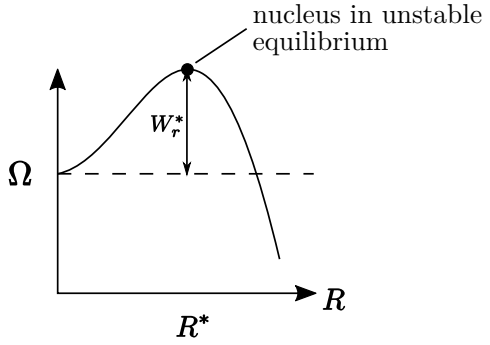


Figure 45.12: Free energy representation for a spherical, incompressible precipitate.

At this point it is useful to make a couple observations about A 3-dimensional stress has units of force/area or energy/volume:

$$\frac{force \cdot length}{area \cdot length} = \frac{energy}{volume} = stress\ (or\ pressure)$$

Similarly, an interfacial free energy (γ) has units of force/length or energy/area:

$$\frac{energy}{area} = \frac{force \cdot distance}{length \cdot length}$$

45.4.1 Nucleation in binary systems.

The critical nucleus of a phase β^* is in unstable equilibrium with the parent phase α. In equilibrium, the chemical potentials of components 1 and 2 are equal in both phases.

$$\begin{aligned} \mu_1^\alpha(P^\alpha, X_2^\alpha) &= \mu_1^\beta(P^{\beta^*}, X_2^{\beta^*}) \\ \mu_2^\alpha(P^\alpha, X_2^\alpha) &= \mu_2^\beta(P^{\beta^*}, X_2^{\beta^*}) \end{aligned} \tag{45.39}$$

If P^α and X_2^α are known, the above system of equations uniquely specifies P^β and $X_2^{\beta^*}$. We can therefore determine ΔP, W_R^*, and R^*.
Equation (45.39) can be interpreted graphically as the common tangent between $G_m^\alpha(P^\alpha, X_2^\alpha)$ and $G_m^\beta(P^{\beta^*}, X_2^{\beta^*})$, as shown in Figure xx.
Note that if T and N_i are held constant:

$$\left(\frac{\partial G_m}{\partial P}\right)_{T, n_i = const.} = V_m \tag{45.40}$$

In the limit where V_m is independent of pressure P and composition X_2, we can write

$$\frac{\Delta G_m}{\Delta P} = V_m \tag{45.41}$$

$$\Leftrightarrow G_m^\beta(P^{\beta^*}, X_2) - G_m^\beta(P^\alpha, X_2) = V_m(P^{\beta^*} - P^\alpha) \tag{45.42}$$

, meaning that the G_m^β is vertically displaced to higher value by a constant amount $V_m \Delta P$. Consequently, to determine $X_2^{\beta^*}$, one can construct a tangent to G_m^β that is parallel to the tangent to G_m^α at X_2^α. The point of tangency then has the abscissa $X_2^{\beta^*}$, as shown in as shown in Figurexx

45.5 Homogenous nucleation rate:

$\dot{N}_v \propto \exp\left[-\frac{W_R^*}{k_B T}\right]$ (nuclei per cm3 per second)
liquid contains C_0 atoms per unit volume ($C_0 \approx 10^{22}$atoms/cm3)
C* = number of clusters that have reached critical size:
$C^* \propto \exp -\frac{W_R^*}{k_B T}$
addition of a single atom converts critical nucleus to stable nucleus - this happens with a frequency
$v_0 (v_0 \approx 10^{11} s^{-1})$

$$N_v = v_0 C_0 \exp\left[-\frac{W_R^*}{k_B T}\right] = v_0 C_0 \exp\left[\frac{-16\Pi\gamma^3 (V_m^s)^2 T_m^2}{3 N_{av} k_B T (\Delta H_m \Delta T)^2}\right] \tag{45.43}$$

$v_0 C_0 \approx 10^{33} cm^{-3} s^{-1}$
The temperature dependence of the nucleation rate is enormous. As a general rule of thumb, we can assume that homogenous nucleation occurs when $W_R^* \approx 75 k_B T$ (see figure 45.13

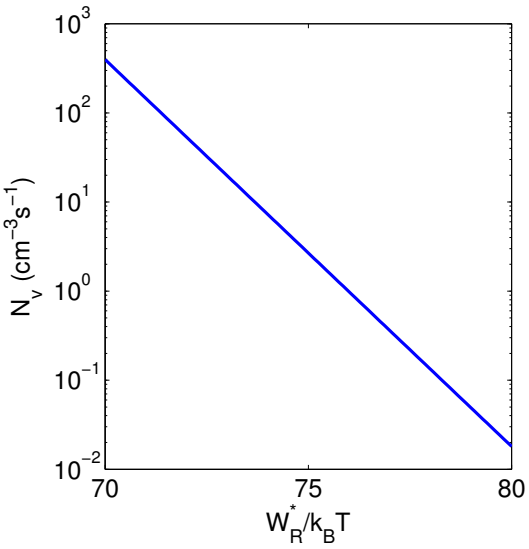


Figure 45.13: Homogeneous nucleation rate.

45.6 Heterogeneous Nucleation

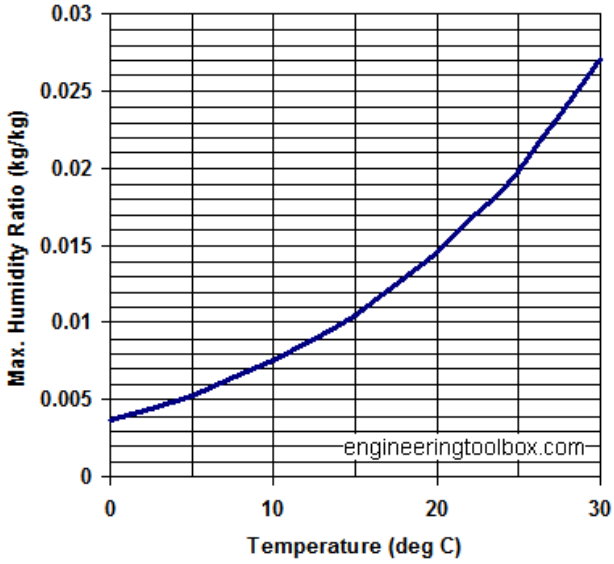


Figure 45.14: Mass of water/mass of air at saturation (http://www.engineeringtoolbox.com/water-vapor-saturation-pressure-air-d_689.html).

Figure 45.15: Fog in San Francisco (upi.com).

Figure 45.16: Condensation of water droplets from a supersaturated vapor phase onto a surface (http://www.noupe.com/photography/stunning-morning-dew-photography.html).

In most cases the amount of undercooling depends on the impurity level. This is because nucleation is often greatly accelerated when the solid phase nucleates at the interface with a different solid phase (including the container of the liquid). To understand why this is we need to consider the equilibrium shapes between three phases: the solid nucleus (δ in our example), the liquid phase (α) and a second solid phase, β, which can be dust particle in the sample, a wall of the sample container, etc. The way that the
Consider droplet of δin matrix of α, near boundary with another phase β:
δ= solid, α= liquid, β= surface (dust in previous example)
4 cases:

45.6.1 Nonwetting case

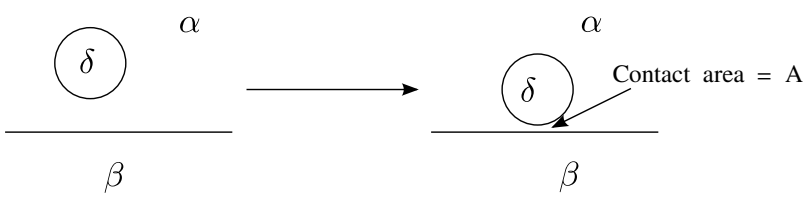


Figure 45.17: Schematic representation of the non-wetting case

$$\Delta\Omega = A(\gamma_{\beta\delta} - \gamma_{\alpha\delta} - \gamma_{\alpha\beta}) \tag{45.44}$$

$$\text{occurs when } \Omega > 0: \ \gamma_{\beta\delta} > \gamma_{\alpha\delta} + \gamma_{\alpha\beta}$$

Change in free energy:$\Delta\Omega = A(\gamma_{\beta\delta} - \gamma_{\alpha\delta} - \gamma_{\alpha\beta})$
Droplet is repelled (non-wetting, or non 'spreading') if $\gamma_{\beta\delta} > \gamma_{\alpha\delta} + \gamma_{\alpha\beta}(\Delta\Omega > 0)$

45.6.2 Complete wetting case

:δ spreads and forms a thin film along α/δboundary:

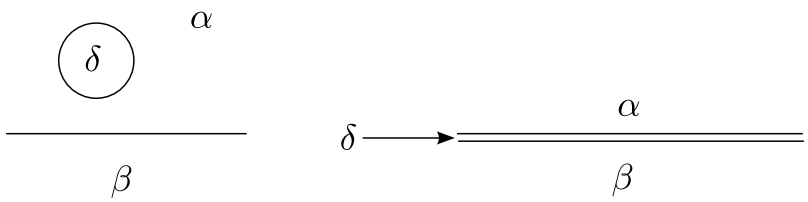


Figure 45.18: Schematic representation of the complete wetting case.

Note: α/β interfacial area much larger than α/δinterfacial area
$\Delta\Omega = A(\gamma_{\alpha\delta} + \gamma_{\beta\delta} - \gamma_{\alpha\beta})$ complete wetting for $\gamma_{\alpha\beta} > \gamma_{\alpha\delta} + \gamma_{\beta\delta}$

45.6.3 Partial wetting (intermediate case)

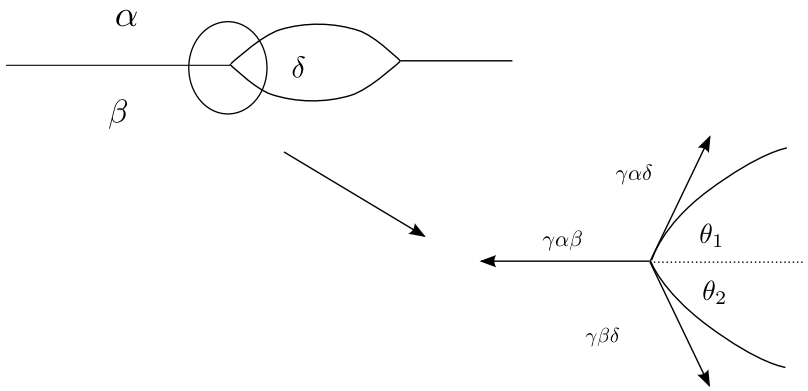


Figure 45.19: Schematic representation of the partial wetting case.

Horizontal force balance:

$$\gamma_{\alpha\beta} = \gamma_{\alpha\delta} cos\theta_1 + \gamma_{\beta\delta} cos\theta_2 \tag{45.45}$$

Vertical force balance:

$$\gamma_{\alpha\delta} sin\theta_1 = \gamma_{\beta\delta} sin\theta_2 \tag{45.46}$$

Partial wetting at a rigid surface:

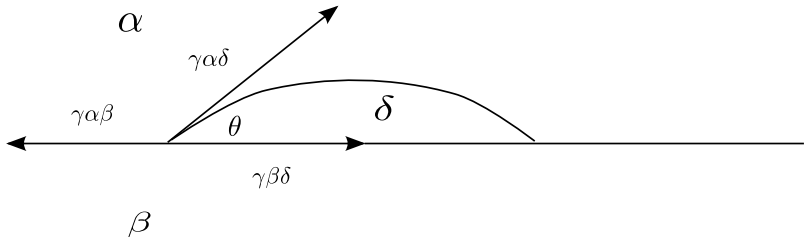


Figure 45.20: Partial wetting of a liquid droplet on a rigid surface.

Horizontal force balance:

$$\gamma_{\alpha\beta} = \gamma_{\alpha\delta} cos\theta + \gamma_{\beta\delta} \tag{45.47}$$

(Substrate not deformed by vertical force component)
Note: forces due to surface energies. Consider a soap film between
Nucleation of solid against an impurity: (δ= solid phase, s, α= liquid phase)

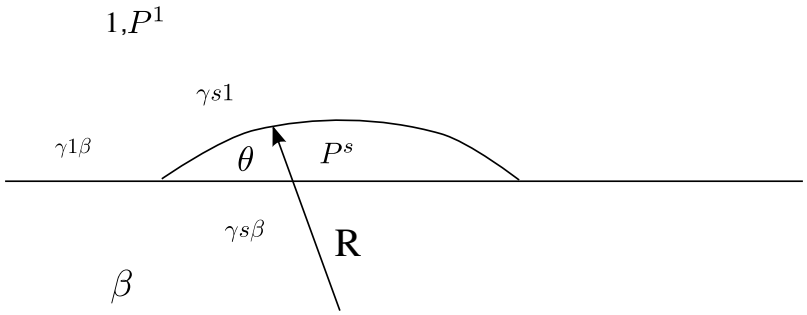


Figure 45.21: Partial wetting on a rigid surface - radius of curvature.

$$W_R^* = V^S(P^\ell - P^S) + \gamma_{sl} A_{sl} + (\gamma_{s\beta} - \gamma_{l\beta}) A^{S\beta} \tag{45.48}$$

$P^\ell - P^S$related to bulk thermodynamics (as before):

$$P^\ell - P^S = \frac{\Delta H_m (T - T_m)}{V_m^s} \tag{45.49}$$

Some geometry:

$$A_{sl} = 2\pi R^2 (1 - cos\theta) \tag{45.50}$$

$$A_{s\beta} = \pi R^2 sin^2\theta \tag{45.51}$$

$$V_s = \pi R^3 (2 + cos\theta)(1 - cos\theta)^2/3 \tag{45.52}$$

Solve for W_R^* :

$$W_R^* = \frac{16\pi\gamma^3}{3(P^s - P^\ell)^2} S(\theta) \;\; S(\theta) = (2 + cos\theta)(1 - cos\theta)^2/4 \tag{45.53}$$

This is exactly the same as same as the homogeneous nucleation case, but with $S(\theta)$ included. The functional form of $S(\theta)$is shown in Figure 45.22.

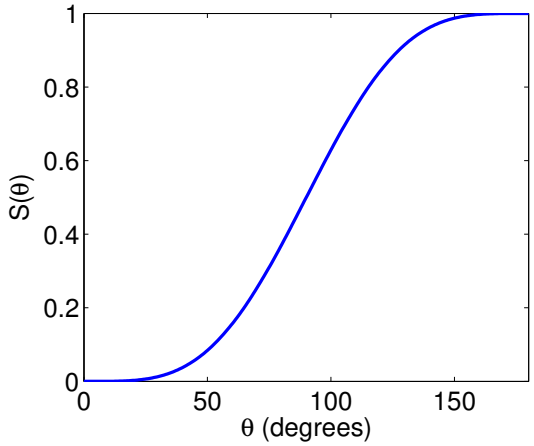


Figure 45.22: Functional form of $S(\theta)$.

What about size of critical nucleus: $R^* = \frac{P^S - P^\ell}{\gamma_{s\ell}}$ unaffected byβ phase
Some limiting cases:
Complete wetting: $\theta = 0$, $R \to \infty$, $S(\theta) \to 0$), no undercooling required
Non-wetting: $\theta = 180^\circ, A_{s\ell} = 4\pi R^2, A_{s\beta} = 0, V_s = 4\pi R^3/3,\ S(\theta) = 1$,surface has no effect
$\theta = 90^\circ:\ A_{s\ell} = 2\pi R^2,\ A_{s\beta} = \pi R^2,\ V_s = 2\pi R^3/3,\ S(\theta) = 1/2$
Hetergeneous nucleation rate:

$$\dot{N}_v = v_1 C_1 \exp\left[-\frac{W_R^*}{k_B T}\right] \tag{45.54}$$

$v_0 \approx 10^{11} s^{-1}$ (as before).
C_1= concentration of atoms in contact with a nucleation site:
$C_1 \approx 10^{15} cm^{-2}(A_n/V)$
A_n= total surface area of nucleation sites
Note: The exact value of A_n/V is not that important. It's generally good enough to say that A_n/V is approximately equal to any characteristic dimension of the droplet.

45.7 Nucleation in a Binary System

If the system is incompressible $(\Delta V = 0)$ or the pressure is very small, then:

$$\frac{\Delta G_m}{V_m} \equiv \Delta G_v = (P^\alpha - P^\beta) \tag{45.55}$$

So we can rewrite expressions for W_R^*, R^*

$$W_R^* = \frac{16\pi\gamma^3}{3\left(\Delta G_v\right)^2}, R^* = \frac{-2\gamma}{\Delta G_m/V_m} \tag{45.56}$$

Consider nucleation from an αphase that has a supersaturation of B

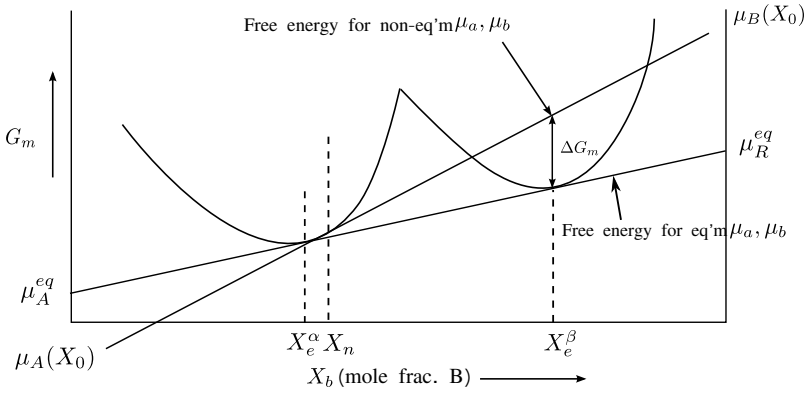


Figure 45.23: Effect of precipitate size on the free energy curves.

For $X_0 > X_e^\alpha, \mu_B > \mu_B^{eq}, \mu_A < \mu_A^{eq}$
Note: $G_m = \mu_a X_a + \mu_b X_b$

$$G_m = \mu_a(1 - X_b) + \mu_b X_b \tag{45.57}$$

$$G_m = \mu_a + X_b(\mu_b - \mu_a) \tag{45.58}$$

Points along straight lines have constant chemical potential
In mathematical terms:
$\frac{\partial G_m}{\partial X_b} = \mu_b - \mu_a, \mu_{ab} \equiv \mu_b - \mu_a$= exchange chemical potential - free energy required to replace an A atom with a B atom

$$G_m(X_{eq}^\beta) = G_m(X_{eq}^\alpha) + \mu_{ab}(X_{eq}^\beta - X_{eq}^\alpha) \tag{45.59}$$

$$\Delta G_m(X_0) = -\Delta\mu_{ab}(X_0)(X_{eq}^\beta - X_{eq}^\alpha) \tag{45.60}$$

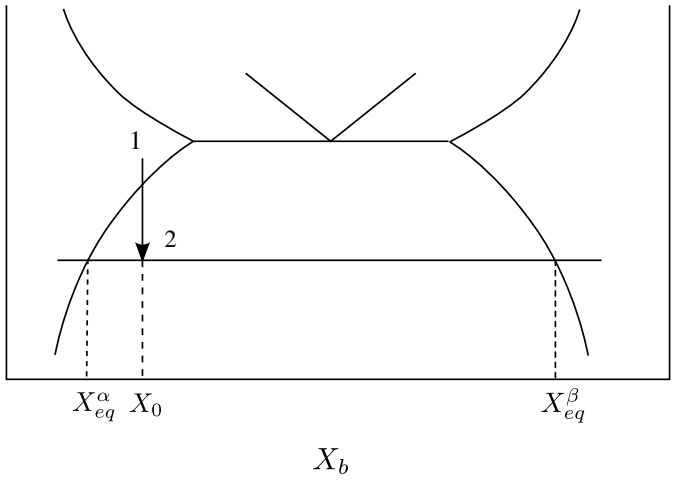


Figure 45.24: Precipitation in a binary system.

Quench from 1 to 2
Note:

(1) As $X_0 \to X_{eq}^\alpha, \Delta G_m \to 0$

(2) $\Delta\mu_{ab} > 0, \Delta G_m < 0$ for $X_0 > X_{eq}^\alpha$

(3) As $X_0 - X_{eq}^\alpha$increases, magnitude of ΔG_mincreases

$W_R^* \propto \frac{1}{\Delta G_m^2}$, $W_R^* \to \infty$as $X_0 \to X_{eq}^\alpha$
$R^* \propto -\frac{1}{\Delta G_m}$, $R^* \to \infty$as $X_0 \to X_{eq}^\alpha$
$\Delta G_m$is determined by curvature of free energy function in vicinity of X_{eq}^α:
$\Delta G_m(X_0) = -\Delta\mu_{ab}(X_0)(X_{eq}^\beta - X_{eq}^\alpha)$
Assume that $X_0$is close to X_{eq}^α: (small supersaturation of βin αphase):

$$\Delta\mu_{ab} \approx \left.\frac{\partial\mu_{ab}}{\partial X_b}\right|_{X_b=X_{eq}^\alpha} (X_0 - X_{eq}^\alpha)$$

but $\mu_{ab} = \frac{\partial G_m}{\partial X_b}$, so $\Delta\mu_{ab}(X_0) \approx \left.\frac{\partial^2 G_m}{\partial X_b}\right|_{X_b=X_{eq}^\alpha} (X_0 - X_{eq}^\alpha)$

$$\Delta G_m(X_0) \approx -\left.\frac{\partial^2 G_m}{\partial X_b^2}\right|_{X_b=X_{eq}^\alpha} (X_0 - X_{eq}^\alpha)(X_{eq}^\beta - X_{eq}^\alpha)$$

45.8 Effects of Strain Energy on Nucleation Barrier

Nucleation in the Solid State
need to account for strain energy
(ice example)

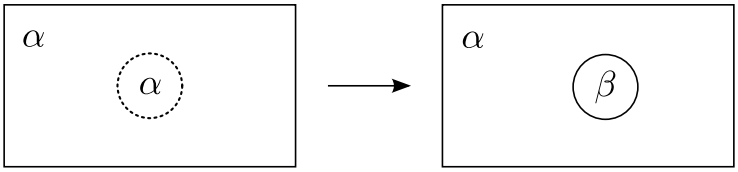


Figure 45.25: Replacement of a matrix phase (α) with a precipitate phase (β).

If volumes of $\alpha$and β are not the same, then both phases are stressed - energy penalty

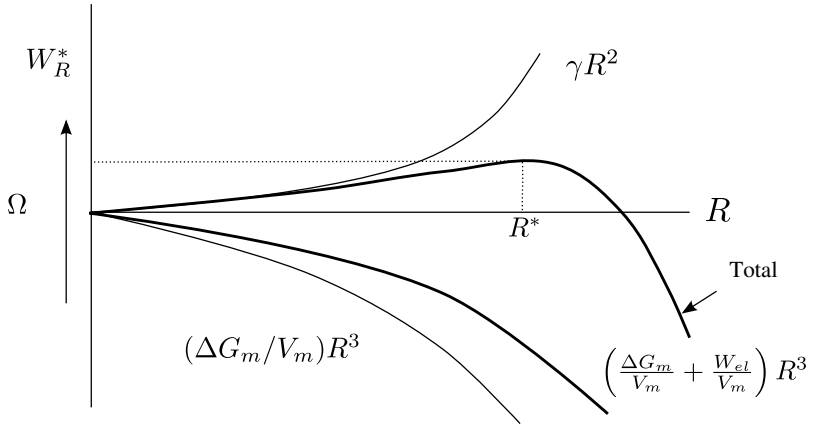


Figure 45.26: Graphical representation of the effect of strain energy.

$W_{e\ell}$ always positive and is proportional to volume of the beta phase - decreases magnitude of volume contribution - increases $W_R^*$and R^*.
For a spherical, coherent precipitate in an elastically isotropic medium:

$$W_{el} = V^\beta \left[\frac{18\varepsilon^2\mu^\alpha K^\beta}{3K^\beta + 4\mu^\alpha}\right] \tag{45.61}$$

K^β= bulk modulus of β phase: $K^\beta = V\frac{\partial P}{\partial V^\beta}$
μ^α= shear modulus of αphase: $\mu = \frac{shear\ stress}{shear\ strain}$

ε= misfit: $\varepsilon = \frac{1}{3}\left[\frac{V_m^\beta - V_m^\alpha}{V_m^\alpha}\right]$
2 limiting cases:

(1) compressible precipitate in a rigid matrix: $\mu^\alpha >> K^\beta$

$$\frac{W_{el}}{V^\beta} = \frac{9\varepsilon^2 K^\beta}{2} \tag{45.62}$$

(2) rigid precipitate in a deformable matrix: $K^\beta >> \mu^\alpha$

$$\frac{W_{el}}{V^\beta} = \frac{18\varepsilon^2 \mu^\alpha}{3} \tag{45.63}$$

For a coherent precipitate of a cubic material: $V_m^\alpha \propto (a^\alpha)^3$, $V_m^\beta \propto (a^\beta)^3$
$a^\alpha$and $a^\beta$are lattice parameters

$$\varepsilon = \left(\frac{a^\beta - a^\alpha}{a^\alpha}\right) \tag{45.64}$$

Comments:

(1) $W_{e\ell} > 0$ for any non-zero ε $\left(W_{e\ell} \propto \varepsilon^2\right)$

(2) Elastic energy scales with volume, not surface area →becomes more important for large values of R:
$\frac{V}{A} \propto \frac{R^3}{R^2} \propto R$
Surface energy is more important for very small precipitates
Coherent nuclei have smaller values of γ, larger values of $W_{e\ell}$than incoherent nuclei

(3) Stages in the growth of a nucleus:

a Critical nuclei are coherent (smallest R)
b Dislocations form at the interface to relieve interfacial mismathch - (larger R)
c Nuclei become incoherent (largest R)

Different processes increase γ and decrease $W_{e\ell}$ as precipitate grows
Total free energy of nucleus of β phase:

$$\Delta G = V^\beta \underbrace{\left\{\frac{\Delta G_m}{V_m}\right.}_{\text{Free energy of mixing term}} + \underbrace{\left.\frac{W_{el}}{V^\beta}\right\}}_{\text{Elastic term}} + A^\beta \gamma$$

Figure 45.27: Contributions to the overall free energy due to precipitation.

critical particle radius corresponds to maximum in ΔG.

$$\left.\frac{\partial \Delta G}{\partial R}\right|_{R=R^*} = 0 = 4\pi (R^*)^2 \left\{\frac{\Delta G_m}{V_m} + \frac{W_{el}}{V^\beta}\right\} + 8\pi R^* \gamma \tag{45.65}$$

$$R^* = \frac{-2\gamma}{\left\{\frac{\Delta G_m}{V_m} + \frac{W_{el}}{V^\beta}\right\}} \tag{45.66}$$

$$W_R^* = \Delta G(R^*) = \frac{16\pi\gamma^3}{3\left(\frac{\Delta G_m}{V_m} + \frac{W_{el}}{V^\beta}\right)^2} = \frac{16\pi\gamma^3}{3\left(\frac{\Delta G_m}{V_m} + \frac{18\varepsilon^2 \mu^\alpha K^\beta}{3K^\beta + 4\mu^\alpha}\right)^2} \tag{45.67}$$

$W_{e\ell}$increases $R^*$and W_R^*
Incoherent nucleus:
$W_{e\ell} = 0$, nucleation feasible when $\Delta G_m < 0$
this occurs when $X_0 > X_{eq}^{\alpha}$- incoherent solubility limit (solvus)
Coherent nucleus:
$W_{e\ell} > 0$, nucleation feasible when $\frac{\Delta G_m}{V_m} < -\frac{W_{e\ell}}{V^{\beta}}$
This occurs when $X_0 > X_{eq}^{\alpha} + \delta X$ - coherent solvus

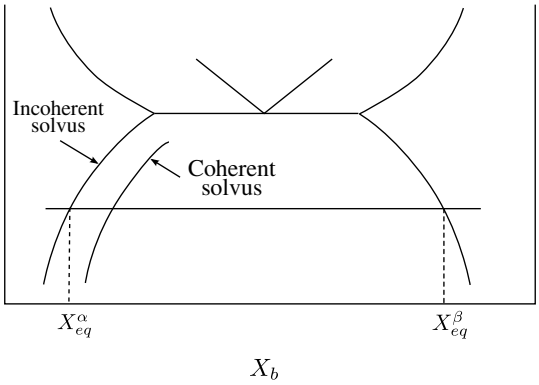


Figure 45.28: The Coherent Solvus

Coherent nucleation:

$$W_R^* = \frac{16\pi\gamma^3}{3(\underbrace{\frac{\Delta G_m}{V_m} + \frac{W_{el}}{V^B}}_{\text{This term < 0 for nucleation to occur}})^2}$$

Figure 45.29: Expression for the coherent solvus.

(1) Coherent interfaces have lower γ - decreases W_R^*

(2) Coherent interfaces have lower magnitude of $W_R^* = \frac{\Delta G_m}{V_m} + \frac{W_{e\ell}}{V^{\beta}}$ - increases W_R^*

Plot W_R^*vs. supersaturation $\left(X_0 - X_{eq}^{\beta}\right)$for both cases:

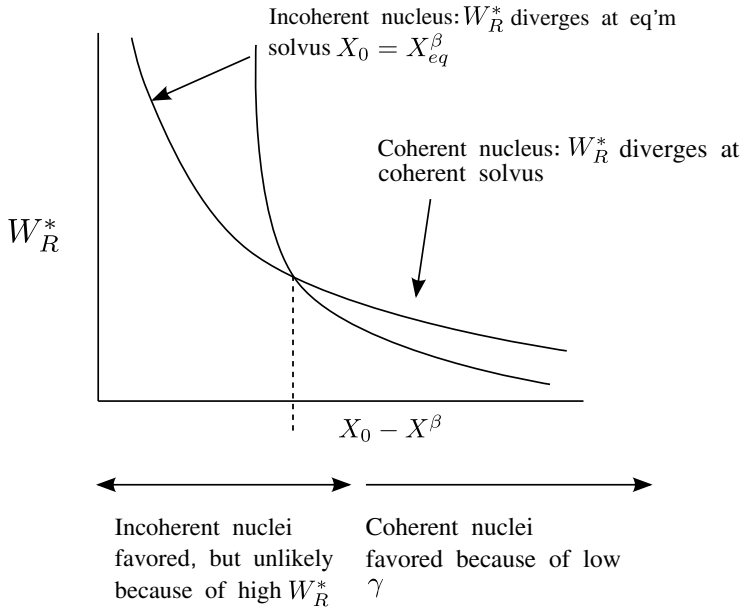


Figure 45.30: Composition dependence of the nucleation rates near the coherent solvus.

Homogeneous nucleation rate:
$\dot{N}_v = v_0 C_0 \exp\left(\frac{-W_R^*}{k_B T}\right)$
v_0 =frequency at which atoms are added to a critical nucleus
One component system: $v_0 \approx 10^{11} s^{-1}$(independent of temp.)
Two component system: atoms must diffuse to critical nucleus - thermally activated process - activation energy = ΔG_d (per atom)

$$\nu_0 = \omega \exp\left(\frac{-\Delta G_d}{k_B T}\right)$$

$$\dot{N}_v = \omega C_o \exp\left(\frac{-\Delta G_d}{k_B T}\right) \exp\left(\frac{-W_R^*}{k_B T}\right)$$

Decreases with decreasing temperature

Increases with decreasing temp

Figure 45.31: Temperature dependence of the nucleation rate in the solid state.

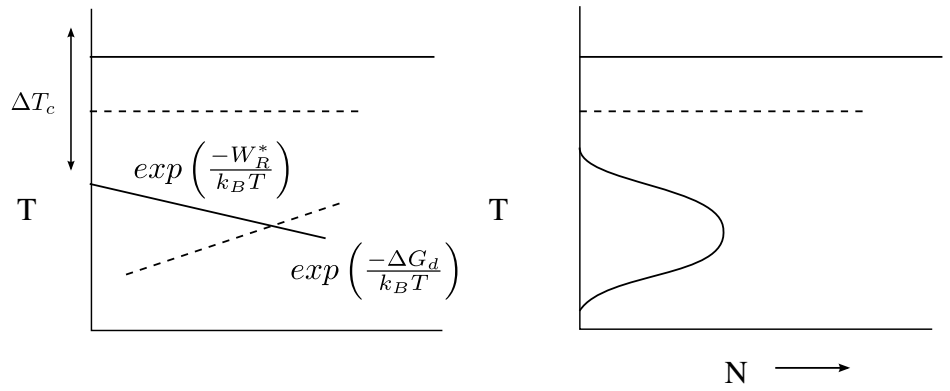


Figure 45.32: Graphical representation of the temperature dependence of the nucleation rate.

Effect of alloy composition:

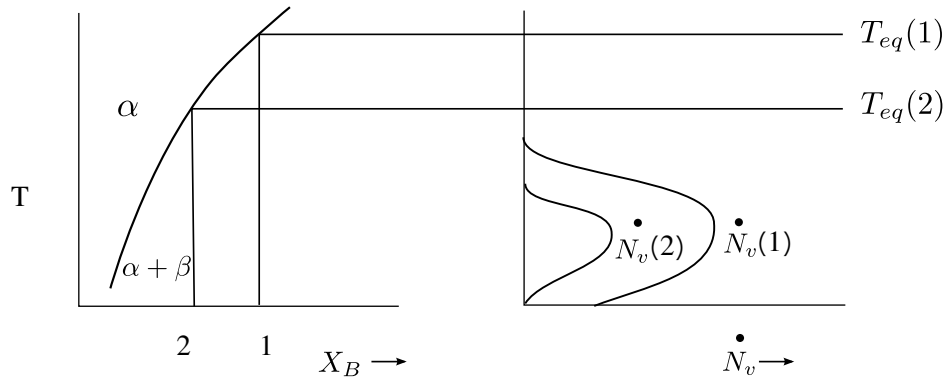


Figure 45.33: Depdendence of the nucleation rate on the supersaturation.

Alloy 1 always has higher nucleation rate: by the time undercooling for alloy 2 is large enough, nucleation is controlled by kinetic term (diffusion)

45.9 Heterogeneous Nucleation

Heterogeneous Nucleation A variety of defects exist in the parent phase:

(1) Vacancies

(2) Dislocations

(3) Stacking faults

(4) Grain boundaries and phase boundaries

(5) Free surfaces

All of these **increase** the free energy of the material

If the creation of a nucleus destroys a defect, some free energy (ΔG_{def}) will be released, reducing activation barrier:

$$\Delta G = V^\beta \left\{ \frac{\Delta G_m}{V_m} + \frac{W_{el}}{V^\beta} \right\} + A^\beta \gamma - \Delta G_{def} \tag{45.68}$$

Nucleation at Grain Boundaries: Eq'm shape for an incoherent grain boundary nucleus: 2 spherical caps:

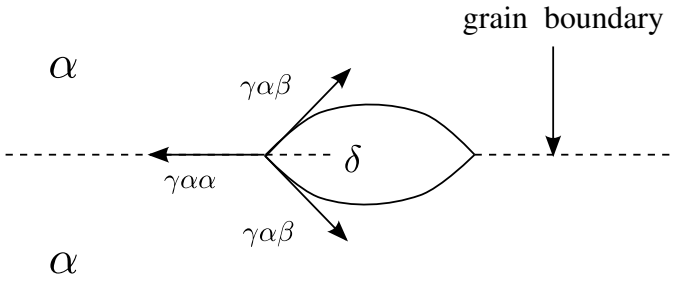


Figure 45.34: A grain boundary precipitate.

$\gamma_{\alpha\alpha} = 2\gamma_{\alpha\beta}cos\theta$, $cos\theta = \frac{\gamma_{\alpha\alpha}}{2\gamma_{\alpha\beta}}$
$\gamma_{\alpha\alpha}$= grain boundary free energy, $\gamma_{\alpha\beta} = \alpha/\beta$ interfacial free energy
$\Delta G = V^{\beta}\left\{\frac{\Delta G_m}{V_m}\right\} + A^{\alpha\beta}\gamma_{\alpha\beta} - A^{\alpha\alpha}\gamma_{\alpha\alpha}$ ($\Delta G_{e\ell} = 0$ for incoherent precipitate)
more convenient to rewrite in terms of $\gamma_{\alpha\beta}$and θ.

$$\Delta G = V^{\beta}\left\{\frac{\Delta G_m}{V_m}\right\} + \gamma_{\alpha\beta}\left(A^{\alpha\beta} - 2A^{\alpha\alpha}\gamma_{\alpha\beta}cos\theta\right) \tag{45.69}$$

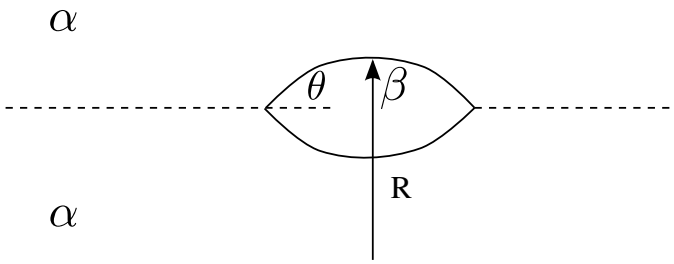


Figure 45.35: Geometry of a grain boundary precipitate.

Some geometry:

$$A^{\alpha\alpha} = \pi R^2 sin^2\theta,\ A^{\alpha\beta} = 2\pi R^2(1 - cos\theta),\ V^{\beta} = \pi R^3(2 + cos\theta)(1 - cos\theta)^2/4$$

Substitute for V^{β}, $A^{\alpha\beta}$and ΔG_{def}:

$$\Delta G = \underbrace{\frac{\pi R^3}{4}(2+\cos\theta)(1-\cos\theta)^2\left\{\frac{\Delta G_m}{V_m}\right\}}_{\text{Volume Term}} + \underbrace{\pi R^2\gamma_{\alpha\beta}\left(2-2\cos\theta-2\sin^2\theta\cos\theta\right)}_{\text{Surface Term}}$$

Figure 45.36: Surface and volume terms for a grain boundary precipitate.

Similar to previous result for heterogeneous nucleation (Draw free energy curve)
Recall:

$$R^* = \frac{-2\gamma_{\alpha\beta}}{\Delta G_m / V_m}$$

$$W_R^* = \frac{16\pi\gamma^3}{3(\Delta G_m/V_m)^2} S(\theta),\ S(\theta) = \frac{1}{2}(2 + cos\theta)(1 - cos\theta)^2$$

Ability of grain boundary to enhance nucleation depends on $cos\theta = \frac{\gamma_{\alpha\alpha}}{2\gamma_{\alpha\beta}}$
No nucleation barrier if $\theta = 0$, $cos\theta = 1$, $\gamma_{\alpha\alpha} = 2\gamma_{\alpha\beta}$ ($2\alpha/\beta$ interfaces replace single grain boundary)
$R^*$and W_R^*reduced even further by nucleation at grain edge or corner (see Fig. 5.7, 5.8, 5.9)
$S(\theta)$ can be calculated for these different geometries.
Grain boundary nucleation favored for high $\gamma_{\alpha\alpha}$(high angle boundary) and for low $\gamma_{\alpha\beta}$(coherent precipitates with one grain) Fig. 5.10
Show figures 5.7 - 5.10

Nucleation at Dislocations 3 effects:

(1) Reduction in $W_{e\ell}$ due to preexisting strain in vicinity of dislocations (most important effect):
ex. edge dislocation:

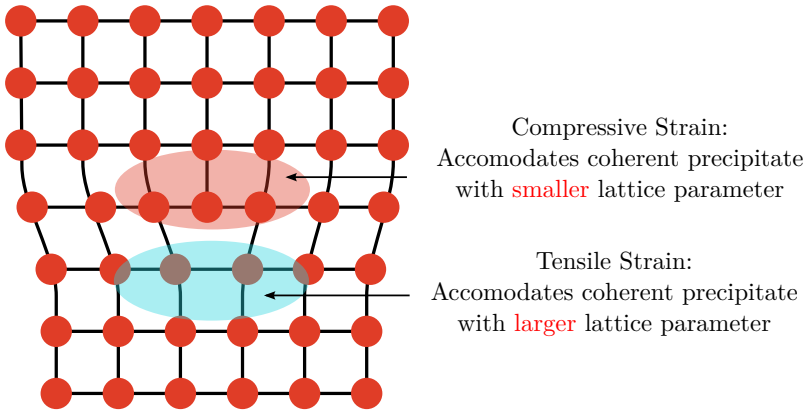


Figure 45.37: Nucleation at an edge dislocation.

Lattice mismatch for coherent precipitate more readily accomodated in vicinity of dislocation

(2) Preferential segregation of impurities to the dislocation can provide 'head start' to the formation of the precipitate

(3) Dislocations can enhance the kinetic contribution associated with impurity diffusion to the nucleus:

$$\dot{N}_v = \omega C_0 \exp\left(\frac{-\Delta G_d}{k_B T}\right) \exp\left(\frac{-W_R^*}{k_B T}\right) \tag{45.70}$$

ΔG_d for diffusion along a dislocation is less than for diffusion through perfect crystal

Vacancy Effects If an alloy is heat treated and rapidly quenched from a high temp, excess vacancies can be quenched in:
$X_v^{eq} = \exp\left(\frac{-\Delta G_v}{k_B T}\right)$increases with increasing temperature
X_v^{eq} = eq'm mole fraction of vacancies
ΔG_v= free energy of formation for a vacancy
$X_v$can exceed eq'm value if sample is rapidly cooled from high temperature $(5 \times 10^4 K/s)$
Excess vacancies enhance nucleation by:

(1) increasing diffusion in parent phase

(2) relieving misfit strain

46 Coarsening Kinetics

Suppose that precipitates with a variety of sizes have appeared:

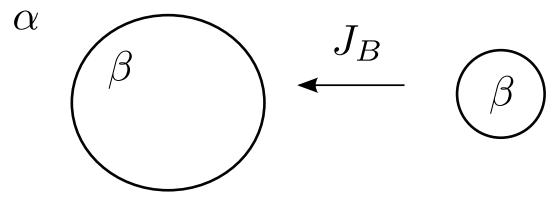


Figure 46.1: Flux of atoms from a small precipitate to a larger one.

For solution obeying Henry's law: $\frac{C_r}{C_e} = 1 + \frac{2\gamma V_m}{RTr}$
Diffusive flux of atoms from small ppts. to large ppt. -
Small ppts. disappear at the expense of larger ones
Average number of precipitates decreases, total fraction of β remains constant

Concentration profile in vicinity of a ppt.:

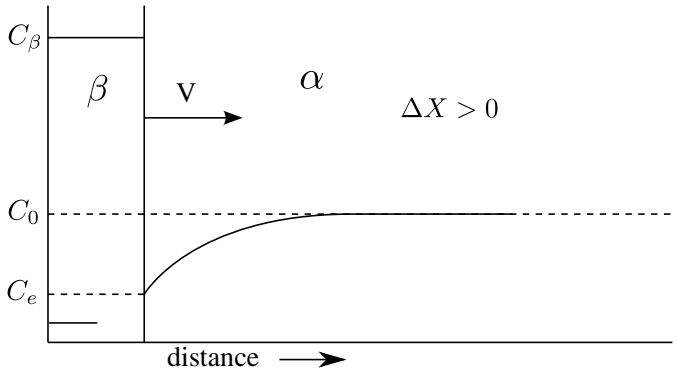


Figure 46.2: Concentration profile near a large precipitate.

$$V = \frac{D}{C_\beta - C_r}\frac{C_0 - C_r}{kr} \tag{46.1}$$

$k = 1$ for spheres, so $V = \frac{dr}{dt} = \frac{1}{r}\frac{D(C_0 - C_r)}{(C_\beta - C_r)}$
C_0 is alloy composition in equilibrium with ppt. of average size, $\bar{r}$:

$$\frac{C_r}{C_e} = 1 + \frac{2\gamma V_m}{RTr} \tag{46.2}$$

$$\frac{C_0}{C_e} = 1 + \frac{2\gamma V_m}{RT\bar{r}} \tag{46.3}$$

If $r > \bar{r}$, the $C_r < C_0$and a ppt. grows (see Figure 46.2).
If $r < \bar{r}$, the $C_r > C_0$and a ppt. shrinks:

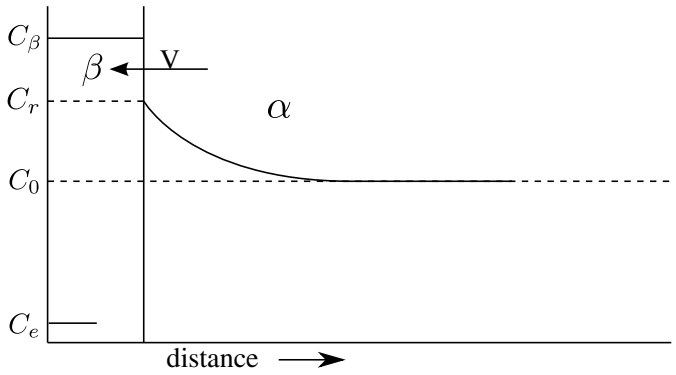


Figure 46.3: Concentration profile near a small precipitate ($r < \bar{r}$)

$$V = \frac{dr}{dt} = \frac{1}{r}\frac{D(C_0 - C_r)}{(C_\beta - C_r)}\,, C_r = C_e + \frac{2\gamma V_m C_e}{RTr}, C_0 = C_e + \frac{2\gamma V_m C_e}{RT\bar{r}}$$

$\frac{dr}{dt} = \frac{2\gamma V_m D C_e}{RTr(C_\beta - C_r)}\left(\frac{1}{r} - \frac{1}{\bar{r}}\right) C_r - C_e$is small, so $C_\beta - C_r \approx C_\beta - C_e$:
$\frac{dr}{dt} = \frac{2\gamma V_m D C_e}{RTr^2(C_\beta - C_e)}\left(\frac{r}{\bar{r}} - 1\right)$ppts. with $r/\bar{r} > 1$ grow, those with $r/\bar{r} < 1$ shrink
Comments:

- $\frac{dr}{dt} \propto \gamma$
- $\frac{dr}{dt} \propto D$
- $\frac{dr}{dt} \propto C_e$
- $\frac{dr}{dt} \propto 1/r^2$- the smallest precipitates grow the fastest

Once can show that the distribution of ppt. sizes relaxes a universal function of $r/\bar{r}$:

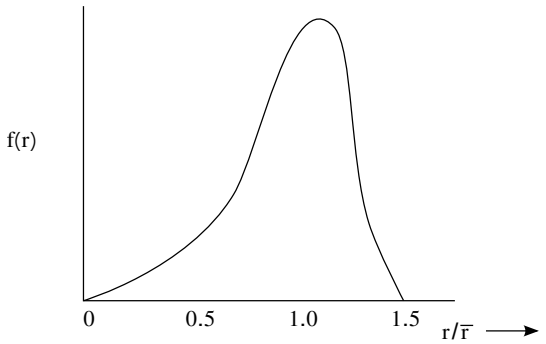


Figure 46.4: Normalized particle size distribution for a coarsening process.

There are no ppts. with radii above $1.5\bar{r}$
Coarsening rate is determined by growth of average ppt size:

$$r^3(t) = r^3(0) + \frac{8\gamma V_m D C_e}{9RT(C_\beta - C_e)}t \tag{46.4}$$

The beginning of coarsening ($t = 0$) corresponds to the elimination of all supersaturation except that associated with curvature, i.e., C_0 is reduced to the point where $\frac{C_0}{C_e} = 1 + \frac{2\gamma V_m}{RTr}$
Assumptions of the theory (Lifshitz Slyozov theory):

(1) Coarsening is diffusion controlled (not interface controlled)

(2) Particles are well separated so that depletion zones do not overlap

(3) Diffusion is uniform and isotropic - short circuit diffusion paths are not important.

47 Precipitate Growth

Nuclei usually bounded by a combination of coherent or semicoherent facets, and smoothly curved incoherent interfaces:

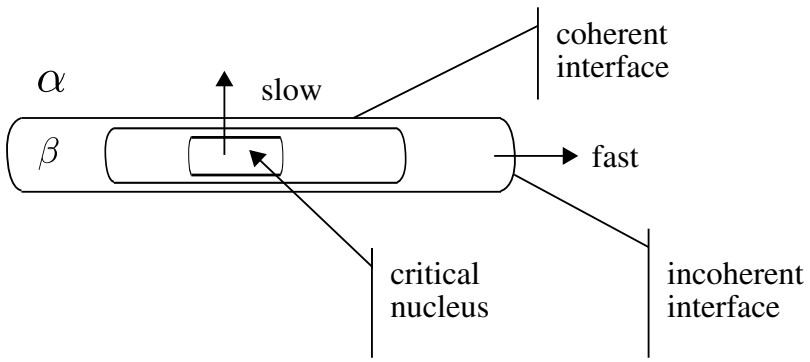


Figure 47.1: Growth of a precipitate with coherent and incoherent interfaces.

If αandβ have different crystal structures, coherent interfaces have low mobility (interface control) and move by a ledge mechanism.
Incoherent interfaces are diffusion controlled, and can move much more quickly.
Precipitates grow as flat plates in this case,

47.1 Growth of Planar Incoherent Interfaces

Nucleation may occur at grain boundaries - many nuclei can coalesce to form a slab:

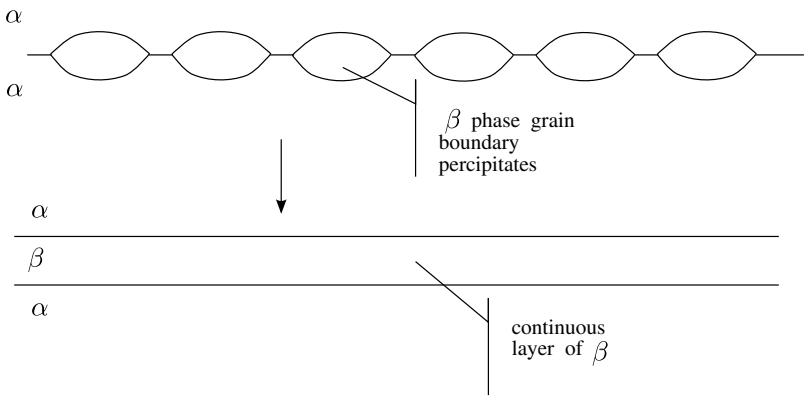


Figure 47.2: Coalescence of grain boundary precipitates.

Phase diagram (work in terms of concentrations): **Draw diagrams on side board**

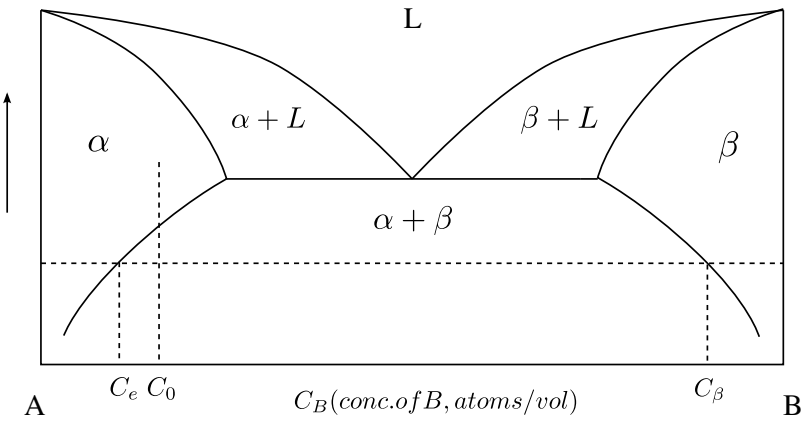


Figure 47.3: Concentration-based eutectic phase diagram.

Driving force for growth is supersaturation of B in α phase $(C_0 > C_e)$
Diffusion controlled growth $\rightarrow$assume local equilibrium at interface $(C = C_e$at interface)

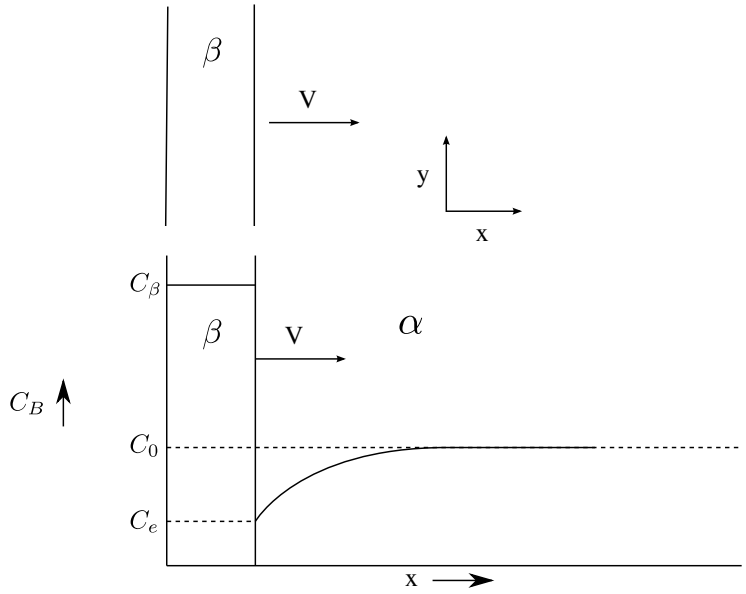


Figure 47.4: Precipitate growth.

Assume quasi steady state: C_0does not change with time.
Fick's first law: $J\left(\frac{atoms}{cm^2 sec}\right) = -D\left(\frac{cm^2}{sec}\right)\frac{\partial C_B}{\partial x}\left(\frac{atoms}{cm^3}\frac{1}{cm}\right)$
Flux depends on conc. gradient $\left(\frac{\partial C_B}{\partial x}\right)$ at α/β interface
Two steps in determination of growth rate:

(1) Relate V to J (and $\left(\frac{\partial C_B}{\partial x}\right)$)

(2) Estimate $\left(\frac{\partial C_B}{\partial x}\right)$

Relationship between V and J:
Consider a section of interface with an area of A.
Motion by a distance dx means that volume Adx is converted from α(with $C_\beta = C_e$) to β (with $C_B = C_\beta$)
Total increase of B atoms in region of width dx $Adx(cm^3)(C_\beta - C_e)(atoms/cm^3)$

These atoms are supplied by the flux, J, given by: $J\left(\frac{atoms}{cm^2 sec}\right) A(cm^2) dt(sec)$
Equating these expressions gives:

$$JAdt = Adx(C_\beta - C_e) \tag{47.1}$$

$$V = \frac{dx}{dt} = \frac{J}{C_\beta - C_e} = \frac{D}{C_\beta - C_e}\frac{\partial C_B}{\partial x} \tag{47.2}$$

Estimation for $\frac{\partial C_B}{\partial x}$:
Simplify concentration profile

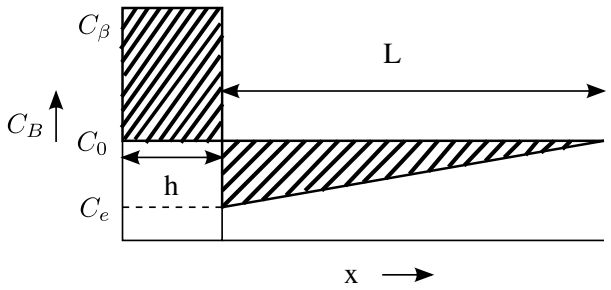


Figure 47.5: Simplification of the concentration profile in the vicinity of a growing precipitate.

$\frac{\partial C_B}{\partial x} = \frac{\Delta C_0}{\ell}$ $\Delta C_0 = C_0 - C_e$ (supersaturation)
ℓ = width of the 'depletion zone' in the α phase ($\sim D/V$)
h = thickness of the β ppt. (half the total thickness if it is growing in the negative x direction as well)

Shaded areas must be equal: B atoms from depletion zone are added to the growing precipitate:

$$(C_\beta - C_0)h = L\Delta C_0/2 \tag{47.3}$$

Rearranging gives the following expression for the width of the depletion zone:

$$L = \frac{2(C_\beta - C_0)h}{\Delta C_0} \tag{47.4}$$

$$V = \frac{D}{C_\beta - C_e}\frac{dC_B}{dx} = \frac{D(\Delta C_0)^2}{2(C_\beta - C_e)(C_\beta - C_{0)}h} \tag{47.5}$$

If molar volume (V_m) does not depend on concentration, we can replace concentrations by mole fractions ($C = X/V_m$):
also assume $C_\beta - C_0 \approx C_\beta - C_e$

$$V = \frac{dh}{dt} = \frac{D(\Delta X_0)^2}{2(X_\beta - X_e)^2}\frac{1}{h} \tag{47.6}$$

integrate: $\int_0^h h dn = \frac{D(\Delta X_0)^2}{2(X_\beta - X_e)^2}\int_0^t dt$

$$h^2 = \frac{D(\Delta X_0)^2}{(X_\beta - X_e)^2}t \text{ (parabolic behavior)}$$

$$h = \frac{\Delta X_0}{X_\beta - X_e}\sqrt{Dt} \tag{47.7}$$

$$V = \frac{dh}{dt} = \frac{1}{2}\frac{\Delta X_0}{X_\beta - X_e}\sqrt{\frac{D}{t}} \tag{47.8}$$

Comments:

(1) $x \propto \sqrt{Dt}$ parabolic growth law for diffusion limited process

(2) $V \propto \Delta X_0$ at constant time, growth rate is proportional to the supersaturation

(3) Diffusion is a thermally activated process, and becomes slow at sufficiently low temperatures

Overall temperature dependence:

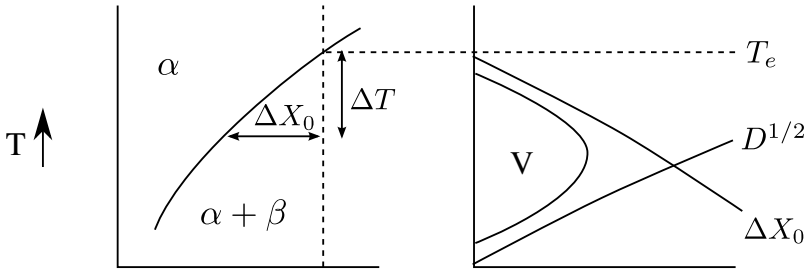


Figure 47.6: Temperature dependence of the precipitate growth velocity.

V has a maximum at an intermediate undercooling:

- D is too small at low temperatures (high undercoolings)
- ΔX_0 is to small at high temperatures (low undercoolings)

When depletion zone (ℓ) becomes comparable to spacing precipitates, V decreases even more:

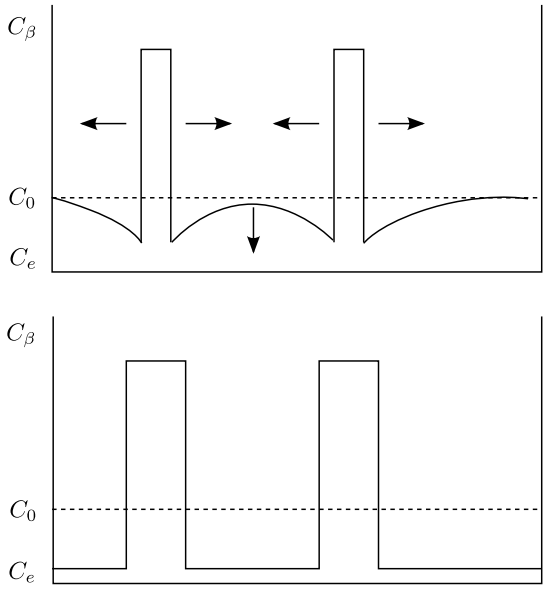


Figure 47.7: Saturation of the growth process

Precipitate growth stops when matrix concentration equals $C_e$everywhere

Grain boundary precipitates may grow much faster because of fast diffusion paths:

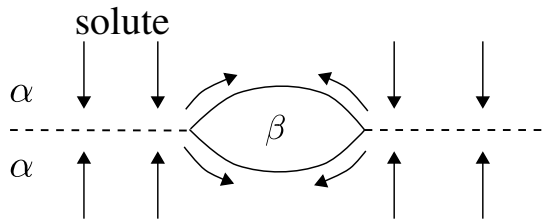


Figure 47.8: Short circuit diffusion paths along grain boundaries.

Steps involved in precipitate growth:

(1) Volume diffusion of solute to grain boundary

(2) Diffusion of solute along GB with some attachment to precipitate

(3) Diffusion along α/β interfaces, allowing accelerated thickening

note: most important for substitutional impurity diffusion
Diffusion Controlled Lengthening of Plates or Needles
Recall reasons for anisotropic growth rates

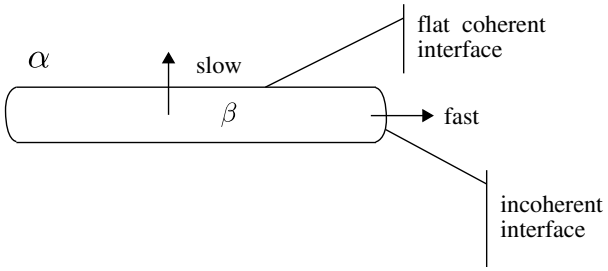


Figure 47.9: Growth of a plate-like precipitate.

Assume that the coherent interface is not mobile - consider growth of curved, incoherent interface

Equilibrium compositions are affected by the curvature:

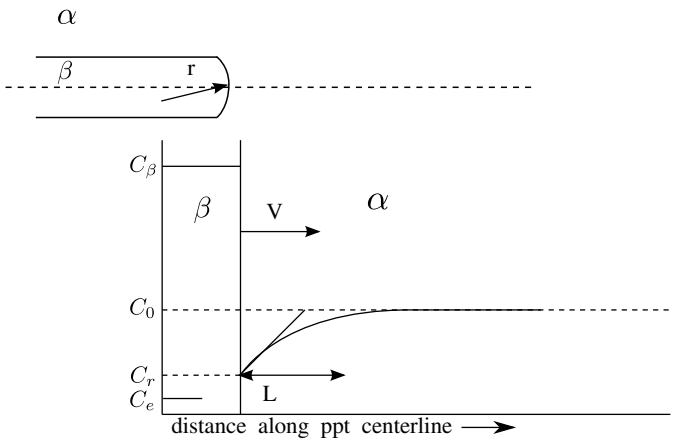


Figure 47.10: Curvature effect on precipitate growth.

Just like previous drawing, except that equilibrium concentration at the interface depends on radius of curvature (Gibbs-Thomson effect)
For a flat interface: $C_r = C_e$
For an interface curved outward (like drawing): $C_r > C_e$
For an inward curvature: $C_r < C_e$

47.2 Curvature Effects

Curvature Effect

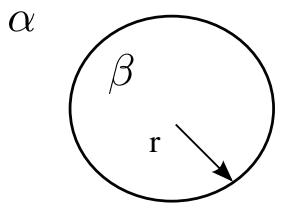


Figure 47.11: Spherical precipitate.

$G_m(P) = G_m(P_0) + V_m(P - P_0)$ Assuming incompressibility, V_m not a function of P)
Laplace Young eq'n: $P^\beta - P^\alpha = \frac{2\gamma}{r}$
For a flate interface ($r = \infty$): $P^\beta = P^\alpha \equiv P^\beta_\infty$
so $P^\beta = P^\beta_\infty + \frac{2\gamma}{r}$
How does this affect G_m:

$$G^\beta_m(P^\beta) = G^\beta_m(P^\beta_\infty) + V_m\left(\frac{2\gamma}{r}\right) \tag{47.9}$$

The interfacial free energy increases G^β_m for a curved interface

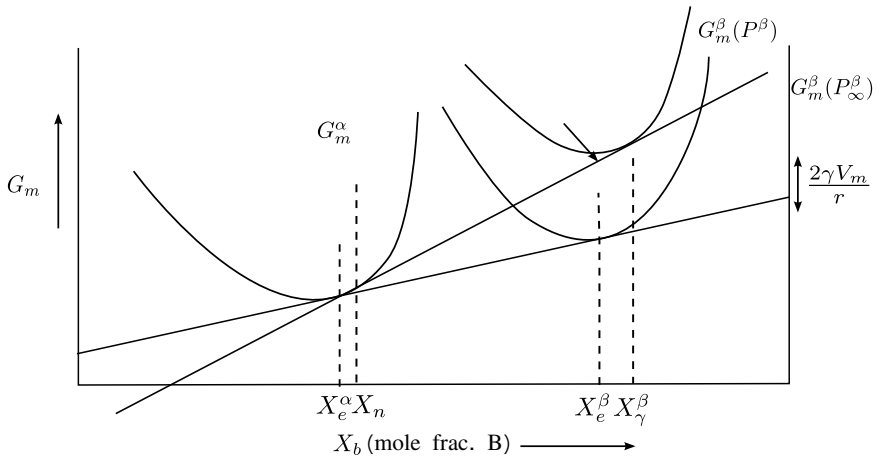


Figure 47.12: Curvature effect on the free energy.

Definitions and assumptions:
Assume nearly pure β: $X^\beta_e \approx X^\beta_r \equiv X_\beta \approx 1$

$$X^\alpha_e \equiv X_e$$

We previously developed an expression between X and ΔG_m:

$$\Delta G_m(X_r) = \frac{2\gamma V_m}{r} = \left.\frac{\partial \mu_{ab}}{\partial X_b}\right|_{X_b = X_e^\alpha} (X_r - X_e)(X_\beta - X_e) \tag{47.10}$$

exchange chemical potential:

$$\mu_{ab} = \frac{\partial G_m}{\partial X_b} = \mu_b - \mu_a = RT(\ln a_b - \ln a_a) \tag{47.11}$$

Here a_b,a_a are activity coefficients
For low X_b(nearly pure A in α phase):

$$a_a = 1 - X_b \approx 1 \text{ (Rault's law)}$$

$$a_b = \Gamma X_b \text{ (Henry's law, } \Gamma\text{=Henry's law coefficient)}$$

$$\mu_{ab} = RT(\ln \Gamma + \ln X_b) \tag{47.12}$$

$$\frac{\partial \mu_{ab}}{\partial X_b} = \frac{RT}{X_b}$$

Substitute:

$$\frac{2\gamma V_m}{r} = \frac{RT}{X_e}(X_r - X_e)(X_\beta - X_e)$$

$$X_\beta \approx 1,\ X_e << 1, \text{ so } X_\beta - X_e \approx 1$$

$$\frac{2\gamma V_m}{r} = RT\left(\frac{X_r}{X_e} - 1\right) \tag{47.13}$$

$$\frac{X_r}{X_e} = 1 + \frac{2\gamma V_m}{RTr} \tag{47.14}$$

Solubility limit depends on radius:

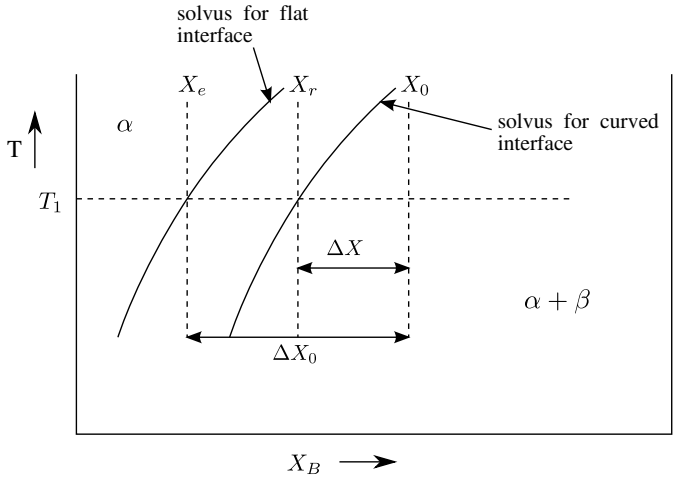


Figure 47.13: Curvature effect on the solvus.

$\Delta X = X_0 - X_r$= supersaturation for curved interface
$\Delta X_0 = X_0 - X_r$ = supersaturation for flat interface

Curvature **decreases** supersaturation:
critical radius (r^*) is value of $r^*$where $\Delta X = 0$, growth stops

$X_0 > X_r (r > r^*)$, concentration gradient drives grown of ppt:

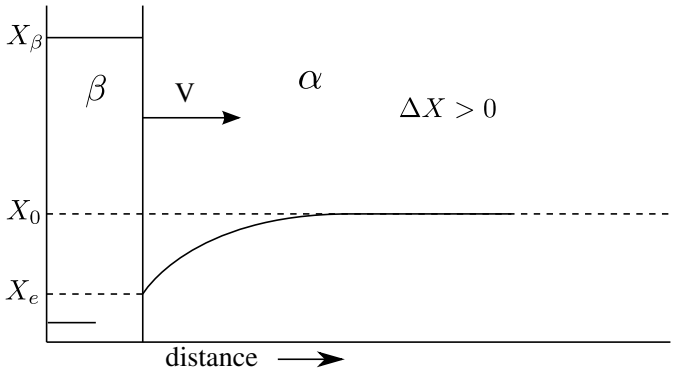


Figure 47.14: Concentration profile for $r > r^*$

$X_0 < X_r (r < r^*)$, concentration gradient drives shrinkage of ppt:

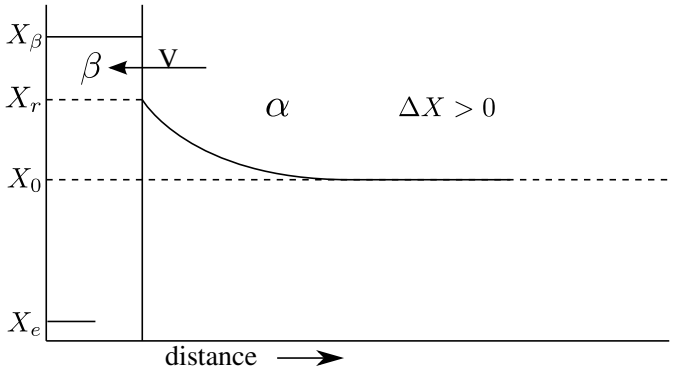


Figure 47.15: Concentration profile for $r < r^*$

$X_0 = X_r (r = r^*)$, critical nucleus - no net tendency to grow or shrink:

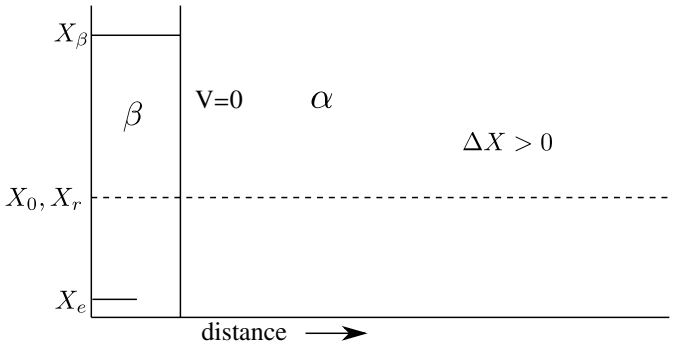


Figure 47.16: Concentration profile for $r = r^*$

For solution obeying Henry's law: $\frac{X_r}{X_e} = 1 + \frac{2\gamma V_m}{RT_r}$

$$X_r = X_e + \frac{2\gamma V_m X_e}{RTr} \tag{47.15}$$

for $r = r^*$, $X_r = X_0$:

$$X_0 = X_e + \frac{2\gamma V_m X_e}{RTr^*} \tag{47.16}$$

Defining ΔX_0 as the difference between X_0 and X_e gives:

$$\Delta X_0 = X_0 - X_e = \frac{2\gamma V_m X_e}{RTr^*} \tag{47.17}$$

$$\Delta X = X_0 - X_r = \frac{2\gamma V_m X_e}{RT}\left(\frac{1}{r^*} - \frac{1}{r}\right) = \frac{2\gamma V_m X_e}{RTr^*}\left(1 - \frac{r^*}{r}\right) = \Delta X_0\left(1 - \frac{r^*}{r}\right)$$

Diffusion Controlled Growth of Plates
(draw schematic of growing plate)
Growth velocity for lengthening of plates: $V = \frac{D}{C_\beta - C_r}\frac{dC}{dx}$ (same as for thickening of plates)
but for lengthening of plates: $\ell \approx kr$, $k \approx 1$(independent of time), $dC/dx \approx \Delta C/kr$

$$V = \frac{D}{C_\beta - C_r}\frac{\Delta C}{kr}$$

for constant V_m: $V = \frac{D}{X_\beta - X_r}\frac{\Delta X}{kr}$,
for $X_\beta \approx 1$, $X_r << 1$, ΔX=$\Delta X_0\left(1 - \frac{r^*}{r}\right)$
$V = D\frac{\Delta X_0}{kr}\left(1 - \frac{r^*}{r}\right)$ ($k \approx 1$, takes into account geometrical effects of diffusion)
constant growth velocity, goes to zero for sufficiently small r because ΔC goes to zero.

48 Case Study: Precipitation of Cu-rich Precipitates from Al:

Al is soft - can be precipitation hardened by adding Cu
Pure Al α phase (κ in above drawing): FCC, a = 4.04Å, valence = 3+
Pure Cu: FCC, a = 3.62Å, valence =1+
Equilibrium precipitate when Al is supersaturated with Cu: $\theta - CuAl_2$- tetragonal

Al-rich region of phase diagram:

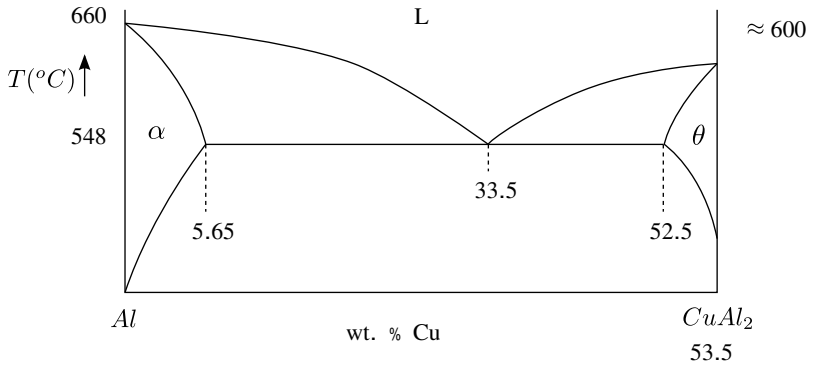


Figure 48.1: Cu-Al phase diagram.

θcannot form coherent interfaces with α, γ for this interface is high:
$W_r^* \propto \gamma^3$, $\dot{N}_v \propto \exp(-W_r^*/k_BT)$, nucleation rate of θ is very slow

Other non-equilibrium precipitates - Guinier-Preston (GP) zones, θ'', θ'
Al: FCC, a = 4.04Å

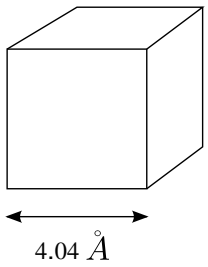


Figure 48.2: Structure of a GP zone

GP Zones: FCC, totally coherent, very small disks: 2 atomic layers thick, 50 atoms in diameter

θ'': tetragonal, FCC distorted in one direction

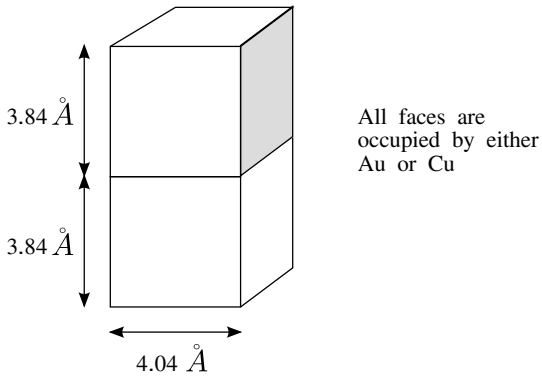


Figure 48.3: Structure of the θ'' phase.

all interfaces coherent with α phase

θ': tegragonal, only one interface coherent with α phase:

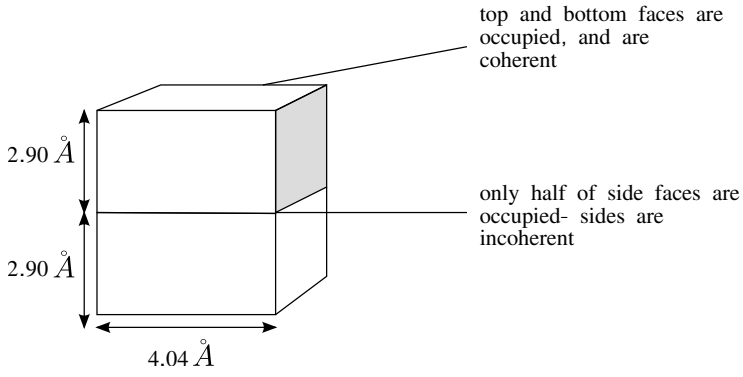


Figure 48.4: Structure of the θ' phase.

θphase: totally incoherent

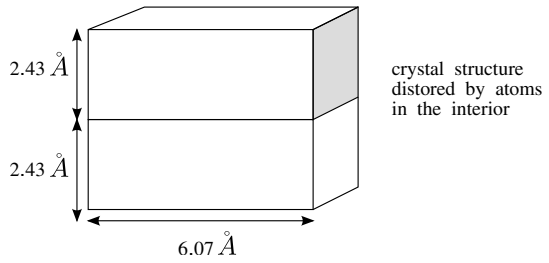


Figure 48.5: Structure of the θ phase.

Rank the values of γ, nucleation rate and thermodynamic stability:

precipitate	γ	$\dot{N}_v$	G_m
GP Zone	1	4	4
θ''	2	3	3
θ'	3	2	2
θ	4	1	1

Table 48.1: Order of the surface energy, nucleation rate and free energy for different Cu-rich precipitate phases in Al, ranked from lowest (1) to highest (4).

Most unstable phase appears first because of higher nucleation rate:
Map time dependence with TTT diagram:

General time dependence for a transformation:

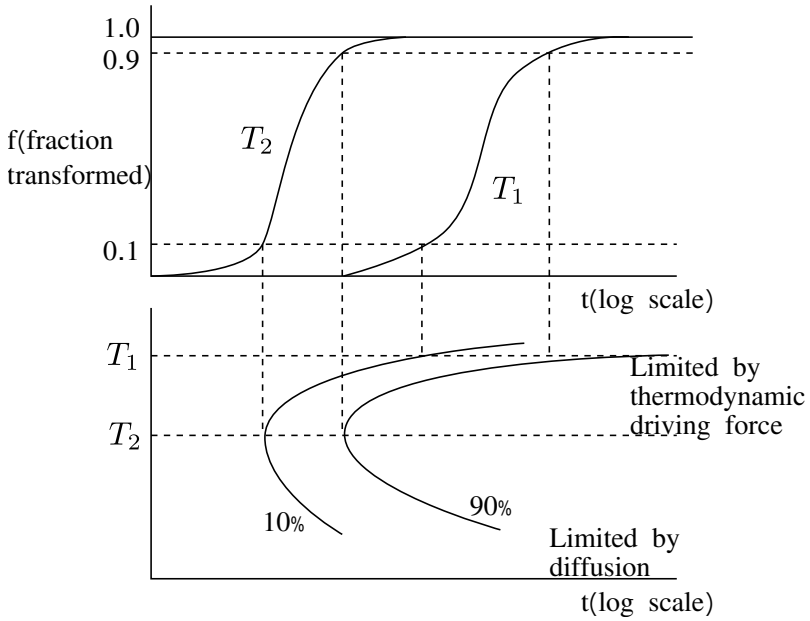


Figure 48.6: Generic TTT diagram.

Unstable compounds only appear at larger undercoolings - different solvus for each phase

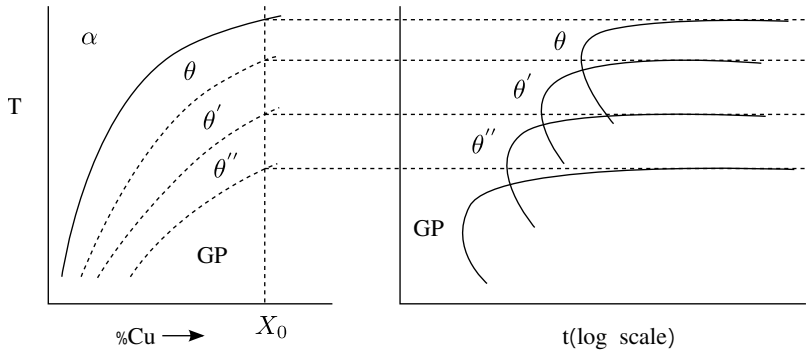


Figure 48.7: TTT diagram for precipitate formation in the Cu-Al system.

At a given undercooling, each phase is in eq'm with a different concentration:

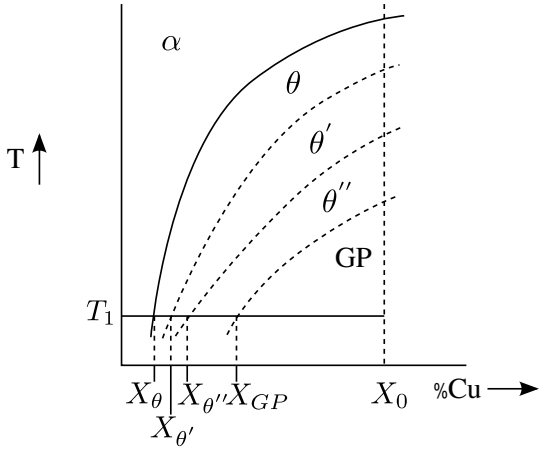


Figure 48.8: Schematic representation of the solvus lines for different precipitate phases in the Al-Cu system.

Each concentration is affected by a local curvature:

$$X_r = X_e \left(1 + \frac{2\gamma V_m}{RTr}\right) \tag{48.1}$$

Once an equilibrium phase is formed, it grows at the expense of a non-eq'm phase because of the lower concentration of Cu in eq'm with it:

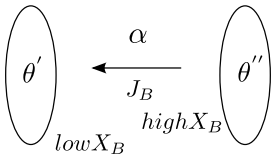


Figure 48.9: Cu flux from a more soluble to a less-soluble precipitate phase.

Matrix concentration of B decreases from $X_0$to $X_{GB}$to $X_{\theta''}$to $X_{\theta'}$to $X_\theta$as different precipitates appear and grow to an appreciable size.

Types of precipitation:

(1) All nucleation occurs at the beginning: site saturation (heterogeneous nucleation sites used up initially)

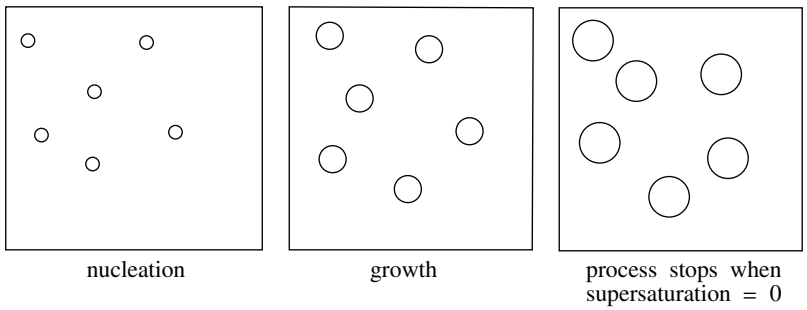


Figure 48.10: Site saturation.

(2) Constant nucleation rate throughout transformation:

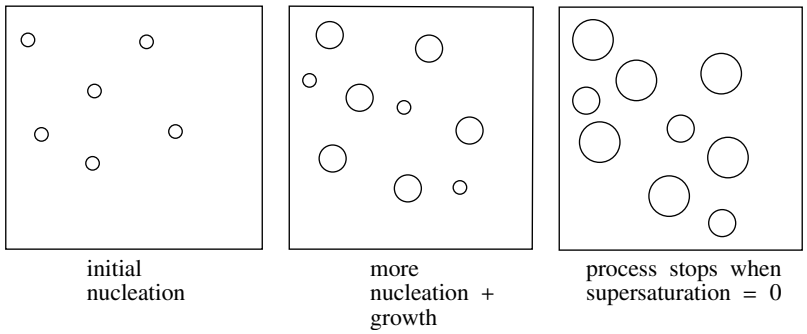


Figure 48.11: Constant nucleation rate throughout precipitation.

(3) Cellular precipitation: entire parent phase is consumed (recrystallization, formation of pearlite)

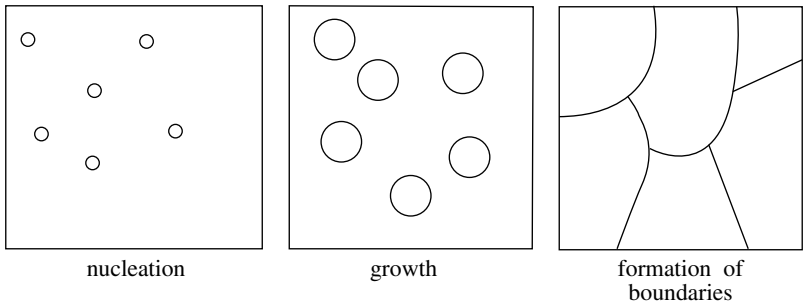


Figure 48.12: Cellular precipitation.

What is time dependence of transformation?
$f(t) = \frac{fraction\ transformed\ at\ time\ t}{final\ fraction\ transformed}$ (depends on temperature)

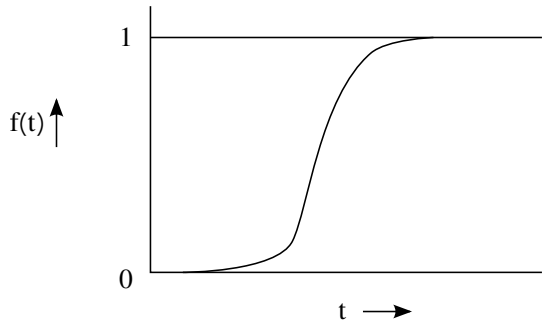


Figure 48.13: Time dependence of transformation

Ex: derivation of $f(t)$for constant $\dot{N}_v$

Vol = volume of a single cell
v = growth velocity
r = cell radius

$Vol(t) = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi(vt)^3$ (nucleation occurs at $t = 0$)

If nucleation does not occur until $t = \tau$:

$$Vol(t,\tau) = \frac{4}{3}\pi v^3(t-\tau)^3 \tag{48.2}$$

$\dot{N}_v d\tau$ = number of nuclei formed during time increment $d\tau$.

$$f(t) = \dot{N}_v \int_0^t Vol(t,\tau)d\tau = \dot{N}_v \frac{4}{3}\pi v^3 \int_0^t (t-\tau)^3 d\tau = \dot{N}_v \frac{\pi}{3} v^3 t^4 \tag{48.3}$$

for long t, $f(t) = 1$:
detailed solution gives:
$f(t) = 1 - \exp\left(-\frac{\pi}{3}\dot{N}_v v^3 t^4\right)$ Johnson-Mehl-Avrami equation
(in agreement with previous expression at small t, since $1 - \exp(-x) \approx x$ for small x
This is a specific example of the following more general expression:

$$f(t) = 1 - \exp(-kt^n) \tag{48.4}$$

n is found to vary between 1 and 4
temperature dependence is in k:
This is the simplest equation that has the basic behavior observed experimentally:

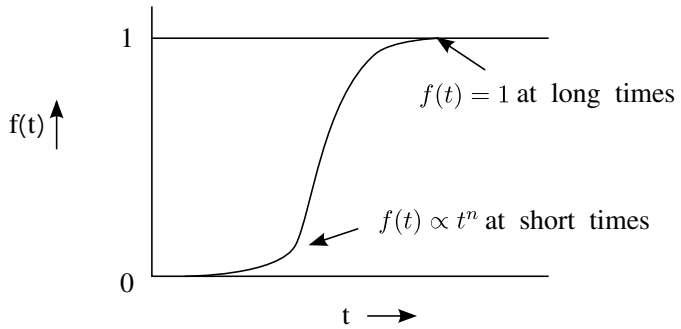


Figure 48.14: Behavior of the Johnson-Mehl Avrami equation

k is low at high temp. because N_v is small (low undercooling)
k is low at low temp. because v is small (slow diffusion)

49 Case Study: Mineralization from Solution

Basic Carbonate forming reaction: (all concentrations in moles/liter)

$$\mathrm{CaCO_3} \leftrightarrow \mathrm{Ca^{2+}} + \mathrm{CO_3^{2-}} \quad \left[\mathrm{Ca^{2+}}\right]\left[\mathrm{CO_3^{2-}}\right] = 6\mathrm{x}10^{-9} \tag{49.1}$$

$\left[Ca^{2+}\right]\left[CO_3^{2-}\right] > 6x10^{-9}$ for aragonite, vaterite
CO_3^{2-} - concentration affected by pH

$$HCO_3^- \leftrightarrow H^+ + CO_3^{2-} \frac{\left[H^+\right]\left[CO_3^{2-}\right]}{\left[HCO_3^-\right]} = 5.61x10^{-11} \tag{49.2}$$

$$H_2CO_3 \leftrightarrow H^+ + HCO_3^- \quad \frac{\left[H^+\right]\left[HCO_3^-\right]}{\left[H_2CO_3\right]} = 1.5x10^{-4} \tag{49.3}$$

Dissolved CO_2 in equilibrium with carbonic acid

$$H_2O + CO_2 \leftrightarrow H_2CO_3 \quad \frac{\left[H_2CO_3\right]}{\left[CO_2\right]} = 1.70x10^{-3} \tag{49.4}$$

Dissolved CO_2 in equilibrium with atmospheric CO_2

$$dissolved\, CO_2 \leftrightarrow atmospheric\, CO_2 \quad \frac{P_{CO_2}}{\left[CO_2\right]} = 29.76\, atm/mol/l \tag{49.5}$$

Basic water equilibrium

$$H_2O \leftrightarrow H^+ + OH^- \quad \left[H^+\right]\left[OH^-\right] = 10^{-14} \tag{49.6}$$

Charge neutrality

$$2\left[Ca^{2+}\right] + \left[H^+\right] = \left[HCO_3^-\right] + 2\left[CO_3^{2-}\right] + \left[OH^-\right] + \left[Cl^-\right] \tag{49.7}$$

Other equilibria. If Chloride ions are present:

$$HCl \leftrightarrow H^+ + Cl^- \quad \frac{\left[H^+\right]\left[Cl^-\right]}{HCl} = 10^{-8} \tag{49.8}$$

So we have 8 equations and 10 unknowns:
$\left[Ca^{2+}\right]$, $\left[CO_3^{2-}\right]$, $[H^+]$, $\left[HCO_3^-\right]$, $[H_2CO_3]$, $[CO_2]$, PCO_2, $[OH^-]$, $[HC\ell]$, $[C\ell^-]$
So we need to specify 2 quantities, generally $\left[Ca^{2+}\right]$ and P_{CO2} to completely specify the problem. Note: $pH = -\ell og\left[H^+\right]$
Increasing $[CO_2]$ increases carbonic acid concentration (Eq.(4)), decreases pH, increases $[H^+]$

50 Case Study: Nanoparticle Nucleation and Growth

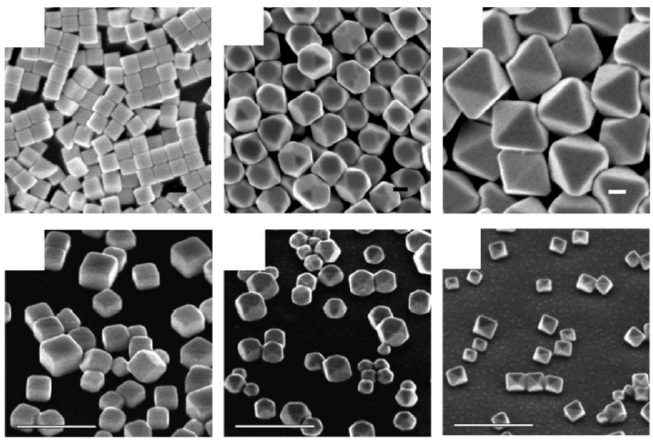


Figure 50.1: Metallic Nanoparticles.[33]

50.1 Motivation

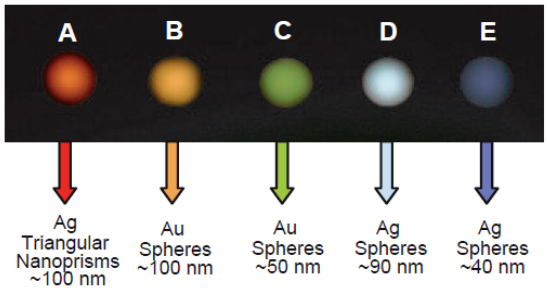


Fig. 5. Rayleigh light-scattering of particles deposited on a microscope glass slide. The slide is used as a planar waveguide, which is illuminated with a tungsten source. The image was taken with a digital camera.

Figure 50.2: Size dependent optical properties of metallic nanoparticles.

50.2 Wulff Constuction

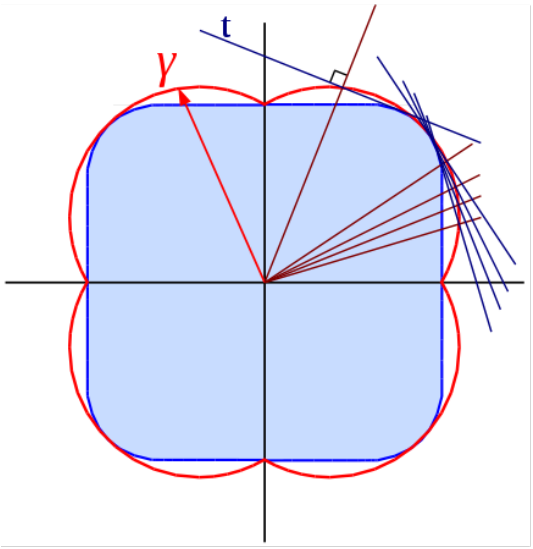


Figure 50.3: The Wulff Construction.

(1) Create polar plot of surface energy (red)

a Find intersections with lines drawn from the origin (brown)

b Draw perpendicular planes (black)

c Eq'm shape defined by innermost planes (blue)

50.3 Kinetic Control of Particle Shape

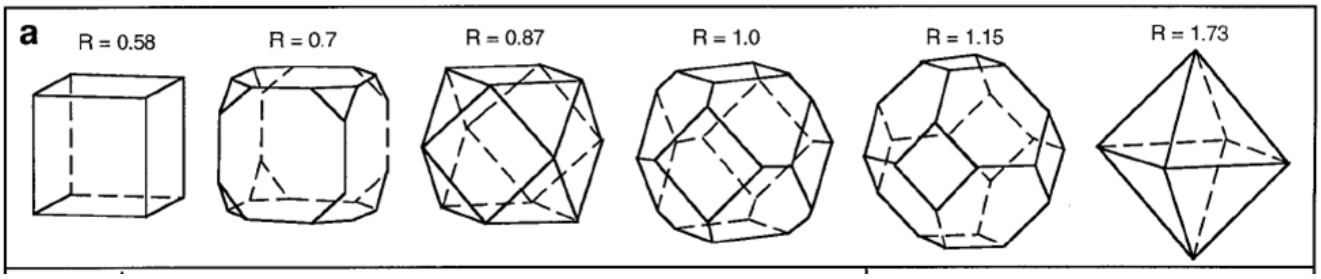


Figure 50.4: Dependence of particle shape on the growth rate ratio.

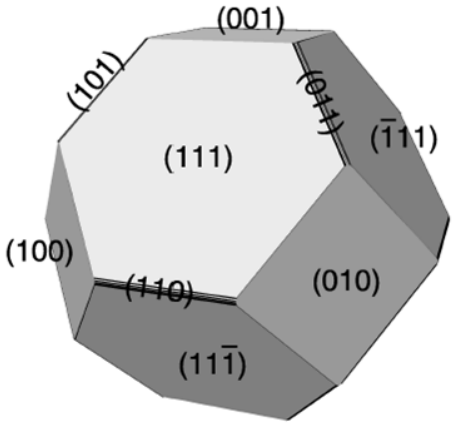


Figure 50.5: Crystal facets for a cubic material.

50.4 Controlling Growth Rates

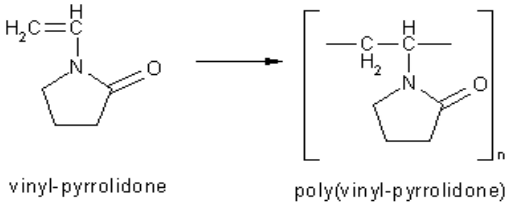


Figure 50.6: Chemical structure of poly(vinyl pyrrolidone).

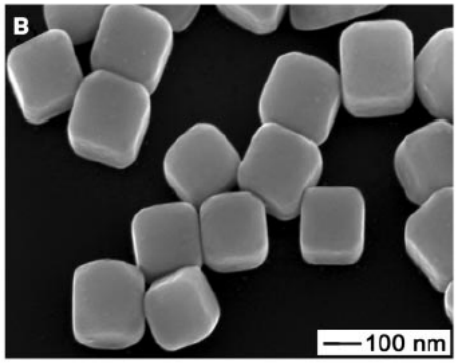


Figure 50.7: Cubic particles form rapid growth in $\langle 111 \rangle$ directions.

- Polymer adsorbs differently to $\{111\}$ and $\{100\}$ facets.
  - Growth in $\langle 111 \rangle$ direction is favored.
  - Cubes result.

Nucleation Rate and Size Distribution

- $W_R^* \propto \frac{1}{(\Delta C)^2}$
- Uniform particle size can be obtained if all nuclei form at same time

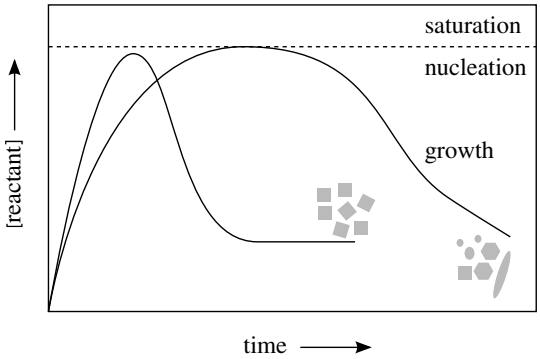


Figure 50.8: Time dependent reactant concentration profiles leading to two types of particle size distributions.

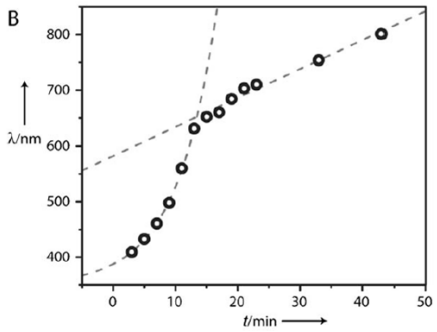


Figure 50.9: Optical Determination of Nucleation and Growth.

- Wavelength (λ) coupled to total nanoparticle volume
- Exponential regime: nucleation and growth
- Linear regime: growth only

51 Interface Stability

In this section we are concerned with factors that cause the interface between a solid and a liquid to propagate in a non planar fashion, resulting in a dendritic microstructure of the sort shown below in Figure. 51.1.

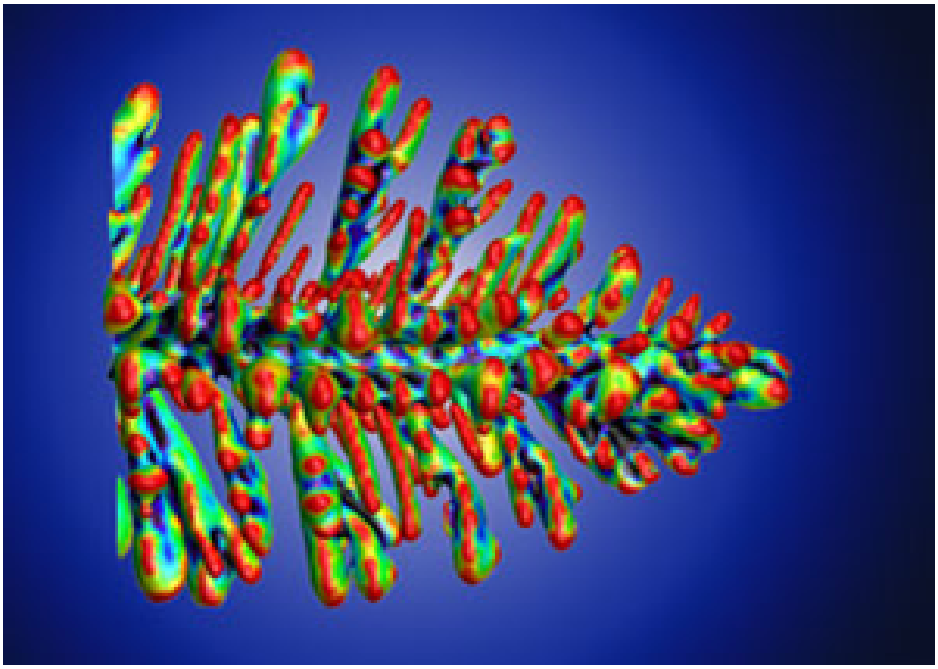


Figure 51.1: Image of a growth highlighting the curvature.

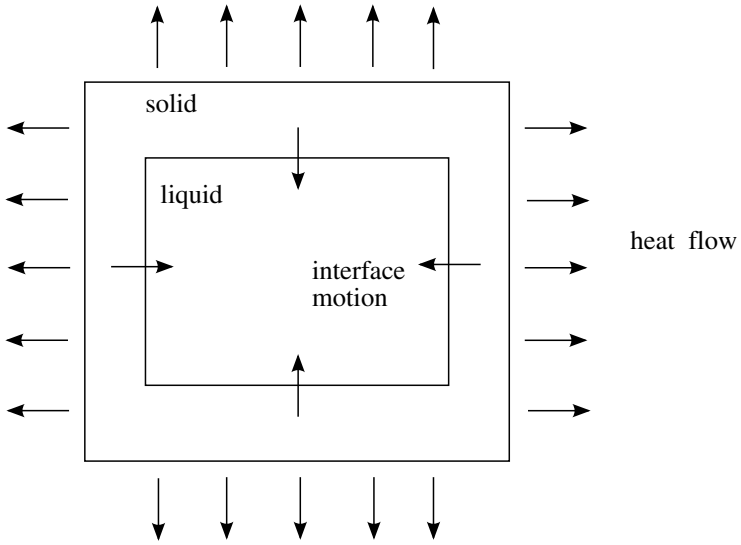


Figure 51.2: Solidification in a cold mold.

Heat flow governed by thermal conductivity: $q = -k\frac{dT}{dx}$
q = heat flux, k = thermal conductivity
Growth into supercooled liquid:

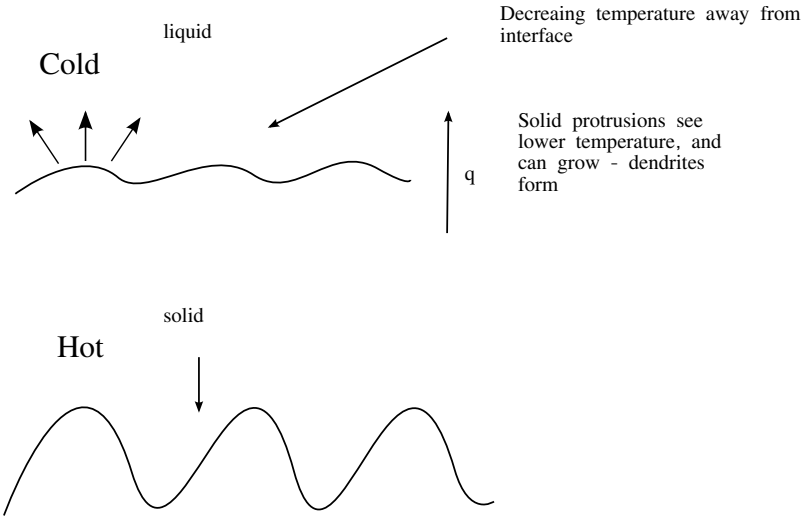


Figure 51.3: Interface instability due to solidification into a supercooled liquid.

Growth by removal of heat from the wall (more common)

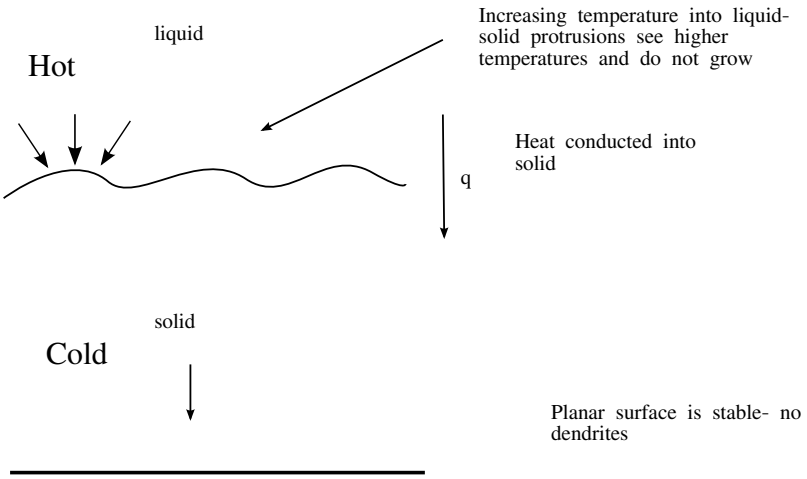


Figure 51.4: Solidification against a cold surface.

Alloys: similar, except formation of dendrites can be controlled by movement of atoms; dendrites can form even when T increases into liquid
Solidify a material with overall composition C_0: higher impurity concentration in liquid phase

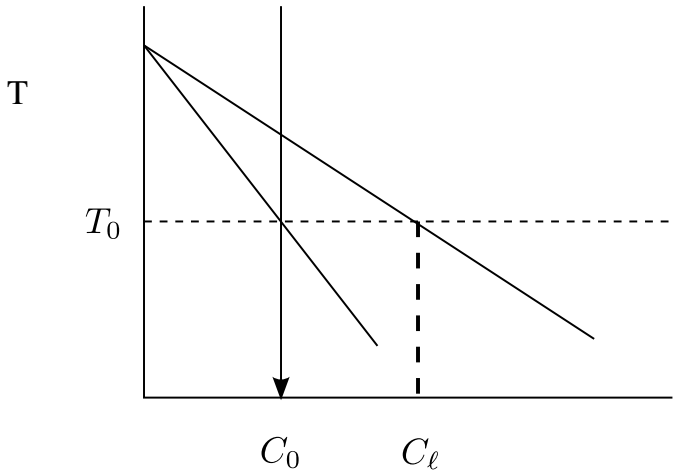


Figure 51.5: Phase diagram for immiscible solid phases.

Composition profile:

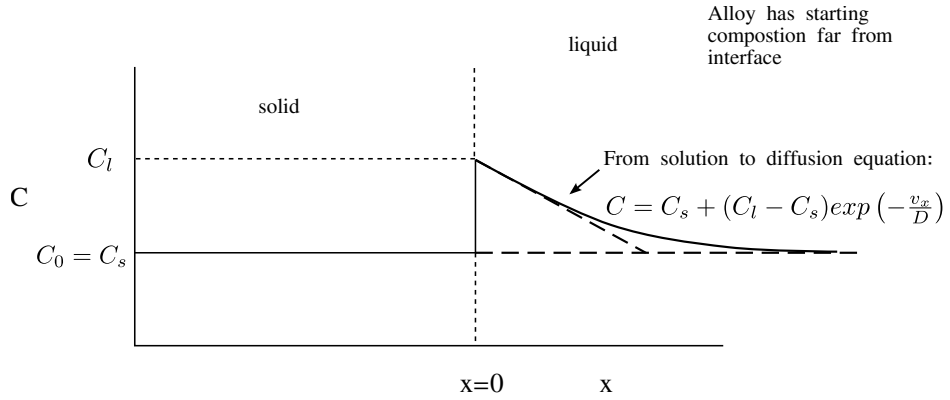


Figure 51.6: Impurity concentration profile at a solid/liquid interface.

V = velocity of solid/liquid interface
D = diffusion of coefficient of impruity in liquid
How does this impurity buildup affect stability of the interface (tendency to form dendrites)
Use concentration profile to define a liquidus temperature that depends on position:

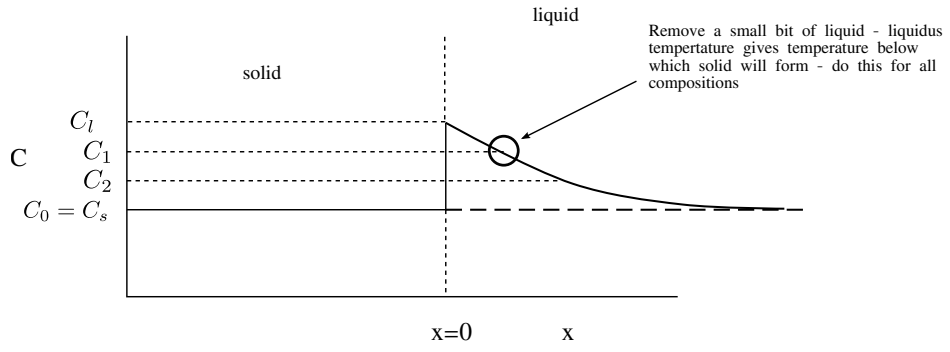


Figure 51.7: Significance of local liquid temperature.

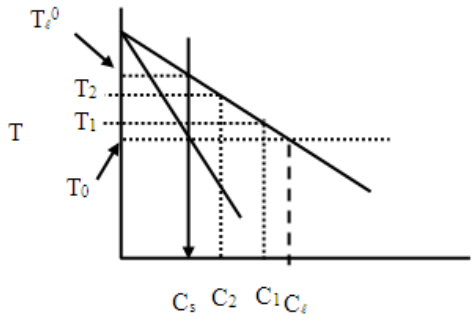


Figure 51.8: Mapping local concentrations to local liquidus temperature.

So we can convert concentration profile to liquidus temperature as a function of distance:

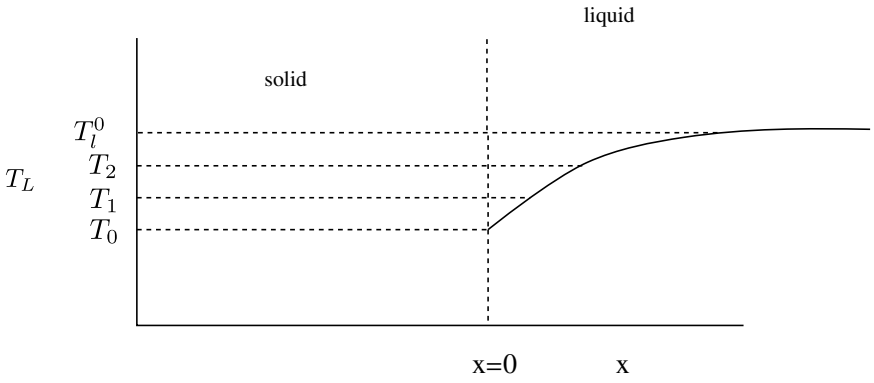


Figure 51.9: Plot of the local liquidus temperature.

If the temperature is less than the local T_ℓ,solid will form, if $T > T_\ell$, no solid forms
Stability criterion determined by comparing actual temperature profiles

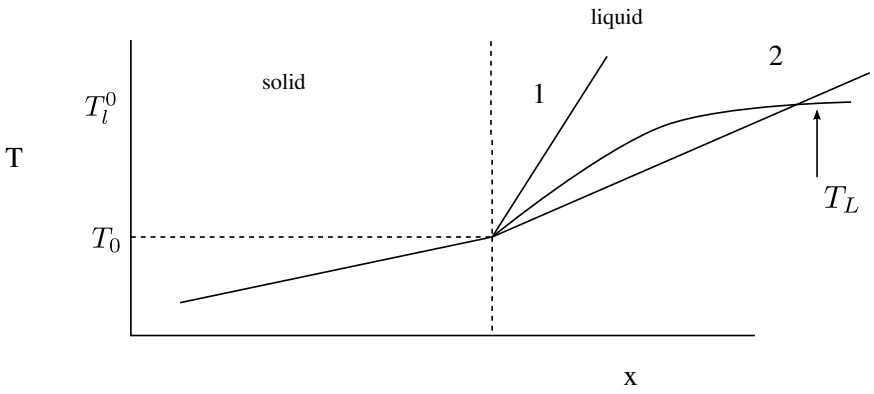


Figure 51.10: Graphical criterion for interface stability.

Case 1: Actual T always above T_ℓ= stable interface
Case 2: Actual T below T_ℓ,solid protrusion can grow:

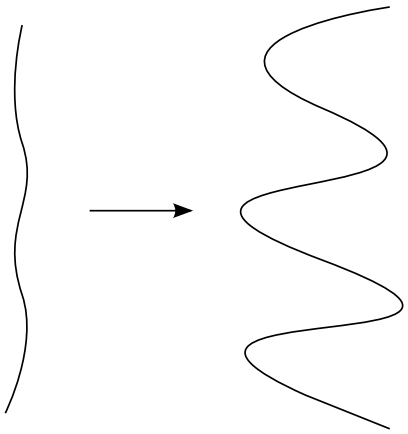


Figure 51.11: Growth of an unstable interface.

Note: interface is always stable for a sufficiently large temperature gradient into the sample.

52 Eutectic Solidification

Thermodynamics review - melting point depression
Thermodynamics: Equation for liquidus line
$X_\ell^\beta(T)$ = liquidus comp. for α side of diagram
$X_S^\beta(T)$ = solidus comp. for α side of diagram
G_B^ℓ = partial molar free energy of liquid B
G_B^S = partial molar free energy of solid B
T_M^A, T_M^B - melting points of pure components

At eq'm:

$$G_B^L(X_L^\beta(T)) = G_B^S(X_S^\beta(T)) \tag{52.1}$$

If $X_S^\beta(T) \approx 1$ (nearly pure solid phase):

$$G_B^S(X_S^\beta(T)) = G_B^L X_S^\beta(T) - \Delta H_B(1 - T/T_m^B) \tag{52.2}$$

Combine (1) and (2):

$$G_B^\ell(X_\ell^\beta) - G_B^\ell(X_S^\beta) = -\Delta H_B\left(1 - \frac{T}{T_m^B}\right) \tag{52.3}$$

Assume liquids form an ideal solution:

$$G_B^L(X_L^\beta) - G_B^L(X_S^\beta) = RT \ln X_L^\beta \tag{52.4}$$

Combine (3) and (4):

$$RT \ln X_\ell^\beta = -\Delta H_B\left(1 - \frac{T}{T_m^B}\right) \tag{52.5}$$

Melting point reduction due to increased entropy of liquid phase
Assumptions:

(1) Ideal mixing in liquid (no enthalpy of mixing)

(2) Solid phases are nearly pure

Optimum phase size for eutectic microstructure: (draw alternating α, β, show wavelength, λ)
Overall enthalpy of melting of solid at composition X_e:
$\Delta H_e = X_B \Delta H_B + (1 - X_B)\Delta H_A$ (weighted average of component enthalpies)

Two contributions to free energy of solidification:

(1) Bulk term:
$\frac{\Delta G_{bulk}}{V} = \frac{-\Delta H_e \Delta T}{V_m T_e}$, ΔT = undercooling $(T_e - T)$

(2) Interfacial term:
within one wavelength - total interfacial area = $2A$ (A = cross section) - 2 α/β interfaces per wavelength)
volume = λA
$\frac{\Delta G_{int}}{V} = \frac{2\gamma_{\alpha\beta}}{\lambda}$
$\frac{\Delta G_{total}}{V} = \frac{2\gamma_{\alpha\beta}}{\lambda} - \frac{\Delta H_e \Delta T}{V_m T_e}$
$\Delta G_{total} < 0$ for $\lambda > \lambda^*$, $\lambda^* = \frac{2\gamma_{\alpha\gamma} V_m T_e}{\Delta H \Delta T}$
compare to $r^*$for nucleation of a solid phase: $r^* = \frac{-2\gamma}{(\Delta G_{bulk}/V)} = \frac{2\gamma V_m T_e}{\Delta H \Delta T}$

Alternate Approach: growth of curved lamellae

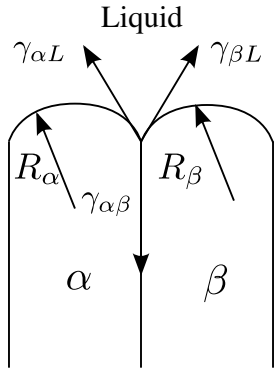


Figure 52.1

$R_\alpha R_\beta \sim \lambda$, details depend on relative values of $\gamma_{\alpha\beta}$, $\gamma_{a\ell}$, $\gamma_{\beta\ell}$.
Supercooled eutectic alloy (composition $X_B = X_E$) - consider α phase

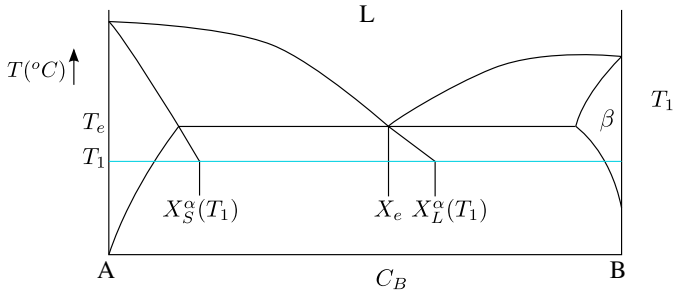


Figure 52.2

Conc. gradient drives B **away** from α/ℓ interface:

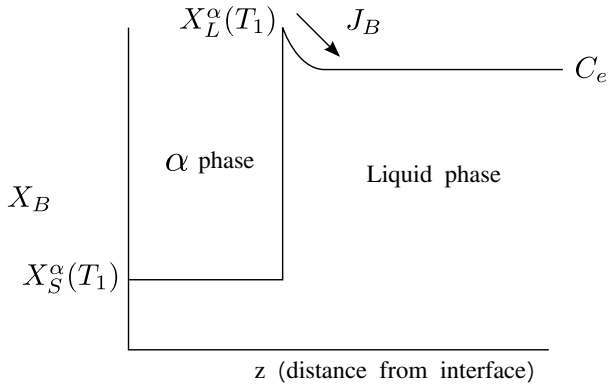


Figure 52.3

Now consider β phase
Supercooled eutectic alloy (composition $X_B = X_E$) - consider α phase

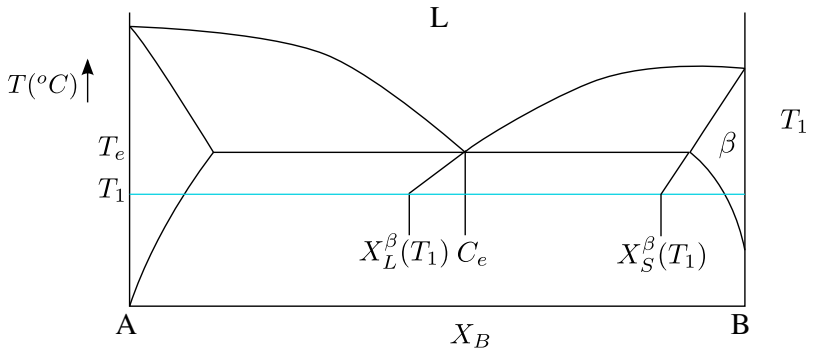


Figure 52.4

Conc. gradient drives B **toward** from β/ℓ interface:

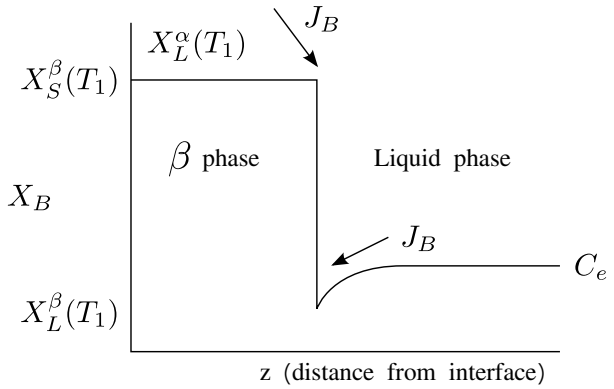


Figure 52.5

Simultaneous growth of $\alpha$and β minimizes required diffusion distance:
In liquid: B is enhanced near α, depleted near β.

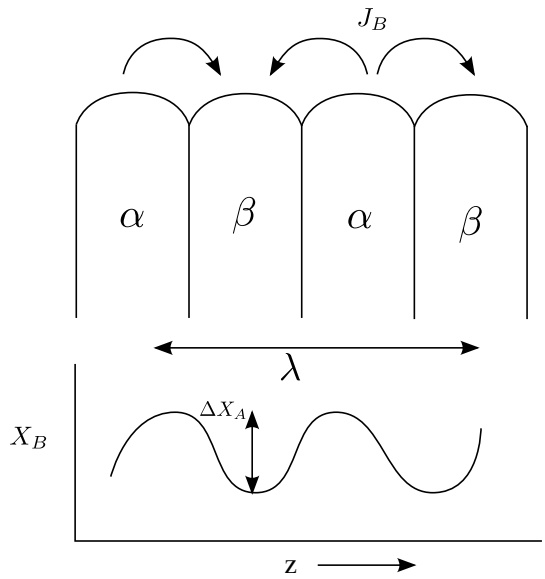


Figure 52.6

Concentration gradient in liquid: $\propto \frac{\Delta X_\lambda}{\lambda}$
Diffusive flux in liquid: $\propto \frac{D\Delta X}{\lambda}$
What determines $\Delta X\lambda$?
Liquidus lines for finite $\lambda$are suppressed in comparison to eq'm ($\lambda = \infty$) values.
Eutectic point of phase diagram:

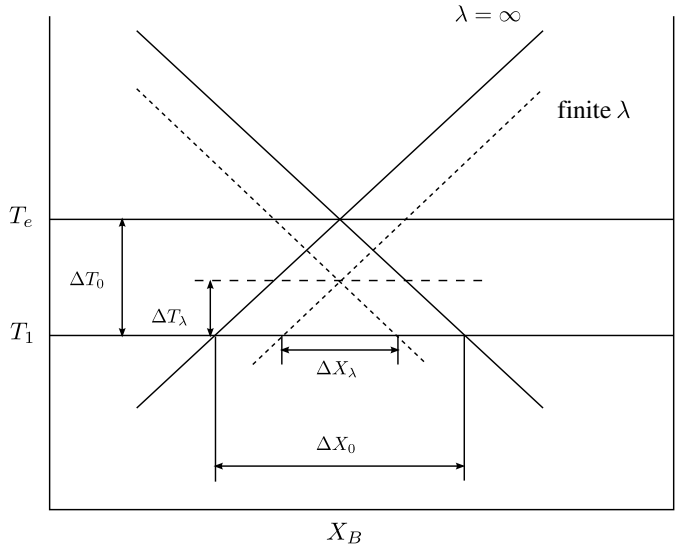


Figure 52.7

(1) Assume liquidus lines are linear in this regime - Must have ΔT_0linearly related to ΔT_λ.

(2) Must have $\Delta X_\lambda = 0$ for $\lambda = \lambda^*$

(3) Must have $\Delta X_\lambda = \Delta X_0$ for $\lambda =$

(4) $\lambda^* \propto \frac{1}{\Delta T_0}$

All this gives: $\Delta X_\lambda = \Delta X_0 \left(1 - \frac{\lambda^*}{\lambda}\right)$
If growth is diffusion limited:

$$V \propto \Delta X_0 \frac{D}{\lambda}\left(1 - \frac{\lambda^*}{\lambda}\right) \tag{52.6}$$

Maximum growth velocity for $\lambda = 2\lambda^*$.

53 Eutectoid Reactions

Eutectic transformation:

$$V \propto \Delta X_0 \frac{D}{\lambda}\left(1 - \frac{\lambda^*}{\lambda}\right)$$

$$\Delta X_0 \propto \Delta T_0$$

Maximum V for $\lambda = 2\lambda^*$.

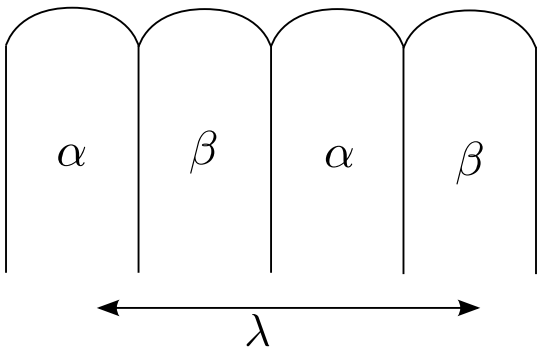


Figure 53.1

Eutectoid transformation: High temp. phase is a solid

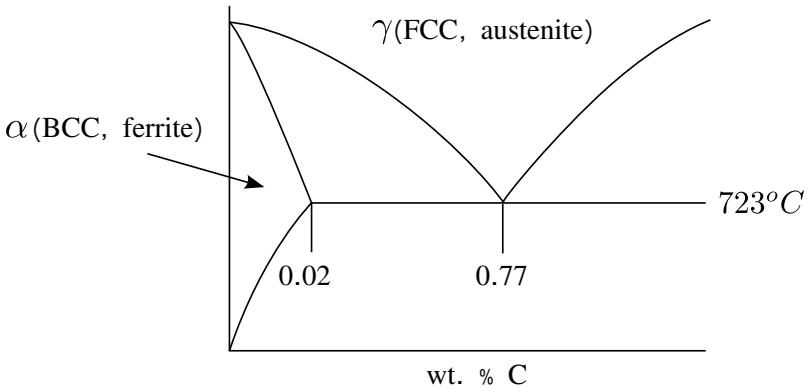


Figure 53.2

$$\gamma \rightarrow \alpha + Fe_3C$$

Lamellar microstructure of α, Fe_3C = pearlite
From lever rule: mostly ferrite
Nucleation occurs preferentially at γ grain boundaries:

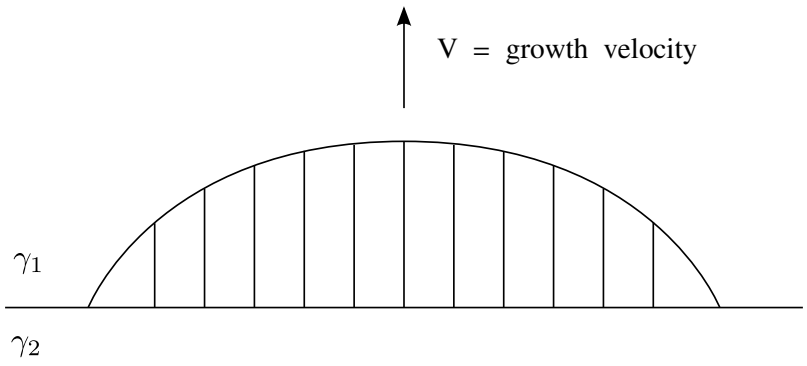


Figure 53.3

Nucleation:

(1) Fe_3C nucleates first

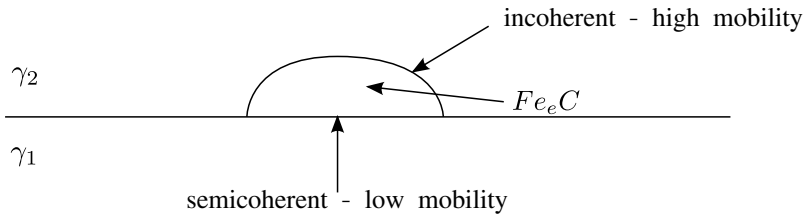


Figure 53.4

(2) region near Fe_3C ppt. depleted in C

(3) α nucleates

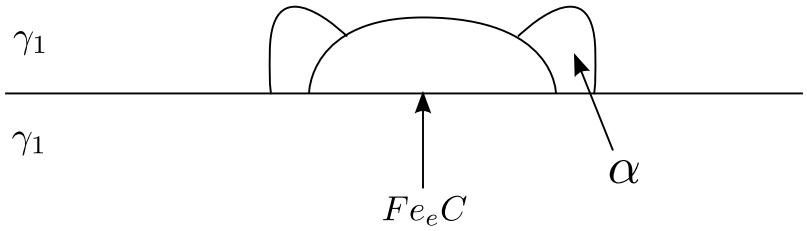


Figure 53.5

Incoherent interfaces of α/γ_2 and Fe_3C/γ_2

Coherent interfaces for α/γ_1 and Fe_2C/γ_1.

Fe_3C is orthorhombic: $a \neq b \neq c$, $\alpha = \beta = \gamma = 90^o$- γ is FCC

Preferred orientation for coherent $Fe_3C\gamma$ interfaces:
$(100)_{Fe_3C}\,(111)_\gamma$;$(010)_{Fe_3C}\,(110)_\gamma$; $(001)_{Fe_3C}\,(112)_\gamma$
Growth rate limited by C diffusion in γ phase (like eutectic case).

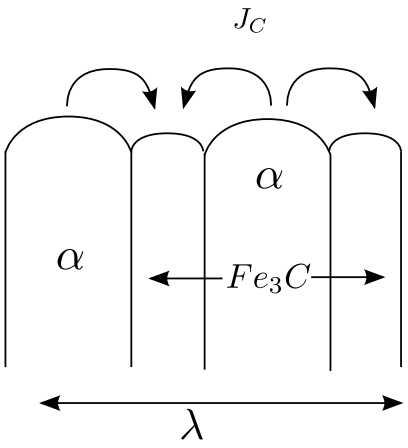


Figure 53.6

Data: Growth rate (Hull, 1942)

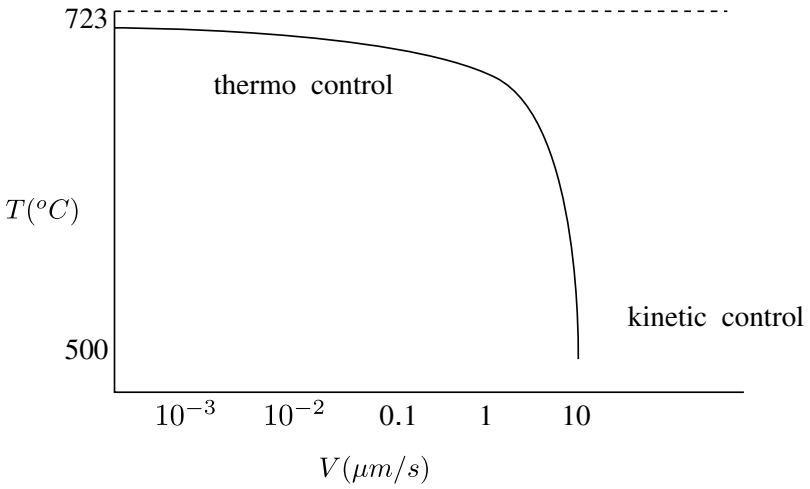


Figure 53.7

Wavelength (Pellissier, 1942)

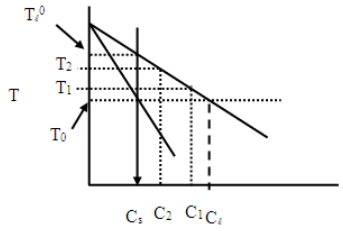
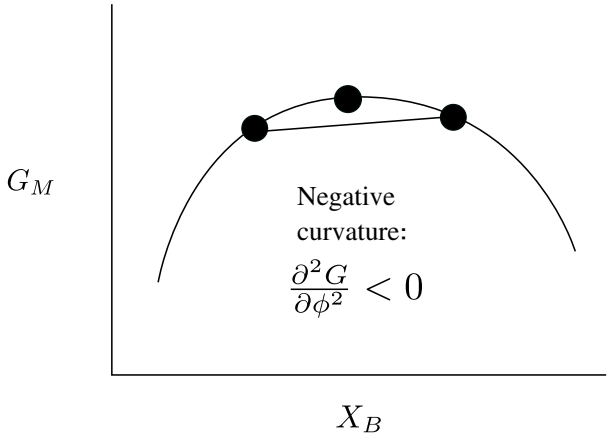


Figure 53.8

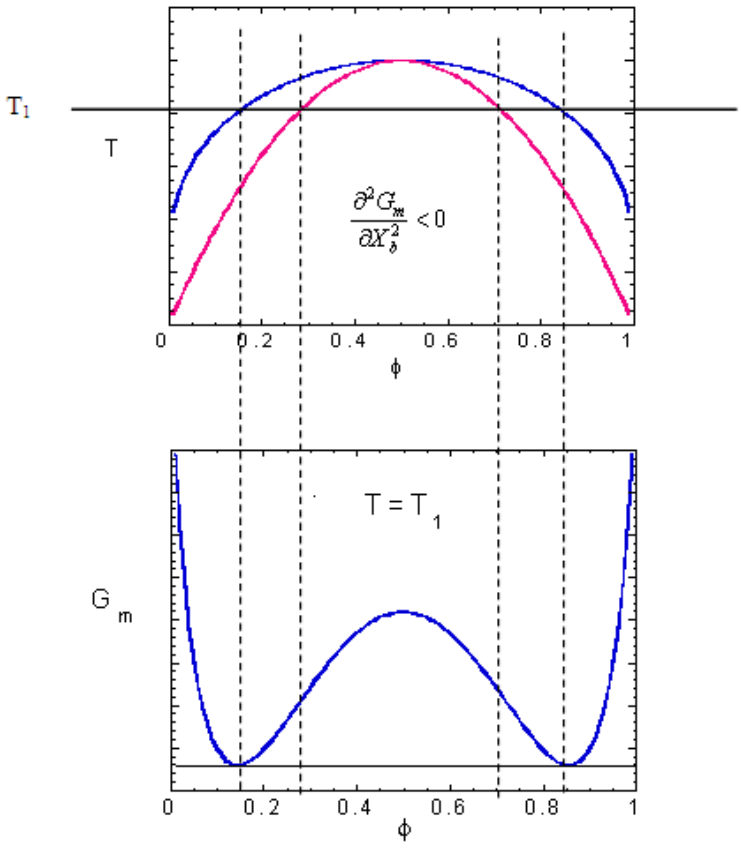
54 Spinodal Decomposition

Definitions
G_v= free energy per volume (zero for pure components)
G_{ext}: extensive free energy
S_v: entropy per volume

H_v: enthalpy per volume
χ: enthalpy of mixing parameter
V_0= reference volume
N_aV_0, N_bV_0= molar volumes of A and B
n_a, n_b: moles of A and B present
ϕ_a, ϕ_b volume fractions of A and B
ϕ_1: eq'm ϕ_b in phase 1
ϕ_2 : eq'm ϕ_b in phase 2
Spionodal Decomposition occurs when $\frac{\partial^2 G}{\partial X^2} < 0$, as illustrated in the following figure:



Example Phase diagram



Spinodal and Binodal Curves

Spinodal curve:

$$\frac{d^2 G_m}{d\phi^2} = 0 \tag{54.1}$$

Binodal curve (common tangent construction) :

$$\mu_a(\phi_1) = \mu_a(\phi_2); \quad \mu_b(\phi_1) = \mu_b(\phi_2) \tag{54.2}$$

Regular solution (quasichemical) model:

$$\frac{V_0 G_v}{RT} = \frac{\phi_a \ln \phi_a}{N_a} + \frac{\phi_b \ln \phi_b}{N_b} + \chi \phi_a \phi_b \tag{54.3}$$

(χdecreases with increasing temp)

Free energy derivatives

$$\phi \equiv \phi_b; \quad \phi_a = 1 - \phi$$

$$\frac{G_v V_0}{RT} = \frac{(1-\phi)\ln(1-\phi)}{N_a} + \frac{\phi \ln \phi}{N_b} + \chi \phi (1-\phi) \tag{54.4}$$

$$\frac{\partial}{\partial \phi}\left(\frac{G_v V_0}{RT}\right) = \frac{1}{N_b} + \frac{\ln \phi}{N_b} - \frac{1}{N_a} - \frac{\ln(1-\phi)}{N_a} + \chi - 2\chi\phi \tag{54.5}$$

$$\frac{\partial^2}{\partial \phi^2}\left(\frac{G_v V_0}{RT}\right) = \frac{1}{\phi N_b} + \frac{1}{(1-\phi)\, N_a} - 2\chi \tag{54.6}$$

Calculation of spinodal curve

$\frac{\partial^2}{\partial \phi^2}\left(\frac{G_v V_0}{RT}\right) = 0$ for $\chi = \chi_s$:

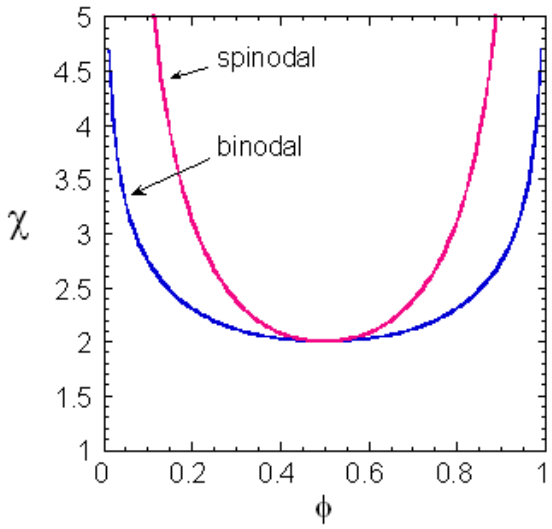
$$\chi_s = \frac{1}{2\phi N_b} + \frac{1}{2(1-\phi)N_a} \tag{54.7}$$

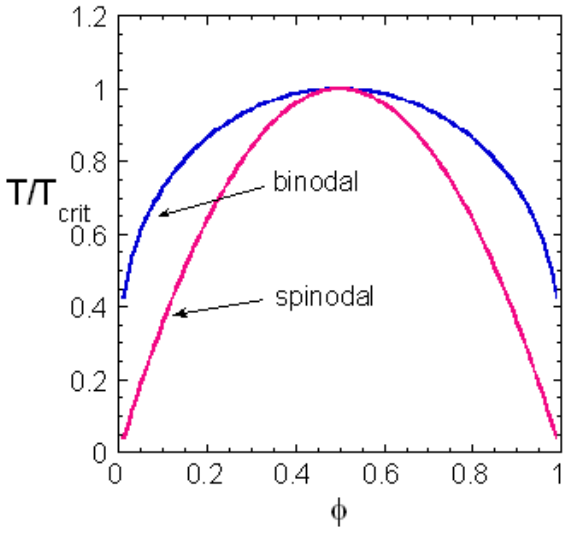
Critical χ is lowest value of χ_s on plot of χ_s vs. ϕ:
$\frac{\partial \chi_s}{\partial \phi} = 0$ for $\phi = \phi_{crit}$,$\chi = \chi_{crit}$

$$\phi_{crit} = \frac{\sqrt{N_a}}{\sqrt{N_a} + \sqrt{N_b}}; \quad \chi_{crit} = \frac{N_a + N_b + 2\sqrt{N_a N_b}}{2 N_a N_b} \tag{54.8}$$

Behavior for $N_a = N_b = 1$

$$\phi_{crit} = 0.5, \chi_{crit} = 2$$





Enthalpy and Entropy (of mixing)

$$\frac{V_0 G_v}{RT} = \frac{\phi_a \ln \phi_a}{N_a} + \frac{\phi_b \ln \phi_b}{N_b} + \chi \phi_a \phi_b \tag{54.9}$$

$$G_v = H_v - TS_v \tag{54.10}$$

$$H_v = \frac{\chi RT}{V_0} \phi_a \phi_b; \;\; S_v = -\frac{R}{V_0} \left[\frac{\phi_a \ln \phi_a}{N_a} + \frac{\phi_b \ln \phi_b}{N_b} \right] \tag{54.11}$$

Relationship between χ and T

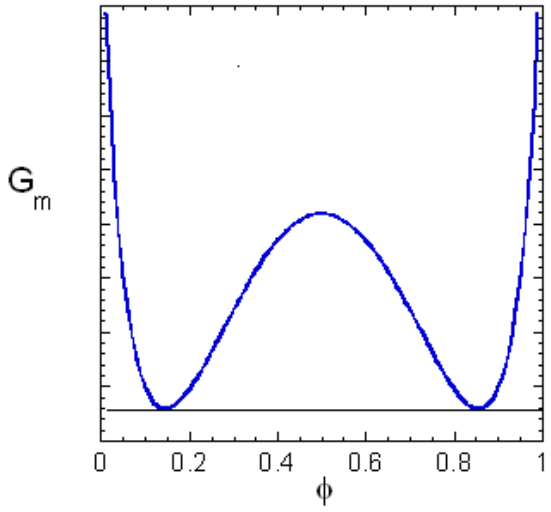
For ideal regular solution model:

$$\chi \propto \frac{1}{T}$$

$\chi = \chi_{crit}$ when $T = T_{crit}$, so

$$\chi = \frac{\chi_{crit} T_{crit}}{T} \tag{54.12}$$

Behavior for $N_a = N_b$
binodal can be calculated analytically because free energy curves are symmetric about $\phi = 0.5$:



$$\frac{\partial G}{\partial \phi} = 0 \text{ for } \chi = \chi_b$$

$$N_a = N_b = N$$

$$\frac{\partial}{\partial \phi}\left(\frac{G_v V_0}{RT}\right) = 0 = \frac{\ln \phi}{N} - \frac{\ln(1-\phi)}{N} + \chi - 2\chi\phi \tag{54.13}$$

$$N\chi_b = \frac{\ln\left(\frac{1-\phi}{\phi}\right)}{1-2\phi} \tag{54.14}$$

$$N\chi_s = \frac{1}{2\phi} + \frac{1}{2(1-\phi)} \tag{54.15}$$

Behavior for $N_a \neq N_b$

- Need a numerical solution (*i.e.* MATLAB)
- obtain excessive free energy by multiplying by system volume:

$$G_{ext} = G_v V$$

$$V = V_0(n_a N_a + n_b N_b)$$

n_a, n_b= total numbers of A and B molecules

- Rewrite volume fractions in terms of n_a, n_b :

$$\phi_a = \frac{n_a N_a}{n_a N_a + n_b N_b}$$

Chemical Potential Expressions

After a bunch of algebra...

$$\frac{G_{ext}}{RT} = n_a \ln \phi_a + n_b \ln \phi_b + \chi n_a N_a \phi_b$$

differentiate G_{ext} to get chemical potentials

$$\mu_a = \frac{\partial G_{ext}}{\partial n_a}; \quad \frac{\mu_a}{RT} = \ln \phi_a + \phi_b \left(1 - \frac{N_a}{N_b}\right) + \chi N_a \phi_b^2$$

$$\frac{\mu_b}{RT} = \ln \phi_b + \phi_a \left(1 - \frac{N_b}{N_a}\right) + \chi N_b \phi_a^2 \tag{54.16}$$

Equate Chemical Potentials
phase 1 has $\phi_b = \phi_1$, $\phi_a = 1 - \phi_1$; phase 2 has $\phi_b = \phi_2$, $\phi_a = 1 - \phi_2$
Equate μ_a in both phases:

$$\ln(1-\phi_1) + \phi_1 \left(1 - \frac{N_a}{N_b}\right) + \chi N_a \phi_1^2 = \ln(1-\phi_2) + \phi_2 \left(1 - \frac{N_a}{N_b}\right) + \chi N_a \phi_2^2 \tag{54.17}$$

Equate μ_b in both phases:

$$\ln \phi_1 + (1-\phi_1)\left(1 - \frac{N_b}{N_a}\right) + \chi N_b (1-\phi_1)^2 = \ln \phi_2 + (1-\phi_2)\left(1 - \frac{N_b}{N_a}\right) + \chi N_b (1-\phi_2)^2 \tag{54.18}$$

2 equations, 2 unknowns (ϕ_1 and ϕ_2) - use fsolve
Cahn Hilliard Theory
- evolution of composition in the spinodal regime

Composition Fluctuations about some average value

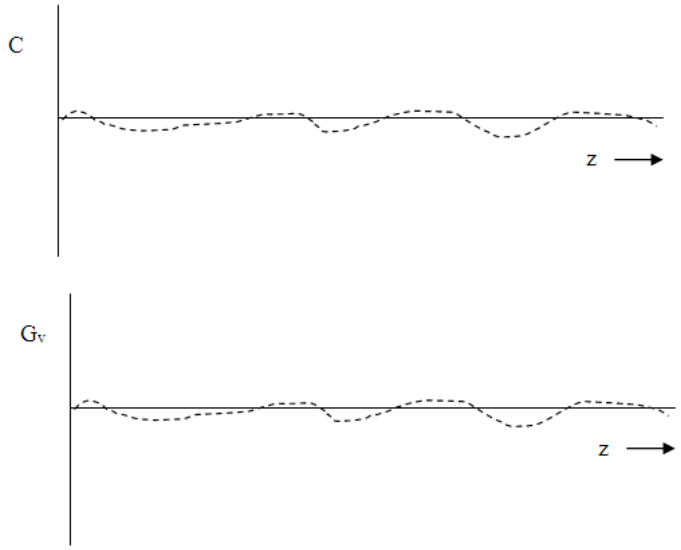


Figure 54.1

Cross sectional area = A
Total free energy:

Assume a sinusoidal perturbation:

$$\phi(z) = \phi_0 + \delta sin(2\pi z/\lambda); \tag{54.19}$$

$\delta << \phi_0$,λ= wavelength
Taylor expansion for G_v:

$$G_v(\phi) = G_v(\phi_0) + \frac{\partial G_v}{\partial C}(\phi - \phi_0) + \frac{1}{2}\frac{\partial^2 G_v}{\partial \phi^2}(\phi - \phi_0)^2 \tag{54.20}$$

ΔG = free energy difference associated with composition oscillation:

$$\Delta G_v = \frac{\Delta G}{A\lambda} = \frac{1}{\lambda}\int_0^\lambda G_v(\phi)dz - G_v(\phi_0) \tag{54.21}$$

Combine (1) - (4) and integrate over one wavelength:

$$\Delta G_v = \frac{\partial G_v}{\partial \phi}\frac{\delta}{\lambda}\underbrace{\int_0^\lambda sin(2\pi z/\lambda)dz}_{0} + \frac{1}{2}\frac{\partial^2 G_v}{\partial \phi^2}\frac{\delta^2}{\lambda}\underbrace{\int_0^\lambda sin^2(2\pi z/\lambda)dz}_{\lambda/2}$$

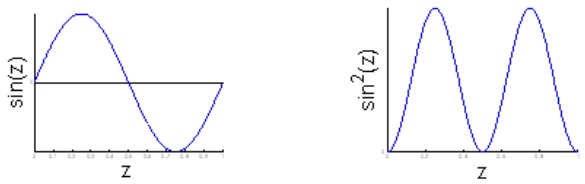


Figure 54.2

$\Delta G_v = \frac{\delta^2}{4}\frac{\partial^2 G_v}{\partial \phi^2}\Big|_{\phi=\phi_0}$, $\Delta G_v < 0$ if $\frac{\partial^2 G_v}{\partial \phi^2} < 0$ (Inside spinodal region of phase diagram)
Prediction: any perturbation is favorable, regardless of λ
Reality: concentration gradients are like interfaces - not energetically favorable
Add gradient term to free energy:

$$\frac{G}{A} = \int \left[G_v(\phi(z)) + \kappa\left(\frac{d\phi}{dz}\right)^2\right]dz \tag{54.22}$$

Local free energy depends on local ϕ **and** local variations in ϕ through nearest neighbor interactions
Contribution to G/A from gradient term:

$$\phi(z) = \phi_0 + \delta \sin(2\pi z/\lambda) \tag{54.23}$$

$$\frac{\partial \phi}{\partial z} = \frac{2\pi\delta}{\lambda}\cos(2\pi z/\lambda) \tag{54.24}$$

$$\left(\frac{\partial \phi}{\partial z}\right)^2 = \delta^2 \left(\frac{2\pi}{\lambda}\right) \cos^2(2\pi z/\lambda) \tag{54.25}$$

$\int_0^\lambda cos^2(2\pi z/\lambda)dz = \lambda/2$
Overall ΔG_v :

$$\Delta G_v = \frac{\delta^2}{2}\left\{\underbrace{\frac{1}{2}\frac{\partial^2 G_v}{\partial \phi^2}}_{<0 \text{ inside spinodal}} + \underbrace{\kappa\left(\frac{2\pi}{\lambda}\right)^2}_{>0 \text{ (large for sharp interfaces - low } \lambda)}\right\}$$

Figure 54.3

$\Delta G_v = 0$ for $\lambda = \lambda_c$ (lowest possible wavelength)

$$\frac{1}{2}\frac{\partial^2 G}{\partial \phi^2} = -\kappa\left(\frac{2\pi}{\lambda_c}\right)^2 \rightarrow \lambda_c = 2\pi\left(-\frac{1}{2\kappa}\frac{\partial^2 G}{\partial \phi^2}\right)^{-\frac{1}{2}} \tag{54.26}$$

$$\lambda > \lambda_c : \Delta G < 0$$

$$\lambda < \lambda_c : \Delta G > 0$$

$\lambda_c \sim 50$ nm in most metallic systems
What about strain energy:
Crystalline systems - strain energy for atoms with different sizes
Recall: $\frac{W_{e\ell}}{V^\beta} = \frac{18\varepsilon^2\mu^\alpha K^\beta}{3K^\beta + 4\mu^\alpha}$
Assume incompressible β: $K^\beta \rightarrow \infty$:
$\frac{W_{e\ell}}{V^\beta} = \frac{18\varepsilon^2\mu^\alpha K^\beta}{3K^\beta} = 6\varepsilon^2\mu$ (assume μ independent of ϕ)
$\varepsilon = \Delta a/a$; $\Delta a = \frac{\partial a}{\partial \phi}\Delta\phi$
Define $\eta \equiv \frac{1}{a}\frac{\partial a}{\partial \phi}$, $\varepsilon = \eta\Delta\phi$

$$\frac{W_{el}}{V} = 6\varepsilon^2(\Delta\phi)^2\mu = 6\eta^2(\phi - \phi_0)^2\mu \tag{54.27}$$

$$\frac{G}{A} = \int \underbrace{\left[G_v(\phi) + \frac{W_{el}}{V}(\phi) \right]}_{\text{volume term modified by elastic contribution}} + \kappa \left(\frac{d\phi}{dz} \right)^2 dz$$

$$\Delta G_v = \frac{\delta^2}{2} \left\{ \underbrace{\frac{1}{2} \frac{\partial^2 G_v}{\partial \phi^2}}_{<0} + \underbrace{6\eta^2 \mu}_{>0} + \underbrace{\kappa \left(\frac{2\pi}{\lambda} \right)^2}_{>0} \right\}$$

Figure 54.4

λ_c is modified by elastic contribution:

$$\lambda_c = \frac{2\pi}{\left(\underbrace{-\frac{1}{2\kappa} \frac{\partial^2 G}{\partial \phi^2}}_{>0} - \frac{6\mu\eta^2}{\kappa} \right)^{1/2}}$$

Figure 54.5

fluctuations with $\lambda < \lambda_c$ cannot grow because they increase the free energy
fluctuations with $\lambda >> \lambda_c$ cannot grow because diffusion is too slow
Overall preferred wavelength: $-2\lambda_c$
strain energy stabilizes homogeneous phase:

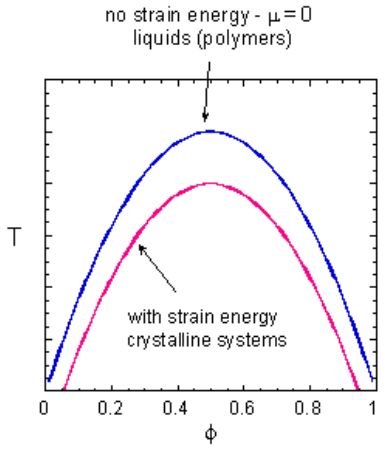


Figure 54.6

(mention recycling codes)
Inside spinodal: $\frac{\partial^2 G}{\partial \phi^2} < 0$, μ_b decreases as ϕ_b increases:

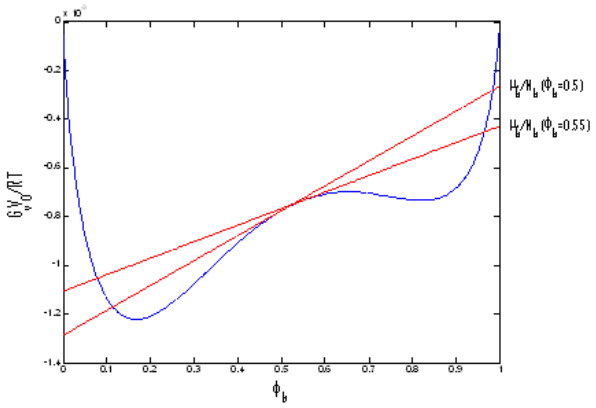


Figure 54.7

Diffusion is always **down** a chemical potential gradient
But this is **up** a concentration gradient when $\frac{\partial^2 G}{\partial \phi^2} < 0$
Fick's first law: $J = -D\frac{\partial C}{\partial x}$: $D < 0$ inside spinodal (uphill diffusion)
Anisoptropic Solids:
Fluctuations grow in the softest direction (lowest shear modulus, μ, lowest λ_c).
Usually <100> directions for cubic metals
(draw resulting microstructure - alternating bands of two phases along preferred crystallographic directions)

Table 54.1

Nucleation and Growth	Spinodal Decomposition
1) Fluctuations large in magnitude but small in number	1) Fluctuations large in number but small in magnitude
2) Barrier to nucleation = W_R^*	2) No nucleation barrier
3) Critical radius = r^* $r^* \propto \frac{1}{\Delta T}$ or $\frac{1}{\Delta C}$ (ΔC = supersaturation	3) Critical wavelength = λ_c

55 Review Questions

- What controls the size of the depletion zone in front of a flat or curved precipitate that is growing?
  - How does it evolve with time?
- What limits the growth velocity of a precipitate phase boundary at high and low temperatures?
  - Which of these limits are connected to the phase diagram?
- Why are flat, plate-like precipitates sometimes formed?
- How does the molar free energy depend on the radius of curvature of a precipitate?
- How does curvature effect the equilibrium concentration of solute that is in equilibrium with a precipitate?
- What does the concentration dependence look like for precipitates that are larger than r*?
  - What if the precipitate is smaller than r*?
- What are transition phases, and why do they form?
  - What is the mechanism by which transition phases shrink at the expense of equilibrium phases?
- How are TTT curves for transition phases related to the phase diagram?
- What are the basic physical assumptions of the Lifshitz/Slyosov coarsening theory discussed in class?
  - What do the depletion zones look like?
  - What determines the average solute concentration in the matrix phase?
- What does the distribution of precipitates look like if coarsening occurs by the Lifshitz/Slyosov mechanism?
- What do the binodal and spinodal curves look like for the regular solution model?
  - What is the critical temperature?
- What determines the size of the characteristic phase size when phase separation occurs by spinodal decomposition?
- What is meant by uphill diffusion?
  - When is it observed?
- How is this phase size modified (in qualitative terms) by coherent strains?
- How do these strains modify the phase diagram to give coherent spinodal and binodal curves?
- How can the liquidus lines be estimated for an ideal eutectic system?
  - What are the assumptions made in the approximation?

- What determines the size of the individual phases for eutectic solidification?
- What determines the size of the individual phases for a eutectoid transformation?
- What is the physical significance of the squared gradient term in the free energy expression?
- How can the shapes and sizes of metallic nanoparticles be controlled?

56 Case study: Damascus Steel

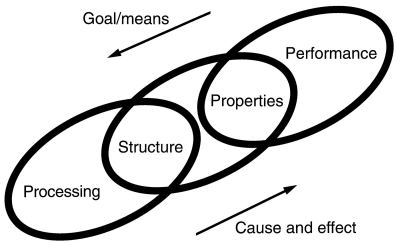


Figure 56.1: Materials Science Paradigm

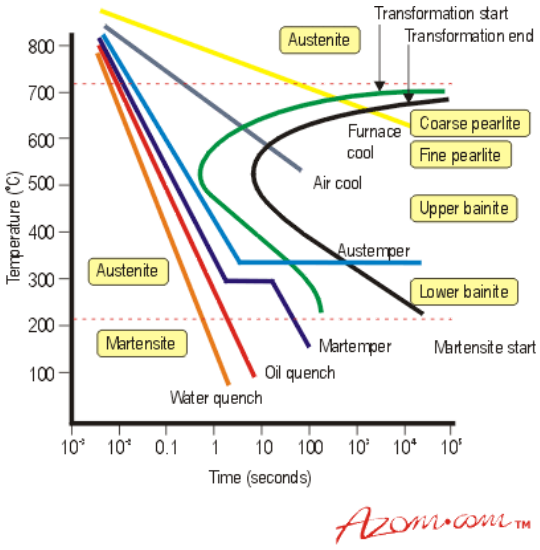


Figure 56.2: TTT diagram for steel.

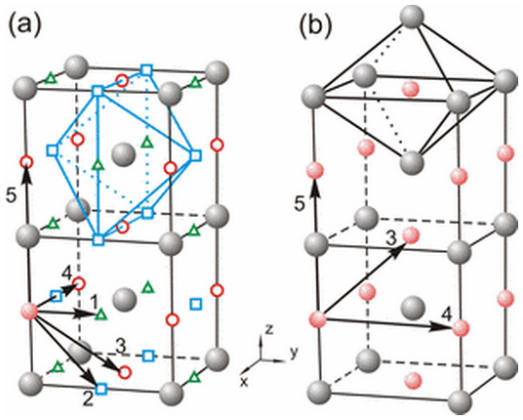


Figure 56.3: Structure of martensite.

Figure 56.4: Steelmaking in Damascus, c. 1900 (http://en.wikipedia.org/wiki/Damascus_steel).

Figure 56.5: A Damascus blade.

Figure 56.6: Damascus blade: Closeup

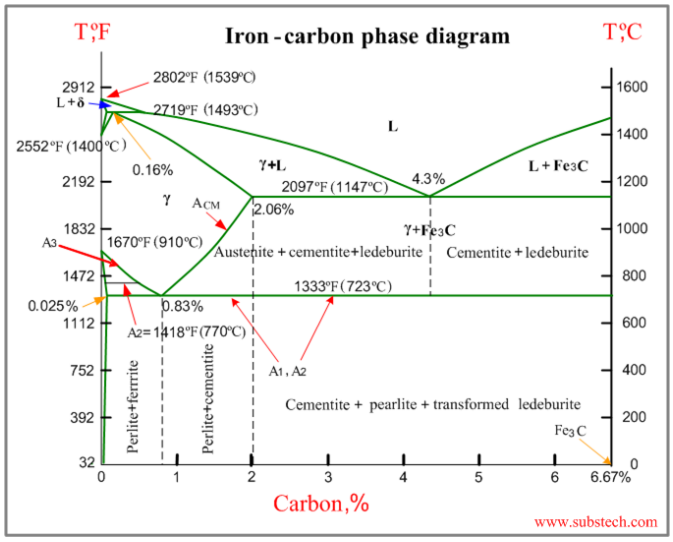


Figure 56.7: Fe-C phase diagram.

57 Matlab Example Scripts

58 Lab Examples

In this section we take a closer look at some of the data from the laboratories.

58.1 Coarsening

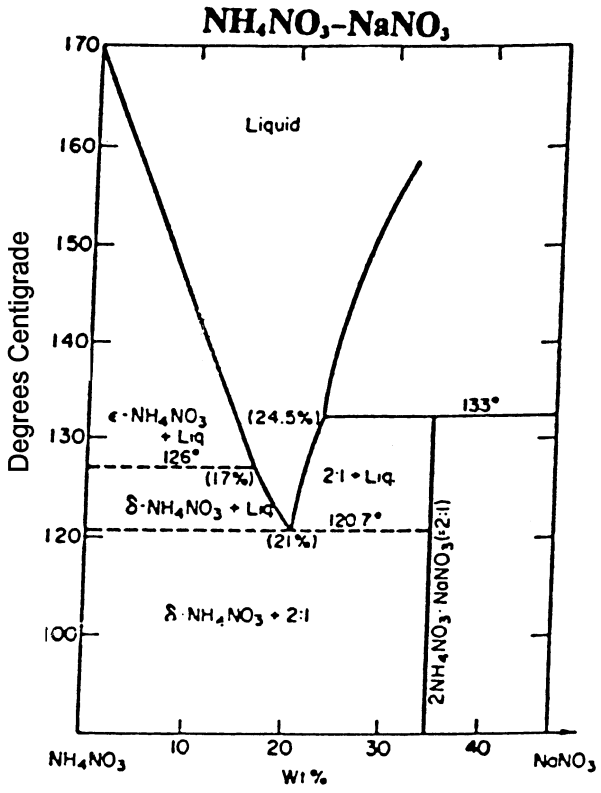


Figure 58.1: Phase diagram for the $NH_4NO_3/NaNO_3$ system.

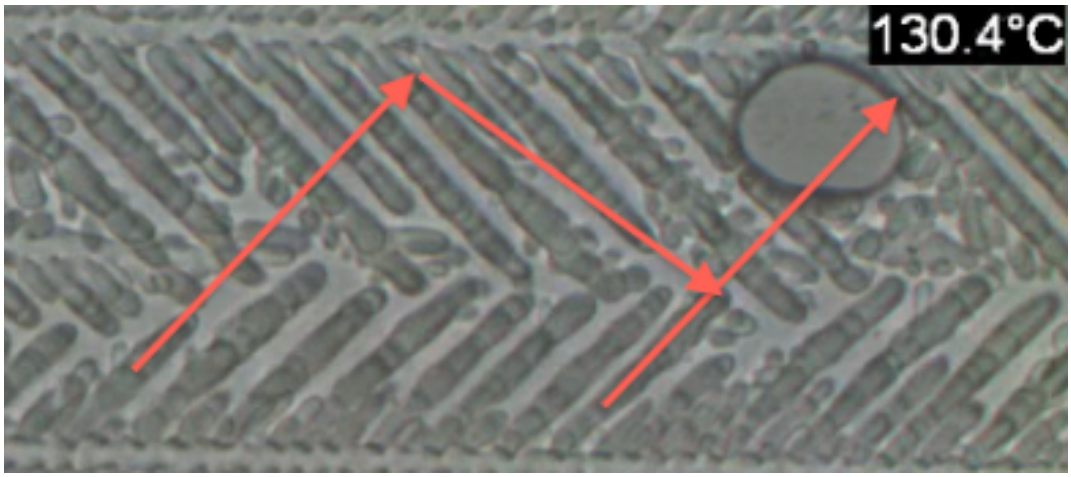


Figure 58.2: Observed Dendritic Morphology

Note: λ_1 (primary arm spacing) $\approx 10\lambda_2$ (secondary arm spacing)

58.2 Secondary Arm Spacing

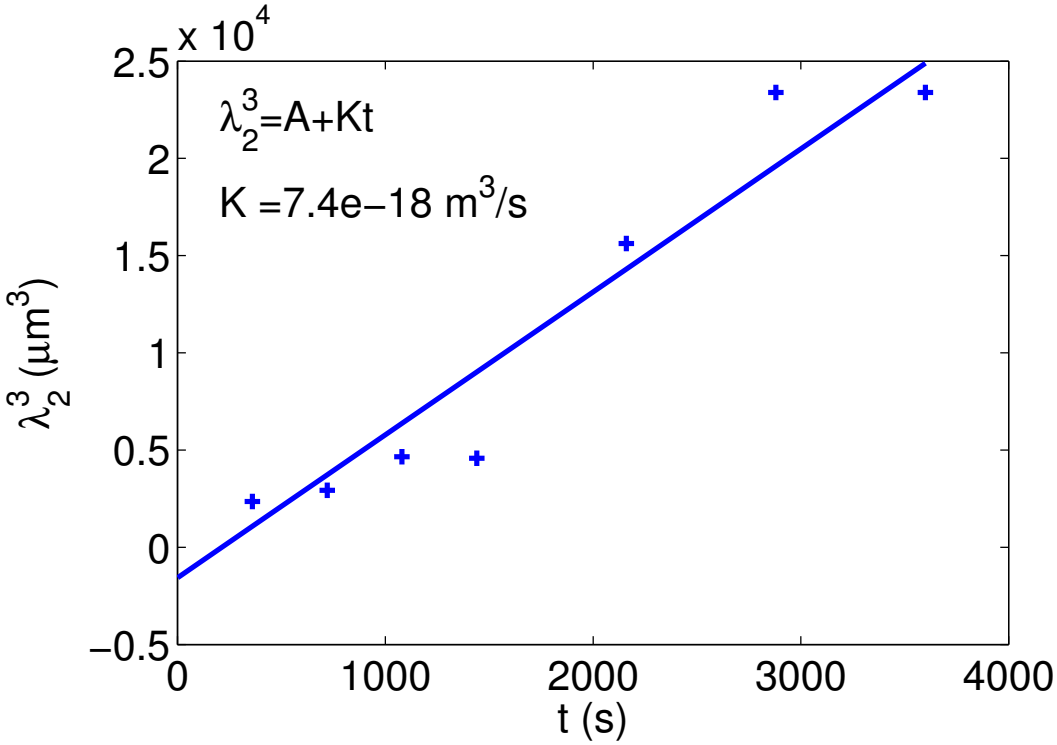


Figure 58.3: Measured Values of the Secondary Arm Spacing

58.3 Surface to Volume Ratio

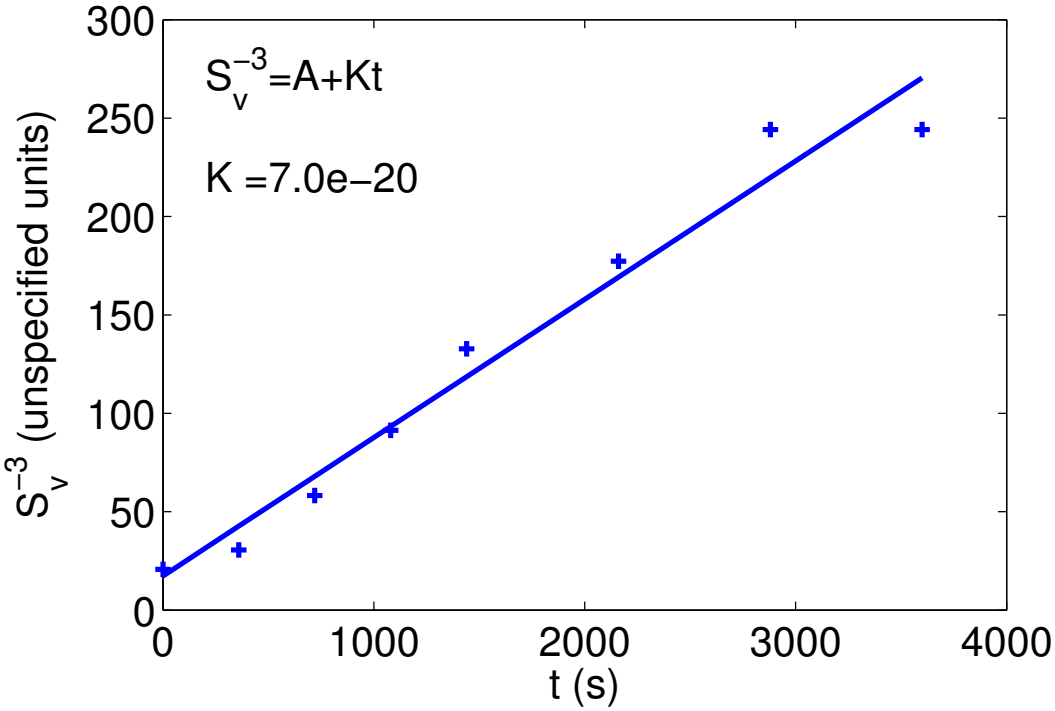


Figure 58.4: Measured surface to volume ratio (units missing)

58.4 Temperature Dependence: K

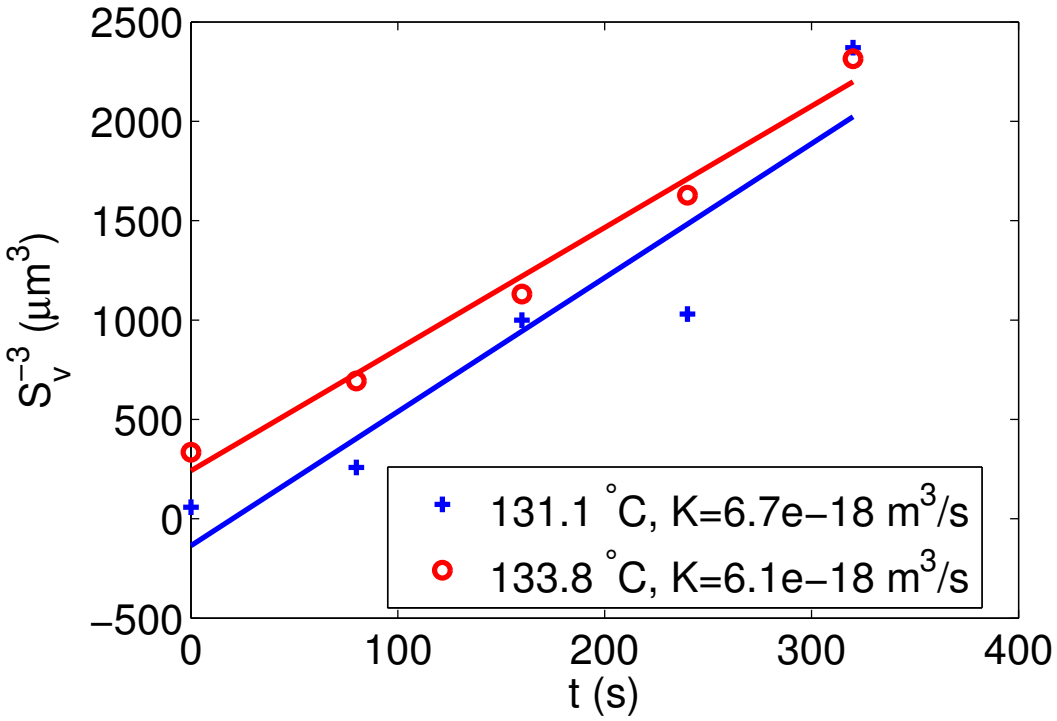


Figure 58.5: Temperature Dependence of the Coarsening Constant

```
%some initialization stuff
clear all
close all
set(0, 'DefaultAxesFontSize', 16);
set(0, 'DefaultTextFontSize', 16);
set(0, 'DefaultLineLineWidth', 2);
set(0, 'DefaultFigurePaperposition',[0 0 6 4])
set(0, 'DefaultFigurePaperSize',[6 4])

%input the data
t=60*[0 6 12 18 24 36 48 60];
l=[7.94 13.3 14.3 16.7 16.6 25 28.6 28.6];
sv=[0.364 0.32 0.258 0.222 0.196 0.178 0.16 0.16];

% create the first plot
plot(t, l.^3, '+b');
hold on
spacingfit=polyfit(t, l.^3, 1); % generates a linear curve fit
plot(t,polyval(spacingfit,t),'b-'); % plots the fit
hold off
xlabel('t (s)')
ylabel('\lambda_{2}^{3} (\mum^{3})')
text(200,2.2e4,'\lambda_{2}^{3}=A+Kt')
ktext=num2str(1e-18*spacingfit(1),'%5.1e');
text(200,1.8e4,['K =' ktext ' m^3/s'])
print(gcf,'-depsc2','../316-2_figures/lab_dendrite_spacing.eps')

% now make the sv plot
figure % creates a new figure window
plot(t, sv.^-3, 'b+');
hold on
svfit=polyfit(t, sv.^(-3),1);
plot(t,polyval(svfit,t))
```

```
xlabel('t (s)')
ylabel('S_{v}^{-3} (unspecified units)')
text(200,270,'S_{v}^{-3}=A+Kt')
ktext=num2str(1e-18*svfit(1),'%5.1e');
text(200,220,['K =' ktext])
print(gcf,'-depsc2','../316-2_figures/lab_dendrite_svplot.eps')

% now include another dataset (from Monday PM group)
figure
t=[0 80 160 240 320]; % time in seconds
sv1311=[0.258 0.157 0.10 0.099 0.075];
sv1338=[0.144 0.113 0.096 0.085 0.0756];
plot(t, sv1311.^-3, 'b+', t, sv1338.^-3, 'ro');
svfit1=polyfit(t, sv1311.^(-3),1);
svfit2=polyfit(t, sv1338.^(-3),1);
hold on
plot(t,polyval(svfit1,t),'b-', t, polyval(svfit2,t),'r-')
xlabel('t (s)')
ylabel('S_{v}^{-3} (\mum^{3})')
% now generate legend info
ktext1=num2str(1e-18*svfit1(1),'%5.1e');
ktext2=num2str(1e-18*svfit2(1),'%5.1e');
legendtext{1}=['131.1 ^{\circ}C, K=' ktext1 ' m^{3}/s'];
legendtext{2}=['133.8 ^{\circ}C, K=' ktext2 ' m^{3}/s'];
legend(legendtext,'location','best')
print(gcf,'-depsc2','../316-2_figures/lab_dendrite_svplotv2.eps')
```

58.5 LSW Theory

$$K = \frac{8\gamma V_m D C_e}{9RT(C_\beta - C_e)}$$

- Does this work (at least approximately)?
- What does D correspond to, and how can we estimate it?
- What about $\frac{C_e}{C_\beta - C_e}$?
- Do we really care about the factor of 8/9?

58.6 Estimating D

$$\lambda_1 \approx \frac{D}{V}$$

$$D \approx V\lambda_1 \approx 10\lambda_2 V$$

T (°C)	V (μm/s)	λ_2 (μm)	$10\lambda_2 V$ (m^{2}s)
130.4	150	9.3	1.4x10^{-9}
130.9	130	11	1.4x10^{-9}
133.6	86	13	1.1x10^{-9}
135	45	16	0.7x10^{-9}

58.7 Estimating the Expected Coarsening Constant

$$K \approx \frac{\gamma V_m D}{RT}$$

$$V_m = \left(\frac{80\,\text{g}}{\text{mol}}\right)\left(\frac{\text{cm}^3}{1.7\,\text{g}}\right) = 47\,\text{cm}^3/\text{mol} = 4.7\text{x}10^{-5}\,\text{m}^3/\text{mol}$$

$$D \approx 10^{-9}\,\text{m}^2/\text{s}$$

$$\gamma \approx 0.1\,\mathrm{J/m^2}$$

$$K \approx \frac{(0.1\,\mathrm{J/m^2})(4.7x10^{-5}\,\mathrm{m^3/mo\ell})(10^{-9}\,\mathrm{m^2/s})}{(8.314\,\mathrm{J.mo\ell{-}K})(400\,\mathrm{K})} = 1.4x10^{-18}\,\mathrm{m^3/s}$$

- Not bad, given all the approximations that we have made.

58.8 Heat treatment of Al alloys

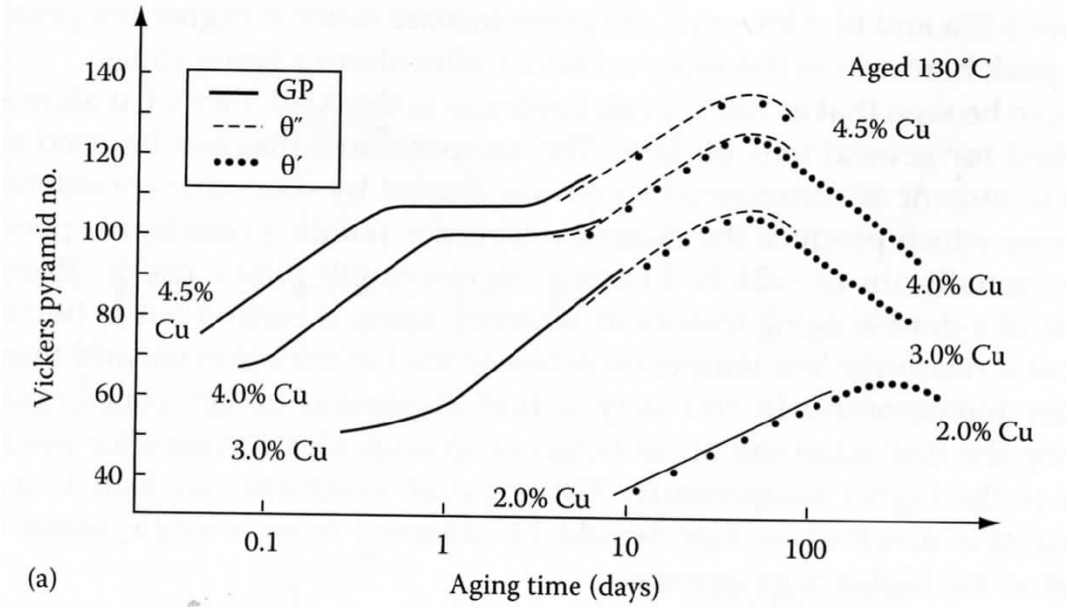


Figure 58.6: Heat Treatment of Al Alloys (from P&E)

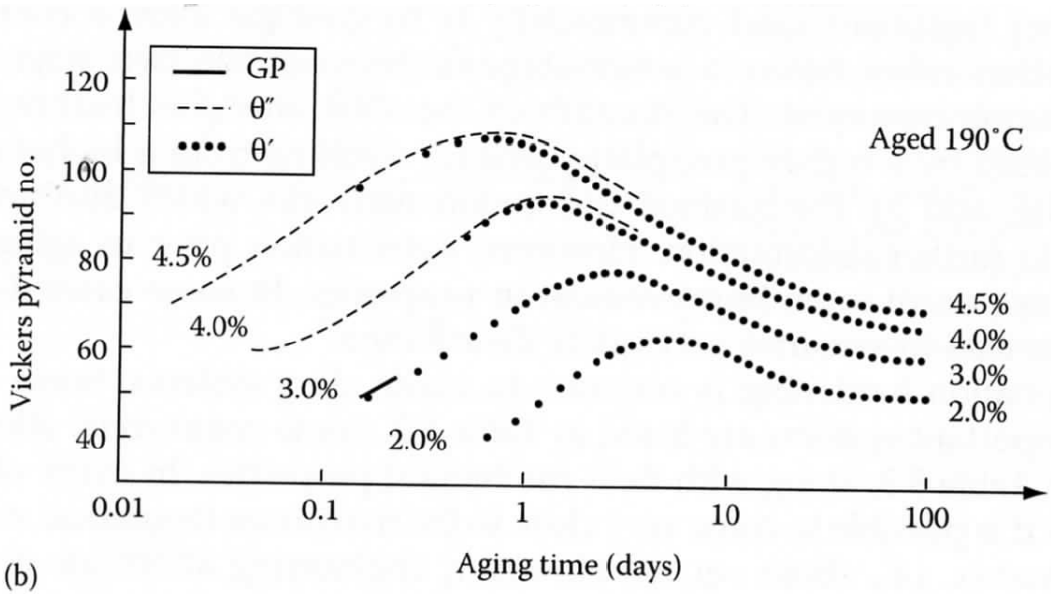


Figure 58.7: (P&E 5.5.4)

59 Appendix: Thermodynamic Data for Cu

```
%% Phase transformations in elemental copper
% This script uses NIST thermochemical data to
%       - create plots of thermodynamic functions of state
%       - calculate phase transformation temperatures, and the changes in
%       standard enthalpy, entropy, and Gibbs free energy associated with
%       the phase transformation
%       - compare the change in Gibbs free energy calculated from NIST data
%       to the linearized Gibbs free energy estimate near the phase
%       transformation temperature.

close all;  clear all; filename = 'Cu_Shomate_';
%% NIST Data
% The NIST Chemistry webbook (URL: http://webbook.nist.gov/chemistry/)
% gives Shomate coefficients that can be used to calculate the heat
% capacity, standard enthalpy, absolute entropy, and standard Gibbs free
% energy as a function of temperature. The Shomate coefficients are define
% only over a certain range of T. Extrapolation beyond this range is
% dangerous.

% NIST Chemistry webbok data for iron can be found here:
% http://webbook.nist.gov/cgi/cbook.cgi?Name=copper&Units=SI
% This is actually a fairly straightforward case, as there only two phase
% transformations: solid (alpha, hcp) -> L and L -> gas

Cu.phase={'s','L','g'};
Cu.Tmin = [298, 1358, 2843.261];
Cu.Tmax = [1358, 2843, 6000];

S=[[17.72891,  32.84450,   -80.48635];
   [28.09870,  -0.000084,   49.35865];
   [-31.25289,  0.000032,   -7.578061];
   [13.97243,  -0.000004,    0.404960];
   [0.068611,  -0.000028,  133.3382];
   [-6.056591, -1.804901,  519.9331];
   [47.89592,  73.92310,   193.5351];
   [0,         11.85730,   337.6003];];

%% Calculating Thermodynamic Functions of State from Shomate Coefficients
% We use here anonymous functions to calculate the thermodynamic functions
% of state. Input arguments are a column vector of exactly 8 Shomate
% coefficients (A-H) and a row vector with temperature values where t = T/1000 [K]. Please note
    that H0
% is really H0, not H0-H0_298! Cp0 and S0 are given in units of J mol^-1 K^-1, H0
% and G0 in units of kJ mol^-1.

Cp0 = @(Shomate,t) Shomate*[0*t+1 ; t     ; t.^2  ; t.^3  ;  t.^-2    ; 0*t  ; 0*t  ; 0*t]; %[J
    mol^-1 K^-1]

H0  = @(Shomate,t) Shomate*[t     ; t.^2/2; t.^3/3; t.^4/4; -t.^-1    ; 0*t+1; 0*t  ; 0*t]; %[
  kJ mol^-1]

S0  = @(Shomate,t) Shomate*[log(t); t     ; t.^2/2; t.^3/3; -0.5*t.^-2; 0*t  ; 0*t+1; 0*t]; %[J
    mol^-1 K^-1]

G0 = @(Shomate, T) H0(Shomate, T/1000)-T/1000.*S0(Shomate, T/1000); % [kJ mol^-1]

%% Calculate phase transformation temperatures
% We define the Gibbs free energy change for four different
% transformations in the system as anonymous functions. To solve for the
% phase transformation tempreature at which the Gibbs free energy change is
% zero, we find the root of the anonymous functions using a guess based on
% the phase transformation temperatures given in phase diagrams.

% vaporization: L->gas
DeltaG_Lg= @(T) G0(S(:,3)',T)-G0(S(:,2)',T);
T_Lg = fzero(DeltaG_Lg,3130);

% melting/fusion: s->L
```

```
DeltaG_sL= @(T) G0(S(:,2)',T)-G0(S(:,1)',T);
T_sL = fzero(DeltaG_sL,1800);


%% Plot Functions of State

% T-ranges
T1=linspace(298,T_sL,1000);
T2=linspace(T_sL,T_Lg,1000);
T3=linspace(T_Lg,6000,1000);

h1=figure;
set(gcf,'DefaultLineLineWidth',1.5)

c=colormap(lines(3));
plot(T1,Cp0(S(:,1)',T1/1000),'Color',c(1,:));hold on;
plot(T2,Cp0(S(:,2)',T2/1000),'Color',c(2,:));hold on;
plot(T3,Cp0(S(:,3)',T3/1000),'Color',c(3,:));hold on;

title('Heat capacity (specific heat) of elemental Cu');
ylabel('C_p^0 [J mol^-^1 K^-^1]');
xlabel('T [K]');
legend(Cu.phase);
set(gca,'TickDir','out');
grid on;

h2=figure;
set(gcf,'DefaultLineLineWidth',1.5)
plot(T1,H0(S(:,1)',T1/1000),'Color',c(1,:));hold on;
plot(T2,H0(S(:,2)',T2/1000),'Color',c(2,:));hold on;
plot(T3,H0(S(:,3)',T3/1000),'Color',c(3,:));hold on;

title('Standard Enthalpy of elemental Cu');
ylabel('H^0-H_2_9_8 [kJ mol^-^1]');
xlabel('T [K]');
legend(Cu.phase);
set(gca,'TickDir','out');
grid on;

h3=figure;
set(gcf,'DefaultLineLineWidth',1.5)
plot(T1,S0(S(:,1)',T1/1000),'Color',c(1,:));hold on;
plot(T2,S0(S(:,2)',T2/1000),'Color',c(2,:));hold on;
plot(T3,S0(S(:,3)',T3/1000),'Color',c(3,:));hold on;

title('Absolute Entropy of elemental Cu');
ylabel('S^0 [J mol^-^1 K^-^1]');
xlabel('T [K]');
legend(Cu.phase);
set(gca,'TickDir','out');
grid on;

h4=figure;
set(gcf,'DefaultLineLineWidth',1.5)
plot(T1,G0(S(:,1)',T1),'Color',c(1,:));hold on;
plot(T2,G0(S(:,2)',T2),'Color',c(2,:));hold on;
plot(T3,G0(S(:,3)',T3),'Color',c(3,:));hold on;

title('Gibbs Free Energy of elemental Cu');
ylabel('G^0 [kJ mol^-^1]');
xlabel('T [K]');
legend(Cu.phase);
set(gca,'TickDir','out');
grid on;


%% Plot Close Ups of Phase Transformations
h5=figure;
set(gcf,'DefaultLineLineWidth',1.5)
% solid-liquid transition
DT = 2;
```

```
T_tr = T_sL;
p1=1;
p2=2;

Tr_left = (T_tr-DT):0.1:T_tr;
Tr_right = T_tr:0.1:(T_tr+DT);
Tr = (T_tr-DT):0.1:(T_tr+DT);

subplot(4,2,1);
ycent=mean([H0(S(:,p1)',Tr_left(end)/1000), H0(S(:,p2)',Tr_left(end)/1000)]);
yrange=abs(H0(S(:,p1)',Tr_left(end)/1000)-ycent);
DeltaH0(1) = H0(S(:,p2)',T_tr/1000)-H0(S(:,p1)',T_tr/1000);

plot(Tr_left,H0(S(:,p1)',Tr_left/1000),'Color',c(p1,:));hold on;
plot(Tr_right,H0(S(:,p1)',Tr_right/1000),'Color',c(p1,:),'LineStyle',':');
plot(Tr_left,H0(S(:,p2)',Tr_left/1000),'Color',c(p2,:),'LineStyle',':');hold on;
plot(Tr_right,H0(S(:,p2)',Tr_right/1000),'Color',c(p2,:));
plot([T_tr T_tr],[0 600],'k:');
plot([T_tr T_tr],[H0(S(:,p1)',Tr_left(end)/1000), H0(S(:,p2)',Tr_left(end)/1000)],'m','
  LineWidth',2);
xlim([Tr(1) Tr(end)]);
ylim(ycent+[-yrange yrange]*1.5);
ylabel('H^0-H^0_2_9_8 [kJ/mol]');
xlabel('T [K]');
set(gca,'TickDir','out');

subplot(4,2,3);
ycent=mean([S0(S(:,p1)',Tr_left(end)/1000), S0(S(:,p2)',Tr_left(end)/1000)]);
yrange=abs(S0(S(:,p1)',Tr_left(end)/1000)-ycent);
DeltaS0(1) = S0(S(:,p2)',T_tr/1000)-S0(S(:,p1)',T_tr/1000);

plot(Tr_left,S0(S(:,p1)',Tr_left/1000),'Color',c(p1,:));hold on;
plot(Tr_right,S0(S(:,p1)',Tr_right/1000),'Color',c(p1,:),'LineStyle',':');
plot(Tr_left,S0(S(:,p2)',Tr_left/1000),'Color',c(p2,:),'LineStyle',':');hold on;
plot(Tr_right,S0(S(:,p2)',Tr_right/1000),'Color',c(p2,:));
plot([T_tr T_tr],[0 600],'k:');
plot([T_tr T_tr],[S0(S(:,p1)',Tr_left(end)/1000), S0(S(:,p2)',Tr_left(end)/1000)],'m','
  LineWidth',2);
xlim([Tr(1) Tr(end)]);
ylim(ycent+[-yrange yrange]*1.5);
ylabel('S^0 [J/(mol K)]');
xlabel('T [K]');
set(gca,'TickDir','out');

subplot(4,2,5);
plot(Tr_left,G0(S(:,p1)',Tr_left),'Color',c(p1,:));hold on;
plot(Tr_right,G0(S(:,p1)',Tr_right),'Color',c(p1,:),'LineStyle',':');
plot(Tr_left,G0(S(:,p2)',Tr_left),'Color',c(p2,:),'LineStyle',':');hold on;
plot(Tr_right,G0(S(:,p2)',Tr_right),'Color',c(p2,:));
plot([T_tr T_tr],[0 -600],'k:');
xlim([Tr(1) Tr(end)]);
ylim(G0(S(:,p1)',Tr_left(end))+[-.2 .2]);
ylabel('G^0 [kJ/mol]');
xlabel('T [K]');
set(gca,'TickDir','out');

subplot(4,2,7);
y1=G0(S(:,p2)',Tr(1))-G0(S(:,p1)',Tr(1));
y2=G0(S(:,p2)',Tr(end))-G0(S(:,p1)',Tr(end));

plot(Tr,G0(S(:,p2)',Tr)-G0(S(:,p1)',Tr),'b');hold on; % Note: could be replaced with DeltaG
  function
plot(Tr,-DeltaS0(1)/1000*(Tr-T_tr),'r:');

plot([Tr(1) Tr(end)],[0,0],'k:');
plot([T_tr T_tr],[y2 y1],'k:');
xlim([Tr(1) Tr(end)]);
ylim([y2 y1]);

ylabel('\DeltaG [kJ/mol]');
xlabel('T [K]');
```

```
204 set(gca,'TickDir','out');
205
206 % liguid-gas transition
207 DT = 2;
208 T_tr = T_Lg;
209 p1=2;
210 p2=3;
211
212 Tr_left = (T_tr-DT):0.1:T_tr;
213 Tr_right = T_tr:0.1:(T_tr+DT);
214 Tr = (T_tr-DT):0.1:(T_tr+DT);
215
216 subplot(4,2,2);
217 ycent=mean([H0(S(:,p1)',Tr_left(end)/1000), H0(S(:,p2)',Tr_left(end)/1000)]);
218 yrange=abs(H0(S(:,p1)',Tr_left(end)/1000)-ycent);
219 DeltaH0(2) = H0(S(:,p2)',T_tr/1000)-H0(S(:,p1)',T_tr/1000);
220
221 plot(Tr_left,H0(S(:,p1)',Tr_left/1000),'Color',c(p1,:));hold on;
222 plot(Tr_right,H0(S(:,p1)',Tr_right/1000),'Color',c(p1,:),'LineStyle',':');
223 plot(Tr_left,H0(S(:,p2)',Tr_left/1000),'Color',c(p2,:),'LineStyle',':');hold on;
224 plot(Tr_right,H0(S(:,p2)',Tr_right/1000),'Color',c(p2,:));
225 plot([T_tr T_tr],[0 600],'k:');
226 plot([T_tr T_tr],[H0(S(:,p1)',Tr_left(end)/1000), H0(S(:,p2)',Tr_left(end)/1000)],'m','
    LineWidth',2);
227 xlim([Tr(1) Tr(end)]);
228 ylim(ycent+[-yrange yrange]*1.5);
229 ylabel('H^0-H^0_2_9_8 [kJ/mol]');
230 xlabel('T [K]');
231 set(gca,'TickDir','out');
232
233 subplot(4,2,4);
234 ycent=mean([S0(S(:,p1)',Tr_left(end)/1000), S0(S(:,p2)',Tr_left(end)/1000)]);
235 yrange=abs(S0(S(:,p1)',Tr_left(end)/1000)-ycent);
236 DeltaS0(2) = S0(S(:,p2)',T_tr/1000)-S0(S(:,p1)',T_tr/1000);
237
238 plot(Tr_left,S0(S(:,p1)',Tr_left/1000),'Color',c(p1,:));hold on;
239 plot(Tr_right,S0(S(:,p1)',Tr_right/1000),'Color',c(p1,:),'LineStyle',':');
240 plot(Tr_left,S0(S(:,p2)',Tr_left/1000),'Color',c(p2,:),'LineStyle',':');hold on;
241 plot(Tr_right,S0(S(:,p2)',Tr_right/1000),'Color',c(p2,:));
242 plot([T_tr T_tr],[0 600],'k:');
243 plot([T_tr T_tr],[S0(S(:,p1)',Tr_left(end)/1000), S0(S(:,p2)',Tr_left(end)/1000)],'m','
    LineWidth',2);
244 xlim([Tr(1) Tr(end)]);
245 ylim(ycent+[-yrange yrange]*1.5);
246 ylabel('S^0 [J/(mol K)]');
247 xlabel('T [K]');
248 set(gca,'TickDir','out');
249
250 subplot(4,2,6);
251 plot(Tr_left,G0(S(:,p1)',Tr_left),'Color',c(p1,:));hold on;
252 plot(Tr_right,G0(S(:,p1)',Tr_right),'Color',c(p1,:),'LineStyle',':');
253 plot(Tr_left,G0(S(:,p2)',Tr_left),'Color',c(p2,:),'LineStyle',':');hold on;
254 plot(Tr_right,G0(S(:,p2)',Tr_right),'Color',c(p2,:));
255 plot([T_tr T_tr],[0 -600],'k:');
256 xlim([Tr(1) Tr(end)]);
257 ylim(G0(S(:,p1)',Tr_left(end))+[-.2 .2]);
258 ylabel('G^0 [kJ/mol]');
259 xlabel('T [K]');
260 set(gca,'TickDir','out');
261
262 subplot(4,2,8);
263 y1=G0(S(:,p2)',Tr(1))-G0(S(:,p1)',Tr(1));
264 y2=G0(S(:,p2)',Tr(end))-G0(S(:,p1)',Tr(end));
265
266 plot(Tr,G0(S(:,p2)',Tr)-G0(S(:,p1)',Tr),'b');hold on;
267 plot(Tr,-DeltaS0(2)/1000*(Tr-T_tr),'r:');
268
269 plot([Tr(1) Tr(end)],[0,0],'k:');
270 plot([T_tr T_tr],[y2 y1],'k:');
271 xlim([Tr(1) Tr(end)]);
272 ylim([y2 y1]);
```

```
273
274 ylabel('\DeltaG [kJ/mol]');
275 xlabel('T [K]');
276 set(gca,'TickDir','out');
277
278 %% Output data
279 clc;
280 fprintf('Phase transformation:\n');
281 fprintf('s->L:\n');
282 fprintf('T_tr    = %4.2f K.\n',T_sL);
283 fprintf('DeltaH0 = %4.2f kJ mol^-1.\n',DeltaH0(1));
284 fprintf('DeltaS0 = %4.2f J mol^-1 K^-1.\n\n',DeltaS0(1));
285
286 fprintf('L->gas:\n');
287 fprintf('T_tr    = %4.2f K.\n',T_Lg);
288 fprintf('DeltaH0 = %4.2f kJ mol^-1.\n',DeltaH0(2));
289 fprintf('DeltaS0 = %4.2f J mol^-1 K^-1.\n\n',DeltaS0(2));
290
291 %% How good is the linear approximation of DeltaG near T_tr?
292 % careful here: extrapolation beyond the range over which Shomate
293 % coefficients are defined may not represent reality very well, or at all.
294 h6=figure;
295 set(gcf,'DefaultLineLineWidth',1.5)
296 DT = 50;
297
298 T_tr = T_Lg;
299 p1=2;
300 p2=3;
301 i=2;
302 Tr = (T_tr-DT):0.1:(T_tr+DT);
303
304 subplot(223);
305 y1=G0(S(:,p2)',Tr(1))-G0(S(:,p1)',Tr(1));
306 y2=G0(S(:,p2)',Tr(end))-G0(S(:,p1)',Tr(end));
307
308 plot(Tr,G0(S(:,p2)',Tr)-G0(S(:,p1)',Tr),'b');hold on;
309 plot(Tr,-DeltaS0(i)/1000*(Tr-T_tr),'r:');
310 plot([Tr(1) Tr(end)],[0,0],'k:');
311 plot([T_tr T_tr],[y2 y1],'k:');
312 xlim([Tr(1) Tr(end)]);
313 ylim([y2 y1]);
314
315 ylabel('\DeltaG [kJ/mol]');
316 xlabel('T [K]');
317 set(gca,'TickDir','out');grid on;
318
319 subplot(224);
320 plot(Tr,100*(G0(S(:,p2)',Tr)-G0(S(:,p1)',Tr)+DeltaS0(i)/1000*(Tr-T_tr))./(-DeltaS0(i)/1000*(Tr-
    T_tr)),'b');hold on;
321 grid on;
322 %plot([T_tr T_tr],[y2 y1],'k:');
323 xlim([Tr(1) Tr(end)]);
324 %ylim([y2 y1]);
325 ylabel('(\DeltaG-\DeltaG_e_s_t)/\DeltaG_e_s_t [%]');
326 xlabel('T [K]');
327 set(gca,'TickDir','out');
328
329 %
330 T_tr = T_sL;
331 p1=1;
332 p2=2;
333 i=1;
334 Tr = (T_tr-DT):0.1:(T_tr+DT);
335 subplot(221);
336 y1=G0(S(:,p2)',Tr(1))-G0(S(:,p1)',Tr(1));
337 y2=G0(S(:,p2)',Tr(end))-G0(S(:,p1)',Tr(end));
338
339 plot(Tr,G0(S(:,p2)',Tr)-G0(S(:,p1)',Tr),'b');hold on;
340 plot(Tr,-DeltaS0(i)/1000*(Tr-T_tr),'r:');
341 plot([Tr(1) Tr(end)],[0,0],'k:');
342 plot([T_tr T_tr],[y2 y1],'k:');
```

```
xlim([Tr(1) Tr(end)]);
ylim([y2 y1]);

ylabel('\DeltaG [kJ/mol]');
xlabel('T [K]');
set(gca,'TickDir','out');grid on;

subplot(222);
plot(Tr,100*(G0(S(:,p2)',Tr)-G0(S(:,p1)',Tr)+DeltaS0(i)/1000*(Tr-T_tr))./(-DeltaS0(i)/1000*(Tr-
  T_tr)),'b');hold on;
grid on;
%plot([T_tr T_tr],[y2 y1],'k:');
xlim([Tr(1) Tr(end)]);
%ylim([y2 y1]);
ylabel('(\DeltaG-\DeltaG_e_s_t)/\DeltaG_e_s_t [%]');
xlabel('T [K]');
set(gca,'TickDir','out');

%% save figures h1=h6 as .eps files

print(h1,[filename,'Cp'],'-deps','-cmyk','-opengl');
print(h2,[filename,'S'],'-deps','-cmyk','-opengl');
print(h3,[filename,'H'],'-deps','-cmyk','-opengl');
print(h4,[filename,'G'],'-deps','-cmyk','-opengl');
print(h5,[filename,'PTs'],'-deps','-cmyk','-opengl');
print(h6,[filename,'Linearity'],'-deps','-cmyk','-opengl');
```

60 316-2 Problems

60.1 General

(1) Write a paragraph discussing the relevance of phase transformations in your daily life.

60.2 Laplace Pressure Derivation

(2) Derive the expression for the Laplace pressure inside a long cylinder of radius R.

60.3 Homogeneous Nucleation

(3) Consider the following data for nickel:

Melting point	1452 °C
Molar entropy of solid at T_m	56.07 J/K
Molar entropy of liquid at T_m	66.27 J/K
Solid density	8.9 g/cm^3
Molar mass	58.7

In their classic experiment Turnbull and Cech studied the undercooling of small droplets for a number of different metals . Assuming that nucleation in the droplets occurs homogeneously and using the data given below calculate the following at 1100 °C and 1200 °C:

a The molar volume of nickel.

b The work of nucleation (W_R^*).

c The dimensionless ratio, W_R^*/k_BT.

d The radius of the critical nucleus.

e The pressure of the critical nucleus in pascals (assume the surrounding liquid is at atmospheric pressure).

f The molar enthalpy of melting at T_m.

g Suppose a Ni droplet with a volume of about 100 μm^3 is solidified. Approximate the temperature to which the droplet must be cooled in order for solidification to occur by homogeneous nucleation.

(4) Import the file labeled ElementData.mat that includes the required data for various elements on the periodic table into Matlab and:

a Derive the expressions for ΔP, R^*, W_R^*, and W_R^*/k_BT in terms of T_m, ΔT, V_m, ΔS_f, and γ.

b Plot V_m, ΔS_f, γ, ΔP, R^*, W_R^*, and W_R^*/k_BT using $\Delta T = 100K$ versus atomic number (Z) and label all axes including units and each data point with the chemical symbol corresponding to the element. Hint: You should only consider those elements for which the values of γ are included in the ElementData.mat file. For both W_R^* and W_R^*/k_BT plot the y axis on a log scale. Also, in order to label the data points with the chemical symbol you will need to use the text(x, y, 'string') function. You may want to use subplots.

The ElementData.mat file has the following format:

```
ElementData =
Name: {118x1 cell}
Symbol: {118x1 cell}
DeltaH0f: [118x1 double]
Tm: [118x1 double]
Z: [118x1 double]
Aw: [118x1 double]
rho: [118x1 double]
gamma: [118x1 double]
Vm: [118x1 double]
DeltaS0f: [118x1 double]
Structure: {118x1 cell}
Units: {1x8 cell}
```

c Discuss the plots from part (b) with respect to trends in the periodic table, which variables are really important, outliers, and rules of thumb i.e. typical range of values or average value. Does homogeneous nucleation ever really happen?

d Now replot the data for both $\Delta T = 352K$ and $\Delta T = 252K$ and compare the R^* and W_R^* values obtained for Ni to those you calculated in question (3).

(5) Derive expressions for R^* and W_R^* for a cuboidal nucleus.

(6) In the derivations for nucleation in this course we assume that the nucleus is incompressible. Show that this is a valid assumption for solidification of Ni with $\gamma = 2.38 J/m^2$ and $R^* = 1nm$. Hint: Assume that the material is linearly elastic and isotropic. Therefore, you can calculate the bulk modulus using a simple relationship which is a function of Young's modulus and Poisson's ratio. Please cite your source for the values of E and ν that you use.

60.4 Surface and Interface Effects

(7) The surface free energy of solid gold at its melting point (1063ºC) is $1.400 J/m^2$. The surface energy of liquid gold at this temperature is $1.128 J/m^2$, and the interfacial energy for the gold solid/liquid interface is = $0.132 J/m^2$. The latent heat of fusion for gold is $1.2x10^9 J/m^3$.

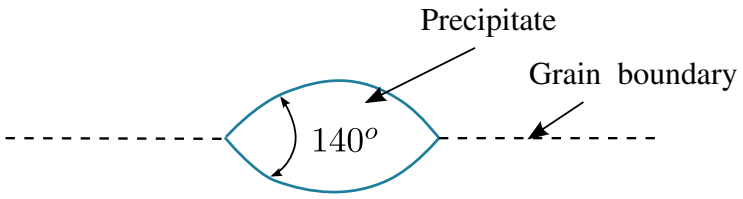
a What is the contact angle for liquid gold on a solid gold surface at 1063ºC ?

b Is there thermodynamic barrier for the melting of a gold surface?

c Suppose a thin liquid gold layer of thickness δ exists at the surface of gold at 1058 °C (5 ° below the equilibrium melting point). By comparing to the free energy of a gold surface that does not have this liquid layer, estimate the maximum thickness of the liquid layer that will be thermodynamically stable at this temperature.

d Very small gold particles have melting points that differ from the melting point of bulk gold. From the analysis given above, do you expect the melting point of a particle with a diameter of 2 nm to be higher or lower than the melting point of bulk gold? Give a brief explanation for your answer.

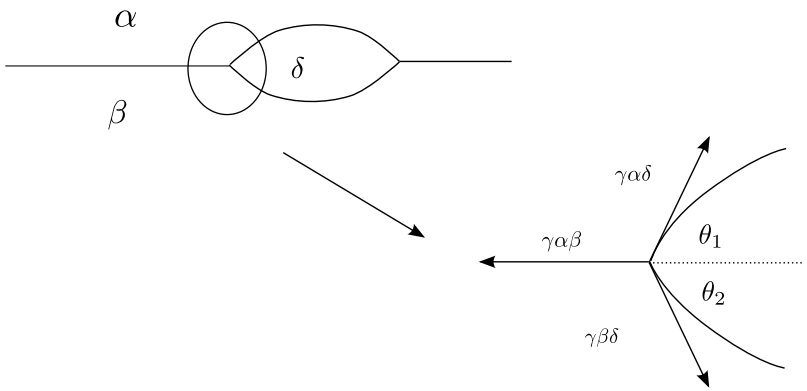
(8) Suppose precipitates form at grain boundaries within the matrix phase, with geometries that look like the following:



What is the ratio of the grain boundary free energy to the interfacial energy between the precipitate and the matrix phase?

(9) Water beads up on a freshly waxed car to form droplets with a contract angle of 80°. What is the interfacial free energy for the wax/water interface, if the surface energy of the wax is $0.025\,\mathrm{J/m^2}$? (Note: you'll need to look up the surface energy of water to do this problem).

(10) An oil droplet (δ phase) is placed on the water surface (phase β) in contact with air (phase α). The schematic of the cross section of the droplet is as describe in class (and repeated below). The surface free energy of water (against air) is $0.072\ J/m^2$. If the measured values of θ_1 and θ_2 in the figure below are 37° and 23°, respectively, what are the values of the oil surface energy and the oil/water interfacial energy.



60.5 Heterogeneous Nucleation

(11) Derive the structure factor, $S(\theta)$.

(12) Suppose that nucleation of a solid, single component metal occurs heterogeneously at a wall. Based on the values given for Ni in problem (3), what contact angle for the critical nucleus must be obtained in order to increase the minimum temperature required for solidification by 50°C?

60.6 Nucleation in a Binary System

(13) 3. Consider the formation of a nucleus β^* with composition $X^{\beta*}$ from metastable α with composition X_0^α.At temperature T, the composition of stable α is X_{eq}^α, that of stable β is X_{eq}^β (all X refer to X_1). In class we derived an expression for the molar Gibbs free energy of formation for the nucleus:

$$\Delta G_m^{\alpha\to\beta*} = G_m^\beta(X^{\beta*}) - G_m^\alpha(X_0^\alpha) - \left.\frac{\partial G^\alpha}{\partial X}\right|_{X_0^\alpha}(X^{\beta*} - X_0^\alpha) \tag{60.1}$$

Show that for $X_0^\alpha - X_{eq}^\alpha \to 0$ and $X^{\beta*} - X_{eq}^\beta \to 0$, Eq. (1) can be rewritten in the following form:

$$\Delta G_m^{\alpha\to\beta*} = -\left.\frac{\delta^2 G^\alpha}{\delta X^2}\right|_{X_0^\alpha}(X_0^\alpha - X_{eq}^\alpha)(X_{eq}^\beta - X_{eq}^\alpha)$$

Hint: Express $G_m^\beta(X^{\beta*})$ in terms of G_m^α. Approximate all terms at non-equilibrium compositions as Taylor expansions around suitable equilibrium values.

(14) In class we used the definition of the misfit parameter for a β nucleus in an α matrix as

$$\varepsilon = \frac{1}{3}\left(\frac{V_m^\beta - V_m^\alpha}{V_m^\alpha}\right)$$

i.e. one third of the volume strain. Show that for cubic systems, the misfit parameter can be approximated as

$$\varepsilon_{cubic} = \frac{a^\beta - a^\alpha}{a^\alpha}$$

where a is the lattice parameter. Hint: Write ΔV in terms of ε_{cubic} and look at the behavior as$\varepsilon_{cubic} \to 0$.

(15) A coherent precipitate nucleates much more easily than does an incoherent particle of the same precipitate. To illustrate this:

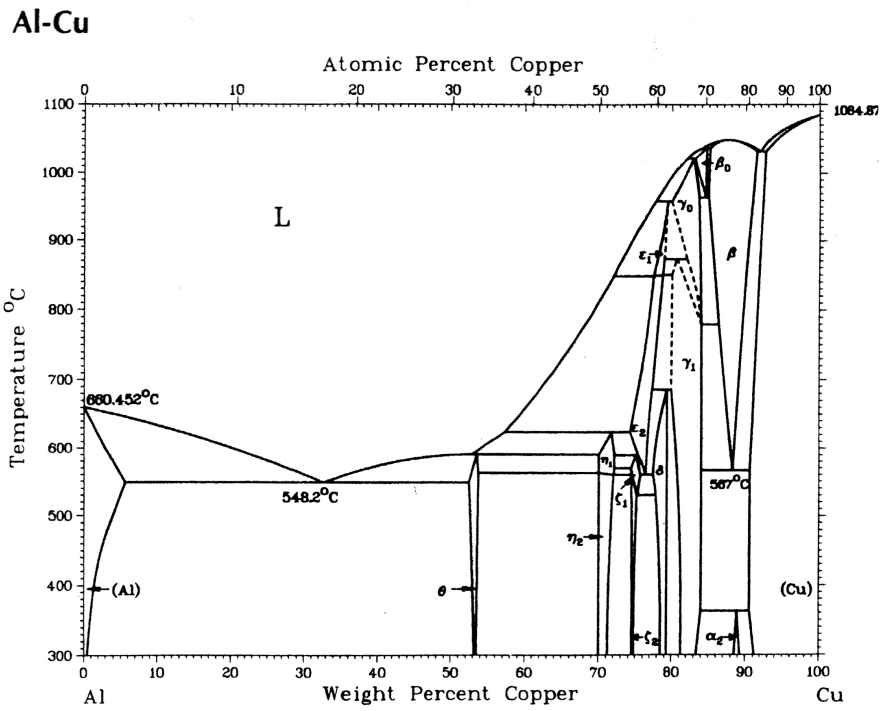
a What is the ratio of W_R^* for the two types of precipitate if $\gamma_{coherent} = 30\,\text{ergs/cm}^2$ and $\gamma_{\text{incoherent}} = 300\,\text{ergs/cm}^2$? Assume that the precipitate is unstrained.

b If the chemical driving force (ΔG_v) is given by$-50\Delta T/T_e\,\text{cal/cm}^3$, $T_e = 1000\,\text{K}$, the misfit strain is 0.001 for the coherent precipitate and zero for the incoherent precipitate, at what ΔT are the W_R^*'s for the two equal? Assume a shear modulus of the matrix of $5.46x10^{10}$ Pa and bulk modulus of the precipitate of $15x10^{10}$ Pa.

c Repeat the previous calculation using a misfit strains of 0.01 and 0.1.

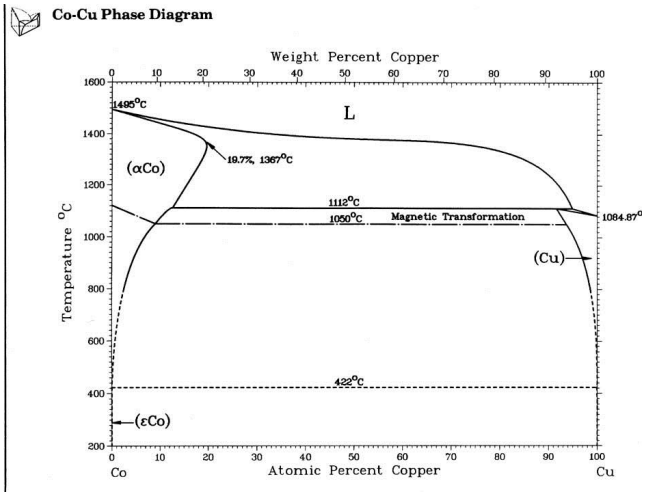
d If the number of nuclei formed per cubic centimeter per second is given by $N = 10^{27} \backslash exp(-W_R^*/kT)$, what is the rate of coherent nucleation at $\Delta T = 25K$ and $250K$ with a misfit of 0.01? What is it for incoherent nucleation at these same values of ΔT?

(16) Consider the following Al-Cu phase diagram:



Suppose that a dispersion of roughly spherical θ precipitates is formed at 300 °C. Estimate the precipitate radius for which Cu solubility in the α phase (the Al-rich phase) will be increased by 25% in comparison to a flat α/θ interface. Assume an interfacial free energy for theα/θ interface of 0.3 J/m^2 and a molar volume for the α and β phases of 7 cm^3.

(17) Consider the Co-Cu phase diagram shown below:



a Plot the equilibrium activity of Cobalt as a function of composition across the entire phase diagram at 900ºC.

b Suppose the interfacial free energy for the Cu/Co interface is 300 mJ/m^2. Develop an expression for r^*, the critical radius for a cobalt precipitate, as function of the atomic % cobalt in the alloy.

c Calculate W_r^* for a Copper rich alloy at 900ºC with a cobalt composition that exceeds the equilibrium composition by a factor of 1.15.

60.7 Spinodal Decomposition

(18) A and B form a regular solution with a positive heat of mixing so that the A-B phase diagram contains a miscibility gap.

a Starting from $G = X_A G_A + X_B G_B + \Omega X_A X_B + RT\left(X_A lnX_A + X_B lnX_B\right)$, derive an equation for d^2G/dX_B^2, assuming $G_\mathrm{A} = G_B = 0$.

b Use the above equation to calculate the temperature at the top of the miscibility gap T_c in terms of Ω.

c Using MATLAB plot the miscibility gap for this system.

d On the same diagram plot the chemical spinodal.

(19) For a homogeneous alloy of composition X_0 decomposes into two parts, one with composition $X_0 + \Delta X$ and the other with composition $X_0 - \Delta X$, show that the total chemical free energy will change by an amount ΔG_c given by

$$\Delta G_c = \frac{1}{2}\frac{d^2G}{dX^2}(\Delta X)^2$$

Hint: Express $G(X_0 + \Delta X)$ and $G(X_0 - \Delta X)$ as Taylor series.

(20) Describe the effect of each of the following, and briefly explain your answer.

a The effect of coherent strains on the characteristic wavelength of the two-phase structure formed by spinodal decomposition.

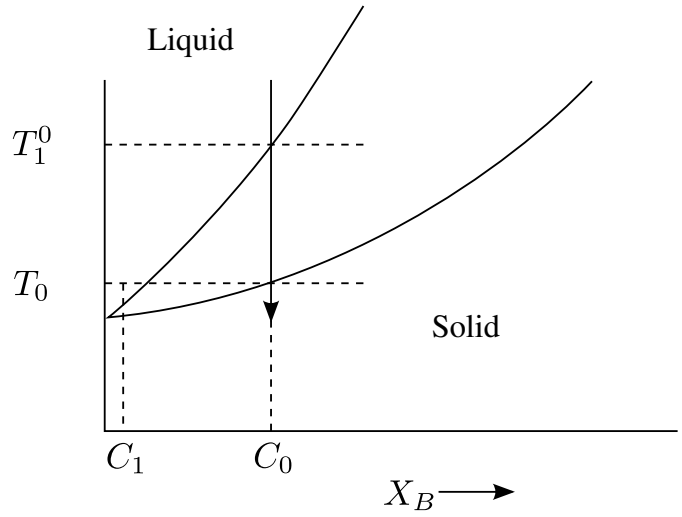
b The effect of a reduction of the surface free energy on the nucleation rate.

c The effect of a decrease in the contact angle of a precipitate on its heterogeneous nucleation rate.

d Can a diffusion coefficient ever be negative? If so, when is this the case?

60.8 Constitutional Undercooling and the 'Mushy Zone'

(21) In our classroom discussion of interface stability, we considered the case where impurities decrease the melting point. Suppose that the impurities increase the melting point, so that the phase diagram looks like this:



Suppose sample with the composition indicated by the arrow is solidified, so that the front moves forward with a certain velocity.

a Sketch the behavior of the impurity concentration in the liquid phase just ahead of the solidification front. Reference any specific compositions to the corresponding compositions on the phase diagram.

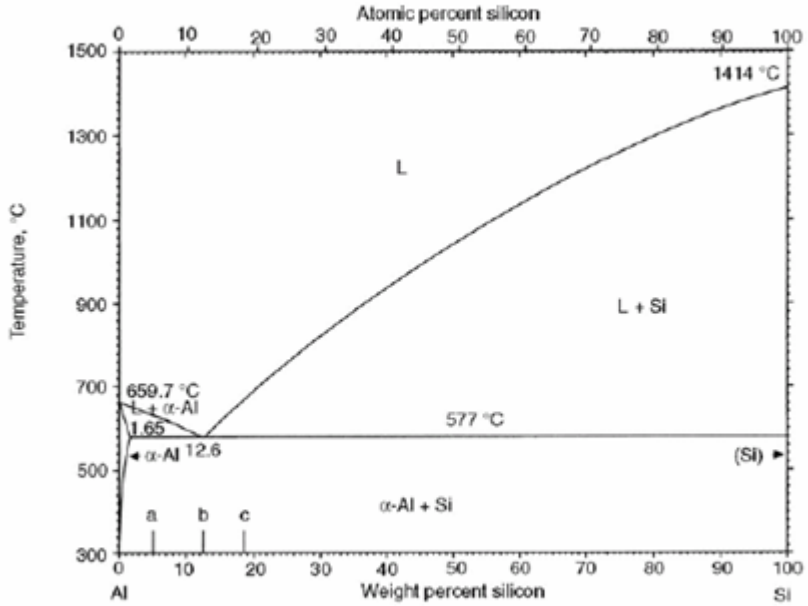
b On a separate figure, sketch the liquidus temperature in the liquid phase just ahead of the solidification front. Reference any specific temperatures to the corresponding temperatures on the phase diagram.

c Comment on the types of temperature profiles that can lead to the formation of a dendritic microstructure for this type of phase diagram. Is the criterion for interface stability qualitatively different from the criterion discussed in class?

(22) Consider the Al/Si phase diagram shown below, along with the following thermodynamic and kinetic data:

Heat of fusion for Al: 10.790 kJ/mol

Diffusion coefficient for impurities in liquid Al: $\sim 5x10^{-9}\,\mathrm{m}^2/\mathrm{s}$



Suppose an alloy with 0.8 wt. % Si is solidified at a rate of $5\mu m/s$. (This is the velocity at which the solid/liquid interface is moving.)

a What is the interface temperature in the steady state?

b What is the thickness of the diffusion layer (*i.e.* the distance into the liquid phase, measured from the solid/liquid interface, over which the liquid composition differs from the average bulk composition far from the interface?

c Estimate the temperature gradient required to eliminate the appearance of a 'mushy zone'.

60.9 Coarsening

(23) The size of Co clusters in Cu vs. aging time at several temperatures was measured using a magnetic technique. At 600 °C the data indicate the following: 10 minute aging, average particle radius = 18 Å, 100 minutes, 35 Å, 1,000 minutes, 70 Å.

a Assuming that the coarsening kinetics are consistent with Lifshitz-Slyozov-Wagner theory that was discussed in class, estimate the size of particles at $t = 0$, the end of the precipitation stage, where the cobalt supersaturation was first in equilibrium with the average size of the cobalt clusters.

b Using the data for the Co/Cu system given in the previous homework, determine the difference in the average mole fraction of Co in the Cu phase at aging times of 100 minutes and 1,000 minutes.

c Use the data given to estimate the diffusion coefficient for Co in Cu at 600 °C.

(24) Assume the following "law" for the kinetics of precipitation:

$$X(t) = 1 - \exp\left[-(t/\tau)^m\right]$$

Consider the following experimental data for the formation of Gunier-Preston zones in Al-2 wt.% Cu at 0°C for $X(t)$ less than 0.25. .

Time in hours	$X(t)$
0.4	0.08
0.7	0.10
1	0.14
2	0.17
4	0.23
6	0.28

a Determine the value of the exponent m in the above equation by plotting this equation in an appropriate fashion. (Hint: you need to rearrange the equation and take logarithms so that m is the slope of the plot).

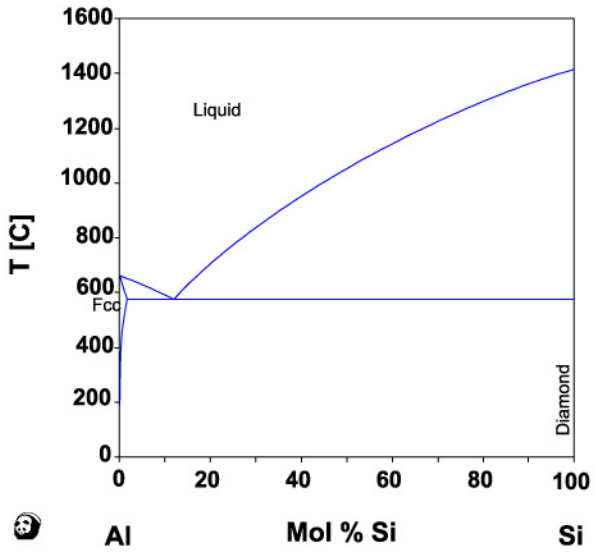
b Plot the qualitative temperature dependence that you would expect for the time constant, τ. Note that you cannot obtain this from the data provided – you need to make some assumptions about what you expect this to look like) Comment on the factors that cause τ to become very large at high and low temperatures. From Fig. 5.25 in Porter and Easterling, what can you say about the behavior of τ in the high temperature regime (i.e., at what temperature must τ diverge to infinity)?

60.10 Eutectic Solidification

(25) Refer to the Al/Si phase diagram and thermodynamic data below to answer the following questions.

Heat of fusion for Al: 395 J/g

Heat of fusion for Si: 1408 J/g



Calculated Al-Si phase diagram assessed by Murray and McAlister

NIMS

a Obtain an estimate for the heat of fusion for the Al/Si eutectic (Joules per cm^3 of eutectic).

b Calculate the bulk free energy gain (ignoring the energy associated with the Al/Si interfaces) associated with the solidification of 1 cm3 of eutectic at 560 °C.

c Calculate the width of the Al and Si phases in a lamellar eutectic for the case where the total free energy change (including the energy associated with the interfaces) on solidification at 560 °C is equal to zero. Assume an interfacial free energy for the Al/Si interface of 350 mJ/m^2.

d Compare the phase widths from part c to the critical radii for the solidification of pure Al and pure Si at an undercooling of 20 °C. Assume that the solid liquid interfacial free energies are similar in magnitude to the Al/Si (solid/solid) interfacial free energy.

e How good is the assumption of ideal liquid mixing in this case? Plot the liquidus lines for the Al-rich and Si-rich phases, using the equation that was developed in class. Compare the location of these lines with the location of the actual eutectic point, and comment on the agreement that you observe.

60.11 Eutectoid Transormations

(26) Imagine the Fe-0.15 wt% C alloy in the figure below is austenitized above A_3, and then quenched to800°C where ferrite nucleates and covers the austenite grain boundaries.

a Draw a composition profile normal to the α/γ interface after partial transformation assuming diffusion-controlled growth.

b Derive an approximate expression for the thickness of the ferrite slabs as a function of time.

(27) The eutectoid temperature for the Fe/C phase diagram is 723 °C. Pearlite formed at 713 °C has a lamellar period (λ) of 1μm.

a Calculate the lamellar period for pearlite that you would expect if the pearlite were formed at a temperature of 623 °C.

b Pearlite forms initially at grain boundaries within the parent austenite phase. Briefly describe why this is so.

c Describe what happens to the microstructure of the steel and to the hardness as increasingly large cooling rates are used. Discuss the role of carbon diffusion, and the role of both equilibrium and non-equilibrium phases.

(28) In the reading about the Wright Flyer Crankcase, the authors assert "In an Al-Cu alloy with significant supersaturation, GP zones develop by spinodal decomposition. The spacing between zones (before coarsening) is determined by the fastest growing wavelength during decomposition. The favored wavelength is inversely related to the second derivative of the free energy versus composition function,which is zero at the spinodal line (located inside but near the GP zone solvus curve) and increases(negatively) with an increase in Cu or a decrease in temperature. Thus, the favored wavelength in the region with a large amount of Cu is smaller than in the regions with small amounts of Cu, and the resulting spacing between zones is smaller."

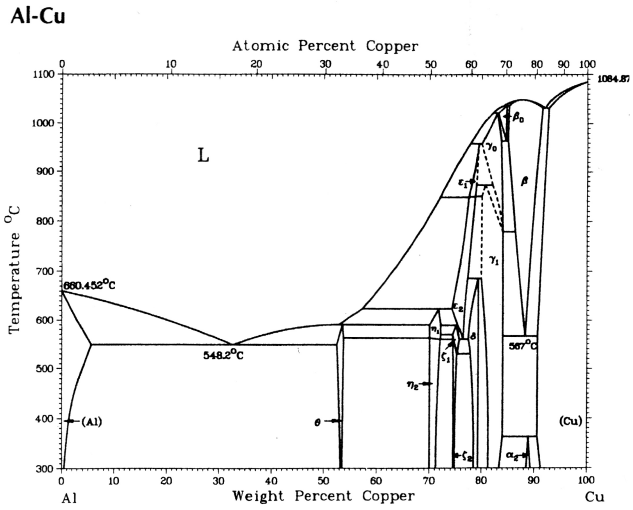
d Support their argument using the equations derived in class for spinodal decomposition. A good way to approach this is to postulate a spinodal line and then consider two cases, i.e. $X_o = 2.5$ wt.% Cu and $X_o = 4.5$ wt.% Cu, in detail, based on this spinodal. Be sure to also explain why the authors made the parenthetical statement "(before coarsening)".

e Is their argument entirely valid or do certain conditions need to be met? If so, what qualifications should be made to make it more accurate?

f The authors claim that "The growth of [GP] zones is ultimately limited by solute depletion in the matrix. Despite its high solute concentration, the region with a large amount of Cu is depleted of solute by the time the zones have grown to about 10 nm." Let's assume with them the GP zones grow at 100°C from a matrix with 4.5 wt.% Cu. How do the authors arrive at this statement, what evidence did they likely use, and what calculations did they perform to arrive at this statement?

60.12 Transitional Phases

(29) Suppose an alloy containing 97 wt. % aluminum and 3 wt. % copper is poured into a mold and solidified by extracting heat from the external surfaces of the mold.



a What phase (or phases) do you expect to be present in the solid immediately after the solidification reaction?

b What phase (or phases) will be present at equilibrium?

c A variety of non-equilibrium phases are observed at intermediate stages in the transformation process. Why are these phases observed?

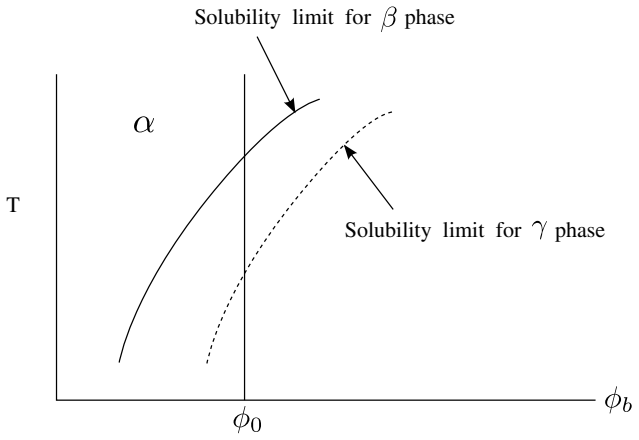
d Once the equilibrium phase is formed, its rate of growth is found to decrease with time. Why is this? What is the rate limiting step in the transformation?

(30) Porter and Easterling, prob. 5.6

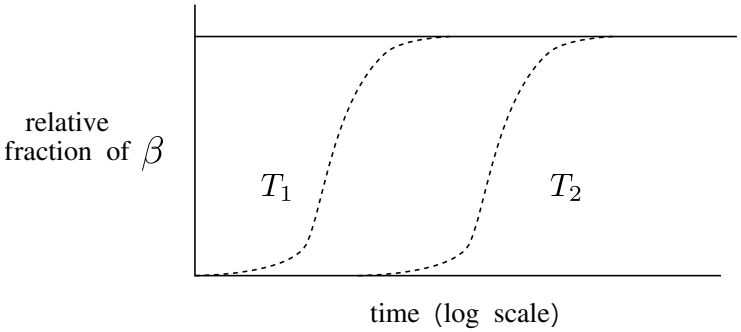
(31) Suppose that in the system of interest, χ is inversely proportional to the absolute temperature, and the critical temperature for this system is 350 K. Replot the phase diagram from part a with temperature on the vertical axis.

60.13 TTT diagrams

(32) Consider the following blowup of the low concentration region of a phase diagram similar to the Co-Cu diagram shown above:



a At an average alloy composition ϕ_0 shown on this plot, it is determined that only β precipitates form (no γ is ever observed) at two temperatures, T_1 and T_2. The time dependence of the appearance of β is plotted at these two temperatures as shown below. On the plot above, indicate locations of $T_1$and T_2 that are consistent with these curves, and briefly describe your reasoning.



b Draw TTT curves for the precipitation of both β and γ for an alloy with the composition of ϕ_0, making connections to specific temperatures from the phase diagram where possible.

60.14 Mineralization

(33) Calculate the CO_3^{2-} concentration in equilibrium with seawater and with each of the following three forms of calcium carbonate: calcite, aragonite, vaterite. You'll need to use the solubility products provided in class, and look up the calcium concentration in seawater.

60.15 Review Questions

- What does the liquid composition look like in front of an advancing solid phase?
- What controls length scale of the composition variation?
- What is the criteria for interface stability with respect to the formation of dendrites?
- Under what conditions are dendrites formed during the solidification of a pure material?
- What is the qualitative behavior of $S(\theta)$ for nucleation at a flat interface, and at grain boundary surfaces, edges and corners?
- How are equilibrium contact angles related to surface and interfacial free energies?
- What is meant by complete wetting?
  - What is the effect on nucleation for the complete wetting case?
- What are the characteristic frequencies and concentrations (C_0, ν_0) that determine the homogeneous and heterogeneous nucleation rates?
- Why is coherent nucleation generally the favored homogeneous nucleation mechanism?
- How is the work to form the critical nucleus calculated?
- What is the 'incoherent solvus', and how does it relate to the expressions listed above?
- Where does the Laplace pressure come from?
- How do you know that kinetic factors must be controlling complex morphologies (dendrite formation, shapes of snowflakes, etc.)?
- What controls the size of the depletion zone in front of a flat or curved precipitate that is growing?
- How does it evolve with time?
- What limits the growth velocity of a precipitate phase boundary at high and low temperatures?
  - Which of these limits are connected to the phase diagram?
- Why are flat, plate-like precipitates sometimes formed?
- How does the molar free energy depend on the radius of curvature of a precipitate?
- How does curvature effect the equilibrium concentration of solute that is in equilibrium with a precipitate?
- What does the concentration dependence look like for precipitates that are larger than r*?
  - What if the precipitate is smaller than r*?
- What are transition phases, and why do they form?
  - What is the mechanism by which transition phases shrink at the expense of equilibrium phases?
- How are TTT curves for transition phases related to the phase diagram?
- What are the basic physical assumptions of the Lifshitz/Slyosov coarsening theory discussed in class?
  - What do the depletion zones look like?
  - What determines the average solute concentration in the matrix phase?
- What does the distribution of precipitates look like if coarsening occurs by the Lifshitz/Slyosov mechanism?
- What do the binodal and spinodal curves look like for the regular solution model?

  - What is the critical temperature?

- What determines the size of the characteristic phase size when phase separation occurs by spinodal decomposition?
- What is meant by uphill diffusion?
  - When is it observed?
- How is this phase size modified (in qualitative terms) by coherent strains?
- How do these strains modify the phase diagram to give coherent spinodal and binodal curves?
- How can the liquidus lines be estimated for an ideal eutectic system?
  - What are the assumptions made in the approximation?
- What determines the size of the individual phases for eutectic solidification?
- What determines the size of the individual phases for a eutectoid transformation?
- What is the physical significance of the squared gradient term in the free energy expression?
- How can the shapes and sizes of metallic nanoparticles be controlled?
- What is the growth mechanism of Si nanowires catalyzed by Gold?
  - What is the importance of the Au/Si phase diagram.
- How does the solubility of calcite compare to the solubility of aragonite or vaterite, and why?
  - How is the concept of the solubility product used?
- In the two-phase mixture of n-type and p-type materials used to form an organic solar cell, what sort of phase morphology is desired, and why?

Part VI
332: Mechanical Behavior of Materials

61 Introduction

In Figure 61.1, we show the materials available to mankind from the beginning of time until the current day.[17] The data are plotted as so-called 'Ashby plots', where two material properties are chosen for the x and y axes. In Figure 61.1 we use density to separate different materials along the x axis, and the fracture strength, σ_f, to separate materials on the y axis. When the story begins, all we had to work with are the materials we could find around us or dig out of the ground. Over time, we figured out how to produce more materials for all kinds of purposes. The biggest developments come after World War II, a situation that coincides with the emergence of materials science as a discipline. We no longer rely on empiricism for materials development, but can actually begin to design new materials with the properties we want, or at least design materials that re better than anything we had before. The principles underlying the development of materials with the mechanical response forms the basis for this course.

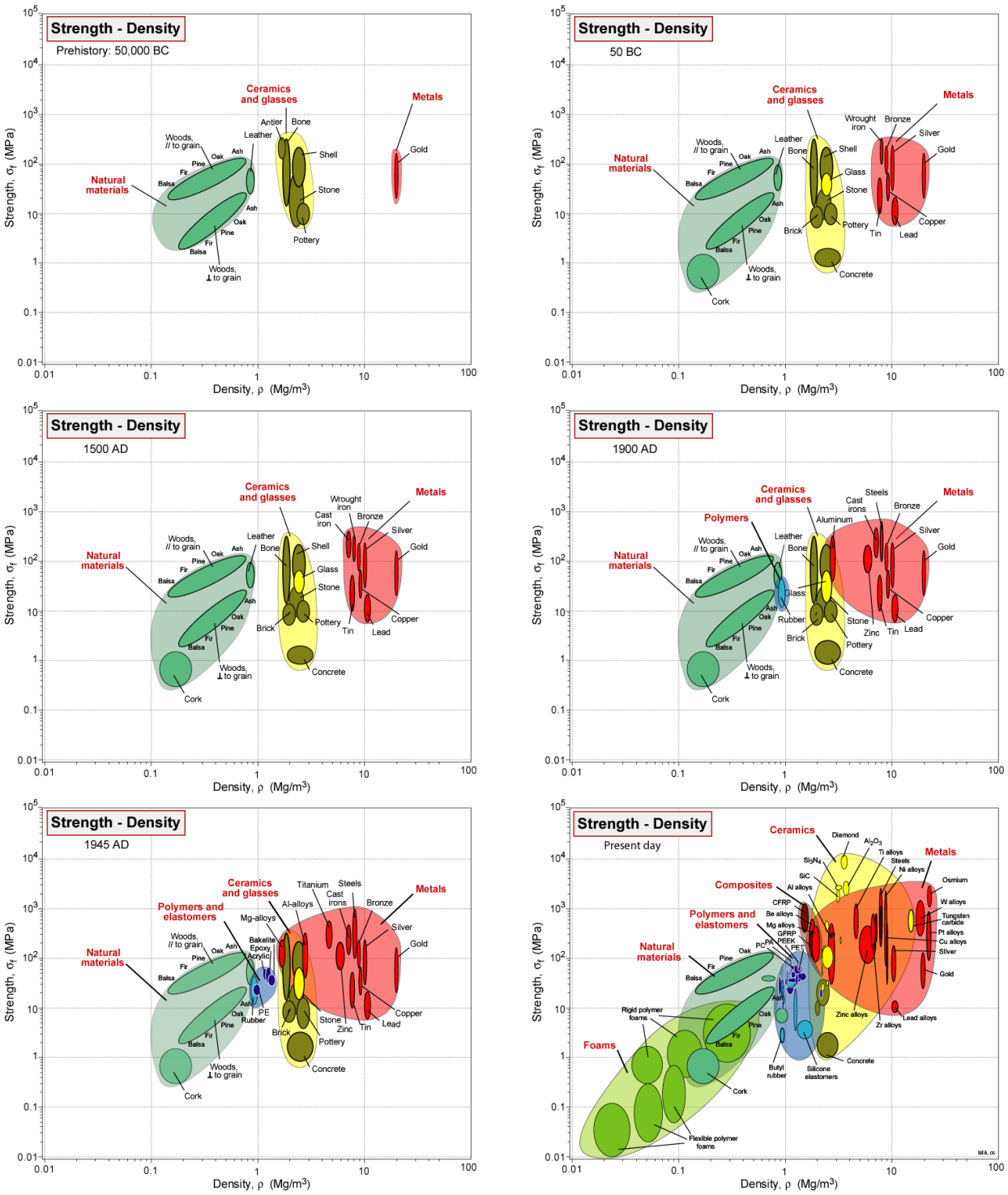


Figure 61.1: Available materials throughout history (from ref.[17]).

62 Stress and Strain

The mechanical properties of a material are defined in terms of the strain response of material after a certain stress is applied. In order to properly understand mechanical properties, we have to have a good understanding of stress and strain, so that's where we begin.

62.1 Tensor Representation of Stress

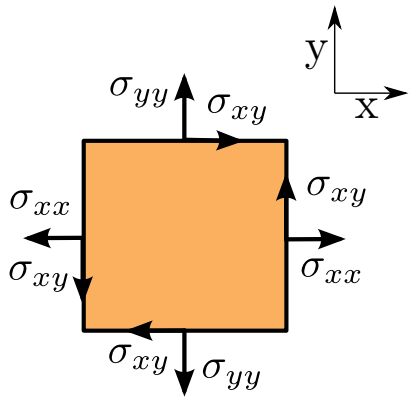


Figure 62.1: 2-dimensional stress tensor

σ_{ij} (stress) = force/area - force and area area vectors (direction of area vector is normal to the surface). The indices, i and j, of the stress tensor σ_{ij} have the following significance:

- i: surface normal (i= x, y, z)
- j: direction of force (j=x, y, z);

σ_{eng} Engineering stress: areas are of undeformed object
True stress (areas defined in deformed state) is a tensor. In two dimensions it has the following form:

$$\sigma = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} \tag{62.1}$$

The tensor must symmetric, with $\sigma_{xy} = \sigma_{yx}$. If this were not the case, the torques on the volume element shown above in Figure 91.1 would not balance, and the material would not be in static equilibrium. As a result the two dimensional stress state is specified by three components of the stress tensor:

- 2 normal stresses, σ_{xx}, σ_{yy}. These are referred to as 'normal' stresses because the force acts perpendicular to the plane that it is referred to.
- A single shear stress, σ_{xy}.

In three dimensions we add a z axis to the existing x and y axes, so the stress state is defined by a symmetric 3x3 matrix. The matrix can be used to define the stresses acting on any given plane. To simplify the notation a bit we label the three orthogonal directions by numbers (1, 2 and 3) instead of letters (x, y and z). The stress tensor gives the components of the force (P_1, P_2 and P_3) acting on a given plane. The plane is specified by the orientation of the unit vector, $\hat{n}$ that is perpendicular to the plane. This vector has components n_1, n_2 and n_3 in the 1, 2 and 3 directions, respectively. It's a 'unit' vector because the length of the vector is 1, *i.e.* $\left(n_1^2 + n_2^2 + n_3^2\right)^{1/2} = 1$. The relationship between $\vec{P}$, σ and $\vec{n}$ is as follows:

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = A \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \tag{62.2}$$

Here A is the total cross sectional area of the plane that we are interested in. (If you need a refresher on matrix multiplication, the Wikipedia page is very helpful).

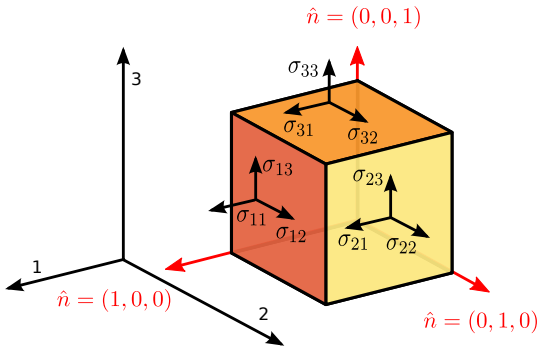


Figure 62.2: 3 dimensional stress tensor

In graphical form the relationship is as shown in Figure 62.2. Like the 2-dimensional stress tensor mentioned above, the 3-dimensional stress tensor must also be symmetric in order for static equilibrium to be achieved. There are therefore 6 independent components of the three-dimensional stress tensor:

- 3 normal stresses, σ_{11}, σ_{22} and σ_{33}, describing stresses applied perpendicular to the 1, 2 and 3 faces of the cubic volume element.
- 3 shear stresses: σ_{12}, σ_{23} and σ_{13}.

The three dimensional stress tensor is a 3x3 matrix with 9 elements (though only 6 are independent), corresponding to the three stress components acting on each of the three orthogonal faces of cube in the Cartesian coordinate system used to define the stress components. The 1 face has n_1=1, $n_2 = 0$ and $n_3 = 0$. by setting $\vec{n} = (1, 0, 0)$ in Eq. 62.2, we get the following for the stresses acting on the 1 face of the volume element:

$$\begin{aligned} P_1/A &= \sigma_{11} \\ P_2/A &= \sigma_{12} \\ P_3/A &= \sigma_{13} \end{aligned} \tag{62.3}$$

Equivalent expressions can be obtained for the stresses acting on the 2 and 3 faces, by setting $\vec{n} = (0, 1, 0)$ and $\vec{n} = (0, 0, 1)$, respectively.

62.2 Tensor Transformation Law

The stress experienced by a material does not depend on the coordinate system used to define the stress state. The stress tensor will look very different if we chose a different set of coordinate axes to describe it, however, an it is important to understand how changing the coordinate system changes the stress tensor.

62.2.1 Specification of the Transformation Matrix

In general, we consider the case where our 3 axes (which we refer to simply as axes 1, 2 and 3) are moved about the origin to define a new set of coordinate axes that we refer to as $1'$, $2'$and $3'$. As an example, consider the simple rotation around the 3 axis by an angle ϕ, shown schematically in Figure 62.3. In general, the relative orientation of the transformed (rotated) and untransformed coordinate axes are given by a set of 9 angles between the 3 untransformed axes and the three transformed axes. In our notation we specify these angles as θ_{ij}, where i specifies the transformed axes ($1'$, $2'$ or $3'$) and j specifies the transformed axis (1, 2 or 3). In our simple example, the angle between the 1 and $1'$ axes is ϕ, so $\theta_{11} = \phi$. The angle between the 2 and $2'$ axes is also ϕ, so $\theta_{22} = \phi$. The 3 axis remains unchanged in our rotation example, so $\theta_{33} = 0$. The $3/3'$ axis remains perpendicular to the 1, $1'$, 2, $2'$ axes, so we have $\theta_{31} = \theta_{32} = \theta_{13} = \theta_{23} = 90°$. Finally, we see that the angle between the $1'$ and the 2 axis is $90 - \phi$ and the angle between the $2'$ and 1 axis is $90 + \phi$. The full θ_{ij} matrix is as follows:

$$\theta_{ij} = \begin{bmatrix} \phi & 90-\phi & 90 \\ 90+\phi & \phi & 90 \\ 90 & 90 & 0 \end{bmatrix} \tag{62.4}$$

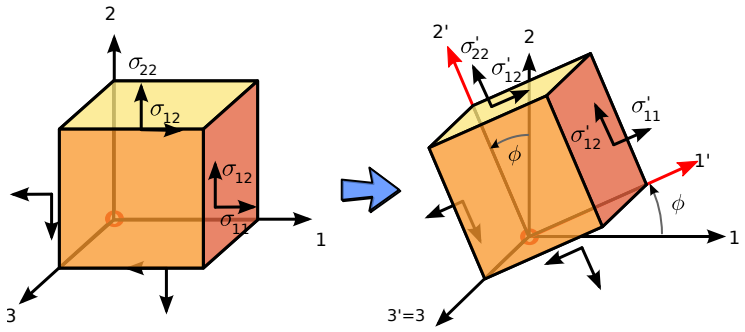


Figure 62.3: Rotation of coordinate system. The coordinate system is rotated by θ about the 3 axis, transforming the 1 axis to $1'$ and the 2 axis to $2'$.

62.2.2 Expressions for the Stress Components

Once we specify the values of θ_{ij}, we can use the following general expression to obtain the stresses in the new (primed) coordinate system as a function of the stresses in the original coordinate system:

$$\sigma'_{ij} = \sum_{k,l} \cos\theta_{kj} \cos\theta_{li} \sigma_{kl} \tag{62.5}$$

For each component of the stress tensor, we have to sum 9 individual terms (all combinations of k and l from 1 to 3). For example, σ'_{12} is given as follows:

$$\begin{aligned}\sigma'_{12} = &\cos\theta_{12}\cos\theta_{11}\sigma_{11} + \cos\theta_{12}\cos\theta_{22}\sigma_{12} + \cos\theta_{12}\cos\theta_{31}\sigma_{13} + \\ &\cos\theta_{22}\cos\theta_{11}\sigma_{21} + \cos\theta_{22}\cos\theta_{21}\sigma_{22} + \cos\theta_{22}\cos\theta_{31}\sigma_{23} + \\ &\cos\theta_{32}\cos\theta_{11}\sigma_{31} + \cos\theta_{32}\cos\theta_{21}\sigma_{32} + \cos\theta_{32}\cos\theta_{31}\sigma_{33}\end{aligned} \tag{62.6}$$

The calculation is breathtakingly tedious if we do it all by hand , so it makes sense to automate this and do the calculation via computer, in our case with MATLAB. In this example we'll start with a simple stress state corresponding to uniaxial extension in the 1 direction, with the following untransformed stress tensor:

$$\sigma_{ij} = \begin{bmatrix} 5x10^6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{62.7}$$

Suppose we want to obtain the stress tensor in the transformed coordinate system obtained from a 45 ° counterclockwise rotation around the z axis. The rotation matrix is given by Eq. 62.4, with $\phi = 45°$. The following MATLAB code solves for the full transformed tensor, using these values for θ_{ij}, with σ_{ij} given by Eq. 62.7:

```
%rotate45.m
sig=zeros(3); % create stress tensor and set to zero
sig(1,1)=5e6; % this is the only nonzero component

sigp=zeros(3); % initalize rotated streses to zero
phi=45;
theta=[phi,90-phi,90;90+phi,phi,90;90,90,0];
for i=1:3
    for j=1:3
        for k=1:3
            for l=1:3
                sigp(i,j)=sigp(i,j)+cosd(theta(k,i))*cosd(theta(l,j))*sig(k,l);
            end  % cosd takes degrees instead of radians as the unit
        end
    end
end
sigp %display the transformed tensor components
```

```
sigp =

   1.0e+06 *

    2.5000    2.5000         0
    2.5000    2.5000         0
         0         0         0
```

Figure 62.4: MATLAB output generated by rotate45.m.

The output of generated by MATLAB is shown in Figure 62.4, and corresponds to the following result:

$$\sigma_{ij} = \begin{bmatrix} 2.5x10^6 & 2.5x10^6 & 0 \\ 2.5x10^6 & 2.5x10^6 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{62.8}$$

Note the following:

- The normal stresses in the 1 and 2 directions are equal to one another.
- The transformed shear stress in the 1-2 plane is half the original tensile stress.
- The sum of the normal stresses (the sum of the diagonal components of the stress tensor) is unchanged by the coordinate transformation

62.2.3 An Easier Way: Transformation by Direct Matrix Multiplication

A much easier way to write the transformation is to do it is to use a little bit of matrix math. The approach we used is described in a very nice web page put together by Bob McGinty[25]: The transformation matrix, Q, is the obtained by taking the cosines of all of these angles describing the relationship between the transformed and untransformed coordinate axes:

$$Q_{ij} = \cos\theta_{ij} \tag{62.9}$$

For the simple case of rotation about the z axis, the angles are given by Eq. 62.4, and Q is given as:

$$Q = \begin{bmatrix} \cos\phi & \cos(90-\phi) & \cos 90 \\ \cos(90+\phi) & \cos\phi & \cos 90 \\ \cos 90 & \cos 90 & \cos 0 \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{62.10}$$

The transformed stress is now obtained by the following simple matrix multiplication:

$$\sigma' = Q\sigma Q^T \tag{62.11}$$

where the Q^T is the transpose of Q:

$$Q^T(i,j) = Q(j,i) \tag{62.12}$$

For the rotation by ϕ around the z axis, Q^T is given by the following:

$$Q^T = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{62.13}$$

Equation62.11 is much easier to deal with than Eq. 62.6, at least in MATLAB, where matrices can be multiplied as regular numbers. The MATLAB code to take a uniaxial stress state and rotate the coordinate system by 45 $^\circ$ about the 3 axis looks like this if we base it on Eq. 62.11:

```
%rotate45_matrix.m
sig=zeros(3); % create stress tensor and set to zero
sig(1,1)=5e6; % this is the only nonzero component
phi=45;
theta=[phi,90+phi,90;90-phi,phi,90;90,90,0];
Q=cosd(theta);
QT=transpose(Q);
sigp=Q*sig*QT
```

Running this script gives the output shown in Figure 62.4.

62.3 Principle Stresses

Any stress state (true stress) can be written in terms of three principal stresses σ_1, σ_2, σ_3 applied in three perpendicular directions.

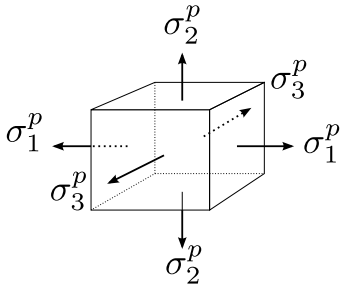


Figure 62.5: Principal Stresses

Stress state fully specified by principal stresses, principal axes. The stress tensor depends on our definition of the axes, but it is always possible to chose the axes so that all of the shear components of the stress tensor vanish, so that the stress tensor looks like the following:

$$\sigma_{ij} = \begin{bmatrix} \sigma_1^p & 0 & 0 \\ 0 & \sigma_2^p & 0 \\ 0 & 0 & \sigma_3^p \end{bmatrix} \tag{62.14}$$

In order to gain some insight into the three points mentioned above, it is useful to consider a range of rotation angles, and not just a singe rotation angle of 45°. One way to dothis is to use the symbolic math capability of MATLAB to obtain the full stress tensor as a function of the rotation angle. We'll use a slightly more complicated stress tensor as our input, which has principal stresses in all three untransformed directions, but without shear stresses, as shown below in Eq. 62.14.
The following MATLAB script results in the output shown in Figure 62.6.

```
%rotatesym.m
clear all
syms phi; % declare as symbolic variable
sig=sym(zeros(3)); % create sybolic stress tensor and set to zero
syms sig1p sig2p sig3p % these are the  normal stresses
sig(1,1)=sig1p; % put normal stresses into the tensor
sig(2,2)=sig2p;
sig(3,3)=sig3p;
theta=[phi,pi/2+phi,pi/2;pi/2-phi,phi,pi/2;-pi/2,-pi/2,0];
Q=cos(theta);
QT=transpose(Q);
sigp=Q*sig*QT;
sigp=simplify(sigp);
pretty(sigp);
```

The components of the stress tensor in the rotated coordinate system are:

$$\sigma'_{ij} = \begin{bmatrix} \sigma_1^p - \sigma_1^p \sin^2\phi + \sigma_2^p \sin^2\phi & \sin(2\phi)\left(\sigma_1^p - \sigma_2^p\right)/2 & 0 \\ \sin(2\phi)\left(\sigma_1^p - \sigma_2^p\right)/2 & \sigma_2^p + \sigma_1^p \sin^2\phi - \sigma_2^p \sin^2\phi & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{62.15}$$

```
+-                                                                                -+
|                          2                    2          sin(2 phi) (sig1p - sig2p)                |
|  sig1p - sig1p sin(phi)  + sig2p sin(phi) ,          --------------------------,         0        |
|                                                                    2                              |
|                                                                                                   |
|          sin(2 phi) (sig1p - sig2p)                                  2                 2          |
|          --------------------------,        sig2p + sig1p sin(phi)  - sig2p sin(phi) ,   0        |
|                     2                                                                             |
|                                                                                                   |
|                     0,                                  0,                        sig3p       |
+-                                                                                -+
```

Figure 62.6: MATLAB output generated by rotatesym_matrix.m.

This is not yet a very illuminating result, but it is the basis for the Mohr circle construction, which provides a very useful way to visualize two dimensional stress states. This construction is described in more detail in the following Section .

62.3.1 Mohr's Circle Construction

The Mohr circle is a graphical construction that can be used to describe a two dimensional stress state. A two dimensional stress state is specified by two principal stresses, σ_1^p and σ_2^p, and by the orientation of the principal axes. The Mohr circle is drawn diameter = $\sigma_1^p - \sigma_2^p$, centered at $\left(\sigma_1^p + \sigma_2^p\right)$. Continue symbolic calculation with the following script:

```
% Mohr's circle derivation - beginning from previous calculation of sigp
center=(sig1p+sig2p)/2;  % the center of Mohr's circle
radius=(sig1p-sig2p)/2;  % radius of Mohr's circle
% now we shift the origin to the center of Mohr's circle, and
% normalize by the radius of Mohrs circle
sigp(1,1)=(sigp(1,1)-center)/radius;
sigp(2,2)=(sigp(2,2)-center)/radius;
% shear stresses are plotted on the y axis, so we don't need to shift the
% origin, we just need to normlized by the radius
sigp(1,2)=sigp(1,2)/radius;
sigp(2,1)=sigp(1,2);
simplify(sigp)
```

The output of Mohr_circle.m is shown In Figure 62.7.

```
[ cos(2*phi),  sin(2*phi), 0]
[ sin(2*phi), -cos(2*phi), 0]
[          0,           0, 0]
```

Figure 62.7: MATLAB output generated by mohr_circle.m.

In the Mohr's circle construction the strain shear components of the stress tensor are plotted on the x axis and the shear component of the stress tensor is plotted on the y axis. An example is shown in Figure 62.8, where each of the

components of the stress tensor is assumed to be a positive number. In this particular example, the coordinate axes need to be rotated by 2θ to get obtain the principal axes. Whether this rotation is clockwise or counterclockwise depends on the sign convention in the definition of the shear stress. We're not going to worry about it hear, but you can refer to the Wikipedia article for the details (see the Section on the sign convention).

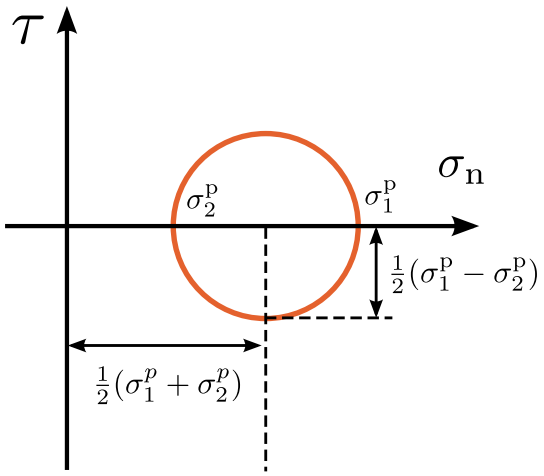


Figure 62.8: Mohr's circle construction.

http://en.wikipedia.org/wiki/Mohr_circle

In general, their are three principal stresses, σ_1^p, σ_2^p and σ_3^p, and we can draw the Mohr circle construction with any combination of these 3 stresses. We end up with 3 different circles, as shown in Figure 62.9. Note that the convention is that σ_1^p is the largest principal stress and that σ_3^p is the smallest principal stress,*i.e.* $\sigma_1^p > \sigma_2^p > \sigma_3^p$. An important result is that the largest shear stress, τ_{max}, is given by the difference between the largest principal stress and the smallest one:

$$\tau_{max} = \frac{1}{2}\left(\sigma_1^p - \sigma_3^p\right) \tag{62.16}$$

This maximum shear stress is an important quantity, because it determines when a material will deform plastically (much more on this later). In order to determine this maximum shear stress, we need to first Figure out what the principal stresses are. In some cases this is easy. In a uniaxial tensile test, one of the principals stresses is the applied stress, and the other two principal stresses are equal to zero.

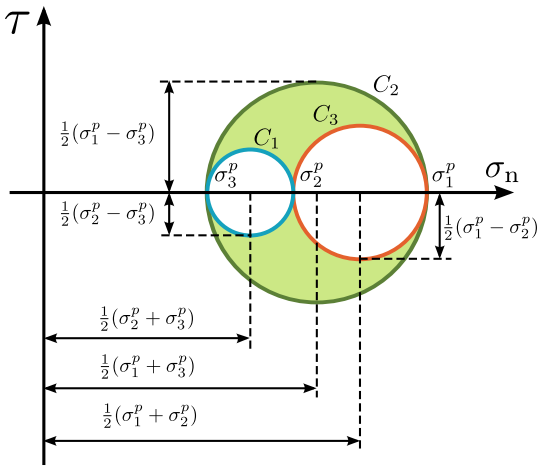


Figure 62.9: Three-dimensional Mohr's Circle.

62.3.2 Principal Stress Calculation

Principal stresses can easily by calculated for any stress state just by obtaining the eigenvalues of the stress tensor. In addition, the orientation of the principal axes (the coordinate system for which there are no off-diagonal components in the stress tensor). If you need a refresher on what eigenvalues and eigenvectors actually are, take a look at the appropriate Wikipedia page (http://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors). We'll use MATLAB to do this, using the 'eigs' command .
To illustrate, we'll start with the stress state specified by Eq. 62.8. which we got by starting with a simple uniaxial extension in the 1 direction, and rotating the coordinate system by 45 ° about the 3 axis. The MATLAB script to do this is very simple and is as follows:

```
clear all % clear the workspace of previously defined variables
sig=1e6*[2.5,2.5,0;2.5,2.5,0;0,0,0];
[directions,principalstresses]=eigs(sig);
% the columns in 'directions' correspond to the dot product of the
% principal axes with the orignal coordinate system
% The rotation angles are obtained by calculating the inverse cosines
theta=acosd(directions)
principalstresses
```

Here's the output generated by this script:

```
theta =

   45.0000   90.0000  135.0000
   45.0000   90.0000   45.0000
   90.0000         0   90.0000


principalstresses =

    5000000         0         0
          0         0         0
          0         0         0
```

Figure 62.10: MATLAB output generated by principal_stress_calc.m.

3 principal axes returned as column vectors. In this case there is a single normal stress, acting in a direction midway between the original x and y axes. The original uniaxial stress state is recovered in this example, as it should be. To summarize:

- Principal Stresses: Eigenvalues of the stress tensor
- Principal Stress directions: Eigenvectors of the stress tensor

62.3.3 Stress Invariants

Some quantities are invariant to choice of axes. The most important one is the hydrostatic pressure, p, given by summing the diagonal components of the stress tensor and dividing by 3:

$$p = -\frac{1}{3}\left(\sigma_{11} + \sigma_{22} + \sigma_{33}\right) = -\frac{1}{3}\left(\sigma_1^p + \sigma_2^p + \sigma_3^p\right) \tag{62.17}$$

The negative sign appears because a positive pressure is compressive, but positive stresses are tensile. The hydrostatic pressure is closely related to a quantity referred to as the 'firststress invariant', I_1:

$$I_1 = \sigma_{11} + \sigma_{22} + \sigma_{33} \tag{62.18}$$

The second and third stress invariants, I_2 and I_3, are also independent of the way the axes are defined:

$$I_2 = \sigma_{11}\sigma_{22} + \sigma_{22}\sigma_{33} + \sigma_{33}\sigma_{11} - \sigma_{12}^2 - \sigma_{13}^2 - \sigma_{23}^2 \tag{62.19}$$

$$I_3 = \sigma_{11}\sigma_{22}\sigma_{33} - \sigma_{11}\sigma_{23}^2 - \sigma_{22}\sigma_{13}^2 - \sigma_{33}\sigma_{12}^2 + 2\sigma_{12}\sigma_{13}\sigma_{23} \tag{62.20}$$

The fact that these invariants exist is useful, but we're not going to do a lot more with them in this class. The most important invariant is the hydrostatic pressure, p.

62.4 Strain

There are 3 related definitions of the strain:

(1) Engineering strain

(2) Tensor strain

(3) Generalized strain (large deformations-

Each of these definitions of strain describe the way different points an object move relative to one another when the material is deformed. Consider two points P_1 and P_2, separated initially by the increments d x, d y and d z along the x, y and z directions. After the deformation is applied, these points move by the following amounts, as illustrated in Figure 62.11:

- P_1 moves by an amount (u, v, w)
- P_2 moves by $(u + du, v + dv, w + dw)$

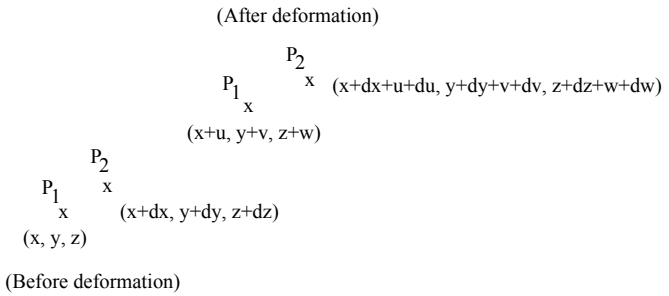


Figure 62.11: Location of two points, P_1 and P_2, before and after an applied deformation.

62.5 Small Strains

Strain describes how much farther point to moved in three different directions, as a function of how far P_2 was from P_1 initially. For small strains we can ignore higher order terms in a Taylor expansion for du, dv and dw and maintain only the first, partial derivative terms as follows:

$$\begin{aligned} du &= \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial z}dz \\ dv &= \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy + \frac{\partial v}{\partial z}dz \\ dw &= \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy + \frac{\partial w}{\partial z}dz \end{aligned} \tag{62.21}$$

The three normal components of the strain correspond to the change in the displacement in a given direction corresponds to a change in initial separation between the points of interest in the same direction:

$$\begin{aligned} e_{xx} &= \frac{\partial u}{\partial x} \\ e_{yy} &= \frac{\partial v}{\partial y} \\ e_{zz} &= \frac{\partial w}{\partial z} \end{aligned} \tag{62.22}$$

The engineering shear strains are defined as follows:

$$\begin{aligned} e_{yz} &= \frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} \\ e_{xz} &= \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \\ e_{xy} &= \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \end{aligned} \tag{62.23}$$

Note: shear strains are often represented by the lower case Greek gamma to distinguish them from normal strains:

$$\gamma_{xy} \equiv e_{xy};\ \gamma_{yz} \equiv e_{yz};\ \gamma_{xz} = e_{xz} \tag{62.24}$$

62.5.1 Tensor Shear Strains

Engineering strains relate two vectors: (dx, dy, dz) and (du, dv, dw), just as a tensor does, but the the transformation law between different coordinate systems is not obeyed for the engineering strains. This is because the All we need to do to get change engineering strains to tensor strains is to divide the shear components by 2. In our notation we use e to indicate engineering strain and ϵ to indicate tensor strains. The tensor normal strains are exactly the same as the engineering normal strains:

$$\begin{aligned} \epsilon_{xx} &= e_{xx} \\ \epsilon_{yy} &= e_{yy} \\ \epsilon_{zz} &= e_{zz} \end{aligned} \tag{62.25}$$

engineering shear strains (e_{yz}, e_{xz}, e_{zy}) are divided by two to give tensor shear strains:

$$\begin{aligned} \epsilon_{yz} &= e_{yz}/2 \\ \epsilon_{xz} &= e_{xz}/2 \\ \epsilon_{xy} &= e_{xy}/2 \end{aligned} \tag{62.26}$$

Note that the tensor strains must be used in coordinate transformations (axis rotation, calculation of principal strains, e_1^p, e_2^p, e_3^p).

62.5.2 Generalized Strain

We can also define the strain by considering a cube of side ℓ that is deformed into a parallelepiped with dimensions of (along principal strain axes). After deformation, the cube has dimensions of $\lambda_1\ell$, $\lambda_2\ell$, $\lambda_3\ell$. Alternatively, a sphere of radius r_0 is deformed into and ellipsoid with principal axes of $\lambda_1 r_0$, $\lambda_2 r_0$ and $\lambda_3 r_0$, as shown in Figure 91.3. The quantities λ_1, λ_2, λ_3 areextension ratios, and are related to the principal strains as follows:

$$\begin{aligned} \lambda_1 &= 1 + e_1^P \\ \lambda_2 &= 1 + e_2^P \\ \lambda_3 &= 1 + e_3^P \end{aligned} \tag{62.27}$$

The true strains, e_1^t, are obtained as by taking the natural log of the relevant extension ratio. For example, for a uniaxial tensile test, the true strain in the tensile direction (assumed to be the 1 direction here) is:

$$e_1^t = \ln(\lambda_1) \tag{62.28}$$

This expression for the true strain can be obtained by recognizing that the incremental strain is always given by the fractional increase in length $d\ell/\ell$, where ℓ is the current length of the material as it is being deformed. If the initial length is ℓ_0 and the final, deformed length is ℓ_f, then the total true strain, e_1^{true} is obtained by integrating the incremental strains accumulated throughout the entire deformation history:

$$e_1^t = \int_{\ell_0}^{\ell_f} \frac{d\ell}{\ell} = \ln \ell \Big|_{\ell_0}^{\ell_f} = \ln\left(\frac{\ell_f}{\ell_0}\right) = \ln \lambda_1 \tag{62.29}$$

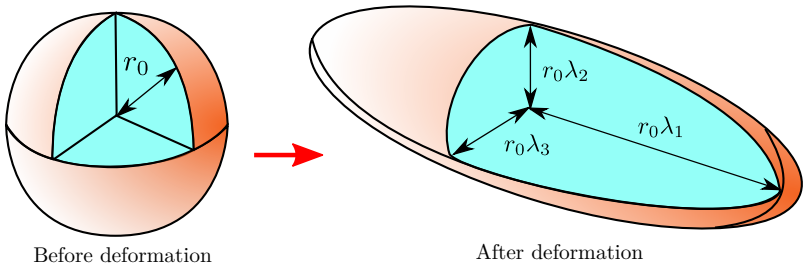


Figure 62.12: Unit sphere deformed into a strain ellipsoid with dimensions of λ_1, λ_2, λ_3.

The extension ratios provide a useful description of the strain for both small and large values of the strain. A material with isotropic mechanical properties has the same coordinate axes for the principal stresses and the principal strains.

62.6 Deformation Modes

Now that we've formally defined stress and strain we can give some specific examples where these definitions are used, and begin to define some elastic constants. We'll begin with the two most fundamental deformation states: simple shear and hydrostatic compression. These are complementary strain states - for an isotropic material simple shear changes the shape but not the volume, and hydrostatic compression changes the volume but not the shape. We'll eventually show that for an isotropic material there are only two independent elastic constants, so if we know how an isotropic material behaves in response to these two stress states, we have a complete understanding of the elastic properties of the material.

62.6.1 Simple Shear

Simple shear is a two dimensional strain state, which means that one of the principal strains is zero (or one of the principal extension ratios is 1).

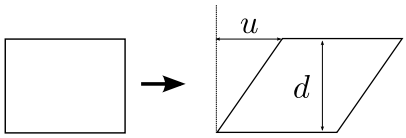


Figure 62.13: Schematic representation of simple shear.

The stress tensor looks like this:

$$\sigma = \begin{bmatrix} 0 & \sigma_{xy} & 0 \\ \sigma_{xy} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{62.30}$$

From the definition of the engineering shear strain (Eq. 62.23) we have:

$$e_{xy} = \frac{u}{d} \tag{62.31}$$

We need to divide the engineering shear strains by 2 to get the tensor strains, so we get the following:

$$\epsilon = \begin{bmatrix} 0 & e_{xy}/2 & 0 \\ e_{xy}/2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{62.32}$$

Theshear modulus is simply the ratio of the shear stress to the shear strain.

$$G\,(\text{shear modulus}) = \frac{\text{Œ}_{xy}}{\text{e}_{xy}} \tag{62.33}$$

Note that the volume of the material is not changed, but it's shape has. In very general terms we can view the shear modulus of a material as a measure of its resistance to a change in shape under conditions where the volume remains constant.

62.6.2 Simple Shear and the Mohr's Circle Construction for Strains

Mohr's circle for strain looks just like Mohr circle for stress, provided that we use the appropriate tensor components. That means that we need to plot $e_{xy}/2$ on the vertical axis and the normal strains on the vertical axis, as shown in Figure 62.14 (where we have used the common notation for the simple shear geometry, with $e_{xy} = \gamma$). One thing that we notice from this plot is that that γ is simply given by the difference between the two principal strains:

$$\gamma = e_1^p - e_2^p = \lambda_1 - \lambda_2 \tag{62.34}$$

For simple shear this relationship is valid, even for large strains, even though there are other aspects of the Mohr's circle construction that no longer work at large strains.

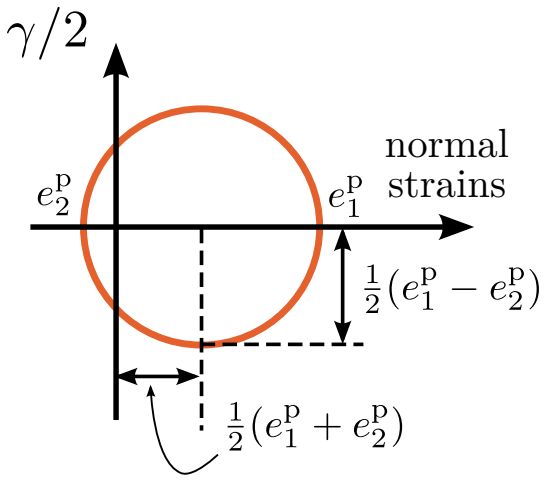


Figure 62.14: Mohr's circle construction for SMALL strains.

62.6.3 Hydrostatic Compression

The bulk compressive modulus, K_b, of a material describes its resistance to a change in density. Formally, it is defined in terms of the dependence of the volume of the material on the hydrostatic pressure, p:

$$K_b = -V\frac{dp}{dV} \tag{62.35}$$

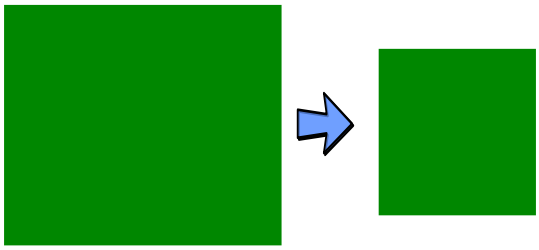


Figure 62.15: Hydrostatic Compression.

62.6.4 Uniaxial Extension

In this case we apply a normal stress along one direction, which we define as the 3 direction so that the stress tensor looks like this:

$$\sigma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix} \tag{62.36}$$

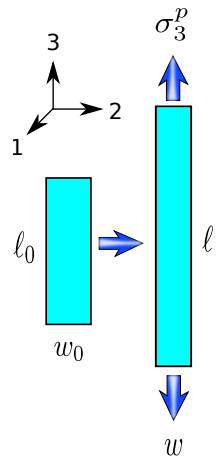


Figure 62.16: Uniaxial tensile deformation.

We can measure two separate strains from this experiment: the longitudinal strain in the same direction that we apply the stress, and the transverse strain, e_{22}, measured in the direction perpendicular to the applied stress (we'll assume that the sample is isotropic in the 1-2 plane, so $e_{11} = e_{22}$. The strains are given by the fractional changes in the length and width of the sample:

$$e_{33} = \frac{\Delta \ell}{\ell_0} \tag{62.37}$$

$$e_{22} = \frac{\Delta w}{w} \tag{62.38}$$

From these strains we can define Young's modulus, E, and Poisson's ratio ν:

$$E = \sigma_{33}/e_{33} \tag{62.39}$$

$$\nu = -e_{22}/e_{33}$$

62.6.5 Longitudinal Compression

A final strain state that we will consider is longitudinal compression. In this state all of the compression is in one direction, which we will specify as the 3 direction. The strains in the other two direction are constrained to be zero, so the strain state is as follows:

$$e = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & e_{33} \end{bmatrix}$$

Note that the strain state is similar to that of uniaxial extension or compression (Figure 91.4), but in the current case we have a single nonzero strain instead of a single non-zero stress. Finite values of σ_{11} and σ_{22} must exist in order for sample in order for this strain state to be maintained, but we're not going to worry about those for now. Instead, we'll use the following relationship for the longitudinal elastic modulus, E_ℓ which is the ratio of σ_{33} to e_{33} for this strain state:

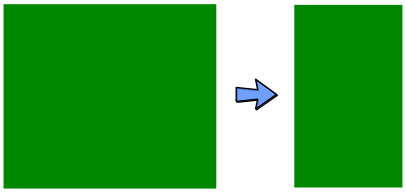


Figure 62.17: Longitudinal Compression

Changes shape and volume, so the relevant modulus involves G and K_b. For an isotropic material:

$$E_\ell = \frac{\sigma_{33}}{e_{33}} = K_b + \frac{4}{3}G \tag{62.40}$$

62.7 Representative Moduli

A few typical values for G and K are listed in Table on this page. Liquids do not have a shear modulus, but they do have a bulk modulus.

Table 62.1: Representative elastic moduli for different materials.

Material	G (Pa)	K_b (Pa)
Air	0	1.0x10 ^{5}
Water	0	2.2x10 ^{9}
Jello	$\approx 10^4$	2.2x10 ^{9}
Plastic	$\approx 10^9$	$\approx$2x10 ^{9}
Steel	8x10 ^{10}	1.6x10 ^{11}

62.8 Case Study: Speed of Sound

The speed of sound, or sound velocity, v_{sound} is actually a mechanical property. Is is related to a modulus, M in the following way:

$$v_{sound} = \sqrt{\frac{M}{\rho}} \tag{62.41}$$

Here ρ is the density of the material. The modulus that we need to use depends on the type of sound wave that is propagating. The two most common are a shear wave and a longitudinal compressional wave:

- Longitudinal compressional wave: $M = E_\ell$
- Shear wave: $M = G$

In a liquid or gas (like air), $G = 0$ and shear waves cannot propagate. In this case there is a single sound velocity obtained by setting $M = K_b$. For an ideal gas:

$$P = \frac{n}{V}RT \tag{62.42}$$

If the compression is applied slowly enough so that the temperature of the gas can equilibrate, we have:

$$K_b = -V\frac{dP}{dV} = \frac{n}{V}RT = P \tag{62.43}$$

So we expect that for a gas, the compressive modulus just equal to the pressure. The situation is a bit more complicated for gas, since we need to use the adiabatic modulus, which is about 40% higher than the pressure itself. (For a detailed explanation, see the Wikipedia article on the speed of soundhttp://en.wikipedia.org/wiki/Speed_of_sound. The brief explanation is that for sound propagation, the derivative in Eq. 62.43 needs to be evaluated at constant entropy and not constant temperature, because the sound oscillation is so fast that the heat does not have time to escape). With ρ=1.2 kg/m ^{3} and $K = 1.4x10^5$ Pa, we end up with a sound velocity of 344 m/s.

63 Contact Mechanics

In a simple tensile test involving a sample with a uniform cross section, the stresses and strains are both uniform throughout the entire sample. In almost any real application where we care about mechanical properties, this is not the case however. A simple example of this is the case where we press a rigid, cylinder into s soft, compliant material as shown in Figure

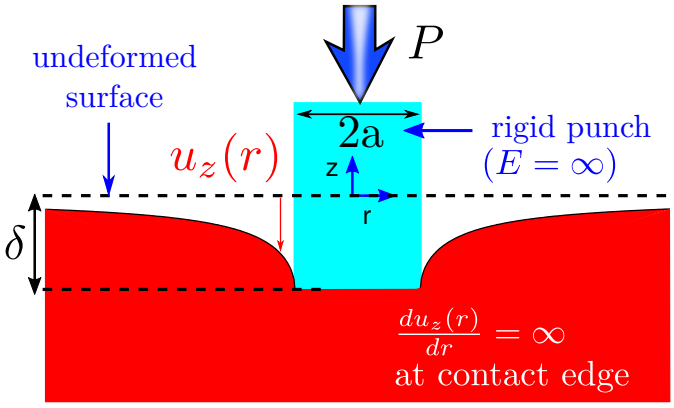


Figure 63.1: Indentation if a soft surface with a rigid, flat-ended cylindrical punch of radius a_0.

63.1 Sign conventions

Sign conventions have a tendency to lead to confusion. This issue is particularly problematics in contact mechanics because compressive loads are considered to be positive, but a compressive stress is negative. Here's a summary of the sign conventions relevant to our treatment of contact mechanics:

- P (force): a positive force is compressive
- δ (displacement): a positive displacement is compressive
- σ (stress): a positive stress is tensile
- e (strain): a positive strain is tensile

In order not to get too hung up in issues related to the sign, we define P_t and δ_t as the tensile loads and displacements:

$$\begin{aligned} P_t &= -P \\ \delta_t &= -\delta \end{aligned} \tag{63.1}$$

63.2 Flat Punch Indentation

(Note: Many of issues presented here are discussed in more detail in a published review article: see ref.[30]). Consider a flat-ended cylindrical punch with a radius of a in contact with another material of thickness , h, as shown schematically in Figure 63.2. The material being indented by a punch rests on a rigid substrate. We are interested in the compressive force, P, that accompanies a compressive displacement, δ, applied to the indenter.

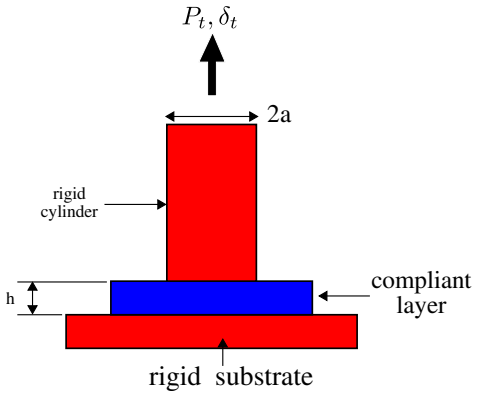


Figure 63.2: Flat punch contact geometry. For an elastic half space, $h \to \infty$.

63.2.1 Flat Punch: Approximate Result for an elastic half space.

For an elastic half space ($h \to \infty$), the strain field under the indenter is nonuniform. The largest strains are confined to a region with characteristic dimensions defined by the punch radius, a. We can get a very approximate expression for the relationship between the compressive load, P, and the compressive displacement, δ from the following approximate concepts:

- The average strain in the highly deformed region of the sample must increase linearly with δ. Because strain is dimensionless we need to divide δ by some length scale in the problem to get a strain. For an elastic half space with $h = \infty$ the only length scale in the problem is the punch radius, a. So the strain fields must depend on δ/a. We'll take this one step further and assume that δ/a is an average in a region of volume $\approx a^3$ under the punch:

$$e_{avg} = -\delta/a \tag{63.2}$$

- The average contact stress, σ_{avg} under the punch can be quantitatively defined by dividing the compressive load by the stress:

$$\sigma_{avg} = -P/\pi a^2 \tag{63.3}$$

- An approximate relationship between P and δ is obtained by assuming that the stress and strain are related through the elastic modulus,*i.e.* $\sigma_{avg} = Ee_{avg}$. Using the equations above for the compliance, C_0:

$$C_0 \equiv \frac{\delta}{P} \approx \frac{1}{\pi E a} \tag{63.4}$$

63.2.2 Flat punch - Detailed Result

In a more general situation both of the contacting materials (the indenter and the substrate) may deform to some extent, so the compliance depends on the properties of both materials. If the materials have Young's moduli of E_1 and E_2 and Poisson's ratios of ν_1 and ν_2, then the expression for C_0 is:

$$C_0 \equiv \frac{\delta}{P} = \frac{\delta_t}{P_t} = \frac{1}{2E_r a} \tag{63.5}$$

where E_r is the following reduced modulus:

$$\frac{1}{E_r} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \tag{63.6}$$

Note that for a stiff indenter, stiff indenter ($E_2 \gg E_1$): $E_r = \frac{E_1}{1-\nu_1^2}$. This is the plane strain modulus that was used previously to relate the energy release rate, $\mathcal{G}$, to the stress intensity factor, K.

63.2.3 Detachment from a Flat Punch

If adhesive forces cause the punch to stick to the substrate, we can use fracture mechanics to understand the force required for detachment to occur. The situation is as shown in Figure 63.3 for a flat-ended cylindrical punch with a radius of a_0. The surface profile of the substrate (assumed in this case to be an elastic half space,*i.e.*, $h = \infty$) is given by the following expression[22]:

$$u_z = (2\delta_t/\pi)\arcsin(a/r) \tag{63.7}$$

where δ_t is the applied tensile displacement and a is the actual radius of the contact area between the punch and the substrate. In Figure 63.3 we compare the shapes of the surface for the following two cases:

- $a = a_0$: this is the initial contact condition, where the substrate is in contact with the full surface of the indenter.
- $a = a_0/2$: the contact radius has been reduced to half its initial value

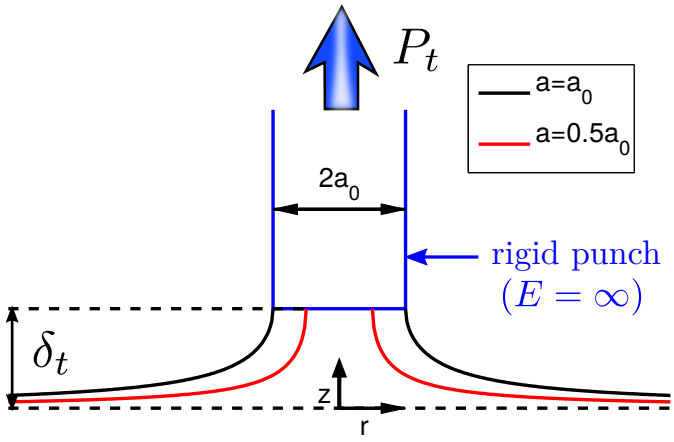


Figure 63.3: Surface profile under a flat punch, from Eq. 63.7.

The decrease in a from a_0 to $a_0/2$ is accompanied by a decrease in the stored elastic strain energy, and this strain energy is what drives the decrease in the contact area. The edge of the contact is a crack, which advances as a decreases. We can use Eq. 64.17 for the energy release rate to obtain the following:

$$\mathcal{G} = -\frac{P_t^2}{2}\frac{dC}{dA} = -\frac{P_t^2}{4\pi a}\frac{dC}{da} \tag{63.8}$$

where we have assumed that the contact area remains circular, with $A = \pi a^2$. Not that A in this expression is the contact area between the indenter and the substrate, and NOT the crack area. The negative sign in Eq. 63.8 emerges from the fact an decrease in contact area corresponds to an equivalent increase in the crack area, so we have:

$$\frac{dC}{dA} = -\frac{dC}{dA_c} \tag{63.9}$$

With $C = C_0 = 1/2E_r a$ (Eq. 63.5) we obtain the following expression for $\mathcal{G}$:

$$\mathcal{G} = \frac{P_t^2}{8\pi E_r a^3} \tag{63.10}$$

In some situations it is more convenient to express the energy release rate in terms of the tensile displacement, δ_t, which we do by using Eq. 63.5 to substitute δ_t for P_t:

$$\mathcal{G} = \frac{E_r \delta_t^2}{2\pi a} \tag{63.11}$$

It is useful at this point to make the following general observations:

- In fracture mechanic terms the contact edge is an interfacial crack. An advancing crack corresponds to a reduction in a, and a receding crack corresponds to an increase in a.
- In general, $\mathcal{G}$ determined by applied load and the geometry.
- $\mathcal{G}_c$ is a property of the interface. The crack moves forward (a decreases) when the value of $\mathcal{G}$ determined by the loading conditions exceeds $\mathcal{G}_c$. The detachment criterion is that the energy release rate, $\mathcal{G}$, is equal to the critical energy release rate, G_c, when the applied tensile force is equal to the critical pull-off force, P_c:

  $$\mathcal{G}_c = \frac{P_c^2}{8\pi E_r a^3} \tag{63.12}$$

  which can be rearranged to give the following for the pull-off force:

  $$P_c = \left(8\pi E_r a^3 \mathcal{G}_c\right)^{1/2} \tag{63.13}$$

- Detachment from of a flat punch from an elastic half space ($a/h = 0$) is unstable for load controlled (constant P_t) OR displacement controlled (constant δ_t) conditions. In each of these cases the contact radius, a, is equal to the punch radius, a_0, at the beginning of an experiment. As the tensile load and displacement are increased,

63.2.4 Detachment: Size Scaling

An interesting aspect of Eq. 63.13 is that the pull-off force scales with $a^{3/2}$, whereas the punch cross sectional area scales more strongly with a ($A = \pi a^2$). This behavior has some interesting consequences, which we can obtain by dividing P_c by the punch cross sectional area to obtain a critical pull-off stress, σ_c:

$$\sigma_c = \frac{P_c}{\pi a^2} = \left(\frac{8 E_r \mathcal{G}_c}{\pi a}\right)^{1/2} \tag{63.14}$$

Note that the detachment stress increases with decreasing punch size.

Figure 63.4: Electron micrograph of Gecko setae.

Let's put in some typical numbers to see what sort of average stresses we end up with:

- $E_r \approx 10^9$ Pa (typical of glassy polymer)
- $\mathcal{G}_c \approx 0.1\,\mathrm{J/m^2}$ (twice the surface energy of a typical organic material)

- $a \approx 100$ nm (smallest reasonably possible value)
- $\sigma_c \approx 50$ MPa

In order for stresses to be obtained, the pillars must be separated so that the stress fields in substrate don't overlap. This decreases the maximum detachment stress from the previous calculation by about a factor of 10, so that the largest stress we could reasonably expect is ≈5 MPa. That's still a pretty enormous stress, corresponding to 500 N (49 Kg) over a 1 cm ^{2} area. This is still difficult to achieve, however, because it requires that the pillar array be extremely well aligned with the surface of interest, a requirement that is very difficult to meet in practice. Nevertheless, improvements in the pull-off forces can be realized by structuring the adhesive layer, and this effect is largely responsible for the adhesive behavior of geckos and other creatures with highly structured surfaces.

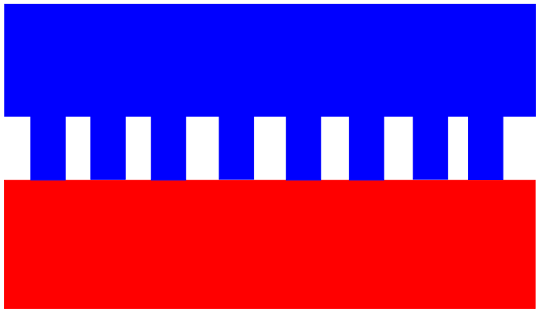


Figure 63.5: Schematic representation of an array of pillars in contact with a flat surface.

63.2.5 Stress Distribution under the Flat Punch

For the flat punch case, the following analytic expression exists for the shape of the normal tress distribution directly under the punch (the plane with $z = \delta_t$ in Figure 63.3):

$$\frac{\sigma_{zz}}{\sigma_{avg}} = 0.5\left(1 - (r/a)^2\right)^{-1/2} \tag{63.15}$$

Here the average stress, σ_{avg} is defined as follows:

$$\sigma_{avg} \equiv \frac{P}{\pi a^2} \tag{63.16}$$

Note that σ_{zz} diverges at the edge of the punch ($r = a$). We know that this must be the case because of the stress concentration that exists at the edge of the punch. To get an expression for stress concentration, K_I at this edge, we first define d as the distance from the punch edge:

$$d \equiv a - r \tag{63.17}$$

Substituting d for r in Eq. 63.15 gives:

$$\frac{\sigma_{zz}}{\sigma_{avg}} = \frac{1}{2}\left[2d/a - (d/a)^2\right]^{-1/2} \tag{63.18}$$

The stress intensity factor describes the stress field near the contact edge, where a/d is small. We can ignore the term involving the square of a/d to obtain the following expression σ_{zz} that is valid near the contact edge:

$$\frac{\sigma_{zz}}{\sigma_{avg}} \approx \frac{1}{2^{3/2}}\left(\frac{d}{a}\right)^{-1/2} \tag{63.19}$$

by comparing to Eq. 64.5 for K_I we obtain the following:

$$K_I = \frac{1}{2}\sigma_{avg}\left(\pi a\right)^{1/2} \tag{63.20}$$

Now we can use the following equation to obtain the following expression for $\mathcal{G}$ (assuming $K_{II} = K_{III} = 0$):

$$\mathcal{G} = \frac{K_I^2}{2E_r} = \frac{\pi a \sigma_{avg}^2}{8E_r} = \frac{P_t^2}{8\pi E_r a^3} \tag{63.21}$$

This is the same result that we got above (see Eq. 63.10), so everything checks out okay. Note that the extra factor of two in the relationship between $\mathcal{G}$ and K_I in Eq. 63.21 comes from the fact the punch is rigid, so it has no stored elastic energy. Because elastic energy is stored only on one side of the interface, the value of $\mathcal{G}$ for crack propagation at the interface with the rigid indenter (Figure 63.3) is half the value of $\mathcal{G}$ for a crack propagating through an elastic material (Figure 64.5, for example).
Eq. 64.22

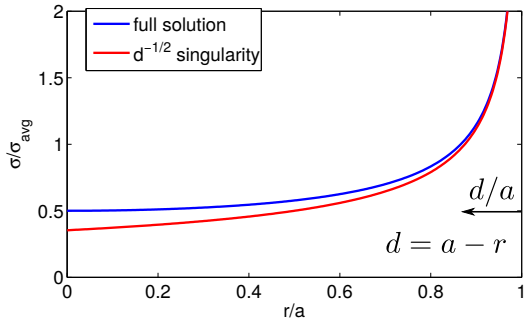


Figure 63.6: Comparison of the full solution for the flat punch contact stresses (Eq. 63.15) with the $d^{-1/2}$ singularity obtained from K_I (Eq. 63.20).

63.2.6 Thickness Effects

When the thickness, h, of the compliant layer between a rigid cylindrical punch punch and a rigid, flat substrate decreases, the mechanics change in a way that makes it more difficult to pull the indenter out of contact with the compliant layer. In general we can write the compliance of the material in the different way:

$$C = \frac{1}{2E_r a} f_C \tag{63.22}$$

For an elastic half space ($h \to \infty$) $f_C = 1$. The factor f_C accounts for changes in the compliance due to the decreased thickness of the layer. In general it depends on Poisson's ratio for the compliant layer and the confinement ratio, a/h (the ratio of the punch radius to the thickness of the layer). For an incompressible compliant layer with $\nu = 0.5$ we have:

$$f_C = \left[1 + 1.33\left(a/h\right) + 1.33\left(a/h\right)^3\right]^{-1} \tag{63.23}$$

We can use Eq. 63.8 to get an expression for the energy release rate, which can be written in the following form:

$$\mathcal{G} = -\frac{P_t^2}{4\pi a}\frac{dC}{da} = \frac{P_t^2}{8\pi E_r a^3} f_{Gp} \tag{63.24}$$

We can use the the fact that $P_t = \delta_t / C$ to get a similar expression for $\mathcal{G}$ in terms of the tensile displacement, δ_t:

$$\mathcal{G} = \frac{E_r \delta_t^2}{2\pi a} f_{G\delta} \tag{63.25}$$

The confinement functions f_{Gp} and $f_{G\delta}$ are both equal to one for $a/h = 0$ and are plotted as a function of a/h for $\nu = 0.5$ in Figure 63.8. A practical consequence of the decrease in f_{Gp} with decreased h is that a larger tensile force is required in order to remove the cylinder from its contact with the compliant layer. With a small value of f_{Gp}, a larger tensile load needs to be applied in order for $\mathcal{G}$ to exceed the critical energy release rate, $\mathcal{G}_c$.

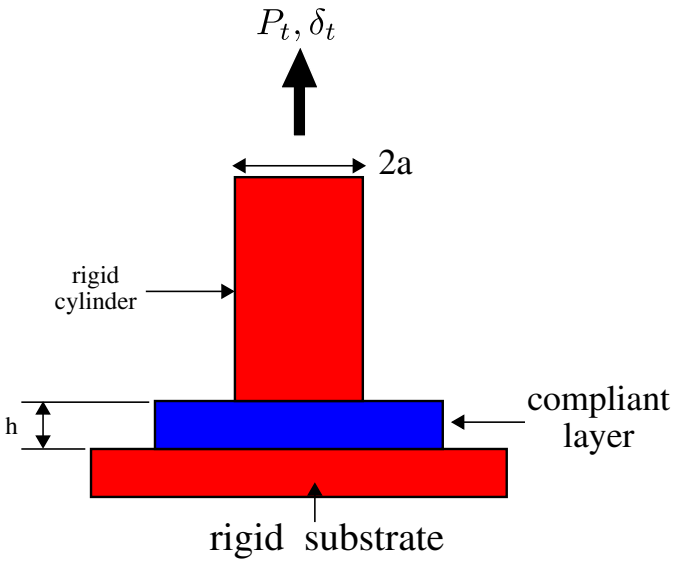


Figure 63.7: A thin, compliant layer being indented with a rigid, cylindrical punch.

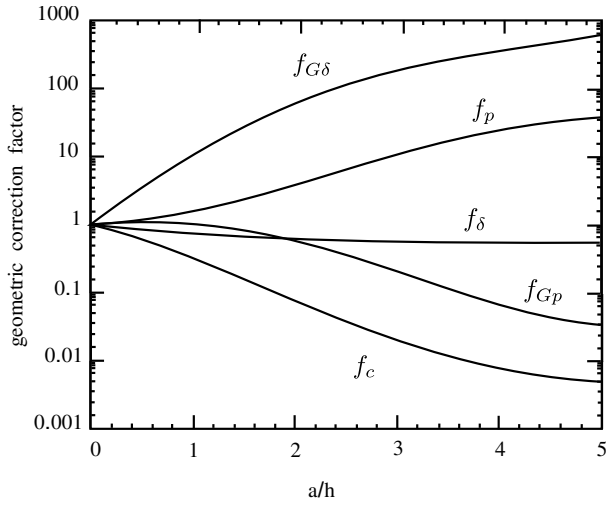


Figure 63.8: Geometric correction factors for the flat punch geometry.

Decreasing the thickness of the compliant layer also changes the distribution of normal stresses in contact with the layer. These normal stresses are plotted for different values of a/h in Figure 63.9.

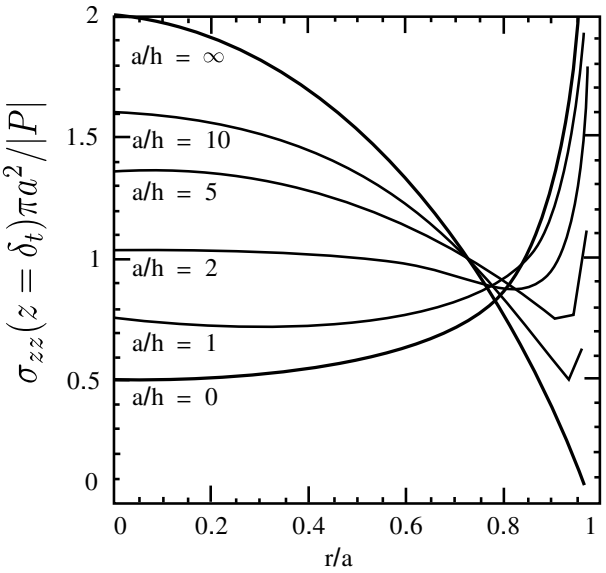


Figure 63.9: Dependence of the stress distribution under a flat punch for different values of the confinement ration, a/h.

- Max. stress in center for thin, incompressible layers ($\nu = 0.5$).
- Decrease in edge stress singularity (decrease K_I) for thin layers.

Since for very thin layers failure does not initiate from the edge because the driving force for this 'edge crack' vanishes as h becomes very thin. Instead, failure initiates from small defects within the central part of the contact zone where σ_{zz} is the highest. The mode I stress intensity factor for a circular, internal crack of radius a_c is given by the following expression:

$$K_I = \frac{2}{\sqrt{\pi}}\sigma_{zz}\sqrt{a_c} \tag{63.26}$$

Note that the prefactor in this expression is slightly different than what is given in Eq. 64.7 because of the different crack geometries. Eq. 64.7 is for a rectangular crack and Eq. 63.26 is for a circular crack. The energy release rate for the circular crack is given by using the relationship between K_I and $\mathcal{G}$ valid for a mode I crack at the interface between compliant and rigid materials (Eq. 63.21):

$$\mathcal{G} = \frac{K_I^2}{2E_r} = \frac{2a_c\sigma_{zz}^2}{\pi E_r} \tag{63.27}$$

63.3 Contact of Paraboloids

63.3.1 Non-Adhesive Case

Suppose that the indenter is not flat, but has a parabolic profile that can be described by the following expression:

$$z = A_p r^2 \tag{63.28}$$

Here z is the vertical distance from the apex of the parabola, r is the radial distance from symmetry axis for the paraboloid and A_p is a constant that defines the shape of the paraboloid. A sphere has a parabolic shape near the apex, which can be seen by considering the equation for a sphere of Radius R that has it's center at $r = 0$, $z = R$:

$$r^2 + (z - R)^2 = R^2 \tag{63.29}$$

Solving Eq. 63.29 for z gives:

$$z = R\left(1 \pm \sqrt{1 - (r/R)^2}\right)$$

For small x, $\sqrt{1-x} \approx 1 - x/2$, so for $r \ll R$ we have:

$$z = \frac{r^2}{2R} \tag{63.30}$$

where we have taken the solution with the smaller value of z, corresponding to the bottom of the sphere. From a comparison of Eqs.63.29 and63.30, we see the paraboloid is a good approximation for the shape of a sphere, with the sphere radius given by $1/2A$. For this reason we use R instead of A to characterize the parabolic shape, since the results can be applied to contact of spheres, provided that the the contact dimensions are much smaller than R. (Generally we are in good shape as long as $r/R < 4$.

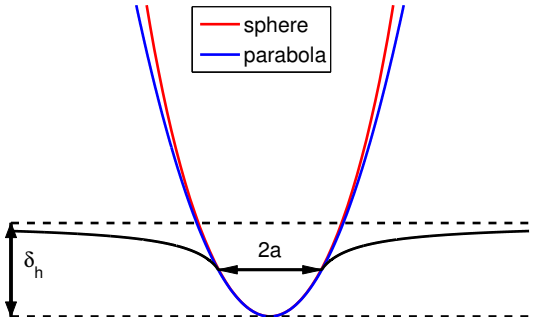


Figure 63.10: Non-adhesive contact of a rigid, parabolic indenter into an elastic material.

The compressive a rigid parabolic indenter into the surface of the material (δ_h in Figure 63.10) is given by the following expression:

$$\delta_h = a^2/R \tag{63.31}$$

Note that this is a completely geometric relationship that does not depend on the modulus of the material that is being indented. The compressive force required to establish a contact radius of a is referred to as P_h, and is given by the following expression:

$$P_h = \frac{4E_r a^3}{3R} \tag{63.32}$$

We can use Eq. 63.31 to substitute for a and obtain a relationship between P_h and δ_h:

$$P_h = \frac{4E_r}{3} R^{1/2} \delta_h^{3/2} \tag{63.33}$$

The assumption here is that there is no adhesion between the indenter and the substrate,*i.e.*, $\mathcal{G} = K_I = 0$. The fact that there is no stress concentration at the interface is consistent with the fact that the slope of the surface profile of the compliant material is continuous at $r = a$. This surface profile is plotted in Figure 63.10 and is given by the following expression:[22]

$$u_z = \frac{\delta}{\pi} \left\{ \left(2 - (r/a)^2\right) \arcsin\left(a/r\right) + (r/a) \left(1 - (a/r)^2\right)^{1/2} \right\} \tag{63.34}$$

63.3.2 Effects of Adhesion on Contact

The easiest way to understand the effect of adhesion on the contact between a parabolic is to consider a hypothetical situation where we turn off the adhesion and bring the indenter into contact with the surface, resulting in the deformation illustrated in Figure 63.10. Now we we turn on the adhesion, and begin retracting the indenter from the surface, maintaining a fixed projected contact radius a. The situation for the case where we have retracted the tip to the point where the tip apex is level with the undeformed surface ($\delta_t = 0$) is illustrated in Figure 63.11. The compressive load is decreased by P_h by an amount given by the difference between the actual compressive displacement, δ, and the non-adhesive displacement, δ_h (given by Eq. 63.31):

$$\delta = \delta_h + C_0 \left(P - P_h \right) \tag{63.35}$$

where C_0 is the compliance given by Eq. 63.5. Using this expression for C_0, Eq. 63.31 for δ_h and Eq. 63.32 for P_h gives:

$$\delta = \frac{a^2}{3R} + \frac{P}{2E_r a} \tag{63.36}$$

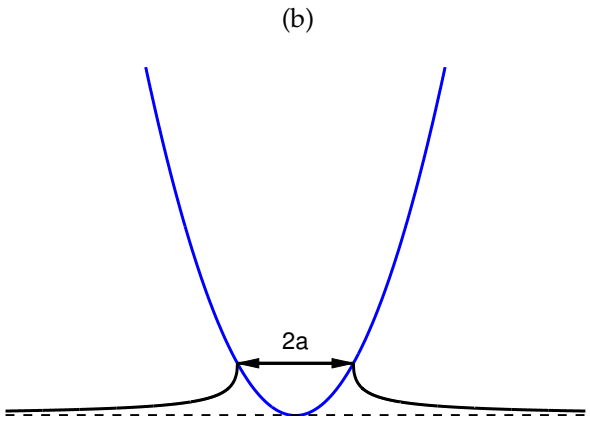


Figure 63.11: An example of the surface profile for adhesive contact for the case where $\delta_t = 0$.

We obtain the expression for the energy release rate for curved object in contact with surface in a way that is very similar to what we did for the flat punch in Section 63.2.3. The only difference is that in the absence of adhesion we need to apply a compressive load, P_h (given by Eq. 63.32):

$$\mathcal{G} = -\frac{(P_t + P_h)^2}{2}\frac{dC}{dA} = \frac{(P_t + P_h)^2}{8\pi E_r a^3} \tag{63.37}$$

This equation can be rearranged to give a^3 as a function of the compressive load, P ($P = -P_t$), to give an expression that was derived in 1971 by Johnson, Kendall and Roberts and commonly referred to as theJKR equation:

$$a^3 = \frac{3R}{4E_r}\left(P + 3\pi\mathcal{G}R + \left(6\pi\mathcal{G}RP + (3\pi\mathcal{G}R)^2 \right)^{1/2} \right) \tag{63.38}$$

63.4 Indentation with Berkovich Trips

Parabolic tips are often used in measurements of adhesion or of the elastic properties of materials. For Hardness measurements tips with sharp corners are more commonly used. One example is theBerkovich tip shown in Figure 63.12.

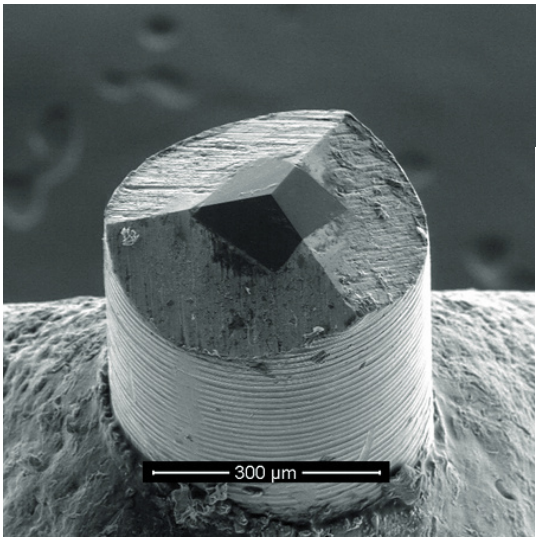


Figure 63.12: Geometry of a Berkovich tip commonly used in indentation experiments. The angle, a, is 65.35°.

The hardness, H, of a material is given by the ratio of the load to the projected contact area of the non-recoverable indent made in the material by the indenter. In our case we obtain the hardness from the maximum load, P_{max} (illustrated in Figure 63.13), and from the corresponding projected area, A, of the hardness impression:

$$H = \frac{P_{max}}{A} \tag{63.39}$$

The projected area is related to the contact depth, δ_c, by a relationship that depends on the shape of the indenter[**?**]. For a Berkovich tip the appropriate relationship is:

$$A = 24.5\delta_c^2 \tag{63.40}$$

The procedure for determining the contact depth was developed by Oliver and Pharr, where the following expression is used to estimate the contact depth:

$$\delta_c = \delta_{max} - 0.75\frac{P_{max}}{S} \tag{63.41}$$

where δ_{max} is the maximum penetration depth of the indenter tip and S is the contact stiffness, determined experimentally as the initial slope of the linear portion of unloading curve (see Figure 63.13). From the measured values of S, P_{max} and δ_{max}, we use Equations63.40 and63.41 to determine A. The reduced modulus is then obtained from the following expression for the contact stiffness, assuming a value for the contact stiffness that is the same for a circular contact of the same area:

$$S = \frac{2}{\sqrt{\pi}} E_r \sqrt{A} \tag{63.42}$$

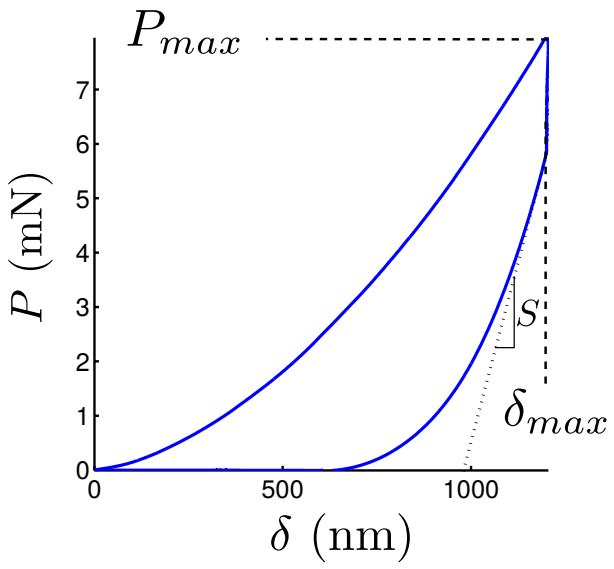


Figure 63.13: Typical load-displacement curve for indentation of the polyester resin used to embed the paint samples, labeled to illustrate the values of P_{max}, h_{max} and S.

64 Fracture

The stress-strain behavior for a many material can exhibit a range of phenomena, depending on the temperature. This is particularly true of many polymers, which can show the range of behaviors in a uniaxial tensile test shown in Figure 64.1. While not all of these behaviors are necessarily observed in the same material, the following general regimes can often be identified:

- T_1: Brittle behavior
- T_2: Ductile behavior (yield before fracture)
- T_3: cold drawing (stable neck)
- T_4: uniform deformation

Here we are concerned with brittle behavior (T_1), or in some cases situations where there is a small degree of ductility in the sample (T_2).

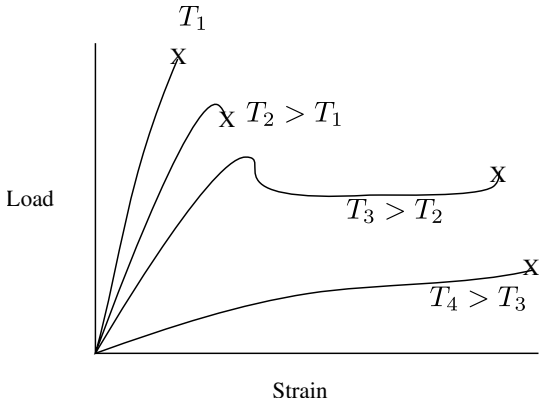


Figure 64.1: Typical generic temperature behavior at different temperatures.

64.1 Fracture Modes

Different fracture modes are defined by the relationship between the applied stress and the crack geometry. These are illustrated schematically in Figure 64.2 Fracture of a homogeneous material fracture generally occurs under Mode I conditions, and this is the most important condition. Mode II conditions, where a shear stress is applied in the direction perpendicular to the crack front, is often important for interfacial fracture, including the adhesive bonding of materials with different properties. Mode III is generally not important for our purposes.

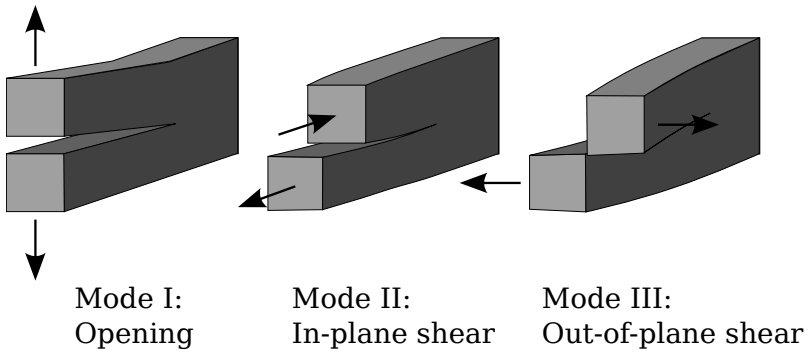


Figure 64.2: Fracture Modes .

64.2 Brittle behavior (fracture mechanics)

http://en.wikipedia.org/wiki/Fracture_mechanics
Mode I crack (stress normal to crack)
2 equivalent approaches:

(1) Irwin model (stress-based approach)

(2) Griffith model (energy-based approach)

64.2.1 The Irwin Model and the Stress Concentration

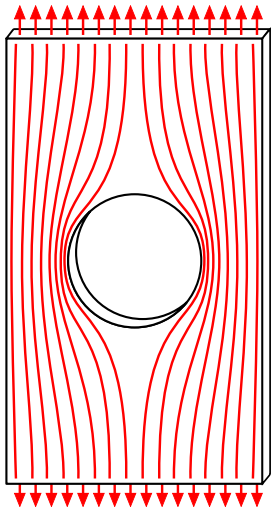


Figure 64.3: Force lines around a circular defect.

For an ellipse of with axis a_c perpendicular to the applied stress and axis b_c parallel to the applied stress (see Figure 64.4), the point of maximum stress is given by the following expression:

$$\sigma_{max} = \sigma_0 \left(1 + 2\frac{a_c}{b_c}\right) \tag{64.1}$$

We can also write this in terms of the radius of curvature of the ellipse, ρ_c, at the point of maximum stress:

$$\rho_c = \frac{b_c^2}{a_c} \tag{64.2}$$

Combination of Eqs.64.1 and64.2 gives:

$$\sigma_{max} = \sigma_0 \left(1 + 2\sqrt{a_c/\rho_c}\right) \tag{64.3}$$

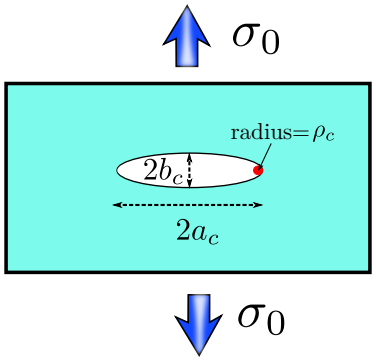


Figure 64.4: Elliptical crack with a crack tip radius of curvature, ρ_c.

64.2.2 Stress Intensity Factor

Consider a planar crack in the x-z plane, as shown conceptually Figure 64.5. The stress in the vicinity of the crack tip can be expressed in the following form:

$$\sigma = \frac{K}{\sqrt{2\pi d}} f(\theta) \tag{64.4}$$

where d is the distance from the crack tip and $f(\theta)$ is some function of the angle θ that reduces to 1 for the direction directly in front of a crack ($\theta = 0$). Different functional forms exist for $f(\theta)$ for the different stress components σ_{xx}, σ_{yy}, etc. The detailed stress fields depend on the loading mode (Mode I, II or II, or some combination of these), and the corresponding stress fields are specified by the appropriate value of K (K_I or mode I, K_{II} for mode II or K_{III} for mode III.

Figure 64.5: Cartesian (a) and polar (b) coordinate axes use d to define stresses in the vicinity of a crack tip.

Mode I loading

The stresses in the vicinity of a mode I crack are given by the following[38]:

$$\begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{xy} \end{pmatrix} = \frac{K_I}{\sqrt{2\pi d}} \cos\frac{\theta}{2} \begin{pmatrix} 1 - \sin\frac{\theta}{2}\sin\frac{3\theta}{2} \\ 1 + \sin\frac{\theta}{2}\sin\frac{3\theta}{2} \\ \cos\frac{3\theta}{2}\sin\frac{\theta}{2} \end{pmatrix} \tag{64.5}$$

This compact notation is used to specify the three relevant values of $f(\theta)$. For example, for σ_{xx} we have the following:

$$\sigma_{xx} = \left(K_I/\sqrt{2\pi d}\right)\cos(\theta/2)\left(1-\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right) \tag{64.6}$$

These expressions assume that the crack tip is very sharp, with a very small radius of curvature, ρ_c. If d is comparable to ρ_c, these equations no longer apply. Consider for example, the presence of an internal crack of length a_c and radius of curvature ρ_c in a thin sheet of material, shown schematically in Figure 64.6. In this case the stress at the crack edge is σ_{max} as given by Eq. 64.3. An assumption in the use of Eq. 64.5 is that the stresses are substantially less than σ_{max}. In other words, K describes the stress field close to the crack tip, but still at distances away from the crack tip that are larger than the crack trip radius of curvature, ρ_c.
The mode I stress intensity factor for this geometry is given by the applied stress, σ_0 and the crack length a_c:

$$K_I = \sigma_0\sqrt{\pi a_c} \tag{64.7}$$

For values of d that are substantially larger than ρ_c but smaller than a_c, we can determine the stresses from Eq. 64.5, with K_I as given by Eq. 64.7.

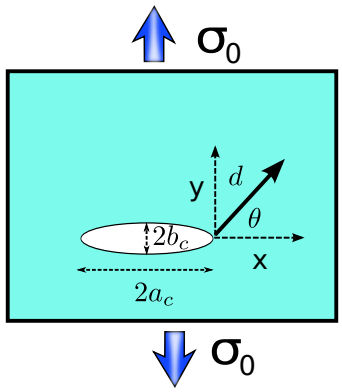


Figure 64.6: An internal crack in a homogeneous solid.

Mode II loading

For mode II loading the crack tip stress fields are given by the following set of expressions[38]:

$$\begin{pmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{xy}\end{pmatrix} = \frac{K_{II}}{\sqrt{2\pi d}}\begin{pmatrix}-\sin\frac{\theta}{2}\left(2+\cos\frac{\theta}{2}\cos\frac{3\theta}{2}\right)\\ \sin\frac{\theta}{2}\cos\frac{\theta}{2}\cos\frac{3\theta}{2}\\ \cos\frac{\theta}{2}\left(1-\sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right)\end{pmatrix} \tag{64.8}$$

It is generally difficult to determine K_{II} in a straightforward way, and finite element methods must often be used to determine it for a given loading condition and experimental geometry. Once K_{II} is known, the crack tip stress fields can be obtained from Eq. 64.8.

Mode III loading

While mode III loading is often encountered in practical applications, it is generally avoided in experiments aimed at assessing the fracture behavior of materials, and is not considered further in this text.

64.2.3 Fracture condition

In the stress-based theory of fracture, the material fails when the stress intensity factor reaches a critical value that depends on the material. For mode I loading, we refer to this critical stress intensity factor as K_{IC}. Setting σ_0 to the fracture stress, σ_f, and setting K_I to K_{IC} in Eq. 64.7 gives:

$$K_{IC} = \sigma_f\sqrt{\pi a_c} \tag{64.9}$$

Rearranging gives:

$$\sigma_f = K_{IC}/\sqrt{\pi a_c} \tag{64.10}$$

So the fracture stress decreases as the flaw size, a_c, increases. This is why a material can appear to be fine, even though small cracks are present in the material. The cracks grow very slowly, but when the reach a critical size for which Eq. 64.10 is satisfied, the material fails catastrophically.
The fracture toughness, K_{IC} has strange units - a stress times the square root of a length. In order to understand where this characteristic stress and the characteristic length actually come from, we need to consider the actual shape of the crack tip. Using Eq. 64.3 we see that the maximum stress in front of the crack tip, σ^f_{max} , at the point of fracture is:

$$\sigma^f_{max} \approx 2\sigma_f\sqrt{a_c/\rho_c} \tag{64.11}$$

where we have assumed that $\sqrt{a_c/\rho_c} \gg 1$, so that we can ignore the extra factor of 1 in Eq. 64.3. Now we can use Eq. 64.9 to substitute $K_{IC}$for σ_f. After rearranging we get:

$$K_{IC} \approx \sigma^f_{max}\frac{\sqrt{\pi}}{2}\sqrt{\rho_c} \approx \sigma^f_{max}\sqrt{\rho_c} \tag{64.12}$$

This expression is really only valid for a crack tip with a well-defined radius of curvature, which is often not the case. Models that aim to predict and understand the fracture toughness of materials are all based on understanding the details of the yielding processes very close to the crack tip, and the resulting crack shape. We'll return to this issue later. For now we can summarize the stress-based approach fracture mechanics as follows:

- With the exception of a very small region near the crack tip, all of the strains are elastic.
- There is a very small plastic zone in the vicinity of a crack tip, with a characteristic dimension, ρ that is determined by the details of the way the material plastically deforms.
- Fracture occurs when the stress field defined by K_I reaches a critical value.

64.3 Griffith Model and energy release rate

Specifying the stress field is the same as specifying the stored elastic energy. One expects that an energy-based fracture criterion could also be derived. This leads us to the second approach, which is generally referred to as the Griffith approach.
Fracture occurs when available energy is sufficient to drive crack forward.

- W= work done on system by external stresses
- U_{el}= elastically stored energy
- $W - U_{el}$= energy available to drive crack forward.

Theenergy release rate, $\mathcal{G}$ describes the amount of energy that is used to move a crack forward by some incremental distance. Formally it is described in the following way:

$$\mathcal{G} = \frac{d}{dA_c}(W - U_{el}) \tag{64.13}$$

where A_c is the crack area. Fracture occurs when the applied energy release rate exceeds a critical value characteristic of the material, defined as the critical energy release rate, $\mathcal{G}_C$. The fracture condition is therefore:

$$\mathcal{G} = \mathcal{G}_c \tag{64.14}$$

The lowest possible value of $\mathcal{G}_c$ is 2 γ where γ is surface energy of the material. That's because the minimal energy to break a material into two pieces is the thermodynamic energy associated with the two surfaces. Some typical values for the surface energy of different materials are listed below (note that 1 mJ/m^2 = 1 erg/cm^2 = 1 dyne/cm).

- Polymers: 20-50 mJ/m^2 Van der Waals bonding between molecules
- Water: 72 mJ/m^2 Hydrogen bonding between molecules
- Metals: ≈1000 mJ/m^2 Metallic bonding

64.3.1 Calculating the Energy Release Rate

We can derive a simple expression for the energy release rate if we assume that the material has a linear elastic response. Consider, for example, an experiment where we apply a tensile force, P_t,to a sample, resulting in a tensile displacement, δ_t, as illustrated in Figure 64.7a. If the material has a linear elastic response, the behavior is as illustrated in Figure 64.7. Suppose that the crack area remains constant as the material is loaded to a tensile force P_t. The sample compliance, C, is given by the slope of the displacement-force curve:

$$C = \left.\frac{d\delta_t}{dP_t}\right|_{A_c} \tag{64.15}$$

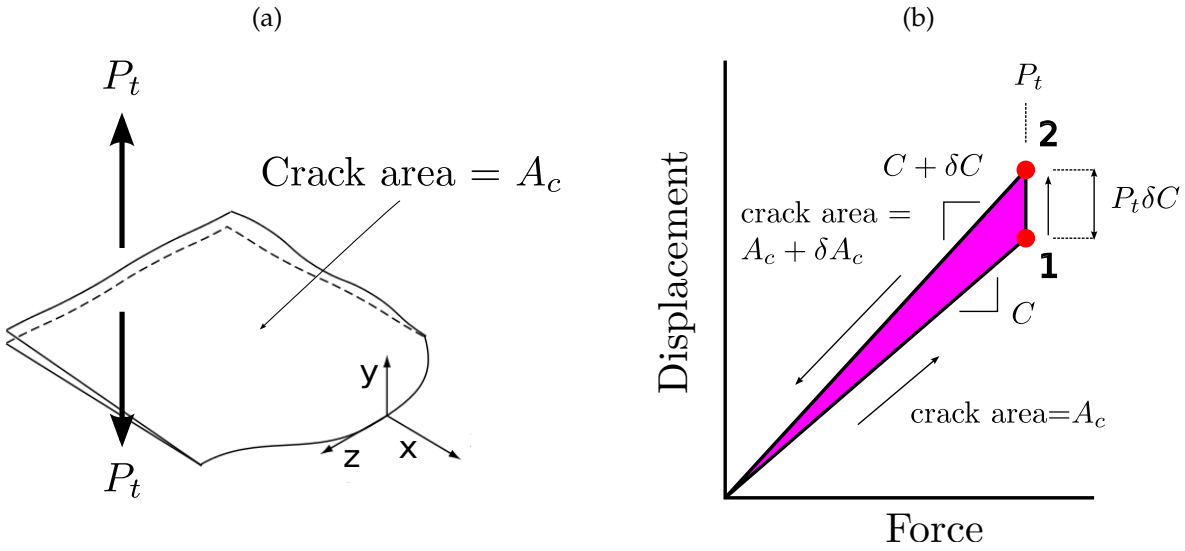


Figure 64.7: Derivation of the compliance expression for $\mathcal{G}$.

Now suppose that the crack area is increased by an amount dA_c while the load remains fixed at P_t,*i.e.* the system moves from point 1 to point two on Figure 64.7b. This increases the compliance by an amount dC, resulting in corresponding increase in the displacement of $P_t\delta C$. If now unload the sample from point 2 back to the origin, the slope of this unloading curve is given by the enhanced compliance, $C + \delta C$. At the end of this loading cycle be have put energy into the sample equal to the shaded area in Figure 64.7b. This is the total work done on the system by the external stresses (W in Eq. 64.13), and given by the following expression:

$$\delta W = \frac{1}{2} P_t^2 \delta C \tag{64.16}$$

Because the sample at the beginning an end of this process is unstrained, we have U_{el} =0. We can therefore take the limit where δA_c becomes very small to replace $\delta W/\delta C$ with the derivative, dW/dC to obtain:

$$\mathcal{G} = \lim_{\delta A_c \to 0} \frac{\delta W}{\delta A_c} = \frac{P_t^2}{2}\frac{dC}{dA_c} \tag{64.17}$$

64.3.2 Double cantilever beam geometry

The double cantilever beam geometry illustrated in Figure 64.8 is a common test used to measure crack propagation in materials. It is commonly used to measure the adhesion between two materials that have been glued together. It consists of to beams, each with width, w, and thickness t. The crack length, a_c in this geometry is the distance between the parts of the beam where the force is applied and the beginning of the region of the sample where the two beams are in contact with one another.

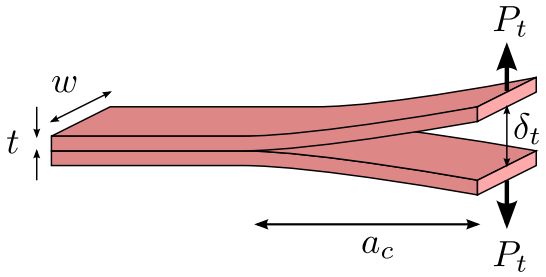


Figure 64.8: Double cantilever beam geometry.

For the double cantilever beam geometry the compliance is given by the following expression:

$$C \equiv \frac{\delta_t}{P_t} = \frac{8a_c^3}{Ewt^3} \tag{64.18}$$

The crack area, A_c is obtained by the crack length, a_c by the width of the sample:

$$A_c = wa_c \tag{64.19}$$

So we have:

$$\frac{dC}{dA_c} = \frac{1}{w}\frac{dC}{da_c}$$

We obtain the following for the energy release rate:

$$\mathcal{G} = \frac{P_t^2}{2w}\frac{dC}{da_c} = \frac{12a_c^2P_t^2}{Ew^2t^3} \tag{64.20}$$

At fixed load, $\mathcal{G}$ increases as the crack length increases - unstable geometry!
use Eq. 64.18 to substitute δ_t for P_t:

$$\mathcal{G} = \frac{3\delta_t^2t^3E}{16a_c^4} \tag{64.21}$$

At a fixed displacement, crack will grow until $\mathcal{G} = \mathcal{G}_c$ and then stop. This is a better way to do the experiment.

64.3.3 General relationship between K and $\mathcal{G}$

The stress intensity factor and the energy release rate are related to one another through the following expression:

$$\mathcal{G} = \frac{K_I^2 + K_{II}^2 + K_{III}^2}{E_r} \tag{64.22}$$

Here E_r is the reduced modulus that is slightly different for plane stress and plane strain conditions:

$$\begin{aligned} E_r &= E \text{ (Plane stress conditions)} \\ E_r &= E^* \text{ (Plane strain conditions)} \end{aligned} \tag{64.23}$$

The quantity E_r is a combination of E and ν that generally appears in plane strain problems:

$$E_r = \frac{E}{1-\nu^2} \tag{64.24}$$

Plane stress conditions generally apply for very thin samples, whereas plane stress conditions apply for thick samples.
The fact that $\mathcal{G} \propto K_I^2$ for a mode I fracture experiment is illustrated in Figure 64.9, which we use to show the relationship between stress and stored elastic energy for an elastically deformed sample. The energy input to the sample up to the point of fracture, which we refer to as U_f, is the area under the stress strain curve:

$$U_f = \frac{1}{2}\sigma_f e_f = \frac{1}{2}\frac{\sigma_f^2}{E_r} \tag{64.25}$$

The stress intensity factor, K_I at the fracture point is proportional to the stress, σ_f, and the strain energy release rate, $\mathcal{G}$, at the point of failure is proportional to the total stored elastic energy, U_f. This means that the following proportionality must hold:

$$\mathcal{G} \propto \frac{K_I^2}{E_r} \tag{64.26}$$

This is consistent with Eq. 64.22, but we need to do a more detailed analysis to get the prefactor exactly right.

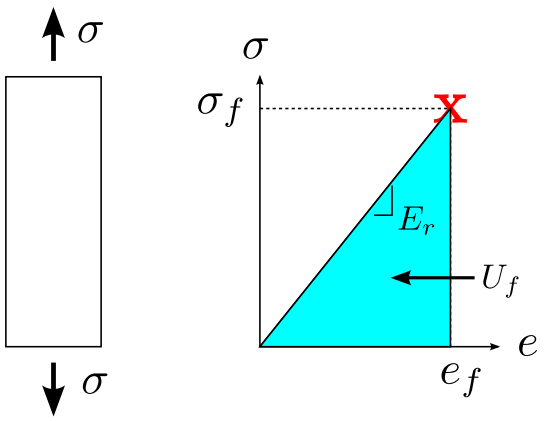


Figure 64.9: Schematic stress/strain curve for a brittle material in the presence of a crack.

64.3.4 Plane strain modulus.

The plane strain modulus, E^* describes the response of a material when it cannot contract in one of the directions that is perpendicular to an applied tensile stress. It's easy to derive this by using the compliance matrix for an amorphous material, which must look like this;

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} = \frac{1}{E}\begin{bmatrix} 1 & -\nu & -\nu & 0 & 0 & 0 \\ -\nu & 1 & -\nu & 0 & 0 & 0 \\ -\nu & -\nu & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2(1+\nu) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2(1+\nu) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(1+\nu) \end{bmatrix}\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \tag{64.27}$$

In writing the compliance matrix this way, we have used the fact that Young's modulus is $1/s_{11}$ and the Poisson's ratio is $-s_{12}/s_{11,}$ so we have $s_{11} = 1/E$ and $s_{12} = -\nu/E$. Suppose we apply a stress in the 1 direction, and require that the strain in the 2 direction is 0. This requires that a non-zero stress develop in the 2 direction. If we assume that $\sigma_3 = 0$, we have:

$$e_2 = \frac{-\nu}{E}\sigma_1 + \frac{\sigma_2}{E} \tag{64.28}$$

If e_2 is constrained to be zero, we have:

$$\sigma_2 = \nu\sigma_1 \tag{64.29}$$

Now we can put this value back into Eq. 64.27 and solve for e_1:

$$e_1 = \frac{1}{E}\left(\sigma_1 - \nu^2\sigma_1\right) \tag{64.30}$$

The plane strain modulus relates σ_1 to e_1, which for the case assumed above ($e_2 = 0$) gives:

$$E_r = \frac{\sigma_1}{e_1} = \frac{E}{1-\nu^2} \tag{64.31}$$

64.4 Fracture Toughness of Materials

In the Griffith (energy-based) model of fracture, material fracture occurs when the applied energy release rate, $\mathcal{G}$, exceeds a threshold value, $\mathcal{G}_C$, which is characteristic of the material. This value is called the critical strain energy release rate, and is a measure of the fracture toughness of the material, just as the critical stress intensity factor is a measure of the fracture toughness. The critical values of K and $\mathcal{G}$ are related to one another through Eq. 64.22. For mode I fracture, $K_{II} = K_{III} = 0$, and this equation reduced to the following:

$$\mathcal{G}_{IC} = \frac{K_{IC}^2}{E_r} \tag{64.32}$$

Note that we have added the subscript 'I' to $\mathcal{G}$ to remind ourselves that this number corresponds to mode I fracture conditions.
Values of $\mathcal{G}_C$ are a bit easier to understand conceptually than values of K_{IC}, since $\mathcal{G}_c$ is simply the energy required to break a sample. We can obtain some estimates of $\mathcal{G}_C$ by making some assumptions about where energy goes. Typical values of $\mathcal{G}_c$ are as follows:

- $\mathcal{G}_C = 2\gamma$ $\left(\approx 0.1\,\mathrm{J/m^2}\right)$ if only work during fracture is to break Van der Waals bonds
- $\mathcal{G}_C \approx 1 - 2\,\mathrm{J/m^2}$ if only work during fracture is to break covalent or metallic bonds across interface
- $\mathcal{G}_C \gg 1\,\mathrm{J/m^2}$ if fracture is accompanied by significant plastic deformation of the sample. For the whole fracture mechanics formulation we are using to be valid, the zone of plastic deformation where this energy dissipation is occurring should be small compared to the overall sample size.

Actual values of $\mathcal{G}_C$ are much larger than these values (see Table64.1) because a significant amount of plastic deformation occurs near the crack tip.

Table 64.1: Typical fracture toughness values (plane strain) for different material.

Material	E (GPa)	K_{IC} (MPa $\sqrt{\mathrm{m}}$)	G_{IC} J/m ^{2}
Steel	200	50	12,000
Glass	70	0.7	7
High M polystyrene or PMMA	3	1.5	750
High Impact Polystyrene	2.1	5.8	16,000
Epoxy Resin	2.8	0.5	100
Rubber Toughened Epoxy	2.4	2.2	2,000
Glass Filled Epoxy Resin	7.5	1.4	300

We can use numbers from Table64.1 to say something about the size of the plastically deformed zone in front of a crack tip. We know that in the elastic region directly in front of a propagating crack, the stress scales as $K_I/\sqrt{2\pi d}$, where d is the distance in front of the crack tip. If the maximum stress is equal to the yield stress, σ_y then this the material must be yielded for values of r which give a stress that exceeds σ_y. A propagating crack has $K_I = K_{IC}$, so if the size of the plastic zone is h_p, we have:

$$\sigma_y \approx \frac{K_{IC}}{\sqrt{2\pi h_p}} \tag{64.33}$$

Rearranging gives:

$$h_p \approx \frac{K_I^2}{2\pi\sigma_y^2} \approx \frac{\mathcal{G}_{IC}}{\sigma_y}\frac{E}{2\pi\sigma_y} \tag{64.34}$$

This formula is approximate because it neglects the fact that yielding of the material actually changes the stress distribution. The details of what is going on in the plastic zone depend on the materials system of interest. Below we give a case study for what happens for some common amorphous, glassy polymers (non-crystalline polymers deformed at temperatures below their glass transition temperature).

64.5 Case Study: Fracture toughness of glassy polymers

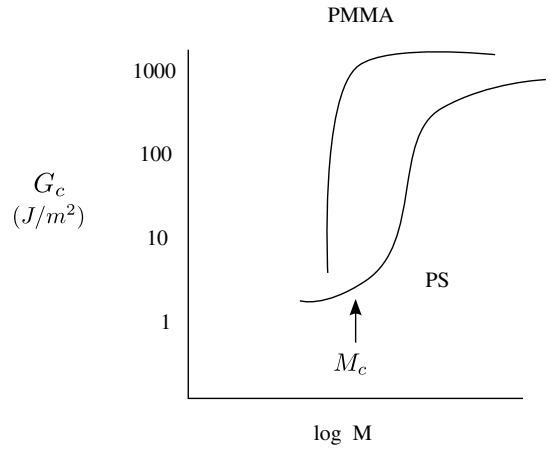


Figure 64.10: Molecular weight dependence of fracture toughness for polystyrene (PS) and poly(methyl methacrylate) (PMMA).

Deformation is significant, but $\mathcal{G}_{Ic}$ is still small compared to other engineering materials.

64.5.1 Deformation Mechanisms

Suppose we do a simple stress strain experiment on polystyrene. Polystyrene deforms by one of two different mechanisms:

(1) Shear band s due to strain softening (decrease in true stress after yield in shear).

(2) Crazing - requires net dilation of sample (fracture mechanism for PS and PMMA).

Figure 64.11

Crazes are load bearing - but they break down to form cracks - failure of specimen.

64.5.2 Crazing

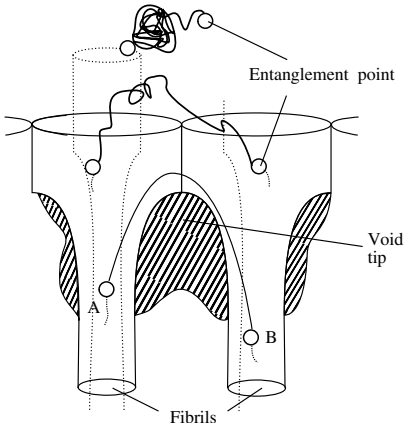


Figure 64.12: Conceptual drawing of fibrils at the interface between the crazed and uncrazed material (from ref.[23])

Fibrils are cold drawn polymer. Extension ratio remains constant as craze widens
Crack propagation:

(1) 1) new fibrils are created at the craze tip

(2) 2) fibrils break to form a true crack at the crack tip

(3) Crazing requires a stress field with a tensile hydrostatic component $\sigma_1 + \sigma_2 + \sigma_3 > 0$ (crazes have voids between fibrils)

(4) Crazing occurs first for PMMA in uniaxial extension ($\sigma_2 = 0$)

(5) $\mathcal{G}_{Ic}$ is determined by energy required to form a craze ($\approx 1000\,\mathrm{J/m^2}$)

(6) Crazing requires strain hardening of fibrils - material must be entangled ($M > M_c$), M_c typically ≈30,000 g/mol.

(7) In general, shear yielding competes with crazing at the crack tip

Meniscus instability mechanism (fibril formation at craze tip)
Material near the craze tip is strain softened, and can flow like a fluid between two plates.

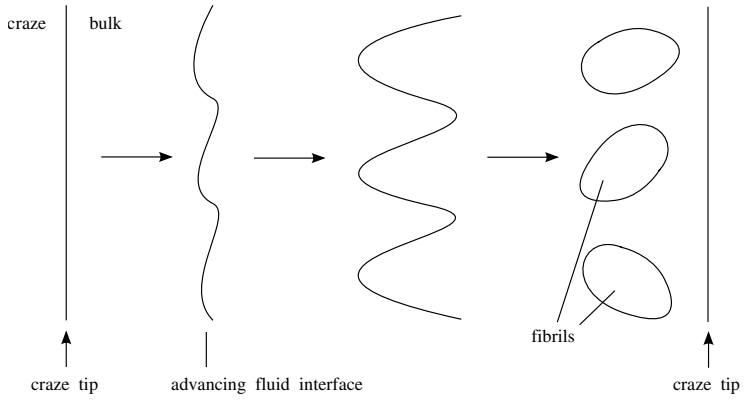


Figure 64.13: Meniscus instability mechanism for the formation of craze fibrils.

http://n-e-r-v-o-u-s.com/blog/?p=1556

Competition between Shear Deformation and Crazing Shear deformation is preferable to crazing for producing high toughness.
Plane stress - shear yielding and crazing criteria (for PMMA)

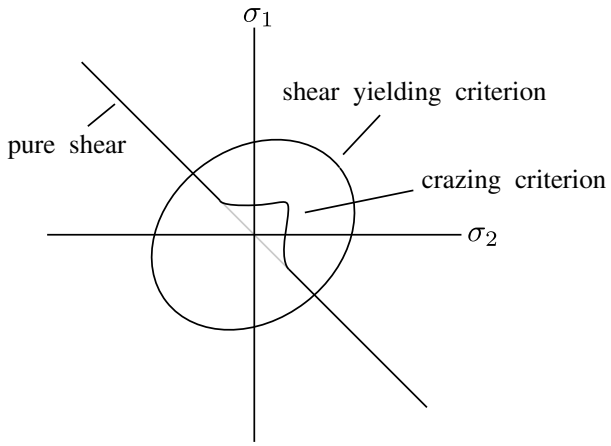


Figure 64.14: Deformation map for the shear yielding and crazing for plane stress conditions ($\sigma_3^p = 0$)

64.5.3 Dugdale model

Assumptions:
1) Tensile stress throughout plastic zone is constant value, σ_c
2) This stress acts to produce a crack opening displacement δ_c

$$\mathcal{G}_{IC} = \delta_c \sigma_c \tag{64.35}$$

The Dugdale zone (the craze in our polystyrene example) modifies the stress field so that it doesn't actually diverge to infinity, since infinite stresses are not really possible.

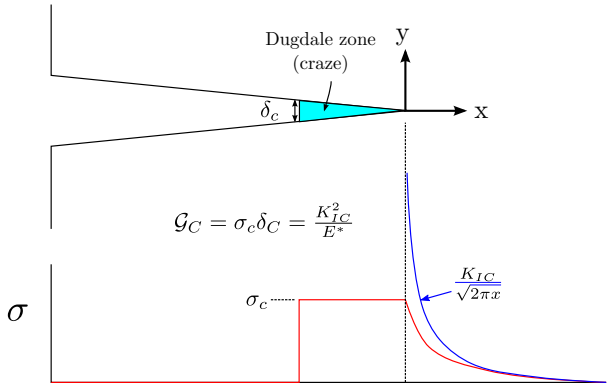


Figure 64.15: Crack tip stresses in the Dugdale model.

High Impact Polystyrene: Polystyrene (PS) is a big business - how do we make it tougher? High impact polystyrene (HIPS) is a toughened version of polystyrene produced by incorporating small, micron-sized rubber particles in the material. The morphology is shown in Figure 64.16, and consists small PS inclusions embedded in rubber particles that are in turn embedded in the PS matrix materials. The rubber particles act as stress concentrators that act as nucleation points for crazes.

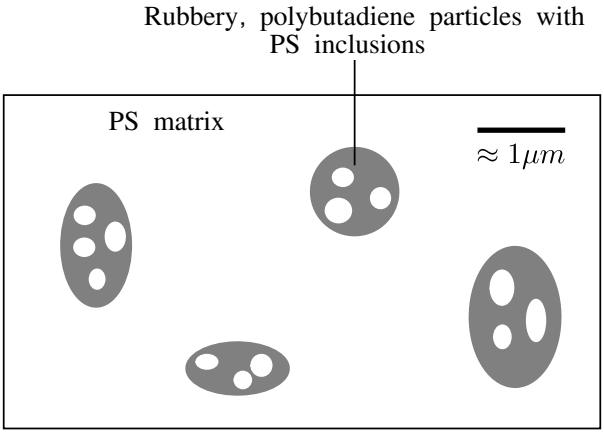


Figure 64.16: Morphology of high impact polystyrene.

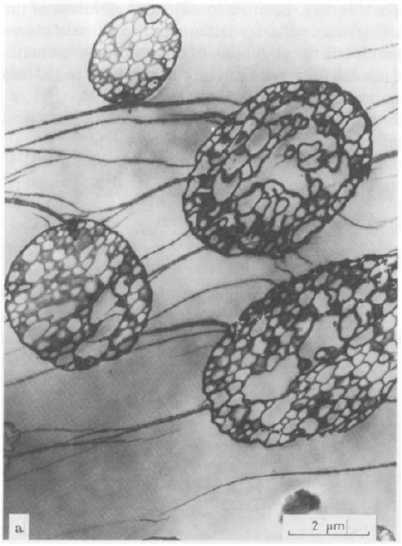


Figure 64.17: Multiple crazes in high impact polystyrene.

Tensile Behavior: In Figure 64.18 we compare the tensile behavior of normal polystyrene (PS) and high impact polystyrene (HIPS). The rubber content in the material reduces the modulus but substantially increases the integrated area under the stress/strain curve up to the point of failure, which is a measure of the toughness of the material. We can summarize the differences between PS and HIPS as follows:

- PS is brittle, with E =3GPa and relatively low fracture toughness
- HIPS is ductile, with a lower modulus $E = 2.1$ GPa and a much larger energy to fracture.

This area under the stress/strain curve is not a very quantitative measure of toughness because we don't have any information about the flaw size responsible for the eventual failure of the material.

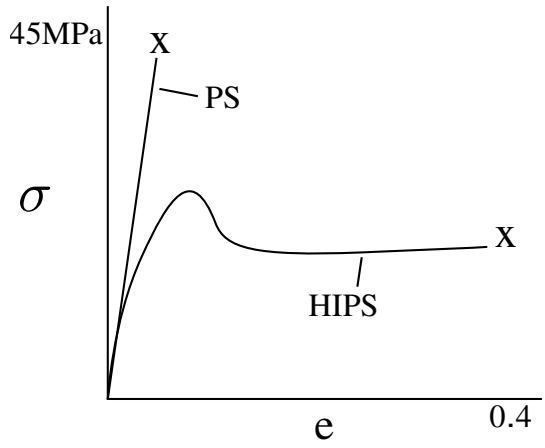


Figure 64.18: Schematic stress-strain curves for polystyrene (PS) and high impact polystyrene (HIPS) in the absence of a crack.

deformation via crazing in vicinity of rubber particles (stress concentrators) throughout sample

Samples with Precrack:

(measurement of K_{IC} or G_{IC})

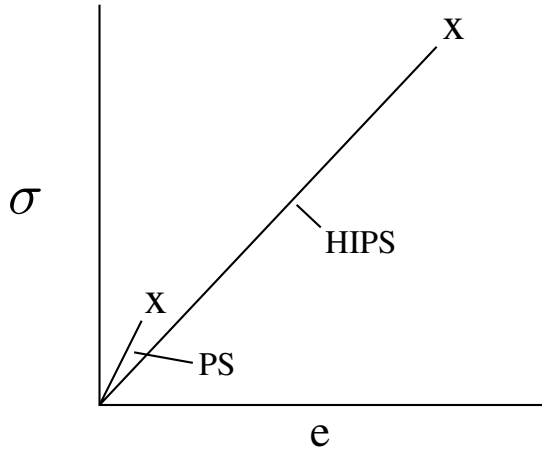


Figure 64.19: Schematic stress-strain curves for polystyrene (PS) and high impact polystyrene (HIPS) in the presence of a crack.

- Deformation limited to region around crack tip
- Much more deformation for HIPS - higher toughness

Impact Tests Impact tests are designed to investigate the failure of materials at high rates. Two common standardized are the Izod impact test and the Charpy impact test. They both involve measuring the loss in kinetic energy of a swinging pendulum as it fractures a sample. The geometry of a Charpy impact test is illustrated in Figure 64.20 Decrease in pendulum velocity after breaking sample gives impact toughness. For a useful discussion of a Charpy impact test, see https://www.youtube.com/watch?v=tpGhqQvftAo.

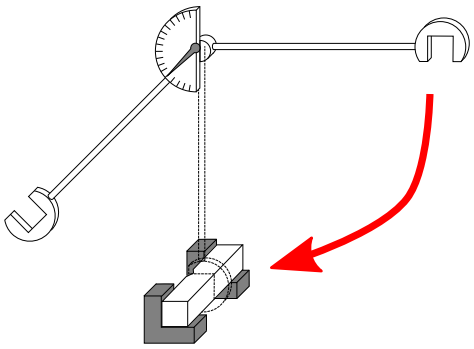


Figure 64.20: Charpy impact test.

For most materials the fracture toughness is rate dependent, but same general features for toughening materials often apply at both high and low fracture rates. For example, high impact polystyrene is much tougher than polystyrene at high strain rates, for the same general reasons outlined in Section 64.5.3.

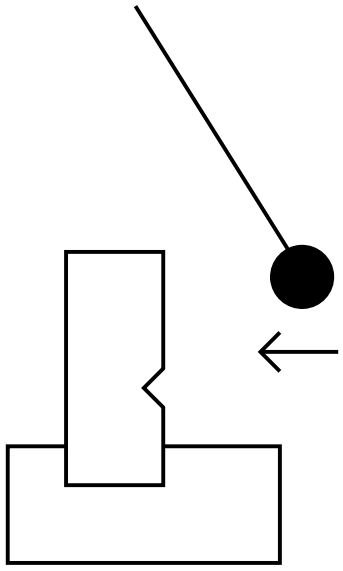


Figure 64.21: Izod impact geometry.

65 Matrix representation of Stress and Strain

As usual, we begin by replacing the directions (x, y, and z) with numbers: $x \rightarrow 1$, $y \rightarrow 2$, $z \rightarrow 3$. Once we do this we have 6 stress components, and six strain components. We then number these components from 1-6, so that 1, 2 and 3 are the normal components and 4, 5 and 6 are the shear components. We do this for both stress and strain as shown in Table65.1.

Table 65.1: Definition of the matrix components of stress and strain.

Engineering stress	Matrix Stress	Engineering Strain	Matrix Strain
σ_{11}	σ_1	e_{11}	e_1
σ_{22}	σ_2	e_{22}	e_2
σ_{33}	σ_3	e_{33}	e_3
σ_{23}	σ_4	e_{23}	e_4
σ_{13}	σ_5	e_{13}	e_5
σ_{12}	σ_6	e_{12}	e_6

A series of elastic constants relate the stresses to the strains. We can do calculations in either of the following two ways:

(1) Start with a column vector consisting of the 6 elements of an applied stress, and use the compliance matrix to calculate the strains.

(2) Start with a column vector consisting of the 6 elements of an applied strain, and use the stiffness matrix to calculate the stresses.

In each case we use a 6×6 matrix to relate two 6-element column vectors to one another. The procedure in each case is outlined below.

65.1 Compliance matrix

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} & s_{46} \\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} & s_{56} \\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & s_{66} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \tag{65.1}$$

The matrix must be symmetric, with $s_{ij} = s_{ji}$, so there are a maximum of 21 independent compliance coefficients:

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{12} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{13} & s_{23} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{14} & s_{24} & s_{34} & s_{44} & s_{45} & s_{46} \\ s_{15} & s_{25} & s_{35} & s_{45} & s_{55} & s_{56} \\ s_{16} & s_{26} & s_{36} & s_{46} & s_{56} & s_{66} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \tag{65.2}$$

Note that the compliance coefficients have the units of an inverse stress (Pa $^{-1}$).

65.2 Stiffness Matrix

The stiffness matrix (c) is the inverse of compliance matrix (note backwards notation). The stiffness coefficients have units of stress.

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} \\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} \tag{65.3}$$

65.3 Symmetry requirements on the compliance (or stiffness) matrix.

65.3.1 Orthorhombic symmetry

Extruded polymer sheets, like the one shown schematically in Figure 65.1 have orthorhombic symmetry, with different elastic properties in the extrusion, thickness and width directions. These materials have orthorhombic symmetry.

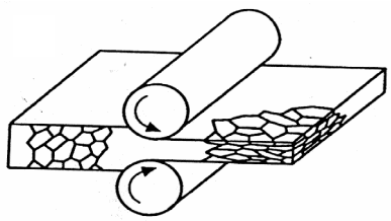


Figure 65.1: Schematic representation of an extruded sheet.

For materials with orthorhombic symmetry, the principal axes of stress and strain are identical, and all compliance components relating a shear strain (e_4, e_5 or e_6) to normal stresses (σ_1, σ_2 or σ_3) or to another shear stress must be zero. The stiffness matrix is as shown in Eq. 65.4 below, and there are 9 independent elastic constants. These 9 elastic can be identified as follows:

- $E_1 = 1/s_{11}$, $E_2 = 1/s_{22}$ and $E_3 = 1/s_{33}$, Young's moduli for extension in the 1, 2 and 3 directions, respectively.
- $G_1 = 1/s_{44}$, $G_2 = 1/s_{55}$ and $G_3 = 1/s_{66}$, Shear moduli for shear in the planes perpendicular to the 1, 2 and 3 directions, respectively.
- s_{12}, s_{13} and s_{23}, which relate stresses in one direction to strains in the perpendicular direction.

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & 0 & 0 & 0 \\ s_{12} & s_{22} & s_{23} & 0 & 0 & 0 \\ s_{13} & s_{23} & s_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{66} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \tag{65.4}$$

65.3.2 Fiber Symmetry

For a material with fiber symmetry, one of the axes is unique (in this case the 3 axis) and the material is isotropic in the orthogonal plane. Since the 1 and 2 axes are identical, there are now 5 independent elastic constants s_{33}, s_{13}, s_{44}, s_{11} s_{12}:

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} & 0 & 0 & 0 \\ s_{12} & s_{11} & s_{13} & 0 & 0 & 0 \\ s_{13} & s_{13} & s_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(s_{11} - s_{12}) \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \tag{65.5}$$

Examples of materials with fiber symmetry include the following:

(1) Many liquid crystalline polymers (e.g. Kevlar)

(2) Materials after cold drawing (plastic deformation to high strains, carried out below the glass transition temperature or melting temperature of the material.)

To better understand the significance of the 5 elastic constants for fiber symmetry, it is useful to consider the types of experiment we would need to conduct to measure each of them for a cylindrical fiber. The necessary experiments are described below.

Fiber extension along 3 direction: measurement of s_{33} and s_{13} We obtain s_{33} and s_{13} by performing a tensile test along the fiber axis (the 3 direction) as shown in Figure 65.2. The strain in the 3 direction is given by the fractional change in the length of the fiber after application of the load, and the strains in the 1 and 2 directions are given by the fractional changes in the diameter of the fiber:

$$e_3 = \Delta\ell/\ell; \; e_1 = e_2 = \Delta d/d \tag{65.6}$$

We then obtain s_{33} and s_{13} from Eq. 65.5, recalling that σ_3 is the only non-zero stress component in this situation:

$$\begin{aligned} s_{33} &= e_{33}/\sigma_3 \\ s_{13} &= e_1/\sigma_3 \end{aligned} \tag{65.7}$$

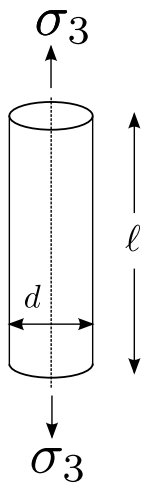


Figure 65.2: Fiber extension.

Fiber Torsion: Measurement of s_{44} We obtain the shear modulus by looking at the torsional stiffness of the fiber, *i.e.*, the Torque, T required to rotate the top and bottom of the fiber by an angle θ_0, as illustrated in Figure 65.3

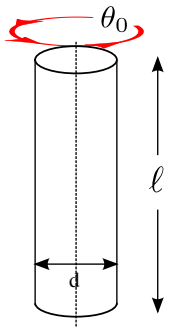


Figure 65.3: Fiber torsion.

We define a cylindrical system with a z axis along the fiber axis. The other axes in this coordinate system are the distance r from this axis of symmetry, and the angle θ around the z axis. The shear strain in the $\theta - z$ plane depends only on r, and is given by :

$$e_{\theta z} = r\frac{d\theta}{dz} = r\frac{\theta_0}{\ell} \tag{65.8}$$

The corresponding shear stress is obtained by multiplying by the shear modulus, G characterizing deformation in the 1-2 and 2-3 planes:

$$\sigma_{\theta z} = Gr\theta_0/\ell \tag{65.9}$$

We integrate the shear stress to give the torque, T:

$$T = \int_0^{d/2} r\sigma_{\theta z} 2\pi r dr = \frac{\pi G\theta_0 d^4}{32\ell} \tag{65.10}$$

So once we know the torsional stiffness of the fiber (the relationship between the applied torque T and the resulting rotation, θ_0) we know the shear modulus, G. This shear modulus is simply the inverse of s_{44}:

$$G = \frac{1}{s_{44}} \tag{65.11}$$

65.3.3 Fiber compression in 1-2 plane: determination of s_{11} and s_{12}

The last two elastic constants for a material with fiber symmetry can be determined from an experiment where the fiber is confined between two surfaces and compressed as shown in Figure 65.4. The elastic constants can be determined by measuring the width of the contact region between the fiber and the confining surface. If the confining surfaces are much stiffer than the fiber itself, than this contact width, $2b$, is determined by the elastic deformation of the material in the 1-2 plane. If there is no friction between the fiber and the confining surfaces the fiber is allowed to extend in the 3 direction as it is compressed and the contact width is given by the following expression:

$$b^2 = \frac{2Fd_0 s_{11}}{\ell} \tag{65.12}$$

where P is the force applied to the fiber, d_0 is its undeformed diameter and ℓ is its length.
A practical situation that is often observed is that friction between the fiber and confining surfaces keeps the fiber length from increasing, so the strain in the three direction must be zero. If we express the stress state in terms of the principle stresses in 1 and 2 directions, we have the following from Eq. 65.5:

$$e_3 = s_{13}\left(\sigma_1^p + \sigma_2^p\right) + s_{33}\sigma_3 \tag{65.13}$$

Setting e_3 to zero in this equation gives the following from σ_3:

$$\sigma_3 = \frac{-s_{13}}{s_{33}}\left(\sigma_1^p + \sigma_2^p\right) \tag{65.14}$$

A consequence of this stress is that the frictionless expression for b gets modified to the following:

$$b^2 = \frac{2Fd_0}{\ell}\left(s_{11} - \frac{s_{13}^2}{s_{33}}\right) \tag{65.15}$$

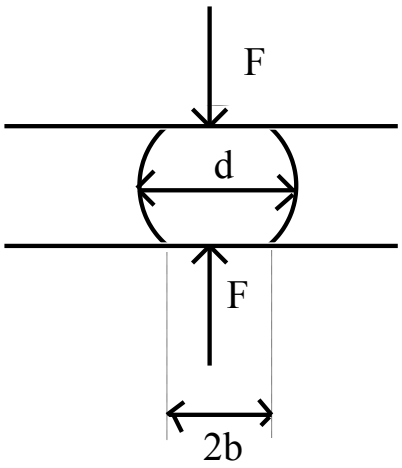


Figure 65.4: Transverse deformation of a fiber.

The remaining constant, s_{12}, is determined from a measurement of the measurement of d/d_0, the ratio of the fiber width at the midplane to the original width of the fiber. This relationship is complicated, and involves several of the different elastic constants.

$$\frac{d}{d_0} = f\left(F, s_{11}, s_{13}, s_{33}, s_{12}\right) \tag{65.16}$$

65.3.4 Cubic Symmetry

For a material with cubic symmetry the 1,2 and 3 axes are all identical to one another, and we end up with 3 independent elastic constants:

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{11} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{12} & s_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{44} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \tag{65.17}$$

65.3.5 Isotropic systems

For an isotropic material all axes are equivalent, and the properties are invariant to any rotation of the coordinate axes. In this case there are two independent elastic constants, and the compliance matrix looks like this:

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{11} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{12} & s_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2(s_{11}-s_{12}) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2(s_{11}-s_{12}) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2(s_{11}-s_{12}) \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} \tag{65.18}$$

The requirement that the material properties be invariant with respect to any rotation of the coordinate axes results in the requirement that $s_{44} = 2(s_{11} - s_{12})$, so there are two independent elastic constants. The shear modulus, G, Young's modulus E and Poisson's ratio, ν are given as follows:

$$G = 1/2(s_{11} - s_{12}); \; E = 1/s_{11}; \; \nu = -s_{12}/s_{11} \tag{65.19}$$

Bulk Modulus for an Isotropic Material The bulk modulus, K_b, of a material describes it's resistance to a change in volume (or density) when we apply a hydrostatic pressure, p. It is defined in the following way:

$$K_b = \frac{1}{V}\frac{dP}{dV} \tag{65.20}$$

The stress state in this case is as follows:

$$\sigma = \begin{bmatrix} -p & 0 & 0 \\ 0 & -p & 0 \\ 0 & 0 & -p \end{bmatrix} \tag{65.21}$$

From the compliance matrix (Eq. 65.18) we get $e_1 = e_2 = e_3 = -p(s_{11} + 2s_{12})$. The change in volume, ΔV can be written in terms of the three principal extension ratios, λ_1, λ_2 and λ_3:

$$\frac{\Delta V}{V_0} = \frac{V}{V_0} - 1 = \lambda_1\lambda_2\lambda_3 - 1 = (1+e_1)(1+e_2)(1+e_3) - 1 \approx e_1 + e_2 + e_3 \tag{65.22}$$

Now we use the fact that for small x, $(1+x)^3 \approx 3x$, so we have:

$$\frac{\Delta V}{V_0} = e_1 + e_2 + e_3 = -3p(s_{11} + 2s_{12}) \tag{65.23}$$

Recognizing that the derivative dP/dV in the definition of K_b can be written as the limit of $p/\Delta V$ for very small p allows us to obtain the expression we want for K_b:

$$K_b = \lim_{p\to 0}\frac{-p}{\Delta V/V_0} = \frac{1}{3(s_{11} + 2s_{12})} \tag{65.24}$$

Relationship between the Isotropic Elastic Constants: Because there are only two independent elastic constants for an isotropic system E and ν can be expressed in terms of some combination of K_b and G. For E the relevant relationship is as follows.

$$E = \frac{9G}{3 + G/K_b} = 2G\,(1 + \nu) \tag{65.25}$$

We can also equate the two expressions for G in Eq. 65.25 to get the following expression for ν:

$$\nu = \frac{3 - 2G/K_b}{6 + 2G/K_b} \tag{65.26}$$

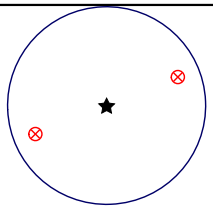
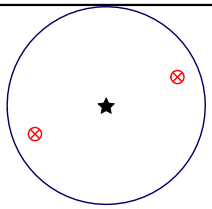
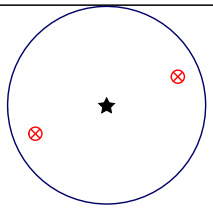
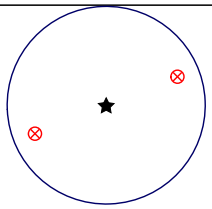
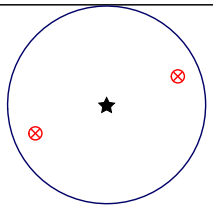
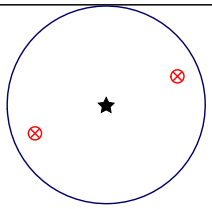
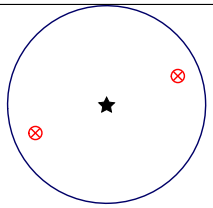
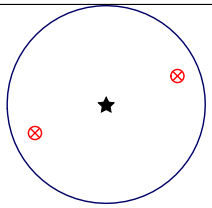
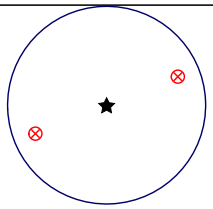
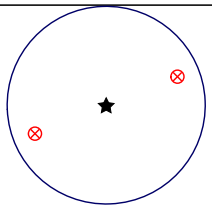
Note that if $K_b \gg G$, $E \approx 3G$ and $\nu \approx 0.5$.

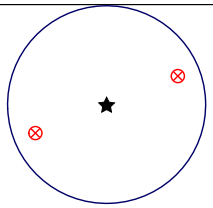
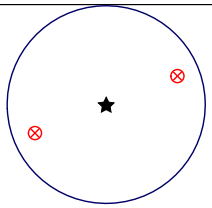
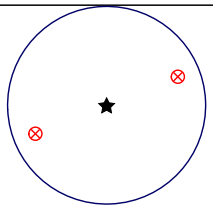
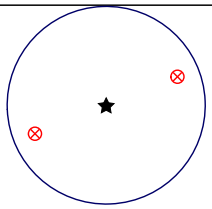
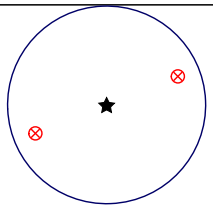
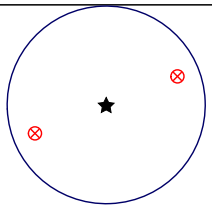
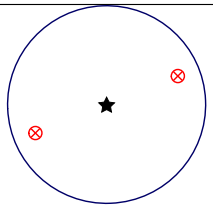
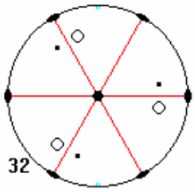
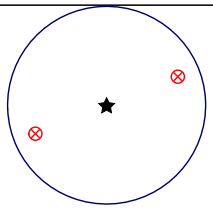
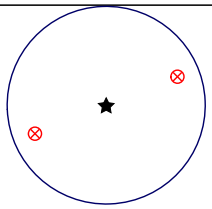
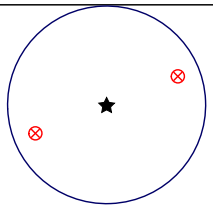
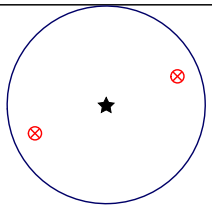
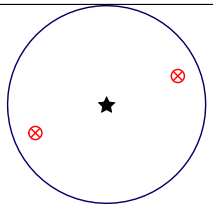
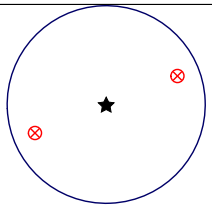
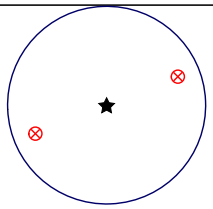
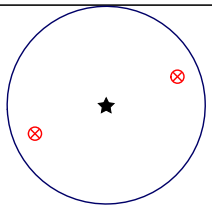
66 Other Symmetry-Related Constitutive Relationships

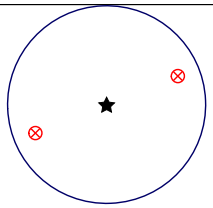
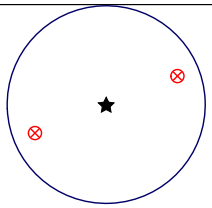
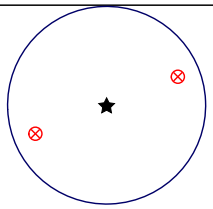
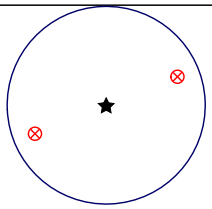
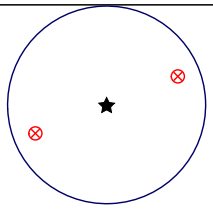
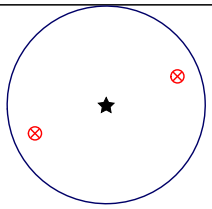
(something about piezoelectricity, thermal expansion, polarizability, etc.)

67 Crystal Systems and Point Groups

All the stereopgraphic projections need to be completed for these.

Point Group	System	Projection	Point Group	System	Projection
1	triclinlic		$\bar{1}$	triclinic	
2	monoclinic		m	monoclinic	
2/m	monoclinic		222	orthorhombic	
mm2	orthorhombic		mmm	orthorhombic	
4	tetragonal		$\bar{4}$	tetragonal	

Point Group	System	Projection	Point Group	System	Projection
4/m	tetragonal		422	tetragonal	
4mm	tetragonal		$\bar{4}2m$	tetragonal	
4/mmm	tetragonal		3	trigonal	
$\bar{3}$	trigonal		32	trigonal	
3m	trigonal		$\bar{3}m$	trigonal	
6	hexagonal		$\bar{6}$	hexagonal	
6/m	hexagonal		622	hexagonal	
6mm	hexagonal		$\bar{6}m2$	hexagonal	

Point Group	System	Projection	Point Group	System	Projection
6/mmm	hexagonal		23	cubic	
m $\bar{3}$	cubic		432	cubic	
$\bar{4}$3m	cubic		m $\bar{3}$m	cubic	

Need to introduce these diagrams and include them for the different point groups. Most likely as an appendix.

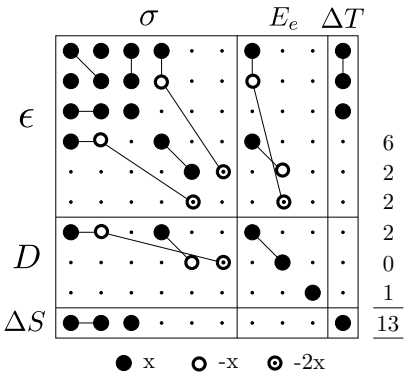


Figure 67.1: Symmetry relationships between matrix elements for the 32 point group.

68 Yield Behavior

The yield point of a material corresponds to the onset of permanent deformation, originating for example from the movement of dislocations. The stress/strain curve for a simple tensile test is shown in Figure 68.1, with the tensile yield stress, σ_y, corresponding to the onset of permanent, irreversible deformation in the material. For most materials this corresponds to the onset of nonlinearity in the stress-strain curve (rubber is the exception, and that case is discussed in more detail in 331). What we need is a generalized criterion that can be used to determine the onset of yield for any stress state. These yield criteria all focus on the importance of the shear stress, and we begin with the discussion of yield in single crystals in the following Section .

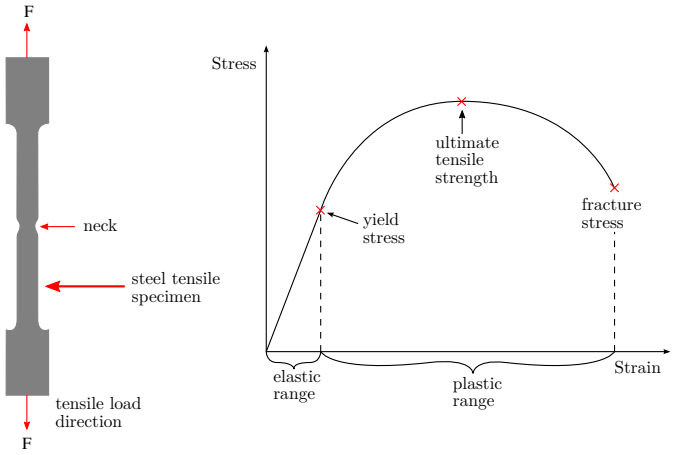


Figure 68.1: Tensile testing sample and representative data for a ductile sample.

Table 68.1: Shear moduli and values of τ_{crss} for different metals.

Material	G (GPa)	τ_{crss} (MPa)
Silver	4.6	0.37
Aluminum	4.2	0.78
Copper	7.2	0.49
Nickel	12.2	3.2-7.35
Iron	13.2	27.5
Molybdenum	19	71.6
Niobium	5.8	33.3
Cadmium	3.8	0.57
Magnesium (basal slip)	2.8	0.39
Magnesium (prism slip)	2.8	39.2
Titanium (prism slip)	6.3	13.7
Beryllium (basal slip)	23.4	1.37
Beryllium (prism slip)	23.4	52

68.1 Critical Resolved Shear Stress

The simplest criterion for the yield of a material is that the Resolved shear stress, τ_{RSS}, exceeds a critical value referred to simply as the critical resolved shear stress, τ_{CRSS}. The resolved shear stress is obtained from the applied stress and the orientation of the slip plane as illustrated in Figure 68.2. For a single crystal the relevant resolved shear stress is the shear stress on acting on the slip plane, in the direction of the Burger's vector. In a tensile experiment it is given by the applied tensile stress, σ, the angle ϕ between the slip plane normal and the tensile axis, and the angle λ between the tensile axis and the slip direction:

$$\tau_{RSS} = \sigma \cos\phi \cos\lambda \tag{68.1}$$

In other words, $\tau_{RSS} = \tau_{CRSS}$ when $\sigma = \sigma_y$, where σ_y is the tensile yield strength. Substituting τ_{CRSS} for τ_{RSS} and σ_y for σ in Eq. 68.1 and solving for σ_y gives:

$$\sigma_y = \frac{\tau_{CRSS}}{\cos\phi \cos\lambda} \tag{68.2}$$

The factor $\cos\phi\cos\lambda$ is called the Schmid factor, and slip occurs along the slip system that the largest value of this quantity. The sum of ϕ and λ must be at least 90 °, and it can be shown that $\cos\phi\cos\lambda$ is maximized when $\phi = \lambda = 45°$, in which case the yield stress is twice the critical resolved shear stress. We refer to this value of the yield stress as the minimum value, σ_y^{min}:

$$\sigma_y^{min} = 2\tau_{CRSS} \tag{68.3}$$

Measured values of the critical resolved shear stress for different metals are shown in Table68.1.

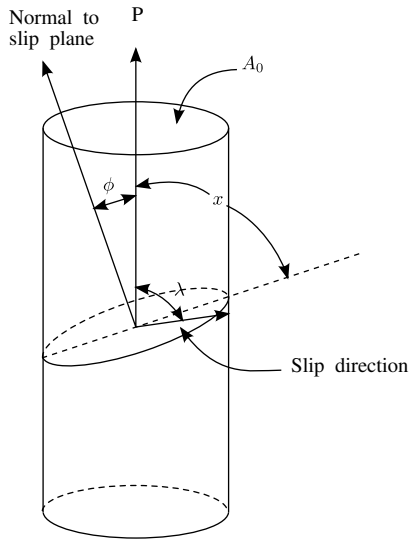


Figure 68.2: Resolved shear stress.

For polycrystalline materials, the situation is more complicated, since all crystal orientations will have a different value of the Schmid factor. In a polycrystalline material, each grain has a different value of the Schmid factor, so the situation is more complicated. However, we can get a reasonable approximation by using an appropriate average value for the Schmid factor instead of the maximum possible value that this value can have. We actually need the quantity, M, which is the average of the reciprocal Schmid factor for all of the grains in a material:

$$M = \left\langle \frac{1}{\cos\phi \cos\lambda} \right\rangle \tag{68.4}$$

Here the average is taken over all possible grain orientations of the material. The tensile yield stress is given by the following expression:

$$\sigma_y = M\tau_{CRSS} = \frac{M}{2}\sigma_y^{min} \tag{68.5}$$

This value of M is typically between 2.7 and 3, so we end up with a yield stress in a polycrystalline material that is between 35 and 50 percent larger than the minimum possible value given by Eq. 68.3.

68.2 Yield Surfaces

The full stress state of a material is defined by the 3 principal stresses, σ_1^p, σ_2^p and σ_3^p. The stress state of a material can therefore be specified on a 3-dimensionsal space where the values of these 3 principal stresses are plotted on three orthogonal axes. The yield surface is the surface in this space that separates the stress states where yielding will occur from those where it will not occur. Here we describe two of the most common yield surfaces, those defined by the Tresca and Von Mises yield criteria.

68.2.1 Tresca Yield Criterion

TheTresca yield criterion is the simplest one in that we just assume that shear occurs whenever the maximum shear stress in the sample exceeds some critical value value τ_{crit}. In mathematical terms, yield occurs under the following conditions:

$$\frac{\left|\sigma_i^p - \sigma_j^p\right|_{max}}{2} > \tau_{crit} \tag{68.6}$$

where the 'max' subscript indicates that we take the principal stress difference ($\sigma_1^p - \sigma_2^p$, $\sigma_2^p - \sigma_3^p$ or $\sigma_1^p - \sigma_3^p$) with the largest magnitude. The yield stress, σ_y, is typically measured in a uniaxial tensile experiment, where

$\sigma_2^p = \sigma_3^p = 0$, and plastic yielding of the material occurs when $\sigma_1^p > \sigma_y$. In a uniaxial tensile experiment, the maximum shear stress is half the applied tensile stress, so $\tau_{crit} = \sigma_y/2$.

68.2.2 Von Mises Yield Criterion

A more complicated criterion is that theVon Mises stress, σ_e, stress is larger than some critical value. The Von Mises stress is given as follows:

$$\sigma_e = \frac{\sqrt{2}}{2}\sqrt{\left(\sigma_2^p - \sigma_1^p\right)^2 + \left(\sigma_3^p - \sigma_1^p\right)^2 + \left(\sigma_3^p - \sigma_2^p\right)^2} \tag{68.7}$$

Yielding occurs when $\sigma_e > \sigma_y$. The Tresca and Von Mises yield surfaces for a two dimensional stress state (σ_3^p=0) are shown in Figure 68.3. Yielding does not occur inside the surface, but does occur outside the surface.

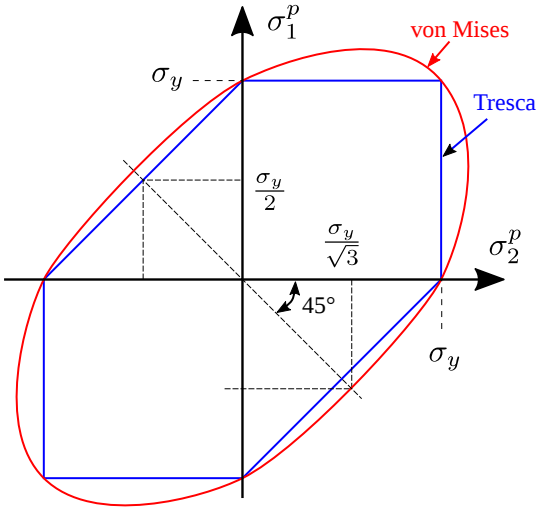
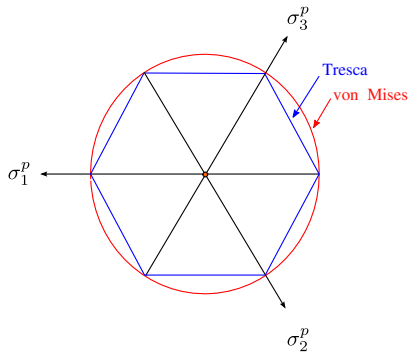


Figure 68.3: Cross sections of the Tresca and Von Mises yield surfaces at the $\sigma_3^p = 0$ plane.

Viewed down the hydrostatic line ($\sigma_1^p = \sigma_2^p = \sigma_3^p$), this is what the criteria look like:



68.3 Localized Deformation

A material that obeys the strain hardening law of Eq. 68.16 will fail when a portion of the sample becomes thinner than the remainder of the sample. The overall behavior of the sample is a balance between the fact that the strain is larger in this region, and can therefore support a larger true stress, but the cross section is larger, so that a larger true stress is needed just to maintain a constant force along the length of the sample. To understand how to think about this we need to consider the relationship between the true stress and the engineering stress.
Consider a sample that is being deformed in uniaxial extension, as illustrated in Figure 68.4. A sample with an undeformed cross sectional area of $A_0$and undeformed length of ℓ_0 is stretched with a force, P. The engineering tensile stress, σ_{eng}, is obtained by dividing the load by the undeformed cross section, and the true tensile stress, σ_t is obtained by dividing the load by the actual cross sectional area of the deformed sample:

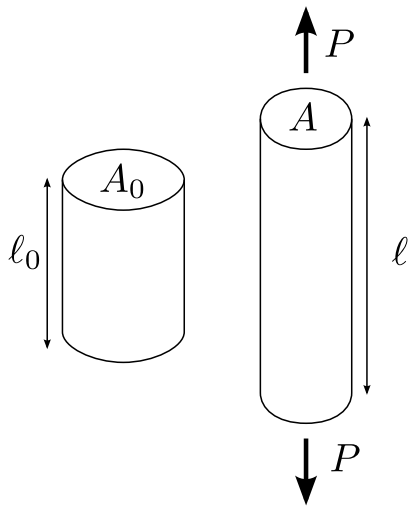


Figure 68.4: Uniaxial tensile test.

$$\sigma_{eng} = P/A_0 \tag{68.8}$$

$$\sigma_{true} = P/A \tag{68.9}$$

In general, the bulk modulus of a material is much larger than its yield stress, so the applied stresses associated with yield phenomena are not large enough to significantly change the volume. As a result, the sample deforms at constant volume, so we have:

$$A\ell = A_0\ell_0 \tag{68.10}$$

The relationship between the true stress and the engineering stress is therefore as follows:

$$\sigma_t = \sigma_{eng}\ell/\ell_0 = \sigma_{eng}\lambda \tag{68.11}$$

68.3.1 Considére Construction

TheConsidére construction is a simple construction that can be used to determine the stability of regions in a sample at large tensile deformations. It can be used in to distinguish between unstable and unstable necking of a sample, illustrated schematically in Figure 68.5. We begin by considering a region of the sample that has a slightly thinner cross section than the rest of the sample. The true stress in this region of the sample will be higher than the rest of the sample because we are dividing the applied load by a lower cross section . Two things can happen at this point, the first possibility is that the larger stress in this region of the sample leads to greater deformation, and the sample breaks as the necked region begins to thin down. This is the unstable necking condition illustrated on the left side of Figure 68.5. In this case the maximum force, P, applied to the sample is the force where the neck begins to form.
The second possibility is that the increased strain in the necked region leads to substantial strain hardening, so that this region of the sample is able to support the larger true stress in that region. Under the appropriate conditions the cross section of the necked region will stabilize at a value that is determined by the stress/strain relationship for the material. The sample deforms by 'drawing' new material into this necked region, as illustrated on the right side of Figure 68.5.

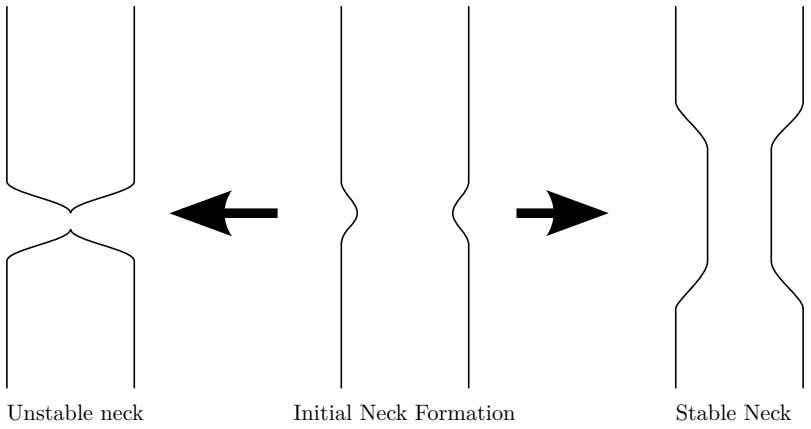


Figure 68.5: Schematic representation of stable and unstable necking of a sample under tensile loading conditions.

To understand when stable or unstable necking occur, we begin by recognizing that the onset of neck formation corresponds to a maximum tensile force that the material is able to sustain. In mathematical terms:

$$\frac{dP}{d\lambda} = 0 \tag{68.12}$$

Since the engineering stress is the load divided by the undeformed cross sectional area, which is a constant, we can write Eq. 68.12 as follows:

$$\frac{d\sigma_{eng}}{d\lambda} = 0 \tag{68.13}$$

We can rewrite this expression in terms of the true stress by recognizing that $\sigma_{eng} = \sigma_t/\lambda$ (see Eq. 68.11), from which we obtain:

$$\lambda\frac{d\sigma_t}{d\lambda} - \sigma_t = 0 \tag{68.14}$$

which we rearrange to the following:

$$\frac{d\sigma_t}{d\lambda} = \frac{\sigma_t}{\lambda} \tag{68.15}$$

This condition is met when a line drawn fro the origin of a plot of σ_t vs. λ is tangent to the curve.

68.3.2 Stable and Unstable Necking

use of the Considére construction is illustrated in Figure 68.6, where we show curves of the true stress vs. extension ratio for a material that does not form a stable neck (part a) and one that does form a stable neck (part b). In part a it is only possible to draw one line originating from the origin that is tangent to the stress-strain curve (point A in Figure 68.6). A necking instability forms when the true stress reaches this value, resulting in a thinned-down region of the sample. This region continues to thin down until the sample breaks. The maximum engineering stress that the sample sees prior to failure is given by the slope of the tangent line in Figure 68.6a.
In Figure 68.6b, it is possible to draw two lines from the origin that are tangent to the curve, with tangent points labeled as A and B. The tangent at point B represents a maximum in the applied force (the stress-strain curve lies below the tangent line) and the tangent at point B represents a minimum in the applied force (the stress-strain curve lies above the tangent line). At point B the neck stabilizes. Additional material is drawn into the necked region with a characteristic draw ratio that given by the value of λ at point B. The engineering stress at which this drawing occurs is less than the engineering stress required to form the neck in the first place. This means that the load during the drawing process to form the stable neck is lower than the stress required to form the neck in the first place. This strain softening phenomenon is generally observed in glassy polymeric materials ($T < T_g$) or semicrystalline polymers for which stable neck formation is observed. If there is not strain softening, then the drawing process occurs at the same value of the engineering stress for which neck formation was first observed. This requires that the tangent curve through point A intersect the stress-strain curve again at a much higher value of λ. This intersection is shown as point C in Figure 68.6b. This point describes the stress and strain in the necked region if the engineering stress remains at the same value that it did when the neck formation process started.

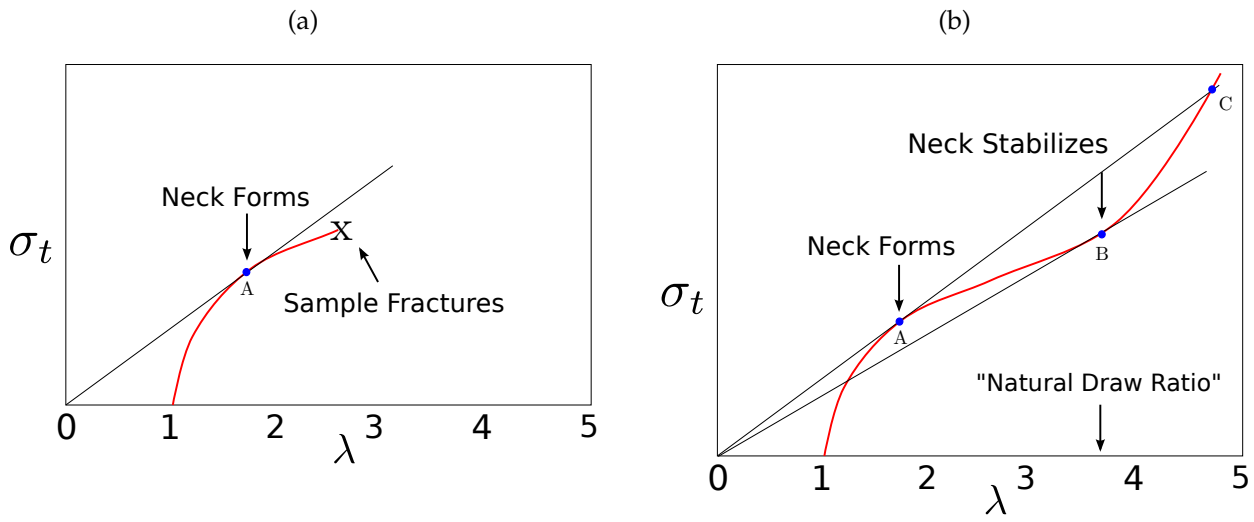


Figure 68.6: Considére construction for a material that does not form a stable neck (a) and a material that does form a stable neck (b).

68.3.3 Case Study - Power Law Strain Hardening

The following equation is often used to describe the behavior of a material after the yield point:

$$\sigma_t = K e_t^n \tag{68.16}$$

where:

- σ_t= true stress (force over actual cross sectional area in a tensile test)
- e_t= true strain ($\ln(\ell/\ell_0) = \ln\lambda$ in a tensile test - see Eq. 62.29).
- K= strength coefficient (true stress at true strain of 1)
- n= strain hardening coefficient (dimensionless)

Limiting behavior: $n = 1$ is perfectly elastic behavior, whereas $n = 0$ corresponds to perfectly plastic behavior. Actual values of n fall somewhere between these two extremes, and are listed in Table68.2 for a variety of metals.

Table 68.2: Strain Hardening Coefficients for Various Materials (From Hertzberg, Table 2.8).

Material	Strain Hardening Coefficient
Stainless Steel	0.45-0.55
Brass	0.35-0.4
Copper	0.3-0.35
Aluminum	0.15-0.25
Iron	0.05-0.15

68.3.4 Necking Instability in a Power-Law Strain Hardening Material

If we write the true stress appearing in Eq. 68.16 as $\ln\lambda$ (see Eq. 62.29), we have:

$$\sigma_t = K \ln((\lambda))^n \tag{68.17}$$

Differentiating with respect to λ gives:

$$\frac{d\sigma_t}{d\lambda} = \frac{nK(\ln(\lambda))^{n-1}}{\lambda} = \frac{nKe_t^{n-1}}{\lambda}$$

and then using Eq. 68.15 to equate $d\sigma_\lambda/d\lambda$ to σ/λ gives the following:

$$\frac{nK\left(\ln\left(\lambda\right)\right)^{n-1}}{\lambda} = \frac{\sigma}{\lambda} = \frac{Ke_t^n}{\lambda}$$

which simplifies to the following

$$n = \epsilon_t \tag{68.18}$$

So a measurement of the strain where the necking instability is observed can be used to determine the strain hardening exponent.

68.4 Deformation Twinning

Deformation twinning is a mechanism of deformation in crystalline materials that doesn't involve dislocation motion. Instead, it involves the concerted, simultaneous motion of may atoms to form a twin boundary in the material. Twin boundaries were discussed in 316-1, and are low-energy, coherent grain boundaries corresponding to a change in stacking sequence in an FCC lattice. The easiest ones to visualize correspond to changes in the stacking sequence of the close-packed layers in face centered cubic (FCC) and hexagonal close packed (HCP) crystals. An example of a twin boundary in and FCC crystal is shown in Figure 68.7.

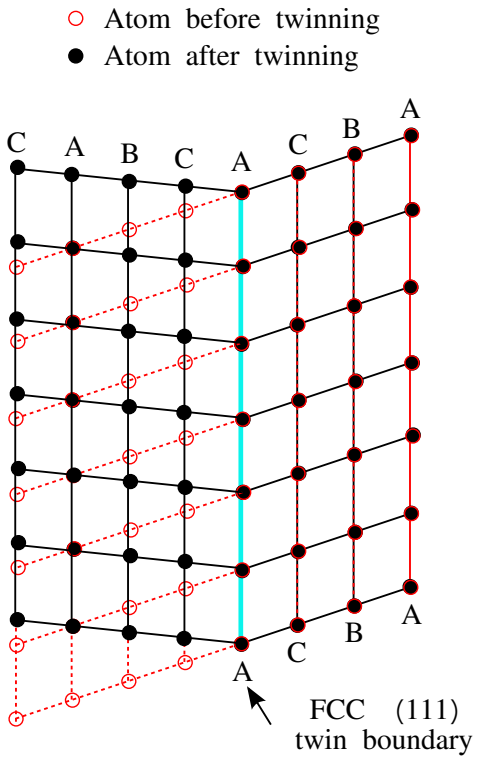


Figure 68.7: Example of a series of twin boundaries in an FCC crystal (from ref.[?], adapted with permission from Macmillan Publishers Ltd.).

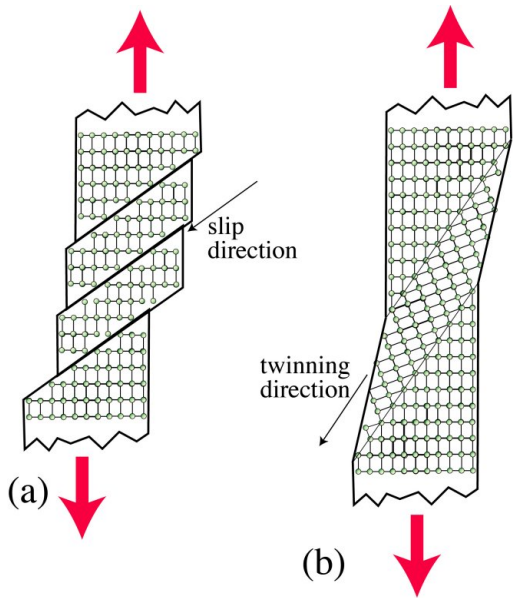


Figure 68.8: Undeformed material (a) and deformed material after slip (b) and deformation twinning (c).

A useful demonstration of deformation twinning is available from here:http://www.doitpoms.ac.uk/ldplib/acoustic/intro.php

68.5 Deformation of Polymers

The deformation mechanisms in polymeric materials are completely different from those in metals and ceramics, and (almost) never have anything to do with the motion of dislocations. To begin with, we can separate into the polymers that are partially crystalline and those that are not. As one would expect, the structures of these non-crystalline (amorphous) and semicrystalline polymers are very different, as shown schematically in Figure Crystallization and glass formation are the two most important concepts underlying the physical properties of polymers. The ways in which the molecules are organized in non-crystalline (amorphous) polymers and semicrystalline polymers are very different, as illustrated in Figure s Polymers crystallize at temperatures below T_m (melting temperature) and form glasses at temperatures below T_g (glass transition temperature). All polymers will form glasses under the appropriate conditions, but not all polymers are able to crystallize. The classification scheme shown in Figure Polymers classification scheme. divides polymeric materials based on the locations of T_g and T_m (relative to the use temperature, T) and is a good place to start when understanding different types of polymers.

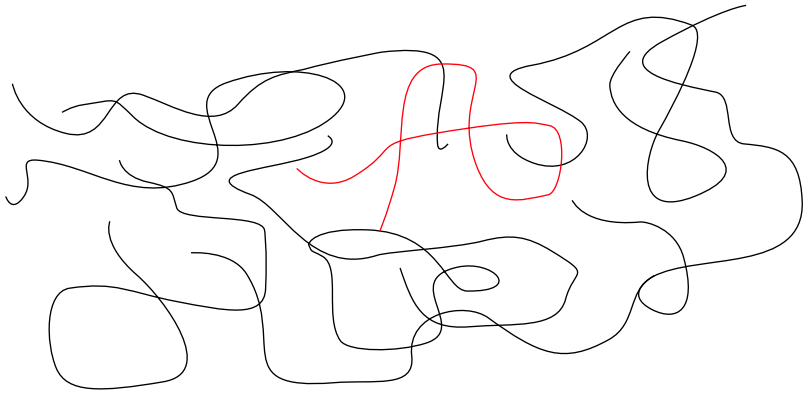


Figure 68.9: Structure of an amorphous polymer.

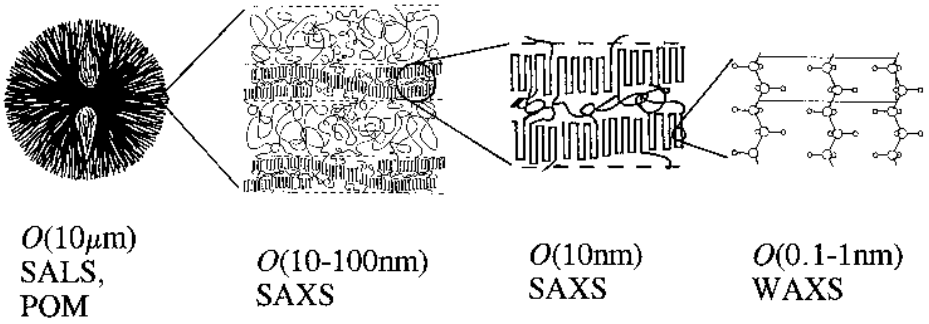


Figure 68.10: Structure at different length scales for a semicrystalline polymer.

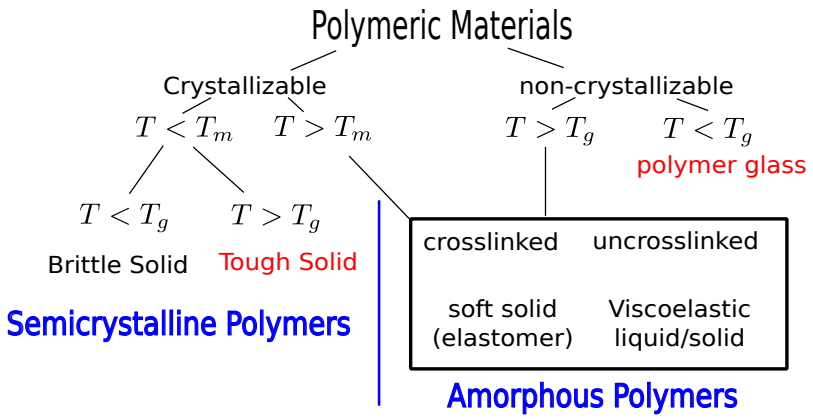


Figure 68.11: Polymers classification scheme.

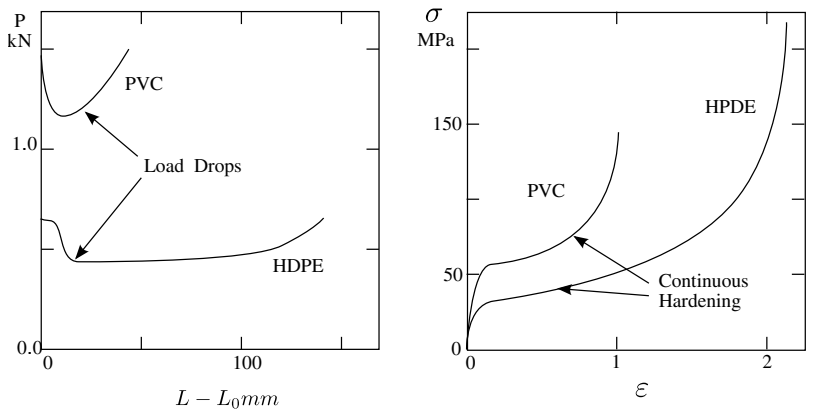


Figure 68.12: Load-Displacement and true stress-true strain curves for PVC (polyvinyl chloride) and HDPE (high density polyethylene).

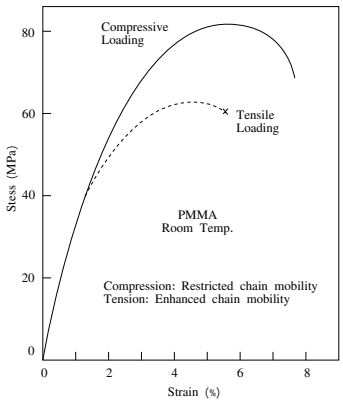


Figure 68.13: Stress-strain curves for PMMA (polymethylmethacrylate) under tensile and compressive loading conditions.

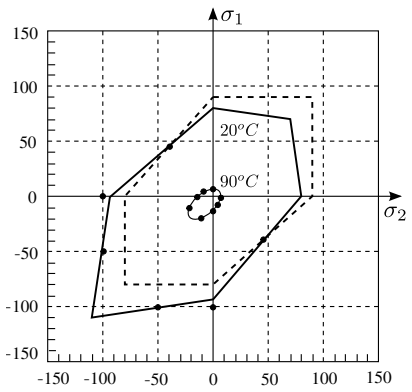


Figure 68.14: Yield surface for PMMA at 20 °C and at 90 °C. For comparison a map of the Tresca yield criterion (where normal forces do not matter) shown as the dashed line.

69 Strengthening Mechanisms in Metals

Yield in metals occurs by dislocation motion. This means that if we want to increase the yield stress of a material we have two choices:

(1) We can make a material with a very low dislocation density, thereby eliminating dislocation as a yield mechanism.

(2) We can do something to the material to make it more difficult for the dislocations to move.

In almost all cases, we choose option two. This is because mechanisms exist for dislocations to be created during the deformation of a bulk material. More specifically, dislocations multiply when they are pinned in some way, as illustrated by the representation of the operation of a Frank-Read dislocation source in Figure 69.1. (See the 316-1 text to review Frank-Read sources if you don't recall the details.) As a result, even if we manage to create a material with a very low dislocation density, a sufficient number of dislocations will always be created once deformation starts. An exception to this rule is single crystalline whiskers, which are so small that the dislocation density can be reduced essentially to zero. This is because it is energetically favorable for dislocations to move to the free surface of the material, where they are annihilated.

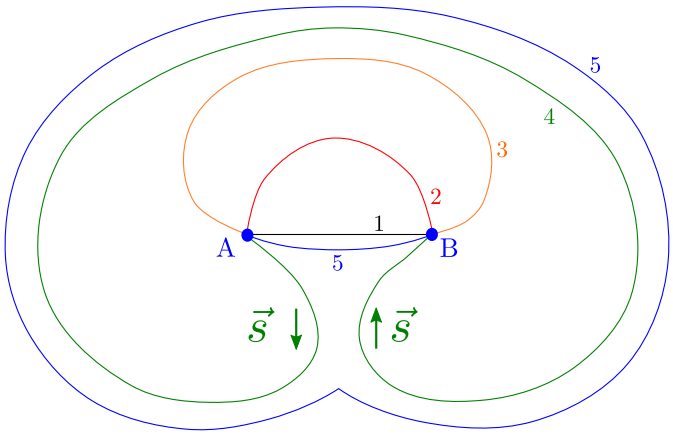


Figure 69.1: Schematic illustration of a Frank-Read source.

Strengthening schemes based on reducing the mobility of dislocations can be divided into the following general categories, which we will discuss in more detail below:

- Work hardening
- Grain boundary strengthening
- Solid solution hardening
- Precipitation hardening

69.1 Work Hardening

Work hardening originates from interactions between dislocations. Dislocations impede the motion of other dislocations, so the multiplication of dislocations that occurs as a material is deformed results in a harder material. The effect is illustrated in Figure 69.2, which shows the relationship between the shear strength of a material and the dislocation density. In the absence of dislocations, the strength approaches the maximum theoretical shear strength of $\approx G/6$, implying a tensile strength of $\approx G/12$.

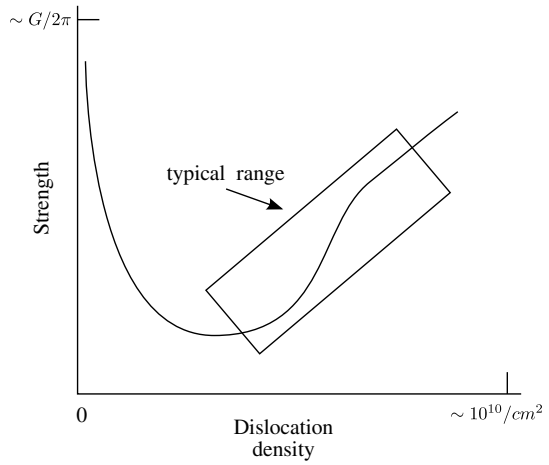


Figure 69.2: Qualitative relationship between strength and dislocation density.

70 Linear Viscoelasticity

Silly Putty is an obvious example of a viscoelastic material, with mechanical properties that depend on the time scale of the measurement. The simplest situation to start with is the small strain region, where the stresses are proportional to the strains, but depend on the time of the experiment. This linear viscoelastic regime is the focus of this Section .

70.1 Time and Frequency Dependent Moduli

As discussed in Section62.6, a variety of elastic constants can be defined for an isotropic material, although only two of these are independent. All of these elastic constants can be time dependent, although in this Section we focus on the specific case of a uniaxial tensile test, used to define the Young's modulus, E, for an elastic material. Viscoelastic materials have mechanical properties that depend on the timescale of the measurement. The easiest way to think about this is to imagine a tensile experiment where a strain of e_0 is instantaneously applied to the sample (see the time-dependent strain for the relaxation experiment in Figure 70.1). In a viscoelastic material, the resulting stress will decay with time while we maintain the strain at this fixed value. The time dependence of the resulting tensile stress, σ, enables us to define a time-dependent relaxation modulus, $E(t)$:

$$E(t) = \frac{\sigma(t)}{e_0} \tag{70.1}$$

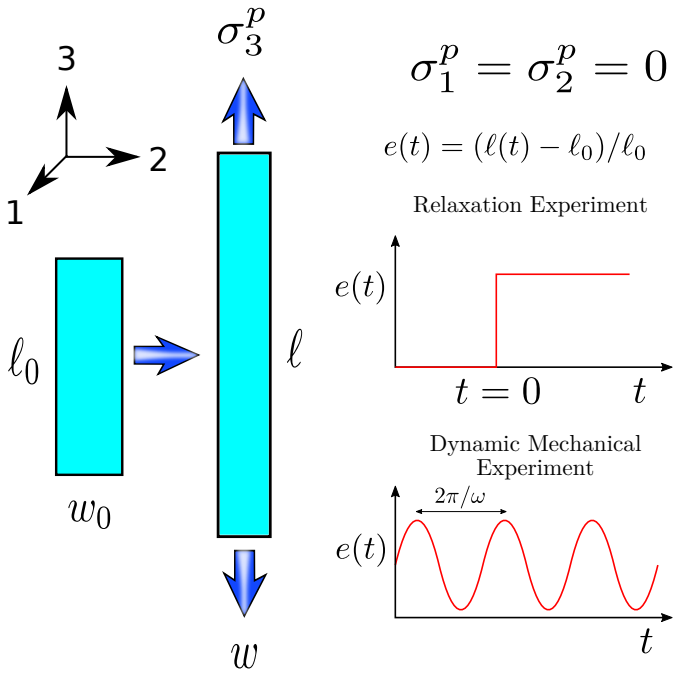


Figure 70.1: Experimental geometry and time dependent strain for for the determination of the time-dependent relaxation modulus and the frequency-dependent dynamic moduli.

We are often interested in the application of an oscillatory strain to a material. Examples include the propagation of sound waves, where wave propagation is determined by the response of the material at the relevant frequency of the acoustic wave that is propagating through the material. In an oscillatory experiment, referred to as a dynamic mechanical experiment in Figure 70.1, the applied shear strain is an oscillatory function with an angular frequency, ω, and an amplitude, e_0:

$$e(t) = e_0 \sin(\omega t) \tag{70.2}$$

Note that the strain rate, $\frac{de}{dt}$, is also an oscillatory function, with the same angular frequency, but shifted with respect to the strain by 90 °:

$$\frac{de}{dt} = e_0 \omega \cos(\omega t) = \gamma_0 \omega \sin(\omega t + \pi/2) \tag{70.3}$$

The resulting stress is also an oscillatory function with an angular frequency of ω, and is described by its amplitude and by the phase shift of the relative to the applied strain:

$$\sigma(t) = \sigma_0 \sin(\omega t + \phi) \tag{70.4}$$

Now we can define a complex modulus with real and imaginary components as follows:

$$E^* = E' + iE'' = |E^*| e^{i\phi} \tag{70.5}$$

There are a couple different ways to think about the complex modulus, E^*. As a complex number we can express it either in terms of its real and imaginary components (E' and E'', respectively), or in terms of its magnitude, $|E^*|$ and phase, ϕ. The magnitude of the complex modulus is simply the stress amplitude normalized by the strain amplitude:

$$|E^*| = \sigma_0 / e_0 \tag{70.6}$$

The phase angle, ϕ, describes the lag between the stress and strain in the sample. Examples for $\phi = 90°$ (the maximum value, characteristic of a liquid) and $\phi = 45°$ are shown in Figure 70.2.

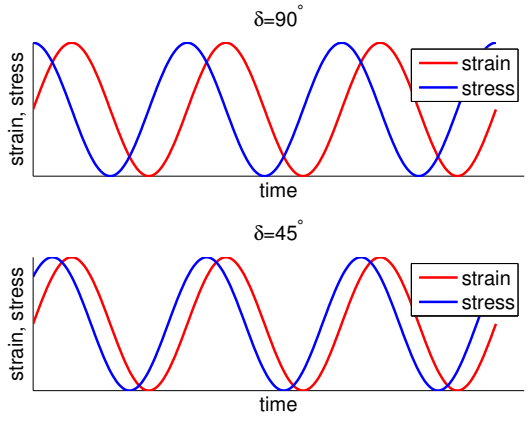


Figure 70.2: Time dependent stress and strain in a dynamic mechanical experiment.

In order to understand the significance of the real and imaginary components of E^* we begin with the Euler formula for the exponential of an imaginary number:

$$e^{i\phi} = \cos\phi + i \sin\phi \tag{70.7}$$

use this expression in Eq. 70.5, we see that thestorage modulus, E' gives the stress that is in phase with the strain (the solid-like part), and is given by the following expression:

$$E' = |E^*| \cos\phi \tag{70.8}$$

Similarly, theloss modulus, E'', gives the response of the material that is in phase with the strain rate (the liquid-like part):

$$E'' = |E^*| \sin\phi \tag{70.9}$$

70.2 Torsional Resonator

We can now consider the specific example of a torsional resonator, shown below in Figure 70.3. The equation of motion for the spring as it is being twisted is[10]:

$$T = K_t\theta + I\frac{d^2\theta}{dt^2} \tag{70.10}$$

where θ is the rotation angle, T is the torque on the resonator, K_t is the torsional stiffness and $I$is the moment of inertia of the system, which depends on the details of the inertial bar used in the experiment. For a cylindrical fiber K_t is obtained from Eq. 91.11:

$$K_t \equiv \frac{T}{\theta} = \frac{\pi G d^4}{32\ell} \tag{70.11}$$

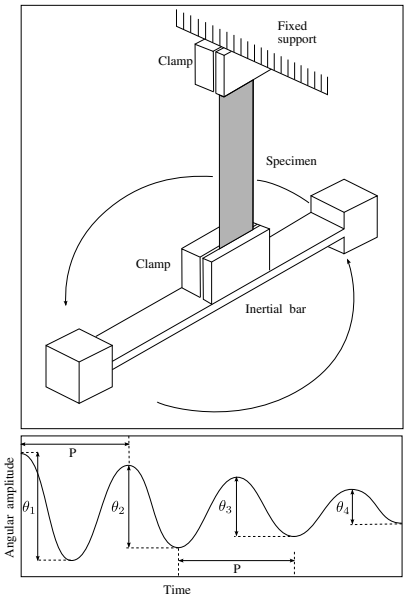


Figure 70.3: Torsional resonator.

The resonant angular frequency of the oscillator, ω_n is given by the following expression:

$$\omega_n = \sqrt{\frac{K_t}{I}} \tag{70.12}$$

The solution to Eq. 70.10 for the case where $T = 0$ and we just let the fiber move (by twisting to a certain angle and letting it go, for example) is as follows:

$$\theta = \theta_0 \cos(\omega_n t) \tag{70.13}$$

This analysis assumes that the system is entirely elastic. What if the spring has some viscoelastic character to it? We begin by using Euler's formula (Eq. 70.7) to rewrite Eq. 70.13 in the following way:

$$\theta = \theta_0 \text{Real}\left(\exp\left(i\omega_n t\right)\right) \tag{70.14}$$

Now all we need to do is to replace the modulus, G, with the complex modulus, G^*, and everything works out fine. Here's what happens:

- The torsional stiffness, K_t gets transformed to a complex torsional stiffness K_t^*
- The resonant frequency, ω_n gets transformed to a complex resonant frequency, ω_n^*
- The imaginary part of the complex resonant frequency becomes an exponential damping function

To understand this last point, suppose write ω_n in the following way:

$$\omega_n = \omega_n' + i\omega_n'' \tag{70.15}$$

Putting this back into Eq. 70.14 gives:

$$\theta = \theta_0 \text{Real}\left(\exp\left(i\omega_n' t\right)\right) \exp\left(-\omega_n'' t\right) \tag{70.16}$$

Using the Euler formula (Eq. 70.7) to expand $\exp\left(i\omega_n' t\right)$ and taking the real part gives:

$$\theta = \theta_0 \cos\left(\omega_n' t\right) \exp\left(-\omega_n'' t\right) \tag{70.17}$$

So the imaginary part of the complex frequency describes the decay of oscillation with time.

70.3 Viscoelastic Models

Models of viscoelasticity can be represented visually by connecting solid-like and liquid-like elements together. In materials science we're used to the solid-like elements. They are simply elastic springs where the force is proportional to the extension of the spring. In terms of stress and strain, the stress, σ is proportional to the strain, e. There is not time-dependence to the behavior of an ideal spring. The stress is proportional to the strain, regardless of how fast or slow the strain is applied. In many real materials the rate at which the strain is applied matters as well, and this is where the dashpots come in. A dashpot is a liquid-like element where the stress is proportional to the rate at which the strain is changing. The proportionality constant is the viscosity, η. Schematic representations of springs and dashpots are shown in Figure 70.4.

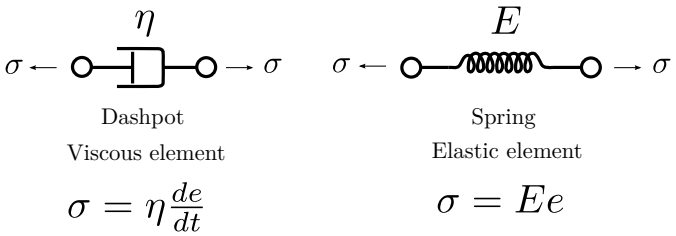


Figure 70.4: Schematic representations of springs and dashpots used to represent the time-dependent response of viscoelastic materials.

As an simple illustration of the significance of the viscosity, we can calculate the time it takes for a small sphere of metal to drop to the bottom of pool of water. Suppose we use an iron particle with a radius of 1 μm. The situation is as shown in Figure 70.5. The gravitational force, F, causing the ball to sink is proportional to the volume of the sphere, and the difference in densities between the solid and liquid:

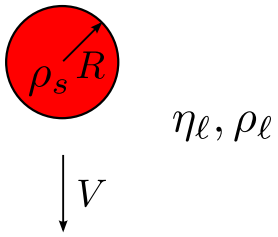


Figure 70.5: Terminal velocity of a sphere descending in a viscous fluid.

$$F = g\frac{4}{3}\pi R^3\left(\rho_s - \rho_\ell\right) \tag{70.18}$$

Here ρ_s is the density of the solid sphere, ρ_ℓ is the density of the surrounding liquid, R is the radius of the sphere and g is the gravitational acceleration (9.8 m/s ^{2}). This force is balanced by the viscous force exerted by the water on the sphere as the sphere moves the liquid with a velocity, v. For a viscous liquid this force is proportional

to the velocity. It is also proportional to the viscosity of the liquid and to the radius of the sphere. The specific relationship is as follows:

$$F = 6\pi V \eta R \tag{70.19}$$

At steady state, the velocity reaches a value that is obtained by equating the forces in Eqs.70.18 and70.19. In this way we obtain:

$$V = \frac{2}{9}\frac{gR^2}{\eta}\left(\rho_s - \rho_\ell\right) \tag{70.20}$$

Iron has a density of 7.87 g/cm ^{3} (7870 kg/m ^{3}) and water has a density of 1000 kg/m ^{3} and a viscosity of 10 $^{-3}$ Pa-sec. From this we get a velocity of 16 μm/s.

70.3.1 Maxwell model

The simplest viscoelastic model that contains both liquid-like and solid-like elements is the Maxwell model consisting of a linear dashpot with viscosity η in series with a linear spring with a modulus, E (see Figure 92.9). Because the elements are in series the strain is the same in each one of them. This stress can be related to the strain e_1 in the dashpot and the strain e_2 in the spring through the following expressions:

$$\sigma = \eta \frac{de_1}{dt} \tag{70.21}$$

$$\sigma = Ee_2 \tag{70.22}$$

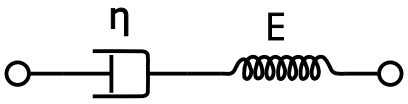


Figure 70.6: Maxwell model.

The total strain, e is $e_1 + e_2$, so the time derivative of the total strain is:

$$\frac{de}{dt} = \frac{de_1}{dt} + \frac{de_2}{dt} = \frac{\sigma}{\eta} + \frac{1}{E}\frac{d\sigma}{dt} \tag{70.23}$$

In a stress relaxation experiment the strain instantaneous increased to an initial value of e_0 and we follow the relaxation of the stress at this fixed strain. Because the dashpot cannot respond instantaneously, all of the initial strain is in the spring,*i.e.*, $e_0 = e_2$ and the initial stress, σ_0 is equal to Ee_0. The solution to Eq. 70.23 for this initial condition and for $de/dt = 0$ is:

$$\sigma\left(t\right) = Ee_0 \exp\left(-t/\tau\right) \tag{70.24}$$

with $\tau = \eta/E$. We divide by e_0 to obtain the relaxation modulus:

$$E\left(t\right) \equiv \sigma\left(t\right)/e_0 = E\exp\left(-t/\tau\right) \tag{70.25}$$

We can also obtain the solution for the stress in the case where we apply an oscillatory strain. The math is a bit more complicated, but we end up with the following complex modulus:

$$E^*\left(\omega\right) = \frac{Ei\omega\tau}{1 + i\omega\tau} \tag{70.26}$$

We can obtain the real and imaginary parts of E^* by multiplying the top and bottom of Eq. 70.26 by $1 + i/\left(\omega\tau\right)$, remembering that $i^2 = -1$. Taking the real and imaginary components of E^* gives E' and E'':

$$E'\left(\omega\right) = \frac{E\left(\omega\tau\right)^2}{1 + \left(\omega\tau\right)^2} \tag{70.27}$$

$$E''\left(\omega\right) = \frac{E\omega\tau}{1 + \left(\omega\tau\right)^2} \tag{70.28}$$

70.3.2 Standard Linear Solid

For a single maxwell element the relaxation modulus decays to zero at long times ($t \gg \tau$). In many real systems we are interested in describing the day of the relaxation modulus from a large value at short times to a much smaller value at larger times. Thestandard linear solid accounts for this by putting an elastic element with a modulus of E_r (the 'relaxed' modulus). In parallel with a single Maxwell element as shown in Figure 70.7. Extra spring adds directly to stress response, so all we have to do is add E_r to the expressions for $E(t)$ and $E^*(w)$ that were obtained from the Maxwell model:

$$E(t) = E_r + E_1 \exp(-t/\tau_1) \tag{70.29}$$

$$E^*(\omega) = E_r + \frac{E_1 i\omega\tau_1}{1 + i\omega\tau_1} \tag{70.30}$$

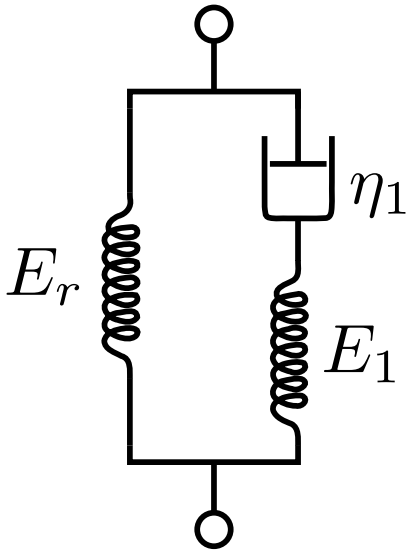


Figure 70.7: Standard linear solid.

So we see that the standard linear solid describes the relaxation of the modulus from an initial value of $E_1 + E_r$ at very short times to a relaxed value of E_r at long times. The standard linear solid is the simplest model for describing the transition of an amorphous, crosslinked polymer from glassy polymer behavior to rubbery behavior with typical values of E_1 and E_r being 10^9 Pa and 10^6 Pa, respectively.

70.3.3 Generalized Maxwell Model

For real systems, the $E(t)$ and $E^*(\omega)$ are usually much more complicated than given by the predictions of the Maxwell model or standard linear solid. To describe the behavior for real materials, we can add an arbitrary number of Maxwell elements in parallel, resulting in the generalized Maxwell model shown in Figure 70.8. The stresses for each of the parallel elements are additive, so $E(t)$ and $E^*(w)$ are given by the following expressions:

$$E(t) = E_r + \sum_j E_j \exp\left(-t/\tau_j\right) \tag{70.31}$$

$$E^*(\omega) = E_r + \sum_j \frac{E_j i\omega\tau_j}{1 + i\omega\tau_j} \tag{70.32}$$

with $\tau_j = \eta_j / E_j$.

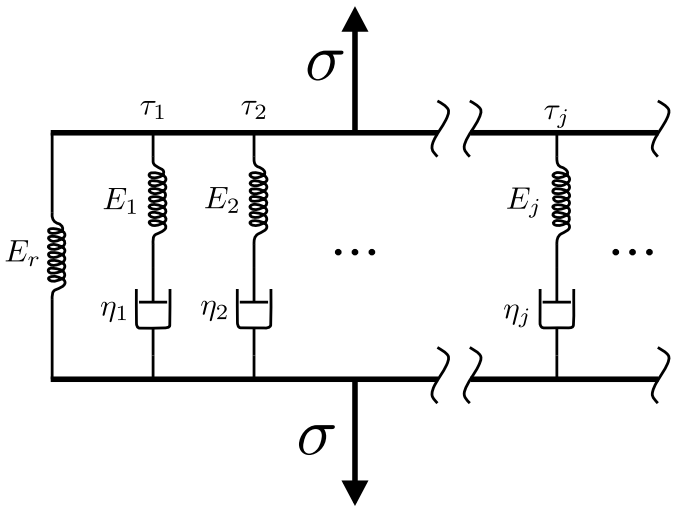


Figure 70.8: Generalized Maxwell model.

70.3.4 Kelvin-Voigt Model

The Kelvin-Voigt model consists of a spring and dashpot in parallel, as shown in Figure 70.9. In this case the stresses in two elements are additive and the strains in the two elements are the same:

$$\sigma = \eta \frac{de}{dt} + Ee \tag{70.33}$$

The Kelvin-Voigt model is the simplest model for the description of a creep experiment, where the stress jumps instantaneously from 0 to $\sigma = \sigma_0$ and we track the strain as a function of time. We have:

$$e(t) = \frac{\sigma_0}{E}\left(1 - \exp\left(t/\tau\right)\right); \quad \tau = \eta/E \tag{70.34}$$

The dynamic modulus for the Kelvin-Voigt element is given by the following expression:

$$E^* = E + i\omega\eta \tag{70.35}$$

So that E' is simply equal to E, and E'' is equal to $\omega\eta$.

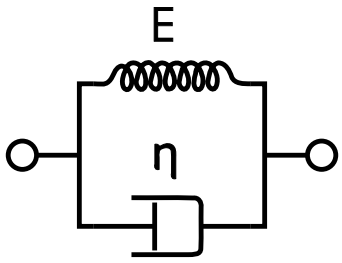


Figure 70.9: Kelvin-Voigt Model.

71 Creep Behavior

In a creep experiment we apply a fixed stress to a material and monitor the strain as a function of time at this fixed stress. This strain is generally a function of both the applied stress, σ and the time, t, since the load was applied.

71.1 Creep in the Linear Viscoelastic Regime

In the linear viscoelastic regime the strain is linearly dependent on the applied stress, σ, allowing us to define thecreep compliance function, $J(t)$, in the following way:

$$J(t) \equiv \frac{e(\sigma, t)}{\sigma} \tag{71.1}$$

An example of the sort of spring/dashpot model used to describe creep behavior of a linear viscoelastic material is shown in Figure 71.1. The model consists of a single Kelvin-Voigt element in series with a spring and a dashpot. In this example the strain obtained in response to a jump in the stress from 0 to σ at $t = 0$ can be represented as the sum of three contributions, e_1, e_2, and e_3:

$$e(\sigma, t) = e_1(\sigma) + e_2(\sigma, t) + e_3(\sigma, t) \tag{71.2}$$

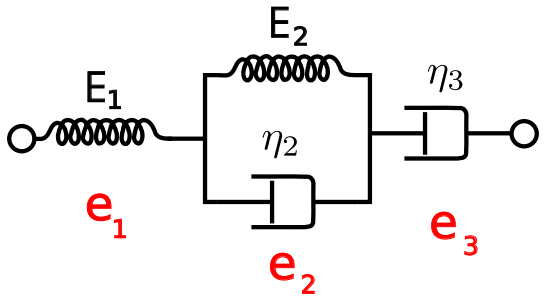


Figure 71.1: Linear creep model.

In this equation e_1 corresponds to an instantaneous elastic strain, e_2 is the recoverable viscoelastic strain and e_3 is the plastic strain, with the three different components illustrated in Figure 71.2, and given by the following expressions:

$$\begin{aligned} e_1 &= \tfrac{\sigma}{E_1} \\ e_2 &= \tfrac{\sigma}{E_2}\left[1 - \exp\left(-t/\tau_2\right)\right] \\ e_3 &= \tfrac{\sigma}{\eta_3} t \end{aligned} \tag{71.3}$$

In many cases we are interested in situations where the strain does not increase linearly with the applied stress. The yield point is one obvious example of nonlinear behavior. In ideally plastic system, we only have elastic strains for stresses below the yield stress, but above the yield stress the strains are much larger.

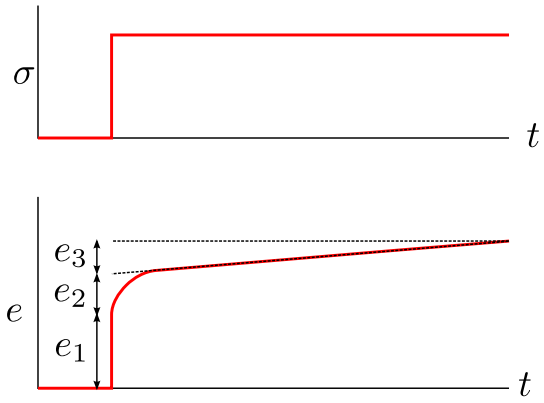


Figure 71.2: Strain contributions in a nonlinear creep model.

71.2 Nonlinear Creep: Potential Separability of Stress and Time Behaviors

The situation becomes much more complicated in the nonlinear regime, where it is no-longer possible to define a stress-independent creep compliance function. In general the strain in the nonlinear regime is a complex function of both the time and the applied stress. In some cases, however, we can separate the stress dependence from the time dependence and write the stress in the following way:

$$e(\sigma, t) = f(\sigma) J(t) \tag{71.4}$$

If this sort of separability holds, then it is possible to make predictions of creep based on limited experimental data. The procedure is illustrated in Figure 71.3 and involves the following steps:

(1) Measure $e(t)$ at different different stresses (σ_1 and σ_2) in Figure 71.3. If the ratio $e(\sigma_1, t) / e(\sigma_2, t)$ is constant for all value of t, then the separability into stress dependent and time-dependent functions works (at least for these two stress levels). These experiment enable us to obtain the time dependent function $J(t)$.

(2) To get the stress-dependent function, $f(\sigma)$, it is sufficient to make a series of measurements at a single, experimentally convenient time.

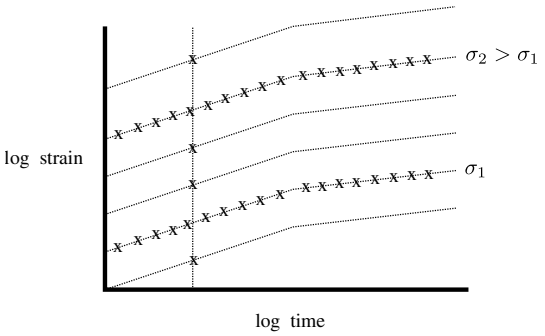


Figure 71.3: Creep response of a material with separable dependencies on stress and time.

71.3 Use of empirical, analytic expressions

The procedure outlined in the previous Section only works if the ratio $e(\sigma_1, t) / e(\sigma_2, t)$ is independent of the time. This is not always the case. However, it is often possible to fit the data to relatively simple models. These models are similar to the spring and dashpot models of Section 70, but a linear stress response is not necessarily assumed. One example corresponds to a nonlinear version of the linear model shown in Figure 71.1. As with the linear model, the strain components are assumed to consist of an elastic strain, e_1, a recoverable viscoelastic strain, e_2 and a plastic strain, e_3. However, we now use nonlinear elements to describe e_2 and e_3, with these two strain components given by the following expressions:

$$\begin{aligned} e_1 &= \tfrac{\sigma}{E_1} \\ e_2 &= C_1 \sigma^n \left[1 - \exp\left(-C_2 t\right)\right] \\ e_3 &= C_3 \sigma^n t \end{aligned} \tag{71.5}$$

Not that the material behavior is specified by 5 constants, E, C_1, C_2, C_3, n, that we need to obtain by fitting to actual experimental data. For a linear response ($n = 1$) we can make a connection to the spring and dashpot models described earlier. In this case the behavior of the material is represented by the model of linear viscoelastic elements shown in Figure 71.4, and the constants appearing in Eq. 71.5 correspond to the following linear viscoelastic elements from Figure 71.1:

$$\begin{aligned} E &= E_1 \\ C_1 &= 1/\eta_2 \\ C_2 &= E_2/\eta_2 \\ C_3 &= 1/\eta_3 \end{aligned} \tag{71.6}$$

This is just one possible nonlinear model that can be used. An additional nonlinear element is obtained from Eyring rate theory, and is described in the following Section.

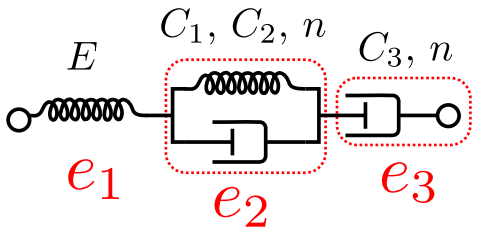


Figure 71.4: Nonlinear creep model.

71.4 Eyring Model of Steady State Creep

TheEyring rate model of creep is a very general model aimed at understanding the effect of the stress on the flow properties of a material. It can be used to describe a very wide range of materials, and is based on the modification of the activation energy for material for flow by the applied stress.

71.4.1 Material Deformation as a Thermally Activated Process

Our starting point is to realize that material deformation is a thermally activated process, meaning that there is some energy barrier that needs to be overcome in order for deformation to occur. The general idea is illustrated in Figure 71.5. The stress does an amount of work on the system equal to σv, where σ is the applied stress and v is the volume of the element that moves in response to this applied stress. The quantity v is typically referred to as anactivation volume. The net result of the application of the stress is to reduce the activation barrier for motion in the stress direction by an amount equal to $v\sigma/2$ and to increase the activation barrier in the opposite direction by this same amount.

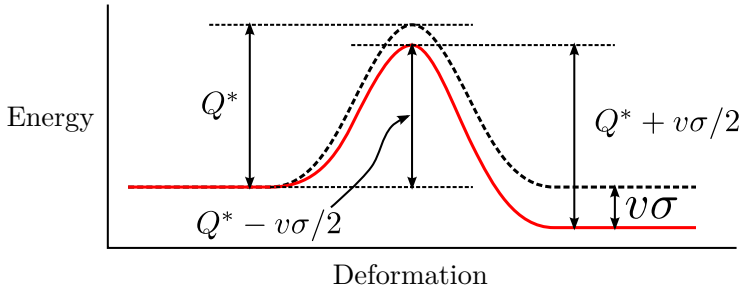


Figure 71.5: Effect of an applied stress on a thermally activated creep process.

Eyring Rate Law We can develop an expression for the strain rate by recognizing that the net strain rate is given by the net frequency of hops in the forward direction. The frequency of hops in the forward and reverse directions, which we refer to as f_1 and f_2, respectively, are given as follows:

$$f_1 = f_0 \exp\left(-\frac{Q^* - v\sigma/2}{k_B T}\right) = \exp\left(-\frac{Q^*}{k_B T}\right) \exp\left(\frac{v\sigma}{2k_B T}\right) \tag{71.7}$$

$$f_2 = f_0 \exp\left(-\frac{Q^* + v\sigma/2}{k_B T}\right) = \exp\left(-\frac{Q^*}{k_B T}\right) \exp\left(\frac{-v\sigma}{2k_B T}\right) \tag{71.8}$$

The net rate of strain is proportional to $f_1 - f_2$, the net frequency of hops in the forward direction:

$$\frac{de}{dt} = A\left(f_1 - f_2\right) = A f_0 \exp\left(-\frac{Q^*}{k_B T}\right)\left[\exp\left(\frac{v\sigma}{2k_B T}\right) - \exp\left(\frac{v\sigma}{2k_B T}\right)\right] \tag{71.9}$$

Where A is a dimensionless constant. Using the following definition of the hyperbolic sine function,

$$\sinh\left(x\right) = \left(e^x - e^{-x}\right)/2 \tag{71.10}$$

we obtain the following expression for the strain rate:

$$\frac{de}{dt} = 2A f_0 \exp\left(-\frac{Q^*}{k_B T}\right) \sinh\left(\frac{v\sigma}{2k_B T}\right) \tag{71.11}$$

Low Stress Regime In the low-stress regime we can use the approximation $\sinh(x) \approx x$ to get the following expression, valid for $v\sigma \ll k_B T$.

$$\frac{de}{dt} = \frac{Af_0 v\sigma}{k_B T} \exp\left(-\frac{Q^*}{k_B T}\right) \tag{71.12}$$

In this regime the strain rate is linear in stress, which means that we can define a stress-independent viscosity from the following expression:

$$\frac{de}{dt} = \frac{\sigma}{\eta} \tag{71.13}$$

Comparing Eqs.71.12 and71.13 gives the following for the viscosity:

$$\eta = \frac{k_B T}{Af_0 v} \exp\left(-\frac{Q^*}{k_B T}\right) \tag{71.14}$$

So the Eyring theory reduces to and Arrhenius viscosity behavior in the linear, low-stress regime.

High Stress Regime In the high-stress regime, we use the fact that $\sinh(x) \approx \frac{\exp(x)}{2}$ for large x to obtain the following expression for $v\sigma \gg k_B T$:

$$\frac{de}{dt} = Af_0 \exp\left(-\frac{Q^*}{k_B T}\right) \exp\left(\frac{v\sigma}{2k_B T}\right) \tag{71.15}$$

Equivalently, we can write the following:

$$\frac{de}{dt} = A_0 \exp\left(-\frac{Q^* - v\sigma/2}{k_B T}\right) \tag{71.16}$$

where $A_0 \equiv Af_0$. The effective activation energy decreases linearly with increasing stress, giving a very nonlinear response. In practical terms the activation volume is obtained by plotting $\ln(de/dt)$ vs. σ, with the slope being equal to $v/2k_B T$.

Physical significance of v Eyring rate models are most often used for polymeric systems. In this case the activation volume can be viewed as the volume swept out the portions of a polymer molecule which move during a fundamental creep event, as illustrated schematically in Figure 71.6. A large activation volume means that cooperative deformation of a large region of the material is required in order for the material to flow. Low values of the activation volume indicate that the deformation is controlled by a very localized event, corresponding, for example, to the rupture of a single covalent bond.

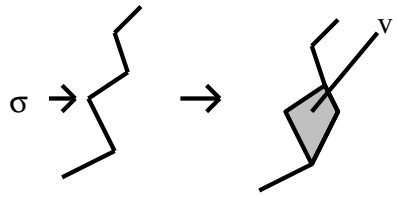


Figure 71.6: Illustration of the activation volume.

71.4.2 Additional Nonlinear Dashpot Elements

Nonlinear elements based on the Eyring rate model can also be included in our spring and dashpot models of viscoelastic behavior. For example, we can use an Eyring rate model to describe the stress dependence of the steady state creep of a nonlinear model, in which case Figure 71.4 gets modified to the following. The steady state component is specified by the prefactor A_0, the activation energy $Q^*$and the activation volume, v.

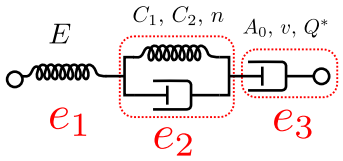


Figure 71.7: Nonlinear creep model, including an Eyring rate element for the steady-state creep component.

71.5 High Temperature Creep in Metals

by 'high' temperature for a metal, we generally mean temperatures above half the absolute melting temperature, which we list below in Figure 71.8.

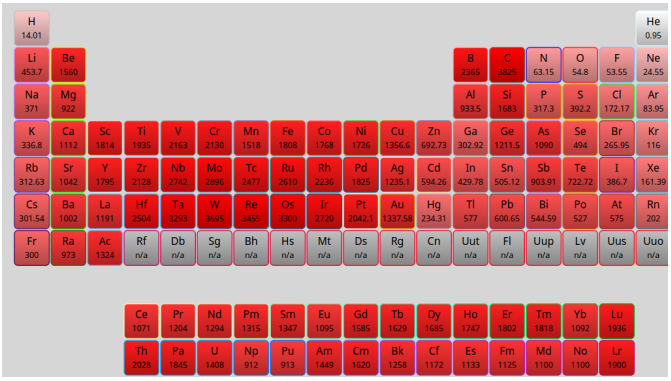


Figure 71.8: Absolute melting temperatures (K) for the elements.

Typical creep response for a metal is shown in Figure 71.9. The easiest way to do the experiment is under constant load conditions. Because the cross sectional area of the sample decreases as the material deforms, the true stress increases during a constant load creep experiment. If the experiment is done under conditions where the load is reduced to maintain a constant true stress, the creep rate does not increase as fast at the longer times. The increase in the creep rate at long times for the constant load test can be attributed to the increased true stress in the sample at the later stages of the creep experiment. Failure of the sample during a constant load test is referred to as thecreep rupturepoint. The creep rupture time decreases with increasing temperature and increasing applied engineering stress, as illustrated by the data for an iron alloy in Figure 71.10.

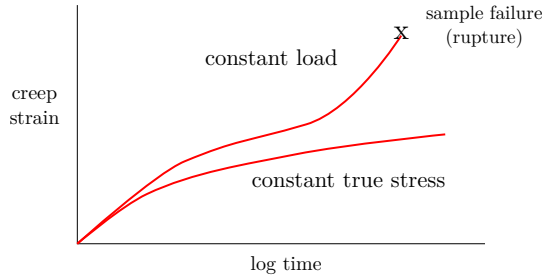


Figure 71.9: Qualitative difference between creep behavior under constant load and constant stress conditions.

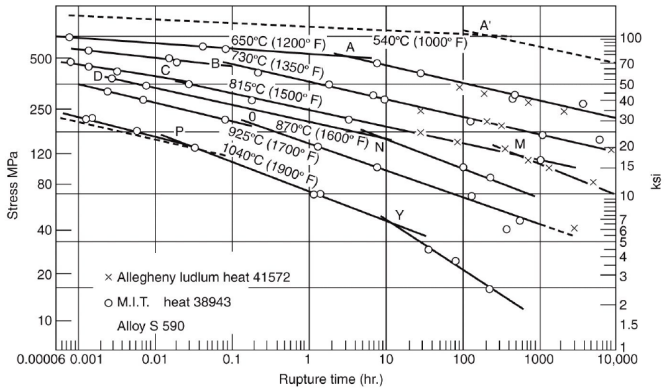


Figure 71.10: Creep rupture data for an iron-based alloy.

Different creep regimes during an engineering (constant load) creep rupture test are illustrated in Figure 71.11. At the beginning off the experiment the sample strain hardens, and the strain rate decreases with time (stage I). Following this there is often a linear regime (stage II in Figure 71.11) where the creep strain increases linearly in time, with a slope equal to the steady state creep rate, $\dot{\epsilon}_{ss}$. Steady state creep data for TiO_2 at different times and temperatures are shown in Figure 71.12. The data indicate that the effects of stress and temperature are separable, with the same activation energy of 280 J/mol obtained for each of the different stress levels. The overall steady state creep rate in this case can be expressed in the following way:

$$\dot{\epsilon}_{II} = f(\sigma) \exp\left(-\frac{Q_c}{k_B T}\right) \tag{71.17}$$

For the specific case of power law creep, the stress dependence is given by a simple power law:

$$f(\sigma) = A\sigma^n \tag{71.18}$$

where n is an empirically determined exponent, with typical values ranging between 1 and 7.

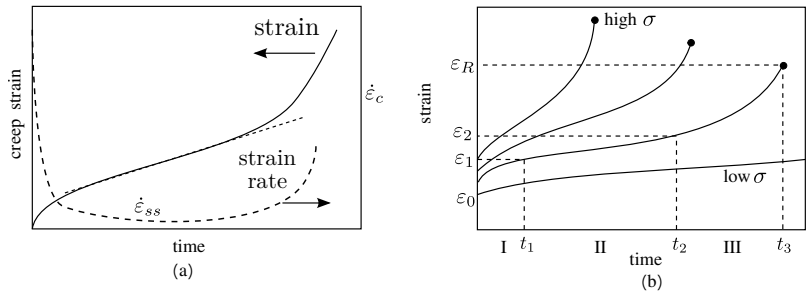


Figure 71.11: Commonly observed creep regimes, illustrating the steady state creep rate, $\dot{\epsilon_{ss}}$.

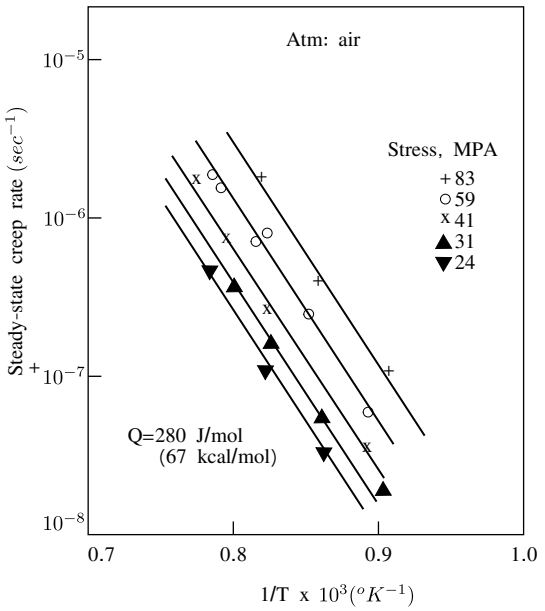


Figure 71.12: Creep behavior of TiO_2 (from ref.[18]).

In the high temperature regime, the activation energy for creep, Q_c, is very similar to the activation energy for self-diffusion in the metal (see Figure 71.13), indicating that high temperature creep is controlled by the same processes that control self diffusion. As we describe below, this common process is vacancy diffusion. Vacancy diffusion can contribute to the creep behavior by enabling dislocation climb, or by changing the shapes of individual grains in a manner that does not require the existence of dislocations. These mechanisms are Nabarro-Herring creep and Coble Creep, as described below.

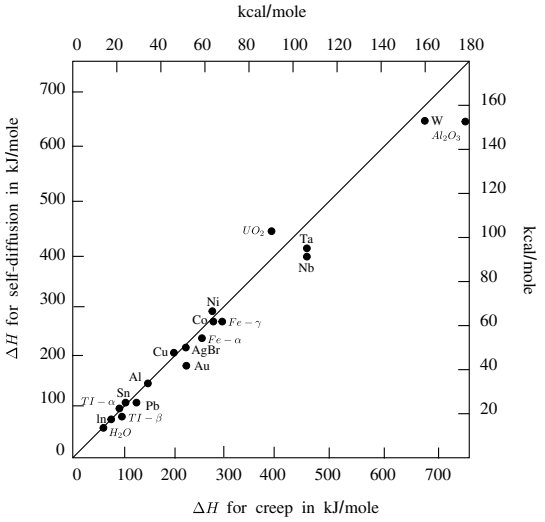


Figure 71.13: Activation energies for creep (for $T/T_m > 0.5$) and atomic diffusion (from ref.[35]). a

71.5.1 Nabarro-Herring Creep: Bulk Diffusion Within a Grain

Nabarro-Herring creep is the mechanism by which vacancies move away from regions of the largest tensile stress in the sample. The mechanism is as illustrated in Figure 71.14. The mechanism is based on the effect that stresses have on the local concentration of voids in the system.

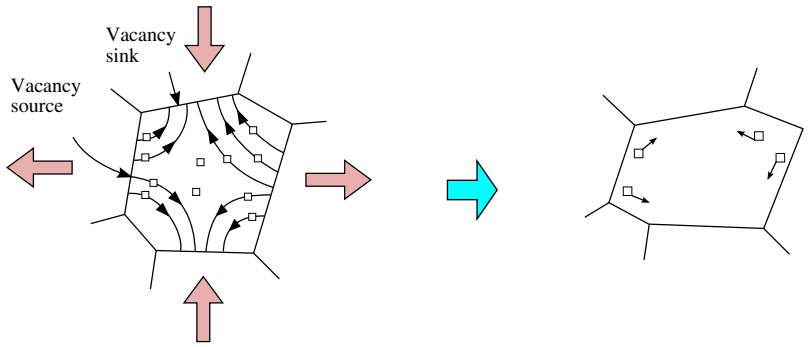


Figure 71.14: Conceptual illustration of Nabarro-Herring creep.

In the absence of an applied stress the equilibrium number of concentration of vacancies, C_v^0 is given by the vacancy formation energy, G_v:

$$C_v^0 = \frac{1}{\Omega} \exp\left(-\frac{G_v}{k_B T}\right) \tag{71.19}$$

where Ω is the atomic volume (also equal to the volume of a vacancy). Now suppose we have a grain that is experiencing a tensile stress of magnitude σ at the top and bottom, and a compressive force of this same magnitude at the sides. We'll call the vacancy concentrations in the tensile and compressive regions of the sample C_v^t and C_v^c, respectively, so we have:

$$\begin{aligned} C_v^t &= C_v^0 \exp\left(\sigma\Omega/k_B T\right) \\ C_v^c &= C_v^0 \exp\left(-\sigma\Omega/k_B T\right) \end{aligned} \tag{71.20}$$

We can now use Fick's first law to to calculate the steady state flux of vacancies, J_v, from the tensile region to the compressive region:

$$J_v = -D_v \frac{dC_v}{dx} \tag{71.21}$$

We approximate the the vacancy concentration gradient, dC_v/dx, as the difference in vacancy concentrations in the tensile and compressive regions of the grain, divided by the grain size, d:

$$\frac{dC_v}{dx} \propto \frac{C_v^t - C_v^c}{d} \tag{71.22}$$

Combining Eqs.71.21 and71.22 gives:

$$J_v \propto -D_v \left(\frac{C_v^t - C_v^c}{d}\right) \tag{71.23}$$

In order to relate the flux (units of vacancies/m ^{2}·s) to the velocity of the grain edge (in m/s), we need to multiply J_v by the atomic volume, Ω:

$$v = \Omega J_v \propto \Omega D_v \left(\frac{C_v^t - C_v^c}{d}\right) \tag{71.24}$$

Now we get the strain rate by dividing this velocity by d :

$$\frac{de}{dt} = \frac{\Omega J_v}{d} \propto \Omega D_v \left(\frac{C_v^t - C_v^c}{d^2}\right) = \frac{2D_v}{d^2} \sinh\left(\frac{\sigma\Omega}{k_B T}\right) \exp\left(-\frac{G_v}{k_B T}\right) \tag{71.25}$$

For real systems the stresses are small enough so that $\sigma\Omega \ll k_B T$, so we can use the fact that $\sinh(x) \approx x$ for small x to obtain:

$$\frac{de}{dt} \propto \frac{2D_v}{d^2} \frac{\sigma\Omega}{k_B T} \exp\left(-\frac{G_v}{k_B T}\right) \tag{71.26}$$

71.5.2 Coble Creep: Grain Boundary Diffusion

If diffusion occurs predominantly along the grain boundaries, the mechanisms is slightly different and is referred to asCoble Creep. The expression for the strain rate is very similar to the expression for Nabarro-Herring creep, with two difference. First, we replace the bulk vacancy diffusion coefficient, D_v in Eq. 71.21 with the grain boundary diffusion coefficient, D_{gb}, which is generally larger than D_v. In addition, we modify the relationship between J_v to account for the fact that vacancies are not diffusing throughout the entire grain, but only along a narrow grain of width d_{gb}. As a result we need to multiply the velocity in Eq. 71.24 by the overall volume fraction of the grain boundary, which is proportional to d_{gb}/d. The net result is the following expression for the strain rate for Coble creep:

$$\frac{de}{dt} \propto \frac{2D_{gb}d_{gb}}{d^3}\frac{\sigma\Omega}{k_BT}\exp\left(-\frac{G_v}{k_BT}\right) \tag{71.27}$$

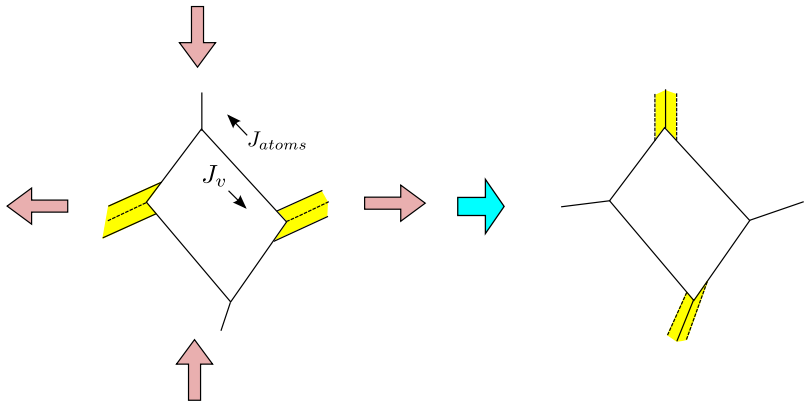


Figure 71.15: Conceptual illustration of Coble creep.

71.5.3 Dislocation Creep (Weertman model):

$$\frac{de}{dt} = A\sigma^3 \sinh\left(B\sigma^{1.5}/k_BT\right)\exp\left(-Q/k_BT\right) \tag{71.28}$$

71.6 Deformation mechanism maps

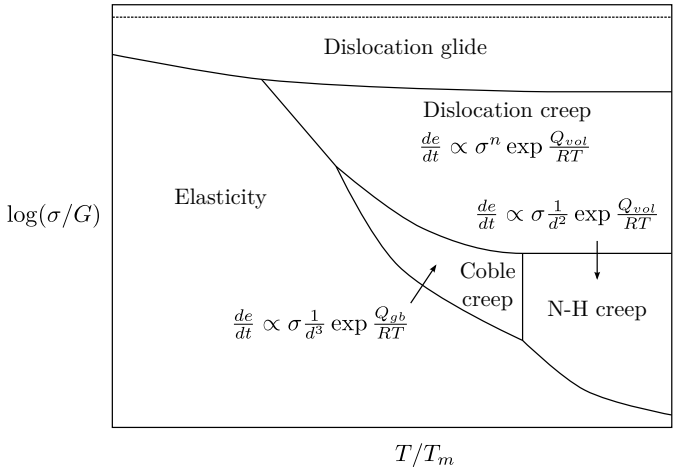


Figure 71.16: Deformation mechanism map.

72 Weibull Analysis of Failure

Failure of brittle materials is determined by the largest flaw size, since the largest flaw size will have the largest value of K for a given applied stress. As an example of the use of Weibull statistics, consider the adhesive transfer of a thin layer of a material from a flat, flexible substrate to a rigid, curved indenter as illustrated in Figure 72.1. The basic geometry of the experiment is illustrated in Figure 72.1a, and consists of a thin, viscoelastic film that is coated on an elastomeric substrate. A hemispherical glass indenter is brought into contact with the film and is then pulled away from the surface. The system is designed so that the adhesion of the film to the glass indenter is stronger than the adhesion to the elastomeric substrate, the film will be transfered from the substrate to the indenter. The process occurs by the sequence of steps illustrated in parts b-e of Figure 72.1:

- b) A crack is nucleated at a defect site at the interface between the film and the substrate at a region where the hydrostatic pressure is maximized.
- c) This crack propagates under the indenter as the material is the indenter is pulled away from the substrate.
- d) Eventually the entire film in has detached from the substrate over the region where it is contact with the indenter. The remainder of the film begins to peel away from the substrate surface.
- e) The film breaks at the edge of the area of contact with the indenter.

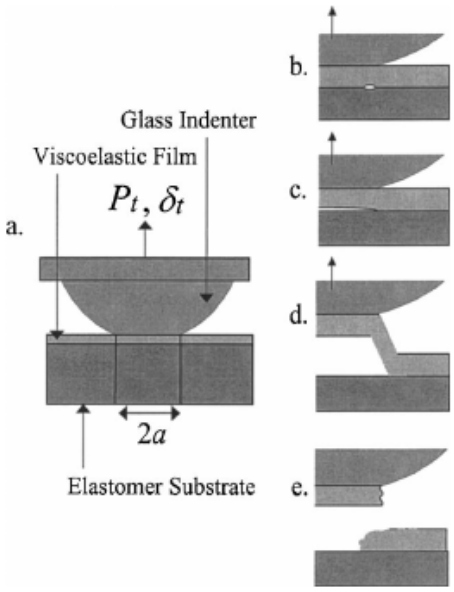


Figure 72.1: Adhesive transfer of a thin, viscoelastic film.

Details of this experiment can be found in reference[26]. The most important thing to keep in mind is that the whole transfer process is controlled by the initial appearance of a cavity at the indenter/substrate interface (Figure 72.1b). This happens when p_{max}, the maximum hydrostatic tension at the film/substrate interface, reaches a critical value that we refer to as p_{cav}. Qualitatively we find that p_{cav} is close to E_{sub}, the elastic modulus of the substrate, but cavitation occurs for different values of p_{cav}. TheWeibull distribution for the survival probability, P_s (the probability that cavitation has NOT occurred) is as follows:

$$P_s = \exp\left[-\left(\frac{p_{max}}{p_{cav}}\right)^M\right] \tag{72.1}$$

Here M is the Weibull modulus, which is a measure of the distribution of failure probabilities. We generally want M to be large, so that the distribution stresses is very narrow. For example, if $M \to \infty$, $P_s = 0$ for $p_{max} > p_{cav}$ and $P_s = 1$ for $p_{max} < p_{cav}$.
We can take natural logs of both sides a couple of times to convert Eq. 72.1 to the following:

$$\ln\left[\ln\left(1/P_s\right)\right] = M \ln p_{max} - M \ln p_{cav} \tag{72.2}$$

73 Intrinsic and Extrinsic Fracture Toughness

Intrinsic Toughness: characteristic value of a material. K_{IC} independent of crack length.
Extrinsic Toughness: factors that affect the nature of the stress distribution in advance of a crack. K_{IC} generally increases with crack length.

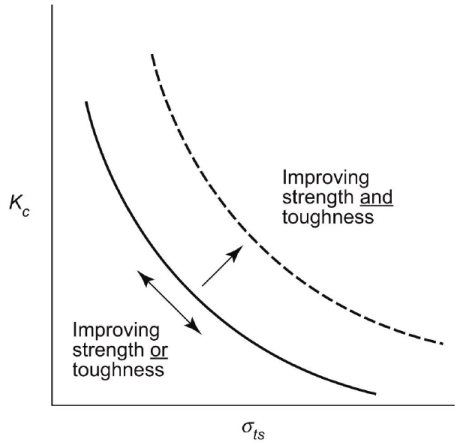


Figure 73.1: Schematic representation of toughness vs. tensile strength.

73.1 Extrinsic Toughening

Extrinsic toughening is based on the modification of the stress field in the region of a growing crack. We can describe it by dividing the macroscopic stress intensity factor, K_I, into two parts, K_{tip} and K_s:

$$K_I = \sigma_0\sqrt{\pi a} = K_{tip} + K_s$$

Here K_{tip} describes the stress distribution very close to the crack tip in the usual way. For example, the tensile stress directly in front of a crack tip is given as:

$$\sigma_{zz}(d) = \frac{K_{tip}}{\sqrt{2\pi d}}$$

The factor K_s describes a reduction in K_{tip} compared to K_s because of other internal stresses that are applied in some way to the sample in the vicinity of the crack tip. We'll consider two examples. In transformation toughening, compressive stresses result from a phase transition in front of the crack tip to a lower density (higher volume) phase, thereby accommodating some of the strains in the stress field defined by K_I, so that $K_{tip} < K_I$ ($K_s > 0$). The second example is a conceptually similar crack bridging example, where fibers bridge the crack, (much like craze fibrils in the polymer crazing example discussed above).

73.1.1 Transformation toughening

Zirconia:
First of all, we do we care about a material like zirconia (ZrO_2)? It makes a pretty artificial diamond if we can get it in its cubic crystal form, but it is also an important material in a certain class of fuel cells.

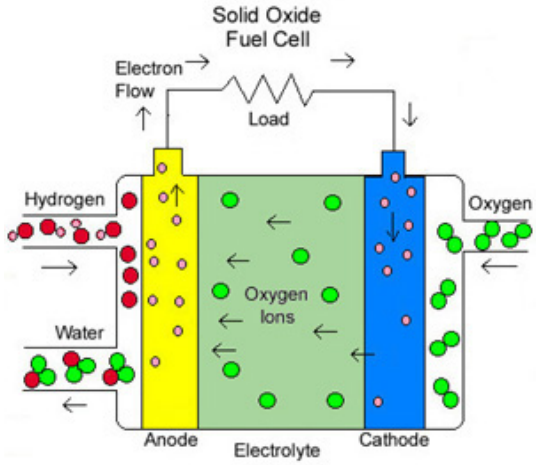


Figure 73.2: Solid oxide fuel cell utilizing a zirconia solid electrolyte (image from http://mypages.iit.edu/~smart/garrear/fuelcells.htm).

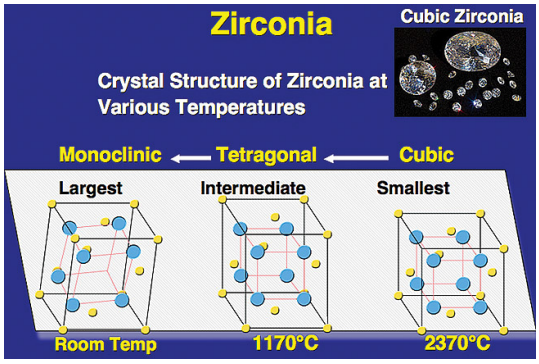


Figure 73.3: Crystal structures of zirconia (from http://www.dentalaegis.com/idt/2012/03/machinable-zirconia)

Figure 73.4: Cracking of a graze (in this case intentional) induced by thermal stresses. (from http://ceramicartsdaily.org/ceramic-art-and -artists/criticism-aesthetics/longquan-celadon-a-revival/).

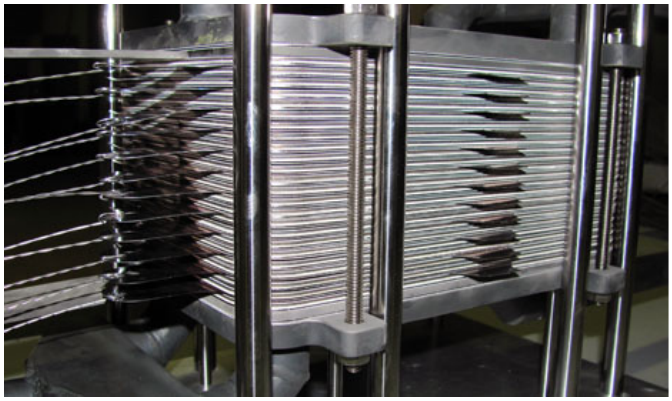


Figure 73.5: Solid oxide fuel cell stack (from http://www.nextechmaterials.com/energy/index.php?option=com_content&view=article&ic

Yttria stabilized Zirconia:

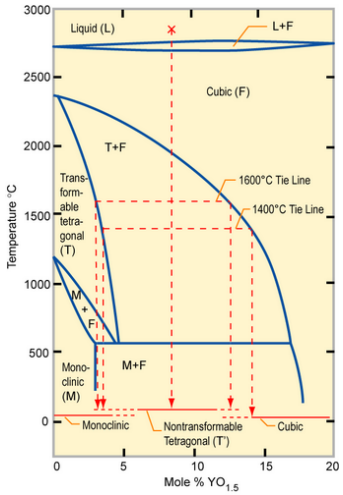


Figure 73.6: ZrO_2/Y_2O_3 Phase Diagram (from http://www.doitpoms.ac.uk/tlplib/fuel-cells/sofc_electrolyte.php?printable=1).

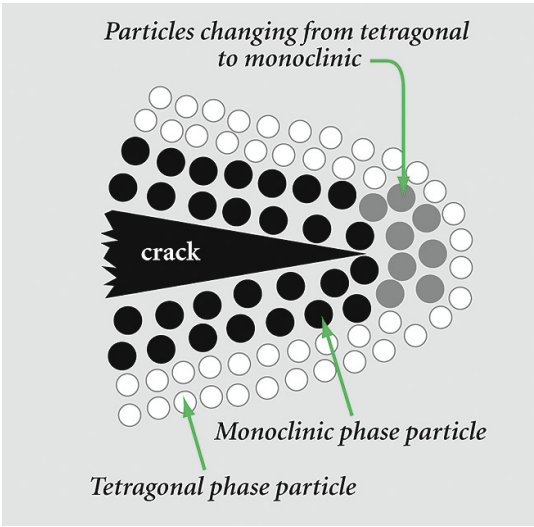


Figure 73.7: Illustration of the transformation toughening mechanism of Zirconia.

73.1.2 Fiber Composites

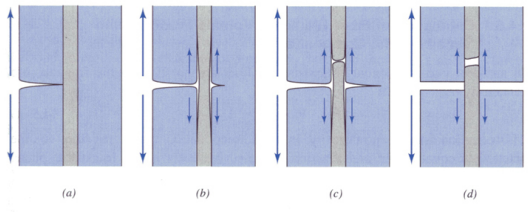


Figure 73.8: Schematic representation of crack propagation through a fiber-reinforced composite. (from http://www.aml.engineering.columbia.edu/ntm/level2/ch05/html/l2c05s02.html)

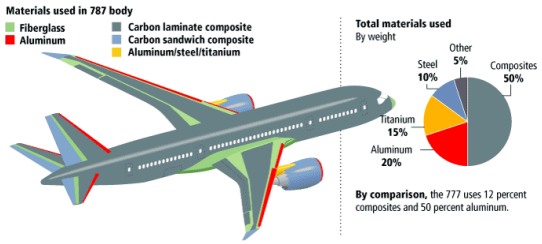


Figure 73.9: use of Composites in the Boeing 787 Dreamliner

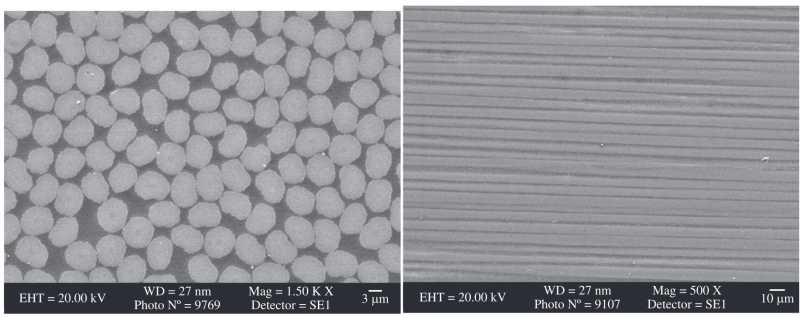


Figure 73.10: Views of a carbon fiber/epoxy composite perpendicular (left) and parallel (right) to the fiber direction.(from http://www.scielo.br/scielo.php?pid=S1516-14392010000300022&script=sci_arttext).

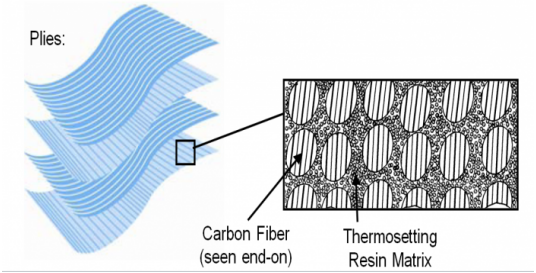


Figure 73.11: Laminate plies in a carbon fiber composite.

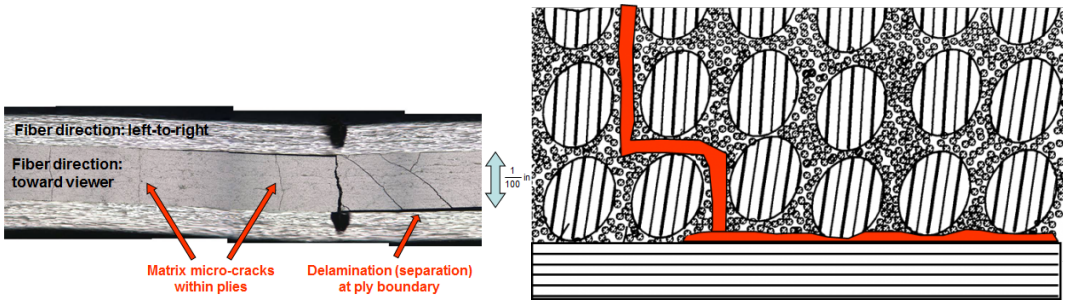


Figure 73.12: SEM image of failure of crack propagation in a carbon fiber composite (left), and a corresponding schematic representation of the crack path (right).

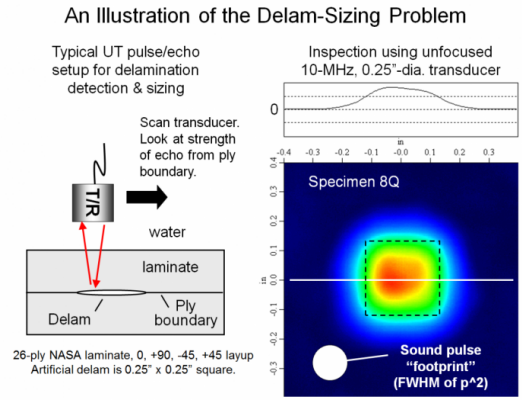


Figure 73.13: Nondestructive evaluation of damage in a carbon fiber composite.

Nondestructive Testing of a Fiber Laminate Composite:

73.1.3 Tempered Glass

For an entertaining and useful description oftempered glass, see the following video:
https://www.youtube.com/watch?v=xe-f4gokRBs

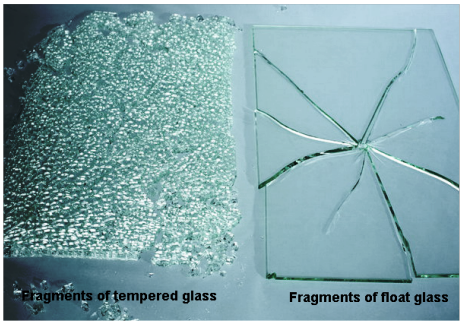


Figure 73.14: Comparison of tempered glass and float glass samples that have been fractured. (from http://www.bharatsafety.com/tempered_glass.html).

Figure 73.15: Laminated safety glass.

73.1.4 Hydraulic Fracturing (Fracking)

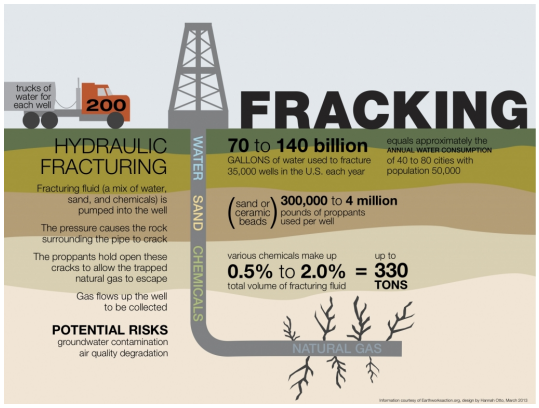


Figure 73.16: The hydraulic fracturing (fracking) process.

https://www.youtube.com/watch?v=VY34PQUiwOQ

74 Materials Selection: A Case Study

The concept of a space elevator is illustrated in Figure 74.1. The idea is that we run a cable directly from the earth out to a point in space. If the center of mass is at a geosynchronous orbit, the entire assembly orbits the earth at the same angular velocity at which the earth is rotating. To get into space, we no longer need to use a rocket. We can simply 'climb' up the cable at any velocity that we want. The concept is certainly appealing if we can get it to work. But could the concept ever actually work? That is determined by the availability (or lack thereof) of materials with sufficient strength for the cable. We'll need to start with some analysis to figure out what sort of properties are needed.

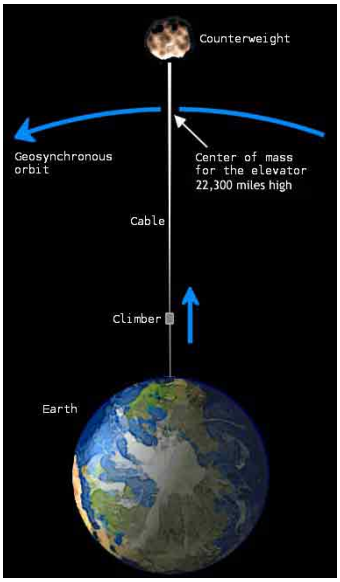


Figure 74.1: Schematic representation of the space elevator.

Consider a mass, m, that is located a distance r from the center of the earth, as illustrated in Figure 74.2. The net force on the object is a centripetal force acting outward (positive in our sign convention) and a gravitational force acting inward (negative in our sign convention):

$$F = m\omega^2 r - G_{gr} M_e m / r^2 + \tag{74.1}$$

where G_{gr} is the gravitational constant and M_e is the mass of the earth. The angular velocity of the earth is 2 π radians per day, or in more useful units:

$$\omega = \frac{2\pi}{(24\,\text{hr})\,(3600\,\text{s/hr})} = 7.3x10^{-5}\,\text{s}^{-1}$$

It is convenient to rewrite the first term in terms of g_0, the gravitational acceleration at $r = r_0$ (at the earth's surface):

$$g_0 = \frac{G_{gr} M_e}{r_0^2} \tag{74.2}$$

The net force on the mass at r can be written as follows:

$$F = m\omega^2 r - m g_0 \left(\frac{r_0}{r}\right)^2 \tag{74.3}$$

The net force is zero for $r = r_s$ (geosynchronous orbit):

$$r_s = \left(\frac{g_0 r_0^2}{\omega^2}\right)^{1/3} = 4.1x10^7\,\text{m}\,(22,000\,\text{miles}) \tag{74.4}$$

This is an easier number to remember than the angular velocity of the earth, so we use this expression for r_s to eliminate ω from Eq. 74.3, obtaining the following:

$$F = m g_0 r_0^2 \left(\frac{r}{r_s^3} - \frac{1}{r^2}\right) \tag{74.5}$$

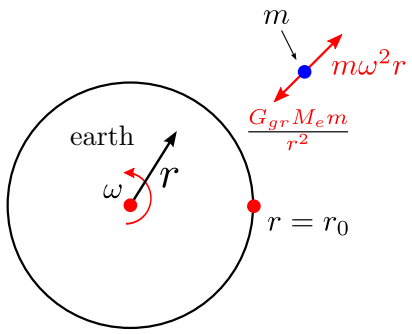


Figure 74.2: Radial forces acting on an orbiting mass.

Now consider a cable that extends from the earths surface $(r = r_0)$ to a distance r_ℓ from the earth's surface, as shown in Figure 74.3. We need the cable to be in tension everywhere so that it doesn't buckle. If we get the tension at the earth's surface ($r = r_0$) we're in good shape. We obtain the total force, F_0 at the earth's surface by integrating the net forces over the whole length of the cable:

$$F_0 = \rho A g_0 r_0^2 \int_{r_0}^{r_\ell} \left(\frac{r}{r_s^3} - \frac{1}{r^2} \right) dr \tag{74.6}$$

After integration we get:

$$F_0 = \rho A g_0 r_0^2 \left[\frac{1}{2 r_s^3} \left(r_\ell^2 - r_0^2 \right) + \left(\frac{1}{r_\ell} - \frac{1}{r_0} \right) \right] \tag{74.7}$$

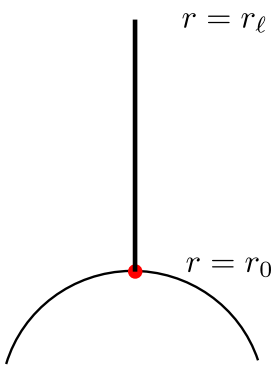


Figure 74.3: A cable extending from the surface of the earth to a distance r_ℓ away from the earth's center.

F_0 is plotted in Figure 74.4:

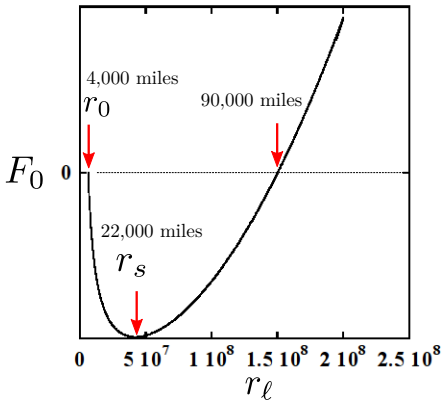


Figure 74.4: F_0 as given by Eq. 74.7.

The maximum tension is at $r = r_s$, and is obtained by integrating the force from r_s to r_ℓ:

$$F_{max} = \rho A g_0 r_0^2 \left[\left(\frac{1}{r_\ell} - \frac{1}{r_s} \right) + \frac{1}{2r_s^3} \left(r_\ell^2 - r_s^2 \right) \right] \tag{74.8}$$

We have the following numbers:

- $r_0 = 6.4x10^6$ m (4,000 miles)
- $r_s = 4.1x10^7$ m (22,000 miles)
- $r_\ell = 1.5x10^8$ m (90,000 miles)
- $g_0 = 9.8$ m/s ^{2}

From these numbers we get $\frac{F_{max}}{\rho A} = \frac{\sigma_{max}}{\rho} = 4.8x10^7 \frac{\text{N/m}^2}{\text{Kg/m}^3}$
Let's compare that to the best materials that are actually available. An Ashby plot of tensile strength (σ_f) and density is shown in Figure 74.5. A line with a slope of 1 on this double logarithmic plot corresponds to a range of materials with a constant value of σ_f/ρ. The line drawn on Figure 74.5 corresponds to σ_f/ρ is ≈2.8x10 ^{6} (in Si units), which is the most optimistic value possible for any known material - corresponding to the best attainable properties for diamond. Forgetting about any issues of cost, fracture toughness, etc., we could imagine that we see that we are a factor of 20 below the design requirement. So there's no way this is ever going to work, no matter how good your team of materials scientists is.

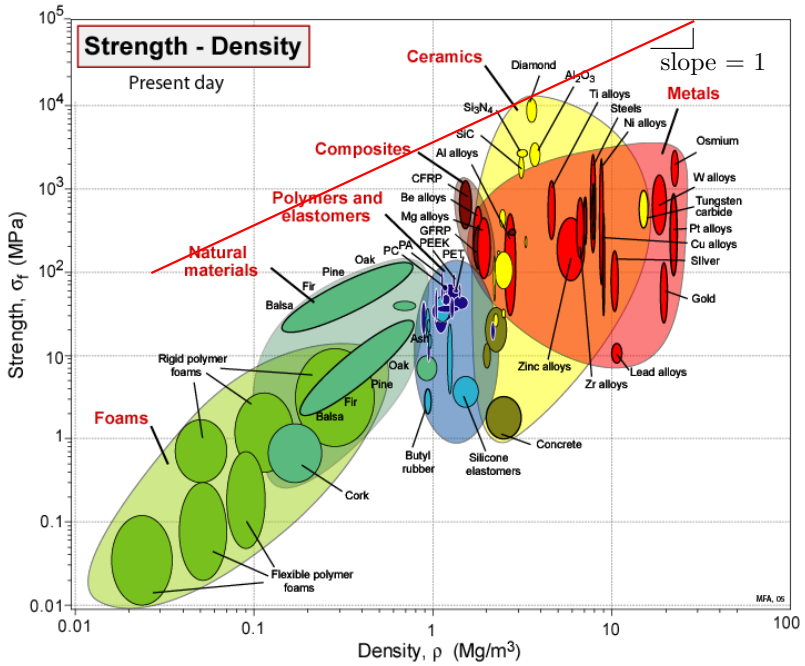


Figure 74.5: Ashby plot of tensile strength and density.

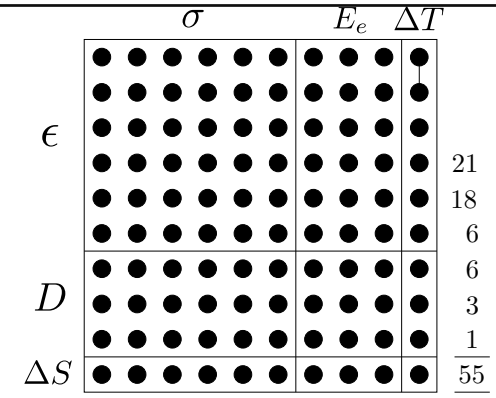
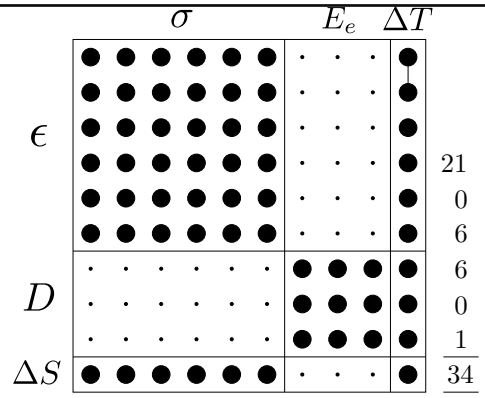
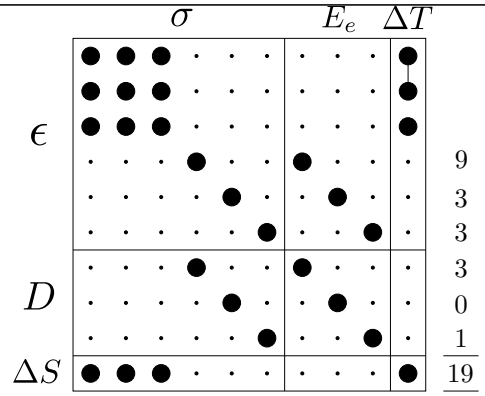
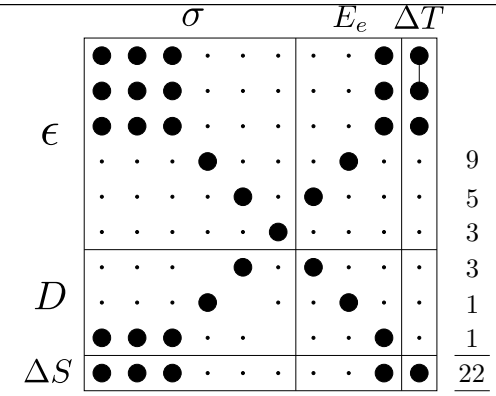
All is not lost yet, however, since we really haven't optimized the geometry. The design we considered above has a constant radius, which we would really not want to do. What if we optimize the geometry so that material has the largest cross section at $r = r_s$ (where the load is maximized). We'll consider a design where the actual cross section varies in a way that keeps the stress (tensile force divided by cross sectional area) constant. The analysis is a bit more complicated than we want to bother with here, but we get a simple expression for the maximum cross sectional area, A_s (at $r = r_s$), to the cross sectional area at the earth's surface (A_0, at $r = r_0$) :

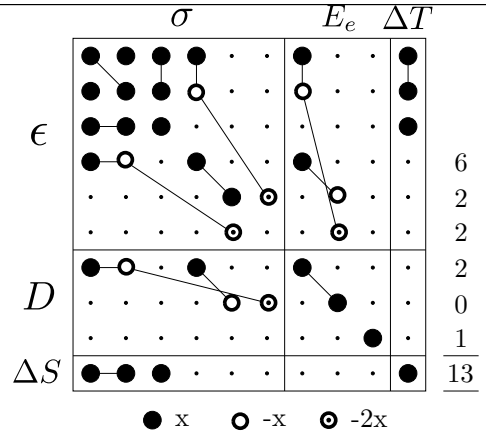
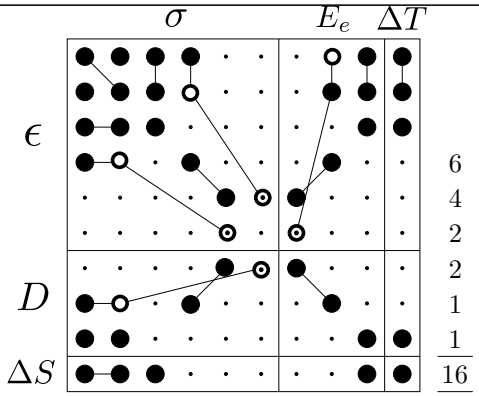
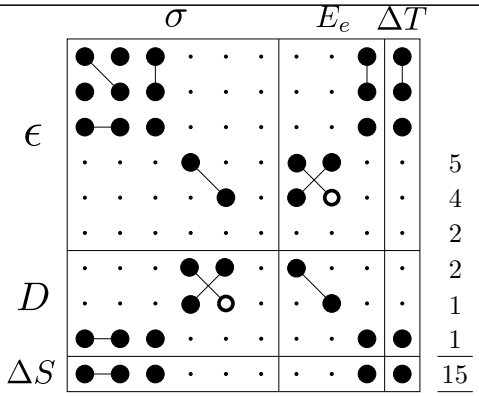
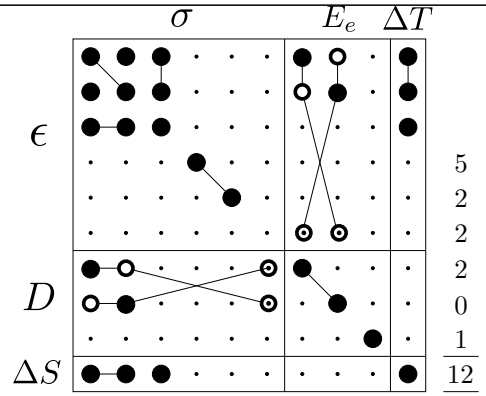
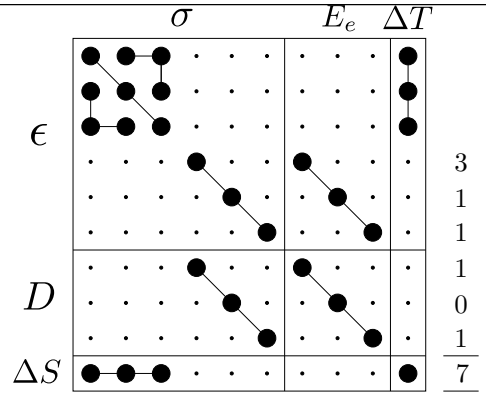
$$\frac{A_s}{A_0} = \exp\left(\frac{0.77 r_0 \rho g_0}{\sigma} \right) \tag{74.9}$$

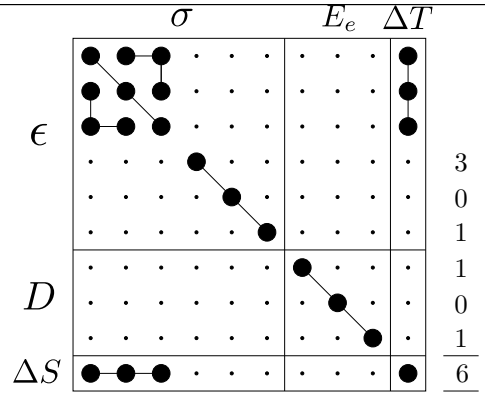
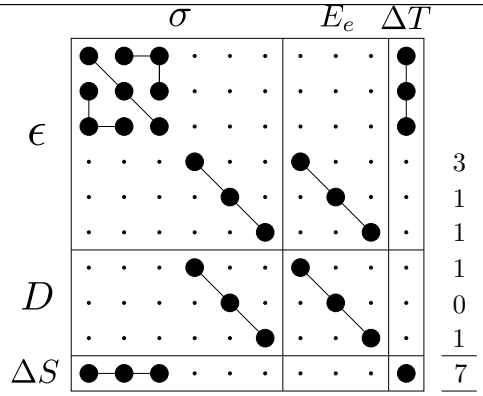
If we assume $\sigma/\rho = 2.8x10^6 \frac{\text{N/m}^2}{\text{Kg/m}^3}$, so that the system is operating at the value of σ/f corresponding to the solid line in Figure 74.5 gives $\frac{A_s}{A_0} = 110$. So in this case cable with a diameter of 1 cm at $r = r_0$ needs to have a

diameter of $\approx$ 10 cm at $r = r_s$. We don't have much leeway in decreasing σ/ρ, however. If the best we can do is $\sigma/\rho = 1.0x10^6 \frac{\mathrm{N/m^2}}{\mathrm{Kg/m^3}}$, we get $\frac{A_s}{A_0} = 10^{21}$, which is clearly not workable. So we are stuck with the requirement that the cable have a tensile strength corresponding to the best material on earth over a length of 90,000 miles, without a single critical defect that would cause the material to fail. Good luck with that.

75 Appendix 1: Symmetry Relationships for All 32 Point Groups

Point Group	Coefficient Map	Point Group	Coefficient Map
1		$\bar{1}$	
2		m	
2/m		222	
mm2		mmm	
4		$\bar{4}$	
4/m		422	
4mm		$\bar{4}$2m	
4/mmm		3	

Point Group	Coefficient Map	Point Group	Coefficient Map
$\bar{3}$		32	
3m		$\bar{3}$m	
6		$\bar{6}$	
6/m		622	
6mm		$\bar{6}$m2	
6/mmm		23	

Point Group	Coefficient Map	Point Group	Coefficient Map
m $\bar{3}$		432	
$\bar{4}$3m		m $\bar{3}$m	

76 Appendix 2: MATLAB Code Examples

(1) Rotation of a Tensor: on page 431

(2) Tensor Rotation (Symbolic Algebra): on page 433

(3) Mohr Circle Derivation (Symbolic Algebra): on page 434

(4) Principal Stress Calculation: on page 436

77 Appendix 3: Introduction to Abaqus

In this part, we will work together to go, step by step, through the Finite Element Analysis (FEA) method for a cantilever beam. You will use the following exercise as a guide for your individual portions of the laboratory exercise. The following is adapted from *SIMULIA*'s *Introduction to Abaqus* as well as Jee Rim's *Introduction to Finite Element Analysis.*

77.1 A First Look at Abaqus/CAE

The main window, shown in Figure 77.1, contains the workspace and all of the tools necessary for preprocessing, numerical analysis, and assessment of a FEA problem. Any general purpose FEA software involves those steps:

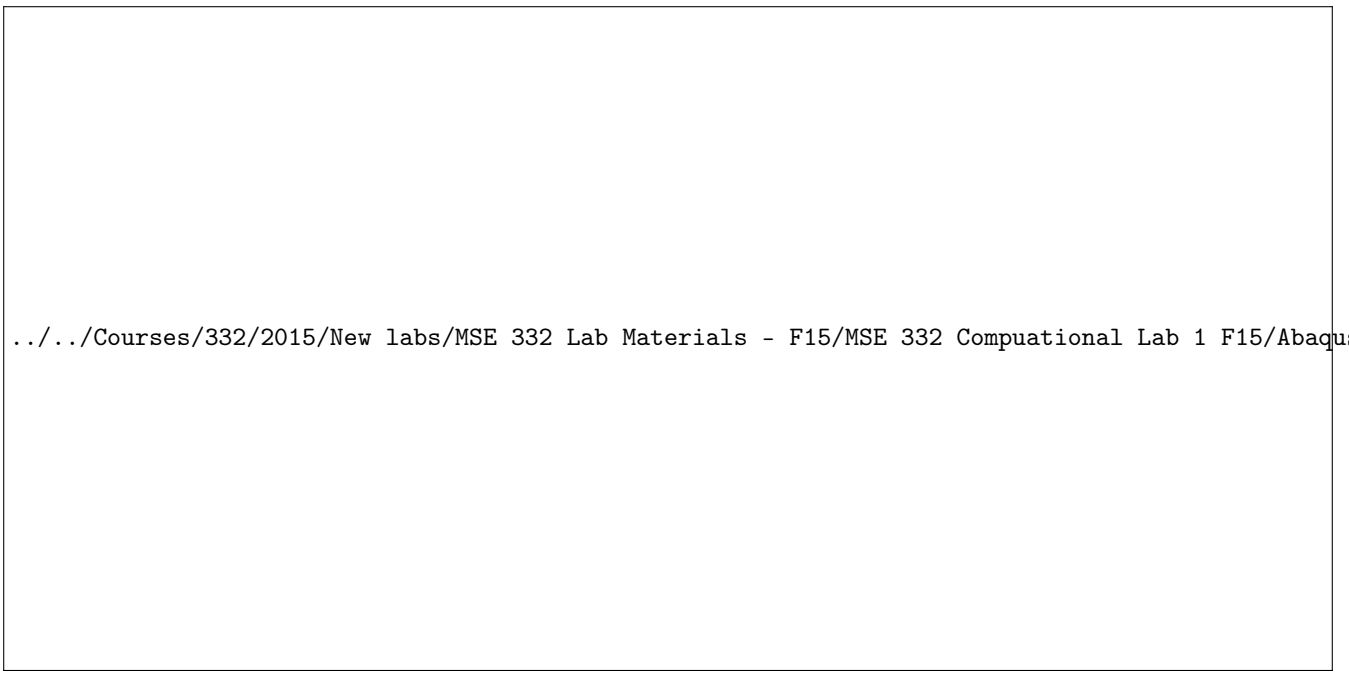


Figure 77.1: The starting workspace in *Abaqus/CAE*.

(1) *Preprocessing*: This involves setting up the FEA problem. You input the geometry, materials properties, loads and boundary conditions. Meshing (discretization) is typically done automatically by the computer, but the analyst must choose element formations that suit the mathematical model, and decide on mesh density.

(2) *Numerical Analysis*: The software will automatically generate the matrices (system of equations) that will define the behavior of each element and merges them into a global matrix equation. It then solves (or attempts to solve) these equations to determine the values of field quantities, like displacements and stresses, at the nodes.

(3) *Assessment*: Many FEA software packages now include visualization and analysis. This is the user-directed selection of how the field values are displayed. We will show examples of deformation visualization, contour plotting, and X-Y data extraction.

77.2 The Model Tree:

During preprocessing an analyst works with various modules in the Model Tree. These can be accessed through the model tree itself (using the mouse) or through the Module drop-down box in the context bar. The important"branches" (for this class) of the Model Tree are:

(1) *Parts*: Here, using CAD design methodology, you create, modify, and set up symmetries for your part.

(2) *Materials*: Here you define your materials properties. We will limit ourselves to the elastic behavior of materials in these labs, but it is possible to define many different mechanical properties of materials (including using experimental data input).

(3) *Sections*: Here you create sections that will be used to assign materials properties to your part.

(4) *Assembly*: Here is where you define your instances - the FEA object that the computation will be performed on. In problems with many parts (we will have only one), you use this utility to combine and assign relationships between parts.

(5) *Steps*: These are the points in which some action happens. In this class, we will always have a initial step followed by a step in which a load is applied statically.

(6) *Loads*: Here we define the forces that are applied to the instances in the model.

(7) *BCs*: Here we define the boundary conditions of the model.

(8) *Jobs*: Here we design the computational job. We use this branch to submit and monitor jobs, as well.

Remember! All units in Abaqus *must be consistent*. Please follow the table below. We will use SI (mm).

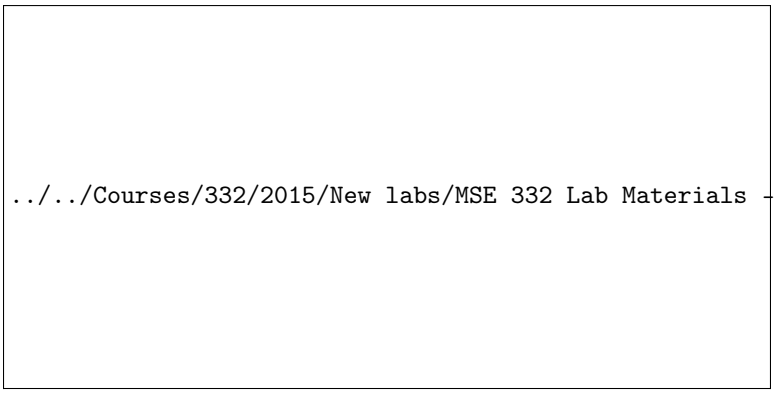


Figure 77.2: Consistent units in Abaqus/CAE.

Taking *Abaqus/CAE* Further:

Here, we'll outline some further useful abilities beyond *Abaqus/CAE*'s simple introduction that will help with your problems.

77.3 Partitioning the Part

In some of the labs, a part may consist of regions with different materials properties. For these projects, your model will still only require one part, and the best way to handle this is to *partition* your part and assign different sections to each partition. See, for example, Figure 77.3.

We should not have to start a new material from scratch - to modify an existing model, simply right-click on the model and select **Copy Model**. A new model tree will appear.

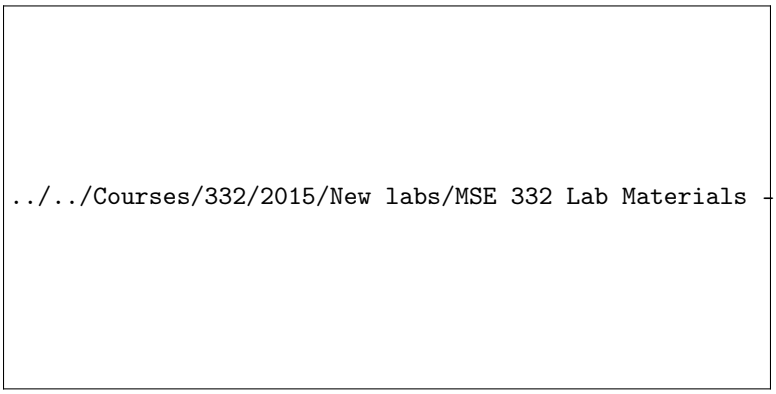


Figure 77.3: A cantilever beam of two materials

There are many ways to partition a part. One way is to use the **Create a Data Point** toolset under the *Property* module. Note - holding down the left-click button will allow you to access other ways to create partitions. Then, after a datum point is defined, partition the cell using **Partition Cell:Define Cutting Plane**. Try using the point and the edge normal that defines the plane slicing through the beam.

Now that there exists a partition, we can assign two different materials properties to each. We can create a new material"Aluminum" with a Young's modulus of 69×10^3 MPa and Poisson's ratio of 0.35. Then you can assign the two sections you created with the materials properties you've input.

77.4 Meshing with Uneven Density

For the most efficiency, the mesh should be finest where the stress gradients are high, such as near discontinuities or other areas of stress concentration. In order to obtain a mesh that is unevenly spaced, you can seed the mesh along each edge separately. In the *Mesh* module, select **Seed** $\rightarrow$ **Edges**. When you have picked the edges you want to seed (shift-click) and clicked **done** in the prompt area, you will be asked to determine the parameters for the creation of Local Seeds. Selected the **By Size** Method, and choose **Single** for Bias. Enter**5** and**15** for Minimum and Maximum Size, respectively. The red arrow that appears will point towards the direction where the seeds become spaced closely. Note that you can flip this vector. Make sure that all the edges are seeded. You should get something that looks like Figure 77.4.

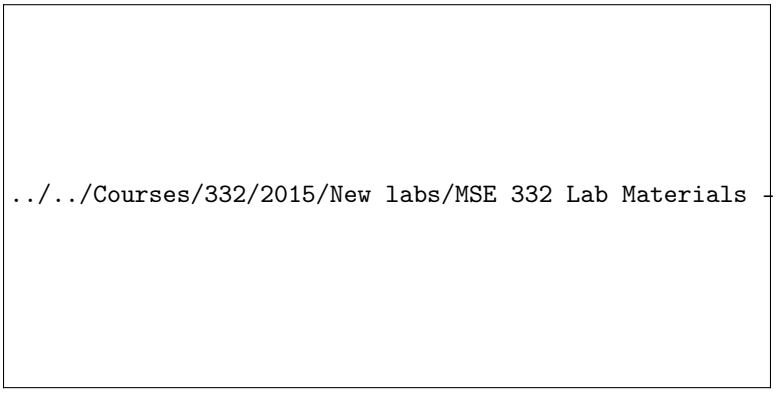


Figure 77.4: Seeding a Mesh with Bias

Note! The student version is only capable of performing jobs on instances with 1,000 or fewer nodes. so, while denser nodes may give better results, typically, you have a very hard limit.

77.5 Data Visualization

There are many ways to visualize your field output when you have successfully completed a job. This is done in the **Visualization** module. First, select the **Common Options** from the toolset. Play around with the options in the Render Style and Visible Edges to make the result depiction to your liking. Note, Abaqus auto-computes a deformation scale factor in order to improve visualizations of small deformations. You may or may not want to do the same.

Another useful display option is the **Superimpose Options**. This makes it easier to visualize the original and deformed shape. An important option here is under Other, which has a tab that allows you to tune your transparency.

To view the deformed and undeformed shapes, toggle the meshed cartoons in the toolset. Contours are valuable visualization tools, so select the colored deformed shape to see color-coded field outputs. Hold the color-coded field output down to select an deformed/undeformed overlay. The default is banded contours, but other options are available, depending on your preference. My preference is shown in Figure 77.5.

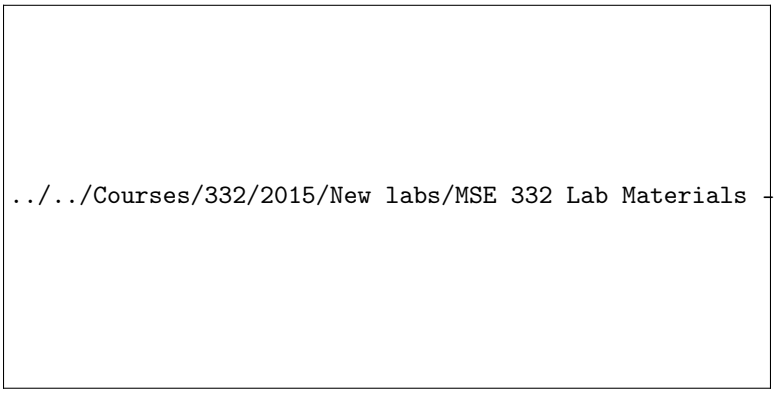


Figure 77.5: Visualization of a deformed beam (Abaqus transparency setting apparently doesn't write well to PDF...)

Now that you understand how to visualize the field outputs, use the Primary field dropdown selections to change the field values. Selection of **U** allow for visualization of Cartesian nodal displacements of the U Magnitude, U1, U2 and U3. Note, Abaqus/CAE is defaulted to scale these values. Look at the legend to assess the magnitude of the strains and strains, and pay attention to the Deformation Scale Factor value in the workspace.

The default options for the visualization are organized under the field values of **S**. These include simple and shear stresses (S_{11}, S_{22}, S_{33}, S_{11}, S_{13}, S_{23}), Principle Stresses, as well as Mises and Tresca stresses.

An X-Y data object is a collection of ordered pairs that Abaqus/CAE stores in two columns - an X-column and a Y-column. The X-Y data can come from the output database, or be entered from a keyboard. Abaqus/CAE can display X-Y data in the form of an X-Y plot, a two-dimensional graph of one variable vs. another.

Let's say you want to create a plot of the displacement of the beam along the length of the beam. The X-Y data will be specified from the Field **Output** of the database, along a path that you specify. First, to create the path. In the Results tree, double-click on Path to bring up a"create path" dialog box. Name your path, and select **Node List** as the Type. In the next dialog box, select **Viewport Selections** to be **Add After**, and then choose the path along a long edge of the deformed cantilever (shift-select). once that path is defined, double-click XYdata in the Results tree and select Path as the source. Choose the path you just defined. From the Frame option choose **Step/Frame** for the frame you want to read the data from (e.g., the deformed beam). Then choose the Field Output to be the **Spatial Displacement Node U2**. Once you complete this, click on Plot and your data will be displayed similar to Figure 77.6. You can also write out the data by clicking **Report** → **XY**. You can overlap plots by selecting multiple datasets and clicking **Add to Plot**.

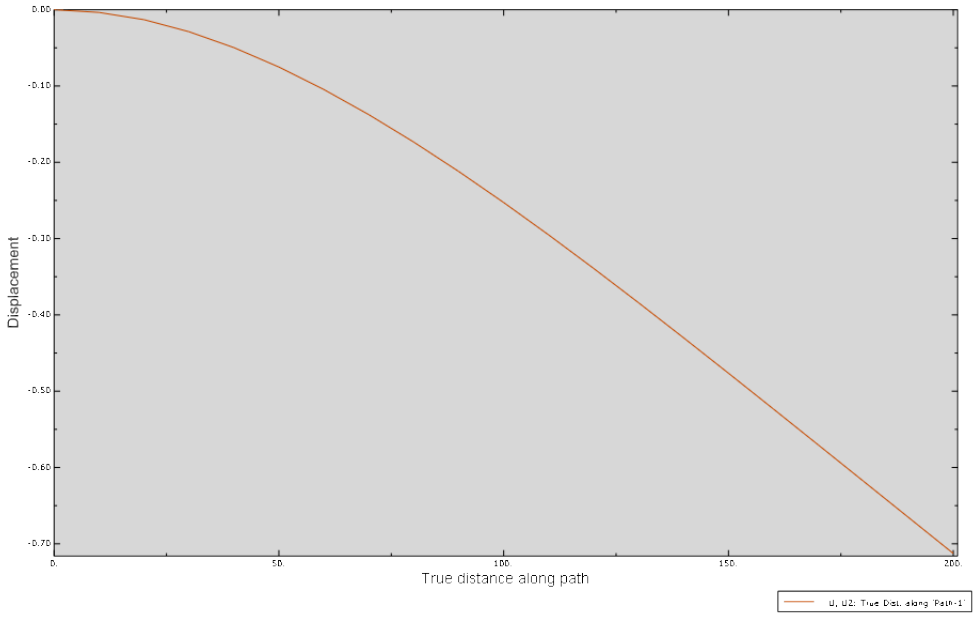


Figure 77.6: Displacement output for the cantilever beam problem.

78 332 Problems

78.1 Course Organization

If you are interested in the 3-d printing option of the laboratory, send an email to the TA. Identify any specific people you would like to work with and have them send an email too.

78.2 The Stress Tensor

Consider the following stress tensor:

$$\sigma_{ij} = \begin{bmatrix} 4 & 3 & 0 \\ 3 & 1 & 2 \\ 0 & 2 & 6 \end{bmatrix} x10^6 \text{ Pa}$$

(1) Calculate the stress tensor for coordinate axes rotated by 30° about the z axis (the 3 axis).

(2) Repeat the calculation for a 30° rotation around the x axis (the 1 axis).

(3) Calculate the three principal stresses.

(4) Calculate the maximum shear stress in the sample.

Consider the following stress tensor:

$$\sigma_{ij} = \begin{bmatrix} -2 & 1.4 & 0 \\ 1.4 & 6 & 0 \\ 0 & 0 & 2 \end{bmatrix} x10^6 \text{ Pa}$$

(1) Draw a Mohr circle representation of the stress contributions in the xy plane

(2) What are the three principal stresses?

78.3 Strains

An engineering shear strain of 1 (100%) is applied to a rubber cube with dimensions of 1cm×1cm×1cm. Young's modulus for the rubber sample is 10^6 Pa, and you can assume it is incompressible.

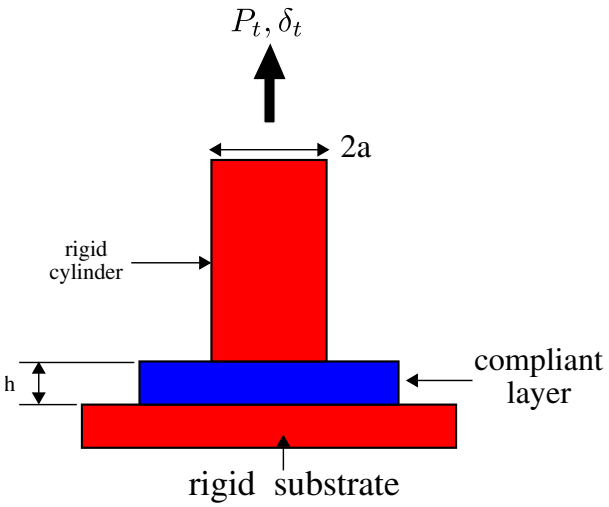
(3) Sketch the shape of the object after the strain is applied, indicating the dimensions quantitatively.

(4) On your sketch, indicate the magnitude and directions of the forces that are applied to the object in order to reach the desired strain.

(5) Calculate the 3 principal extension ratios characterizing the final strain state?

78.4 Contact Mechanics

(6) Consider the contact of a flat rigid punch with a thin elastic layer, as shown schematically below:



Suppose the compliant layer is incompressible gel ($\nu = 0.5$), with a Young's modulus,E, of 10^4 Pa. The critical energy release rate for failure at the gel/punch interface is 0.1 J/m^2. The punch radius,a, is 3 mm.

a What is the tensile force required to separate the punch from the layer if the layer is infinitely thick?

b What is the stress intensity factor,K_I, just prior to detachment of the punch from the surface?

c How close to the punch edge to you need to be for the tensile stress at the punch/layer interface to be equal to the modulus of the layer?

d Describe in qualitative terms what happens to the following quantities as the thickness,h, of the compliant layer decreases:

i. The overall compliance of the system.

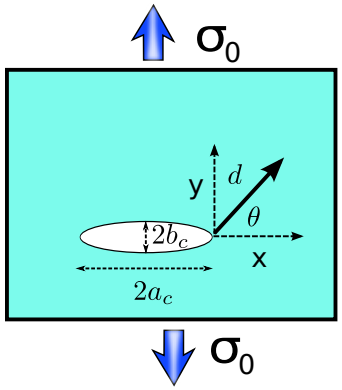
ii. The load required to detach the indenter from the substrate.

iii. The displacement at which the indenter detaches from the substrate.

iv. The shape of the tensile stress distribution at the punch/substrate interface.

78.5 Fracture Mechanics

(7) The stress fields in the vicinity of a crack tip in a material are determined by the distance,d, from the crack, and the polar angleθ, defined in the following diagram.



a For a fixed value of$d$, plot the behavior ofσ_{xx},σ_{yy} andσ_{xy} for a mode I crack as a function ofθ.

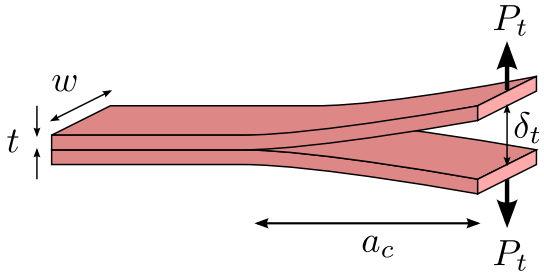
b What happens to the stresses for$\theta = 180°$? Why does this make sense?

c A mode I crack will travel in the direction for which the normal stress acting across the crack surfaces is maximized. What direction is this?

(8) Look up the fracture toughness (K_{IC}) and Young's modulus(E) for window glass. Assuming that the maximum local stress is$\approx E/10$, estimate the crack tip radius of curvature for a crack propagating through window glass.

(9) As a crack advances, what happens to the stiffness of the cracked body? What happens to the compliance?

(10) A set of double cantilever beam adhesion test specimens was fabricated from a set of aluminum alloy samples. The geometry as as shown below:



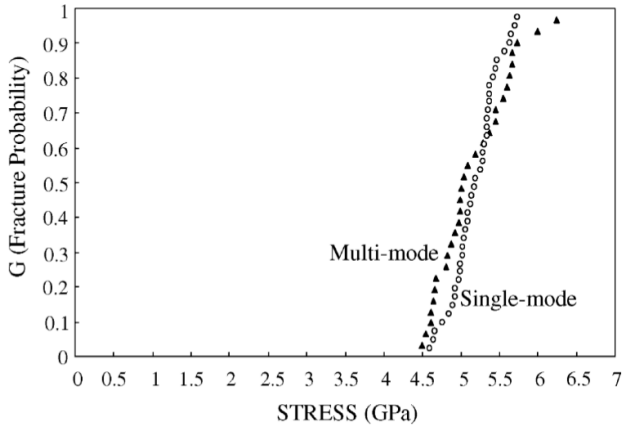
Suppose each of the two beams has a thickness (t) of 10 mm, a width(w) of 20 mm and a length of 80 mm. The double cantilever beam sample was produced by using an adhesive to glue the two beams together. Assume the precrack with a length,a, of 10 mm. The critical energy release rate for the adhesive is 65J/m^2.

d Calculate the values of the tensile load,P, and the displacement,Δ, where the crack starts to propagate.

e In a load-controlled experiment,P is held constant once the crack starts to propagate, and in a displacement controlled testΔ is held constant once the crack starts to propagate. From the relationship between$\mathcal{G}$ and$P$,Δ and$a$, describe why the load controlled experiment results in unstable crack growth, but the displacement controlled experiment results in stable crack growth.

f From your answer to part b, describe how you would design an experiment where you measured the energy release rate required to propagate the crack at a specified velocity.

(11) The following data was obtained for the fracture stress of a series of silica glass fibers used for optical communications:



The graph shows the distribution of failure probabilities as a function of the applied tensile stress. None of the samples had fractured at a stress of 4.5 GPa, but they had all fractured at a stress of 6 GPa. From these data, and from fracture toughnesses given for inorganic glass in class (and in the course notes), estimate the intrinsic flaw sizes that are present at the surface of the glass fibers. Comment on these sizes, and if you think the fracture mechanics analysis makes sense to use in this case.

(12) Silicones containing resin fillers are used as an encapsulant materials in light emitting diodes (LEDs) in order to protect the electronics from harsh environments and to aid in heat dissipation. Near the surface of the electronic components, temperatures can go as high as200 °C for extended time periods.

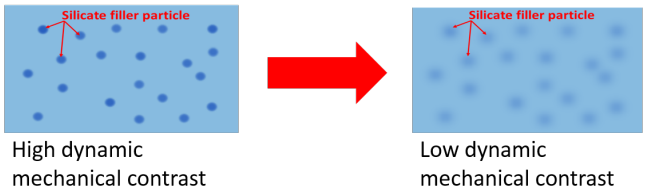


Figure 78.1: High dynamic mechanical contrast is important

a Given that a high dynamic mechanical contrast is desirable in creating a soft material with high fracture toughness, what would you suggest as a design criteria in order to maintain high dynamic mechanical contrast at high temperatures? (Hint: think about the role of the Tg of the matrix and filler content.)

b Thermal mismatch at the interface between the encapsulant and electronic can lead to residual stresses that promote crack propagation. In assessing the performance of the encapsulant at the interface, should a failure stress or a failure strain criteria be used? Why?

c From a thermal management and mechanics perspective, why do you think a soft encapsulant (e.g. silicone) will be more preferable over a hard encapsulant (e.g. glass)?

78.6 Matrix Representation of Stress and Strains

For an isotropic system there are only two independent elastic constants,s_{12} ands_{11}. This is because if properties are anisotropic in the 1-2 plane, the compliance coefficient describing shear in this plane,s_{44}, is equal to2 $(s_{11} - s_{12})$. We can use the Mohr's circle construction to figure out why this equality must exist.

(1) Start with the following pure shear stress and strain states:

$$\sigma = \begin{bmatrix} 0 & \sigma_{12} & 0 \\ \sigma_{12} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \; e = \begin{bmatrix} 0 & e_{12} & 0 \\ e_{12} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Use the matrix formulation to obtain a relationship betweenσ_{12} ande_{12} in this coordinate system.

(2) Rotate the coordinate system by 45° so that the stress state looks like this:

$$\sigma = \begin{bmatrix} \sigma_1^p & 0 & 0 \\ 0 & \sigma_2^p & 0 \\ 0 & 0 & 0 \end{bmatrix} ; \; e = \begin{bmatrix} e_1^p & 0 & 0 \\ 0 & e_2^p & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Use the Mohr's circle construction to write these principal stresses and strains in terms ofσ_{12} ande_{12}. Then use the matrix formulation to obtain an expression betweenσ_{12} ande_{12} in this rotated coordinate system.

(3) For an isotropic system, the relationship betweenσ_{12} ande_{12} should not depend on the orientation of the coordinate axes. Show that the only to reconcile the results from parts a and b is for$s_{44}$to be equal to2 $(s_{11} - s_{12})$.

78.7 Typical Moduli

(4) Calculate the sound velocities for shear and longitudinal waves traveling through the materials listed in the table in the course notes.

78.8 Yield Criteria

(5) A cube of material is loaded triaxially resulting in the following principal stresses at the point of plastic yielding:σ_1^p =140 MPa,σ_2^p =20 MPa, andσ_3^p =35 MPa.

a What is the shear strength of the material according to the Tresca yield criterion?

b If the the value ofσ_3^p were increased to 70 MPa, how does this change your result? Explain.

(6) From the work of D. C. Jillson, Trans. AIME 188, 1129 (1950), the following data were taken relating to the deformation of zinc single crystals:

Table 78.1: Applied tensile force for slip of single crystal Zn.

ϕ	λ	P (N)
83.5	18	203.1
70.5	29	77.1
60	30.5	51.7
50	40	45.1
29	62.5	54.9
13	78	109.0
4	86	318.5

In this tableϕ is the angle between the loading axis and the normal to the slip plane,λ is the angle between the loading axis and the slip direction and$P$ is the force acting on the crystal when yielding begins. The crystals have a cross-sectional area,A_0, of 122x10^{-6}m^2.

a What is the slip system for this material.

b For each combination of$P$,ϕ andλ, calculate the resolved shear stress,τ_{RSS} and normal stress,σ_N acting on the slip plane when yielding begins.

c From your calculations, doesτ_{RSS} orσ_N control yielding?

d Plot the Schmid factor versus the normal stress,P/A_0, acting on the rod. At what Schmid factor value are these experimentally-measured yield loads at a minimum? Does this make sense?

(7) What is the effect of the resolved normal stress on the yield behavior of crystalline metals and ceramics? What about polymers? Describe any differences between the two cases.

(8) A sample of pure iron has a uniaxial tensile yield strength of 100 MPa. Assume that the yield behavior is described by the Von Mises yield criterion.

a What do you expect for the yield strength of the material in a state of uniaxial compression?

b What will the yield strength be under a stress state of shear hydrostatic pressure?

c What is the shear yield strength of the material.

78.9 Strengthening Mechanisms

(9) Explain which has a larger effect on solid solution strengthening — symmetrical or asymmetrical point defects — and identify which specific defects lead to symmetrical or asymmetrical stress fields. List at least one example of an engineering material in which this factor comes into play.

(10) The lower yield point for a certain plain carbon steel bar is found to be 135 MPa, while a second bar of the same composition yields at 260 MPa. Metallographic analysis shows that the average grain diameter is 50 μm in the first bar and 8 μm in the second bar.

a Predict the grain diameter needed to cause a lower yield point of 205 MPa.

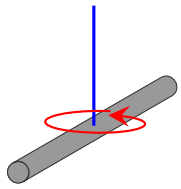
b If the steel could be fabricated to form a stable grain structure of 500 nm grains, what strength would be predicted?

c Why might you expect the upper yield point to be more alike in the first two bars than the lower yield point?

(11) The lattice parameters of Ni and Ni_3Al are 3.52×10^{-10} m and 3.567×10^{-10} m, respectively. The addition of 50 at% Cr to a Ni-Ni_3Al superalloy increases the lattice parameter of the Ni matrix to 3.525 × 10-10 m. Calculate the fractional change in alloy strength associated with the Cr addition, all other things being equal.

78.10 Viscoelasticity

(12) Consider a cylindrical metal bar with a density of 7.6 g/cm^3, a diameter of 1 cm and a length of 10 cm. It is suspended from a polymer fiber with a length, ℓ, 30 cm and a diameter of 1 mm.



a Suppose the fiber is perfectly elastic, with a shear modulus 10^9 Pa. Calculate the natural frequency of the system if the bar is rotating back and forth, causing the fiber to twist.

b Suppose the fiber is viscoelastic, with G' at the frequency calculated from part a equal to 10^9 Pa, and with $G'' = 10^7$ Pa. How many periods of the oscillation will take place before the magnitude of the oscillation decays by a factor of e (2.72)? Note: the rotational moment of inertia for the suspended metal bar in this geometry is $m\ell^2/12$, where m is the total mass of the bar and ℓ is its length.

(13) As mentioned in class, the Maxwell model, with a viscous element and an elastic element in series with one another, is the simplest possible model for a material that transitions from solid-like behavior at short times, to liquid-like behavior at long times. For a shear geometry we refer to the elastic element as G_0 and the viscous element as η.

a For what angular frequency are the storage and loss moduli equal to one another for a material that conforms to the Maxwell model? Express you answer in terms of the relaxation time, τ.

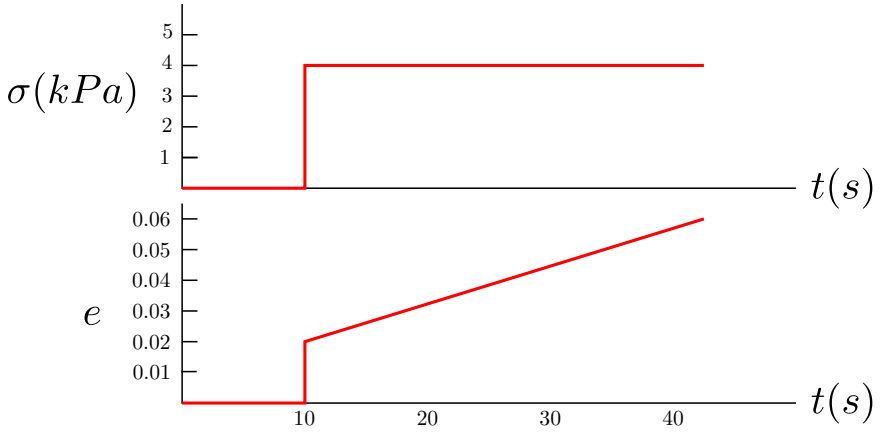
b Suppose the material is oscillated at a frequency that is ten times the frequency you calculated from part a. Does the material act more like a liquid or a solid at this frequency? Describe your reasoning.

c Calculate the values of G' and G'' at the frequency from part b, and calculate the phase angle, ϕ describing the phase lag between stress and strain in an oscillatory experiment. Not that the following expression relates ϕ, G' and G'':

$$\tan\phi = \frac{G''}{G'}$$

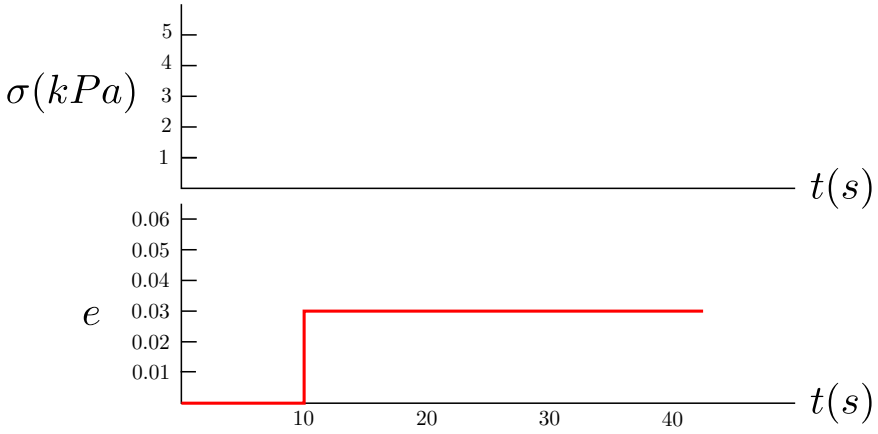
Does this phase angle make sense, given your answer to part b? Compare your value of ϕ to the values you expect for a perfectly elastic solid and a perfect liquid.

(14) The following stress and strain response are observed for a material during the initial stages of a creep experiment.



a Draw a spring/dashpot model that describes this behavior. Label moduli and viscosities as quantitatively as possible.

b A stress relaxation test (strain shown below) is performed on the same material. On the stress axis below, draw the stress response that you expect for the model you have drawn from part a.



(15) The following stress tensor describes the state of stress of a material at its yield point:

$$\sigma = \begin{bmatrix} 0 & 3 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & -5 \end{bmatrix} \text{MPa}$$

Suppose the same material is subjected to stress state of simple shear. At what value of the applied shear stress do you expect yielding to occur, assuming that the material obeys a Tresca yield criterion.

(16) Consider a material with elastic constants given by the following compliance matrix:

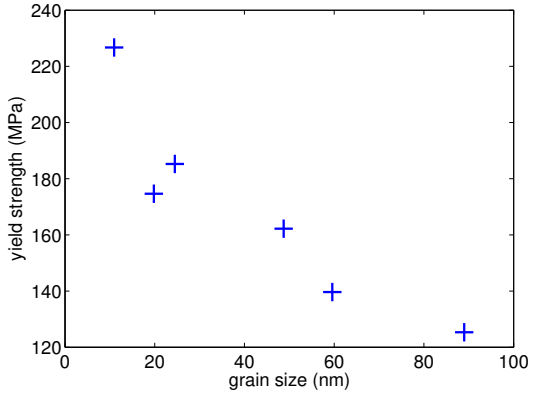
$$s_{ij} = \begin{bmatrix} 14.5 & -4.78 & -0.019 & 0 & 0 & 0 \\ -4.78 & 11.7 & -0.062 & 0 & 0 & 0 \\ -0.019 & -0.062 & 0.317 & 0 & 0 & 0 \\ 0 & 0 & 0 & 31.4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 61.7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 27.6 \end{bmatrix} \text{GPa}^{-1}$$

a Describe the symmetry of this material, and explain your reasoning.

b What is the highest value for Young's modulus that you would expect for this material? What direction does it correspond to?

c Calculate the value of Poisson's ratio obtained from an experiment where the materials is pulled in the 3 direction and change in sample width in the 2 direction is measured.

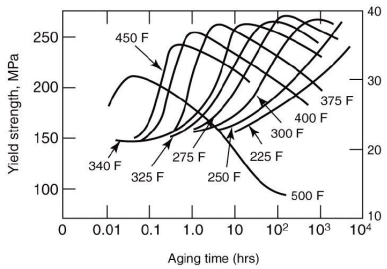
(17) The following plot shows values of the yield strength of copper samples as a function of the grain size of these samples.



a Describe why the yield stress decreases with increasing grain size.

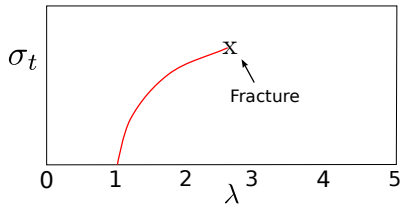
b Describe the procedure you would use to determine the limiting value of the yield strength in the absence of grain boundaries.

(18) The figure below shows the yield strength of a precipitation hardened aluminum alloy as a function of aging time at different temperatures. Note that the yield strength initially goes through maximum and then decreases with time. Explain why this happens in as much detail as possible.

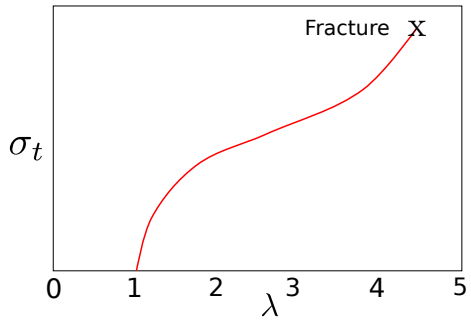


(19) Consider the following two stress-strain curves obtained from a glassy polymer material. In these plots σ_t is the true stress and λ is the extension ratio (1+e, where e is the tensile strain).

(a)

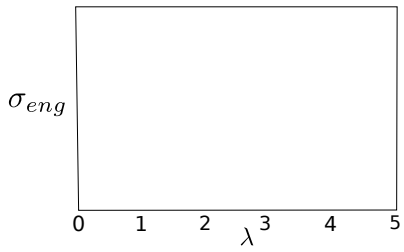


(b)

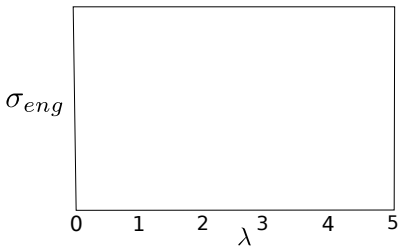


On the axes below, sketch the behavior of the engineering stress (measured tensile force vs. undeformed cross sectional area) vs extension ratio that you expect for each of these samples. Be as quantitative as possible with regard to the x axis.

(a)



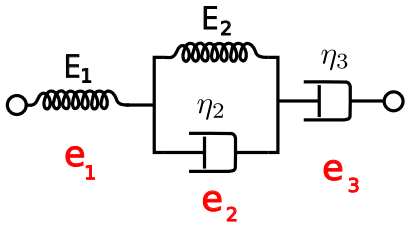
(b)



Briefly describe why you drew the curves the way you did.

78.11 Nonlinear Viscoelasticity and Creep

(20) A step stress (0 for t<0,σ for t>0) is applied to a solid which can be modeled by the following combination of linear springs and dashpots:



a This model is useful because it includes a non-recoverable creep component, a recoverable time dependent creep component, and an instantaneous, recoverable strain. Identify the element (or combination of elements) in the above model which is associated with each of these three contributions to the strain.

b This model has been applied to creep data for oriented polyethylene at room temperature. Model predictions were compared to data obtained from samples of high molecular weight (High M) and low molecular weight (Low M). Both samples were drawn to the same draw ratio. The following values of E_1,E_2,η_2 and η_3 were obtained from experimental data:

Sample	σ (GPa)	E_1 (GPa)	E_2 (GPa)	η_3 (GPa-s)	η_2 (GPa-s)
Low M	0.025	17.4	33.5	$1.8x10^5$	4300
Low M	0.05	13.6	35.6	$6.3x10^4$	5000
Low M	0.1	17.7	26.4	$3.1x10^4$	2200
Low M	0.15	17.7	26.5	$2.6x10^4$	2300
Low M	0.2	16.4	26.8	$1.2x10^4$	2000
High M	0.1	18.3	31.9	$3.1x10^6$	$8.7x10^4$
High M	0.15	16.6	21.3	$1.7x10^6$	$7.3x10^4$
High M	0.2	15.8	32.7	$7.7x10^5$	$3x10^4$
High M	0.3	25.4	39.1	$4.8x10^4$	2800
High M	0.4	25.0	43	$3x10^4$	3000
High M	0.5	21.7	46	$2.5x10^4$	5000

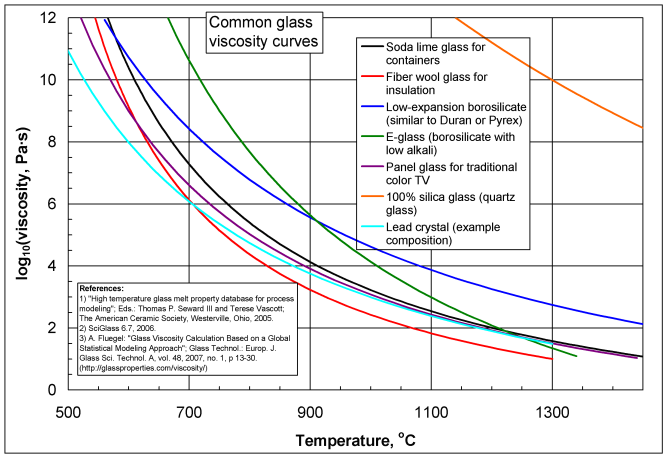
From the values ofη_3 given in this table, determine the stress dependence of the steady state creep rate. From this stress dependence, calculate the activation volume for non-recoverable creep in the high and low molecular weight samples, and compare these values to one another.

(21) In class we briefly talked about the possibility that you could see flow of a glass window over very long periods of time. Can creep of a glass by viscous flow give measurable changes in sample dimensions over a very long period of time? To sort this out, do the following:

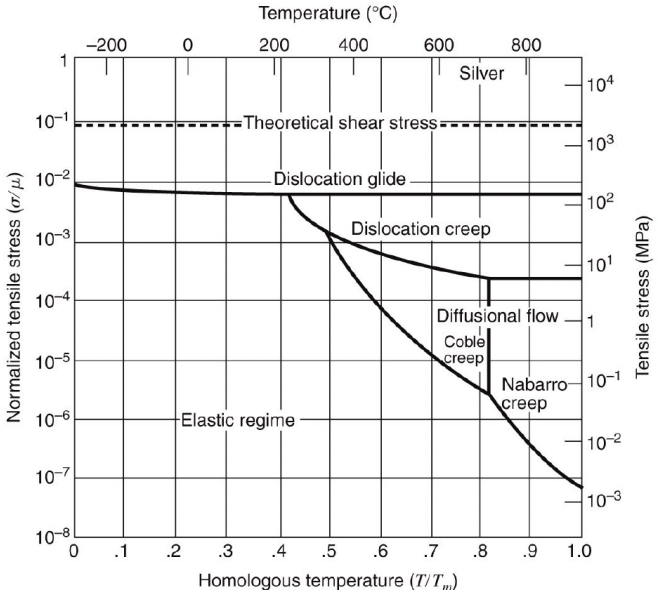
a Estimate the stress at the bottom of a very tall pane of window glass, due to the weight of the window itself. Look up the density of silica glass, and a height of the window that makes sense (choose a big one).

b Estimate the viscosity that would be needed to give a measurable change in sample dimensions after 400 years.

c Using the data below, does it make sense to you that observable flow could noticeably change the dimensions of the window? (You'll need to make some assumptions about how the viscosity will extrapolate to room temperature.

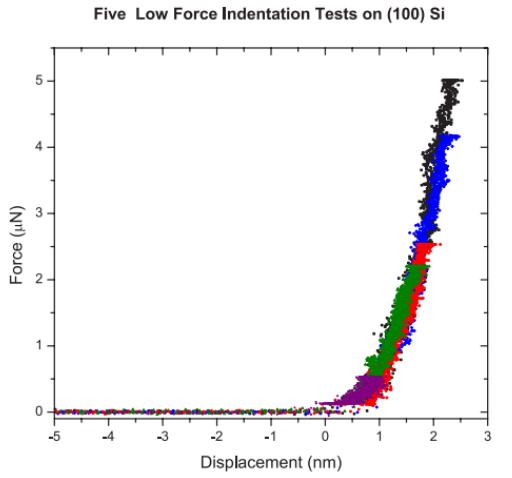


(22) Creep in metals at low stresses occurs by a vacancy diffusion mechanism, which means that the activation volume for these creep mechanisms corresponds to the atomic volume. Show using the data below for silver that we can safely replacesinh $(\sigma v/2k_BT)$ with$\sigma v/2k_BT$, so that the creep rates are linear in stress at all relevant temperatures and stresses where the dominant creep mechanisms involve vacancy diffusion. (You'll need to luck up data you can use to calculate the atomic volume of silver).



78.12 Nanoindentation

(23) Commercial nanoindenters are generally not suitable for the characterization of soft materials. To understand why this is the case, consider the following indentation data from the Hysitron web site (this is for the same instrument that Northwestern has in the NUANCE facility):



A series of five low force indents on (100) Silicon showing the industry leading noise-floor performance and data repeatability of the **TI 950** nanomechanical testing system.

a If the data in this figure are obtained with a spherical indenter of radius R, use the data from this figure to estimate the value of R. Assume that the material is being indented elastically and that adhesion can be neglected. (You'll need to look up mechanical property data for silicon).

b Suppose that the material is replaced by an elastomer with a modulus of 10^6 Pa. What value of R would need to be used to obtain the same Force displacement curve for this much softer material? (Assume that the effects of adhesion can eliminated by performing the indentation in a suitable liquid.

(24) Suppose an elastomeric sphere with a radius of 1 mm and a reduced modulus, E^*, of 10^6 Pa is placed on a flat, rigid substrate. Suppose also that the adhesion between the sphere and the substrate is characterized by a critical energy release rate of 0.05 J/m^2, independent of the crack velocity or direction of crack motion. Calculate the radius of the circular contact area that develops between the elastomer and the surface, assuming that there is no external load applied to the sphere (apart from it's weight).

78.13 Weibull Statistics

(25) A set of glass rods with a Weibull modulus of 30 are fractured in a uniaxial tensile test. If the stress at which 63% of the samples fracture is 100 MPa, what stress should you apply to make sure that less than 1% of the rods fail? What is the maximums stress if you want to make sure that less than one in10^6 rods fail? (Note that 1/e is 0.37). What does the stress need to be to get less than 1 failure in10^6 if the Weibull modulus is 10 instead of 30?

(26) What determines the value of the Weibull modulus in a brittle material?

78.14 Fracture Toughness

(27) Hertzberg 7.5: What is crack tip shielding?

(28) Describe the difference between a crack and a craze? How is the maximum width of a craze related to$\mathcal{G}_c$ andK_{IC}?

(29) Describe how transformation toughening works to increase the toughness of a ceramic material likeZrO_2.

(30) What is a Charpy impact test conducted, and what does it measure?

(31) Hertzberg 7.26: What the difference between the side windows of your car and the windshield? Include the role of tempering, thermal annealing and composite layering, and describe how the desired properties are obtained for the two different applications of glass.

78.15 General Knowledge

(32) How are$G$,ν and$K$ related to one another for an isotropic material?

(33) What are typical values of$G$ and$K$ for metals, ceramics and polymers?

79 332 Computational Exercises

79.1 MAT SCI 332 Report Guidelines

Compile the team's work into a single report. Your report should contain:

(1) A brief introduction.

(2) The work of each *individual* student as a separate section. You can work together on the finite element analysis, but write up the work for your component individually and combine work with labmates later.

(3) A conclusion for the project as a whole.

I am looking for accuracy and deliberation in your modeling and thoughtful consideration of the results. I am not concerned with formatting or the length of your report particularly, but as for any report you write, readability, good citation sourcing, informative figure presentation, and conciseness are valued. Broadly, the body of the report should include the following:

(1) The problem definition: geometry, materials properties, loads and boundary conditions.

(2) Discretization: Element type and mesh parameters.

(3) The results: Provide the results as a function of parameter values and discretization, as required. Show *relevant* results — displacement or stresses — and make sure to note which stress values you are showing. Note how your results change as the parameters or mesh changes.

Hint - well-constructed figures are very important to a good report. You may want to carefully design your figures so that that they are efficient at conveying relevant information.

Note that the Abaqus/CAE viewport can be exported as an image by selecting **File** $\longrightarrow$ **Print**.

Submit an electronic copy of your report, as well as your Abaqus files (.cae and .odb files), by uploading them to CANVAS the date that they are due.

79.2 332 Computational Lab #1: Finite Element Analysis of a Cantilever Beam

Consider the elastic deformation of a beam. From applied mechanics, it is well-known that the cross-sectional shape of the beam has an important effect on its bending stiffness (K), or the resistance of a member to bending deformation. Here we will use the simplest example of a rectangular cross-sectioned 3D beam element to compare the Finite Element Analysis (FEA) and analytical results.
Directions: Model a cantilever beam loaded on its top surface by a *distributed* force P, as shown in Figure 79.1.

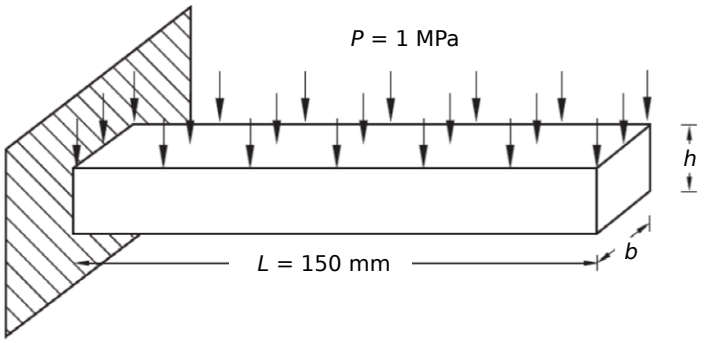


Figure 79.1: Our model for a cantilever beam.

The beam is made of a material with elastic modulus E, and a Poisson's ratio ν. Its dimensions are area $A = h \times b$, where h and b are the rectangle's height and base, respectively. The parameters we will use for this analysis are:

- $A = 500\,\text{mm}^2$
- $L = 150\,\text{mm}$
- $P = 1\,\text{MPa}$
- $E = 200\,\text{GPa}$ ($\sim$ steel)
- $\nu = 0.3$
- h and b = variable (start with 20 mm and 25 mm)

For this exercise, each group will choose 2-3 aspect ratios of $\frac{h}{b}$ (be reasonable please: not too much more than 5/1), keeping the area of the beam approximately constant, and solve for the deformation using Abaqus. For your report, you will address the following questions:

(1) What is the deflection of the free end of the beam for your selected ratios? Compare to exact values from the analytical max deflection: $\delta_{max} = \frac{FL^3}{8EI}$, where F is the force, L is the length of the beam, E is the elastic modulus, and I is the area moment of inertia.

(2) What are the stresses in the beam? Do your beams remain in the elastic regime under this load?

(3) What is the effect of the boundary condition on the stresses in the beam?

(4) What conclusions can you draw regarding the shape of the beam cross-section and the bending stiffness?

79.3 332 Computational Lab #2: Finite Element Analysis of Stress Concentrators

Stresses and strains are concentrated at holes, slots, or changes of dimension in elastic bodies. Here we will use finite elements to compute local stresses near stress concentrators. We will also utilize some new FEA techniques (non-native geometries, segmentation, seed gradients, symmetry boundary conditions, mesh density) to analyze a component that was designed in an external program and used as a source file for a 3D printed part. This computational lab is in many ways analagous to the experimental exercise completed earlier in the course.
Directions: Load and modify a CAD file (previously used for as source file in your 3D printing lab). Modify it to produce an accurate simulation of a stress concentrator (a circular or elliptical hole) as shown in Figure 79.2. As in

the last lab, you can work together, but you are each individually responsible for presenting one model in your report.

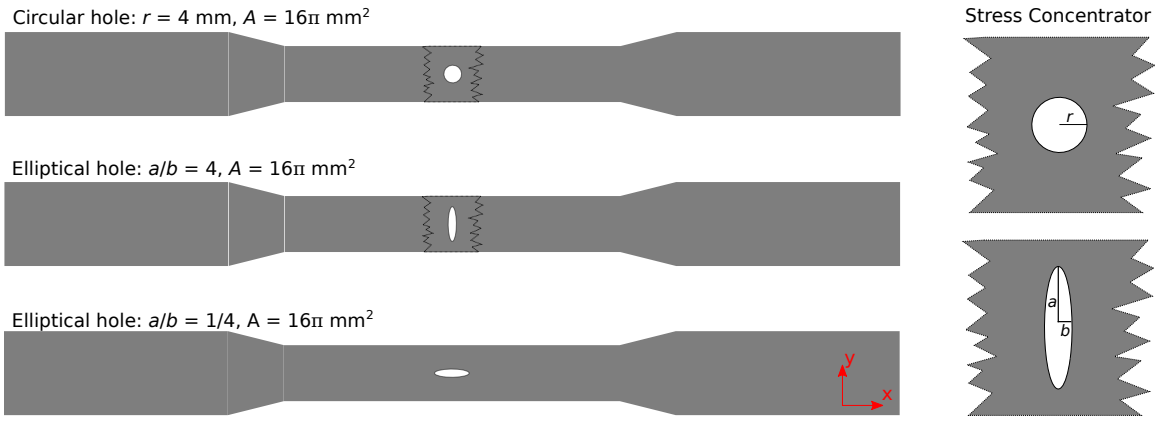


Figure 79.2: Dog-bone specimens with circular (radius r) and elliptical stress concentrators (dimensions a and b).

The 3D-printed dog-bone specimen made from polymer proved to have non-isotropic tensile behavior and is too complex of a problem at this level. Instead, we will consider a dog-bone specimen of identical dimensions made of steel (elastic modulus $E = 209 \times 10^3$ MPa, and a Poisson's ratio of $\nu = 0.3$. We will use the .SAT files to define the dimensions of the material to ASTM standard. For the stress concentrators (holes) make sure that the total extruded area is equal to 16π mm^2 Use a total applied force of 10 MPa and use element type C3D8R.
The directions for this lab will be less explicit than the previous Cantilever Beam Lab. Your goal is to construct as good of a model as possible. You know how to construct an Abaqus simulation, but here are some tips for approaching this more advanced problem:

(1) **Working with Non-native Geometries:** It's perfectly possible to reproduce the ASTM D638 Type I geometry using native design tools in Abaqus/CAE. However, you already have a .SAT file from your earlier work (available on Canvas in ./Computational Lab Documents/Lab #2/). Abaqus can read this file format. Simply go to **File** → **Import-Part** and browse to find the ASTM D638 Type I.SAT file. Rename the part as you like, and accept defaults for the Part Filter. Your part should appear in your workspace.

(2) **Modeling Forces/Boundary Conditions:** You need not perfectly model the serrated grips' interaction with the specimen shoulder (although you could), but carefully consider the forces on your specimen when designing your boundary conditions.

(3) **Symmetry Operations:** One of the best ways to reduce computational expense is to take advantage of symmetry in your model. You can use cleverly defined symmetry conditions to reduce your mesh size by and arrive at the most elegant FEA solution.

(4) **Meshing Algorithms:** The accuracy of FEA simulations are highly dependent on meshing. If possible, you want to avoid irregular meshes with elements that have high skews or low aspect ratios. The native dog-bone geometry, as loaded, requires a sweep mesh algorithm (it is yellow in the mesh module), which often works, but can lead to mesh defects. It is possible to use *partitioning* to create more regular segments that can be meshed using structural meshed algorithms.

(5) **Mesh Density and Seed Gradients:** Your results will be greatly effected by your mesh density in the regions of high stress (near the concentrator). Too coarse of a mesh will yield imprecise stress values. You should systematically adapt your mesh to ensure sufficient mesh density but (somewhat) efficient computation. Use of seed gradients are an excellent way to do this (hint.)

(6) **Python Scripting (totally optional):** Scripting can greatly speed your analyses. If you'd like, look at this guideby J. T. B. Overvelde to get started.

For your write-up, please make sure to address the following:

(1) Describe/show how the specimens deform.

(2) Describe/show the stresses in the specimen. Relate this to your theoretical understanding of stress concentrators as discussed in class.

(3) How do your stresses depend on mesh density? Note, you will have to examine a number of mesh densities per specimen to investigate this.

(4) What conclusions can you draw regarding the stress concentration at hole with various shapes?

(5) Where do the maximum stresses occur in the specimen's long axis? σ_{11}, σ_{22}, or σ_{33}, depending on your coordinate system)? Compare the maximum value with the applied stress.

(6) Prove that your mesh density near the stress concentrators is sufficient. To do this, plot the maximum short-axis σ values and long axis σ values relative to the applied stress as a function of the number of elements around the hole. Compare these relative values with the theoretical solution for the stress concentration near a circular hole in an infinite plate.

And, as always:

(1) Make sure you use units in your report.

(2) Label all graphs, figures, and tables appropriately (legend, axes, units, captions, etc.)

(3) Cite your sources.

(4) Upload the reports to Canvas as a .pdf. Include your (intuitively named) .cae and .odb files.

Part VII
331: Soft Materials

80 Introduction

Since the title of this book is 'Soft Materials', it makes sense to define what we really mean by 'soft'. Here are two ways to think about it:

(5) Soft Materials have Low Elastic Moduli.

By 'low' we mean significantly lower than the moduli of crystalline metals and ceramics. The jellyfish shown in Fig.80.1 is obviously 'soft' in this sense. Metals and ceramics typically have moduli in the range of 100 MPa (see Fig.80.2). While the strength of metals can be adjusted by a variety of mechanisms that affect the nature of dislocation motion in these systems, the the modulus is set by the nature of the interatomic potentials and there nothing that can really be done to significantly affect the modulus of a given material. Polymers are different, however, and have a much broader range of elastic moduli. The stiffest of these (Kevlar™for example) have elastic moduli in at least one direction that are comparable to the modulus of steel.

(1) Thermal Fluctuations Matter in Soft Materials.

At a molecular level, the relevant energy scale that determines a variety of important properties is the thermal energy,k_BT, wherek_B is Boltzmann's constant and$T$ is the absolute temperature. If different molecular arrangements within a material differ in energy by an amount that of the order ofk_BT or less, than these different arrangements will all be experienced by the material. When the free energy of a material is dominated by the entropy associated with the accessibility of these different arrangements, it is possible to calculate the elastic properties of the material from the molecular structure with considerable accuracy. The best example of this is our discussion ofRubber Elasticity.

Figure 80.1: An example of a 'soft' material.

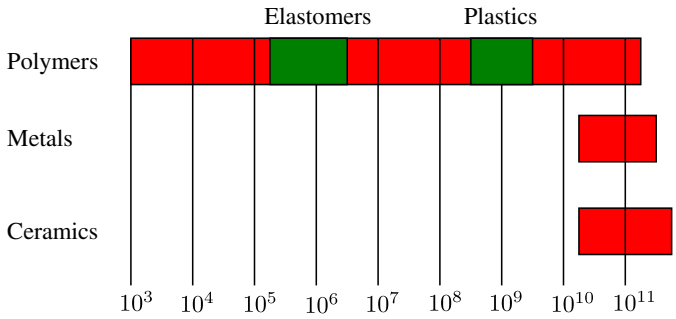


Figure 80.2: Range of Young's moduli (Pa) for different materials classes.

Exercise: How high above the earths surface must a single oxygen molecule be lifted in order for its gravitational potential energy to be increased byk_BT?
Solution: The gravitational potential energy ismgh, where$m$ is the mass of the object,g is the gravitational acceleration (9.8 m/s^2) and$h$ is height. The mass of a singleO_2 molecule is obtained by dividing the molecular weight in g/mole by Avogadro's number:

$$m = 32\left(\frac{\text{g}}{\text{mole}}\right)\left(\frac{\text{mole}}{6.02x10^{23}}\right) = 5.3x10^{-23}\,\text{g} = 5.3x10^{-26}\,\text{kg}$$

The conversion to kg illustrates the first point that we want to make with this simple calculation:**Don't mix units and convert everything to SI units (m-kg-s) to keep yourself sane and avoid unit errors.** Boltmann's constant,k_B needs to be in SI units as well:

$$k_B = 1.38x10^{-23}\,\text{J/K}$$

Now we just need to equatek_BT tomgh, using something reasonable for the absolute temperature,T. I'll use$T$=300 K to get the following:

$$h = \frac{k_BT}{mg} = \frac{(1.38x10^{-23}\,\text{J/K})\,(300\,\text{K})}{(5.3x10^{-26}\,\text{kg})\,(9.8\,\text{m/s}^2)} = 8000\,\text{m} = 8\,\text{km}$$

The full oxygen partial pressure is given proportional to$\exp(-mgh/k_BT)$, so this value of$h$ is the altitude where the oxygen pressure is a factor of$e$ (2.7) less than the value at sea level.

80.1 Synthetic Polymers

Most of this book is devoted to synthetic polymers because of their widespread use and importance, and because they illustrate many of the key concepts that are relevant to natural polymers and other soft materials. We begin with a basic introduction to the synthesis of polymeric materials,, and the z component is between properties of these materials. We conclude with some examples of soft materials made up of building blocks that are held together by relative weak forces.

80.1.1 What is a Polymer

A polymer is a large molecule made from many smallrepeat units or 'mers'. There is an inherent anisotropy at the molecular level because both strong and weak bonding interactions are important:

- Strong covalent bonds are formed within a molecule (between 'mers').
- Weak Van der Waals or hydrogen bonding are formed between molecules, and cause the materials to condense into a solid or liquid phase.

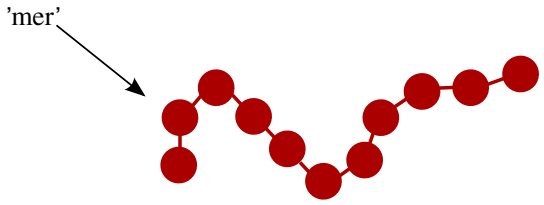


Figure 80.3: Schematic representation of a polymer

80.1.2 Classification Scheme

Crystallization and glass formation are the two most important concepts underlying the physical properties of polymers. Polymers crystallize at temperatures belowT_m (melting temperature) and form glasses at temperatures belowT_g (glass transition temperature). All polymers will form glasses under the appropriate conditions, but not

all polymers are able to crystallize. Theclassification scheme shown in FigureClassification scheme for polymeric materials. divides polymeric materials based on the locations ofT_g andT_m (relative to the use temperature,T) and is a good place to start when understanding different types of polymers.

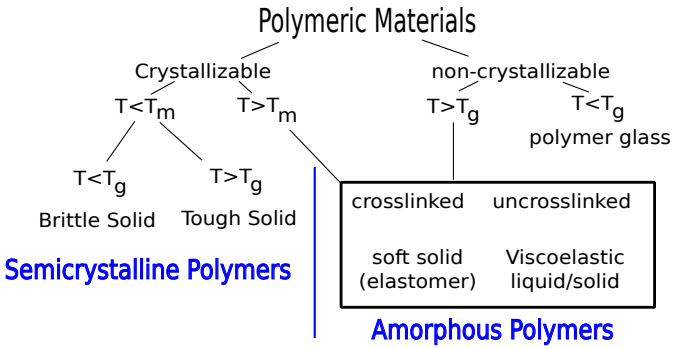


Figure 80.4: Classification scheme for polymeric materials.

Elastomers

Traditional elastomers are amorphous materials with a glass transition temperature less than the use temperature so that they remain flexible. The are generally crosslinked so that they do not flow over long periods of time. Common examples are shown in Figure80.5.

Polychloroprene (Neoprene)

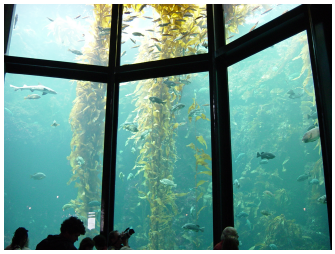
Polyisobutylene

Silicone

Figure 80.5: Examples of elastomeric materials.

Glassy Polymers

Glassy polymers are amorphous like elastomers, but their glass transition temperature is above the use temperature. Because of this they behave as rigid solids, with elastic moduli in the range of10^9 Pa. Glassy polymers do not need to be crosslinked, because belowT_g the molecules and flow of the material is suppressed. Also, because the materials are homogenous over length scales comparable to the wavelength of light, they are transparent. Common examples are illustrated in Figure80.6.



Poly(methyl methacrylate) (PlexiGlas)

Polycarbonate

Poly(phenylene oxide)

Figure 80.6: Examples of glassy polymers.

Semicrystalline Polymers

Semicrystalline polymers must have molecular structures that are compatible with the formation of an ordered lattice. Most atactic polymers are amorphous (non-crystalline) for this reason (with the exception being examples like poly(vinyl alcohol where the side group is very small). Another requirement is that T_g be less than T_m. If the glass transition temperature is higher than T_m the material will form a glass before crystallization can occur. In the glassy polymer the material is kinetically trapped in the glassy state, even though the crystalline state has a lower free energy below T_m. Examples of some semicrystalline polymers are shown in Figure 80.7.

Polyethylene terephthalate.

Poly(tetrafluoroethylene) (Teflon)

Spider Web

Figure 80.7: Examples of semicrystalline polymers.

80.2 Understanding Polymer Chemistry

Crystallization and glass formation processes are central to our understanding of polymeric materials, we must eventually address the following question:

- How is a polymer's tendency to crystallize or form a glass determined by is molecular structure?

Before we answer this question, however, we must answer the following question:

- What determines this molecular structure, and how are our choices limited?

In order to answer this question properly, we need to study the processes by which polymeric materials are made. Polymer synthesis involves organic chemistry. After familiarizing ourselves with some of the relevant polymer chemistry, we will be in a position to study the physical properties of polymers. For this reason, our discussion of molecular structure in polymers will include some chemistry

80.2.1 Covalent Bonding

Polymer molecules consist of atoms (primarily carbon, nitrogen, oxygen and hydrogen) which are covalently bound to one another. It is useful at this point to recall some of the basic principles governing the bonding between these atoms:

- Nitrogen, oxygen and carbon atoms are surrounded by 8 electrons, included shared electrons.
- Hydrogen atoms are surrounded by 2 electrons, included shared electrons.
- A single bond involves two shared electrons, a double bond involves 4 shared electrons, and a triple bond involves 6 shared electrons.

80.2.2 Lewis Diagrams

Lewis diagrams (Wikipedia link) provide a convenient way of keeping track of the valence electrons in covalently bonded compounds. Several examples are given here. Note how the rules given on the previous page are followed in each case.

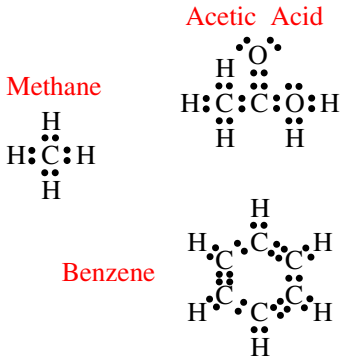


Figure 80.8: Examples of Lewis Diagrams

80.2.3 Bonding

The following principles of covalent bonding in organic materials are very helpful:

- Oxygen: group 6 (6 valence electrons) - 2 more needed to complete shell -oxygen forms 2 bonds with neighboring atoms.
- Nitrogen: group 5 (5 valence electrons) - 3 more needed to complete shell -nitrogen forms 3 bonds with neighboring atoms.
- Carbon: group 4 (4 valence electrons) - 4 more needed to complete shell -4 bonds with neighboring atoms.

in and consider the actual backbone bonds that are connecting the polymriple bond counts as 3 bonds. The chemical structures throughout this book can be seen to obey these rules. For an example, considerPolyamides andPolyesters.

80.2.4 Shorthand Chemical Notation

Most of the chemical structures illustrated in this text are relatively simple, consisting of single and double bonds between atoms. We generally don't bother to write all of the carbons and hydrogens into to the structure. We use the following common conventions.

- If no element is included, element at junctions between different bonds are assumed to be carbon.
- If atoms are missing, so that the rules given above for the number of bonds attached to each atom type are not satisfied, the missing atoms are hydrogens.

An example of this convention is shown below in the structure forbenzene::

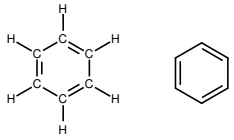


Figure 80.9: Examples of Shorthand Notation when drawing chemical structures

Note the resonance between the two possible ways of drawing the double bonds in the drawing on the left. Molecules with these types of alternating double and single bonds are referred to asaromatic compounds.

81 Polymerization Reactions (week 1)

Polymerization is the process by which small molecules react with one another to form large polymer molecules. Polymerization reactions can be broken up into the following two general categories:

Step Growth Polymerizations: Collections of A and B species react with one another. In linear step growth polymerizations, the ends of molecules react with one another to form longer molecules. A variety of reactions are possible, so you need to know at least a little organic chemistry. We'll focus on just a few of the most common cases.

Chain Growth Polymerizations: Each polymer chain has one reactive site to which additional monomers are added.

Additional Resource: The Macrogalleria web site has some excellent, simple descriptions of polymerization reactions. Specific examples are referenced at different points throughout this book.Click here for Macrogalleria polymerization overview.

81.1 Step-Growth Polymerizations

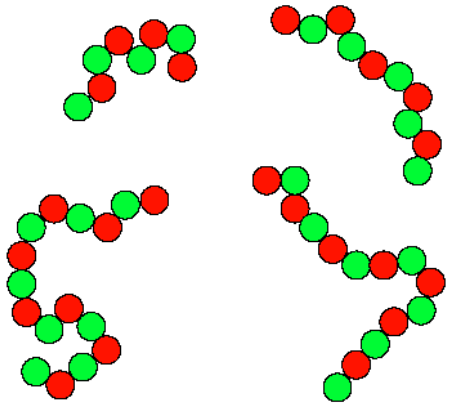


Figure 81.1: Polymers produced by step growth polymerization. The red and green circles correspond to 'A' and 'B' monomers that react with one another to form the polymer.

In this example green and red monomers can only react with each other. Because there are 5 more red monomers than green monomers, there are 5 molecules remaining at the end of the reaction, with each of these molecules possessing two red end groups. The extent of reaction,p is defined as the fraction of available reactive groups which have actually undergone a reaction. Values close to one are needed in order to obtain useful, high molecular weight polymer. A delicate stoichiometric balance generally needs to be maintained (same amount of red and green monomers) in order to obtain high molecular weight This is why it is always amorphousymerization.
idx step-growth animation is a schematic illustration of the buildup of large polymer molecules from small molecule building blocks. The following pages illustrate some of the specific reactions which take place during the polymerization process. To illustrate the concepts involved, we consider the following polymer types, all of which are produced by step-growth polymerization:

(1) Polyamides

(2) Polyesters

(3) Polyurethanes

(4) Epoxies

81.1.1 Polyamides

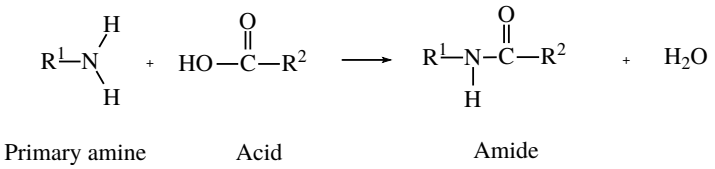


Figure 81.2: Formation of an amide from an amine and a carboxylic acid

In this example, a primary amine reacts with a carboxylic acid to form an amide linkage. Water is liberated during the condensation reaction to form the amide. Primary amines and acid chlorides undergo a similar reaction:

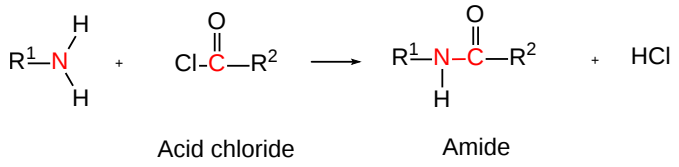


Figure 81.3: Formation of an amide from a primary amine and an acid chloride

Acid chlorides react very rapidly with amines at room temperature, which is very useful for demonstration purposes. Acid chlorides can also react with water to form carboxylic acids, however, and commercial polyamides are generally produced by reaction with carboxylic acids.

81.2 Interfacial Polymerizations

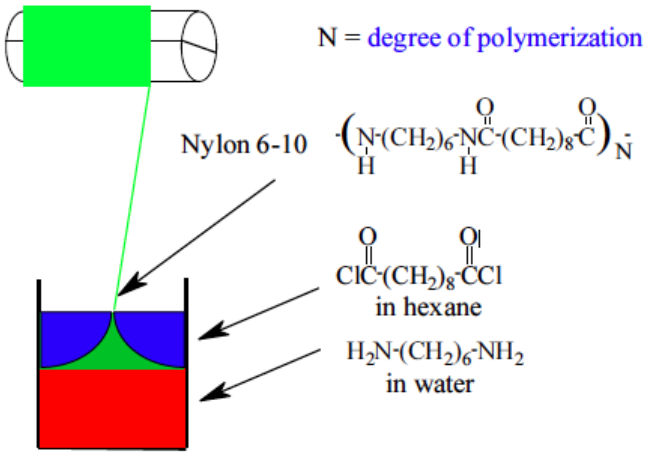


Figure 81.4: Schematic representation of the interfacial polymerization off nylon.

81.2.1 Polyesters

Polyesters can be formed by condensation reactions of alcohols with carboxylic acids:

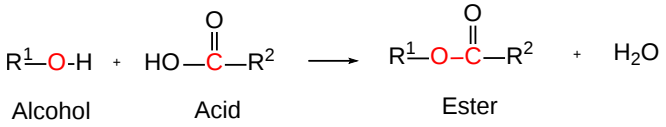


Figure 81.5: Formation of an ester from an alcohol and a carboxylic acid.

Polyethylene terephthalate (Mylar, Dacron, 2L soda bottles) is a common example. Note that under appropriate conditions, the reverse reaction (hydrolysis) reaction can also take place, where the addition of water to an ester bond forms the acid and the alcohol. This reaction is important in a variety of polymers used in biomedical applications, which degrade in the body via hydrolysis of the polymer. Polycaprolactone is one example.

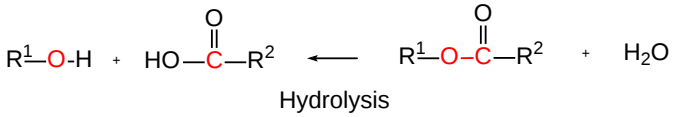


Figure 81.6: Ester hydrolysis: The reverse of the esterification reaction.

81.2.2 Polyurethanes

Polyurethanes are formed by the reaction of isocyanates with alcohols, as shown below:

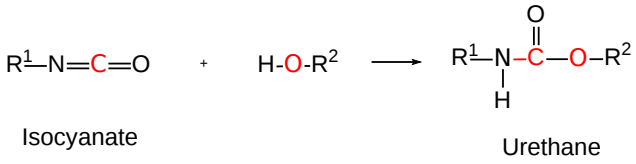


Figure 81.7: Urethane Formation

Note that this is NOT a condensation reaction, since no byproducts are formed during the reaction. Also note that the urethane linkage contains an oxygen atom in the backbone, whereas the amide linkage does not. (TheR^1 andR^2 substituents can have different structures, but are always attached to the illustrated linkages by carbon atoms.)

81.2.3 Epoxies

All epoxies involve reactions of epoxide groups (3-membered rings containing an oxygen atom) with curing agents. Amine curing agents are very common, as illustrated Figure , which shows a primary amine reacint with an epoxide group).

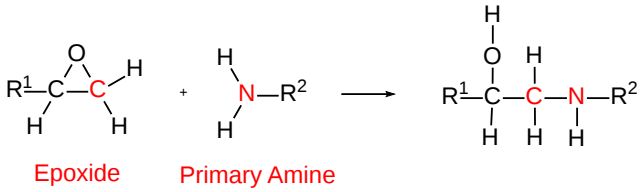


Figure 81.8: Reaction of an epoxide with a primary amine.

The secondary amines which remain can react with additional epoxide groups to form a branched structure as shown in Figure81.9. The reactive functionality of a primary diamine is 4 when the reaction is with an epoxide (as opposed to its functionality of 2 in the case where the diamine reacts with an acid or acid chloride).

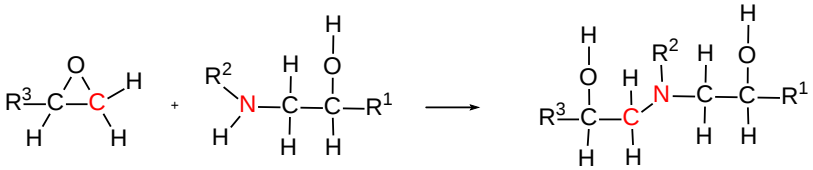


Figure 81.9: Reaction of an epoxide with a secondary amine.

82 Molecular Weight Distributions (week 2)

The repeat unit for this polymer consists of a red and green unit together. The degree of polymerization (N) is the number of repeating units for a single molecule. Because of the statistical nature of the reaction process, not all molecules at any given point in the polymerization will have the same size. This point is illustrated in Figure82.1, which shows schematic representations of polymer with and$N = 10.5$. The molecular weight,M, of a polymer is obtained from$N$ by multiplying byM_0, the molecular weight of a repeating unit of a polymer chain:

$$M = NM_0 \tag{82.1}$$

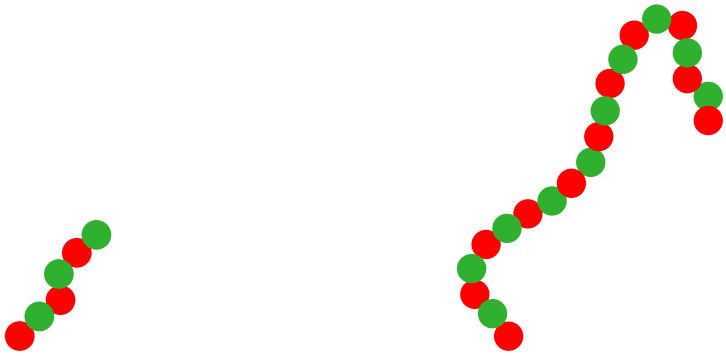


Figure 82.1: Linear, step-growth polymers with$N = 3$ (left) and$N = 10.5$ (right).

Some key definitions:
N: Degree of polymerization - the number of repeat units in a single molecule
M: Molecular weight - molecular weight of a polymer molecule (g/mole)
M_0 : Molecular weight per repeat unit (g/mole)

The complete distribution of molecular sizes can be described by a histogram. There are two different quantities that we can keep track of:n_N andw_N, which are defined as follows:

- n_N: number of molecules with degree of polymerization =N
- w_n: weight of all molecules with degree of polymerization =N

The total number of molecules (n) or total weight (w) is obtained by summing over all possible values of$N$:

$$n = n_1 + n_2 + n_3 + ...n_\infty = \sum_N n_N \tag{82.2}$$

$$w = \sum_N w_N = M_0 \sum_N N n_N \tag{82.3}$$

Note: We've assumed here for simplicity that$N$ takes on integer values. This is true for chain growth polymerizations, but for step growth polymers$N$ can formally have non-integer values (like the value of 10.5 in Figure82.1). In this case you can just imagine that distribution has divided into 'bins' of width '1'.

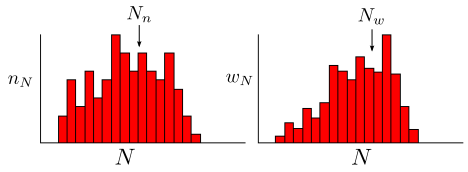


Figure 82.2: Histogram of molecular sizes. The same distribution is shown, with the numbers of molecules of each size (n_N1*2/1tted on the left, and the total weight corresponding to each of the different molecular sizes (w_N53plotted on the right.

82.1 Number Average Molecular Weight

The number average molecular weight,M_n, is simply the total weight of all the polymer, divided by the number of polymer molecules (typically in moles):

$$M_n = w/n \tag{82.4}$$

In more formal terms, we can replace$w$ and by the respective sums from Eqs.82.2 and82.3:

$$M_n = \frac{M_0 \sum_N N n_N}{\sum_N n_N} \tag{82.5}$$

By dividing by the repeat unit molecular weight,M_0, we obtain the number average degree of polymerization,N_n:

$$N_n = \frac{\sum_N N n_N}{\sum_N n_N} \tag{82.6}$$

82.2 Weight Average Molecular Weight

The weight average molecular weight,M_w, is similar to number average, but we use the distribution of weights instead of the distribution of molecules:

$$M_w = \frac{M_0 \sum_N w_N N}{w} = \frac{M_0 \sum_N w_N N}{\sum_N w_N} = \frac{M_0 \sum_N n_N N^2}{\sum_N n_N N} \tag{82.7}$$

The weight average degree of polymerization,N_w, is obtained simply by dividingM_w byM_0,*i.e.*$N_w = M_w/M_0$. The polydispersity index, given byM_w/M_N must be greater than or equal to one. Polymers with a polydispersity index close to 1 are said to be monodisperse, with a narrow distribution of molecular weights. While Eq.82.7 is useful in a formal sense, it's a bit cumbersome to use. Suppose we have a series of molecular weight fractions, withw_i between the weight of each fraction andM_i being the weight average molecular weight of each fraction. The weight average molecular weight for this collection of molecules is given by the following summation:

$$M_w = \sum_i w_i M_i \tag{82.8}$$

Exercise: Suppose 3g of a monodisperse polymer with M=50,000 g/mol is blended with 5g of a monodisperse polymer with M=85,000 g/mol. What are the values ofM_n, and the polydispersity index for the resultant mixture?
Solution: To getM_n, we need to divide the total weight of polymer (3g + 5g = 8g) by n, the total number of moles of polymer molecules:

$$n = \frac{3g}{50,000g/mol} + \frac{5g}{85,000g/mol} = 1.9x10^{-4}mol$$

$$M_n = \frac{8g}{1.19x10^{-4}mol} = 67,200g/mol$$

The average molecular weight is obtained from the definition of this quantity:

$$M_w = \frac{M_0 \sum_{N=1}^{\infty} w_N N}{w} = \frac{1}{w}\sum_{N=1}^{\infty} w_N M_N \tag{82.9}$$

In our example, we obtain the following:

$$M_w = \frac{1}{8g}(3g * 50,000g/mol + 5g * 85,000g/mol) = 71,875g/mol$$

Polydispersity Index =

$$\frac{M_w}{M_n} = \frac{71,875g/mol}{67,200g/mol} = 1.07$$

Exercise: What if the polymers in the previous example are not monodisperse? Suppose the molecular weights quoted are weight averages, the low molecular weight polymer has a polydispersity of 1.6, and the high molecular weight polymer has a polydispersity of 2.5.
Solution: We divide by the polydispersities to get number average molecular weights, and use these values to get the total number of molecules:

$$M_n(poly.a) = \frac{50kg/mol}{1.6} = 31kg/mol$$

$$M_n(poly.b) = \frac{85kg/mol}{2.5} = 34kg/mol$$

$$n = \frac{3g}{31,000g/mol} + \frac{5g}{34,000g/mol} = 2.44x10^{-4}mol$$

$$M_n(blend) = \frac{8g}{2.44x10^{-4}mol} = 32,800g/mol$$

Due to the form of the expression for the average molecular weight, we can modify the standard expression for the Mw by summing over all different components of the blend, and using the weight average molecular weight for each component:

$$M_w(blend) = \frac{1}{w}\sum_i w_i M_{wi} \tag{82.10}$$

In our caseM_w is the same as it was in the previous example:

$$M_w = \frac{1}{8g}(3g * 50,000g/mol + 5g * 85,000g/mol) = 71,875g/mol$$

Polydispersity index =

$$\frac{M_w}{M_n} = \frac{71,875g/mol}{32,800g/mol} = 2.19$$

82.3 Molecular Weight Evolution during Step Growth Polymerization

Supposen_a moles of difunctional A molecules with molecuar weightM_a react withn_b difunctional B molecules with molecular weightM_b in a step growth polymerization. If$n_a > n_b$, then all of the B molecules can react, and$n_a - n_b$ molecules will remain at the end of the polymerization:

$$n^{min} = n_a - n_b$$

For condensation reactions, a mole of condensation product with molecular weightM_c material as a wholeecule of B that reacts. The total weight,w, of polymer produced is:

$$w = n_a M_a + n_b M_b - 2n_b M_c \tag{82.11}$$

The maximum number average molecular weight,M_n^{max}, is obtained by using this value for$w$, an setting$n$ equal ton_{min}:

$$M_n^{max} = \frac{w}{n^{min}} = \frac{n_a M_a + n_b M_b - 2n_b M_c}{n_a - n_b} \tag{82.12}$$

For high molecular weight polymers, we can make some additional approximations, since know that$n_a - n_b$ <<n_b in this case. We can therefore simplifyM_n to the following:

$$M_n^{max} \approx \frac{n_b \left(M_a + M_b - 2M_c\right)}{n_a - n_b} \tag{82.13}$$

The repeat unit molecular weight,M_0 is obtained by adding the molecular weights for the A and B monomers, and subtracting out the molecular weight of the 2 condensation products that are formed (since a repeat unit contains 2 A-B linkages, and therefore involves 2 condensation reactions):

$$M_0 = M_a + M_b - 2M_c \tag{82.14}$$

From this we obtain the following expression forM_n^{max} :

$$M_n^{max} \approx \frac{n_a M_0}{n_a - n_b} \tag{82.15}$$

The next thing we would like to sort out is how$n$, the total number of molecules, varies with the extent of reaction,p. The trick is to realize that every time a reaction takes place, the number of molecules in the system decreases by 1. Each molecule has two groups (one at each end) that can react, and two groups take part in each reaction. For$n$ we obtain the following:

$$n = n_a + n_b - p\left(n_a + n_b\right) = \left(n_a + n_b\right)\left(1 - p\right) \tag{82.16}$$

In this equation$n_a + n_b$ is the initial number of molecules that we start off with (when$p$=0), and$p\left(n_a + n_b\right)$ gives the decrease in the number of molecules resulting from all of the reactions that have taken place for some finite value of$p$. The total mass of the polymer that has been produced is$\approx \left(n_a + n_b\right) M_0/2$ (ignoring the fact that the ends of all the molecules some extra mass that may potentially leave as part of a condensation product). The weight average molecular weight is simplyw/n, from which we obtain the following:

$$M_n = \frac{w}{n} = \frac{(n_a + n_b)M_0}{2(n_a + n_b)(1 - p)} = \frac{M_0}{2(1 - p)} \tag{82.17}$$

As mentioned above, the degree of polymerization that can be obtained from a homogeneous mixture of A and B monomers in a linear stop growth polymerization is limited by stoichiometry. Reaction of all A monomers requires the existence of unreacted B monomers, and*vice versa*. If$n_a \neq n_b$, then some of the functional groups on either the A or B monomers must remain unreacted, so that$p$ has a maximum value somewhat less than 1. Let's consider the case where$n_a - n_b$, in which case we know there are$n_a - n_b$molecules left when all the B groups have reacted. Each of these molecules has an unreacted A group at each end, so there are a total of$2\left(n_a - n_b\right)$ unreacted groups at the end of the reaction. At the beginning of the reaction (when$p$=0) we had$n_a + n_b$ reactive groups, so the total number that have reacted is$2\left(n_a + n_b\right) - 2(n_a - n_b) = 4n_b$. This corresponds to a fraction of reacted groups which we refer to asp^{max}:

$$p^{max} = \frac{4n_b}{2\left(n_a + n_b\right)} = \frac{2n_b}{n_a + n_b} \tag{82.18}$$

Another way to obtainM_n^{max} is to use Eq.82.17 forM_n, with$p = p^{max}$:

$$M_n^{max} = \frac{M_0}{2(1 - p_{max})} = \frac{M_0(n_a + n_b}{2\left((n_a - n_b)\right.} \tag{82.19}$$

which with$n_a \approx n_b$ is equivalent to Eq.82.15.

From the information given so far we know how the number average degree of polymerization evolves with$p$, but in many cases we would like to know what the full molecular weight distribution might look like. To do this we need to account for the statistics, and use the more detailed definitions forM_n (Eq.) andM_w (Eq.) A molecule formed from a linear step growth polymerization that has a degree of polymerization of$N$ was made from$2N$ molecules (N of type A and$N$ of type B). Because these$2N$ original molecules are all bifunctional, there are a total of$4N$ reactive groups in the monomers that reacted together to form a polymer of length$N$. Only two of these groups remain unreacted (the end groups). Because each reaction involves two groups, this means that$2N - 1$ reactions have taken place ($4N - 2$ groups have reacted).

The key assumption is that all groups react with equal probability. This probability must therefore be equal to the fraction of functional groups which has reacted, which has already been defined as the extent of reaction,p. The probability that any single one of these$2N - 1$ reactions has taken place is therefore equal to$p$. We have to multiply all of these probabilities together to get the probability of having a polymer with a degree of polymerization of$N$, so we end up with the following:

$$n_N \propto p^{2N-1} \tag{82.20}$$

The weight,w_N, corresponding to each fraction is proportional to$n_N N$:

$$w_N \alpha N p^{2N-1} \tag{82.21}$$

The shapes of these distributions forn_N andw_N are shown in the plots below, and can be replotted with different values of$p$ in the MATLAB modulue referred to below.

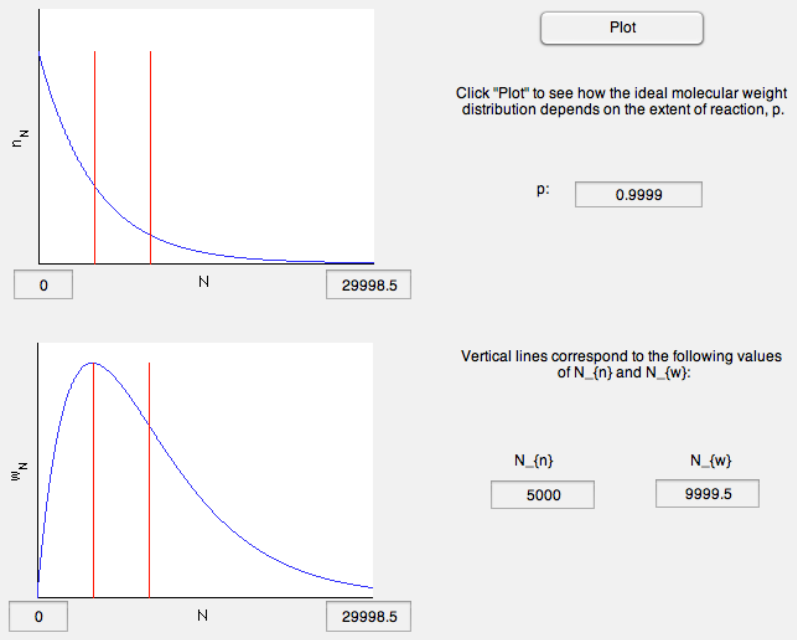


Figure 82.3: Distribution of molecular sizes during an ideal step growth polymerization.

(1) What happens to the polydispersity index as the degree of polymerization increases?

(2) Where are the number and weight averaged degrees of polymerization in comparison to the peak in w_N?

(3) How does the shape of the molecular weight distribution change as N gets very large?

(4) What sizes of molecules are present in the highest numbers?

(5) For what values of p is Nn equal to 0.5, 1, and infinity? Can you develop a simple explanation for these results? What are the values of the polydispersity for these values of p?

Analytic expressions can be derived for number-average and weight average degrees of polymerization for ideal step growth polymerizations that follow the distribution described by Eq.82.20. Substitution of this expression into Eq.82.5 gives the following for M_n:

$$M_n = \frac{M_0 \sum_N N p^{(2N-1)}}{\sum_N p^{(2N-1)}} \tag{82.22}$$

We assume that our material is isotropic N, *i.e.*, N=1/2, 1, 3/2, 2, etc. Half integer values of N are allowed because N is the number of repeat units in the molecule. (A value of 3/2 for N would correspond to a molecule with one A monomer between two monomers, for example). Alternatively, we can replace N by $L/2$, where L takes on all positive integral values:

$$M_n = \frac{M_0}{2} \frac{\sum L p^{(L-1)}}{\sum p^{(L-1)}} \tag{82.23}$$

We can follow a similar procedure to get an expression for M_w. Combination of the ideal, linear step growth size distribution (Eq.82.20) with the definition of M_w (Eq.82.7), gives

$$M_w = \frac{M_0 \sum_N N^2 p^{(2N-1)}}{\sum_N N p^{(2N-1)}} \tag{82.24}$$

we can again let $L = 2N$, so that L takes on all integer values, so that the expression for M_w becomes:

$$M_w = \frac{1}{2} \frac{M_0 \sum_N L^2 p^{(L-1)}}{\sum_N L p^{(L-1)}} \tag{82.25}$$

These expressions for M_n and M_w can be simplified by using the following mathematical formulas:

$$\begin{aligned} \sum_{L=1}^{\infty} p^{L-1} &= \frac{1}{1-p} \\ \sum_{L=1}^{\infty} L p^{L-1} &= \frac{1}{(1-p)^2} \\ \sum_{L=1}^{\infty} L^2 p^{L-1} &= \frac{(1+p)}{(1-p)^3} \end{aligned} \tag{82.26}$$

Substituting the results for these sums into Eqs.82.23 and82.25 results in the following:

$$M_n = \frac{M_0}{2(1-p)}; \;\; N_n = \frac{1}{2(1-p)}; \tag{82.27}$$

$$M_w = \frac{M_0}{2}\left\{\frac{1+p}{1-p}\right\} \tag{82.28}$$

The expression forM_n is the same as the result that we got from the previous section, just by keeping track of the total number of molecules in the system. Calculation ofM_w is a bit more complicated though, and requires the statistical approach used in this section. For the polydispersity index we get:

$$\frac{M_w}{M_n} = 1 + p \tag{82.29}$$

Useful, high molecular weight polymers produced by step-growth polymerization are therefore have a polydispersity index that is close to 2.

83 Non-Linear Step Growth Polymerizations (week 2)

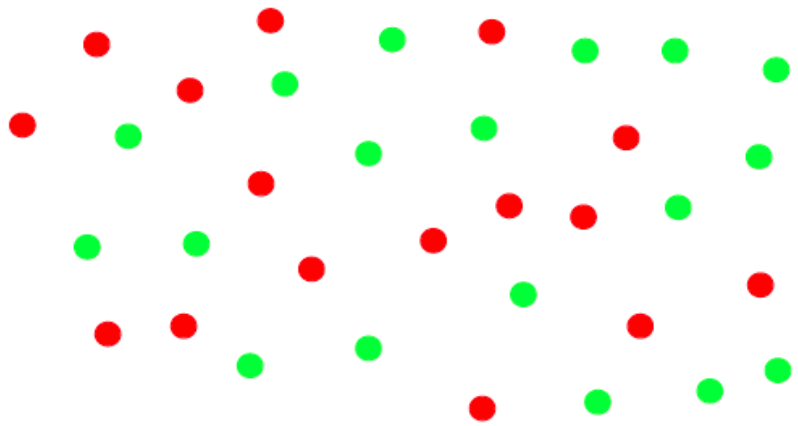


Figure 83.1: Nonlinear Step Growth Polymerization

Formula N_{av}rowth polymerizations involve the use of monomers that have a reactive functionality,f, which is greater than 2.

83.1 Gelation

A network (or gel) is formed when the molecular weight of the largest molecule in the system reaches infinity. Everything which is covalently attached to this single very large molecule is part of the gel fraction.

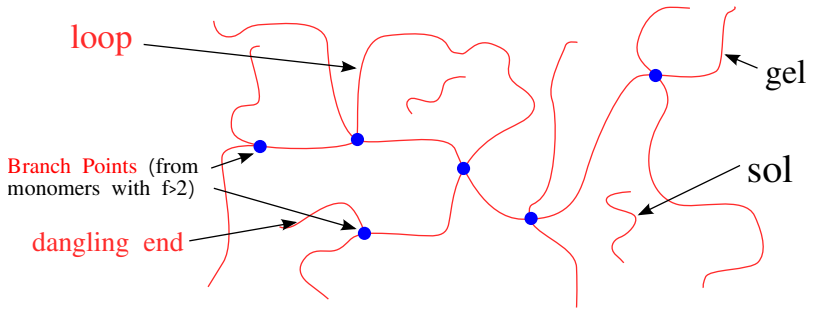


Figure 83.2: Molecular schematic of gel network.

Unattached molecules which are not part of thegel fraction make up thesol fraction. Thegel point refers to the point in the polymerization reaction at which point a gel fraction first appears, and the overall process is referred to asgelation.

83.2 Prepolymers

In many cases, the length of the polymer between branch points is determined by the molecular weight of aprepolymer. Prepolymers generally have a reactive functionality of 2.

Figure 83.3: Example of a Prepolymer

The following factors determine the property of the network:

(1) Molecular weight of prepolymer

(2) Composition of prepolymer

(3) Functionality of curing agent (4 for diamine curing of epoxies)

(4) Extent of reaction (number of dangling ends)

(5) Network structure (loops, entanglements, etc.)

83.3 Corothers Theory of Gelation

We begin withn_a molecules with a reactive functionality off_a that are able to react withn_b molecules with a reactive functionality off_b. For illustrative purposes, consider the perfectly stoichiometric case where the total functionality of all the A monomers is equal to the total functionality of all the B monomers:

$$n_a f_a = n_b f_b \tag{83.1}$$

We define an average reactive functionality,f_{av}, so that$(n_a + n_b)\, f_{av}$ is equal to the total number of functional groups:

$$(n_a + n_b)\, f_{av} = n_a f_a + n_b f_b \tag{83.2}$$

The number of initial molecules corresponding to a given number of functional groups is inversely proportional tof_{av}. For large values off_{av}, p does not have to reach unity in order for the total number of molecules to be reduced to one. An expression forN_n can be obtained substituting p$f_{av}/2$ for$p$ in Eq. , describing the evolution ofM_n with$p$ for linear step growth polymerization:

$$N_n = \frac{1}{2 - pf_{av}} \tag{83.3}$$

The quantitypf_{av} describes the average number of times that any given molecule has reacted. The number average molecular weight is predicted to diverge to infinity when this average number of reactions per molecule is 2. This occurs when$p = p_{gel}$, wherep_{gel}Standardextent of reaction at the endpoint. Setting$p_{gel} f_{av} = 2$ gives:

$$p_{gel} = \frac{2}{f_{av}} \tag{83.4}$$

This simple equation for the gel point is very useful, although it is important to remember that this derivation suffers from the following deficiencies:

(1) Gelation actually occurs for finiteN_n, since$N = \infty$ only for the largest molecule.

(2) The theory neglects loop formation (reactions between two portions of the same molecule). These reactions change the shape of a given molecule, but do not decrease the overall number of molecules.

The first assumption tends to overestimatep_{gel}, and the second assumption tends to underestimatep_{gel}. A more accurate description of gelation requires a much more detailed theory than the one presented here.
When one molecule which spans the entire sample, it can no longer flow like a liquid. One the characteristics of non-linear step growth polymerizations is that the viscosity (resistance to flow) of the reaction mixture increases as the extent of reaction increases, and eventually diverges at the gel point. Also, the reaction rate generally decreases as the reaction proceeds and the reactive molecules become larger and larger. These features are illustrated conceptually in Figure83.4. Here we show schematic representations of the time-dependent extent of reaction, along with the time dependence of the number average degree of polymerization,N_n, and of two physical properties of the material:η_0 andE_∞. Hereη_0 is the limiting viscosity measured at very low shear rates, andE_∞ is the elastic modulus measured at very long times. The viscosity is a characteristic of a liquid material, and characterizes the material for values of the extent of reaction,p, that are less than the extent of reaction corresponding to the gel point,p_{gel}. For values of$p$ that exceedp_{gel}, the material behaves as a solid, and has a finite elastic modulus,E_∞. At the gel point, the viscosity diverges to infinity, since it is no longer possible for the material to flow. Instead, the material becomes an elastic solid, with a value ofE_∞ that increases from 0 at the gel point.
Note that that materials near the gel point are viscoelastic, and can no longer be described by a single value of the viscosity or modulus. Instead, these properties depend on the timescale of the measurement, as described in more detail on the section on viscoelasticity.

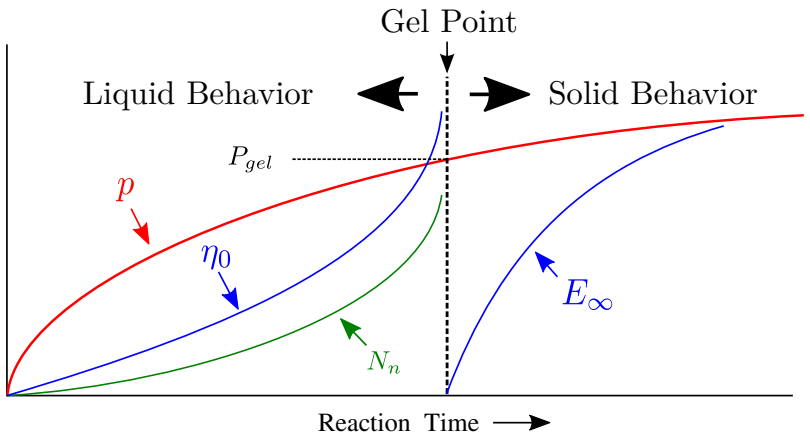
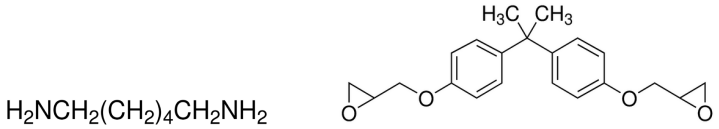


Figure 83.4: Time-dependence of property development in a nonlinear step growth polymerization.

Exercise: A very simple epoxy formulation has bisphenol A diglycidyl ether as the epoxide component and hexamethylene diamine as the crosslinker:



A) Hexamethylenediamine B) Bisphenol A diglycidyl ether

(1) How many grams of hexamethylene diamine should be added to 1g of Bisphenol A diglycidyl ether in order to optimize the polymerization reaction in a homogeneous solution?

(2) Estimate the fraction of amine groups that need to react in order to reach the gel point.

Solution: In this case monomer A (the hexamethylenediamine) has a molecular formula of$C_6H_{16}N_2$, corresponding to a molecular weight,M_a, of 116 g/mole. The functionality of monomer A,f_a us 4 in this case, since each primary amine is able to react with 2 epoxide groups. Monomer B has a molecular formula of$C_{21}H_{24}O_4$, givingM_b = 340 g/mole, and a reactive functionality,f_b of 2. We want the total functionality of the A monomers to be equal to the total functionality of the B monomers in this case, which means that the molar quantities of A and B monomers should be related by Eq.83.1, from which we obtain the following forn_a:

$$n_a = \frac{n_b f_b}{n_a} = \frac{n_b}{2}$$

We have 1g of B, which we divide by it's molecular weight to get the number of moles:
So we know we need about$1.45x10^{-3}$ moles of the hexamethylenediamine. Multiplying by its molecular weight of 116 g/mole gives a total mass of 0.17g. Note that we need a lot less of A than B in this case. If we want our two component epoxy to consist of two parts that we combine in nearly equal volumes, we'll have to add some extra stuff to part A (the diamine part) that doesn't take part in the curing reaction.
To figure out the average functionality, we arrange Eq.83.2 to give the following forf_{av}:

$$f_{av} = \frac{n_a f_a + n_b f_b}{n_a + n_b}$$

Combining this with Eq.83.4 for the gel point gives:

$$p_{gel} = \frac{2\,(n_a + n_b)}{n_a f_a + n_b f_b}$$

Substituting our stochiometric condition that$n_a = n_b/2$ gives$p_{gel} = 0.75$.

84 Chain-Growth Polymerizations

In chain growth polymerization monomers are added one at a time to a reactive site at that typically remains at the end of the monomer that has been most recently added. There are three phases of the reaction that need to be considered in general:

(1) Initiation: In this step the a reactive species is formed that is able to form covalent bond with a monomer, reforming the active species at the end of the growing polymer chain.

(2) Propagation: Monomers are added one at a time to the growing chain.

(3) Termination: The reactive undergoes a reaction of some sort that causes it to lose its reactivity toward other monomers. The polymer chain is 'dead' at this point, and no longer grows in length.

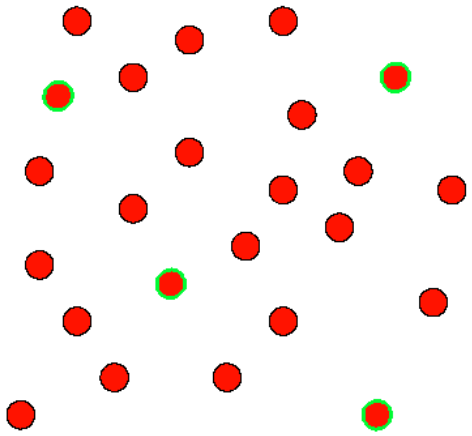


Figure 84.1: Chain growth polymerization.

84.1 Chain Growth Polymerization Mechanisms

Chain growth polymerizations generally occur either by addition to a double bond, or by opening a ring. In both cases the overall number of bonds is conserved, and there are no condensation products. If the cases where chain growth occurs by addition to double bond between two carbon atoms, the following takes place:

- A new bond is formed between the active sight and one if the doubly bonded carbons
- The double bond is shifted to a single bond.

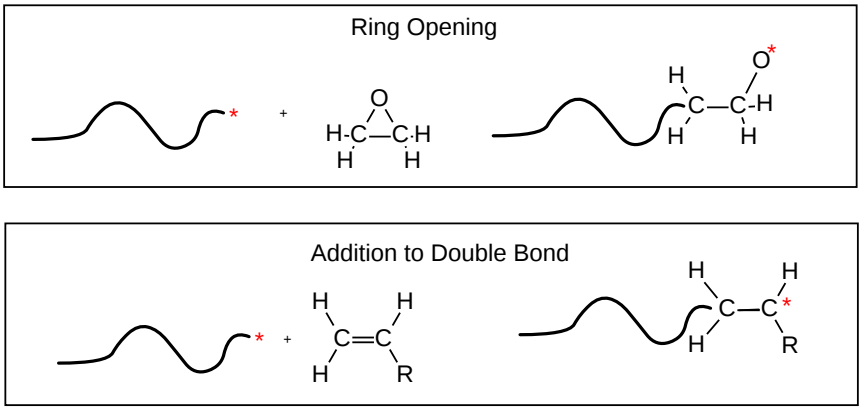


Figure 84.2: The two basic types of chain growth polymerizatoins.

In a ring opening polymerization, the following takes place:

- A new bond is formed between the active species and one of the atoms in a cylic monomer.
- An adjacent bond in the monomer is broken.
- The active site moves to the end of the linear molecule.

84.2 Reactive Species for Chain Growth Polymerization

Three of the most common active sites are a free radical, negatively charged anion, and a positively charged cation, illustrated in Figure84.3 for the addition to the double bond of vinyl polymers. Vinyl polymers have the general structureCH_2CHR, where R is something other than hydrogen. In each of these cases, a single monomer repeat unit is added to the end of the chain, and the reactive site moves to the end of the chain, on the repeat unit that has just been added. Reactions in organic chemistry are all about keeping track of what the bonding

electrons are doing, and Lewis diagrams are very helpful in this sense. In the figures designed to illustrate different propagation and termination reactions, we just show the Lewis structures for some of the bonds, to make it simpler to keep track of the situation before and after the reaction has taken place. For our purpose we are not as interested in the detailed reaction mechanism, which would require that we provide a bit more information about the structure of some of the short-lived reaction intermediate.

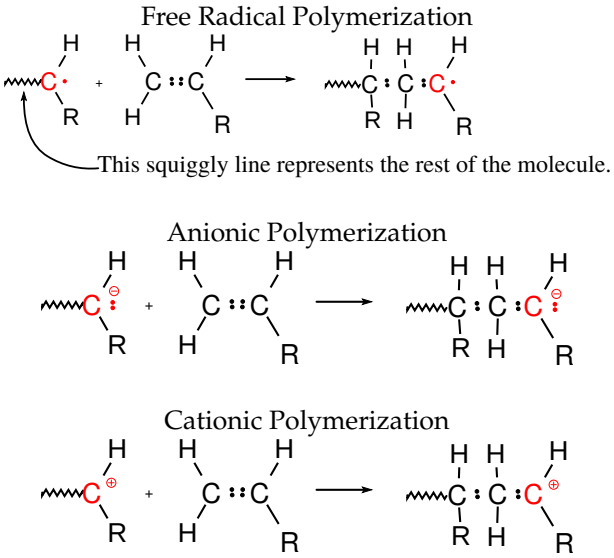


Figure 84.3: Propagation mechanisms for three common chain growth polymerization mechanisms.

84.3 Initiation

The previous section outlines the chemistry of some different propagation reactions, where monomers are added to an active site. Initiation is the step by which the active site is produced at the beginning. Here we give two simple examples, one for initiation of a radical polymerization, and the second being the initiation of a an anionic polymerization.

84.3.1 Initiation of a free radical polymerization

In this example, illustrated in Figure84.4, a peroxide bond between two oxygen atoms splits to form two free radicals. These radicals are able to add to the double bond in a vinyl monomer like styrene to initiate the polymerization reaction.

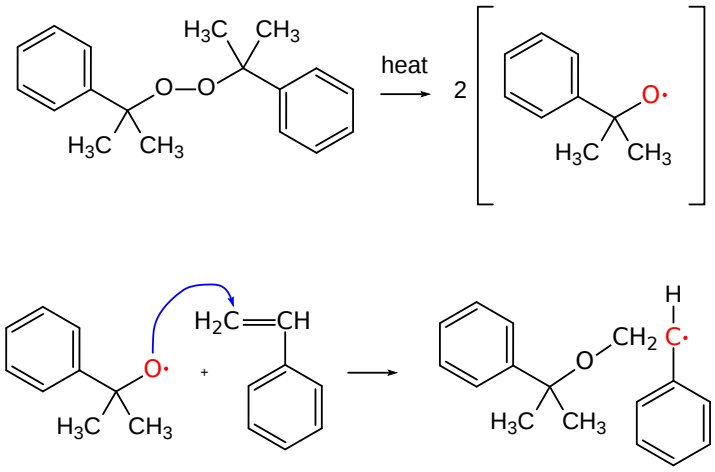


Figure 84.4: Initiation off a radical polymerization reaction with an epoxide that decomposes into two radical species at elevated temperature.

84.3.2 Initiation of an anionic polymerization

In the example shown in Figure84.5 is very simple conceptually. The reaction is done in solution, and a small amount of secondary butyl lithium is simply added to a solution of styrene molecule in an appropriate organic solvent. The carbon-lithium bond is very reactive, and has a lot of ionic character. It can really be viewed as an existing anion that is ready to react directly with the styrene monomer.

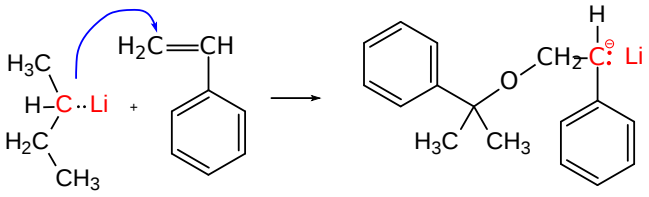


Figure 84.5: Initiation off a radical polymerization reaction with an epoxide that decomposes into two radical species at elevated temperature.

84.4 Termination Mechanisms

Termination is the process by which reactive chain ends become unreactive. Termination reactions can be avoided in 'living polymerizations' like anionic and cationic polymerizations.

84.4.1 Radical Recombination

Termination reactions cannot be eliminated in free radical polymerizations, since two free radicals can always combine directly to form a single bond, as illustrated below in Figure84.6.

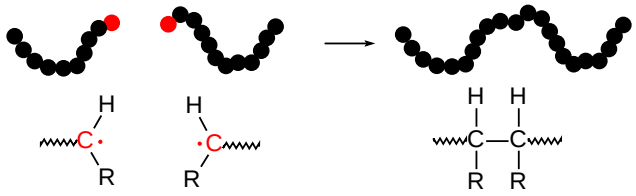


Figure 84.6: Termination by radical recombination.

As unpaired electrons, free radicals are very reactive towards one another. As illustrated in this example, two free radicals can readily combine with one another to form a covalent bond. Combination is therefore one type of termination reaction which is very prominent with free radical polymerizations. Note that the number of molecules decreases by one during the combination reaction.

84.4.2 Disproportionation

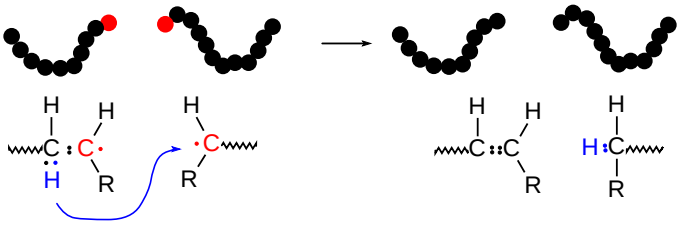


Figure 84.7: Termination by disproportionation. The circled proton, along with an electron, moves from the molecule on the right, terminating the free radical on the right. The remaining electrons on the molecule on the left rearrange themselves into a double bond.

Plain Laytionation reactions can be viewed as the transfer of a proton andEntanglementsfrom one active molecule to another. The animation here illustrates how the transfer results in the termination of both molecules,

with the formation of a double bond. In the final state, the carbon atoms all have eight electrons (including shared electrons) in the valence shell. The molecular weights of the two polymer molecules remain essentially unchanged (with the exception of the transfer of a single proton).

84.4.3 Chain Transfer

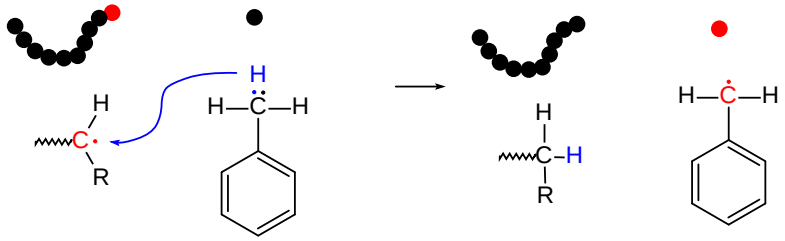


Figure 84.8: Chain transfer to toluene.

Chain transfer refers to the migration of the active free radical from one molecule to the other. In this example, the active radical from the growing polymer chain is terminated by the addition of a proton and an electron from the toluene molecule on the right. The net result is that the free radical is transferred from the polymer molecule to the solvent molecule (toluene), which can then initiate the polymerization of additional monomer. Note that the process is very similar to disproportionation, except that the species which donates the proton and electron does not already have a free radical.

84.5 Intramolecular Chain Transfer

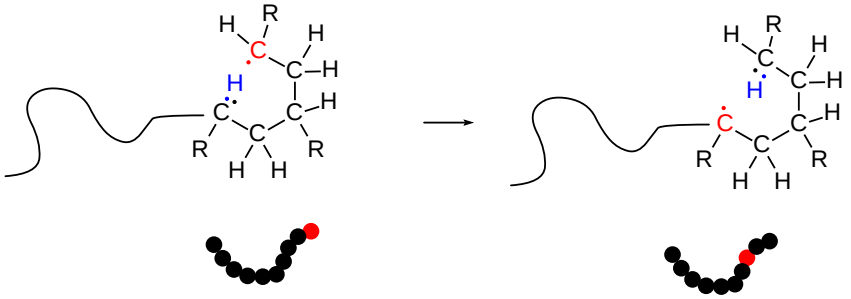


Figure 84.9: An Animation of Intramolecular Chain Transfer

In this example, the active radical moves from the end of the polymer chain to a different portion on the same polymer chain. Polymerization continues from this radical, resulting in the formation of a short branch, in this case consisting of four carbons. Because the branches are randomly placed along the polymer backbone, they interfere with the polymer's ability to crystallize.

Exercise:

(1) Chain transfer agents, like toluene in the previous example, can be intentionally added to give some control over the molecular weights of polymers synthesized by free radical polymerizations. In a qualitative sense, how will the addition of chain transfer agents change the number average molecular weight of the polymer?

(2) How does your answer to this question change if the chain transfer is always to other sites within the same polymer molecule?

(3) Termination reactions involving free radical polymerizations involve the elimination of two radicals to form a new bond. Where does this new bond appear for combination and disproportionation reactions?

84.6 Polymerization of Dienes

Dienes are monomers with two double bonds, with a single bond between them. Examples include butadiene, isoprene and chloroprene. The situation for isoprene is illustrated in Figure84.10. The monomer has 4 carbon atoms in a line, which we number 1 through 4. Double bonds connect carbons 1 to 2 and 3 to 4, and a single bond connects carbon 2 to 3. Carbons 2 and 3 are distinguishable from one another because carbon 2 is bonded to a methyl group that is not present on carbon 2. When a propagating polymer chain interacts with the isoprene monomer, addition can occur in any of 4 different ways, which are illustrated in Figure84.10:

- 1,2 addition: The active chain end adds across the bond between carbons 1 and 2, just as it would in a normal polymerization.
- 3,4 addition: Like 1,2 addition, but the reaction occurs across the bond between carbons 3 and 4.
- 1,4 addition: Here the reactive site attaches to carbon 1 and the active site moves to carbon 4, with a double bond being formed between carbons 2 and 3. This double bond can exist in a 'cis' conformation or 'trans' configurations. These cis and trans configurations represent chemically different structures, resulting in polymers with different properties.

Natural rubber is a naturally occurring version of polyisoprene that is harvested from certain tropical trees. The excellent elastomeric properties of this material arise from the pure cis 1-4 microstructure that is produced during the natural polymerization of this polymer. Because of the zigzag nature of the polymer backbone, it does not readily crystallize. A segment of the backbone of the polymer is shown in Figure84.11. This schematic shows a portion of the molecule in a fully extended form. Because of the easy rotation about single bonds, the molecules actually exist in a collection of random configurations.

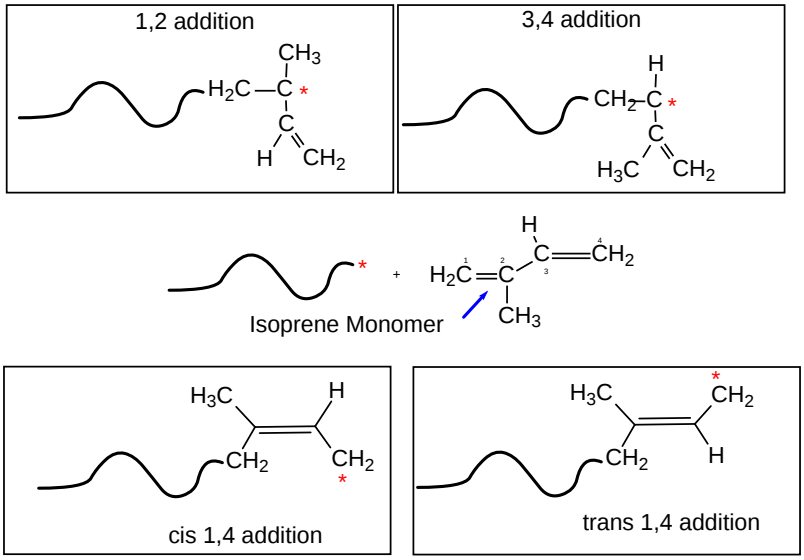


Figure 84.10: Polymerization of Isoprene

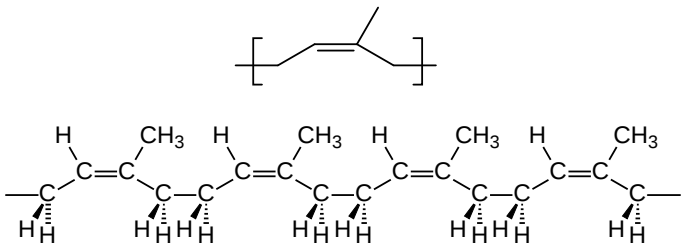


Figure 84.11: The chemical structure of natural rubber (cis 1-4 polyisoprene)..

84.7 Living Polymerizations

Living polymerizations are chain growth polymerizations that proceed without termination or chain transfer reactions. Relatively monodisperse polymers ($M_w/M_N < 1.1$) can typically be obtained when the initiation rate is faster than than the propagation rate. Block copolymers are formed by the sequential addition of two (or more) different monomer types in a living polymerization.

84.8 Vinyl Polymers and Tacticity

Vinyl polymers are an important class of polymers produced by chain growth polymerizations. They are produced from vinyl monomers with the following general structure:

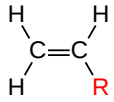


Figure 84.12: Chemical structure of a vinyl monomer.

When a vinyl monomer is polymerized and we stretch the resulting polymer chain out so that the backbone C-C bonds all lie on the same plane, the 'R' groups end up on different sides of the molecule. The distribution of these R groups determines thetacticityVogel equationle, according to the following definitions:

- Isotactic: All R groups on same side of the molecule.
- Syndiotactic: R groups on alternate sides of the molecule.
- Atactic R groups randomly placed.

These different tacticities are illustrated in Figure84.13.

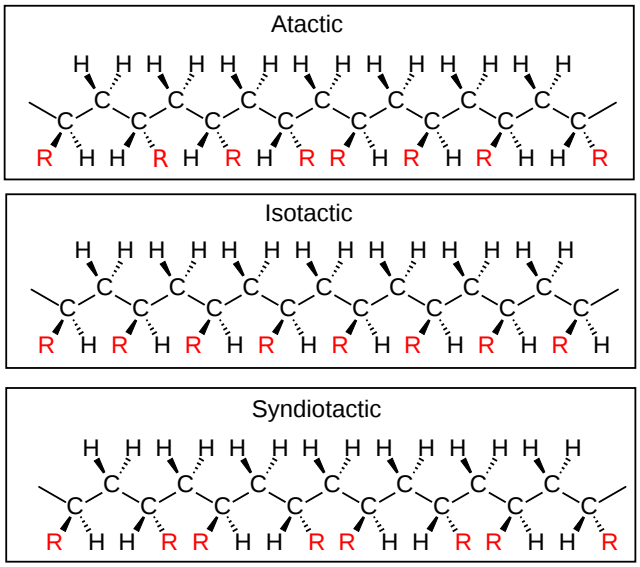


Figure 84.13: Illustration of different polymer tacticities.

Tacticity is significant because it determines the ability of a polymer to crystallize. The disordered structure of an atactic polymer is inconsistent with the ordered structure of a crystalline polymer. As a result, atactic polymers generally cannot crystallize.Standardons to this rule include polymers where the 'R' group is very small, so that this group can be incorporated into an ordered crystalline array, even if it is randomly placed along the polymer chain. For this reason polyvinyl chloride can be partially crystalline even if the polymer is atactic. Atactic polystyrene and atactic polypropylene, however, are always amorphous.

85 Common Polymers

Here we list some common polymer produced by the different synthesis methods introduced in the previous sections.

85.1 Chain Growth: Addition to a Double Bond

85.1.1 Polyethylene

The simplest polymer from a structural standpoint is polyethylene, with the structure shown below in Figure85.1.Dynamicity is not relevant in this case, since there are no substituents other than hydrogen on the carbon backbone.

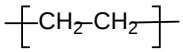


Figure 85.1: Chemical structure of polyethylene.

High Density Polyethylene: High density polyethylene refers to version with very little chain branching, thus resulting in a high degree of crystallinity. Completely linear polyethylene has a melting point of 138 °C, and a glass transition temperature near -100 °C.

$$T_g = \approx -120\,^{\circ}\text{C}, T_m = 138\text{C (perfectly linear)}$$

Low Density Polyethylene: Low density polyethylenes (LDPE's) and high density polyethylenes (HDPE's) are identical in their chemical structure at the atomic level. They are actually structural isomers of one another. Chain branching within low density polyethylene inhibits crystallization, resulting in a material with a melting point lower than 138 degrees C. The decreased crystallinity of LDPE results in a material which is more flexible (lower elastic modulus) than HDPE.

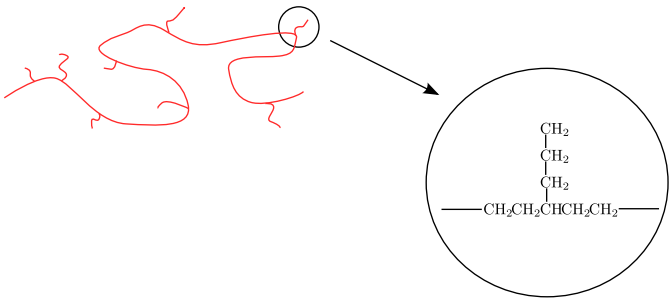


Figure 85.2: Structure of low density polyethylene

The chain branches responsible for inhibiting crystallization in low density polyethylene are typically short. This illustration shows a 3-carbon (propyl) branch, potentially resulting from intramolecular chain transfer during the polymerization reaction.

$$T_g = \approx -120\,^{\circ}\text{C}, T_m < 138\text{C (depending on branching)}$$

The Importance of Molecular Weight: Polymeric materials generally have favorable mechanical properties only when the molecular weight is very large - typically hundreds of thousands of g/mol. The point is illustrated with polyethylene:

- M = 16 g/mol: ethylene gas
- M≅200 g/mol: candle wax
- M≅ $2x10^5$-$5x10^5$ g/mol: milk jugs, etc.
- M≅ $3x10^6$ -$5x10^6$ g/mol: ultrahigh molecular weight polyethylene. This materials as excellent toughness and wear resistance, and is often used as one of the contact surfaces in joint implants.

85.1.2 Polypropylene

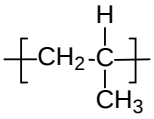


Figure 85.3: Structure of polypropylene

The most widely used form of polypropylene is isostatic, with a melting point of 183 °C, and a glass transition temperature which is well below room temperature. Single crystals of polypropylene have lower moduli than single crystals of polyethylene along the chain direction, because of the helical structure of propylene. The modulus of isotropic semicrystalline polypropylene is often larger than that of high density polyethylene, however, because of the details of the semicrystalline structure that is formed. The uses of polypropylene and high density polypropylene are similar.

85.1.3 Polypropylene

This polymer included to illustrate the evolution o the polymer properties when we continue to make the side chain longer.

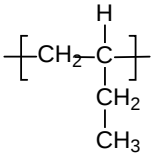
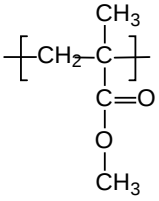


Figure 85.4: Polybutene-1

85.1.4 Poly(methyl methacrylate):

Poly(methyl methacrylate) (PMMA) is one of the most common materials used to make polymer glass. It is commonly known by the DuPont tradename Plexiglas,™ and has a glass transition temperature between 100 °C and 125 °C, depending on the tacticity. It is also forms the basis for many biomaterials, including dental adhesives.

$$T_g = 100 - 125\,^{\circ}\text{C, no} T_m \text{ (atactic)}$$



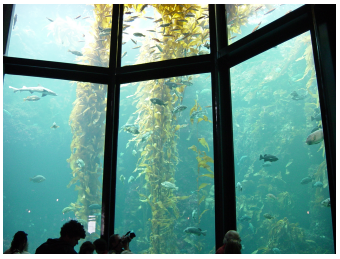


Figure 85.5: Poly(methyl methacrylate)

85.1.5 Poly(methyl acrylate)

Poly(methyl acrylate) is not a widely used polymer, primarily because it's glass transition temperature is too low (≈5°C for the atactic polymer) to be useful as a rigid polymer glass, and too high to be useful as an elastomer. It is included here to illustrate the effect that removing the extra methyl group from the polymer backbone has on the glass transition of the polymer.

$$T_g = 5\,^{\circ}\text{C, no} T_m \text{ (atactic)}$$

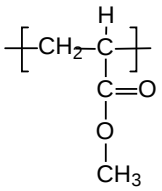


Figure 85.6: Poly(methyl acrylate).

85.1.6 Neoprene

Neoprene, also called polychloroprene, is a material commonly used in wetsuits. Like polyisoprene, it can be polymerized in different forms, corresponding to 1-2, 3-4, cis 1-4 and trans 1-4 addition of the monomer (trans 1-4 addition shown below). This pictured wetsuit has a 0.5mm layer of neoprene sandwiched between layers of nylon and another synthetic material.

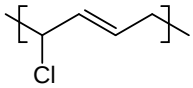


Figure 85.7: Neoprene (polychloroprene).

85.1.7 Polyisobutylene

Polyisobutylene is a common material used to make elastomers, referred to more simply as 'butyl' rubber. It is generally copolymerized by with a small amount of isoprene, so that the resulting double bonds can be used to form a crosslinked material. It is more resistant to solvent penetration than most elastomers, and is often used in applications (like the gloves above) where barrier resistance is needed.

$$T_g = -75\,^{\circ}\mathrm{C}$$

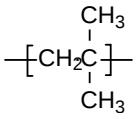


Figure 85.8: Polyisobutylene.

85.1.8 Polystyrene

Polystyrene is almost always atactic, and therefore amorphous. It has a glass transition temperature of 100 °C, and is therefore a glassy polymer at room temperature. Its uses are typically in packing material "Styrofoam", and for making cheap plastic objects, like the vials shown above. When suitably modified by the addition of other types of polymers, it is the basis for relatively high performance plastics such as high impact polystyrene (HIPS) and acrylonitrile-butadiene-styrene (ABS).

$$T_g = 100\,^{\circ}\mathrm{C}, \text{no} T_m \text{ (atactic)}$$

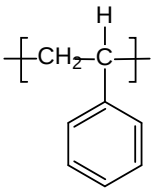


Figure 85.9: Structure of polystyrene.

85.1.9 Poly(tetrafluroethylene) (PTFE)

PTFE is more commonly known by its DuPont trade name, Teflon. Its most outstanding properties are its low surface energy and its low friction against a variety of other materials. It has very poor mechanical properties, however, and is difficult to process by its melting temperature excedes the temperature at which it begins to degrade.

$T_g = 130°C$ (by one report[28]);$T_m \approx 330\,°C$.

Figure 85.10: Monomer unit of poly(tetrafluroethylene) (Image from[5]).

85.1.10 Poly(vinyl acetate)

Poly(vinyl acetate) is often used as base for chewing gum. It is glassy at room temperature but becomes softer at body temperature, which is just aboveT_g.

$T_g = 30\,°C$, noT_m (atactic)

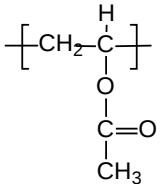


Figure 85.11: Monomer unit of poly(vinyl acetate)

85.1.11 Poly(vinyl chloride) PVC

Poly(vinyl chloride) can be partially crystalline even if the material is atactic, because the "R" group in this case is a chlorine atom, which is relatively small. The glass transition temperature of the material is 85 °C, although the addition of small molecules as "plasticizers" can reduceT_g to below room temperature. When a material is referred to as "vinyl", it is probably PVC. Record albums (before the age of compact disks) and water pipes are commonly made out of poly(vinyl chloride).

$T_g = 85\,°C$, noT_m (atactic)

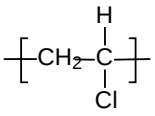


Figure 85.12: Structure of poly(vinyl chloride)

85.1.12 Poly(vinyl pyridine)

Poly(vinyl pyridine) is very similar to the polystyrene, and its physical properties (, entanglement molecular weight, etc.Processiry similar. It exists in one of two forms, poly(2-vinyl pyridine) (P2VP and poly(4-vinyl pyridine) (P4VP), based on the location of the nitrogen in the phenyl ring. Both types interact strongly with metals. The polymers are not used in wide quantities, but have been useful in a range of model studies of polymer behavior, often when incorporated with another material as part of a block copolymer.

$$T_g = 100\,^\circ\text{C}, \text{no} T_m \text{ (atactic)}$$

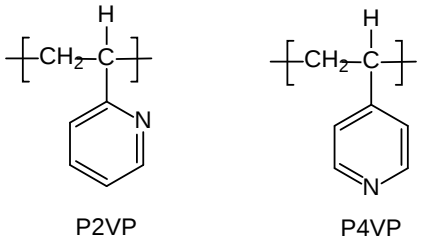


Figure 85.13: Structure of two types of poly(vinyl pyridine)

85.2 Chain Growth: Ring Opening

85.2.1 Poly(ethylene oxide)

Poly(ethylene oxide) (PEO) is generally formed by the ring opening polymerization of ethylene oxide. It is also referred to as polyethylene glycol, although this generally refers to lower molecular weight versions with hydroxyl end groups. PEO is water soluble, and is used in a wide range of biomedical applications, often in a gel form. Lithium salts are also soluble in PEO, and PEO/Li complexes are often used as an electrolyte in battery and fuel cell applications.

$$T_g = -65\,^\circ\text{C}; T_m \approx 65\,^\circ\text{C}$$

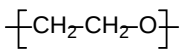


Figure 85.14: Monomer unit of poly(ethylene oxide)

85.2.2 Polycaprolactam:

Polycaprolactam is the polyamide equivalent of polycaprolactone, and is synthesized by the ring opening polymerization of the corresponding cyclic amide. It is often referred to as Nylon 6, since there are 6 carbon atoms in the repeating unit of the polymer. Note that this is different than Nylon 6,6 produced by condensation polymerization, where the repeat unit has 6 carbons originating from each of the two monomers used in the polymerization reaction.

$$T_g \approx 50\,^\circ\text{C}; T_m \approx 220\,^\circ\text{C}$$

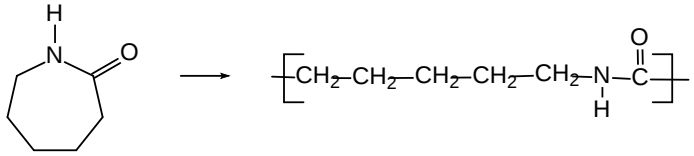


Figure 85.15: Polycaprolactam.

85.2.3 Polycaprolactone:

Polycaprolactone somewhat unique in that is a polyester that is synthesized by ring opening polymerization of a cyclic ester. It can be viewed as a polyester version of the polyamide, polycaprolactam. Contrary to step growth polymerization of polyesters, the ester linkage is not formed during the polymerization reaction, but is already present in the monomer. Polycaprolactone is biodegradable because the polymer slowly degrades by ester hydrolysis over time.

$$T_g \approx -50\,^{\circ}\text{C}; T_m = 60\,^{\circ}\text{C}$$

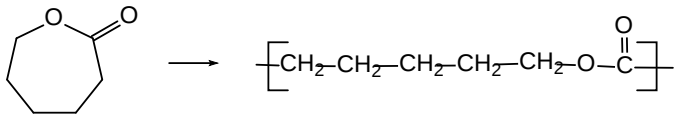


Figure 85.16: Polycaprolactone.

85.3 Step Growth Polymers

A variety of common polymers are discussed briefly in the pages below.

85.3.1 Kevlar™

Kevlar™ is a trademark of DuPont, Inc. The name actually is used to refer to a variety of aromatic polyamides, or aramids. As the name suggests, the polymers have phenyl groups in the backbone of the chain, and the repeat units are joined by amide linkages. The simplest possible aramid has the structure shown in Figure85.17.

T_m: above degradation temperature for the polymer.

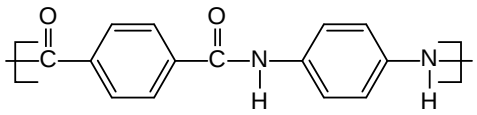


Figure 85.17: Monomer unit of one version of Kevlar

85.3.2 Polycarbonate

A variety of polycarbonates exist. The most common one, (with the GE trademark of Lexan) has a glass transition temperature of 150 °C. It is used for compact disks, eyeglass lenses, and shatterproof glass, like that featured here in the greenhouse. Seehttp://www.pslc.ws/macrog/pcsyn.htm for a good description of the synthesis of polycarbonate via a step growth, condensation reaction involving a phenolic di-alcohol and phosgene$(COCl_2)$.

$$T_g \approx 150\,^{\circ}\text{C}$$

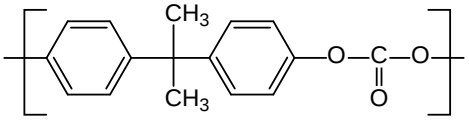


Figure 85.18: Polycarbonate.

85.3.3 Polyethylene Terephthalate (PET)

Polyethylene terephthalate (trade names include Mylar and Dacron) is produced in fiber form for textiles, and in film form for recyclable bottles, etc. Its degree of crystallinity is highly dependent on the the processing conditions, since it can easily be quenched to a glassy state before crystallization is able to occur.

$$T_g = 80\,^{\circ}\mathrm{C}$$

$$T_m = 260\,^{\circ}\mathrm{C}$$

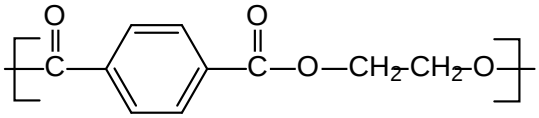


Figure 85.19: Polyethylene Terephthalate

85.3.4 Poly(phenylene oxide)

Polyphenylene oxide is a high performance polymer has many varied uses, largely because of its excellent performance at high temperatures.$T_g \approx 190\,^{\circ}\mathrm{C}$

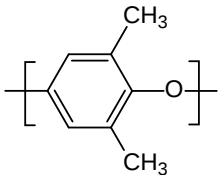
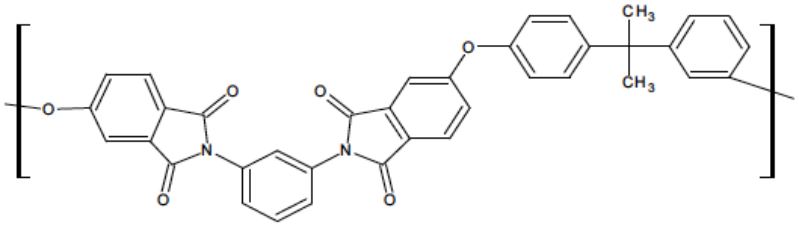


Figure 85.20: Poly(phenylene oxide).http://plastiquarian.com/ppo.htm

85.3.5 Ultem Polyetherimide

Ultem (a trademark of GE) is a form of polyetherimide. It is a high performance polymer that combines high strength and rigidity at elevated temperatures with long term heat resistance ($T_g = 215\,^{\circ}\mathrm{C}$). The repeat unit is illustrative of the complex chemical structure of many modern, high performance polymers.



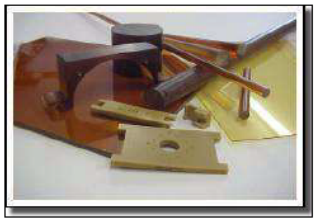


Figure 85.21: Structure of Ultem polyetherimide, with some materials that have been made from it.http://www.alcanairex.com/products/e/100/120p01_e.htm

85.3.6 Silicones

Silicones are an important class of synthetic polymers which do not have carbon in the backbone. Instead, the backbone consists of alternating silicon and oxygen atoms. Different classes of silicones are specified by the

substituents on the silicone atoms. Poly(dimethyl siloxane) (PDMS), with methyl substituents, is the most important silicone. Its glass transition and melting temperatures are very low, so that it remains flexible at very low temperatures. It also has a very low surface energy, and forms a hydrophobic surface that is very water resistant.

$$T_g =\sim -130^oC; T_m = -45^oC$$

Figure 85.22: Poly(dimethyl siloxane).

We list PDMS here as a step growth polymer because it can be produced from a self-condensation of silanol (SiOH) groups. The starting point is actually dimethyl-dichlorosilane. In th presence of water the SiCl bonds hydrolyze to SiOH:

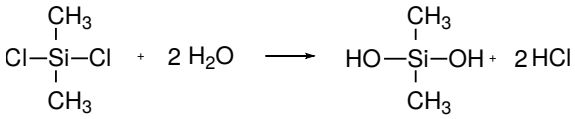


Figure 85.23: Hydrolysis of chlorosilane bonds.

The resultant silanol groups than can then condense by the elimination of water:

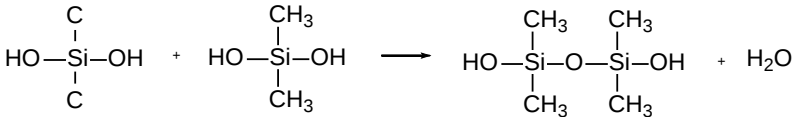


Figure 85.24: Condensation of silanol groups.

One of the interesting features of silicones is that they can also be synthesized by anionic, ring opening polymerization of cyclic, oligomeric forms of PDMS. Here's one example:

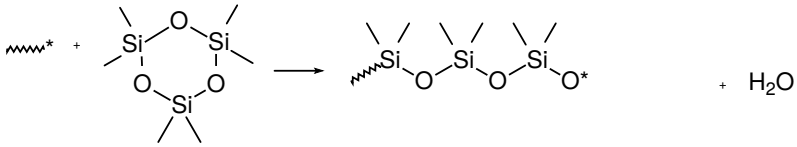


Figure 85.25: Anionic ring opening polymerization of a cyclic silicone oligomer.

86 Crosslinking

Crosslinking (or vulcanization) is the process by which separate molecules in an amorphous polymer are chemically attached to one another, as shown schematically in Figure86.1. The black lines indicate individual molecules, and the red dots represent crosslink points where these molecules are 'tied' together.

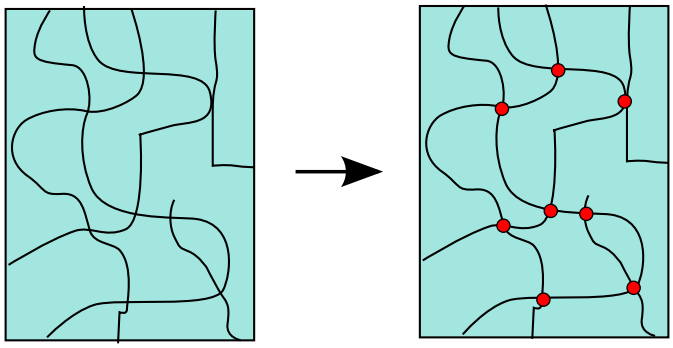


Figure 86.1: Diagram representation of crosslinking.

A representative and commercially important crosslinking system, known by the commercial name of Sylgard 184, consists of the two different parts shown in Figure86.2. Part 1 consists of poly(dimethyl silioxane) polymer with vinyl groups (double bonds) at each end of the polymer molecules. Part 2 contains a silicone crosslinker with several silane groups (Si-H) that are able to react with the vinyl groups as shown in Figure86.3. The reaction can be viewed as a step growth reaction involving a component with a reactive functionality of 2 (part 1) and a second component with a reactive functionality that is much larger than 2 (part 2).

Part 1:$n \approx 60$ Part 2:$n \approx 10$

Figure 86.2: A diagram of silicone crosslinking.

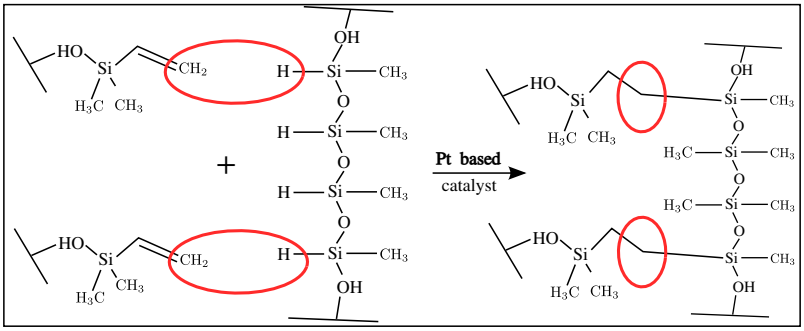


Figure 86.3: Reactionary diagram of silicone crosslinking.

87 Copolymers

The polymers considered so far are homopolymers, which means they consist of a single chemical repeat unit. Copolymers consist of more than one repeat unit. Examples includestatistical copolymers, where the different repeating units are distributed throughout a molecule, as shown schematically in Figure87.1.

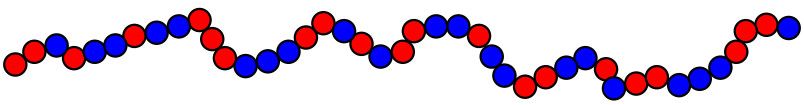


Figure 87.1: Schematic representation of a statistical copolymer.

Inblock copolymers, the different repeating units appear as distinct 'blocks'. A diblock copolymer for example, can be viewed as two different homopolymer molecules that have been covalently joined to one another as

shown in Figure87.2. A wide variety of interesting block copolymer morphologies have been observed. Diblock copolymers are most often synthesized by living, chain growth polymerizations.

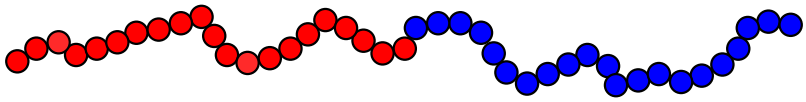


Figure 87.2: Schematic representation of a block copolymer.

Exercise: Describe a procedure for making a diblock copolymer of polystyrene and poly(methyl methacrylate).
Solution: The details of how to do this go a bit beyond the scope of this text, but conceptually we know we need to do the following in sequence: Add and appropriate initiator for a living chain growth polymerization, add monomer 'A' until all of it reacts, add monomer B until all of it reacts, terminate the reaction. Here are the details that can work for this system.

(1) Add an appropriate initiator. We don't want to use a radical initiator, because we need to avoid the inevitable termination reactions. From the information given in this text, either anionic or cationic polymerization. It turns out that the only anionic polymerization works in this case. We'll choose tetrahydrofuran as a solvent, clean it up very well to get rid of any impurities, and add some secondary butyllithium (see Fig.84.5) as the initiator. The molar quantity of initiator will be equal to the number of moles of polymer we get at the end of the reaction.

(2) Add the first monomer to the initiator solution. The order matters in this case. We need to add the styrene first, and choose reaction conditions so that the initiation is much faster than the subsequent polymerization. This ensures that the polymers all have roughly the same length, since they all started polymerizing a the same time.

(3) Once all the styrene reacts, we add the methyl methacyrylate monomer to the solution. Because the polymerization is 'living', the active chain ends from the polystyrene part of the polymerization are still available to react with the methyl methacrylate.

(4) Terminate the reaction. This is very easy to do, and can be done by adding a small amount of methanol to the reaction, to produce lithium methoxide and a 'dead' chain after the proton from alcohol trades places with the lithium that is associated with the negatively charged carbon at the end of the polymer chain.

88 The Glass Transition

The information presented so far about the structure of amorphous polymers applies both to elastomers like natural rubber, in addition to engineering thermoplastics like polystyrene. The properties of rigid plastics and rubber are obviously quite different, in spite of these similarities. The difference is that below a certain temperature, molecular motion is no longer possible, and the material becomes a relatively hard solid. This transition between "soft" and "rigid" behavior occurs at theglass transition temperature,T_g.

88.1 Free volume

A very useful way to think about the glass transition involves the concept of unoccupied space, or 'free volume' in a polymer. This concept is illustrated conceptually in Figure88.1. The red dots connected together by lines represent polymer molecules. In this lattice model of polymers, the red dots are constrained to fit on a lattice. (The lattice is completely artificial, and does not exist in reality). These dots represent the 'occupied volume'. Empty lattice sites representfree volume. The lattice on the left has no free volume, so there is no way for the molecules to move by the 'hopping' of segments into a small region of unoccupied space. The lattice on the right does have free volume, however, so that molecular motion is relatively easy.

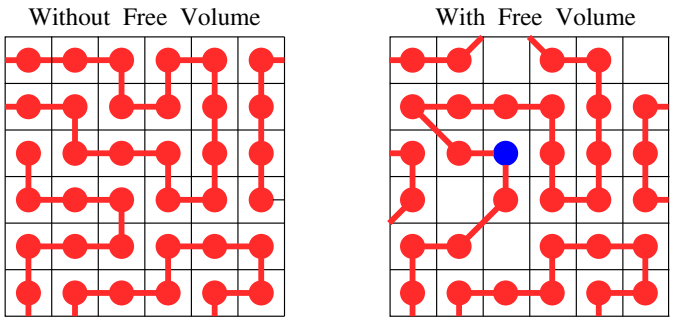


Figure 88.1: Conceptual representations of polymers with and without free volume.

The conceptual free volume model of the glass transition makes the assumption thatvolume thermal expansion coefficient,α is defined in the following manner:

$$\alpha = \frac{1}{V}\frac{dV}{dT} \tag{88.1}$$

Note that the linear thermal expansion coefficient, defined in terms of the linear dimensions of the sample rather than the volume, is equal to $\alpha/3$. If the fractional change in free volume is small ($\Delta V/V \ll 1$) thenα describes the linear relationship between the change in volume in the change in temperature:

$$\frac{\Delta V}{V} = \alpha \Delta T \tag{88.2}$$

In the glassy phase ($T < T_g$), the free volume is assumed to remain constant, so that increases in the volume are governed entirely by increases in the occupied volume:

$$\frac{\Delta V_{occ}}{V} = \alpha \Delta T \tag{88.3}$$

In the liquid phase ($T > T_g$), the free volume and occupied volumes both increase with temperature, and we have:

$$\frac{\Delta V}{V} = \frac{\Delta V_f}{V} + \frac{\Delta V_{occ}}{V} = \alpha_\ell \Delta T \tag{88.4}$$

Combining88.3 and gives the following expression for the free volume in the liquid state:

$$\frac{\Delta V_f}{V} = \left(\alpha_\ell - \alpha_g\right) \Delta T \tag{88.5}$$

Now we define a new temperature,T_∞ which is the theoretical temperature at which the free volume would go to zero, if the material were to continue to follow the liquid behavior belowT_g. It is illustrated schematically in Figure88.2. To determine the fractional free volume at the glass transition, we use Eq.88.5 with$\Delta T = T_g - T_\infty$, recognizing that$V_f/V = 0$ at$T = T_\infty$:

Typically,T_∞ is about 50K below the measured glass transition temperature, and the fractional free volume atT_g is in the range of a couple percent.

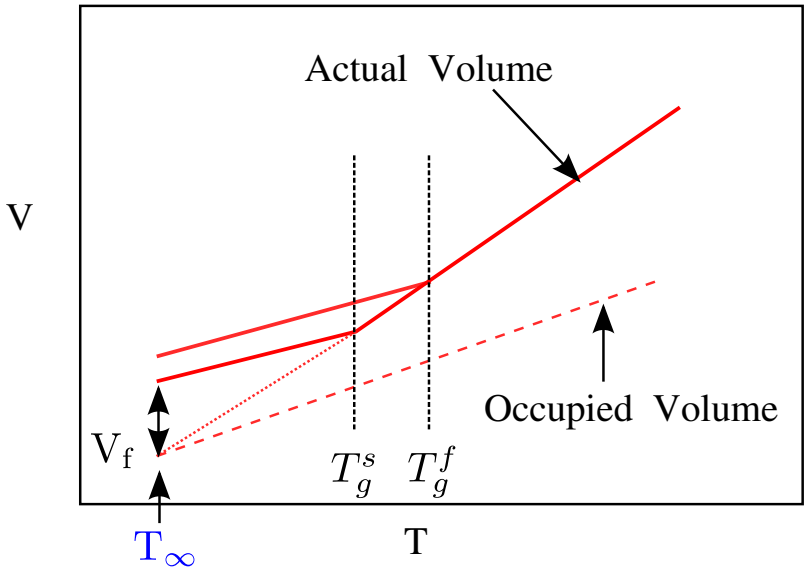


Figure 88.2: The glass transition and the free volume concept.

The concept of free volume can be helpful in sorting out how the glass transition depends on the structure of the polymer, based on the following two guidelines:

(1) Changes to the polymer structure that increase the free volume needed in and r for the backbone of the polymer to move will increaseT_g. Making the backbone very still (by the incorporation of phenyl groups, as inPET increaseT_g. Also, bulky, 'fat' substituents (like methyl or phenyl groups) tend to increaseT_g when added to the backbone.

(2) Changes to the polymer structure that introduce extra free volume will decreaseT_g. This is the case for long, thin additions to the backbone, such as alkyl chains$(-(CH_2)_n-)$.

Without experience it's difficult to know which of these effects is greater. If I add something to the backbone for example, does it create more free volume that it needs to move? The best way to get a handle on this is to take a look at the glass transition temperatures of atactic polymers with the general chemical structure shown in Figure88.3. The glass transition temperatures for these different polymers are listed in Table88.1.

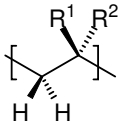


Figure 88.3: Structure of the repeat units for the materials listed in Table88.1.

	R^1=H	R^1=CH_3
R^2: H	-125	-13
R^2: CH_3	-13	-75
R^2:	100	20
R^2:	10	105
R^2:	-54	20

Table 88.1: Glass transition temperatures (in°C) for polymers with the general structure shown in Figure88.3 (From ref.[6])

Exercise: Polystyrene has a volume thermal expansion coefficient of$6 \times 10^{-4}\,\mathrm{K}^{-1}$ for temperatures aboveT_g, and a thermal expansion coefficient of$2 \times 10^{-4}\,\mathrm{K}^{-1}$ for temperatures belowT_g. Estimate the ratio of the free volume to the total volume of sample in the glassy state.StandardthatT_∞ is 50K belowT_g.
Solution: Since the fractional free volume is assumed to be fixed belowT_g equilibr are looking for is simply the value at the glass transition, as given by Eq.88.5. With$\alpha_\ell = 6 \times 10^{-4}\,\mathrm{K}^{-1}$,$a_g = 2 \times 10^{-4}\,\mathrm{K}^{-1}$ and$T_g - T_\infty = 50\,\mathrm{K}$, we obtain:

$$\frac{V_f}{V} = \left(4 \times 10^{-4}\,\mathrm{K}^{-1}\right)(50\,\mathrm{K}) = 0.02$$

This value of 0.02 is a typical value for the fractional free volume in a glassy polymer.

88.2 Enthalpy and Heat Capacity

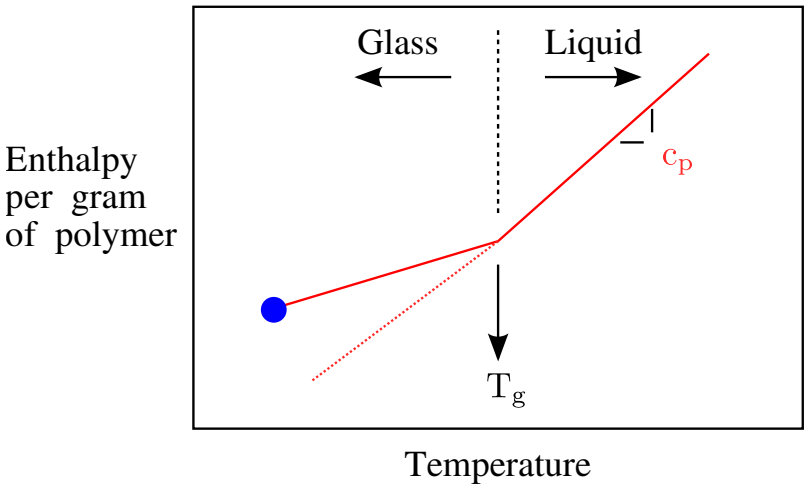


Figure 88.4: A graph of enthalpy vs temperature

The heat capacity (c_p) of the liquid is larger than the heat capacity of the glass. Note that the enthalpy itself is continuous at$T = T_g$. The glass transition is therefore a second order transition - thermodynamic quantities like volume, enthalpy and entropy are continuous at the transition but the derivatives of these quantities with respect to temperature are discontinuous.

88.3 Differential Scanning Calorimetry (DSC)

Glass transition temperatures are most commonly measured by differential scanning calorimetry. The technique can be used to measure the glass transition temperature of a polymer, in addition to the temperature at which a polymer melts or crystallizes. The heat of fusion (heat required to melt a semicrystalline polymer) can also be obtained. Heat flowing into the sample is endothermic, and heat flowing out of the sample is exothermic. We have plotted endothermic heat flow in the positive direction, and exothermic heat flow in the negative direction. Unfortunately, there is no universally followed sign convention for displaying DSC data. Sometimes endothermic heats are plotted in the negative direction. In order to avoid confusion, the endothermic direction is commonly indicated on the DSC plot, as we have done in the curve shown in Figure88.6. Students are encouraged to look at the Macrogalleria web sites on the glass transition[4] and differential scanning calorimetry[3] for discussions at the appropriate level for this class.

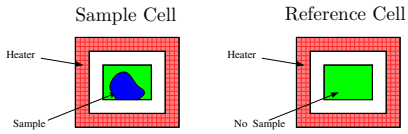


Figure 88.5: Simple schematic representation of a calorimeter.

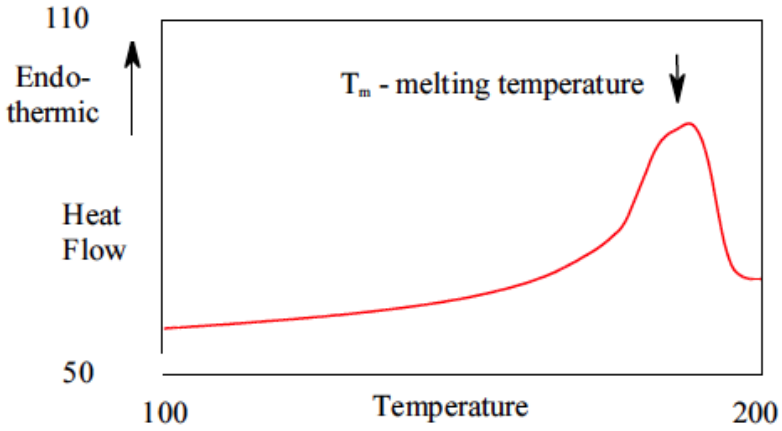


Figure 88.6: Graph of heat flow vs temperature.

89 Structure of Amorphous Polymers

Materials science is the study and application of the relationships between the properties of a material, the structure required to obtain these properties, and the processing methods which can be used to obtain these properties. Our discussion of polymer synthesis was motivated primarily by a need to understand processing methods for polymer materials. We also introduced the structure of different polymers at the atomic level, corresponding to the arrangements of individual atoms in the polymeric repeat units. Our discussion now moves to the structure of polymers on a molecular scale. We begin with amorphous polymers, and follow with a discussion of semicrystalline polymers. Ultimately, we will find that our understanding of many of the important properties of polymers can be related to these structural features by simple, yet remarkably accurate theories. In particular, the theory of rubber elasticity relates the elastic properties of an elastomer to the molecular structure described here.

The squiggly lines in Figure89.1 represent the backbones of individual molecules in an amorphous polymer. (One molecule is highlighted in red) Our goal is to relate the properties of an amorphous polymer to the distribution of shapes of these molecules. Ultimately, we will show that many of the mechanical properties of amorphous polymers at temperatures above the glass transition are related to this distribution of polymer shapes.

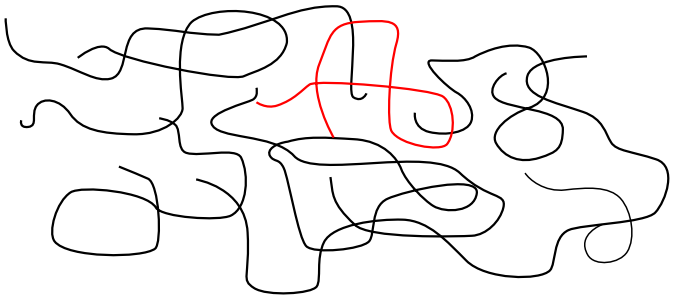


Figure 89.1: A diagram of an amorphous polymer.

89.1 End-to-end Vector for Polymer Molecules

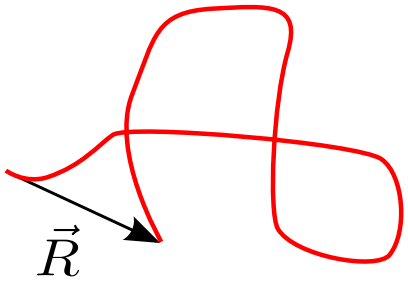


Figure 89.2: The end-to-end vector,$\vec{R}$, for a polymer molecule.

The shape of this particular molecule can only be completely specified by describing the path taken by the molecule from one end to the other. Fortunately, the most important quantity is much simpler. This quantity is theend-to-end vector,$\vec{R}$ which is simply the vector spanning the two ends of a given polymer molecule. This vector will in general be different for different polymer molecules in a sample, but the distribution of vectors can be accurately predicted. The shapes of amorphous polymer molecules are random, and random walk statistics can be used to describe the distributions of end-to-end vectors that are obtained.

89.2 Random Walk in One Dimension

We can understand much of what we need to understand by considering the simplest case of a random walk in 1 dimension. In this case our 'walker' takesN_x steps, each of which is a randomly chosen move to either the left or right. At the end of these$N_x$steps we write down how many steps removed from the starting point our walker is. We refer to this number as$i$. For example, if the walker ends up 4 steps to the right of where he ended up, then$i = 4$. Where our walker ended up after one particular trip is not very important though. We need to have him take lots of trips, and collect statistics to see where he ended up after taking a large number of trips. We'll let him take a total of$M$ trips, and write down the number of times,n, that the walker has ended up$i$ steps from where we started. If the walk is really random, then the relationship between$n$ and$i$ will be approximated by Eq.89.1

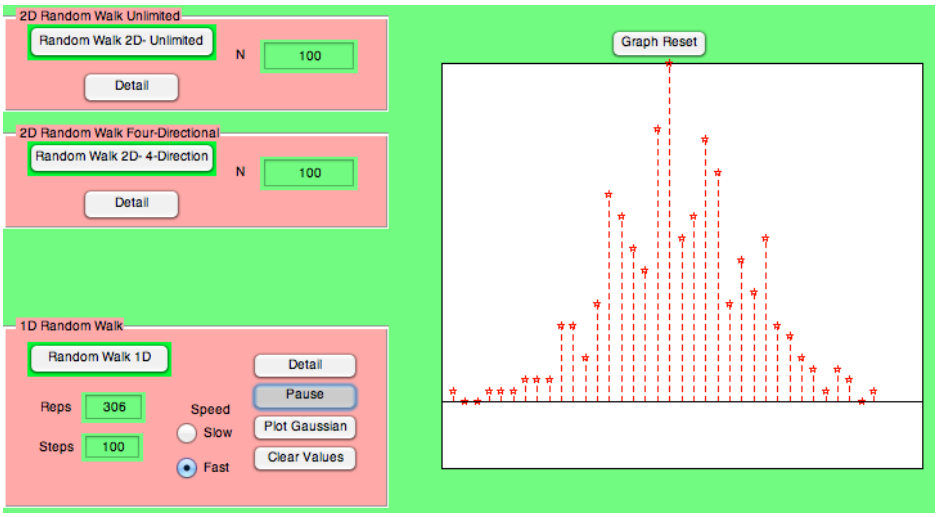


Figure 89.3: Random walk animation

$$n(i) = 2M\left(\frac{1}{2\pi N_x}\right)^{1/2}\exp\left(\frac{-i^2}{2N_x}\right) \tag{89.1}$$

We can check that the summation is appropriately normalized. In other words, is the following expression valid?

$$\sum_i n_i = M \tag{89.2}$$

The summation here is over all possible values of$i$. Note that for a random walk with an even number of steps, only even values of$i$ are possible. Similarly, for a walk with an odd number of steps, only odd values of$i$ are possible. We can get the summation we want by summing over all integer values of$i$ and then dividing by two:

$$\sum_i n(i) = \frac{M}{\sqrt{2\pi N_x}} \sum_{i=-\infty}^{\infty} \exp\left(\frac{-i^2}{2N_x}\right) \approx \frac{M}{\sqrt{2\pi N_x}} \int_{-\infty}^{\infty} \exp\left(\frac{-i^2}{2N_x}\right) di \tag{89.3}$$

Functions of the form$f(x) \propto \exp -(x/A)^2$ (where$A$ is a constant) are called Gaussian functions, and generally describe random processes. The following integrals of a Gaussian function will be very useful for us:

$$\begin{aligned} \int_{-\infty}^{\infty} \exp -(x/A)^2 \, dx &= A\sqrt{\pi} \quad \text{(a)} \\ \int_{-\infty}^{\infty} x \exp -(x/A)^2 \, dx &= 0 \quad \text{(b)} \\ \int_{-\infty}^{\infty} x^2 \exp -(x/A)^2 \, dx &= A^3\sqrt{\pi}/2 \quad \text{(c)} \end{aligned} \tag{89.4}$$

In our case,$A = \sqrt{2N_x}$, so$\sum_i n(i) = M$ and our expression for the Gaussian distribution of random walks is indeed normalized.

89.3 Average of a Function

The entire distribution represented by the Gaussian function will be useful to us, but it is still useful to have some averages. We encountered averages already in our discussion of the number and weight average distribution. In general, if$Pr(x)$ is the probability that a function has a value of$x$, than the average value of$x$ (referred to as$\langle x \rangle$) is obtained from the following expression:

$$\langle x \rangle = \sum_x x Pr(x) \tag{89.5}$$

where the sum is over all possible values of$x$. Similarly, the average ofx^2 is given by:

$$\left\langle x^2 \right\rangle = \sum_x x^2 Pr(x)\, dx \tag{89.6}$$

The procedure can be generalized to calculate the average of any function of$x$:

$$\langle f(x) \rangle = \sum_x f(x) \, Pr(x) \tag{89.7}$$

Suppose I put the following 10 numbers into a drawer: 1, 4, 8, 12, 19, 25, 28, 33, 37, 45 I randomly pick a number from the drawer and then return it. What is the average value of all the numbers that I pick if I continue with this exercise?

For a function of i, where i is only able to take on discrete values, we have:

$$< f(i) >= \sum_i f(i) Pr(i)$$

What if we have a function of a continuous variable, x? In this case, we must replace $Pr(i)$, with $Pr\,(x)\,dx$, where $Pr(x)dx$ is the probability that the continuous variable has a value between x and $x + dx$. We must also replace the summation with the appropriate integral to obtain the following result:

$$\langle f\,(x)\rangle = \int f\,(x)\,Pr\,(x)\,dx \tag{89.8}$$

Calculate the average value of x^2, assuming that x is equally likely to take on all values between 1 and 10, and that no values outside this range are possible.

89.4 Averages for Random Walks

If the length of each step taken by the random walker is a, then the distance between the beginning and the end of the walk is simply i times this step length:

$$R_x = ia$$

We can substitute R_x/a for i and dR_x/a for di in Eq. 89.3 to get an expression for the total number of walks M. We'll generalize a bit further and replace the limits of $-\infty$ and ∞ for i to limits of R_{min} and R_{max} for R_x. We obtain the following for $M\,(R_{min},\,R_{max})$, the total number of walks with values of R_x between R_{min} and R_{max}:

$$M\,(R_{min},\,R_{max}) = M \int_{R_{min}}^{R_{max}} Pr\,(R_x)\,dR_x \tag{89.9}$$

with the probability density function, $Pr\,(R_x)$ given as follows:

$$Pr\,(R_x) = \frac{1}{\sqrt{2\pi N_x a^2}} \exp\left(\frac{-R_x^2}{2N_x a^2}\right) \tag{89.10}$$

For simplicity, it is helpful to define $A = \sqrt{2N_x a^2}$ so that we can write $P\,(R_x)$ in the following way:

$$Pr\,(R_x) = \frac{1}{\sqrt{\pi} A} \exp - (R_x/A)^2 \tag{89.11}$$

To calculate the average values of R_x and R_x^2, referred to as $\langle R_x\rangle$ and $\left\langle R_x^2\right\rangle$, respectively, we use Eq. 89.8 to calculate the averages:

$$\langle R_x\rangle = \int_{-\infty}^{\infty} R_x Pr(R_x) dR_x = \frac{1}{\sqrt{\pi} A} \int_{-\infty}^{\infty} R_x \exp - (R_x/A)^2 \tag{89.12}$$

$$\left\langle R_x^2\right\rangle = \int_{-\infty}^{\infty} R_x^2 Pr(R_x) dR_x == \frac{1}{\sqrt{\pi} A} \int_{-\infty}^{\infty} R_x^2 \exp - (R_x/A)^2 \tag{89.13}$$

Finally, we use the standard integrals to

$$\langle R_x \rangle = 0 \tag{89.14}$$

$$\left\langle R_x^2 \right\rangle = \frac{A^2}{2} = N_x a^2 \tag{89.15}$$

We know that the first result has to be true because the Gaussian distribution function is symmetric: positive and negative values of R_x with equal magnitudes are equally likely. This means that any contribution to $\langle R_x \rangle$ will be exactly offset by a contribution with the opposite sign. For this reason, the mean-square end-to-end distance, $\langle R_x^2 \rangle$, or the root mean square end-to-end distance, $\left\langle R_x^2 \right\rangle$, is typically used to describe the spatial extent of a random walk.

89.5 Random Walks in 3 Dimensions

Now we can extend some of these results to describe the shapes of real polymer molecules. We begin with a discussion of two dimensional random walks. The simplest extension is to give our random walker 4 choices for his movements. In addition to moving to the left and right, he can now move up and down, as illustrated in Figure89.4a. If we move the constraint that there are only 2 directions in which the walker can move, we get the example shown in Figure89.4b. These trajectories are now beginning to look like the shapes of actual polymer molecules, which is why this whole exercise is relevant.

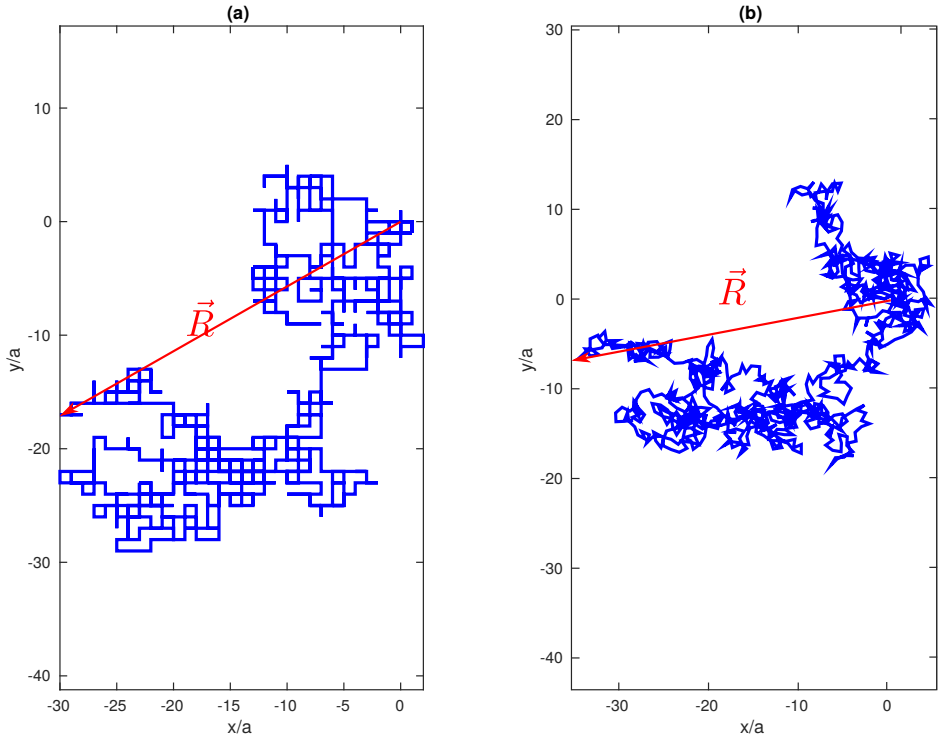


Figure 89.4: 2-D Random Walk of 1000 steps on a square lattice (a) and with random bond angles (b). The end-to-end vector, $\vec{R}$, is shown in each case.

Much of our understanding of amorphous polymers is based on our treatment of individual molecules as random walks in three dimensions. We assume that we can treat a polymer of degree of polymerization N as a random walk of N steps, each with an effective length of a, which we refer to as the Statistical segment length. The

statistical segment length does not correspond to a real polymer dimension that we can get directly by looking at the structure of the polymer molecule. We need more detailed models of chain dimensions for that. For now, you just need to know that it is a length that gives the right overall molecular size. To simplify our treatment and make the connection back to the statistics that we already developed for one-dimensional random walks, we make the following simplifying assumptions:

- We will get the overall statistics right by assuming a simple lattice model (as in Figure89.4), where the a chain is able to movie in any of the x, y or z directions.
- Of the$N$ steps made by the chain,$N/3$ of them move in each of the x, y and z directions.

The value ofR_x that we studied in the one dimensional random walk case is just the x component of the end-to-end vector, which also has y and z components:

$$\vec{R} = R_x\hat{x} + R_y\hat{y} + R_z\hat{z} \tag{89.16}$$

where$\hat{x}$,$\hat{y}$ and$\hat{z}$ are unit vectors in the x, y and z directions, respectively. The probability that a given random walk of N steps in three dimensions has an end-to end vector of with x, y and z compenents ofR_x,R_y andR_z is given by multiplying the probabilities for the individual components:

$$Pr(R_x,\, R_y,\, R_z) = Pr(R_x)Pr(R_y)Pr(R_z) = (2\pi Na^2/3)^{-3/2}\exp\left(\frac{-3R^2}{2Na^2}\right) \tag{89.17}$$

where we have used the following relationships:

$$R^2 = R_x^2 + R_y^2 + R_z^2 \tag{89.18}$$

and

$$N_x = N_y = N_z = N/3 \tag{89.19}$$

This relationship can be further simplified by recalling that$Pr(R_x, R_y, R_z)dR_xdR_ydR_z$ is the probability that x component of the end-to-end vector is betweenR_x and$R_x + dx$, the y component is betweenR_y and$R_y + dy$, and the z component is betweenR_z and$R_z + dz$. Because the probability only depends on$R$ for an isotropic system (as must be the case for a truly random walk), we can substitute$4\pi R^2dR$ for$dR_xdR_ydR_z$ to obtain:

$$Pr(R) = \frac{3(6)^{1/2}R^2}{\pi^{1/2}(Na^2)^{3/2}}\exp\left(\frac{-3R^2}{2Na^2}\right) = \frac{4.15R^2}{R_0^3}\exp\left(\frac{-3R^2}{2R_0^2}\right) \tag{89.20}$$

where$Pr\,(R)\,dR$ is the probability that the magnitude of the end-to-end vector lies between$R$ and$R + dR$, and we have used the following definition ofR_0:

$$R_0 = \sqrt{N}a \tag{89.21}$$

Note thatR_0^2 is the average ofR^2, obtained from the following version of Eq.89.8:

$$\left\langle R^2\right\rangle = \int_0^\infty R^2 Pr\,(R)\,dR = \frac{4.15}{R_0^3}\int_0^\infty R^4\exp\left(\frac{-3R^2}{2R_0^2}\right) = R_0^2 \tag{89.22}$$

We can define a dimensionless probability density$\bar{P}$, which gives the probability that the normalized quantityR/R_0 lies in some range:

$$\bar{P}r\left(\bar{R}\right) = 4.15\left(\bar{R}^2\right)\exp\left(-1.5\bar{R}^2\right) \tag{89.23}$$

Because$\bar{R}$ is a dimensionless ratio, the probability density,$\bar{P}r\left(\bar{R}\right)$, is also dimensionless. The shape of this function is shown in Figure89.5.

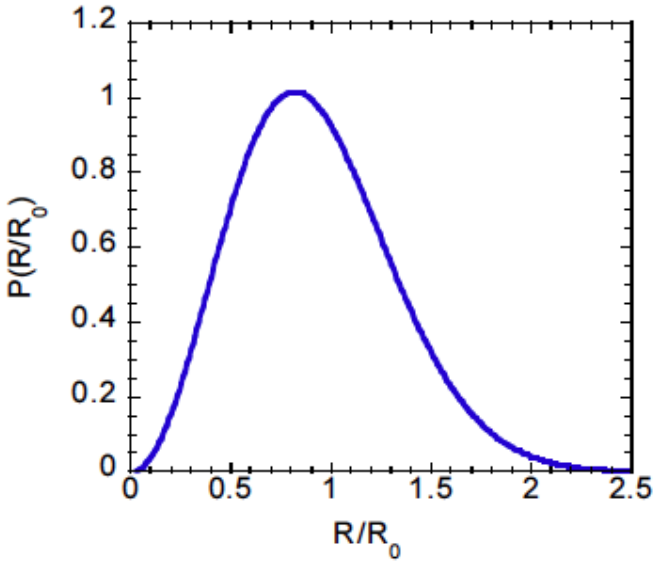


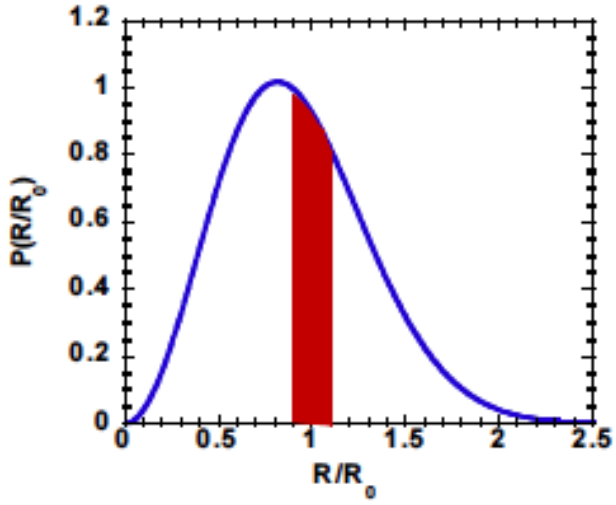
Figure 89.5: Probability density.

Note that the distribution is normalized:$\int_0^\infty \bar{P}rd\bar{R} = 1$.

Exercise: What fraction of molecules in an amorphous polymer have end to end vectors that are within 10% of R_0?
Solution: We need to integrate the probability density from R/R_0=0.9 to R/R_0 = 1.1. With x=R/R_0 we have:

$$Pr_{tot} = 4.15 \int_{0.9}^{1.1} x^2 \exp(-1.5x^2) dx$$

This corresponds to the shaded area in the following figure:



Numerical integration of the equation above gives Pr_{tot} = 0.18, which corresponds to the red shaded area. This is the total fraction of molecules with R between $0.9R_0$ and $1.1R_0$.

90 Models of Chain Dimensions

Our discussion of random walk statistics tells us what the shapes of the molecules look like when we know the statistical segment length, but it doesn't tell us what the statistical segment length actually is for a particular polymer. Clearly the statistical segment length must be related to the detailed chemical structure of the repeat unit itself. For example, we wouldn't expect that polyvinyl chloride and polyethylene terephthalate would have the same statistical segment length, based on the very different chemical structures of these polymers. To obtain a prediction for the statistical segment length we need to zoom in and consider the actual backbone bonds that are connecting the polymer together. In general every bond is vector, $\vec{r}$ with a length and an orientation. As illustrated in Figure90.1, the end-to-end vector is obtained by adding up all N_b bond vectors (N_b=12 in this example):

$$\vec{R} = \sum_{i=1}^{N_b} \vec{r}_i = \vec{r}_1 + \vec{r}_2 + \vec{r}_3 + ... + \vec{r}_{N_b} \tag{90.1}$$

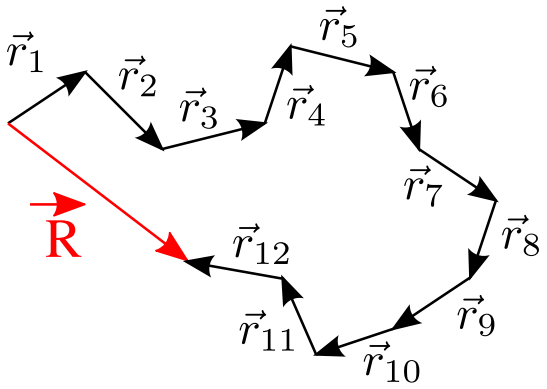


Figure 90.1: Vector diagram of a polymer.

90.1 General Considerations

We are interested in the average magnitude of the end-to-end vector. As discussed in the previous section, the average of$\vec{R}$ itself is not useful, since it just averages to zero. We are interested in the average value of the square of the magnitude of the end-to end vector. This quantity is equal to the dot product of the end-to-end vector with itself:

$$R^2 = \vec{R} \cdot \vec{R} = \left(\vec{r}_1 + \vec{r}_2 + \vec{r}_3 + ... + \vec{r}_{N_b}\right) \cdot \left(\vec{r}_1 + \vec{r}_2 + \vec{r}_3 + ... + \vec{r}_{N_b}\right) \tag{90.2}$$

The dot product hasN_b^2 terms, which can be represented in the following matrix form:

$$R^2 = \left\{ \begin{array}{ccccc} \vec{r}_1 \cdot \vec{r}_1 & +\vec{r}_1 \cdot \vec{r}_2 & +\vec{r}_1 \cdot \vec{r}_3 & +... & +\vec{r}_1 \cdot \vec{r}_{N_b} \\ +\vec{r}_2 \cdot \vec{r}_1 & +\vec{r}_2 \cdot \vec{r}_2 & +\vec{r}_2 \cdot \vec{r}_3 & +... & +\vec{r}_2 \cdot \vec{r}_{N_b} \\ +\vec{r}_3 \cdot \vec{r}_1 & +\vec{r}_3 \cdot \vec{r}_2 & +\vec{r}_3 \cdot \vec{r}_3 & +... & +\vec{r}_3 \cdot \vec{r}_{N_b} \\ +... & +... & +... & +... & +... \\ +\vec{r}_{N_b} \cdot \vec{r}_1 & +\vec{r}_{N_b} \cdot \vec{r}_2 & +\vec{r}_{N_b} \cdot \vec{r}_3 & +... & +\vec{r}_{N_b} \cdot \vec{r}_{N_b} \end{array} \right\} \tag{90.3}$$

The dot product between two vectors,$\vec{A}$ and$\vec{B}$ is given in general by$\left|\vec{A}\right| \left|\vec{B}\right| \cos\theta$ whereθ is the angle between the two vectors, illustrated below in Figure90.2. At this point we will make our first simplifying assumption, which is that the lengths of all bonds along the backbone are identical. This assumption is certainly valid for vinyl polymers, and other polymers which have only C-C single bonds along the backbone.

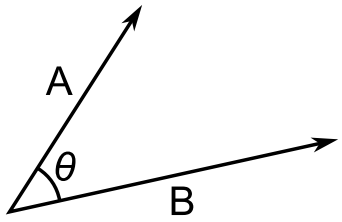


Figure 90.2: Definition of the angle between two vectors.

If this bond length isℓ, thenR^2 can be represented as follows, whereθ_{ij} is the angle between bond$i$ and bond$j$:

$$R^2 = \ell^2 \left\{ \begin{array}{ccccc} \cos\theta_{11} & +\cos\theta_{12} & +\cos\theta_{13} & +... & +\cos\theta_{1N_b} \\ +\cos\theta_{21} & +\cos\theta_{22} & +\cos\theta_{23} & +... & +\cos\theta_{2N_b} \\ +\cos\theta_{31} & +\cos\theta_{32} & +\cos\theta_{33} & +... & +\cos\theta_{3N_b} \\ +... & +... & +... & +... & +... \\ +\cos\theta_{N_b1} & +\cos\theta_{N_b2} & +\cos\theta_{N_b3} & +... & +\cos\theta_{N_bN_b} \end{array} \right\} \tag{90.4}$$

The average value of R, will be determined by the average values of ϑ_{ij} In mathematical terms, we have the following expression:

$$\left\langle R^2 \right\rangle = \ell^2 \left\{ \begin{array}{ccccc} \langle\cos\theta_{11}\rangle & +\langle\cos\theta_{12}\rangle & +\langle\cos\theta_{13}\rangle & +... & +\left\langle\cos\theta_{1N_b}\right\rangle \\ +\langle\cos\theta_{21}\rangle & +\langle\cos\theta_{22}\rangle & +\langle\cos\theta_{23}\rangle & +... & +\left\langle\cos\theta_{2N_b}\right\rangle \\ +\langle\cos\theta_{31}\rangle & +\langle\cos\theta_{32}\rangle & +\langle\cos\theta_{33}\rangle & +... & +\left\langle\cos\theta_{3N_b}\right\rangle \\ +... & +... & +... & +... & +... \\ +\langle\cos\theta_{N_b1}\rangle & +\langle\cos\theta_{N_b2}\rangle & +\langle\cos\theta_{N_b3}\rangle & +... & +\langle\cos\theta_{N_bN_b}\rangle \end{array} \right\} \tag{90.5}$$

In more compact notation we can write:

$$\left\langle R^2 \right\rangle = \ell^2 \sum_{i,j=1}^{N_b} \langle\cos\theta_{ij}\rangle \tag{90.6}$$

So we have reduced the problem to figuring out what the values of$\langle\cos\theta_{ij}\rangle$ are in the system. This requires a specific model of how the bonds are joined to one another. In the following sections we consider two specific models (the freely jointed and freely rotating models), and a general model where the bond rotation angles are constrained.

90.2 Freely Jointed Chain Model

The simplest (and most unrealistic) model of chain dimensions is the freely jointed chain model, where all bond orientations are equally likely. In the freely jointed chain model$\langle\cos\theta_{ij}\rangle = 0$ for$i \neq j$, because orientations giving positive and negative values of$\cos\theta_{ij}$ are equally likely. The only exception is for theN_b diagonal terms for which$i = j$. These represent the projection of a vector onto itself, and all have$\theta = 0$ and$\cos\theta = 1$.

$$\left\langle R^2 \right\rangle = N_b\ell^2 \tag{90.7}$$

90.3 Freely Rotating Chain Model

In this model, bond angles are fixed, but rotation about bonds is possible. Rotation about a bond sweeps out a cone as shown below. The bond rotation angle is defined asϕ, and in the freely rotating model all values ofϕ are assumed to be equally likely.

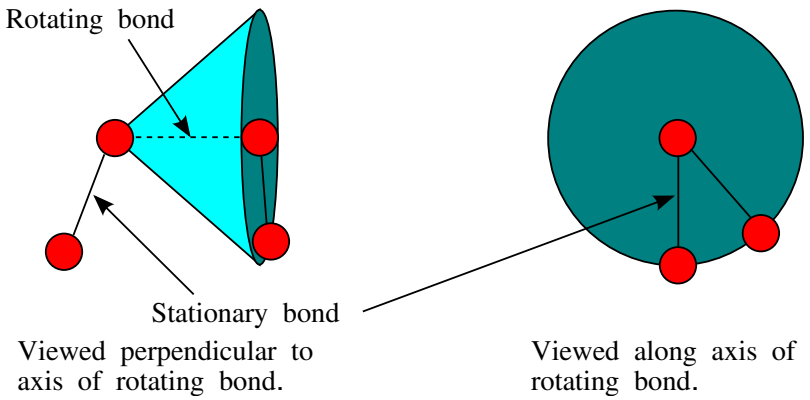


Figure 90.3: Freely Rotating Chain Models with views perpendicular to the axis of the rotating bond and along the axis of the rotating bond

In the freely rotating chain model, the angle between adjacent bonds is fixed at certain angle, referred to here simply asθ (without the subscripts). Adjacent bonds,*i.e.* those with$i - j = 1$ or$i - j = -1$, have$\langle\cos\theta_{ij}\rangle = \cos\theta$. It

can be shown that terms with $i-j$=2 or -2 have $\langle \cos\theta_{ij} \rangle = \cos^2\theta$, terms with $i-j$=3 or -3 have $\langle \cos\theta_{ij} \rangle = \cos\theta$, etc. The matrix of terms making up the summation to give $\langle R^2 \rangle$ therefore has the following form:

$$\left\langle R^2 \right\rangle = \ell^2 \left\{ \begin{array}{ccccc} 1 & +\cos\theta & +\cos^2\theta & +... & +\cos^{(N_b-1)}\theta \\ +\cos\theta & +1 & +\cos\theta & +... & +\cos^{(N_b-2)}\theta \\ +\cos^2\theta & +\cos\theta & +1 & +... & +\cos^{(N_b-3)}\theta \\ +... & +... & +... & +... & +... \\ +\cos^{(N_b-1)}\theta & +\cos^{(N_b-2)}\theta & +\cos^{(N_b-3)}\theta & +... & +1 \end{array} \right\} \tag{90.8}$$

Of the N_b^2 terms in the summation, N_b terms (the diagonal terms with $i=j$) are equal to 1. There are N_b-1 terms with $i-j$=1, and an additional N_b-1 terms with $i-j$=-1. There are therefore 2(N_b-1) terms with a magnitude equal to $\cos\theta$. Similarly, there are 2(N_b -2) terms equal to $\cos^2\theta$, 2(N_b -3) terms equal to $\cos^3\theta$, etc. The expression for $\langle R^2 \rangle$ can therefore be written as follows:

$$\left\langle R^2 \right\rangle = \ell^2 \left\{ N_b + 2(N_b-1)\cos\theta + 2(N_b-2)(\cos\theta)^2 + ... + 2\left(\cos\theta\right)^{N_b-1} \right\}$$

In the more compact notation, we have:

$$\left\langle R^2 \right\rangle = N_b\ell^2 + \ell^2 \sum_{i=1}^{N_b-1} 2\left(N_b - i\right)\left(\cos\theta\right)^i \tag{90.9}$$

Equation 90.9 is an exact expression that is valid for any value of N_b. Unfortunately, it's not a very useful equation. Fortunately, we can get a greatly simplified expression that is very nearly exact for any reasonably large value of N_b. We are able to do this because $\cos\theta$ is less than one, so that $(\cos\theta)^i$ rapidly decreases as i increases. For example, for $\theta = 71^\circ$ (the value corresponding to C-C single bonds), $(\cos\theta)^i$ =1.3x10^{-5} for $i = 10$. For large values of N_b, we only need to consider contributions from values of i which are much smaller than N_b. A very good approximation is obtained by substituting N_b for $N_b - i$ and extending the sum to $i = \infty$, which leads to the following:

$$\left\langle R^2 \right\rangle = N_b\ell^2 \left\{ 1 + 2\sum_{i=1}^{\infty} \left(\cos\theta\right)^i \right\} \tag{90.10}$$

At this point we are in a position to use the following standard sum, valid for any value of x with an absolute magnitude less than 1:

$$\sum_{i=1}^{\infty} x^i = \frac{x}{1-x} \tag{90.11}$$

Use of this expression with $x = \cos\theta$ gives (after a little algebraic rearrangement):

$$\left\langle R^2 \right\rangle = N_b\ell^2 \left\{ \frac{1+\cos\theta}{1-\cos\theta} \right\} \tag{90.12}$$

For singly bonded carbon along the backbone, $\theta = 71^\circ$ and $\langle R^2 \rangle = 2N_b\ell^2$.

90.4 Bond Angle Restrictions

The freely rotating model is overly simplistic, because all bond rotation angles cannot be equally likely. At certain angles, functional groups attached to adjacent carbon atoms are too close to one another, as illustrated here.

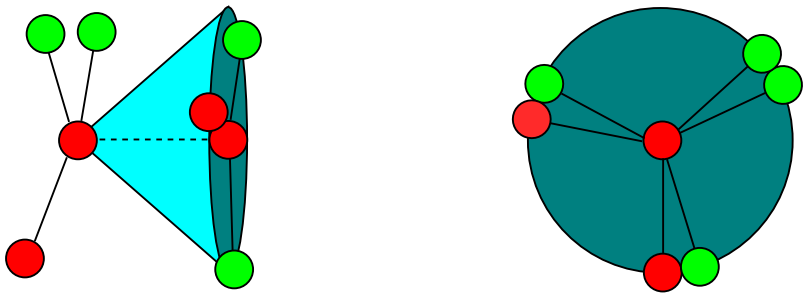


Figure 90.4: Bond angle restrictions shown perpendicular and along the axis of rotation (Newman projection).

Interferences between functional groups cause the free energy to vary as a function of the bond rotation angle,ϕ, with an energy vs.ϕ curve that looks like the one shown in Figure90.5. The trans configuration, where$\phi = 0$, has the lowest free energy, because the backbone carbon bonds are as far from each other as possible in this configuration. Secondary minima in the energy exist for$\phi = \pm 120°$, where there is minimal overlap between the additional groups bonded to the carbon carbon molecules. These configurations are referred to as the gauche configurations.

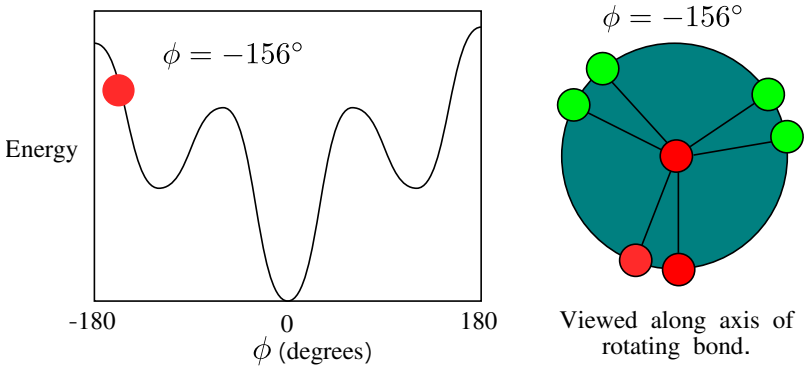


Figure 90.5: Energy as a function of bond rotation angle.

We know that the freely rotating chain model must also be overly simplistic, because some bond rotation angles (those corresponding to the gauche and trans configurations, for example) are more likely than others. In this case$\langle R^2 \rangle$ is given by the following expression, which we present here without proof.

$$\left\langle R^2 \right\rangle = N_b \ell^2 \left\{ \frac{1 + \cos\theta}{1 - \cos\theta} \right\} \left\{ \frac{1 + \langle \cos\phi \rangle}{1 - \langle \cos\phi \rangle} \right\} \tag{90.13}$$

Here$\langle \cos\phi \rangle$ is the average of$\cos\phi$, taking into account the fact that not all angles are equally likely. Note that$\langle \cos\phi \rangle = 0$ for the freely rotating chain model, in which case Eq.90.13 reduces to Eq.90.12.

90.5 Characteristic Ratio

Thecharacteristic ratio,C_∞, is defined as the ratio of the actual value of$\langle R^2 \rangle$ to the value which would be obtained from the freely jointed chain model:

$$C_\infty \equiv \frac{\langle R^2 \rangle}{N_b \ell^2} \tag{90.14}$$

Each model of chain dimensions has it's own prediction for the value ofC_∞. For a chain with fixed bond angles, comparison to Eq.90.13 gives the following:

$$C_\infty = \left\{ \frac{1+\cos\theta}{1-\cos\theta} \right\} \left\{ \frac{1+\langle\cos\phi\rangle}{1-\langle\cos\phi\rangle} \right\} \tag{90.15}$$

Exercise: What would the characteristic ratio be for polystyrene if 70% of the bonds were in trans configurations and the remaining 30% were in gauche configurations?
Solution: We use Eq.90.15 forC_∞,with$\theta = 71^\circ$, the correct bond angle for singly bonded carbon, which we have along the polystyrene backbone. Gauche bonds have$\phi = \pm 120^\circ$ ($\cos\phi = -0.5$) trans bonds have$\phi = 0$ ($\cos\phi = 1$). With our assumed distribution of gauche and trans bonds we obtain the following for$\langle\cos\phi\rangle$:

$$\langle\cos\phi\rangle = 0.7\,(1) + 0.3\,(-0.5) = 0.55$$

This gives$C_\infty = 6.9$, which is actually a pretty reasonable value. Most polymers with singly bonded carbons along the backbone have a value forC_∞ that is between 6 and 10.

Exercise: How is the statistical segment length of a vinyl polymer related toC_∞?
Solution: The characteristic ratio and statistical segment length are both defined in terms of an expression for R2. All we need to do is equate these two expressions (Eqs.89.21 and90.14):

$$C_\infty N_b \ell^2 = N a^2 \tag{90.16}$$

The repeat unit for a vinyl polymer contains two C-C single bonds along the backbone, so$N_b = 2N$. Also, single bonds between carbon atoms have a bond length of 1.54 Å. The statistical segment length is therefore related toC_∞ as follows:

$$a = 2.18\text{Å}\sqrt{C_\infty} \tag{90.17}$$

90.6 Self-Avoiding Random Walks

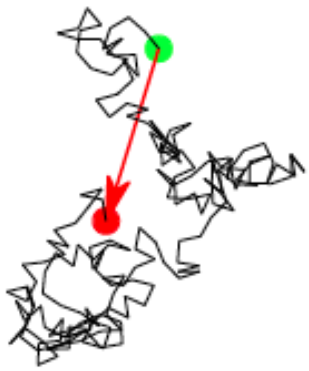


Figure 90.6: Random walk diagram

An isolated polymer molecule in a solvent no longer obeys random walk statistics. If the polymer is surrounded by solvent molecules, these surroundings no longer look the same as other parts of the same molecule. The chain will adopt shapes that avoid direct polymer/polymer contacts, and the end-to-end distance will be larger than the value predicted by random walk models. The rms end-to-end distance increases with the 0.6 power of the chain length, instead of the 0.5 power that was obtained for random walk statistics:

$$R_0 = a_s N^{0.6} \tag{90.18}$$

Note: The value ofa_s is not exactly the same as the statistical segment length describing chain dimensions in pure polymer melts, but it will be close to it.

90.7 Deformation of a Single Polymer Molecule

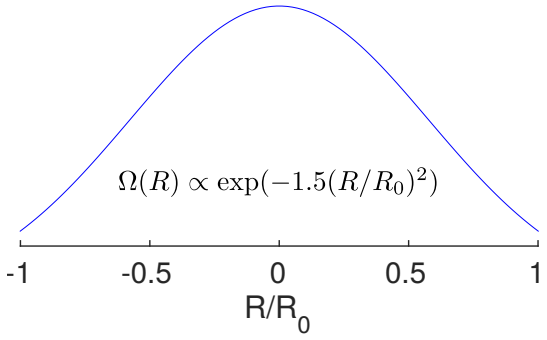


Figure 90.7: Single strand deformation

$$\Omega\left(\vec{R}\right) = C \exp\left(-1.5\left(R/R_0\right)^2\right) \tag{90.19}$$

The entropy is obtained from the number of possible molecular shapes:

$$S_d\left(\vec{R}\right) = k_B \ln \Omega\left(\vec{R}\right) = k_B \ln C - \frac{3k_B R^2}{2Na^2} \tag{90.20}$$

The free energy is assumed to be entirely dominated by this entropic contribution:

$$F_d\left(\vec{R}\right) = -TS_d\left(\vec{R}\right) = -k_B T \ln C + \frac{3k_B T R^2}{2Na^2} \tag{90.21}$$

Note: In reality, there is a small enthalpic contribution to the stretching free energy as well, because stretching the chain increases the relative proportion of higher energy gauche bonds, but we're not going to worry about that correction here. Because we are only every interested in changes in the free energy, constant terms that don't depend on$R$ are going to drop out when we compare free energies, so the detailed value of this constant is not going to matter.

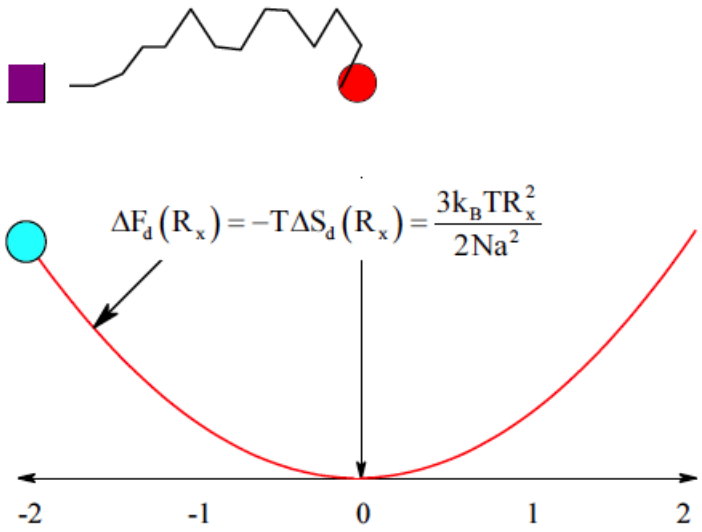


Figure 90.8: Free energy of deformation

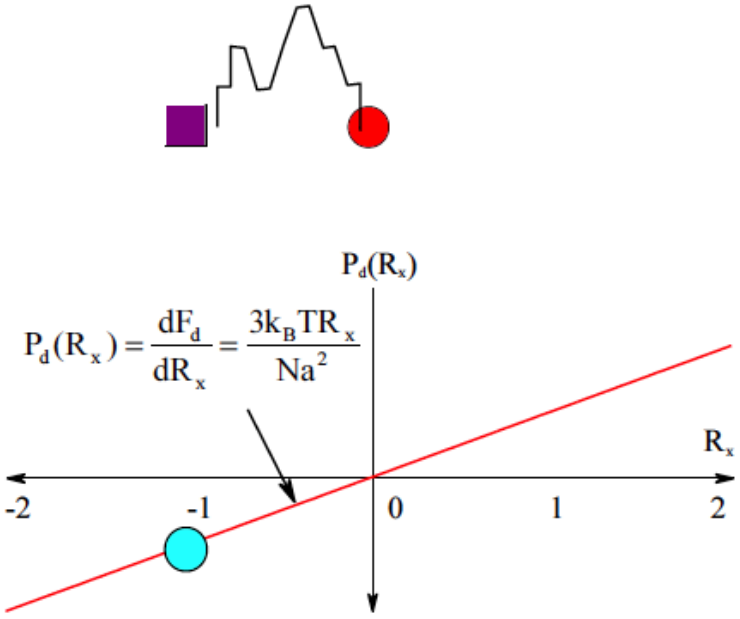


Figure 90.9: Force of deformation

Figure 90.10: Generalized deformation

If the end-to-end vector is not directed purely along the x-axis, we need to consider the effects of the extensions in other directions as well. Because the free energy scales with the square of the separation, only the magnitude of

the end-to-end vector matters. This expression for the elastic free energy of an individual molecule in terms of its end-to-end distance is remarkably simple. It is one of the most important and widely used results in polymer science. In the following section, we will see how it is used to obtain estimates for the elastic modulus of a crosslinked elastomer.

91 Elasticity

In this section, we discuss the mechanical properties of polymers, defined as the relationships which are observed between an applied stress state and the resulting strain. Are focus in this course is on elasticity and viscoelasticity. We begin with a refresher on the fundamental definitions of stress and strain, and then discuss the three most important experimental geometries for mechanical measurements of soft materials.

91.1 Fundamental Definitions

91.1.1 Specification of the State of Stress

The two dimensional stress state is illustrated below in Figure91.1. In two dimensions there are 3 independent components of the stress tensor. Two of these are normal forces where the stress is perpendicular to the relevant surface. These areσ_{xx} andσ_{yy}, with forces acting in the x and y directions, respectively. There is also a shear stress,σ_{xy}, which act parallel to each of the 4 surfaces of a square chunk of material. Note that the magnitudes of the shear forces acting on the 4 different surfaces are all equal to one another, a fact imposed by the requirement that there must be no net force or torques acting on this overall square of material (otherwise the material would be accelerating in response to this net force). The values of these three stress components depend on the way we define the orientation of the coordinate system. It is always possible to define a coordinate system so that the shear components disappear, and we are only left with normal stresses. These are theprincipal stresses, which in two dimensions areσ_1^p andσ_2^p. The same holds for a 3-dimensional stress state: we can always find a coordinate system for any given stress state so that only normal stresses. In three dimensions we have three principal stresses,σ_1^p,σ_2^p andσ_3^p. These act in the three, mutually perpendicular principal stress directions as illustrated in Figure91.2. In summary, any state of stress is completely specified by the three principal stresses (σ_1^p,σ_2^p,σ_3^p), and the orientation of the coordinate system used to define these principal stresses.

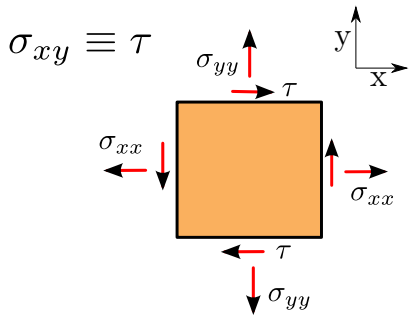


Figure 91.1: Representation of a two-dimensional stress state.

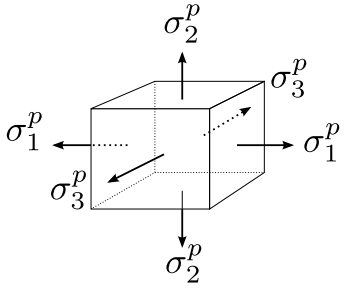


Figure 91.2: 3-D representation of stress, expressed in terms of the principal stresses,σ_1^p,σ_2^p andσ_3^p.

91.1.2 Extension Ratios and the Strain Ellipsoid

A similar concept of principal axes applies to the strain state as well. For polymeric systems (and elastomers in particular), we need a description that can be used to describe both small and large strains. We the concept of thestrain ellipsoid, illustrated in Figure91.3. An undeformed sphere of radiusr_0 is deformed into an ellipsoid with principal radii of$r_0\lambda_1$,$r_0\lambda_2$ and$r_0\lambda_3$. Hereλ_1,λ_2 andλ_3 are theprincipal extension ratios. The values of these principal extension ratios, along with the orientation of the principal axes of the strain ellipsoid, completely specify the state of strain.

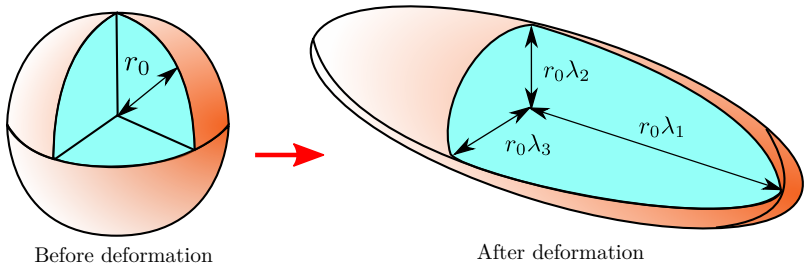


Figure 91.3: Deformation of a sphere of radiusr_0 into an ellipsoid with principal radii of$r_0\lambda_1$,$r_0\lambda_2$ and$r_0\lambda_3$.

91.1.3 Uniaxial Extension or Compression

A common mechanical testing geometry corresponds to the application of a single normal force to the sample, as illustrated in Figure91.4. For our purposes we define the coordinate system so that the normal force is applied in the z direction, and we refer to this force asP_z. We also assume that our sample has isotropic mechanical properties and is in the shape of a cylinder. The result of the applied force is to increase the length of of the sample fromℓ_0 toℓ, and to decrease it's diameter fromd_0 to$d$.

- True stress:$\sigma_z^{true} = P_z/A$,$\sigma_x^{true} = \sigma_y^{true} = 0$ units: N/m^2 (Pa)
- Engineering stress:$\sigma_z^{eng} = P/A_0$ units: N/m^2 (Pa)
- True strains:$\epsilon_z = \ln(\ell/\ell_0)$, $\epsilon_x = \epsilon_y = \ln(d/d_0)$ units: Dimensionless
- Engineering strains:$e_x = e_y = (d-d_0)/d_0$,$e_z = (\ell-\ell_0)/\ell_0$ units: Dimensionless
- Young's Modulus:$E = \sigma_z^{eng}/e_z$ (taken at small strains where$\sigma_z^{eng} \approx \sigma_z^{true}$) units: (Pa)
- Extension ratio:$\lambda_z = \ell/\ell_0 = 1 + e_z$ units: Dimensionless
- Poisson's ratio (for small strains):$v = -e_x/e_z$ units: Dimensionless

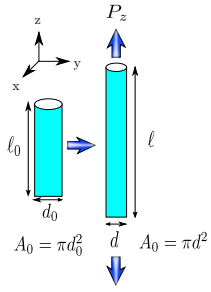


Figure 91.4: Uniaxial tensile deformation.

According to the first law of thermodynamics, the change in free energy of a system is equal to the work done on the system, provided that the process is reversible. In this example, an applied deformation force,P_z, produces a small displacement,$d\ell$. The work done on the system is$P_z d\ell$, which is equal to the change in the deformation free energy change,ΔF_d:

$$P_z d\ell = d(\Delta F_d) \tag{91.1}$$

We can write this force in terms ofλ_z, the extension ratio in this direction:

$$P_z = \frac{d(\Delta F_d)}{d\ell} = \frac{d(\Delta F_d)}{d\lambda_z}\frac{d\lambda_z}{d\ell} \tag{91.2}$$

With$\lambda_z = \ell/\ell_0$, we have$\frac{d\lambda_z}{d\ell} = 1/\ell_0$. We also have$V_0 = A_0\ell_0$, so and we obtain the following for the engineering stress,σ_{eng}:

$$\sigma_{eng} = \frac{P_z}{A_0} = \frac{d}{d\lambda_z}\left(\frac{\Delta F_d}{V_0}\right) \tag{91.3}$$

HereA_0,ℓ_0 andV_0 are the cross sectional area, length and volume of the undeformed material.
The deformation free energy normalized by the undeformed volume is the deformation free energy density,Δf_d:

$$\Delta f_d = \frac{\Delta F_d}{V_0} \tag{91.4}$$

With these definitions the following simple expression gives the engineering stress in terms of the deformation free energy density, for uniaxial deformation along the z axis:

$$\sigma_{eng} = \frac{d}{d\lambda_z}(\Delta f_d) \tag{91.5}$$

91.1.4 Simple Shear

In a simple shear deformation, two parallel surfaces within a material remain at a fixed separation as they slide past one another, as illustrated in Figure91.5. In this example an applied shear force,P_{xy} is applied to the two surfaces, which each have an area,A. The separation between the two surfaces is$h$, and the relative displacement of the two surfaces is$u$. For these quantities we obtain the shear sress,σ_{xy}, the engineering shear strain,e_{xy} orγ, and the shear modulus,G:

- Shear stress:$\sigma_{xy} = P_{xy}/A$,$\sigma_{yz} = \sigma_{xz} = 0$ units: $\mathrm{N/m^2}$ (Pascal)
- Shear strain:$e_{xy} = \gamma = u/d$ units: Dimensionless
- Shear modulus:$G \equiv \sigma_{xy}/\gamma$ (defined in the low-strain regime).

Note thate_{xy} andγ are both commonly used symbols to represent the shear strain.

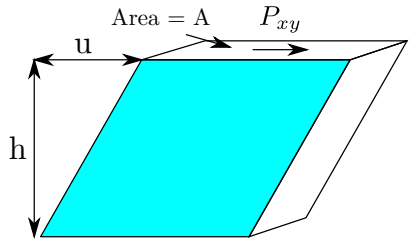


Figure 91.5: Diagrammatic representation of shear deformation.

The force is again given by the derivative of the free energy with respect to the displacement in the direction of the applied force:

$$P_{xy} = \frac{d(\Delta F_d)}{du} = \frac{d(\Delta F_d)}{d\gamma}\frac{d\gamma}{du} = \frac{1}{h}\frac{d}{d\gamma}(\Delta F_d) \tag{91.6}$$

The shear stress is then obtained by dividing by the cross-sectional area (note:$hA = V$,v_s =n_s/V,σ_{xy} =P_{xy}/A :

$$\sigma_{xy} = \frac{d}{\partial d\gamma}(\Delta f_d) \tag{91.7}$$

One complicating factor of simple shear in comparison to uniaxial extension or compression is that the strain ellipsoid is rotated continuously during the shear experiment. As a result we cannot connect the principal axes of the coordinate system in a straightforward way. The situation for shear in the x-y plane is illustrated in Figure91.6. Shear is a two-dimensional stress state, so the material is undeformed in the third direction, in this case the z direction, so$\lambda_z = 1$. The principal strains in the x-y plane areλ_1 andλ_2. While the relationship between the orientation of the strain ellipsoid is a complicated one, the shear strain itself is given by the difference betweenλ_1 andλ_2:

$$\gamma = e_{xy} = \lambda_1 - \lambda_2 \tag{91.8}$$

Eq.91.8 will enable us to determine the deformation free energy of a material undergoing a shear deformation, and is used below in Section91.2.5.

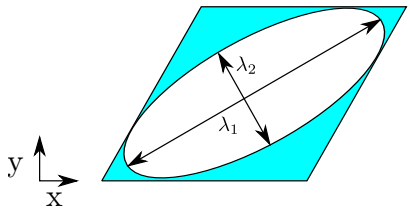


Figure 91.6: Extension ratios for simple shear.

91.1.5 Torsion

An important geometry for characterizing soft materials is the torsional geometry shown in Figure91.7. In this material a cylindrical or disk-shaped material is twisted about an axis of symmetry. The material could be a long, thin fiber (Figure91.7a) or a flat disk sandwiched between two plates (Figure91.7b). We obtain the shear modulus by looking at the torsional stiffness of mateiral,*i.e.*, the Torque,T required to rotate the top and bottom of the fiber by an angleθ_0.

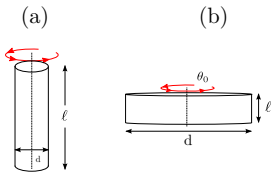


Figure 91.7: Fiber torsion.

We define a cylindrical system with a z axis along the fiber axis. The other axes in this coordinate system are the distance$r$ from this axis of symmetry, and the angleθ around the z axis. The shear strain in the$\theta - z$ plane depends only on$r$, and is given by:

$$e_{\theta z} = r\frac{d\theta}{dz} = r\frac{\theta_0}{\ell} \tag{91.9}$$

The corresponding shear stress is obtained by multiplying by the shear modulus,G:

$$\sigma_{\theta z} = Gr\theta_0/\ell \tag{91.10}$$

We integrate the shear stress to give the torque,T:

$$T = \int_0^{d/2} r\sigma_{\theta z} 2\pi r dr = \frac{\pi G\theta_0 d^4}{32\ell} \tag{91.11}$$

This geometry is commonly used in an oscillatory mode, whereθ is oscillated at a specified frequency. In this case the torque response is obtained by using the dynamic shear moduli defined in Section92.5 in place of$G$.

91.2 Rubber Elasticity

Crosslinked rubbers are unique in that thermodynamic arguments can be used to predict their elastic moduli with remarkable accuracy. Our starting point will be a description of the free energy of an elastomer as a function of its deformation. This free energy is dominated by entropic contributions arising from restrictions on the number of different conformations (or shapes) that polymer strands are able to adopt. The detailed descriptions of polymer chain statistics given earlier were developed so that we would be in a position to describe the mechanical properties of rubbery materials. The pages that follow are proof that statistics and thermodynamics are actually useful!!

91.2.1 Molecular Deformation

Different types of molecular deformations, and the characteristic lengths and forces, are illustrated here.

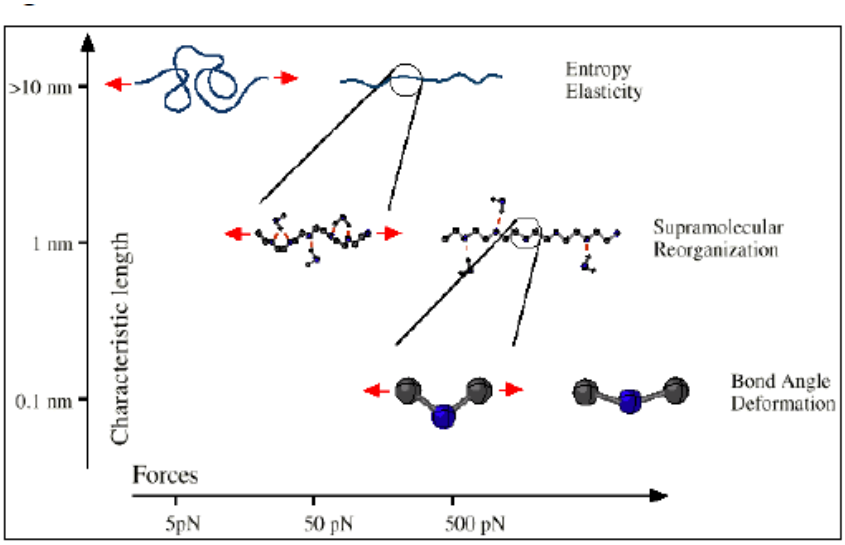


Figure 91.8: Various types of molecular deformation

From Osterhelt et al.:http://www.iop.org/EJ/abstract/1367-2630/1/1/006/

91.2.2 Free energy of a stretched rubber

When a rubber is stretched, the free energy increases. In general the free energy increase has enthalpic and entropic contributions ($F_d = H_d - TS_d$). The basic assumption of rubber elasticity theory is that the free energy increase due to deformation is dominated by the decrease in the entropy. In other words$|T\Delta S_d| >> |\Delta H_d|$. The second assumption we will make is that the deformation at a microscopic level mimics the deformation at a macroscopic level. In other words, relative changes in the spacings between crosslink points are identical to relative changes in the overall sample demonstrations. This assumption is referred to as the affine deformation assumption, and is illustrated on the following page.

91.2.3 Affine Deformation Assumption

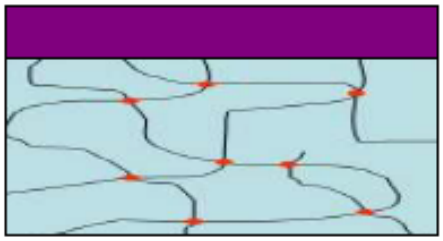


Figure 91.9: A chunk of rubber

The block to the left represents a chunk of rubber. The black lines are individual network strands, and the red circles are crosslink points that tie these strands together. If the molecules are deforming "affinely", the distances between crosslink points change by ratios that are identical to the extension ratios of the material as a whole.

91.2.4 Free energy change due to deformation

Figure 91.10: Contrasting free energies before and after deformation

What really matters when determining the mechanical properties of a material is the change in free energy resulting from a deformation, which we express in terms of the molecular extension ratios in the x, y and z directions. Prior to deformation, we have$\vec{R} = R_x\hat{x} + R_y\hat{y} + R_z\hat{z}$, so that$R^2 = R_x^2 + R_y^2 + R_z^2$. The undeformed free energy,F_0, is given by using this expression forR^2 in Eq.90.21:

$$F_0 = -k_BT\ln C + \frac{3k_BT}{2Na^2}\left\{R_x^2 + R_y^2 + R_z^2\right\} \tag{91.12}$$

After deformation, the values ofR_x,R_y andR_z are each multiplied by the respective extension ratio ($\vec{R} = \lambda_x R_x\hat{x} + \lambda_y R_y\hat{y} + \lambda_z R_z\hat{z}$), so that$R^2 = \lambda_x^2R_x^2 + \lambda_y^2R_y^2 + \lambda_z^2R_z^2$. The deformed free energy,F_{def} is:

$$F_{def} = -k_BT\ln C + \frac{3k_BT}{2Na^2}\left\{\lambda_x^2R_x^2 + \lambda_y^2R_y^2 + \lambda_z^2R_z^2\right\} \tag{91.13}$$

The free energy change due to deformation of the molecule is given as follows:

$$\Delta F_d = F_{def} - F_0 = \frac{3k_BT\left\{(\lambda_x^2-1)R_x^2 + (\lambda_y^2-1)R_y^2 + (\lambda_z^2-1)R_z^2\right\}}{2Na^2} \tag{91.14}$$

This result is for the deformation of a single polymer molecule, which for a crosslinked elastomer corresponds to a segment that connects crosslink points. There are a huge number of these segments in a macroscopic chunk of rubber. To get the free energy change for the material as a whole, we need to replace R_x^2 by $n_s\left\langle R_x^2\right\rangle$, where n_s is the total number of nework strands and $\left\langle R_x^2\right\rangle$ is the average value of R_x^2 for these segments. We need to make similar substitutions for R_y^2 and R_z^2 to obtain the following:

$$\Delta F_d = \frac{3k_BTn_s\left\{(\lambda_x^2-1)\left\langle R_x^2\right\rangle + (\lambda_y^2-1)\left\langle R_y^2\right\rangle + (\lambda_z^2-1)\left\langle R_z^2\right\rangle\right\}}{2Na^2} \tag{91.15}$$

Now we assume that the material was isotropic when it was crosslinked, so that $\left\langle R_x^2\right\rangle = \left\langle R_y^2\right\rangle = \left\langle R_z^2\right\rangle = \left\langle R^2\right\rangle/3$. With this assumption, and with $R_0^2 - Na^2$, we obtain the following result:

$$\Delta F_d = \frac{k_BTn_s\left\langle R^2\right\rangle\left\{\lambda_x^2+\lambda_y^2+\lambda_z^2-3\right\}}{2R_0^2} \tag{91.16}$$

Note that $\left\langle R^2\right\rangle$ is the mean square end-to-end distance of the polymer strands that span the crosslinks, and that R_0^2 is the value of $\left\langle R^2\right\rangle$ when the polymer strands obey random walk statistics. The relationship between these two quantities depends on the conditions of the crosslinking reaction.
It is often useful to work in terms of intensive free energy changes (free energy per unit volume). The free energy of deformation per unit volume Δf_d is obtained very simply from ΔF_d by dividing by V, the volume of a sample. We retain the same expression as shown above, but with the strand concentration, ν_s substituted for the number of strands, n_s:

$$\Delta f_d = \frac{\Delta F_d}{V} = \frac{k_BT\nu_s\beta\left\{\lambda_x^2+\lambda_y^2+\lambda_z^2-3\right\}}{2} \tag{91.17}$$

where ν_s and β are defined as follows:

$$\nu_s = \frac{n_s}{V};\ \beta = \frac{\left\langle R^2\right\rangle}{R_0^2} \tag{91.18}$$

Like the intensive free energy, the strand concentration, ν_s is useful because it does not depend on the overall size and shape of the elastomer. The assumptions used to develop the free energy expression given above correspond to the simplest model of rubber elasticity, which is often referred to as the Neohookean model. In our discussion of the specific cases of shear and uniaxial deformation that we discuss below, we make the additional assumption that the material is incompressible, in which case the product of the three principal extension ratios is one:

$$\lambda_1\lambda_2\lambda_3 = 1 \tag{91.19}$$

91.2.5 Shear deformation of an elastomer

For pure shear we need to relate the shear strain,γ, to the extension ratios,λ_1,λ_2 andλ_3. Pure shear is a two dimensional stress state, with$\lambda_3 = 1$ so the Neohookean strain energy function (Eq.91.17) reduces to the following:

$$\Delta f_d = \frac{k_B T v_s \beta \left\{\lambda_1^2 + \lambda_2^2 - 2\right\}}{2} = \frac{k_B T v_s \beta (\lambda_1 - \lambda_2)^2}{2} = \frac{k_B T v_s \beta \gamma^2}{2} \tag{91.20}$$

The shear stress is obtained by differentiation of the strain energy function with respect toγ (Eq.91.7):

$$\sigma_{xy} = \frac{d}{d\gamma}(\Delta f_d) = G\gamma \tag{91.21}$$

where$G$ is the shear modulus, given as follows:

$$G = k_B T v_s \beta \tag{91.22}$$

Note that the shear stress is proportional to the shear strain, even for large values of the strain. The expression for the shear modulus can be used to write the deformation free energy density in terms of$G$.

91.2.6 Uniaxial deformation of an elastomer

For uniaxial extension or compression the deformation is applied along one axis, which we define as the z axis. We assume that our material is isotropic, so that the extensions in the x and z directions are identical to one another, i.e.$\lambda_x = \lambda_y$. In addition, the material is assumed to be incompressible, so$\lambda_x \lambda_y \lambda_z = 1$. We therefore have$\lambda_x = \lambda_y = \lambda_z^{-1/2}$. We can therefore write the free energy of a deformed elastomer as a function of the single extension ratio,λ_z:

$$\Delta f_d = \frac{G}{2}\left\{\lambda_z^2 + \frac{2}{\lambda_z} - 3\right\} \tag{91.23}$$

The engineering stress is obtained by differentiating with respect toλ_z (Eq.91.5):

$$\sigma_{eng} = \frac{d}{\partial d\lambda}\left(\Delta f_d\right) = G\left\{\lambda_z - \frac{1}{\lambda_z^2}\right\} \tag{91.24}$$

Young's modulus (E) is defined as the derivative of the stress with respect to strain, evaluated at low strain ($\lambda_y = 1$):

$$E \equiv \frac{\partial \sigma_{eng}}{\partial \lambda_z}|_{\lambda_z=1} = G\left\{1 + \frac{2}{\lambda_z^2}\right\}\Big|_{\lambda_z=1} = 3G = 3v_s\beta k_B T, \tag{91.25}$$

The concentration of network strands (v_s) is the inverse of the volume per strand. This can be calculated from the molecular weight and density. DefiningM_s as the number average molecular weight of a network strand (molecular weight between crosslinks) gives:

$$v_s\left(\frac{\text{strands}}{\text{volume}}\right) = \frac{\rho\,(\text{mass/volume})}{M_s\,(\text{mass/mole})} \cdot N_{av}\,(\text{strands/mole}) \tag{91.26}$$

where N_{av} is Avogadro's number ($6.02x10^{23}$). With $R = k_B N_{av} = 8.314\,\mathrm{J/mole \cdot K}$ the expression

$$E = \frac{3\rho\beta RT}{M_s}$$

Finally, note that 1 $\mathrm{J/m^3}$ = 1 Pa. Stresses and elastic moduli have units of force/area or energy/volume.

Exercise: Calculate the expected value of Young's modulus for a crosslinked polyisoprene that has a number average molecular between crosslinks of 4000 g/mole. The density of polyisoprene is 0.9 $\mathrm{g/cm^3}$. Assume that the polymer was crosslinked under equilibrium conditions in the melt state.
Solution: Because the polymer was crosslinked under equilibrium conditions, the network strands obey random walk statistics, with β = 1. To keep everything in SI units, we need the density in units of $\mathrm{kg/m^3}$ (1 $\mathrm{g/cm^3}$ = 1000 $\mathrm{kg/m^3}$), and M_s in kg/mole. In our example, ρ = 900 $\mathrm{kg/m^3}$ and M_s = 4 kg/mole. We'll also assume that we are interested in the elastic modulus near room temperature (T≈ 300 K).

$$E = \frac{3(900\,\mathrm{kg/m^3})(8.314\mathrm{J/mole-K})(300\mathrm{K})}{4\mathrm{kg/mole}} = 1.6\mathrm{x}10^6\,\mathrm{J/m^3}.$$

92 Viscoelasticity

Many of the concepts of rubber elasticity still apply under certain situations to materials that are not actually crosslinked. Consider, for example, the behavior of Silly-Putty. Silly-putty is based on silicones that are not actually crosslinked. Nevertheless, this material still bounces like an ordinary crosslinked rubber ball. This is because over very short times it behaves elastically. The deformation energy stored in the material as it comes into contact with the surface and deforms is available to propel the material back into the air as the deformation relaxes and this strain energy is converted to kinetic energy. If we let the sample sit on a surface for a long period of time, however, it eventually flows and behaves as a liquid.

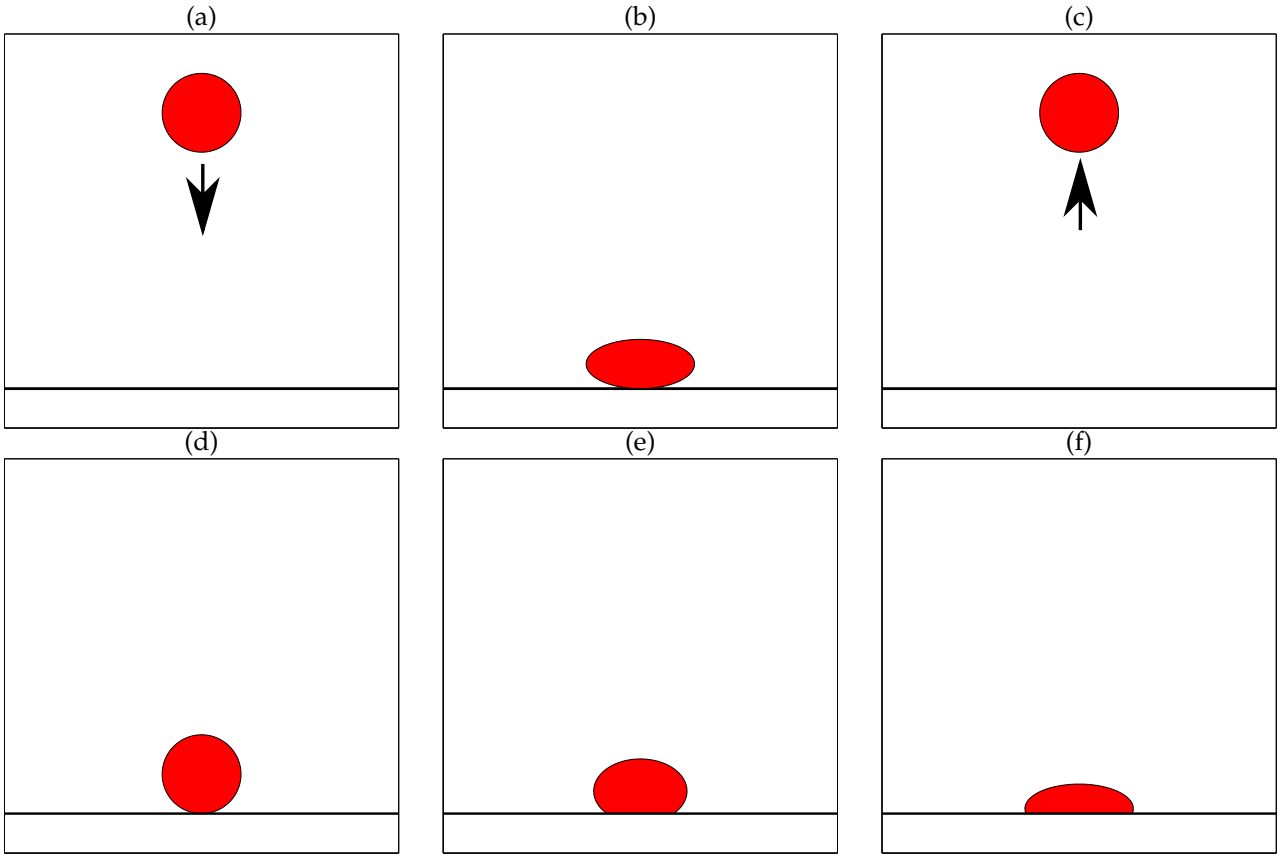


Figure 92.1: Illustration of viscoelastic behavior. Images of a viscoelastic material behaving elastically as it bounces from a surface (a-c) and the same material spreading over a surface at long times (d-f).

Before we discuss the way that the properties of viscoelastic materials are characterized, it is useful to discuss the limiting behaviors of purely elastic solids and perfectly viscous liquids. We illustrate this by showing how differently materials respond to a shear strain,γ that begins to increase at some constant rate to some final value, and is then fixed at this final value as illustrated in the top panel of Figure92.2. The response for solids, liquids and viscoelastic materials is as described below.

Solids: In this case the stress is proportional to the strain, and the rate at which the strain is applied does not matter. As a result the time dependence of the stress looks just like the time dependence of the strain. The slopes of the time-dependent stress and strain curves (the top two curves in Figure92.2) are related to one another by the shear modulus of the material.

Liquids: In liquids the stress is proportional to the rate at which the strain is applied, and is independent of the current strain. Liquids do not store any strain energy, and as soon as the strain stops changing, the stress drops back to zero. The stress for the time dependent strain shown in Figure92.2 is constant while the stress is increasing and is zero otherwise. The stress is given by the shear viscosity,η, as follows:

$$\sigma_{xy} = \eta \frac{d\gamma}{dt} = \eta \dot{\gamma} \tag{92.1}$$

Viscoelastic Materials: Viscoelastic materials have characteristics of both solids and liquids. The shear stress depends on both the shear stress and the strain rate, and is not a simple function of either one. The shear stress actually depends on the details of the previous strain history, as described in more detail below.

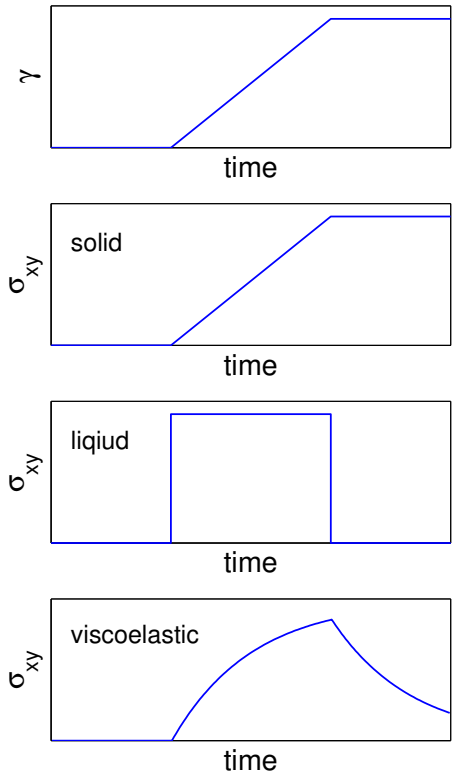


Figure 92.2: Interactive contrast of elastic, viscous, and viscoelastic response.

92.1 Relaxation Modulus

The linear (small strain) properties of viscoelastic materials are described by the same set of moduli used to describe elastic materials (Young's modulus, shear modulus, Poisson's ratio). For viscoelastic materials, however, all of these properties time dependent. Here we illustrate the relevant concepts for a shear geometry, in which case we are interested in the shear relaxation modulus,$G(t)$. This quantity is defined as a ratio of shear stress to shear strain, in the same way that the time-independent elastic shear modulus is defined. In this case, however, the shear stress is not constant but depends on the time. The formal definition is illustrated by considering the instantaneous application of a shear strain,γ_0, to a material as illustrated in Figure92.3. The shear relaxation modulus is defined as the ratio of the resulting time dependent stress to this initial strain:

$$G(t) \equiv \frac{\sigma_{xy}(t)}{\gamma_0} \tag{92.2}$$

The shear modulus is independent ofγ_0, provided thatγ_0 is sufficiently small. This low strain regime, referred to as the linear viscoelastic regime is the focus of our remaining discussion of viscoelasticity.

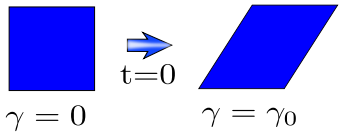


Figure 92.3: Application of an instantaneous shear strain.

92.2 Boltzmann Superposition Principle

In the linear viscoelastic regime the effects of strains applied at different times are additive, a concept known as theBoltzmann superposition principle. A simple example is illustrated in Figure92.4, where the following step strains are applied to the sample:

- A strain of$\Delta\gamma_1$ applied at timet_1
- A strain of$\Delta\gamma_2$ applied at timet_2
- A strain of$\Delta\gamma_3$ applied at timet_3

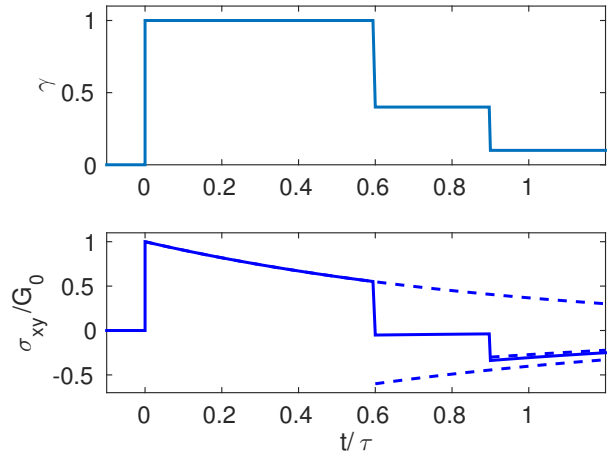


Figure 92.4: Material response to a series of step strains.

The stress contribution from these different strains is obtained by multiplying by the shear relaxation modulus, evaluated at the time that has passed since the strain was applied. If strain was applied att_1, for example, the stress at some later time,t, is obtained by multiplying by$G\left(t-t_1\right)$. Adding up the contributions from the three different step strains in our example leads to the following::

$$\sigma_{xy} = \Delta\gamma_1 G(t-t_1) + \Delta\gamma_2 G(t-t_2) + \Delta\gamma_3 G(t-t_3) \tag{92.3}$$

This expression can easily be extended to include an arbitrary number of step strains:

$$\sigma_{xy} = \sum_i \Delta\gamma_i G(t-t_i) \tag{92.4}$$

By taking a sum of very small strain increments, we can generalize this expression to account for a continuously changing strain. We begin by writing$\Delta\gamma_i$ in the following way:

$$\Delta\gamma_i = \frac{d\gamma}{dt}(t_i)\Delta t_i = \dot{\gamma}(t_i)\Delta t_i \tag{92.5}$$

This substitution leads to the following expression for the stress:

$$\sigma_{xy} = \sum_i \dot{\gamma}(t_i) G(t-t_i)\Delta t_i \tag{92.6}$$

We can write this in integral form by taking$\Delta t_i \to 0$ and replacing the summation by an integral over all times less than the current time:

$$\sigma_{xy}(t) = \int_{-\infty}^{t} \dot{\gamma}(t_i) G(t-t_i) dt_i \tag{92.7}$$

This expression gives the current stress that remains as a result of all of the strains introduced at different times in the past. It is often convenient to change variables so that$s = t - t_i$,$dt_i = -ds$. Note that s =∞ when$t_i = -\infty$, and$s = 0$ when$t_i = t$. The integral can therefore be rewritten as follows:

$$\sigma_{xy}(t) = \int_0^\infty \dot{\gamma}(t-s)G(s)ds \tag{92.8}$$

Note that because$s$ represents time,$G(s)$ and$G(t)$ represent the same time-dependent relaxation modulus. We are just using different variables (s and$t$) to represent time.
The simplest application of Eq. is 'steady shear', where a constant strain rate is applied. Because$\dot{\gamma}$ is independent of time in this case, it can be brought outside the integral, giving the following:

$$\sigma_{xy} = \dot{\gamma}\int_0^\infty G(t)dt \tag{92.9}$$

The following expression is obtained for the viscosity:

$$\eta_0 = \frac{\sigma_{xy}}{\dot{\gamma}} = \int_0^\infty G(t)dt \tag{92.10}$$

We refer to this viscosity as thezero shear viscosity, because it assumes that the material is in the linear viscoelastic regime, where the applied stress is low enough so that the stress is proportional to the magnitude of the applied strain. The subscript '0' is a reminder to us that the zero shear viscosity corresponds to a limiting value of the viscosity that is obtained at a very low strain rate.

92.3 Idealized Relaxation Curves

One thing to keep in mind when looking at the viscoelastic properties of materials is that processes occur over a very large range of time scales. To capture all of these time scales we typically plot the time on a logarithmic scale. The same is true for the frequency-domain experiments discussed below. The relaxation modulus can often vary of several orders of magnitude, so that we also plot the relaxation modulus itself on a logarithmic scale.
Three different idealized forms of the relaxation modulus are shown in Figure92.5. The simplest behavior corresponds to aMaxwell model, where a single relaxation time,τ_g, describes the decay in the modulus from the glassy value ofG_g:

$$G(t) = G_g \exp\left(-t/\tau_g\right) \tag{92.11}$$

This relaxation time could, for example, describe the decay in the modulus from a glassy value of$\approx 10^9$ Pa. This relaxation time can be viewed as a 'glass transition time', and its existence is a clue that some equivalence between time and temperature must exist. The relaxation time characterizing the glass transition depends very strongly on the temperature, but is independent of the polymer molecular weight. If the polymer molecular weight is very high, however, the polymers become entangled with one another, and behave elastically for times that are too short for these entanglements (shown schematically in Figure92.6) to relax by molecular diffusion. In addition toτ_g, there is a second transition time,τ_e, determined by the lifetime of these molecular entanglements. Addition of this second relaxation results in the following expression for the relaxation modulus:

$$G(t) = G_g \exp\left(-t/\tau_g\right) + G_e \exp\left(-t/\tau_e\right) \tag{92.12}$$

HereG_e is the plateau modulus (often referred to asG_N^0 in the literature). Entanglements behave like crosslinks, but they have finite lifetimes as the molecules diffuse and the entanglements are eliminated and reformed

elsewhere. The plateau modulus is given by the concentration of entanglements,v_e, or equivalently, by the number average molecular weight between entanglements,M_e (referred to as the entanglement molecular weight):

$$G_e = v_e k_B T = \frac{\rho RT}{M_e} \tag{92.13}$$

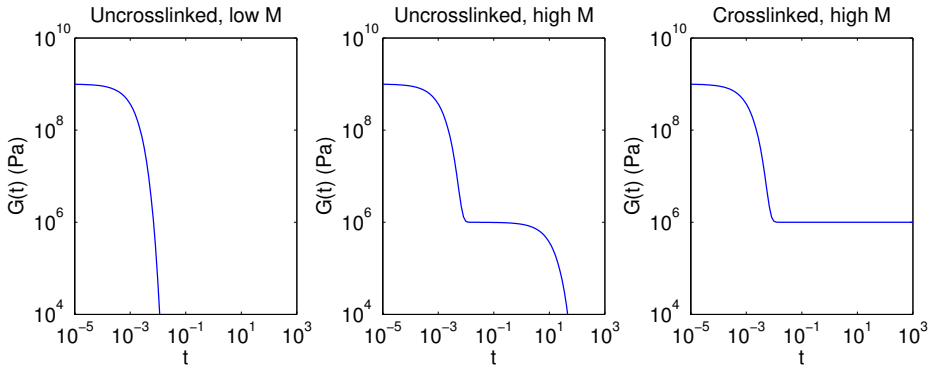


Figure 92.5: Characteristic relaxation curves.

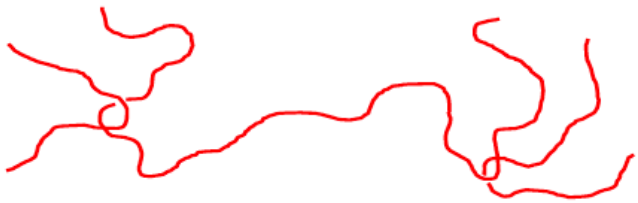


Figure 92.6: Schematic representation of a molecular entanglement.

The actual relaxation behavior of polymeric materials is complex, and can generally not be described in detail by including only one or two relaxation times. The actual relaxation behavior can always be described with sufficient accuracy by ageneralized Maxwell model, where we include a large number of individual exponential relaxation processes:

$$G(t) = \sum_{i=1}^{N_r} G_i \exp(-t/\tau_i) \tag{92.14}$$

Inclusion of a sufficiently large number of relaxation processes (largeN_r) enables very complicated relaxation behavior to be modeled accurately.

92.4 Temperature Dependence

A characteristic of single-component polymeric materials is that the temperature dependence of all of the different relaxation times is the same. The concept is illustrated in Figure92.7, which shows the relaxation modulus at two different temperatures for an idealized material with two different relaxation times. The relaxation times scale with temperature according to the same temperature dependent shift factor,a_T. The shift factor is equal to one at a reference temperature,T_{ref}, and increases as the temperature is decreased. Because both relaxation times are multiplied by the same factor, the curve is shifted linearly along the logarithmic x axis.

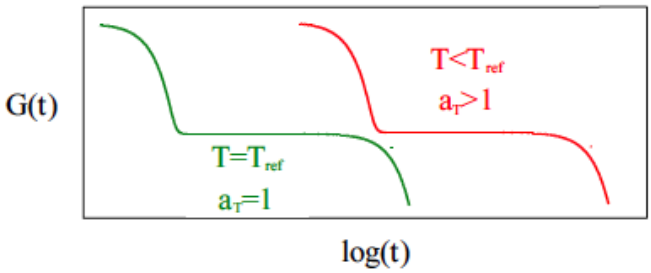


Figure 92.7: Temperature dependence of a material with two relaxation times.

Because all of the relaxation times are multiplied bya_T when the temperature is changed, data obtained at different temperatures superpose when plotted as a function oft/a_T. As a result the relaxation behavior over a large range of times can be obtained by measuring the relaxation spectrum at different times, and shifting the data to the reference temperature.
TheVogel equation is an empirical equation that is often used to describe the temperature dependence of the shift factor,a_T:

$$\log(a_T) = A + \frac{B}{T - T_\infty} \tag{92.15}$$

The factor A is determined from the requirement that$a_T = 1$ at$T = T_{ref}$:

$$A = -\frac{B}{T_{ref} - T_\infty} \tag{92.16}$$

The relaxation times diverge to∞ at$T = T_\infty$, which in free volume theory is the temperature at which the free volume of the equilibrium liquid goes to zero.

92.5 Dynamic Moduli

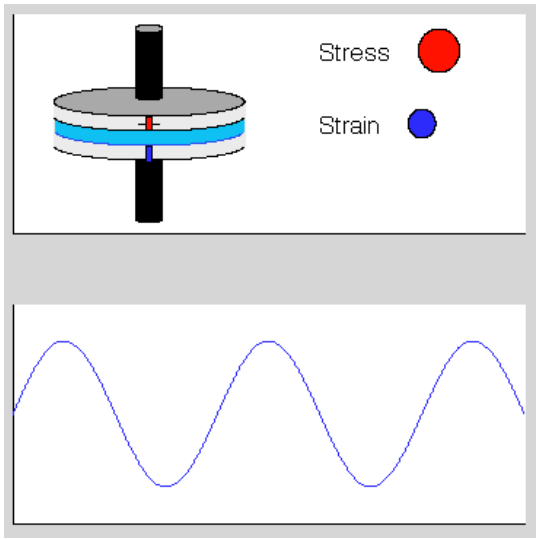


Figure 92.8: Dynamic mechanical spectroscopy.

Experimentally, a wide variation in time scales is accessed by oscillating or vibrating the sample and measuring the frequency response of the material. Consider, for example, an oscillatory shear strain:

$$\gamma = \gamma_0 \sin(\omega t) \tag{92.17}$$

The strain rate is also sinusoidal:

$$\dot{\gamma}(t) = \frac{d\gamma(t)}{dt} = \omega\gamma_0 \cos(\omega t) = \omega\gamma_0 \sin(\omega t + \pi/2) \tag{92.18}$$

Note that the strain and the strain rate are out of phase by$\pi/2$ (90°). This concept of a phase difference is very important in understanding the frequency dependent dynamic moduli. In this case we use Boltzmann superposition to obtain an expression for the stress:

$$\sigma_{xy}(t) = \int_0^\infty \dot{\gamma}(t-s)G(s)ds = -\omega\gamma_0 \int_0^\infty \cos\left\{\omega(t-s)\right\} G(s)ds \tag{92.19}$$

Now we make use of the following trigonometric identity:

$$\cos(a-b) = \sin(a)\sin(b) + \cos(a)\cos(b) \tag{92.20}$$

We can therefore write the time dependent stress in the following way.

$$\sigma_{xy}(t) = \gamma_0\omega \left[\int_0^\infty G(s)\sin(\omega s)ds\right] \sin(\omega t) + \gamma_0\omega \left[\int_0^\infty G(s)\cos(\omega s)ds\right] \cos(\omega t) \tag{92.21}$$

We defineG' (storage modulus) andG'' (loss modulus) such thatG' describes the component of the stress that is in phase with the strain (the elastic component) andG'' describes the component of the stress that is in phase with the strain rate (the viscous component).

$$\frac{\sigma_{xy}(t)}{\gamma_0} = G'(\omega)\sin(\omega t) + G''(\omega)\cos(\omega t) \tag{92.22}$$

By comparing Equations92.21 and92.22, we obtain the following forG' andG'':

$$\begin{aligned} G'(\omega) &= \omega\int_0^\infty G(t)\sin(\omega t)dt \\ G''(\omega) &= \omega\int_0^\infty G(t)\cos(\omega t)dt \end{aligned} \tag{92.23}$$

In is useful to consider the behavior ofG' andG'' in limiting cases where the system is a perfectly elastic solid with no viscous character, and where the material is a Newtonian liquid with no elastic character:

Perfectly Elastic System: In this case the shear modulus is independent of both time and frequency. For$\gamma(t) = \gamma_0 \sin(\omega t)$ we have:

$$\sigma_{xy}(t) = G\gamma(t) = G\gamma_0 \sin(\omega t) \tag{92.24}$$

Comparing to Eq.92.22 gives:

$$\begin{aligned} G' &= G \\ G'' &= 0 \end{aligned} \tag{92.25}$$

Perfectly Viscous System: We again have$\gamma(t) = \gamma_0 \sin(\omega t)$, but this time the shear stress depends only on the strain rate:

$$\sigma_{xy}(t) = \eta \frac{d\gamma(t)}{dt} = \eta\omega\gamma_0 \cos(\omega t) \tag{92.26}$$

Comparing to Eq.92.22 gives:

$$\begin{gathered} G' = 0 \\ G''(\omega) = \omega\eta \end{gathered} \tag{92.27}$$

92.5.1 Phase Angle and Loss Tangent

G' andG'' can be viewed as the real and imaginary components of a a complex modulus,G^*:

$$G^*(\omega) = G'(\omega) + iG''(\omega) \tag{92.28}$$

As with any complex number, we can representG^* by its magnitude, which we refer to as$|G^*|$ and its phase angle, which we refer to asδ:

$$\begin{gathered} |G^*(\omega)| = \left\{ (G'(\omega))^2 + (G''(\omega))^2 \right\}^{1/2} \\ \delta = \tan^{-1}\left(\frac{G''(\omega)}{G'(\omega)}\right) \end{gathered} \tag{92.29}$$

The significance of$|G^*|$ andδ are illustrated by writing the sinusoidal displacement and the resulting stress in the following way:

$$\frac{\gamma(t)}{\gamma_0} = \sin(\omega t) \tag{92.30}$$

$$\frac{\sigma_{xy}(t)}{\gamma_0} = |G^*| \sin(\omega t + \delta) \tag{92.31}$$

$|G^*|$gives the ratio of the stress amplitude to the strain amplitude, andδ is the phase lag between the stress and strain. The following expressions relate$|G^*|$,G',G'' andδ:

$$\begin{gathered} G' = |G^*| \cos(\delta) \\ G'' = |G^*| \sin(\delta) \\ \tan(\delta) = \frac{G''}{G'} \end{gathered} \tag{92.32}$$

TheLoss tangent, $\tan(\delta)$, gives the fraction of energy that is dissipated during each cycle. It is maximized when the inverse of the frequency corresponds to a characteristic relaxation time of the material. For a material with two characteristic relaxation times,τ_1 andτ_2, $\tan(\delta)$ will be maximized at$\omega = 2\pi/\tau_1$ and$\omega = 2\pi/\tau_2$.

92.5.2 Spring and Dashpot Models

As mentioned above, a generalized Maxwell model, where the time-dependent relaxation is described by a sum of exponential relaxations, is often used to describe the viscoelastic behavior of soft materials. A way of visualizing the effect of different relaxation times is to imagine a series of springs (elastic elements) and dashpots (viscous elements) that are linked together in some way. Because it is easy to visualize the extension of springs in a tensile experiment than it is to visualized them in a shear experiment, we consider the time dependent behavior of the tensile modulus,E (Young's modulus in a traditional, static experiment). A single maxwell element

(Figure92.9) has a spring of modulus$E$ in series with a dashpot of viscosityη, and gives and gives an exponential relaxation with a relaxation time,τ, given byη/E.

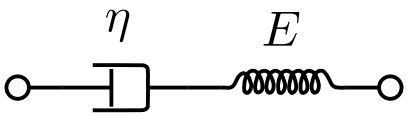


Figure 92.9: Maxwell model.

In a generalized Maxwell the stress is given by the sum of the contributions from all the individual Maxwell elements. In Figure92.10 we include a spring with modulusE_r that is not in series with a dashpot. This represents the 'relaxed' modulus, and describes the stress remaining in the system for$t \to \infty$. The overall stress after the instantaneous application of a tensile strain,e, is:

$$\frac{\sigma}{e} = E_r + \sum_j E_j \exp\left(-t/\tau_j\right), \tag{92.33}$$

with$\tau_j = \eta_j/E_j$. In order to obtainG' andG'' we use Eq.92.23. The sine and cosine transforms of an exponential function can be determined analytically, and we end up with the following for$G'(\omega)$ and$G''(\omega)$ if$G(t) = G_0 \exp(-t/\tau)$:

$$\begin{aligned} G'(\omega) &= \omega G_0 \int_0^\infty \exp(-t/\tau)\sin(\omega t)dt = \frac{G_0\omega\tau}{1+(\omega\tau)^2} \\ G''(\omega) &= \omega G_0 \int_0^\infty \exp(-t/\tau)\cos(\omega t)dt = \frac{G_0(\omega\tau)^2}{1+(\omega\tau)^2} \end{aligned} \tag{92.34}$$

(Note: A MATLAB script that demonstrates the use of the use of MATLAB's symbolic algebra capability to determine these integrals is given in Appendix 1). Plots of these values forG' andG'' are shown in Figure92.11. Note that for$\omega = 1/\tau$,$G' = G''$. For$\omega > 1/\tau$ the material behaves as a liquid, with$G'' > G'$. For angular frequencies above$1/\tau$, the opposite is true, and the material has solid-like behavior with$G' > G''$.

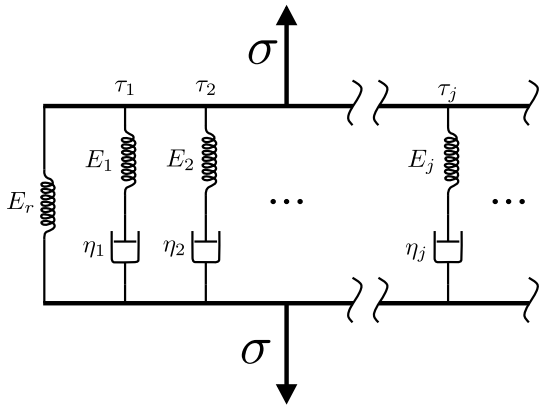


Figure 92.10: Spring-dashpot representation of a generalized Maxwell model.

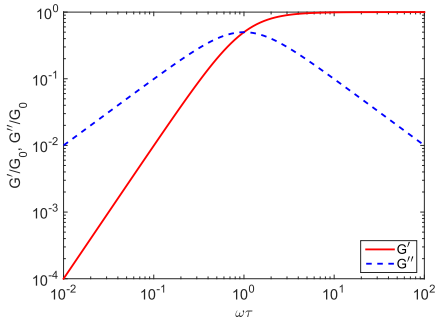


Figure 92.11: G' andG'' for a Maxwell model (Eq.92.34).

Because the generalized Maxwell model is just the sum of individual Maxwell elements, we can apply Eq.92.34 to each of these individual elements, and sum the result to get the storage and loss moduli for the general case where many different relaxation times are needed to fully characterize the linear viscoelastic response of the material:

$$\begin{aligned} G'(\omega) &= E_r + \sum_j \frac{G_j \omega \tau_j}{1+(\omega\tau_j)^2} \\ G''(\omega) &= E_r + \sum_j \frac{G_j (\omega\tau_j)^2}{1+(\omega\tau_j)^2} \end{aligned} \tag{92.35}$$

Zero Shear Viscosity If$E_r = 0$ and$\omega << 1/\tau_0$, whereτ_0 is the longest relaxation time of the material, then the response will be liquid like, with$G'' = \eta\omega$. The zero shear viscosity is obtained from the response in this low frequency regime:

$$\eta_0 \equiv \lim_{\omega\to 0} \frac{G''(\omega)}{\omega} \tag{92.36}$$

This quantity is equal to the zero shear viscosity defined in Eq.92.10.

92.5.3 Time-Temperature Superposition of Dynamic Mechanical Data

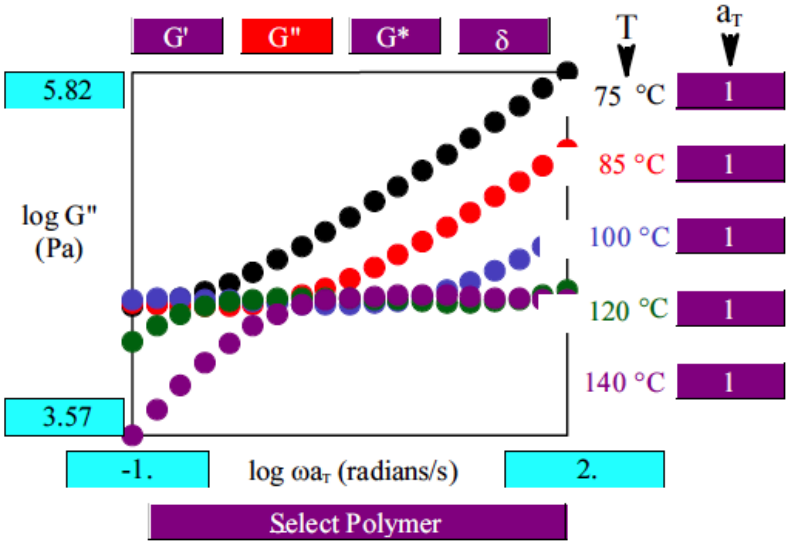


Figure 92.12: An interactive showing dynamic mechanical data

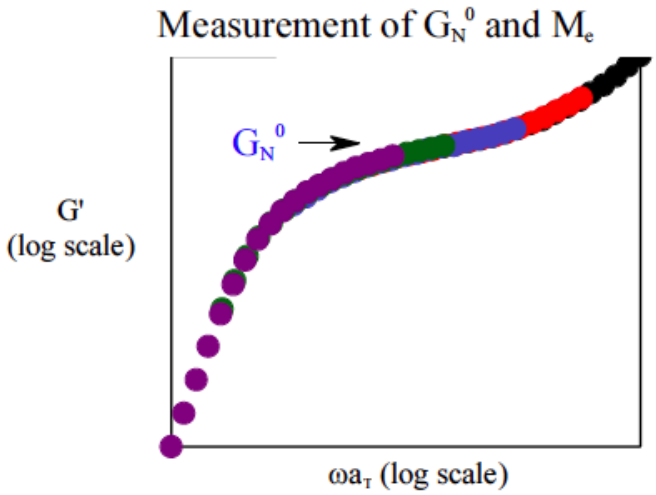


Figure 92.13: Plot showing temperature shifted curve

Measurement ofG_N^0 andM_e In this plot we show the temperature shifted master curve forG′, from the PTBA data. The specific value ofG_N^0 is generally taken as the inflection point in G′, or as the value of G′ at the frequency at whichδ is minimized. Once G′ is known, the entanglement molecular weight, Me can be determined from therelevant expression.

Viscocity from Rheological Data The zero shear viscosity,η_0, at the temperature where$a_T = 1$ is determined from the low frequency data, in the regime where G′ is proportional toω. The full temperature dependence of n is then given by the measured temperature dependence a dependence ofa_T.

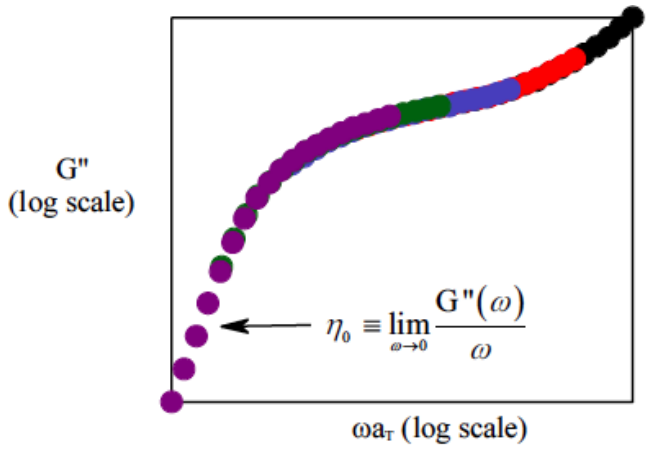


Figure 92.14: Rheological data

93 Semicrystalline Polymers

Many polymers of commercial importance are able to crystallize to a certain extent. The most widespread crystallizable polymers are polyethylene and isotactic polypropylene. As we will see, crystallization is never complete, and there is always a certain amorphous fraction in any polymer. For this reason, many of the concepts which apply to amorphous polymers are still valid for semicrystalline polymers. For example, semicrystalline polymers are often quite brittle at temperatures below the glass transition temperature of the amorphous fraction. In the following pages, we introduce some common semicrystalline polymers, and discuss their basic structural feature polymers, in addition to kinetic issues that are important in their processing.

93.1 Structural Hierarchy in Semicrystalline Polymers

The structure of semicrystalline polymers is much more complex than the structure of amorphous polymers, and our discussion in the following pages is not as quantitative as our discussion of the random walk configurations of polymer chains in amorphous polymers. Instead, we illustrate the features of semicrystalline polymers by considering the following basic structural features:

(1) Helices formed by individual polymer molecules

(2) Perfect crystals formed by the lateral packing of these helices

(3) Lamellar crystallites formed by folded helices

(4) Semicrystalline regions formed by stacking of lamellar crystallites

(5) Spherulitic morphologies formed by radial growth of lamellar crystallites

93.2 The Structural Repeat Unit

The molecular structure is an important factor in the ability of a polymeric material to crystallize. One relevant parameter related to the structure vinyl polymers is the number of backbone carbons in the repeating molecular structure. We refer to this as thestructural repeat unit, which is distinct from themonomeric repeat unit that is defined in terms of the structure of the monomer itself. Polyethylene, example, can be viewed as a repeating string ofCH_2 units, so there one carbon atom per structural repeat unit, although there are two carbon atoms per monomeric repeat unit. Tacticity needs to be accounted for as well. Isotactic vinyl polymers have two backbone carbon atoms per structural repeat, because all non-hydrogen substituent groups are on the same side of the molecule. Syndiotactic vinyl polymers have four backbone carbon atoms per structural repeat, because the substituent groups alternate between different sides of the molecule.

Exercise: What are the structural repeats for each of the following polymers:

(1) Poly(tetrafluoroethylene)

(2) Syndiotactic Polypropylene

(3) Isotactic Polypropylene

(4) Atactic Polystyrene

Solution:

(1) The repeating unit here is simplyCF_2 since I can reconstruct the polymer simply by connecting these units to one another.

(2) The repeating unit here contains 4 backbone carbons and two methyl groups, since there is a methyl group on every other carbon, and these methyl groups alternate from one side of the backbone to the other.

(3) In this case there are two backbone carbons and a single methyl group (a single proplyne unit).

(4) Since the polymer is atactic, it has an inherently disordered strucutre, with no structural repeat. This is why it is always amorphous.

93.3 Helix Formation

The first step in understanding the structure of crystalline polymer is to understand the shapes that the polymer molecules take when they pack within the crystalline structure. The simplest situation is for a molecule like polyethylene, for which the molecules in the crystalline structure are all fully extended, with all of their bonds in the trans configuration as shown in Figure93.1.

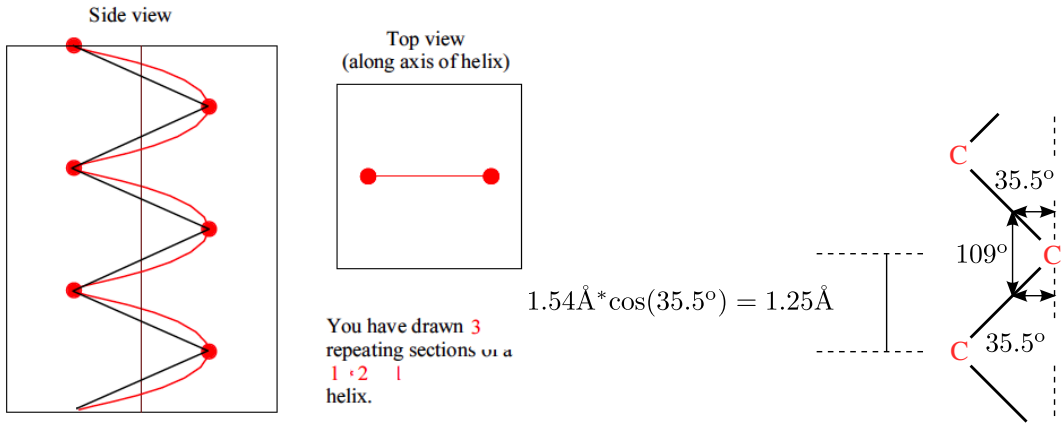


Figure 93.1: Diagram of a planar zigzag (left), with the specific geometry for tetrahedrally bonded carbon (right). Note that the bond angles appear distorted in the leftmost portion of the figure because the horizontal and vertical scales are not necessarily equivalent.)

The planar zigzag configuration is one special case of the shapes of polymers along the chain axis. In the more general case, the backbone carbon atoms all lie along the a helix that is formed about some central axis. The general shape of a helix is illustrated in Figure**??**. Helices of vinyl polymers form by rotation about the C-C single bonds which make up the backbones of the molecules. The curved line in the side view of the planar zigzag in Figure93.1 represents the helix that passes through each of the carbon atoms in a single molecule in the polyethylene crystal structure.

The general notation describing the helical structure is$x * y/z$ where$x$,y and$z$ are defined as follows:

- x: the number of backbone carbons in the structural repeat. This is the only number that we can get by looking at the molecular structure of the polymer itself. We need to know the crystal structure to know what$y$ and$z$ are.
- y: the number of structural repeats per crystallographic repeat. This means that the productxy gives the number of carbon atoms that repeat along the chain direction within one unit cell.
- z: the number of turns of the helix per crystallographic repeat.

The planar zigzag can be formally viewed as a 1*2/1 helix, with 3 turns of the helix shown in Figure93.1. From the top view, we see that all of the backbone atoms of this helix are in the same plane. More complicated helices form because of steric interactions between the other substituents that are placed on the carbon backbones. It's easy for polyethylene molecules to adopt an all-trans configuration because the hydrogen atoms are small and don't interfere with each other when the polymer chain adopts this particular shape. Fluorine atoms are larger than hydrogens, however, so poly(tetrafluorethylene) (Teflon^{TM}) adopts a more complicated, 13/7 helix, which is is shown below in Figure93.2.

Exercise: What is length of a polyethylene molecule with M = 10,000 g/mole, if the molecule is in all-trans (planar zigzag) configuration.
Solution: For bond angle,θ, for singly bonded carbon atoms are 71°, and the C-C bond length is 1.54 Å. From the drawing below, we see that the projected length of this bond along the direction of the zigzag is equal to 1.25 Å. We just need to multiply this length by the number of carbons in the polyethylene chain, which we get by dividing the molecular weight of 10,000 g/mole by the molecular weight of a singleCH_2 unit (14 g/mole). This gives the following for the total length,L:

$$L = 1.25\,\text{Å}\left(\frac{10,000\,\text{g/mole}}{14\,\text{g/mole}}\right) = 960\,\text{Å}$$

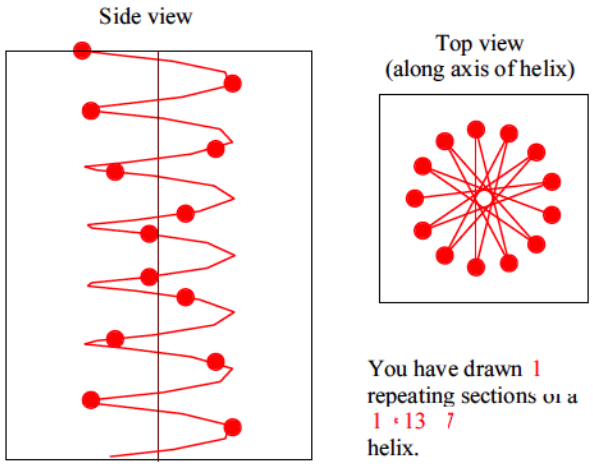


Figure 93.2: Representation of a 13/7 helix.

93.4 Crystalline unit cells

The structure of a perfect, three dimensional crystal, is defined by the lateral packing of polymer helices into repeating structure. As with any crystalline material, the structure is defined by the unit cell dimensions, and by the contents of the unit cell. The dimensions of the unit cell are defined by three lengths (a,b and$c$) and by three angles (α,β andγ), as illustrated in Figure93.3.

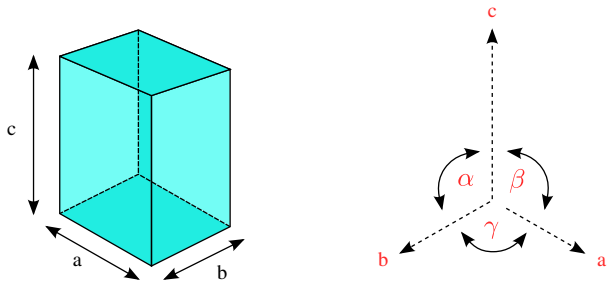
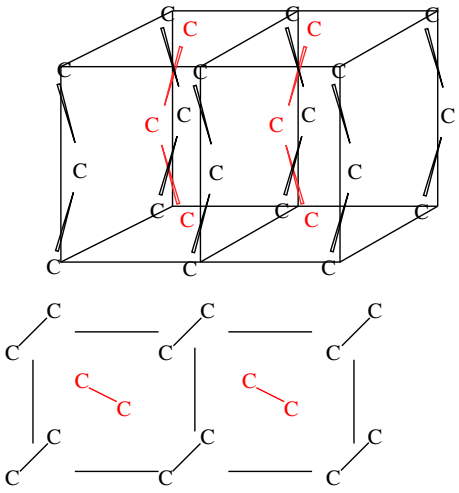


Figure 93.3: Diagram of a crystalline unit cell.

The perfect crystal structure of a polymer is defined by the helix type, unit cell dimensions and by the packing of the helices in the unit cell. Polyethylene can exist in multiple crystal structures. One of these, referred to as polyethylene I, is illustrated in Figure93.4. This crystal structure has two 1*2/1 helices packed into an othorhombic unit cell ($\alpha = \beta = \gamma = 90^\circ$). As is often the case, the helix axis corresponds to the$c$ direction of the unit cell.



Side View

Top View

$\alpha = \beta = \gamma = 90°$

a= 4.946 A

b= 7.418 A

c= 2.546 A

2 helices per unit cell

Figure 93.4: Structure of polyethylene I.

Polymer	Helix Type	a, b, c (Å)	α,β,γ	structural units/cell
polyethylene I	1*2/1	7.41, 4.95, 2.55	90, 90, 90	4
polyethylene II	1*2/1	8.09, 4.79, 2.53	90, 107.9, 90	4
poly(tetrafluoroethylene) I	1*13/6	5.59, 5.59, 16.88	90, 90, 113.3	13
polypropylene (syndiotactic)	4*2/1	14.50, 5.60, 7.40	90, 90, 90	8
polypropylene (isotactic)	2*3/1	6.66, 20.78, 6.50	90, 99.6, 90	12
PET	12*1/1	4.56, 5.96, 10.75	98.5, 118, 112	
Polyisoprene (cis 1,4)	8*1/1	12.46, 8.86, 8.1	90, 90, 90	8

Exercise: Calculate the density of perfectly crystalline polyethylene, given assuming the crystal structure from the previous page.
Solution: Density is mass/volume, so we just need to calculate the volume of a unit cell, and the total mass of material included in the unit cell:
1. Unit cell volume:$a \cdot b \cdot c$=9.341 $\times 10^{-23}$ cm^3
2. Mass: 14 g/mole =2.326 $\times 10^{-23}$ g perCH_2 unit. Since there are 4CH_2 units in a unit cell, mass per unit cell =4(2.326 $\times 10^{-23}$ g) $= 9.302 \times 10^{-23}$ g.
3. Density = mass/volume = 0.996 g/cm^3.
Note that the density is close to 1 g/cm^3, which is generally true for most polymers.

93.5 Stiffness of Polymer Single Crystals

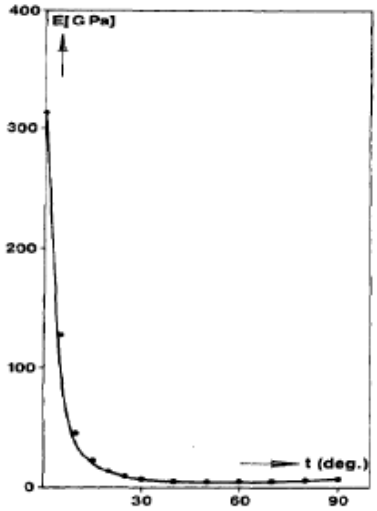


Figure 93.5: Macromol. 23, 2365-2370 (1990).

Single crystal polyethylene has the following calculated values for the modulus along the 3 crystal axes:

- E=315 GPa along c axis
- E=8.0 GPa along a axis
- E=9.9 GPa along b axis

Amorphous fraction: E=0.001 GPa
The actual modulus is almost always much lower than this because of imperfect crsytal alignment, chain folding and the presence of an amorphous fraction. See the following pages for details.

93.6 Chain Folding

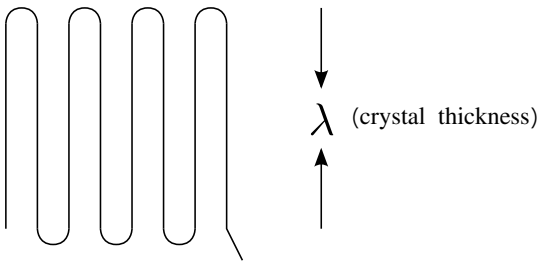


Figure 93.6: Diagram of polymer chain folding

The solid line in this animation represents the axis of a helix. As the polymer crystallizes, the helix folds back and forth on itself, producing a flat, plate-like lamellar crystallite. The chain folds are energetically unfavorable - the configuration with the lowest free energy would be a for the helices to pack in a completely extended manner, with no chain folds at all. However, the chain-folded structure is favored kinetically. Thin crystallites (low values of thecrystal thickness,λ), are able to form much more quickly than crystallites with high values ofλ.

93.6.1 Models of chain folding

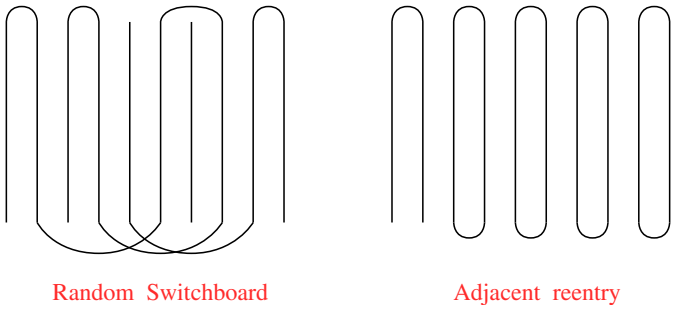


Figure 93.7: Extreme views of chain folded crystals

These drawings illustrate two extreme views of the structure of chain folded crystals. In the adjacent reentry model, the molecules are assumed to fold back on themselves in a very well-organized way. In the random switchboard model, chains which leave the crystal at the tops and bottom surfaces reenter the crystal at random positons. Reality lies somewhere between these two extremes.

93.6.2 Amorphous Fraction and Tie Molecules

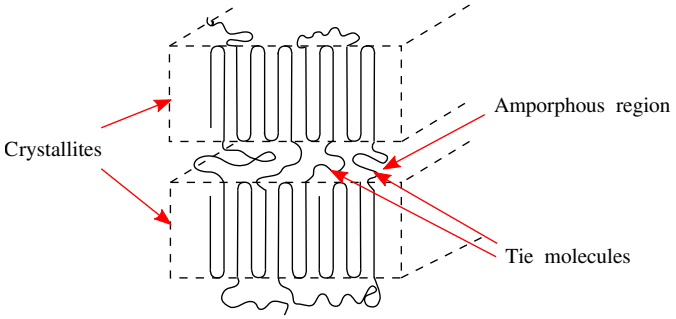


Figure 93.8: Lamellar crystallites

Isolatedlamellar crystallites can be obtained from polymers crystallized from dilute solution, but melt-crystallized polymers often have structures similar to what is illustrated here. The lamellar, platelike crystallites are separated by amorphous regions. Individual crystallites are bridged bytie molecules. These tie molecules span the amorphous region, and are incorporated into adjacent crystallites.

93.6.3 Percent Crystallinity for Lamellar Crystallites

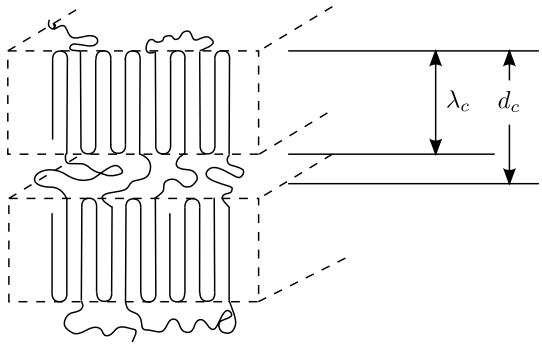


Figure 93.9: Diagram of percent crystallinity of lamellar crystallites

The amorphous fraction in a semicrystalline polymer resides between the lamellar crystallites as depicted here. If the crystallites are assumed to be perfectly crystalline, the percent crystallinity is given by the ratio of thecrystal thickness (λ_c) to thecrystal period (d_c). The crystal period is determined from x-ray scattering measurements. The overall percent crystallintity is determined by measuring the density of the sample.

93.7 Spherulitic morphologies

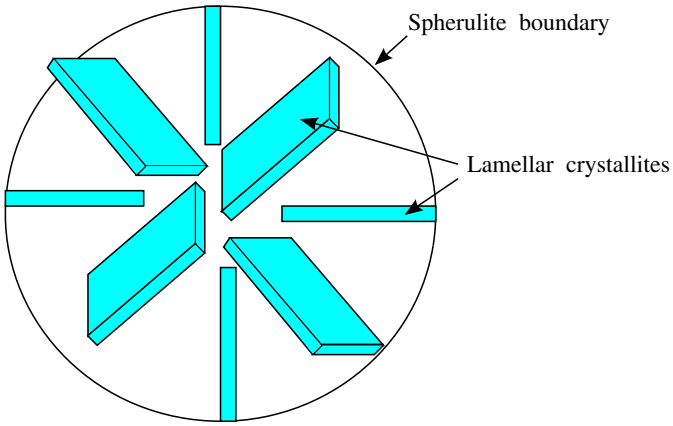


Figure 93.10: A representation of a spherulite

Aspherulite originates from the radial growth of lamellar crystallites from a central nucleation point. Not all semicrystalline polymers have spherulitic morphologies, but this morphology is quite common.

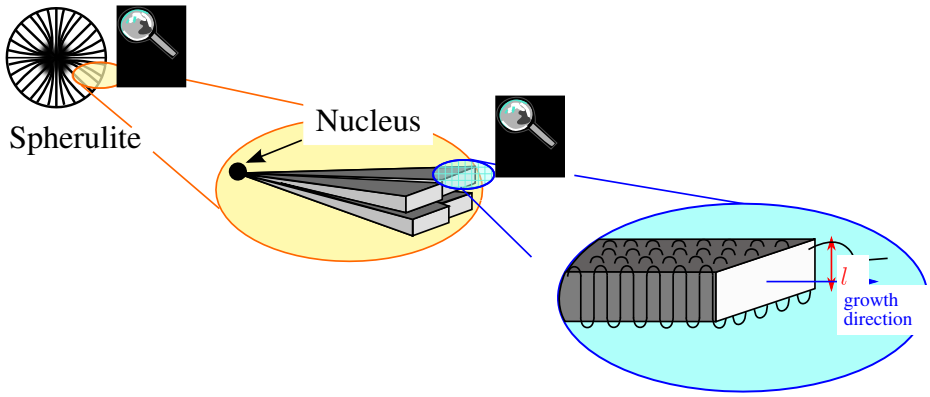


Figure 93.11: Formation of a spherulite

During crystallization, the spherulite grows radially outward, with each branch extending in the growth direction while maintinaing a constant crystal thickness,λ_c.

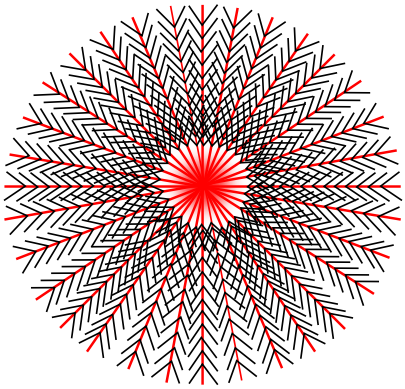


Figure 93.12: Geometric representation of a spherulite

This is the geometric requirement for addition of new lamellae; the additional space taken up by the spherulite as a whole is filled by new lamellae.

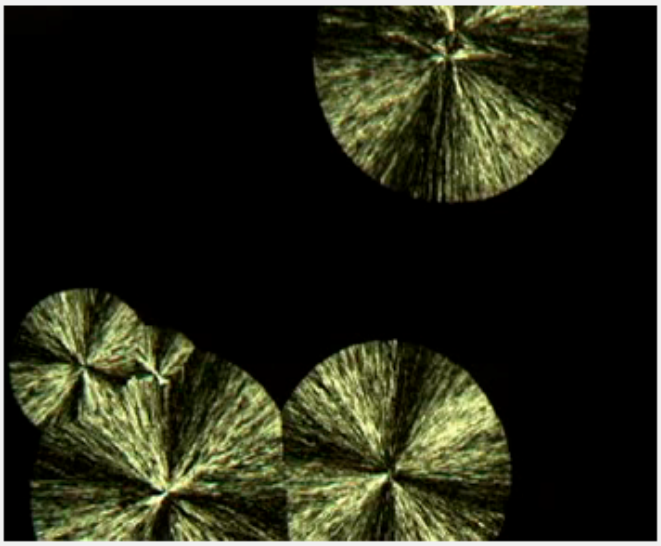


Figure 93.13: Video of the formation of a spherulite

93.8 Birefringence

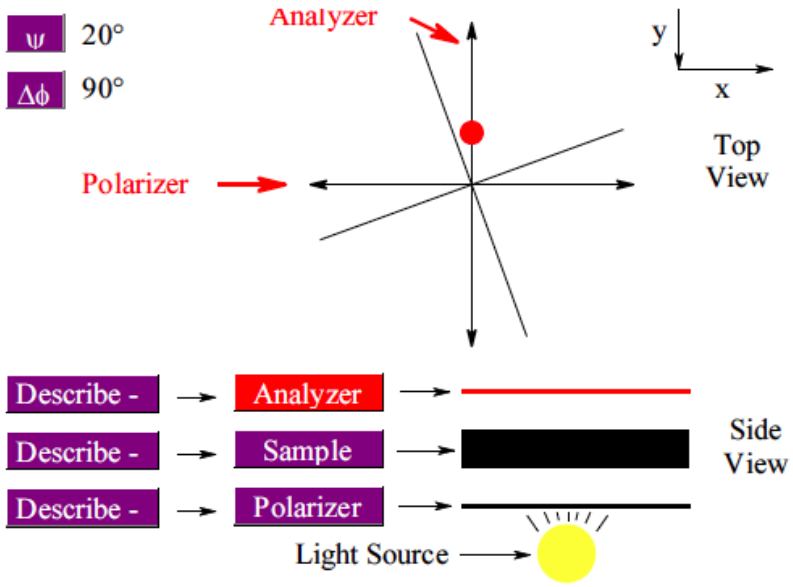


Figure 93.14: Interactive model of birefringence

(1) Light emerging from the sample for$\psi = 45^{o}$ and$\Delta\phi = 90^{o}$ is said to be circularly polarized. Why is this a sensible description of the polarization state you observe under these circumstances?

(2) For what values ofψ is no light at all transmitted through the analyzer?

(3) The maximum light transmission through the analyzer is obtained for$\varphi = 45^{o}$ and$\Delta\phi = 180^{o}$. What fraction of the light emerging from the sample is transmitted under these circumstances?

93.8.1 Birefringence and Radial Symmetry

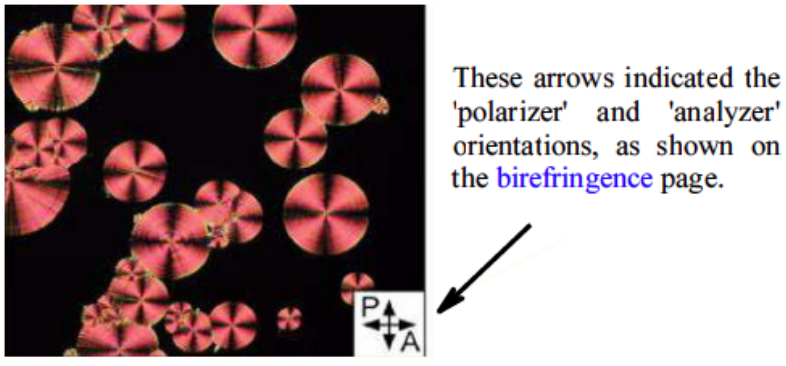


Figure 93.15: Diagram showing birefringence

http://bly.colorado.edu/lcphysics/textures/

If the object between the polarizer and analyzer has spherical symmetry, as is the case with these liquid crystalline droplets, the'Maltesecross' pattern shown here will be obtained. Dark patches correspond toΨ= 0 and 90°, and the bright patches are obtained atΨ= ± 45°.

93.9 Growth of a Lamellar Crystallite

The overall free energy change (ΔF) associated with the formation of a small crystallite has two components:

(1) A surface energy component, which is always positive.

(2) A bulk free energy component, which is negative for temperatures below the equilibrium melting temperature,T_m^0.

The bulk component can be written in terms of the undercooling ($T_m^0 - T$) by writing the bulk free energy change per unit volume (Δf_c) in terms of its entropic and enthalpic components:

$$\Delta f_c = \Delta h_c - T\Delta s_c \tag{93.1}$$

HereΔh_c andΔs_c are the enthalpy and entropy of melting for a perfect crystal, normalized by the volume of the crystal. We can use the fact that$\Delta f_c = 0$ when$T = T_m^0$ to eliminateΔs_c from the equation and writeΔf_c in terms of the enthalpy of melting and the the undercooling,ΔT, defined as$T_m^0 - T$:

$$\Delta f_c = -\frac{\Delta h_c \Delta T}{T_m} \tag{93.2}$$

Note that we have used the following sign conventions (as is generally the case):

- Δh_c is the enthalpy required to melt the crystal, and is always a positive number.
- From its definition,ΔT is positive whenever the temperature is below the equilibrium melting temperature.
- Δf_c is the free energy difference between the crystal and the amorphous material, and is negative for temperatures below the equilibrium melting temperature.

The growth of a lamellar crystallite occurs by the addition of molecular strands (typically corresponding to the c axis of the cyrstalline structure) to the growing crystallite as illustrated in Figure93.16. In addition to the

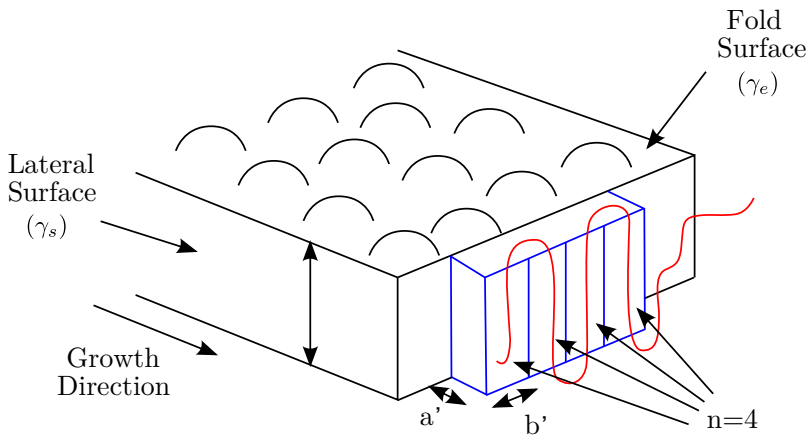


Figure 93.16: Schematic representation of a growing lamellar crystallite.

If we account for the surface energies associated with the new crystal/amorphous interface that is created, the increase in free energy for adding$n$ "strands" to the face of a lamellar crystallite is given by the following expression:

$$\Delta F = 2a'\lambda_c\gamma_s + 2na'b'\gamma_e + na'b'\lambda_c\Delta f_c \tag{93.3}$$

If we assume that$n$ is large we can neglect the term involvingγ_s, so that$\Delta F < 0$ when$\Delta f_c < \frac{2\gamma_e}{\lambda_c}$. With$\Delta f_c$ given by Eq.93.2, we can rewrite the condition that$\Delta F < 0$ in the following form:

$$\lambda_c > \frac{2\gamma_e T_m^0}{\Delta h_c \Delta T} \tag{93.4}$$

A crystal will grow by the addition of additional strands when this criterion is met. Thick crystals (highλ_c) are thermodynamically favorable, but thin crystals (lowλ_c) are kinetically favorable. In this model, the crystal thickness which is actually obtained is the lowest, thermodynamically possible value. The inequality from the previous page therefore becomes an approximate equality:

$$\lambda_c \approx \frac{2\gamma_e T_m^0}{\Delta h \Delta T} \tag{93.5}$$

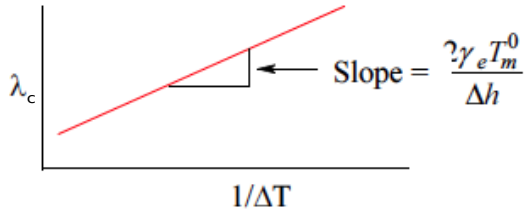


Figure 93.17: Graphical representation of the crystal thickness prediction from Eq.93.5.

Because thick crystals are thermodynamically more stable than thin crystals, the thick crystals will melt at a slightly higher temperature than thin crystals. One therefore expects a relationship between the temperature at which the crystals will form, and the temperature at which the crystals will melt. Crystals crystallized at higher temperatures will also melt at higher temperatures. At an undercooling of zero the actual melting temperature,T_m, will be equal to the equilibrium melting temperature. For very low undercoolings, however, the crystallization rate becomes very slow, so crystallization will never take place at the exact equilibrium melting temperature. Values of the equilibrium melting temperature are obtained from measurements ofT_m at higher undercoolings, and extrapoloating to$\Delta T = 0$ as illustrated in the following graph:

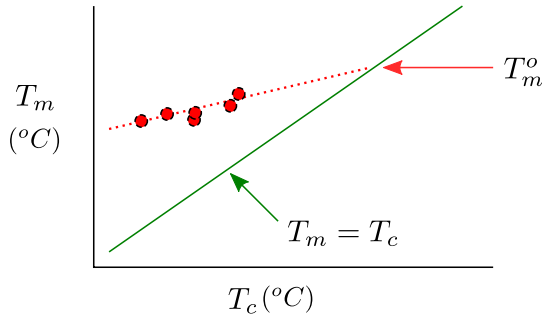


Figure 93.18: Melting temp (T_m) as a function of the crystallization temperature,T_c.

93.10 Density of Semicrystalline Polymers

The percent crystallinity in a polymer sample is generally determined by measuring the density of the material. A linear relationship between density and percent crystallinity is assumed, so that the following relationship holds:

$$\%Crystallinity = \frac{\rho - \rho_{amorphous}}{\rho_{crystalline} - \rho_{amorphous}} x100 \tag{93.6}$$

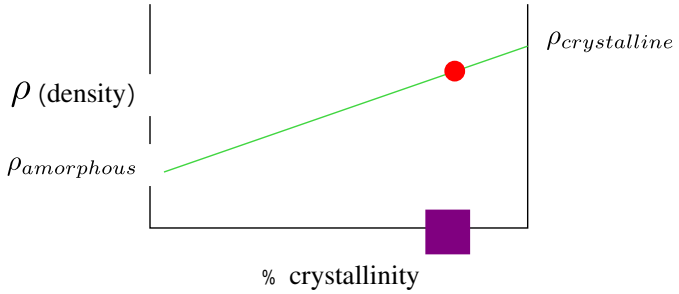


Figure 93.19: Density vs percent crystallinity

94 Liquid Crystals

Liquid crystals are materials that that are intermediate between true traditional crystalline materials, and liquids. The molecules in a liquid crystalline phase are not organized into a full, 3- dimensionally periodic structure, but they do have some degree of orientational or positional order. Examples of molecular organization in some liquid crystalline phases are shown below.

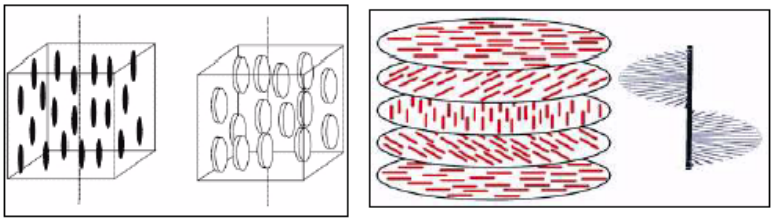


Figure 94.1: Structure of liquid crystals

http://commons.wikimedia.org/wiki/Liquid_crystal

94.1 Nematic Liquid Crystals

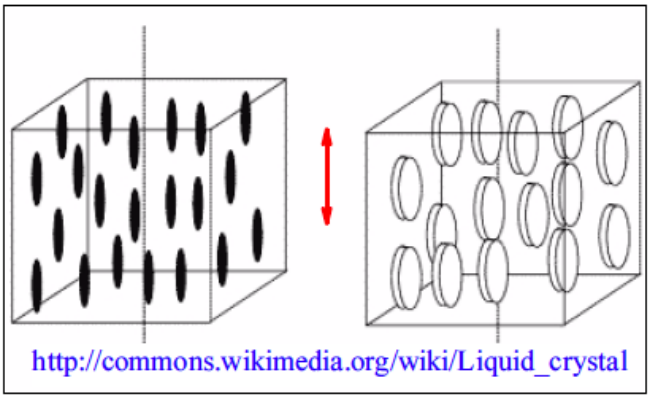


Figure 94.2: Structure of nematic liquid crystals

In anematic liquid crystal phase, the centers of the molecules have liquid-like order, but the molecules are aligned in the same direction, defined by thedirector, represented for these needleshaped and disc-shaped molecules by the red arrow.

94.2 Nematic Liquid Crystals - Texture

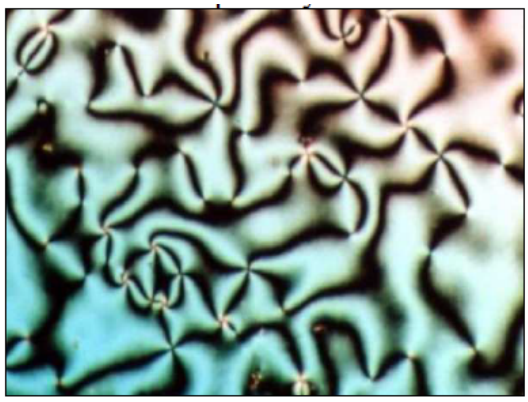


Figure 94.3: Image of a nematic liquid crystal

http://commons.wikimedia.org/wiki/Liquid_crystal

This photograph was taken between crossed polarizers, uing a microscope geometry like that described in the description ofBirefringence.

94.3 Liquid Crystal Displays

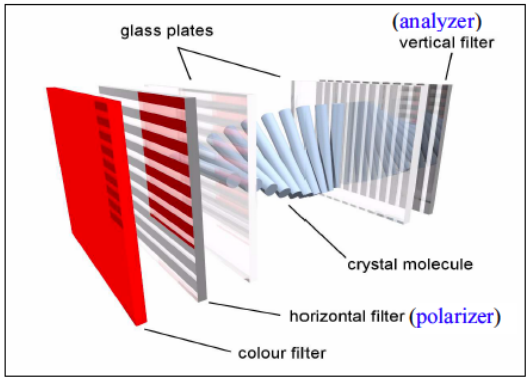


Figure 94.4: Structure of an LCD

http://en.wikipedia.org/wiki/Liquid_crystal_display

95 Solutions and Blends

We now move from our discussion of the shapes of individual polymer molecules in a homogeneous, one component system, to the behavior of solutions and blends consisting of different types of molecules that have been mixed together. We are interested in studying these systems for many reasons, including the following:

(1) Characterization: many of the properties of polymer molecules are determined by the properties of polymer solutions.

(2) Processing: many applications of polymers involve require that the polymer molecules be dissolved in an appropriate solvent.

(3) Improved materials: new materials with improved properties can often be formed by blending or "alloying" different polymers.

(4) Recycling: polymers need to be separated before they can be recycled because the different polymers do not mix favorably.

The key quantity when discussing thermodynamics of polymer solutions (polymer + small-molecule solvent) and polymer blends (polymer A + polymer B) is the free energy of mixing. We therefore begin with a discussion of polymer solution thermodynamics. In general, we consider the case where we mix n_a A molecules with a degree of polymerization N_a with n_b B molecules with a degree of polymerization N_b. Polymer solutions will correspond to the case where one of the degrees of polymerization is very small (typically 1). We will make the assumption that the sizes of the repeat units for the components of the mixture are identical. This assumption is not as restrictive as it would initially appear, since one can always define "effective" repeat units with the desired volume. In this sense, N_a and N_b are determined by the relative molecular volumes of the components in the mixture, and are defined in terms of some reference volume.

Exercise: The room temperature densities of polystyrene and toluene are 1.05g/cm^3 and 0.87g/cm^3, respectively. What reference volume, V_0, should we use to characterize the thermodynamics of polystyrene solutions in toluene? What degrees of polymerization would characterize a solution where the polystyrene molecular weight is 100,000 g/mole?
Solution: The obvious reference volume in this case is the volume of a toluene molecule, which we take as having N=1. This volume is obtained from the density and molecular weight of toluene. Toluene (C_6H_8) has a molecular weight of 92 g/mole.

$$V_0 = \left(\frac{92\,\mathrm{g}}{\mathrm{mole}}\right)\left(\frac{\mathrm{cm}^3}{0.87\,\mathrm{g}}\right) = 106\,\mathrm{cm}^3/\mathrm{mole}$$

Note that we have used the molar volume for V_0. This convention will be followed throughout our discussion. The value of N for the polymer molecules will be given by the molecular volume, divided by the reference volume:

$$N_{polystyrene} = \left(\frac{100,000\,\mathrm{g}}{\mathrm{mole}}\right)\left(\frac{\mathrm{cm}^3}{1.05\,\mathrm{g}}\right)\left(\frac{\mathrm{mole}}{106\,\mathrm{cm}^3}\right) = 898$$

In our discussion of the thermodynamics of mixtures, we will refer to N as the number of "segments", defined in terms of a reference volume as described here. In some cases this reference volume will correspond to a repeat unit of the polymer, but this does not need to be the case in general.

95.1 Chemical Potentials and Free Energy of Mixing

The chemical potential of A molecules, μ_a, is defined as the free energy required to add an additional 'A' molecule to the system. Because all A molecules are equivalent, it is also equal to the free energy per A molecule in the system. The following expressions are helpful

$$F = n_a\mu_a + n_b\mu_b \tag{95.1}$$

$$\begin{aligned}\left.\frac{\partial F}{\partial n_a}\right|_{n_b} &= \mu_a \\ \left.\frac{\partial F}{\partial n_b}\right|_{n_a} &= \mu_b\end{aligned} \tag{95.2}$$

The derivative with respect to n_a is taken with n_b held constant, and vice versa. Note that F is the total extensive free energy of the system, since it increases as the numbers of A and B molecules (n_a and n_b) increases.

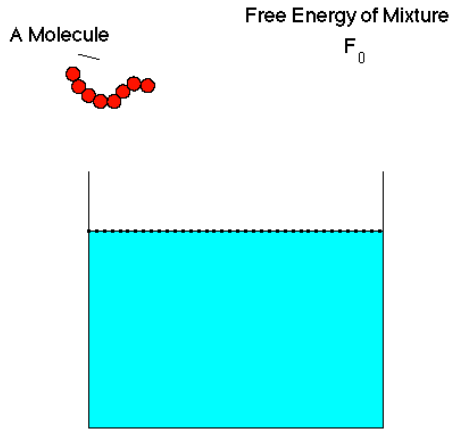


Figure 95.1: Animation of the chemical potential of a molecule

The chemical potentials of A and B molecules of the pure components are defined asu_a^0 andu_b^0. The free energy of the two components before mixing is therefore given by:

$$F_{unmixed} = n_a\mu_a^0 + n_b\mu_b^0 \tag{95.3}$$

The free energy of mixing is equal to the difference in the free energies of the mixed and unmixed systems:

$$\Delta F_{mix} = F - F_{unmixed} = n_a\mu_a + n_b\mu_b - (n_a\mu_a^0 + n_b\mu_b^0) = n_a(\mu_a - \mu_a^0) + n_b(\mu_b - \mu_b^0) \tag{95.4}$$

Because we are only interested in changes in chemical potentials due to mixing, we can set the chemical potentials of the pure, unmixed components to zero ($\mu_a^0 = \mu_b^0 = 0$).
By defining the chemical potentials of the pure components as zero, we obtain:

$$\Delta F_{mix} = n_a\mu_a + n_b\mu_b \tag{95.5}$$

Equation95.5 is completely general, and valid for mixtures of molecules of any size. It's not that useful, however, because we generally don't write down values ofn_a andn_b when we want to specify the composition of a multicomponent mixture. For mixtures of small molecules (metal alloys, for example), we commonly use mole fractions,X_a andX_b, to specify the composition:

$$\begin{aligned} X_a &= \frac{n_a}{n_a+n_b} \\ X_b &= \frac{n_b}{n_a+n_b} \end{aligned} \tag{95.6}$$

This makes sense for metals because the molar volumes of the different components are not that different from one another. In polymer systems, however, we are often dealing with systems where the components have molar volumes that can differ by orders of magnitude. This is the situation if we are interested in the solubility of a high molecular weight polymer (which can have a molecular weight of several hundred thousand g/mole) in a solvent (which will typically have a molecular weight of$\approx$ 100 g/mole). Organic solutions are typically specified by the weight fractions of the different components. In discussions of polymer solution thermodynamics it is more common to work in terms of volume fractions of the different components in the systems.
The concepts in this section can be easily extended to systems with more than two components, but we confine ourselves here to binary solutions of just two components: A molecules of lengthN_a and B molecules of lengthN_b.
In our discussion of polymer thermodynamics it is important to keep in mind$N_a N_b$ are not necessarily true degrees of polymerization as defined earlier in this text. Instead they are obtained by dividing the molar volumes of the different components by some reference volume,V_0. For polymer/solvent mixtures this reference volume is typically taken as the molar volume of the solvent, and for polymer/polymer mixtures it is typically taken as the

volume per repeat unit for one of the polymers. The total volume,V, and volume fractions,ϕ_a andϕ_b are given by the following expressions:

$$V = n_a N_a V_0 + n_b N_b V_0 \tag{95.7}$$

$$\begin{aligned} \phi_a &= \frac{n_a N_a V_0}{V} \\ \phi_b &= \frac{n_b N_b V_0}{V} \end{aligned} \tag{95.8}$$

The free energy of mixing per unit volume,Δf_{mix}, is obtained fromΔF_{mix} by dividing by the volume of the system:

$$\Delta f_{mix} = \frac{n_a \mu_a + n_b \mu_b}{n_a N_a V_0 + n_b N_b V_0} = \frac{\mu_a \phi_a}{N_a V_0} + \frac{\mu_b \phi_b}{N_b V_0} \tag{95.9}$$

By rearranging the equation at the bottom of the previous page we obtain:

$$\Delta f_{mix} V_0 = \frac{\mu_a \phi_a}{N_a} + \frac{\mu_b \phi_b}{N_b} \tag{95.10}$$

The quantitiesΔf_{mix},μ_a / N_a andμ_b / N_b all depend on the composition of the material. The relationship between these quantities is illustrated by the tangent construction shown in Figure95.2, where the composition dependence ofΔf_{mix} is plotted as a function ofϕ_b for a hypothetical material. For a given composition (denotedϕ_b' in Figure95.2), the chemcial potentials are obtained by first drawing a tangent to the free energy curve at$\phi_b = \phi_b'$. The value ofμ_a / N_a is determined by extrapolating this tangent line to$\phi_b = 0$ andμ_a / N_a is obtained by extrapolating the tangent line to$\phi_b = 1$. This result is consistent with Eq.95.10, which shows that normalized free energy of mixing at a certain composition is determined by the weighted average of the normalized chemical potentials. As a result a straight line drawn between the normalized chemical potentials on the left and right axes must hit the free energy curve at$\phi_b = \phi_b'$.

To fully validate the tangent construction illustrated in Figure95.2 we still need to show that the straight line between the chemical potentials must be equal to the slope of the free energy curve. This can done in formal terms by combining Eqs.95.2 and95.8, which after simplification leads to the following relationship:

$$\frac{\partial(\Delta f_{mix} V_0)}{\partial \phi_b} = \frac{\mu_b}{N_b} - \frac{\mu_a}{N_a} = -\frac{\partial(\Delta f_{mix} V_0)}{\partial \phi_a} \tag{95.11}$$

The fact that the tangent must give the difference betweenμ_b / N_b andμ_a / N_b makes sense conceptually, because the only way changing the composition by some amount requires that an A segment be replaced by a B segment.

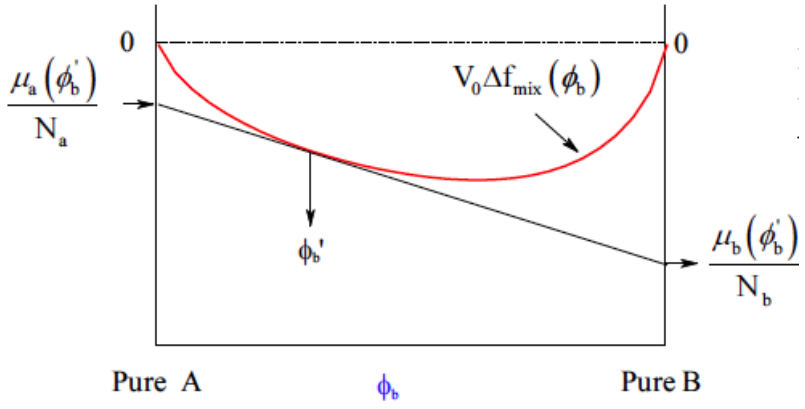


Figure 95.2: Relationship of the free energy of mixing to the the chemical potentials.

95.2 Ideal Entropy of Mixing for Polymers

For small molecules the ideal free energy of mixing is given by the following expression:

$$\frac{\Delta s_{ideal} V_0}{R} = -X_a \ln X_b - X_b \ln X_b \tag{95.12}$$

The quantity$s_{ideal} V_0$ is the entropy of mixing per mole of molecules. If the molar volumes of the A and B molecules are the same, than we can simply replace the mole fractions,X_a andX_b with the corresponding volume fractions,ϕ_a andϕ_b. The entropy of mixing per molecule is independent of the size of the molecule. Therefore, the entropy of mixing per repeat unit is smaller than the entropy of mixing per molecule by a factor of$N$, where$N$ is the degree of polymerization. In general, we have the following expression for the ideal free energy of mixing per volume for polymers with molar volumes of$N_a V_0$ and$N_b V_0$:

$$\frac{s_{ideal} V_0}{R} = -\frac{\phi_a \ln \phi_a}{N_a} - \frac{\phi_b \ln \phi_b}{N_b} \tag{95.13}$$

95.3 Idealized Enthalpy of Mixing

The key result is that the entropy of mixing for a given volume is much smaller for large molecules than it is for small molecules. This result explains why different types of polymers almost never mix with one another in the liquid state. The driving force for two liquids to mix is normally the entropy of mixing, but this driving force is very low for mixtures of very large polymer molecules. In fact, the entropy of mixing for polymer mixtures is sometimes negative. While the ideal entropy of mixing must be positive, it is very small, and can be overwhelmeed by non-ideal contributions to the entropy of mixing which can be positive or negative. Before discussing these effects, however, we will continue in our derivation of the Flory Huggins equation for the free energy of mixing. The Flory-Huggins theory combines the ideal free energy of mixing with a very simple form for the enthalpy of mixing. Unlike the entropy of mixing, the enthalpy of mixing is assumed to be independent of the size of the molecules. We can therefore use small molecules to illustrate the origins of the equation for the enthalpy of mixing.

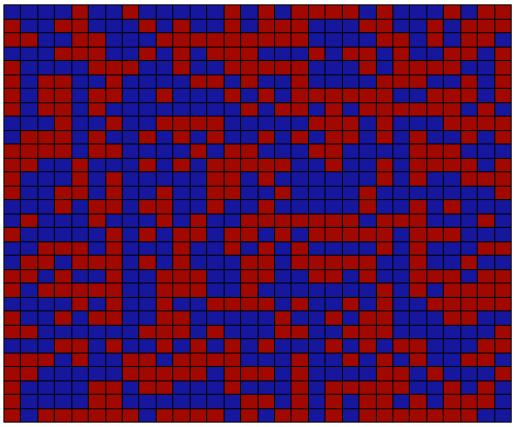


Figure 95.3: An animated simulation of mixing

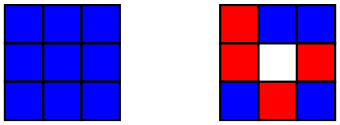


Figure 95.4: Pictoral representations of mixing

In this example, the blue square is placed into an environment where one of its four nearest neighbors is blue, and three are red. Red-blue contacts have an energy (or enthalpy) ofE_{ab}, red-red contacts have an energy ofE_{aa}, and blue-blue contacts have an energy ofE_{bb}. The enthalpy change,ΔH, associated with the removal of the square from a pure blue phase on the left (4 blue neighbors) to the mixture on the right (1 blue neighbor and 3 red neighbors) is therefore given by the following formula:

$$\Delta H = 3E_{ab} + E_{bb} - 4E_{bb} = 3(E_{ab} - E_{bb}) \tag{95.14}$$

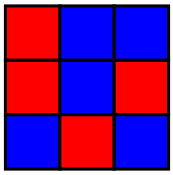


Figure 95.5: Diagram showing contacts

How do we calculate the energy of a randomly mixed collection of A and B molecules? We begin by assuming that the energy of an A-A contact isE_{aa}, the energy of an B-B contact isE_{bb}, and the energy of an A-B contact isE_{ab}. Also assume that each molecule$z$ nearest neighbors ($z = 4$ for the 2-dimensional array of squares shown above.) The number of A-B contacts in a mixture ofn_a A molecules andn_b B molecules is given by:

$$\begin{aligned} n_{aa} &= n_a z p_a/2 \\ n_{bb} &= n_b z p_b/2 \\ n_{ab} &= n_a z p_b \end{aligned} \tag{95.15}$$

wherep_a is the probability that a nearest neighbor is an A molecule. The factor of 2 in the expressions forn_{aa} andn_{bb} is needed to avoid double counting. (For example, ifn_a =2, there is only one interaction between the two molecules, not 2.) Now we invoke the mean field (or random mixing) approximation, which is that$p_a = \phi_a$, and$p_b = \phi_b$. We obtain the follwing.

$$\begin{aligned} n_{aa} &= n_a z \phi_a/2 \\ n_{bb} &= n_b z \phi_b/2 \\ n_{ab} &= n_a z \phi_b \end{aligned} \tag{95.16}$$

The energy of the mixed state is obtained by multiplying the numbers of different contacts by their energy:

$$E_{mixed} = E_{aa} n_{aa} + E_{bb} n_{bb} + E_{ab} n_{ab} \tag{95.17}$$

Substitution forn_{aa},n_{bb} andn_{ab} gives the following:

$$E_{mixed} = E_{aa} n_a z \phi_a/2 + E_{bb} n_b z \phi_b/2 + E_{ab} n_a z \phi_b \tag{95.18}$$

The energy of the unmixed state is obtained in a similar fashion, remembering that in this state all A's are surrounded by other A's, and all B's are surrounded by other B's:

$$E_{unmixed} = E_{aa} n_a z/2 + E_{bb} n_b z/2 \tag{95.19}$$

The energy change due to mixing is obtained by subtracting the energy of the mixed state from the energy of the mixed state:

$$\Delta E_{mix} = E_{mixed} - E_{unmixed} = \frac{E_{aa} n_a z \phi_a}{2} + \frac{E_{bb} n_b z \phi_b}{2} + E_{ab} n_a z \phi_b - \frac{E_{aa} n_a z}{2} - \frac{E_{bb} n_b z}{2} \tag{95.20}$$

This expression can be further simplified by recalling that$n_a = \phi_a n$ and$n_b = \phi_b n$, where$n$ is the total number of molecules ($n = n_a + n_b$):

$$\Delta E_{mix} = nz \left[\frac{E_{aa} \phi_a^2}{2} + \frac{E_{bb} \phi_b^2}{2} + E_{ab} \phi_a \phi_b - \frac{E_{aa} \phi_a}{2} - \frac{E_{bb} \phi_b}{2} \right] \tag{95.21}$$

After some rearrangement we obtain the following expression:

$$\Delta E_{mix} = nz \phi_a \phi_b [E_{ab} - \frac{E_{aa} + E_{bb}}{2}] \tag{95.22}$$

Note that the mixing energy is positive if the energy of a-b contacts is larger than the average energy of a-a and b-b contacts. Because z,E_{aa},E_{bb}, andE_{ab} are actually not measurable quantities, it is common to lump them together into a parameter called theChi parameter represented by the Greek letter 'chi':

$$\chi \equiv \frac{z}{k_B T} \left[E_{ab} - \frac{E_{aa} + E_{bb}}{2} \right] \tag{95.23}$$

With this definition ofχ, the energy of mixing for A and B molecules (with the same size) is:

$$\Delta H_{mix} = \Delta E_{mix} = RT \chi n \phi_a \phi_b \tag{95.24}$$

Strictly speaking, the 'enthalpy' and the 'energy' of a system are not the same thing. The Gibbs free energy involves the enthalpy, and is the appropriate thermodynamic quantity at a fixed temperature and pressure. The Helmholtz free energy involves the energy, and is the appropriate thermodynamic quantity at fixed temperature and volume. For compressible systems the two quantities differ slightly. We use the expression from the previous page for the extensive enthalpy of mixing, and divide by the volume of the system(nV_0) to obtain the intensive enthalpy of mixing:

$$\Delta h_{mix} = \frac{\Delta H_{mix}}{nV_0} = \frac{RT \chi \phi_a \phi_b}{v_0} \tag{95.25}$$

Normalizing as we did for the ideal entropy of mixing gives:

$$\frac{\Delta h_{mix} V_0}{RT} = \chi \phi_a \phi_b \tag{95.26}$$

95.4 Flory-Huggins Free Energy of Mixing

The most commonly used expression used to describe the free energy of mixing for polymer/polymer and polymer/solvent mixtures combines mean-field enthalpy of mixing with the ideal combinatorial entropy of mixing:

$$\Delta f_{mix} = \Delta h_{mix} - T \Delta s_{ideal} \tag{95.27}$$

Combining these forms for the enthalpy and entropy gives the following expression, referred to as the Flory-Huggins free energy of mixing:

$$\frac{\Delta f_{mix} V_0}{RT} = \frac{\phi_a \ln \phi}{N_a} + \frac{\phi_b \ln \phi_b}{N_b} + \chi \phi_a \phi_b \tag{95.28}$$

It is important to recognize that many assumptions went into the derivation of this equation, and that they can't possibly all be true. Nevertheless, this equation is very commonly used in polymer science, and needs to be understood. In fact, this equation is often used as a definition forχ. Based on our earlier definition ofχ in terms of energetic parameters, one would expect that it's magnitude would decrease as the temperature increases. In fact, the actual free energy of mixing can have very different behavior, and the free energy of mixing is generally accurately described by the Flory- Huggins expression only whenχ is allowed to vary with temperature and composition. We begin on the next page by constructing phase diagrams, with the assumption that the free energy of mixing is adequately described by a value ofχ that is independent of composition. Free energy curves for$N_a = N_b = 1$ are shown in Figure95.6 forχ ranging from 0 to 4. For$N_a = N_b = 1$, the critical value ofχ is equal to 2. Whenχ is larger than this critical value there will be a region of compositions where the free energy curve has a negative curvature in the free energy curve.

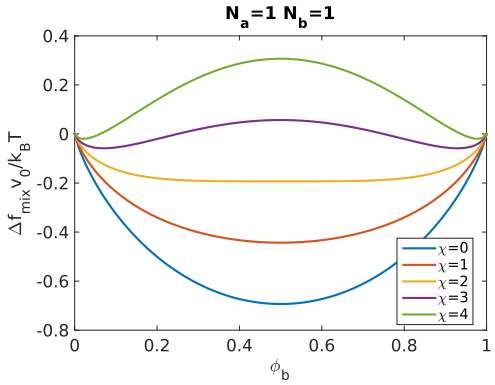


Figure 95.6: Regular solution form or for the free energy of mixing of small molecules.

95.5 The Coexistence (binodal) and Spinodal Curves

The critical value ofχ decreases asN_a andN_b increase. In addition, the free energy curves are no longer symmetric about$\phi_b = 0.5$ when$N_a \neq N_b$. Both of these points are illustrated in Figure95.7, where we plot the Flory-Huggins free energy of mixing (Eq.95.28) for$N_a = 200$,$N_b = 500$ andχ =0.012. A two-phase region will exist for average compositions of the alloy that are between the compositions betweenϕ_1 andϕ_2, the two compositions defined by the common tangent construction in Figure95.7. Because the tangent to the curve at these two compositions are identical, they have the same values ofμ_a and$\mu_{b,}$ which is a requirement for equilibrium. In addition, the average free energy for a mixture of these two phases (dashed line in Figure95.7) is less than the free energy for a homogenous phase of this average composition (solid line in Figure95.7). By plotting these two composition as a function ofχ, we obtain the binodal curve shown as the solid line in Figure95.8. The binodal curve is also referred to as the coexistence curve because it shows the compositions of the phases that are in equilibrium with one another in the two-phase region of the phase diagram.

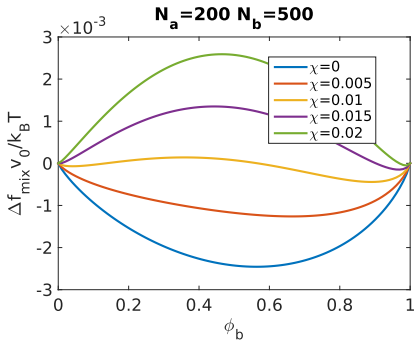


Figure 95.7: Free energy of mixing for a system with asymmetric molecular sizes in the two phase regime.

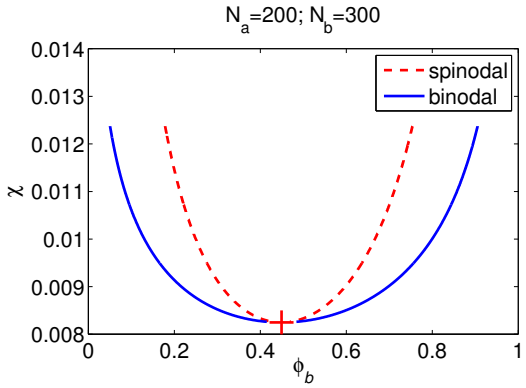


Figure 95.8: Binodal and spinodal curves for$N_a = 200$ and$N_b = 300$.

The spinodal curve (or stability limit) represents the boundary in the phase diagram where the the free energy of mixing changes from a positive curvature to a negative curvature. It occurs when$\chi = \chi_s$, whereχ_s defines the spinodal curve. For$N_a = 200$ and$N_b = 300$ the spinodal curve is shown as the dashed line in Figure95.8. The change in curvature from positive to negative values occurs when the the second derivative of the free energy with respect to composition is equal to zero. We obtain the following expression forχ_s by setting$\phi_b = 1 - \phi_a$ in the free energy expression:

$$\frac{\partial^2}{\partial \phi_a^2}\left(\frac{\phi_a \ln \phi_a}{N_a} + \frac{(1-\phi_a)\,(\ln 1 - \phi_a)}{N_b} + \chi_s \phi_a\,(1-\phi_a)\right) = 0 \tag{95.29}$$

Evaluating the second derivative leads to the following:

$$\frac{1}{\phi_a N_a} + \frac{1}{(1-\phi_a)N_b} - 2\chi_s = 0 \tag{95.30}$$

which can be rearranged to give the following forχ_s:

$$\chi_s = \frac{1}{2\phi_a N_a} + \frac{1}{2\phi_b N_b} \tag{95.31}$$

95.6 Critical Point

The phase diagrams obtained from the free energy of mixing expression show under what conditions polymeric liquids are completely miscible, and under what conditions phase separation will occur. The critical point is the point on the phase diagram where a liquid is just beginning to undergo phase separation. Mathematically, the critical point corresponds to the point on the spinodal curve whereχ_s has the lowest possible value. We obtain this point by setting the derivative ofχ_s to zero:

$$\frac{\partial}{\partial \phi_a}\left(\frac{1}{2\phi_a N_a}+\frac{1}{2(1-\phi_a)N_b}\right)=0 \tag{95.32}$$

The solution to this equation is given by$\phi_a = \phi_{a,crit}$, at which pointχ_s is equal toχ_{crit}:

$$\phi_{a,crit}=\frac{\sqrt{N_b}}{\sqrt{N_b}+\sqrt{N_a}} \tag{95.33}$$

$$\chi_{crit}=\frac{N_a+N_b+2\sqrt{N_a N_b}}{2N_a N_b} \tag{95.34}$$

Exercise: The value of for blends of polystyrene (PS) and poly(methyl methacrylate) (PMMA) is 0.037 at 170 °C, based on a reference volume equal to the reference volume of a polystyrene repeat unit. Suppose monodisperse samples of PS and PMMA of equal molecular weights are mixed with one another. What range of molecular weights will be completely miscible with one another at 170 °C? (Assume the densities of the two polymers are identical).
Solution: First we calculateχ_{crit}:

$$\chi_{crit}=\frac{N_a+N_b+2\sqrt{N_a N_b}}{2N_a N_b}=\frac{2}{N} \tag{95.35}$$

for$N_a = N_b = N$

The polymers are completely miscible forχ<χ_{crit}, which occurs when N < $2/\chi$. Withχ= 0.037, we have complete miscibility for N < 54. The repeat unit molecular weight for polystyrene is 104 g/mol, so the critical molecular weight below which the polymers are completely miscible is 54*104 g/mol = 5616 g/mol

95.7 Spinodal Decomposition

"Spinodal Decomposition" is the mechanism of phase separation where small perturbations in the composition of the single phase grow because this decreases the overall free energy. This happens inside the spinodal curve defined previously. The microstructure is generally characterized by a well defined wavelenth, as illustrated below.

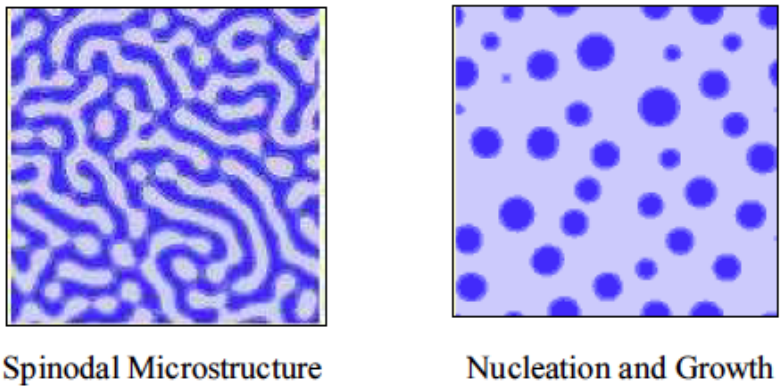


Figure 95.9: Characteristic microstructures for phase separation by spinodal decompositon (left) and nucleation and growth (right).

95.8 Phase Diagrams in Temperature Space

Actual values ofχ for real systems are generally obtained by fitting the Flory Huggins expression to the free energy of mixing to actual experimental data. This procedure can give actual values ofχ which either increase or decreaes with increasing temperature. Smaller values ofχ (where mixing is favored) can therefore be obtained at either high or low temperatures, giving the following two types of phase diagrams:

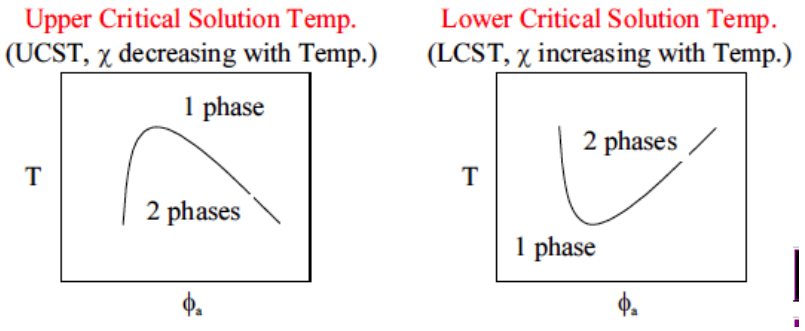


Figure 95.10: Phase diagrams in temperature space.

Exercise: The following equation describes the temperature dependence of χ for the PMMA/butanol polymer/-solvent system:

$$\chi = 1.45 - 0.0115T \tag{95.36}$$

Here T is expressed in°C. For what range of temperatures will high molecular weight PMMA completely dissolve in butanol?
Solution: From the phase diagram, we see that the two components are completely miscible for $\chi < \chi_{crit}$. If the A phase is the solvent ($N_a = 1$) we obtain the following for χ_{crit}:

$$\chi_{crit} = \frac{1 + N_b + 2\sqrt{N_b}}{2N_b} \tag{95.37}$$

As the polymer molecular weight becomes very large ($N_b \to \infty$), $\chi_{crit} \to 0.5$. The temperature at which χ is equal to this critical value of 0.5 for polymer solutions is called the theta temperature. It corresponds to the temperature at a theoretical polymer with an infinite molecular weight is just beginning to precipitate from solution. The general condition for complete miscibility for a polymer/solvent system, for the case where the polymer has a very high molecular weight, is given by the requirement that $\chi < 0.5$. For the data given for the PMMA/Butanol system, this occurs for $T > 82.6\,^{\circ}$C. Because χ decreases with increasing temperature, this system exhibits an upper critical solution temperature (UCST).
A cautionary note: Note that the polymer volume fraction at the critical point is very low (i.e., $\phi_{a,crit}$ is very close to 1 for $N_b >> N_a$). The detailed predictions of the Flory Huggins theory will not be accurate in this regime, for reasons that are discussed more fully in the section on dilute solution thermodynamics.

95.9 Chemical Potentials

Recall that the Chemical Potentials and Free Energy of Mixing of species A is determined by differentiating the extensive free energy with respect to the number of A molecules in the system. The extensive free energy is obtained by multiplying the free energy density, f, by the volume of the system, V. For a binary mixture of A and B polymers we have:

$$\Delta f_{mix} = \frac{RT}{v_0}\left\{\frac{\phi_a \ln \phi_a}{N_a} + \frac{\phi_b \ln \phi_b}{N_b} + \chi \phi_a \phi_b\right\} V \tag{95.38}$$

Using the expressions for ϕ_a and ϕ_b (Eq.95.8) and V (Eq.95.7) we have:

$$\frac{\Delta F_{mix}}{RT} = \{n_a \ln \phi_a + n_b \ln \phi_b + \chi n_a N_a \phi_b\} \tag{95.39}$$

We are now in a position to obtain expressions for the chemical potentials by differentiating the extensive free energy of mixing. We make the usual assumption that the chemical potentials of the pure components are equal to zero (activities of the pure components equal to one):

$$\mu_a = \frac{\partial \Delta F_{mix}}{\partial n_a} = RT \frac{\partial}{\partial n_a}\left(n_a \ln \phi_a + n_b \ln \phi_b + \chi n_a N_a \phi_b\right) \tag{95.40}$$

In this way we obtain the following result:

$$\frac{\mu_a}{RT} = \ln \phi_a + \phi_b\left(1 - \frac{N_a}{N_b}\right) + \chi N_a \phi_b^2 \tag{95.41}$$

When discussing polymer/solvent mixtures we replace the 'a' and 'b' subscripts with 'p' (for polymer) and 's' (for solvent). In addition, we take$N_s = 1$. Formally,N_p is the ratio of the polymer volume to the solvent volume. With these substitutions, we obtain the following expression forμ_s, the solvent chemical potential, which is valid for all concentrations:

$$\frac{\mu_s}{RT} = \ln(1-\phi_p) + \phi_p\left(1 - \frac{1}{N_p}\right) + \chi\phi_p^2 \tag{95.42}$$

95.9.1 Limiting case for high molecular weight polymer

For$N_p \to \infty$ the solvent chemical potential reduces toμ_s^∞:

$$\frac{\mu_s^\infty}{RT} = \ln(1-\phi_p) + \phi_p + \chi\phi_p^2 \tag{95.43}$$

In addition the critical point from Eqs.95.33 and95.34 occurs at$\phi_p = 0$ and$\chi = 0.5$. This means that for$\chi > 0.5$ pure solvent will be in equilibrium with a phase that contains both polymer and solvent. The polymer volume fraction in this phase is determined from the requirement that the solvent chemical potential must be zero. In other words, if a chunk of high molecular weight polymer is immersed in a solvent, solvent diffuses into the polymer untilϕ_p decreases to the point where$\mu_s^\infty = 0$. In general, this solution must be solved numerically. In Figure95.11 we plot the concentration dependence ofμ_s^∞ for$\chi = 0$,$\chi = 0.5$ and$\chi = 1$.

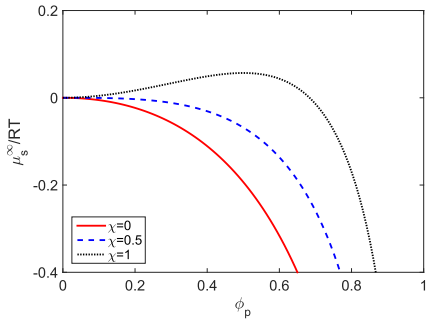


Figure 95.11: Concentration dependence of the solvent chemical potential from Eq. .

95.9.2 Limiting case for dilute solutions

For small values ofϕ_p the log term can be expanded as follows:

$$\ln(1-\phi_p) = -\phi_p - \frac{\phi_p^2}{2} - ... \tag{95.44}$$

Retaining only the first two terms in the expansion gives the following for the solvent chemical potential:

$$\frac{\mu_s}{RT} = -\frac{\phi_p}{N} + (\chi - 1/2)\,\phi_p^2 \tag{95.45}$$

Exercise: What is the polymer volume fraction for a polymer solution with a concentration of 0.05 g/cm^3? Assume the solvent has a density,ρ_s, of 0.85 g/cm^3 and the polymer has a density,ρ_p, of 1.1 g/cm^3 and there is no volume change on mixing.
Solution: We just need to calculate the total weight of solution in a volume of $1cm^3$:

$$c = \phi_p \rho_p = 0.05(1.1\,\mathrm{g/cm^3}) = 0.055\mathrm{g/cm^3} \tag{95.46}$$

Note: Actual solution concentrations are generally reported as weight of polymer per total volume of solution - the quantity that we refer to as c. From a theoretical standpoint, it is generally more convenient to use the polymer volume fraction,f_c. These quanties are linearly related to one another through the polymer density, as illustrated here. The solvent density is not needed for this calculation, although it is needed to determine the ratio of polymer to solvent

95.10 Osmotic Pressure

The concept of osmotic pressure can be illustrated by considering the operation of the membrane osmometer illustrated conceptually in Figure95.12. The device consists of a polymer/solvent mixture that is separated from a bath of pure solvent by a membrane that is permeable to solvent but impermeable to the polymer. The chemical potential of the pure solvent is zero, by definition. Because the solvent chemical potential in the polymer/solvent mixture is less than zero, there is a driving force for solvent to move into this side of the membrane from the pure solvent side. Solvent will continue to move into this side of the device until the hydrostatic pressure generated by the increased depth of the liquid layer provides counteracts the driving force associated with the chemical potential balance. This pressure is the osmotic pressure, and is obtained by dividing the chemical potential difference by the solvent volume:

$$\Pi = -\mu_s / V_s \tag{95.47}$$

Π = osmotic pressure
ρ = solution density
g = gravitational acceleration ($9.8\mathrm{m/s^2}$
V_s = solvent molar volume

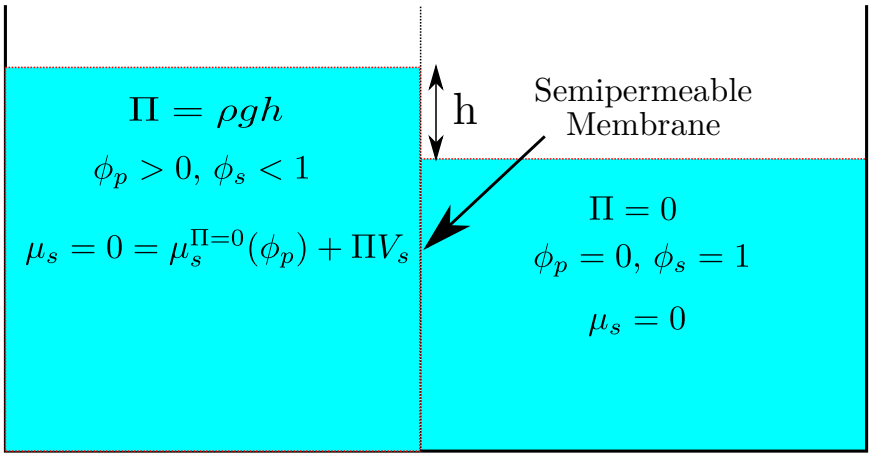


Figure 95.12: Diagram demonstrating the principles of osmotic pressure.

One important aspect of an osmotic pressure measurement is that it can be used to provide a direct measure of the number average molecular weight of the polymer. Osmotic pressure effects are important in many other situations as well. For example, the osmotic pressure is the minimum pressure required to purify a solution by

forcing it through a membrane. The osmotic pressure of aqueous solutions is also important in a wide variety of situations in biology. The height of many trees is limited by the maximum osmotic pressure that can be sustained in the leaves, since this is the driving force that draws water to the upper reaches of the tree. Here we explore these phenomena in more detail, and develop some quantitative descriptions of the osmotic pressure.
By combining the Flory-Huggins expression for the solvent chemical potential (Eq.95.42) with the definition of the osmotic pressure (Eq.95.47), we obtain the following expression forΠ:

$$\Pi = \frac{RT}{V_s}\left(-\ln(1-\phi_p) - \phi_p\left(1-\frac{1}{N_p}\right) - \chi\phi_p^2\right) \tag{95.48}$$

IfN_p is large the osmotic pressure depends only on the polymer concentration, and not on the degree of polymerization. The following result is obtained forN_p >>1:

$$\Pi = \frac{RT}{V_s}\left(-\ln(1-\phi_p) - \phi_p - \chi\phi_p^2\right) \tag{95.49}$$

95.10.1 Osmometry for Molecular Weight Determination

The number average molecular weight can be obtained by measuring the osmotic pressure of a dilute polymer solution. Using the dilute solution form of the Flory-Huggins expression for the solvent chemical potential (Eq.95.45) gives the following forΠ:

$$\frac{\Pi}{RT} = \frac{\phi_p}{N_p V_s} + (1/2-\chi)\frac{\phi_p^2}{V_s} \tag{95.50}$$

The polymer volume fraction and overall solution volume (V) are given by the following expressions:

$$\phi_p = \frac{n_p N_p V_s}{V} \tag{95.51}$$

Use of this expression forϕ_p gives the following forπ:

$$\frac{\Pi}{RT} = \frac{n_p}{V} + (0.5-\chi)N_p^2 V_s\left(\frac{n_p}{V}\right)^2 + \ldots \tag{95.52}$$

Here it is helpful to rewriten_p/V in the following way:

$$\frac{n_p}{V} = \left(\frac{w}{V}\right)\left(\frac{n_p}{w}\right) = \frac{C}{M_n} \tag{95.53}$$

where$c$ is simply the polymer concentration (kg/m^3 in SI units, g/cm^3 in the units that are more commonly used) andM_n is the number average molecular weight (kg/mole in SI units). The osmotic pressure is therefore related to the number average molecular weight and concentration as follows:

$$\frac{\Pi}{RT} = \frac{C}{M_n} + (0.5-\chi)N_p^2 V_s\left(\frac{1}{M_n}\right)^2 C^2 + \ldots \tag{95.54}$$

This is expression is often written as a virial expansion in powers of the concentration:

$$\frac{\Pi}{RT} = \frac{C}{M_n} + A_2C^2 + A_3C^3 + ... \tag{95.55}$$

A_2 is the second virial coefficient, which describes effective two-body interactions between polymer segments within the solvent.$A_2 = 0$ for$\chi = 0.5$, which means that the polymer chain does not 'see' itself under these conditions. For$\chi = 0.5$, the polymer chain behaves as a random walk, with chain dimensions very similar to what is obtained for the pure, amorphous polymer in the absence of solvent. For$\chi > 0.5$, polymer/polymer contacts are more favorable than polymer/solvent contacts. Under these conditions, high molecular weight polymers will not dissolve.

95.10.2 Scaling Theory of Osmotic Pressure

The previous derivation of the expressions for the osmotic pressure are useful, but they are not packed with physical insight regarding the actual meaning of osmotic pressure. Some simple physical arguments are more useful in this sense. We can start by making an analogy to the ideal gas law, which must be valid in the dilute regime. We start with the familiar ideal gas law:

$$PV = nRT \tag{95.56}$$

If we replace the pressure,P, with the osmotic pressure,Π, and replace the number of moles of gas molecules,n, with the number of polymer moleculesn_p. By equatingn_p/V withC/M_n (Eq.95.53), we recover the leading term in virial expansion for the osmotic pressure (Eq.95.55):

$$\Pi = \frac{CRT}{M_n} \tag{95.57}$$

In the dilute regime, molecular collisions are rare, and the pressure and free energy are dominated by the entropic penalty associated with confinement of the molecules into a fixed volume. Virial coefficients quantify the deviations from this ideal, dilute limit.

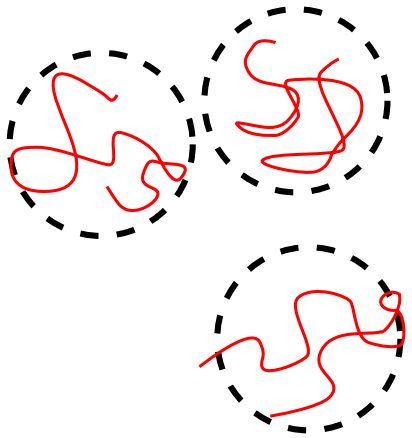


Figure 95.13: Polymer solution at a concentration less than the overlap concentration,c^*. Atc^* the spheres overlap and completely fill the space.

As the solution gets more and more concentrated, we eventually get to the point where the individual polymer molecules overlap with one another. If we treat the polymers as spheres with a volume equal toR_0^3, the concentration at which this happens, defined as theoverlap concentration,c^*, is given by the average concentration in one of these spheres:

$$C^*\left(\mathrm{g/cm^3}\right) = \frac{M\,(\mathrm{g/mole})}{R_0^3\,(\mathrm{cm^3})N_{av}\,(\mathrm{mole^{-1}})} \tag{95.58}$$

Below the overlap concentration, we can specify the polymer concentration byξ, the average distance between polymer molecules in solution, as illustrated in Figure95.14. The concentration of polymer molecules (molecules per volume) is$1/\xi^3$ and the osmotic pressure is given by the following expression:

$$\Pi = \frac{RT}{\xi^3} \tag{95.59}$$

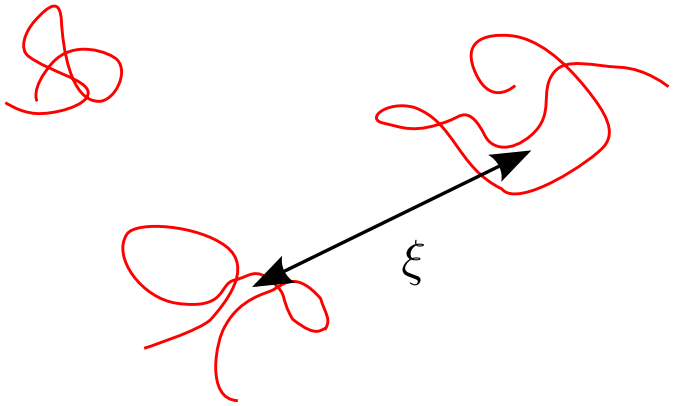


Figure 95.14: Correlation length and dilute solutions

Exercise: What is the overlap concentration poly(methyl methacrylate) with M = 100,000 g/mole in toluene?
Solution: We can combine Eq.95.58 for the overlap concentration with Eq.90.18 forR_0 to obtain the following:

$$C^* = \frac{M}{R_0^3 N_{av}} = \frac{M}{a_s^3 N^{1.8} N_{av}} = \frac{M_0}{a_s^3 N^{0.8} N_{av}} \tag{95.60}$$

Note that we have used the self-avoiding walk expression forR_0, which is the relevant expression for good solvent conditions ($\chi < 0.5$). PMMA hasM_0 = 100 g/mole, so$N = 1000$ for the polymer in our example. With these values of$N$ andM_0, and an assumed value of 7 Å (7×10^{-8} cm) fora_s, we obtain$c^* = 1.9 \times 10^{-3}$ g/cm^{-3}. The density of toluene is slightly less than 1 g/cm^3, so this corresponds to a polymer weight fraction of a couple tenths of a percent.

The basic assumption of scaling theory of osmotic pressure is thatΠ is still given by Eq.95.59 at concentrations above the overlap concentration, but thatξ has a different meaning in this regime. The approach works in the 'semidilute' solution regime where the overall concentration is still low ($\phi_p < 0.25$), but where there is substantial overlap between polymer molecules. Each overlap point represents a constraint that contributes an entropic contribution of ~k_BT to the free energy of the system. The osmotic pressure is therefore ~k_BT times the concentration of overlap points, ork_BT/ξ^3, whereξ is the correlation length, or average distance between overlap points. As the solution gets more concentrated, the average distance between polymer contacts decreases., as shown schematically in Figure95.15.

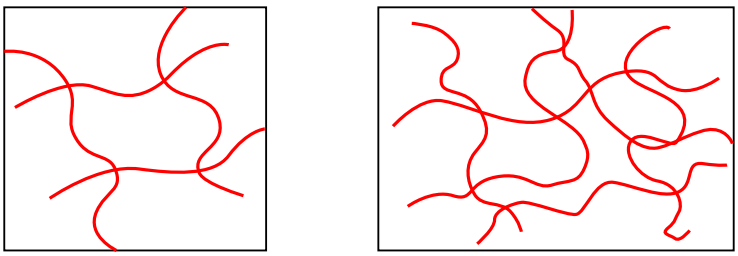


Figure 95.15: Change in correlation length with solution concentration.

To get the concentration dependence of the the correlation length, we can start with the postulate that the corrections to the low-conentration form of the osmotic pressure must depend only onc/c^*, so that the osmotic pressure has the following form:

$$\frac{\Pi M}{CRT} = 1 + f\left(C/C^*\right) \tag{95.61}$$

Here$f\left(C/C^*\right)$ is some function that we still need to figure out. It's easy to sort out what this function must look like, because for$C \gg C^*$ we don't expect that the molecular weight of the polymer to matter any more. At these concentrations the osmotic pressure is dominated by polymer/polymer contacts, and the solution at this local level does not depend at all on the molecular weight. The only way for the osmotic pressure to be independent of the concentration for high concentrations is for$f\left(C/C^*\right)/M$ to be independent of the$M$. Because$C^* \propto M^{-4/5}$, this is only possible if$f\left(C/C^*\right) \propto \left(C/C^*\right)^{5/4}$. This means that for$C > C^*$ we have$\Pi \propto C^{9/4}$, or equivalently,$\Pi \propto \phi_p^{9/4}$. This means that that the correlation length, which is related to the osmotic pressure through Eq.95.59, is given by the following expression for$C > C^*$:

$$\xi = a_0 \phi_p^{-3/4} \tag{95.62}$$

Herea_0 is a typical monomer size, comparable in magnitude to the cube root of the molecular volume, or to the statistical segment length. This value forξ can then be combined with Eq.95.59 to give the following expression for the osmotic pressure in the semidilute concentration regime:

$$\Pi = \frac{k_B T}{a_0^3} \phi_p^{9/4} \tag{95.63}$$

The full expression for the osmotic pressure for dilute and semidilute concentrations is obtained by using this value ofξ for$C > C^*$and adding the osmotic pressure that dominates in the dilute limit:

$$\Pi = \frac{CRT}{M} + \frac{k_B T}{a_0^3} \phi_p^{9/4} \tag{95.64}$$

While derived in completely different ways, Eqs.95.49 and95.63 actually behave similarly in the concentration regime where they both apply. This is illustrated in Figure95.16., where these two expressions are compared for a specific combination ofa_0 andχ.

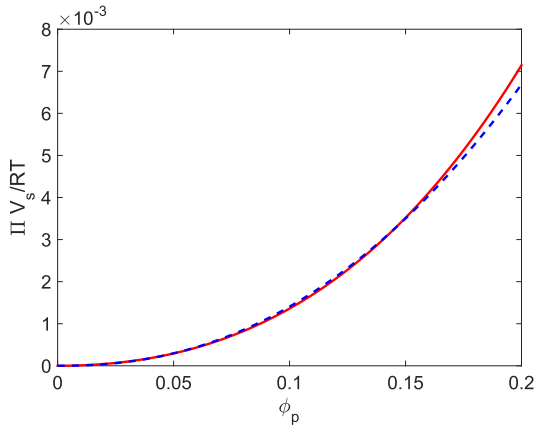


Figure 95.16: Concentration dependence of the Flory-Huggins form for the osmotic pressure (Eq.95.49) with the semidilute scaling form (Eq.95.63), with $\chi = 0.4$ and $a_0^3 = 4v_s$.

95.10.3 Typical Magnitude of the Osmotic Pressure

The fundamental pressure scale for water is the thermal energy, RT divided by the solvent molar volume, V_s. For water, we have (at 25 ºC):

$$V_s = \frac{18\,\mathrm{g}}{\mathrm{mole}}\frac{10^{-6}\mathrm{m}^3}{1\mathrm{g}} = 1.8\mathrm{x}10^{-5}\,\mathrm{m}^3 = 18\,\mathrm{cm}^3 \tag{95.65}$$

$$\frac{RT}{V_s} = \frac{(8.314\,\mathrm{J/mole-K})(298\,\mathrm{K})}{3.0\times10^{-29}\,\mathrm{m}^3} = 1.4\times10^8\,\mathrm{Pa} \tag{95.66}$$

This pressure is relatively large. The actual osmotic pressure is obtained by multiplying by a prefactor involving the Flory-Huggins interaction parameter and the polymer volume fraction. Osmotic pressures of many atmospheres (1 atm =10^5 Pa) can easily be obtained.

Exercise: What is the osmotic pressure of seawater?
Solution: The osmotic pressure arises from all of the dissolved salts in seawater. Typical concentrations are tabulated below.[11] The total concentration of all dissolved salts is 1.12 moles/kg, which after multiplying by the density of seawater (1025 kg/m^3) gives an overall salt concentration of 1.15 moles/m^3. The osmotic pressure is obtained by multiplying this concentration by RT:

$$\Pi = (8.314\,\mathrm{J/mol\cdot K})\,(300\,\mathrm{K})\left(1025\,\mathrm{kg/m^3}\right) = 2.6x10^6\,\mathrm{Pa}$$

Species	moles/kg
H_2O	53.6
Cl^-	0.546
Na^+	0.469
Mg^{2+}	0.0528
SO_4^{2-}	0.0282
Ca^{2+}	0.0103
K^+	0.0102
Br^-	0.000844
Sr^{2+}	0.000091
F^-	0.000068
Total Ions	1.12

95.11 Equilibrium Swelling of a Neohookean Material

A general expression for the pressure difference is obtained from the first law of thermodynamics, using an approach that is similar to the approach used to determine the stresses in deformed elastic solids.

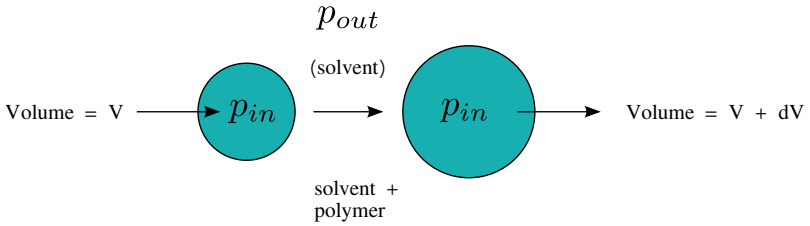


Figure 95.17: Diagram of equilibrium swelling

Work done on system:

$$(p_{in} - p_{out})dV = \Delta p dV \tag{95.67}$$

Change in elastic free energy:

$$d(\Delta F_d) \tag{95.68}$$

$$\Delta p = \frac{\partial}{\partial V}\Delta F_d = V\frac{\partial}{\partial V}\Delta f_d \tag{95.69}$$

For the specific case of a Neohookean material, we begin with the appropriate free energy expression:

$$\Delta f_d = \frac{G\left\{\lambda_x^2 + \lambda_y^2 + \lambda_z^2 - 3\right\}}{2} \tag{95.70}$$

For isotropic swelling,$\lambda_x = \lambda_y = \lambda_z = \lambda$. The shear modulus in this equation is the modulus of the reference state, which has V = V0 andλ = 1. With$V/V_0 = \lambda^3$ we have:

$$\Delta f_d = \frac{3G(\lambda^2 - 1)}{2} = \frac{3G}{2}\left[\left(\frac{V}{V_0}\right)^{2/3} - 1\right] \tag{95.71}$$

The elastic pressure at the reference state is given by differentiating with respect to the volume, and evaluating the result at$V = V_0$

$$\Delta p = V_0 \frac{\partial(\Delta f_d)}{\partial V}|_{V=V_0} = \frac{\partial}{\partial V}\left[\frac{3G}{2}\left(\frac{V}{V_0}\right)^{2/3}\right]_{V=V_0} = G \tag{95.72}$$

The result from the previous page is that the elastic swelling pressure is equal to the shear modulus of the material. This shear modulus is related to the polymer volume fraction in the swollen gel by the expression derived previously. In a swelling experiment, the pressure is simply the osmotic pressure difference associated with the difference in polymer concentrations inside and outside the gel:

$$\Pi_{in} - \Pi_{out} = \frac{f\phi_p^{1/3}\rho RT}{M_s}\frac{\left\langle R_{dry}^2\right\rangle}{R_0^2} \tag{95.73}$$

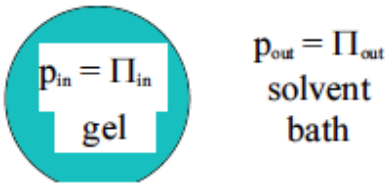


Figure 95.18: Model of equilibrium swelling

The osmotic pressure in a bath of pure solvent is zero (Π_{out} = 0). If a dry material is immersed in the solvent bath, it takes up solvent, reducingϕ_p until the osmotic pressure inside the gel is equal to its shear modulus. The following equation forϕ_p is obtained if we use the scaling expression for the osmotic pressure:

$$\frac{k_B T}{a_0^3}\phi_p^{9/4} = \frac{f\phi_p^{1/3}\rho RT}{M_s}\frac{< R_{dry}^2 >}{R_0^2} \tag{95.74}$$

$$\phi_p = \left(G_{dry}\frac{a_0^3}{k_B T}\right)^{12/23} \tag{95.75}$$

96 Surfaces and Interfaces

96.1 Equilibrium Contact Angle

The shape of a liquid droplet on a solid surface is determined by a balance of forces where the solid, liquid and ambient environment (typically air) come into contact. These three phases meet at the contact line. The surface energy of the liquid (γ_ℓ), the surface energy of the solid (γ_s) and the interfacial energy of the solid/liquid interface (γ_{sl}) are related to one anther by the following expression, obtained by balancing horizontal forces at the contact line.

$$\gamma_s = \gamma_{sl} + \gamma_l \cos\theta_e$$

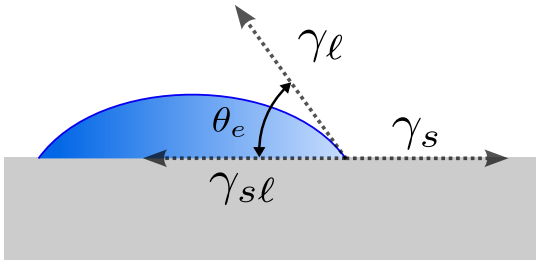


Figure 96.1: Diagram of the equilibrium contact angle

96.2 Polymer/polymer Interfacial Free Energy

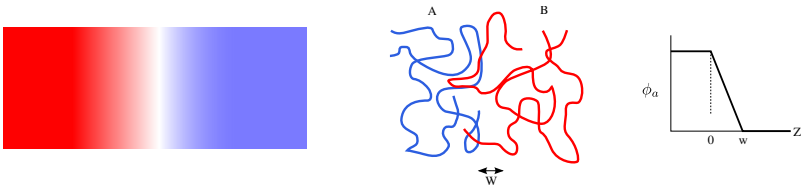


Figure 96.2: Schematic representation of the interface between immiscible polymers.

The interface between two immiscible polymers is characterized by a finite interfacial width,w. The actual value of$w$ is determined by a balance of enthalpic considerations, which favor a small value of$w$, and entropic considerations, which favor a large value of$w$. Below we obtain an expression for the interfacial energy as a function of$w$, and minimize this with respect to$w$ to get expressions for both$w$ and the interfacial free energy between immiscible A and B polymers.

The enthalpic contribution to the interfacial free energy is obtained by integrating the enthalpy per unit volume,Δh, associated with contacts between A and B repeat units in the interfacial zone. Here we use the same regular solution form that forms the basis of the Flory-Huggins expression for the free energy of mixing:

$$\Delta h(z) = \frac{RT}{V_0}\chi\phi_a(z)\phi_b(z) \tag{96.1}$$

If we assume as simple linear expression for the concentration profile in the interfacial zone, we have the following for$0 < z < w$ (as illustrated in Figure96.3).

$$\phi_a(z) = 1 - z/w; \tag{96.2}$$

$$\phi_b(z) = z/w \tag{96.3}$$

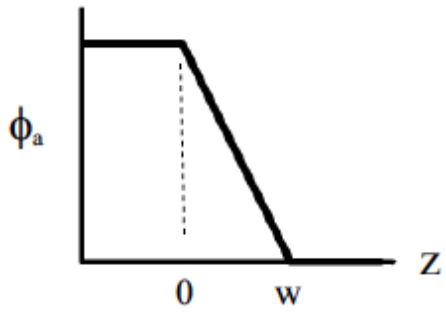


Figure 96.3: Linear approximation to the concentration profile across an interface between immiscible polymers.

These expressions result in the following forγ_χ, the enthalpic contribution to the interfacial free energy. We refer to this contribution asγ_χ because it includes all of the thermodynamic interactions that are included inχ. In addition to all the enthalpy of mixing, this parameter also includes the non-ideal entropy of mixing:

$$\gamma_\chi = \int_0^w \Delta h\left(z\right) dz = \frac{RT\chi}{V_0} \int_0^w \left(\frac{z}{w}\right)\left(1-\frac{z}{w}\right) dz = \frac{RTw\chi}{6V_0} \tag{96.4}$$

A simple argument for the entropic contribution every time a random walk goes from one side of the interface to the other, it must turn around, so that we use one half of the available molecular configurations. If the number of molecular configurations isΩ, the configurational entropy of associated with these configuration is$k_B \ln \Omega$. If the number of configurations decreases by a factor or 2, then the entropy decreases by$k_B \ln\left(2\right)$ and the free energy increases by$k_B T \ln\left(2\right)$. Because our theory is already approximate, we'll assume that this free energy increase per forced turnaround of a polymer molecule is simply$k_B T$.

Our calculation of the entropic contribution to the free energy proceeds in the four steps outlined below:

(1) DetermineN_w, number of repeat units in random walk with$R_0 = w$.

$$a\sqrt{N_w} = w; N_w = \left(w/a\right)^2$$

(2) Calculatev_w, the volume of each of these random walks:

$$v_w = N_w v_0 = v_0\left(w/a\right)^2$$

(3) CalculateΓ_w, the number of these random walks per unit area in the interfacial region:

$$\Gamma_w = \frac{w}{v_w} = \frac{a^2}{wv_0}$$

(4) multiply byΓ_i by$k_B T$ to getγ_s, the entropic contribution to the interfacial free energy:

$$\gamma_s = \frac{k_B T a^2}{w}$$

The overall interfacial free energy,γ_{ab}, is obtained by addingγ_χ andγ_s:

$$\gamma_{ab} = \gamma_\chi + \gamma_s = \frac{k_B T}{v_0}\left(\frac{\chi w}{6} + \frac{a^2}{w}\right)$$

The actual value of$w$ is obtained by minimizingγ with respect to$w$. We obtain the following:

$$w = a\sqrt{\frac{6}{\chi}}$$

$$\gamma_{ab} = \frac{2ak_B T}{v_0}\sqrt{\chi/6} = \frac{2aRT}{V_0}\sqrt{\chi/6}$$

This approximate approach gives the correct scaling ($w \propto a\chi^{-1/2}; \gamma_{ab} \propto \frac{ak_B T}{v_0}\chi^{1/2}$), but the exact numerical prefactors are not obtained with quantitative accuracy with this approach. A more complete treatment results in the following for$w$ andγ:

$$w = \frac{2a}{\sqrt{6\chi}} \tag{96.5}$$

$$\gamma_{ab} = \frac{aRT}{V_0}\sqrt{\chi/6} \tag{96.6}$$

96.3 Block Copolymer Morphologies

Block copolymers consisting of two different polymers characterized by a sufficiently large value of χ form a wide variety of interesting morphologies. In these cases two different polymer types that prefer to separate into different phases are covalently linked to one another along a single molecule. The morphologies observed for simple diblock copolymer molecules (a sequence of A repeat units linked to a sequence of B repeat units) is shown in Figure96.4. If the length of the A block is much longer than the length of the B block, then the B blocks segregate into isolated spheres in a continuous matrix of A polymer formed from the A blocks. If the relative length of the B block is increased, the spheres transition to a hexagonal array of cylinders, and then to a morphology where the B cylinders form an interconnected, 3-dimensional network within the A matrix. If the A and B block lengths are similar in length, then the morphology is a simple alternating array of A and B lamellae, as shown schematically in Figure96.5. An electron micrograph of a lamellar diblock copolymer morphology is shown in Figure96.6.

Figure 96.4: Diblock copolymer morphologies.

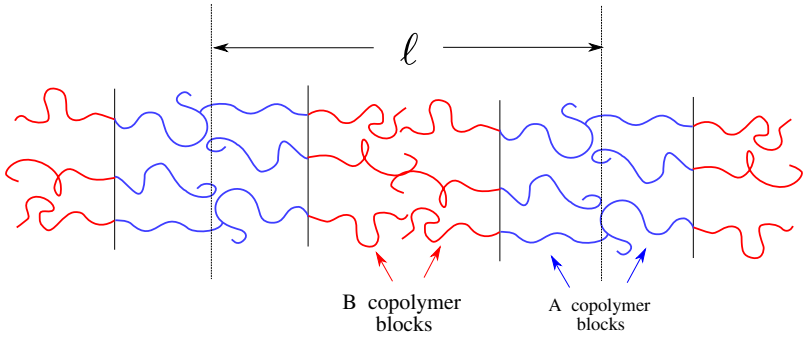


Figure 96.5: Molecular orientation in a lamellar diblock copolymer morphology.

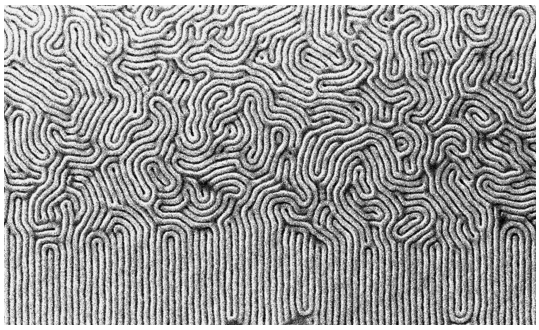


Figure 96.6: Electron micrograph of a lamellar diblock copolymer morphology.

The free energy per molecule of a lamellar-forming diblock copolymer is the sum of a chain stretching free energy, F_s , and a free energy associated with the interface between the A and B domains. We refer to this interfacial free energy per chain as F_γ. For our purposes we assume that the A and B blocks have the same statistical segment length, a. The overall volume fractions of the two different blocks, f_a and f_b, are given by the relative degrees of polymerization of the two blocks:

$$\begin{aligned} f_a &= \frac{N_a}{N_a+N_b} = \frac{N_a}{N} \\ f_a &= \frac{N_b}{N_a+N_b} = \frac{N_b}{N} \end{aligned} \tag{96.7}$$

Here N is the overall degree of polymerization of the molecule, *i.e.*, $N = N_a + N_b$. The overall width of the A region of the block copolymer is $f_a \ell$, and the width of the B region is $f_b \ell$. The stretching free energy is obtained from the requirement that each block of the molecule must have an rms end-to-end distance equal to half the width of the respective region:

$$\begin{aligned} R_0^a &= f_a \ell / 2 \\ R_0^b &= f_b \ell / 2 \end{aligned} \tag{96.8}$$

We define f_a and If the degrees of polymerization of the two blocks are N_a and N_b, and the statistical segment length for each of the two different repeating units is a, we have:

$$F_s = \frac{3}{2}\frac{RT\left(R_0^a\right)^2}{N_a a^2} + \frac{3}{2}\frac{RT\left(R_0^b\right)^2}{N_b a^2} = \frac{3}{8}\frac{\ell^2}{Na^2} \tag{96.9}$$

The interfacial free energy is obtained by multiplying the interfacial free energy by the area per molecule along the A/B interface, which we refer to as Σ:

$$F_\gamma = \gamma_{ab}\Sigma \tag{96.10}$$

This area per molecule is given by dividing the molecular volume, NV_0, by the length per molecule, $\ell/2$:

$$\Sigma = \frac{2NV_0}{\ell} \tag{96.11}$$

The overall free energy per chain, F_{tot}, is given by the sum of F_s and F_γ. Combination of Eq. 96.6 for γ_{ab} with Eqs. 96.9, 96.10 and 96.11 gives:

$$F_{tot} = \frac{3}{8}\frac{RT\ell^2}{Na^2} + \frac{2aNRT}{\ell}\sqrt{\chi/6} \tag{96.12}$$

The equilibrium value of ℓ, which we refer to as ℓ_{eq} is obtained as the value of ℓ for which F_{tot} is minimized. Setting dF_{tot}/d_ℓ to 0 gives:

$$\ell_{eq} = \left(\frac{8}{3\cdot\sqrt{6}}\right)^{1/3} aN^{2/3}\chi^{1/6} = 1.03aN^{2/3}\chi^{1/6} \tag{96.13}$$

We can use the fact that $R_0 = N^{1/2}$a to write the expression for ℓ_{eq} in the following way:

$$\frac{\ell_{eq}}{R_0} = 1.03\left(\chi N\right)^{1/6} \tag{96.14}$$

We can now substitute ℓ_{eq} for ℓ in Eq. 96.12 to obtain the following expression for the equilibrium free energy per molecule in the lamellar microstructure:

$$F_{eq} = 1.19RT\left(\chi N\right)^{1/3} \tag{96.15}$$

Equation 96.15 gives us the molar free energy for a lamellar phase, but it still does not tell us if this lamellar phase has a lower free energy for the disordered phase, where the A and B blocks remain completely mixed. In order to

do this we need to compare to the molar free energy in the disordered phase, where the A and B blocks remain completely mixed with one another. To do this we use the Flory-Huggins free energy of mixing (Eq.95.28). Because the block copolymer is a one component system, there is no entropy of mixing, so we don't include the ideal entropy of mixing terms. We just have the term involvingχ, which accounts for the interactions between the A and B repeat units in the mixed, disordered phase. The free energy per molecule is obtained by multiplying the free energy per unit volume by the molecular volume,NV_0. We get the following simple expression forF_{dis}, the molar free energy of the disordered phase:

$$F_{dis} = RT\chi\phi_a\phi_b = RT\chi f_a f_b \tag{96.16}$$

For a symmetric diblock copolymer with$f_a = f_b = 0.5$,$F_{dis} = F_{eq}$ whenχN is equal to 10.4. WhenχN is larger than this value, the ordered phase free energy given by Eq.96.15 is lower than the disordered phase free energy given by Eq.96.16. As a result we expect that an ordered lamellar phase will be formed for$f_a = 0.5$ whenχN is greater than 10.4. A disordered phase will be formed for$\chi N < 10.4$. This value of 10.5 is remarkably similar to the value of 10.5 that is obtained by a much more detailed calculation. These more detailed calculations are necessary in order to understand how the preferred geometry of the ordered phase (spheres, cylinders, etc.) depends on thef_a. A full calculation of the phase diagram is shown in Figure96.7.

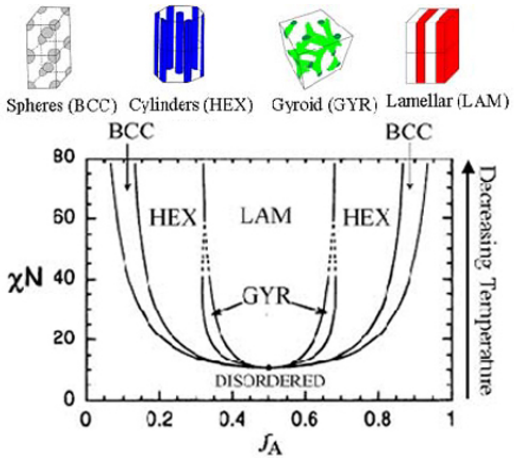


Figure 96.7: Block copolymer phase diagram (adapted from ref.[13]).

96.4 The Effect of Block Copolymers on Phase Behavior

Diblock copolymer molecules can act as macromolecular 'surfactants', segregating preferentially to the interface between the corresponding homopolymers. In the schematic illustration below, A/B diblock copolymer molecules segregate preferentially to the interface between A and B phases, thereby limiting the ability of the morphology to coarsen by coalescence of the B domains. An actual example of this for polystyrene/poly(methyl methacrylate) (PS/PMMA) system is illustrated here.

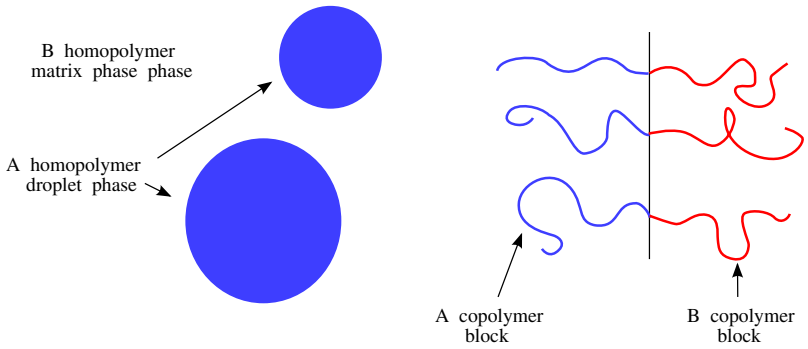


Figure 96.8: Phase behavior of block copolymers

97 Case Study: Thermoreversible Gels from Triblock Copolymer Solutions

97.1 Introduction

Many of the principles of this course are illustrated by the behavior solutions of acrylic triblock copolymers in higher alcohols. The specific molecular structures of the polymer and the solvent are shown in Figure97.1 and consist of a midblock of poly(n-butyl acrylate) (PnBA) between two blocks of poly(methyl methacrylate) (PMMA). The value ofχ for the PnBA/solvent system is less than 0.5 for all temperatures of interest, but theχ for the PMMA/solvent system has the following temperature dependence:

$$\chi = 1.45 - 0.0115T \tag{97.1}$$

with$T$ in°C.
The following points are illustrated by this example:

(1) Polymers are soluble in solvents when$\chi < 0.5$.

(2) For$\chi > 0.5$, the solvent content within the polymer can be obtained from the chemical potential expression.

(3) To understand the temperature dependence of the solubility you need to know the temperature dependence ofχ.

(4) Polymer liquids have higher heat capacities than the corresponding polymer glasses.

(5) Calorimetry can be used to detect the enthalpy recovery peak from an aged polymer glass.

(6) Gelation occurs when the average functionality is≈2.

(7) Time-temperature superposition works when the structure does not change appreciably with Temperature.

(8) The viscosity is obtained by integrating$G(t)$ or from the response at very low frequencies.

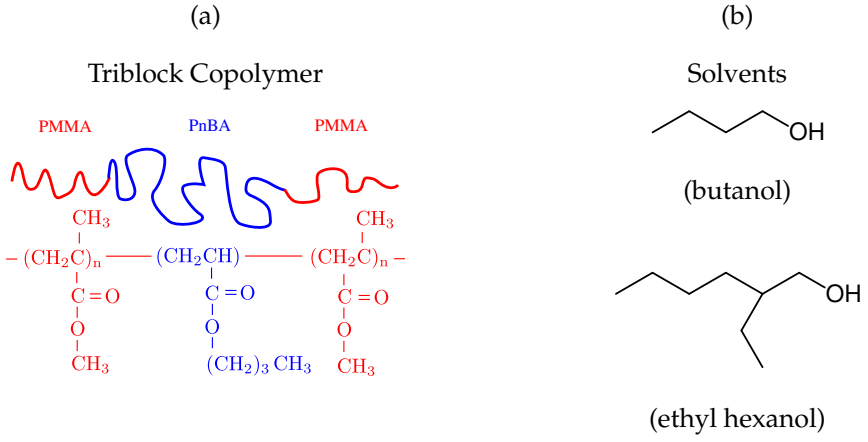


Figure 97.1: Chemical structure of gel-forming triblock copolymer gels (a), and the alcohols used as the solvents (b).

From Eq.97.1 we see that the value ofχ characterizing the PMMA/solvent interaction is greater than 0.5 for$T < 83\,^{\circ}$C. At temperatures above 83°C the PMMA and PnBA blocks are both in good solvent conditions, and the polymer behaves as a normal polymer solution. At lower temperatures, however, the PnBA midblock remains in good solvent conditions, but the PMMA endblocks are no longer soluble. As a result these endblocks aggregate to form the micellar structures illustrated schematically in Figure97.2. As a solution is cooled,χ increases and solvent is expelled from these aggregates. When sufficient solvent has been expelled from the PMMA aggregates, they become glassy, and the system behaves as a solid.

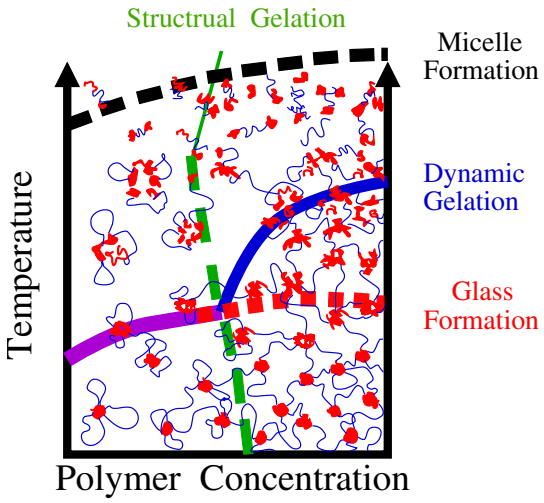


Figure 97.2: Schematic representation of the temperature and concentration dependence of the structure of the triblock copolymer solutions.

97.2 Thermoreversible Gelcasting of Ceramics

The transition from 'liquid-like' to 'solid-like' behavior as the temperature is changes enable these materials to be used in a variety of applications, including the ceramics processing application illustrated in Figures97.3 and97.4. In this application ceramic particles are added to the the triblock copolymer at a relatively high volume fraction (around 50% by volume). The thermoreversible nature of the triblock copolymer solution allows the suspension to be poured into a silicone mold at high tempertures (typically≈ 80 °C). When the system is cooled to room temperature, the triblock copolymer forms a solid gel and the material can be removed from the mold. A ceramic material with the shape of the mold is obtained by letting the solvent evaporate from the suspension, and firing the object at elevated temperature to burn off the polymer and sinter the ceramic particles together.

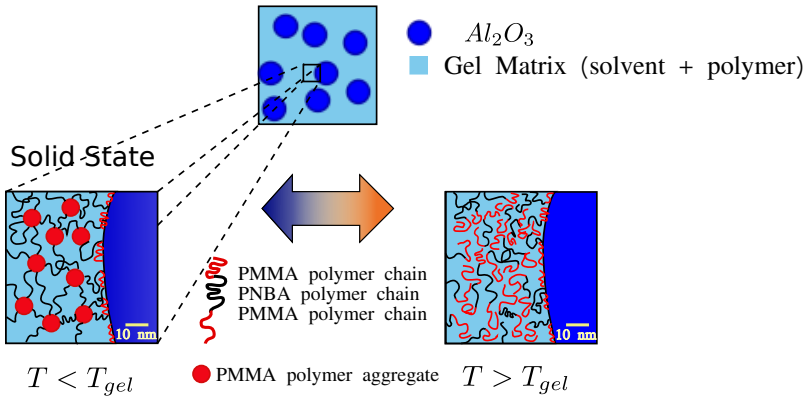


Figure 97.3: Schematic representation of the alumina particles dispersed within a thermoreversible triblock copolymer gel.

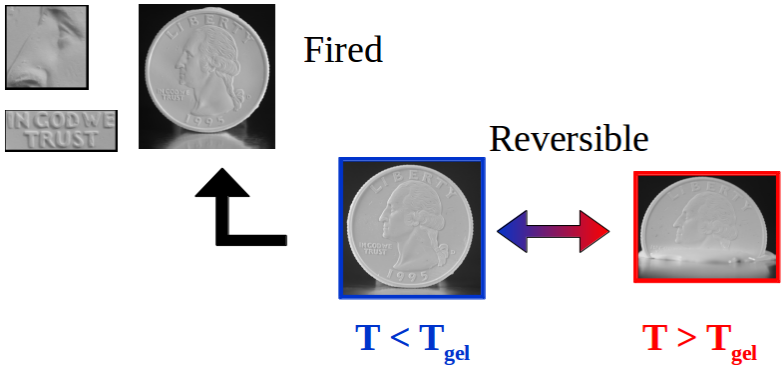


Figure 97.4: Illustration of the thermoreversible gelcasting process.

97.3 Quantifying the Solid/Liquid Transition

The temperature dependence of the mechanical properties of the triblock copolymer solutions are quantified by the frequency dependence of the storage and loss moduli,$G'(\omega)$ and$G''(\omega)$. It is conceptually simpler to think in terms of the time-dependence of the relaxation modulus,$G(t)$, which can be written as a sum of exponential relaxations:

$$G(t) = \sum_i G_i \exp(-t/\tau_i) \tag{97.2}$$

In the frequency domain, the values storage and loss moduli are given by the following expressions:

$$G'(\omega) = \sum_i \frac{G_i \tau_i^2 \omega^2}{1 + \tau_i^2 \omega^2} \tag{97.3}$$

$$G''(\omega) = \sum_i \frac{G_i \tau_i \omega}{1 + \tau_i^2 \omega^2} \tag{97.4}$$

The most important thing for us is the viscosity, which we can obtain in terms of the values ofG_i andτ_i, or in terms of the loss modulus at very low frequencies:

$$\eta = \sum_i G_i \tau_i = \lim_{\omega \to 0} \frac{G''}{\omega} \tag{97.5}$$

We generally obtain data as a master plot (where the frequency is multiplied by the temperature shift factor,a_T). In this case we obtain the temperature dependence of the viscosity by multiplying bya_T:

$$\eta = a_T \lim_{a_T \omega \to 0} \frac{G''}{a_T \omega} \tag{97.6}$$

We can also obtain an expression for the limiting modulus at high frequency (or low temperature):

$$G_0 = \sum_i G_i \tag{97.7}$$

Finally, we define an average relaxation time,τ_{av} in the following way:

$$\tau_{av} = \frac{\eta}{G_0} = \frac{\sum G_i \tau_i}{\sum G_i} \tag{97.8}$$

Values of the storage and loss moduli for the triblock copolymer solutions are shown in Figure97.5, along with the values ofG_i andτ_i used to fit the data. The temperature dependent viscosity for a solution with$\phi_p = 0.15$ is shown in Figure97.6, and the temperature dependent relaxation times for a variety of solution concentrations are shown in Figure97.7.

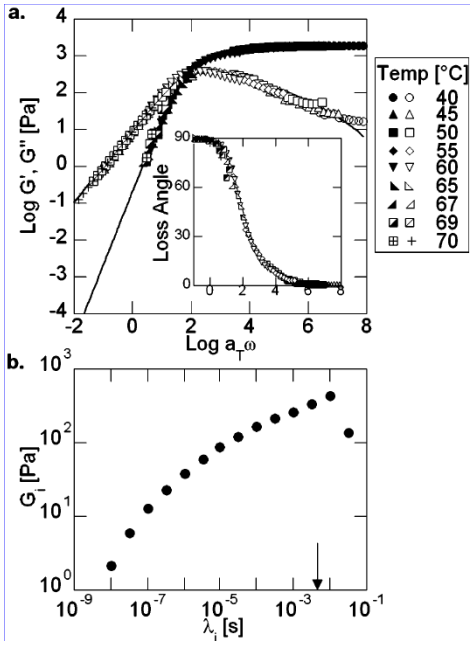


Figure 97.5: a) Master curves forG' andG'' (symbols) with fits to Eqs.97.3 and97.4. b) Values ofG_i andτ_i used to fit the data from part a (the relaxation times,τ_i are referred to byλ_i in the figure). The reference temperature (defined as the temperature where$a_T = 1$) is 65°C for these data.

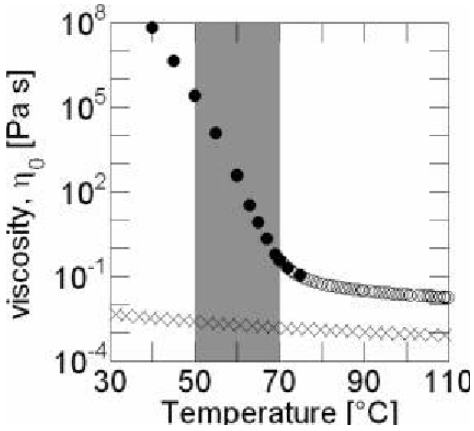


Figure 97.6: Temperature dependence of the viscosity for a triblock copolymer solution with$\phi_p = 0.15$. in 2-ethyl hexanol. The x's represent the solvent viscosity.

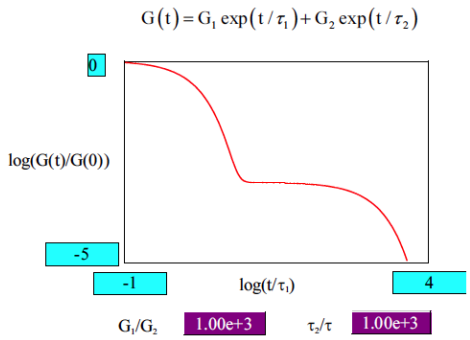


Figure 97.7: Values of the average relaxation time,τ_{av} as a function of temperature. Data are included forϕ_p = 0.05 (•) $\phi_p = 0.10$ (■) $\phi_p = 0.15$ (⊡) $\phi_p = 0.20$ (▲) $\phi_p = 0.25$ (▼) $\phi_p = 0.30$ (◯).

97.4 Characterizing the Glass Transition in PMMA domains

The glass transition is typically observed as step change in heat flow as a sample is cooled through the glass transition. In our case the PMMA domains go through a glass transition during cooling, but the PMMA concentration is so low that the glass transition is not clearly observed during cooling or when when the sample is immediately reheated after cooling (see Figure97.8). However, a large enthalpy peak is observed after the sample has been left to 'age' at room temperature for a long time before reheating the sample. As shown in Figure97.9, the enthalpy peak gets larger and moves to higher temperature as the aging time at room temperature is increased. This peak has the appearance of a melting peak associated with crystallization, but in this case it arises from the glassy PMMA domains.

The origins of the enthalpy peak observed during the aging experiments can be understood by realizing that the glass transition separates an equilibrium, liquid regime at temperatures aboveT_g from a non-equilibrium, glassy regime at temperatures belowT_g. At temperatures below the glass transition, the enthalpy content of the sample is higher than the equilibrium enthalpy content defined by the extrapolation of the liquid behavior (the dashed line in Figure97.10). As a result, the enthalpy slowly decreases toward this equilibrium line as the sample is aged below the glass transition. The decrease in enthalpy during this aging process is labeled asΔH_a in Figure97.10. When the sample is reheated, the enthalpy increases with temperature according to the glassy heat capacity,C_p^g. As a result the enthalpy content eventually crosses the equilibrium line and becomes less than the equilibrium enthalpy content. At a temperature somewhere just aboveT_g, the sample is able to equilibrate, and the enthalpy increases by an amountΔH_r in order to catch up to the equilibrium value. This is the enthalpy corresponding to the peaks in Figure97.9. As the aging time decreases,ΔH_a andΔH_r both decrease, and the temperature at which the enthalpy is recovered moves closer toT_g. The actual value of theT_g can be estimated by measuring the area of the enthalpy recovery peak and plotting against the location of this peak. Doing this for the data shown in Figure97.9 results in an estimate forT_g of the PMMA domains of 35°C.

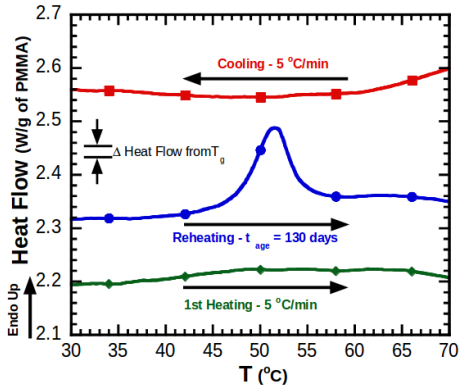


Figure 97.8: Measured heat flow during cooling, immediate reheating, and reheating after room temperature storage for 130 days.

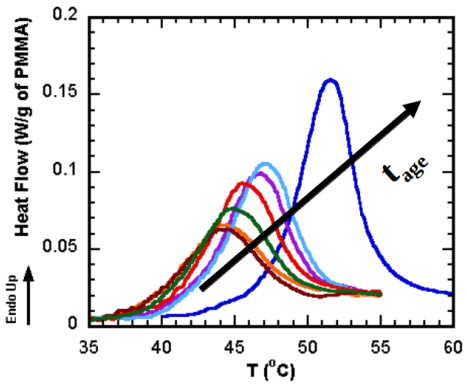


Figure 97.9: Dependence on the enthalpy recovery peak on the aging time.

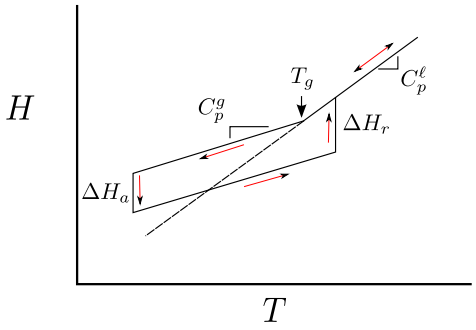


Figure 97.10: Origins of the enthalpy recovery peak.

As a final check on all of this, we can see if it makes sense that the glass transition of the PMMA domains should be around 35°C. The glass transition of PMMA is about 125°C, which is certainly a lot higher than the value of 35°C that we are claiming here. However, we need to remember that the PMMA domains contain a lot of solvent, and this solvent will depressT_g significantly. How much solvent do we expect there to be in the PMMA domains? From Eq.97.1 we obtainχ=1.05 at$T = 35\,^{\circ}$C. The relationship betweenχ andϕ_s, the solvent volume fraction in the PMMA cores of the micelles is obtained by settingN_p to∞ and setting the solvent chemical potential (from Eq95.43) to 0:

$$\begin{aligned} \ln(\phi_s) + 1 - \phi_s + \chi\,(1-\phi_s)^2 &= 0 \\ \ln(1-\phi_p) + \phi_p + \chi\phi_p^2 &= 0 \end{aligned} \tag{97.9}$$

This equation needs to be solved numerically to obtainϕ_s (orϕ_p) from a specified value ofχ. A simple MATLAB script used to do this is given in Section102.2. A value of 1.05 forχ givesϕ_s=0.289. Independent measures of the glass transition temperature of bulk PMMA samples with solvent loadings comparable to this do in fact show measured glass transition temperatures close to 35°C. The relationship betweenχ andϕ_p given by Eq.97.9 is plotted in Figure97.11.

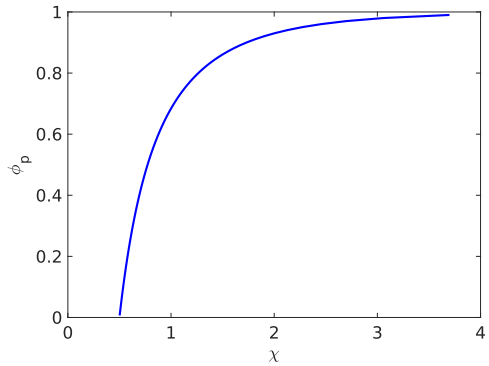


Figure 97.11: Relationship betweenϕ_p andχ obtained from Eq.97.9.

97.5 Concentration of the gel Modulus

Elasticity of the triblock copolymer solutions at low temperatures arises from the fact that the middle, PnBA blocks of the copolymer can span different PMMA aggregates, thereby linking the whole structure together. The PMMA aggregates behave as physical crosslinks with a functionality given by the number of midblocks that that bridge different aggregates. This functionality is obtained by multiplying the aggregation number,*i.e.* the number of PMMA endblocks in a single aggregate, by the probability$f$ that a PnBA midlbock spans two different PMMA aggregates (we also have to divide by two to account for the fact that there are two PMMA blocks on each triblock copolymer molecule). Aggregation numbers depend on the polymer concentration, and are typically very large, as illustrated in Figure97.12. For high molecular weight polymers the percolation threshold, where the average functionality of a micelle is 2, is quite low, corresponding to$\phi_p \approx 0.035$. Above this percolation threshold, the shear modulus is given by the following expression:

$$G_0 = \nu k_B T \frac{D^2}{R_0^2} \tag{97.10}$$

Here$D$ is the average distance between micelle cores, which can be measured experimentally by x-ray scattering. We also have$R_0 = N^{1/2}a$, where$N$ is the midblock degree of polymerization and$a$ is the statistical segment length for the midblock. The quantityν is the conentration of 'load bearing strands', which in this case is the concentration of triblock copolymer chains with bridging midblocks:

$$\nu = \frac{f\phi_p \rho N_{av}}{M}$$

where$M$ is the molecular weight of the triblock copolymer molecule andρ is the polymer density. The modulus is strongly concentration dependent because of the concentration dependence of the bridging fraction,f, which is shown in Figure97.13. Measured and calculated values ofG_0 are shown in Figure97.14.

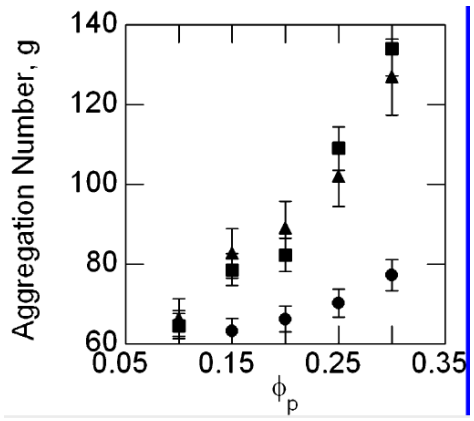


Figure 97.12: Concentration dependence of aggregation number for three different PMMA-PnBA-PMMA block copolymer solutions. The three data sets correspond to three differnt triblock copolymers, with different molecular weights for the PMMA and PnBA blocks.

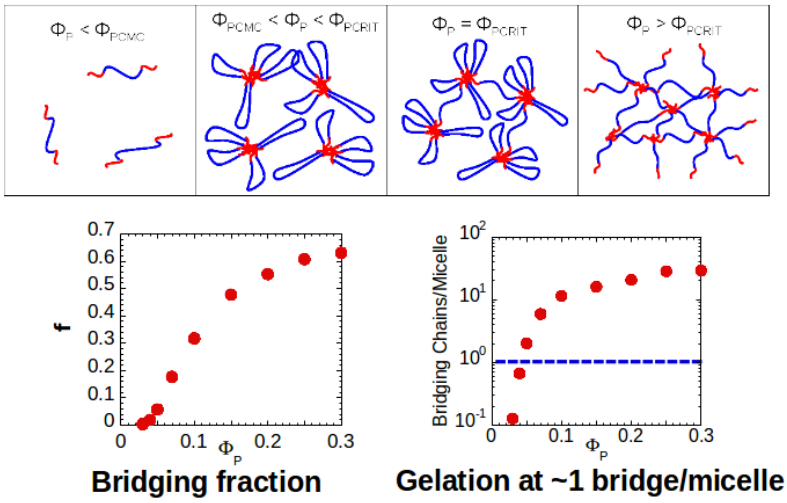


Figure 97.13: Concentration dependence of the network functionality.

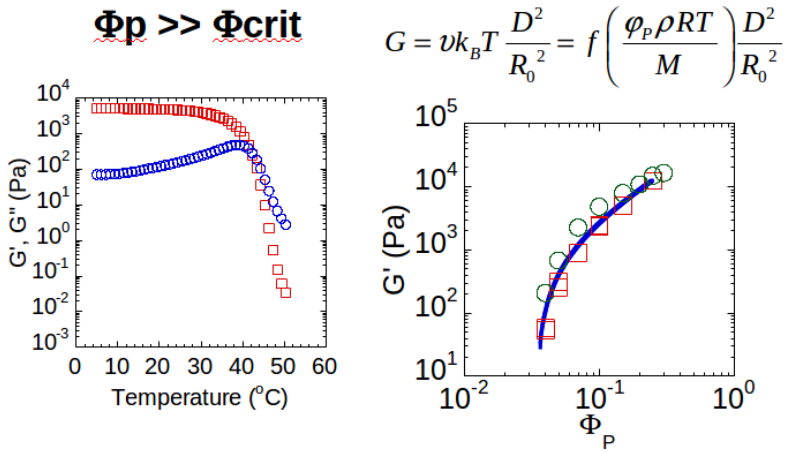


Figure 97.14: Calculated and experimental concentration dependence of the network functionality.

97.6 Hydrogels: Water as the Solvent

As a final illustration of what can be done with these sorts of triblock copolymer gels, we consider materials where the midblock is replaced with poly(acrylic acid), a polymer that is water soluble at neutral pH. The structure of these polymer is shown in Figure97.15. This figure also shows a scheme for forming gels from these materials. Instead of adjustingχ between the solvent and the PMMA endblocks by changing temperature, we do this by adding a small amount of water to the solvent (which is initially dimethyl sulfoxide). Addition of just a small amount of water increases the effective value ofχ characterizing the solvent/PMMA interaction. The result is that the relaxation times for the triblock copolymer solution increase dramatically, as illustrated by the rheological data in Figure97.16. The effects of solvent composition can be illustrated by introducing a shift factor,a_s that depends on the composition of the solvent. It's use is illustrated in Figure97.17. The viscosity at a given temperature is proportional toa_s, so we see that small increases in the water content of the solvent result in an increase in the solution viscosity by several orders of magnitude. This occurs because water induces the aggregation of PMMA blocks in to discrete domains, just as reducing temperature did for the case where alcohol was used as the solvent.

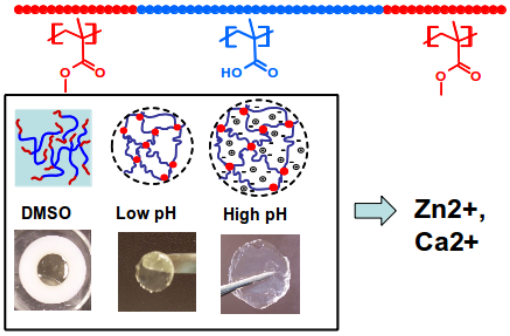


Figure 97.15: Formation and ionic crosslinking of PMMA-PMAA-PMMA triblock copolymer hydrogels (PMAA = poly(methacrylic acid).

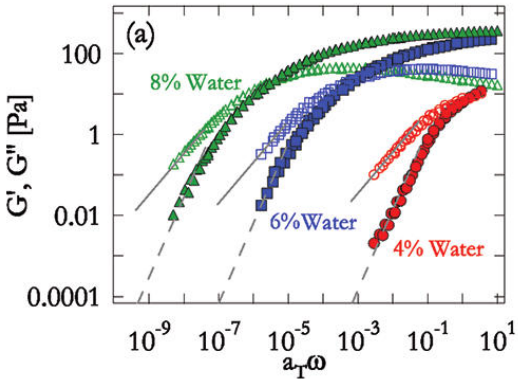


Figure 97.16: Master plots of storage and loss moduli for PMMA-PMAA-PMMA triblock copolymer gels in DMSO/water mixtures, for water contents within the solvent of 4, 6 and 8 wt. %.

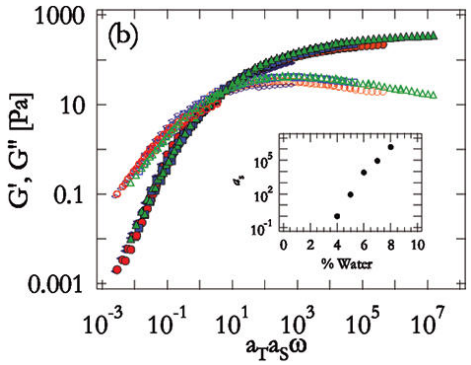


Figure 97.17: 'Super Master' plot of the data from Figure.

A useful feature of the hydrogels formed from the triblock copolymers with the poly(acrylic acid midblocks) is that the properties can be changed dramatically by immersing the gels in solutions containing divalent cations like Ca^{2+} or Zn^{2+}. These ions add additional crosslinks to the material, increasing the modulus by a factor of 100 or more (Figure97.18). The materials also have high toughness, and can be extended to several times their original length prior to failure (Figure97.19).

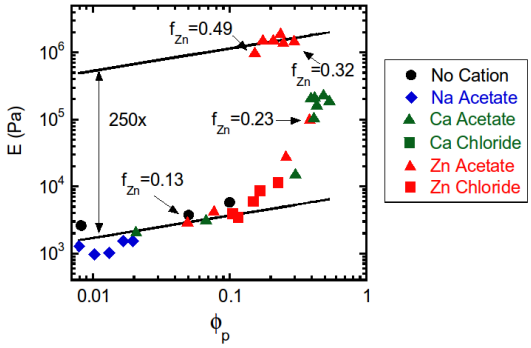


Figure 97.18: Elastic modulus for triblock copolymer gels that have been ionically crosslinked to varying degrees.

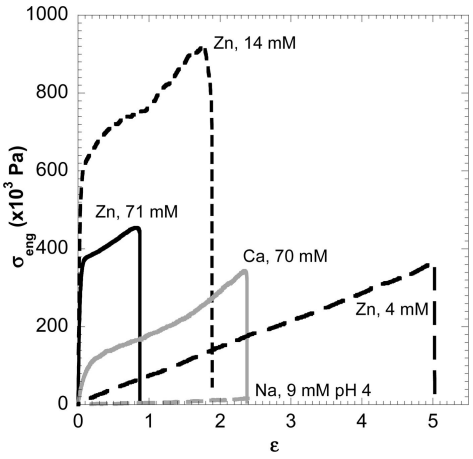


Figure 97.19: Tensile stress/strain curves for ionically crosslinked triblock copolymer gels.

98 Summary

98.1 Classification Scheme

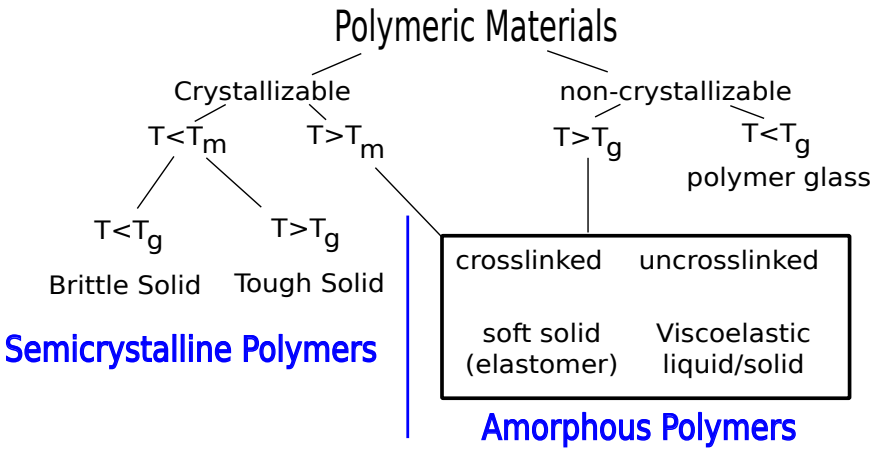


Figure 98.1: Summary of polymer properties and classes

98.2 Molecular Weight Averages

NDegree of polymerization
MMolecular weight
M_0Molecular weight per repeat unit

$$M_n = w/n$$

$$M = NM_0$$

$$M_w = \frac{1}{w}\sum_i w_i M_i = \text{Weight Average Molecular}$$

w = total weight of polymer
n = total number of polymer molecules
w_i = weight of fraction i (grams)
M_i = molecular weight of fraction i

98.3 Polymerizations

98.3.1 Linear step growth polymerization

$$M_n = \frac{M_0}{2(1-p)}$$

$$\frac{M_w}{M_n} = 1 + p$$

$$n = n_a - n_b$$

(p = extent of reaction)

98.3.2 Nonlinear step growth polymerization

$$M_n = \frac{M_0}{2 - pf_{av}}$$

$$(n_a + n_b)f_{av} = n_a f_a + n_b f_b$$

f_a,f_b = functionality of A and B monomers
n_a,n_b = numbers of A and B monomers
pf_{av} = 2 at the gel point

98.3.3 Chain Growth Polymerization

Active site is typically a free radical, anion, or cation.
Propagation corresponds to addition to a double bond, or opening of a ring.
Termination can occur by Disproportionation or combination.
Intermolecular chain transfer reduces molecular weight.
Intramolecular chain transfer results in chain branching.

98.4 Chain Dimensions

98.4.1 Structure of Amorphous Polymers

The probability density distribution for the end to end vector, R (with componentsR_x,R_y andR_z) is:

$$P(R_x, R_y, R_z)\exp(-\frac{3R^2}{2Na^2})$$

$$P(R)R^2\exp(-\frac{3R^2}{2Na^2})$$

Here a is the the statistical segment length, defined by the following expression:

$$R_0 \equiv < R^2 >^{1/2} = N^{1/2}a$$

98.4.2 Structure of amorphous vinyl polymers

$$R_0^2 = C_\infty N_b l^2 = Na^2$$

$$C_\infty = \left\{\frac{1+\cos\theta}{1-\cos\theta}\right\}\left\{\frac{1+ < \cos\phi >}{1- < \cos\phi >}\right\}$$

θ = bond angle

ϕ = rotation angle

98.4.3 Glass Transition

$$\frac{V_f}{V} = (\alpha_l - \alpha_g)(T_g - T_\infty)$$

(Free volume theory)

98.4.4 Differential Scanning Calorimetry

Thermal signatures of melting, crystallization and glass formation

98.5 Crystallization

Polymer Crystallization

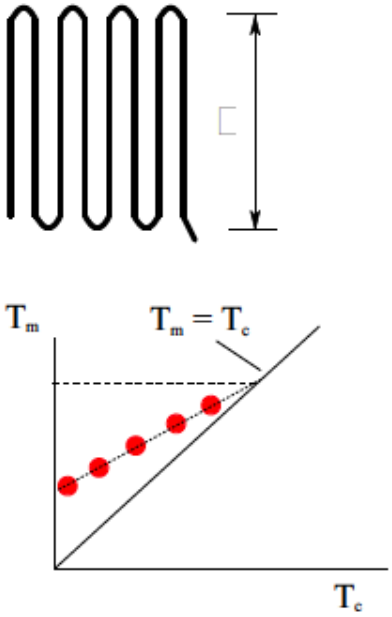


Figure 98.2: Polymer crystallization

$\lambda \approx \frac{2\gamma_e T_m^0}{\Delta h \Delta T}$
% crystallinity $=\ell/d$
x*y/z notation for helix formation
spherulitic morphologies

98.6 Rubber Elasticity

Single Chain:

$$\Delta F_d = \frac{3k_B T R^2}{2Na^2}$$

Bulk Elastomer:

$$\Delta f_d = \frac{v_s k_B T}{2}\left\{\lambda_x^2 + \lambda_y^2 + \lambda_z^2 - 3\right\}$$

Shear:$\sigma_{xy} = G\gamma$;

$$G = v_s k_B T = \frac{\rho RT}{M}$$

Uniaxial Extension or Compression:

$$\sigma_{eng} = G\left\{\lambda_y - \frac{1}{\lambda_y^2}\right\}$$

Δf_d = deformation free energy (per volume)
σ_{eng} = engineering stress
σ_{xy} = shear stress
γ = shear strain
v_s = strand concentration
λ_x,λ_y,λ_z = extension ratios

98.7 Viscoelasticity

Response to step strain:

$$G(t) \equiv \frac{\sigma_{xy}(t)}{\gamma(0)}$$

Boltzmann superposition principle:

$$\sigma_{xy}(t) = \int_0^\infty G(s)\gamma(t-s)ds$$

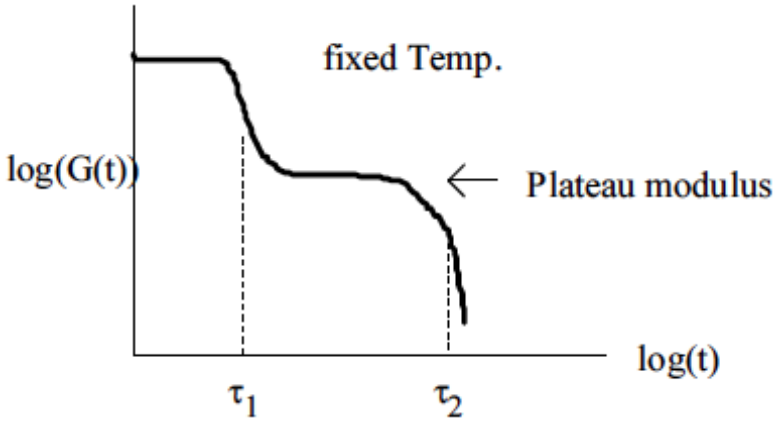


Figure 98.3: Bolzmann superposition principle on a graph

$\dot{\gamma}$ = strain rate= $\frac{d\gamma}{dt}$
Dynamic Mechanical Spectroscopy

$$\gamma = \gamma_0 \sin(\omega t)$$

$$\gamma(t) = \omega\gamma_0 \cos(\omega t)$$

$$\frac{\sigma_{12}(t)}{\gamma_0} = G'(w)\sin(\omega t) + G''(w)\cos(\omega t)$$

$$\tan\delta = \frac{G''(\omega)}{G'(\omega)}$$

G' = storage modulus (elastic response)

G''= loss modulus (viscous response)
δ = phase angle

$$\eta_0 = \lim_{\omega \to 0} \frac{G''(\omega)}{\omega} = \int_0^\infty G(t)dt = A \exp\left(\frac{B}{T - T_\infty}\right)$$

Time-temperature superposition: Changes in temperature shift the G(t), G'(ω) and G"(ω) curves horizontally (when plotted against log(t) or log(ω)), but the shapes of the curves do not

98.8 Solutions and Blends

Free energy of mixing

$$\Delta f_{mix} = \frac{RT}{V_0}\left\{\frac{\phi_a \ln \phi_a}{N_a} + \frac{\phi_b \ln \phi}{N_b} + \chi \phi_a \phi_b\right\}$$

$$\pi = -\frac{\mu_s}{v_0}$$

Chemical potential:

$$\frac{\mu_a}{RT} = \ln \phi_a + \phi_b(1 - \frac{N_a}{N_b}) + \chi N_a \phi_b^2$$

$\pi = \frac{RTc}{M_n}$ in dilute limit

Spinodal curve:

$$\chi_s = \frac{1}{2\phi_a N_a} + \frac{1}{2(1-\phi_a)N_b}$$

Critical point:

$$\phi_{a,crit} = \frac{\sqrt{N_b}}{\sqrt{N_b} + \sqrt{N_a}}$$

$$\chi_{crit} = \frac{N_a + N_b + 2\sqrt{N_a N_b}}{2N_a N_b} = \begin{cases} 0.5 \text{ for } N_a = 1,\ N_b \to \infty \\ \frac{2}{N} for N_a = N_b = N \end{cases}$$

98.9 Interfaces

Polymer/Polymer Interfaces

$$w = \frac{2a}{\sqrt{6\chi}}$$

$$\gamma = \frac{a}{V_0} RT \sqrt{\chi/6}$$

99 Review Questions

99.1 Polymer Synthesis

- Can you move between repeat unit structures and monomer structures for linear polyesters, polyamides and polyurethanes?
- What are the lewis structures of the terminal carbon for anionic, cationic and free radical polymerizations?
  - What happens to these structures during radial recombination, disproportionation and chain transfer reactions?
- What is a condensation reaction?
- Are all step growth reactions condensation reactions?
- What is a living polymerization?
  - Why are free radical polymerizations generally not living?
- How are block copolymers synthesized?

99.2 Molecular Weight Distributions

- How are the degree of polymerization and the molecular weight related to each other?
- What does the molecular size distribution look like for an ideal step growth polymerization? How doM_n,M_w and the polydispersity index depend on the extent of reaction?
- How areM_w,M_n and the polydispersity determined from a given molecular weight distribution?

99.3 Gels and Elastomers

- How do epoxies react with primary amines to give a gel?
- How is the extent of reaction at the gel point determined from the Carothers theory?
- Why is the Carothers theory only an approximation?
- How can monomers be chosen to form a crosslinked polymer by either chain growth or step growth polymerization?

99.4 Glass transition

- What happens to the free volume at the glass transition?
- What happens to the molecular structure and mechanical stiffness of an amorphous polymer at the glass transition temperature?
- How does molecular structure (backbone flexibility, bulky or thin side groups) affect the location of the glass transition temperature?
- What isT_∞ and what dos it have to do with the glass transition?
- In the free volume theory of the glass transition, what do the temperature dependencies of the free volume and occupied volume look like?

99.5 Chain dimensions of amorphous polymers

- What is the characteristic ratioC_∞?
  - What is its definition in terms of real, measurable quantities?
  - How is it related to the statistical segment length?
  - How can it be calculated from a given distribution of bond rotation angles?
  - What is it for the freely jointed and freely rotating chain models?
- What are the bond angles and bond lengths for tetrahedral bonding?
- What does the distribution of end-to-end distances look like for an amorphous polymer?
- How does this depend on the molecular weight?

99.6 Rubber Elasticity

- What is the force required to hold the ends of a polymer molecule at a fixed separation? How is this force determined from the free energy? What are the origins of this free energy, and what is it quantitatively?
- How can one calculate the shear modulus, Young's modulus, and deformation free energy for an elastomer?
- Why does the modulus of an elastomer increase with temperature?
- What happens to the free energy, enthalpy, and entropy of a rubber as it is stretched?
- What are typical values of the elastic moduli for rubbery, glassy and semicrystalline polymers?

99.7 Viscoelasticity

- What do the relaxation moduli for crosslinked and crosslinked polymers look like?
- How do the characteristic relaxation times associated with the glass transition and molecular disentanglement depend on temperature and molecular weight?
- How can data collected at different temperatures be used to obtain data corresponding to a very large range of frequencies?
- What does the temperature dependence of the shift factors (a_T) look like for an amorphous polymer above its glass transition?
- How is the viscosity of a polymer related to the time dependent relaxation modulus (G(t)) and to the dynamic moduli ($G'(\omega)$ and/or$G''(\omega)$)?
- How is the entanglement molecular weight determined from experimental data?

99.8 Semicrystalline Polymers

- What is the definition of tacticity?
- How does tacticity relate to a materials ability to crystallize?
- What does the x*y/z helix notation mean?
- Are the crystalline regions of a semicrystalline polymer more dense or less dense than the amorphous regions?
- How is the density of a perfect polymer crystal calculated?
- How are the orientations of polymer backbones, lamellae, and spherulites related to one another?
- How are crystallization and melting temperatures related, and how can one determine the equilibrium melting point for a polymeric material?

- What is meant by the fold surface energy? How does the crystal thickness depend on undercooling, and what is the qualitative explanation for this effect?
- Why does the melting temperature of a polymeric material sometimes depend on the temperature at which it was crystallized?
- Why aren't semicrystalline polymers transparent?
- What happens to the molecular structure and mechanical stiffness of a polymer sample during crystallization?

99.9 Solution Thermodynamics

- How are phase diagrams for binary (two component) systems related to the free energy of mixing?
- Why is the magnitude of the entropy of mixing for polymer/polymer mixtures so small?
- What is the critical point? Where does it occur for a symmetric polymer blend, or for a polymer/solvent system?
- How can the osmotic pressure of a polymer/solvent system be measured?
- What is the significance of the correlation length, and how is it related to the osmotic pressure?
- How can the equilibrium concentration of solvent in a high molecular weight polymer be predicted for the case where the solvent does not completely dissolve the polymer?
- What combinations of entropy and enthalpy of mixing give USCT (upper critical solution temperature) behavior? What about LCST (lower critical solution temperature) behavior?
- Why is LCST behavior more common than UCST behavior for polymer mixtures?
- Why is USCT behavior more common for polymers in organic solvents, whereas LCST behavior is more common for polymers dissolved in water.

99.10 Interfaces and Self-Assembly

- How do you calculate the interfacial free energy between immiscible polymers?
- What factors determine the equilibrium lamellar period for a diblock copolymer? How is this number calculated quantitatively?

100 331 Problems

100.1 Basic chemistry review

Some basic problems having to do with the chemical structures of polymers.

(1) Calculate the repeat unit molecular weights for each of the 22 polymers shown in the 'common polymers' section of the text.

(2) Draw Lewis diagrams illustrating the valence shell configurations for polystyrene, poly(methyl methacrylate) and poly(dimethylsiloxane).

(3) Draw Lewis diagrams illustrating the valence shell configurations for amide, ester and urethane linkages.

100.2 Polymerization Reactions

(4) Consider the following monomers:

a List all of the monomers that will polymerize by chain growth polymerization.

b Draw the repeat unit structures of 3 different polymer structures that can be formed from (c)

c Which monomers will react with each other to form a linear polyamide?

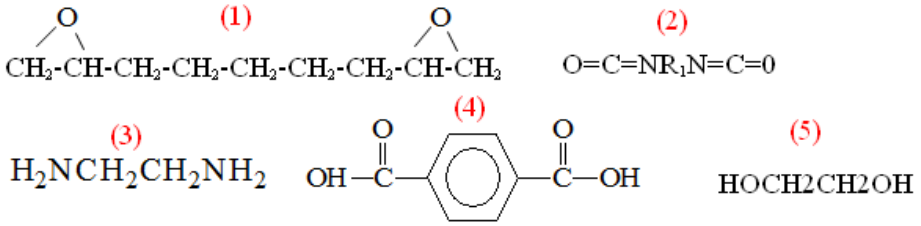
d Which monomers can you combine to produce a highly crosslinked polymer network by step growth polymerization so that no condensation products are produced?

(5) Suppose a polyester is formed by combining 1.00 mol of a diacid (2 acid groups per molecule) with 1.01 mol of a dialcohol in a homogeneous solution. The molecular weight of the diacid is 160g/mol and the molecular weight of the dialcohol is 180 g/mol.

a What is the highest value of M_n that can be expected from this combination?

b How many grams of water will be produced during the polymerization reaction?

(6) Consider the following 5 monomers:



a Draw the repeat units for three linear polymers that can be produced by reactions between the monomers in this list. Identify these polymers according to their type (polyamides, polyesters, etc.), and indicate whether each polymerization reaction is condensation reaction or not.

b What combination of monomers from this list would you choose in order to produce a three-dimensional network?

(7) "Living" polymerizations are chain growth polymerizations that proceed without termination or chain transfer reactions.

a Discuss the conditions that must be met in order for polymers with a low polydispersity to be formed by a living polymerization.

b What type of reactions must be avoided in order to produce block copolymers by radical chain growth polymerization.

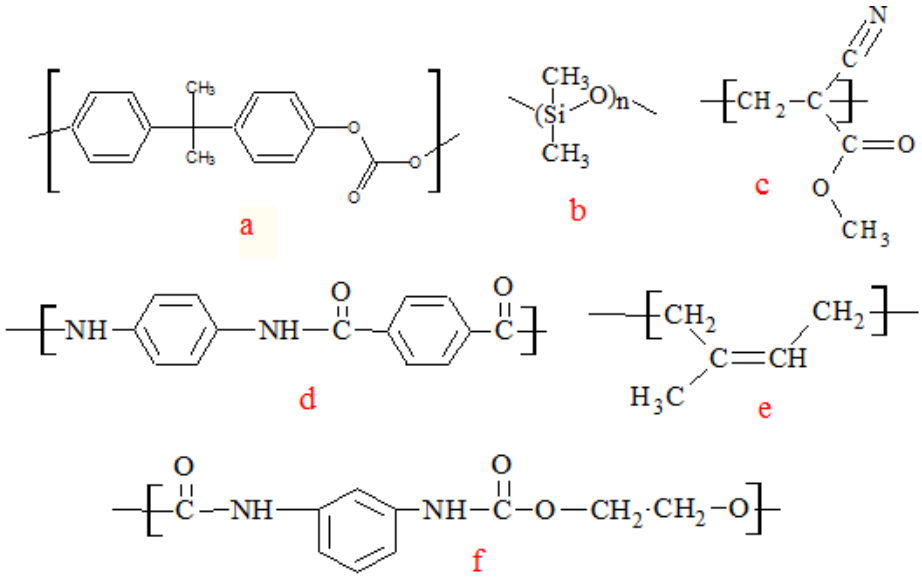
c Proteins are perfectly disperse copolymers, with a polydispersity index of exactly 1. While relatively monodisperse polymers ($M_w/M_n < 1.05$) can be made by living polymerization methods like anionic polymerization, it is not possible to make perfectly monodisperse polymer by these methods, even when termination and chain transfer events are completely eliminated. Why is that?

(10) Suppose nylon is formed in two separate batches by reacting the same sets of monomers with one another. In the first reaction, an extent of reaction of 0.999 is obtained. In the second reaction, an extent of reaction of 0.9999 is obtained. Assume that an ideal distribution of molecular weights is obtained for each reaction.

a What is the ratio of M_w for the second reaction to M_w for the first reaction?

b Suppose equal weights of the polymers formed by these two reactions are mixed with one another. What is the polydispersity index of the resultant mixture?

(11) Draw the chemical structures of the monomers from which the following polymers are formed.



(12) Find answers to the following problems:

a Draw the repeat unit structure of a polyamide formed by a condensation polymerization involving a combination of the above molecules, and identify the condensation product.

b Draw the structure of a polymer formed by ring opening polymerization of one of these molecules.

c Which of the molecules shown above can be polymerized by chain growth polymerization? Draw the chemical structure of the resultant polymers.

(13) Suppose you want to sell a set of cheap plastic mugs which are suitable for drinking coffee. Briefly discuss the potential applicability of the following materials for this purpose:

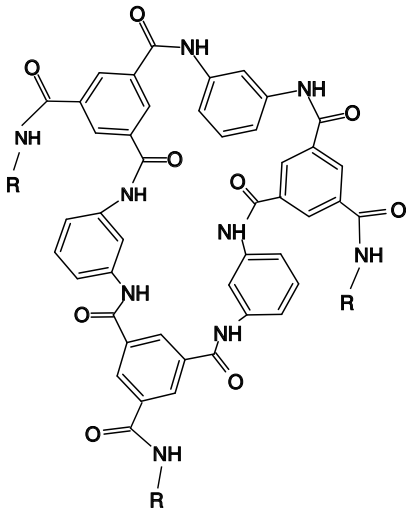
(a) atactic polystyrene

(b) atactic poly(vinyl chloride)

(c) high density polyethylene

(14) Consider the the piece of the highly crosslinked chemical structure shown below.

a Draw the chemical structures of the two different monomers that were used to form this material. (Note that the R groups just represent connections to the rest of the material, they are not part of the structures you need to draw).



b What type of polymer is this (polyester, polyamide, polyurethane, epoxy, etc.)

c Next to each structure that you have drawn in part a, indicate its chemical functionality.

d What is the relative amount of each of the two monomers from part a (in moles) that you should be added to one another to optimize the polymerization?

e Assuming that the two monomers are added together in optimum proportions you mentioned in part d, what fraction of reactive groups need to react to form an insoluble gel.

100.3 The Glass Transition

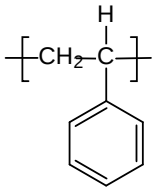
(15) An amorphous polymer has$T_\infty = 50°C$. The thermal expansion coefficient in the glassy state is$3 \times 10^{-4}/°C$, and the thermal expansion coefficient in the liquid state is$4.5 \times 10^{-4}/°C$.

a Calculate the glass transition temperature if the fractional free volume at the glass transition is 0.025.

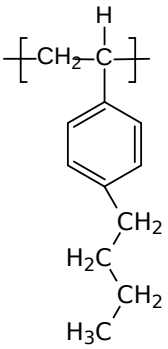
b The glass transition temperature measured at a different cooling rate is found to correspond to a fractional free volume of 0.021. Is this cooling rate higher of lower than the cooling rate from part a?

(16) Rank these three polymers in the order of highT_g to lowT_g, and provide an explanation for your prediction:

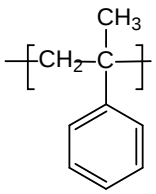
Polystyrene



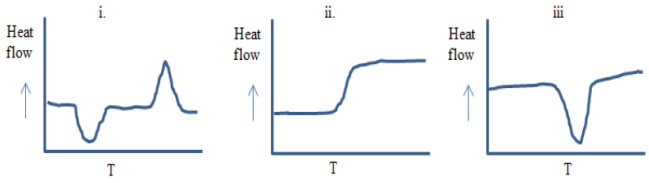
Poly(para butylstyrene)



Poly(α-methylstyrene)



(17) The DSC curves below correspond to the 3 polymers listed below. Identify which is which, and identify on the curves any crystallization (T_c), glass transition (T_g) or melting (T_m) temperatures, and give your reasoning in each case.



a Atactic polystyrene being heated from room temperature to 200^{o}C.

b High density polyethylene being cooled from 200^{o}C to room temperature.

c Poly(ethylene terephthalate) that had originally been cooled rapidly from 300^{o}C to room temperature, and is now being slowly heated from room temperature back up to 300^{o}C.

(18) Short, bulky side groups (like methyl of phenyl groups) increase the fractional free volume required for the motion of polymer chain segments, whereas long, thin groups (like the n-butyl group -$CH_2CH_2CH_2CH_3$) increase the available free volume.

a Describe how each of these two types of groups is expected to affect the glass transition temperature.

b Polymers with very little flexibility along the backbone of the polymer chain have very high glass transition temperatures. Give a brief explanation for this in terms of free volume theory.

100.4 Chain Dimensions

(19) The value ofR_0, the rms (root mean squared) end-to-end distance for a polystyrene molecule in the bulk material (no solvent) is 200Å. The molecular weight, M, of this molecule is 100, 000 g/mol.

a What is the rms end-to-end distance for a polystyrene molecule with M=250,000 g/mol?

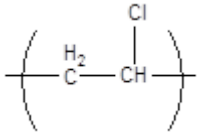
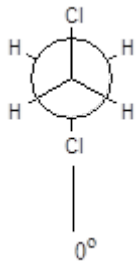
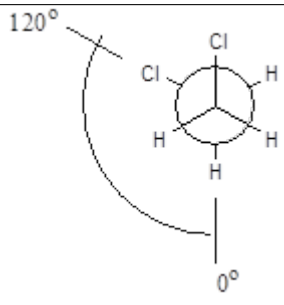
b What is the characteristic ratio (C_∞) for polystyrene?

c What is the statistical segment length (equivalent length of a repeat unit in a random walk model of chain dimensions) for polystyrene?

d The volume pervaded by a single molecule is given approximately by R_0^3. For a polystyrene molecule with M=250, 000 g/mol, compare this volume to the volume actually occupied by the polystyrene molecule. How many different molecules overlap with one another in polystyrene with this molecular weight?

e Consider a simple model of polystyrene for which all of the bonds are in either gauche or trans configurations. What fraction of bonds are in the trans configuration?

(20) Consider a simple model for the chain dimensions of amorphous poly(vinyl chloride) where only trans and gauche bonds are possible.

Poly(vinyl chloride)	Trans Configuration	Gauche Configuration
		

a What fraction of the backbone bonds are in the gauche configuration if C_∞ for the polymer is 8?

b Calculate the statistical segment length of a repeat unit, assuming the value of C_∞ from part a.

(21) What fraction of molecules in an equilibrated amorphous polymer melt have end-to-end vectors with magnitudes larger than the root-mean-square end-to-end distance? (note – this will require a numerical or graphical solution.);

100.5 Mechanical Properties

(22) A thermoset is cooled to a very low temperature. As the temperature is increased, the elastic modulus remains roughly constant until T_g, at which point the elastic modulus rapidly decreases. Above the glass transition temperature the elastic modulus slowly increases. Describe in qualitative terms what is going on at the glass transition, and why the modulus actually increases slightly above T_g.

(23) Values of Young's moduli (E) for elastomers are typically near 10^6 Pa (1 MPa).

a What is the average molecular weight between crosslinks for a polymer with this value of E? Assume the polymer has a specific gravity of 1. What fraction of cis 1-4 polyisoprene repeat units contain crosslinks if the modulus is 10^6 Pa?

b Plot the predicted tensile stress/strain curve for this material for engineering strains from -0.5 to 2 (extension ratios from 0.5 to 3).

c The yield point of a material is often defined as the onset of nonlinear behavior in the stress-strain curve. Can this definition be used for an elastomeric material that is deformed in tension? Why or why not?

(24) Suppose a sample of Sylgard 184 is cured, and the degree of polymerization of the prepolymer is 60. Estimate the value for both the shear modulus and Young's modulus of this polymer. Describe any assumptions that you've made.

(25) Consider the deformation behavior of a uniaxially stretched rubber.

a Plot the free energy of a uniaxially stretched rubber as a function of the extension ratio fromλ = 0.1 (compressive strain toλ =4 (tensile strain). (The extension ratio of interest here corresponds to the direction in which the rubber is stretched.)

b At what value of the strain is the free energy minimized?

c For what compressive value of the extension ratio is the free energy of the rubber equal to the free energy it has for an extension ratio equal to 4?

(26) Calculate the force required to fix the ends of a 1,4 polyisoprene molecule with M = 100,000 g/mol at the separation of 1000Å. AssumeC_∞ = 8, and that the double and single bonds have equal lengths of 1.54 Å.

(27) Consider a material which is defined between two circular plates that are parallel to one another. Each plate has a diameter of 2.5 cm, and the plates are separated by 1mm.

a Calculate the maximum shear strain at the outer periphery of the disks, if the top disk is rotated by$1^\circ$with respect to the bottom disk.

b Calculate the torque required to rotate the top plate by1°, in terms of the magnitude of the complex shear modulus,$G*$. (You will need to integrate contributions from all portions of the disk, keeping in mind that the local stress is equal to the shear modulus multiplied by the local strain.)

c Suppose the material between the two plates is a purely viscous fluid ($G' = 0$ at the appropriate frequency), and that the viscosity of the fluid is 1000 Pa-sec. Calculate the magnitude of the torque which must be applied to give a sinusoidal rotation with a magnitude of1°. Assume that the angular frequency of the applied torque is 100 radians/second.

(28) Consider a polymer with a density of $1 g/cm^3$ that has the following relaxation modulus at room temperature(the molecular weight is much higher than the average molecular weight between entanglements).

$$G(t) = G_1 \exp\left(\frac{-t}{\tau_1}\right) + G_2 \exp\left(\frac{-t}{\tau_2}\right)$$

Suppose$G_1 = 10^9$ Pa ,G_2 =$3x10^5$ Pa,τ_1 =10^{-6} s andτ_2 = 1 s.

a Calculate the zero shear viscosity for this polymer.

b Is this polymer crosslinked? How do you know?

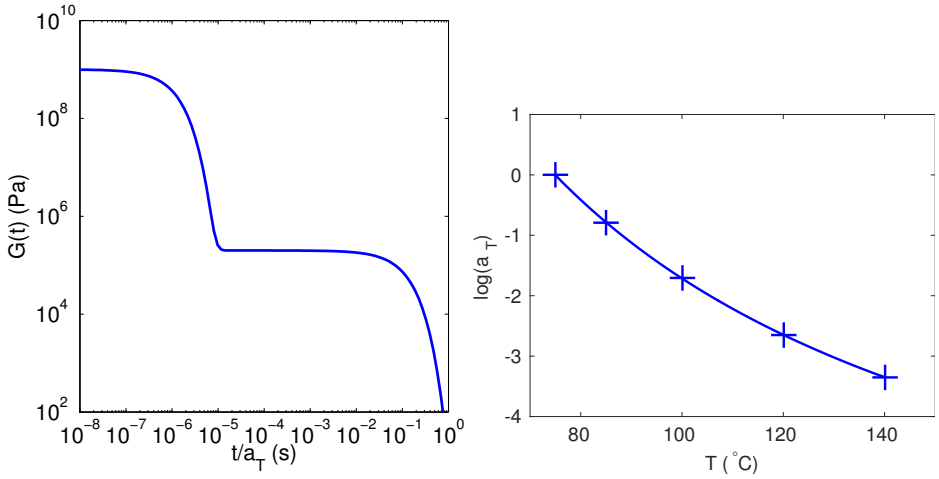
c Calculate the average molecular weight between entanglements.

d Describe what happens to τ_1 and τ_2 when the temperature is increased (do they remain constant, increase or decrease)?

e What vibrational frequencies do you expect to be most efficiently damped by this polymer?

f Describe what happens to τ_1 and τ_2 when the molecular weight of the polymer is increased (do they remain constant, increase or decrease)?

(29) Consider a polymer with the behavior of the relaxation modulus ($G(t)$) and temperature shift factor(a_T) plotted below. Assume the polymer has a density of 1 g/cm^3.



a Describe in qualitative terms what is going on at each of the two characteristic values of t/a_T where $G(t)$ decreases substantially.

b Is this polymer crosslinked? Briefly explain your answer.

c Estimate the entanglement molecular weight for this polymer.

d Estimate the viscosity of this polymer that is obtained at very low shear rates, at a temperature of 40°C.

e What is the ratio of viscosity of the polymer at -40°C to the viscosity at 40°C?

f (5 points) Suppose that $G'(\omega)$ is measured at a very high frequency of 10^9 s^{-1}. What value of G' do you expect to obtain from this experiment.

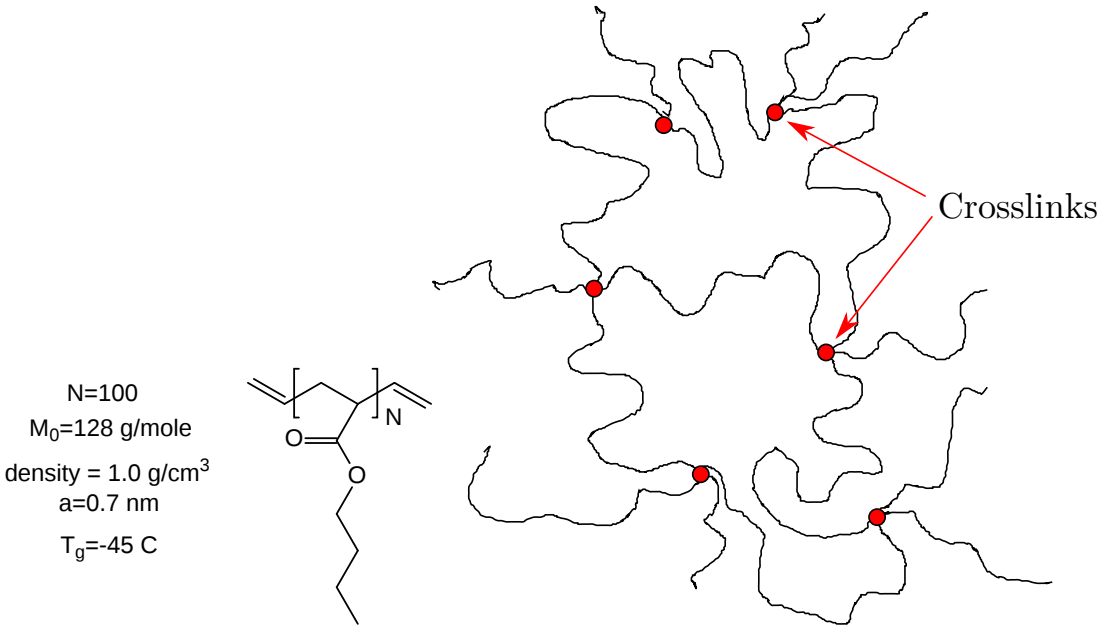
(30) Determine the shift factors as a function of temperature for poly(t-butyl acrylate) (PtBA), using 75°C as the reference temperature. Plot these shift factors as a function of temperature, and determine a value for T_∞ by fitting to the Vogel equation. Compare this value of T_∞ to T_g for PtBA. Values of G' and G'' at different temperatures are included in the data tables at the end of the text.

(31) Determine the zero shear viscosity of PtBA with M=350,000 g/mol at 75°C.

(32) Determine the entanglement molecular weight,M_e, for PTBA. Assume a density of 1 g/cm^{3}for PtBA.

(33) Repeat the previous 3 problems for poly(2-vinylpyridine) with a molecular weight of 82,000 g/mol, using160°C as the reference temperature (where$a_T = 1$), and130° C for the viscosity calculation. (Data included with the data tables at the end of the text).

(34) An elastomer is made from a poly(n-butyl acrylate) precursor molecule with double bonds at either end that are able to react with a tetrafunctional crosslinker. The chemical structure is shown at the left below, and a schematic representation of the crosslinked structure is shown at the right.



g Calculate the shear modulus, G, for this material at room temperature.

h Give Young's modulus (E) and the Poisson's ratio (ν) for this material.

i What is the root mean square end-to-end distance$\langle R^2 \rangle^{1/2}$ for crosslink points at either end of the one of the precursor chains when the material is in its undeformed state?

j What is the value ofC_∞ for poly(n-butyl acrylate)?

k Express the deformation free energies of the elastomer, divided by the shear modulus,G, for the following strain states:

i. Simple shear at a shear strain,γ, of 1.

ii. Uniaxial extension to an engineering strain of 100 % ($\lambda = 2$).

iii. Uniaxial compression to a engineering strain of -50% ($\lambda = 0.5$).

(35) For each if the following pairs of quantities, circle the**largest** one. If you expect them to have the same value, write 'same'.

quantity 1	quantity 2
The thermal expansion coefficient of poly(n-butyl acrylate) at temperatures**above** -45 C.	The thermal expansion coefficient of poly(n-butyl acrylate) at temperatures**below** -45 C.
the free volume for a poly(n-butyl acrylate) sample**above** -45 C.	the free volume for a poly(n-butyl acrylate) sample**below** -45 C.
The**free energy** of an undeformed poly(n-butyl acrylate) sample.	The**free energy** of poly(n-butyl acrylate) sample at a shear strain of 0.3
The**entropy** of an undeformed poly(n-butyl acrylate) sample.	The**entropy** of poly(n-butyl acrylate) sample at a shear strain of 0.3
T_g for poly(n-butyl acrylate)	T_g for poly(methyl acrylate) (where the 4-carbon butyl group in poly(n-butyl acrylate) is replaced by a methyl group.

100.6 Crystallinity

(36) Consider a polyethylene molecule with a molecular weight of 150,000 g/mol.

a Estimate the number of times the molecule folds back on itself within a lamellar crystallite with a thickness of 100Å. Assume for the purposes of this calculation that the molecule remains in the crystallite, and does not move into the amorphous regions of the polymer.

b From your answer to part a, what can you say about the ability to stretch a polymer molecule in a semicrystalline polymer without breaking any bonds?

(37) Describe how axis of the unit cell in polyethylene is oriented with respect to a spherulite, with respect to an individual lamellar crystallite, and with respect to the molecular chain axis.

(38) Consider a polymer that crystallizes with an orthorhombic unit cell (all angle= 90°). The polymer has a carbon backbone, and crystallizes with a 4*1/1 helical structure. The unit cell dimension are a=10.4Å, b=5.3Å with 1.27Å per backbone carbon along the c axis. The molecular weight of a polymeric repeat unit is 125 g/mol, and there are two of these units per unit cell.

c What is the theoretical density of a completely crystalline version of this polymer (100% crystalline)?

d The percent crystallinity of a sample with a density of 1.35 g/cc is 60%. What is the density of the amorphous regions of this polymer?

(39) Poly(ethylene terephthalate) has a relatively high equilibrium melting point (265 °C), but the polymer at room temperature is almost completely amorphous. Based on this result, do you expect thatT_g for PET is above room temperature or below room temperature? Explain your reasoning.

(40) Consider two perfect single crystals of polyethylene and polypropylene. Polypropylene forms a 3/1 helix, and polyethylene forms a 2/1 helix. Which of these single crystals would you expect to have a higher modulus along the C axis? Include in your discussion the relative importance of bond bending and bond rotation during elastic deformation of the two materials. Also mention the importance of the lateral packing (i.e., the number of helices per unit area in the AB plane).

(41) High density polyethylene has an equilibrium melting temperature of 138 °C. Two identical batches of this polymer are crystallized at different temperatures: one at 100 °C and one at 125 °C. Describe one important difference in the resulting structures of these two semicrystalline polymers, and explain the origin of the difference.

(42) A linear homopolymer was crystallized from the melt at crystallization temperatures (T_c) within the range 270K to 330K. Following complete crystallization the following melting temperatures (T_m) were measured by differential scanning calorimetry (DSC).

Given:

T_c(K)	T_m(K)
270	300
280	306.5
290	312.5
300	319.0
310	325.0
320	331.0
330	337.5

Graphically determine the equilibrium melting temperature,T_m^0.

(43) Small angle x-ray scattering experiments using x-rays with a wavelength of 1.54 Å were used to measure the distance between crystalline regions. Peaks corresponding to the lamellar period were obtained at the following angles:

$T_c(K)$	$\theta(degrees)$
270	0.44
280	0.39
290	0.33
300	0.27
310	0.22
320	0.17
330	0.11

a Use the Bragg equation to calculate the values of the lamellar period for each crystallization temperature.

b The degree of crystallinity was measured to be 45% for all samples. Calculate the lamellar thickness in each case, and use this information to graphically obtain the fold surface energy,γ_e. Assume the enthalpy of fusion of the polymer is $150 J/cm^3$, and the value of the equilibrium melting temperature from the previous problem.

(44) Poly(ethylene oxide) (PEO) is a semicrystalline polymer with melting point of 65°C and a glass transition temperature of -60°C. Suppose a series of crystalline samples is made by performing the crystallization at temperatures of 40°C and 60°C.

a Describe any differences you expect rate of crystallization, and the crystal thicknessℓ_c, at these two temperatures. Be as quantitative as possible.

b Suppose the two samples are slowly heated. Describe any differences you expert to observe in the melting temperatures for these two samples.

c PEO is always crystalline at room temperature, but PET is often amorphous at room temperature, even though PET has a much higher melting temperature than PEO. Why is this?

100.7 Liquid Crystals

(45) Consider the liquid crystalline texture from the nematic liquid crystal texture page of a nematic liquid crystalline material confined between two crossed polarizers.

a Give an example of the type of defect that can lie at the places where 4 dark bands come together. (show how the molecules are oriented in these regions)

b Describe why alternating light and dark bands emanate from these defect points.

c Is it possible to obtain a completely black image from a sample that has good nematic order? If so, describe a situation where this would be the case.

100.8 Thermodynamics of Solutions and Blends

(46) Which contribution to the free energy (enthalpy of entropy) dominates at low temperature?

(47) What are the signs of$\triangle h_{mix}$and$\triangle s_{mix}$for a system which has an LCST phase diagram?

(48) For what values ofχ will a polymer be able to completely dissolve in a small-molecular solvent?

(49) Measured values ofχ for a blend of polymer A with polymer B have the following temperature dependence:

$$\chi = \frac{0.326}{T} - 2.3 \times 10^{-4}$$

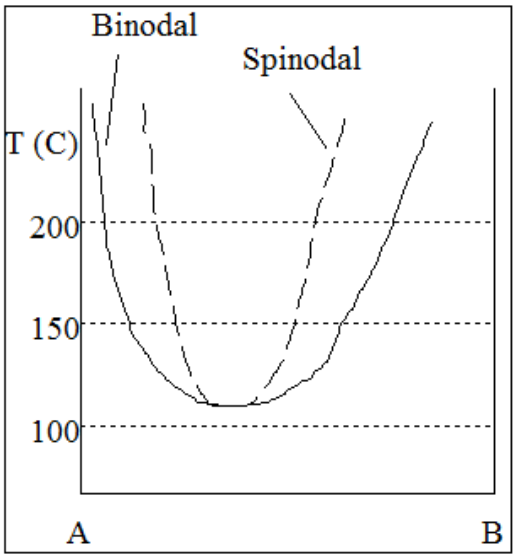
where T is the absolute temperature.

Consider a blend with$N_a = 2500, N_b = 3500$

a What is the predicted critical composition, and the predicted critical temperature for this blend?

b Plot free energy curves (free energy as a function of composition) forχ/χ_{crit}= 1.1, 1.2, 1.3, 1.4, and 1.5. Using the common tangent construction from these five curves, construct a phase diagram withχ/χ_{crit} on the vertical axis and the blend composition on the horizontal axis.

(50) Consider the following phase diagram:



a Is the enthalpy of mixing positive or negative for this system?

b Is the entropy of mixing positive or negative for this system?

c Draw free energy of mixing vs. composition curves at$100^\circ, 150^\circ, 200^\circ$ that are consistent with this phase diagram. Show any common tangents points and points of inflection.

(51) Consider a polymer/solvent mixture for which the ratio of polymer to solvent volume is 10,000.

a Determine the critical value of the chi parameter (χ_{crit})

b What is polymer volume fraction at the critical point?

c Plot the solvent chemical potential for this system as a function of polymer volume fraction for$\chi/\chi_{crit} = 0.9, \chi/\chi_{crit} = 1.0, \chi/\chi_{crit} = 1.2, \chi/\chi_{crit} = 1.5$.

d Suppose that the polymer is dipped into a pure solvent (solvent chemical potential equals zero). Solvent will diffuse into the polymer until the solvent chemical potential in the polymer reaches zero. From the plots made for part c, determine the equilibrium solvent volume fractions for$\chi/\chi_{crit} = 1.2$, and$\chi/\chi_{crit} = 1.5$.

(52) Describe in your own words why most pairs of different types of polymers do not mix with each other in the liquid state.

(53) Most liquids mix better at high temperatures than they do at low temperatures. However, polymeric liquids often mix better at low temperatures than high temperatures. Explain in as much detail as possible why this is the case.

(54) Consider the following osmotic pressure data obtained for polyethylene oxide solutions in water (1MPa =10^6 Pa) at 20 °C.

ϕ_P	$\Pi(MPa)$
0.0457	0.0411
0.0917	0.145
0.138	0.346
0.185	0.685
0.232	1.212
0.280	2.010
0.329	3.161
0.377	4.868
0.427	7.323
0.476	10.77
0.516	14.54

a Plot the osmotic pressure as a function of the polymer volume fraction, using logarithmic axes for both the pressure and the polymer volume fraction. On the same plot, include the Flory Huggins expression forΠ, using the value ofχ that gives the best fit to the data.

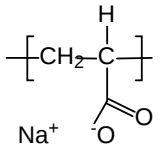
b Repeat part a, but now compare the data to the scaling expression forϕ in the semidilute regime (Eq.), wherea_o^3/v_s is the adjustable parameter. (Note thatv_s is the solvent molecular volume, obtained from the molar volume,V_s, by dividing byN_{av}).

(55) Use the PEO solution data from the previous page to calculate the following:

a The concentration of PEO at$20°C$ that gives an osmotic pressure large enough to support a column of water that is 100 meters tall.

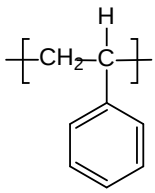
b What is the correlation lengthξ, for a PEO solution in water with$\phi_P = 0.1$?

c For charged molecules, the osmotic pressure is often dominated by the counterions. Suppose a solution of sodium polyacrylate at a concentration of $0.05 g/cm^3$is made in water. Estimate the osmotic pressure, assuming that the counterions can be treated as an ideal gas.

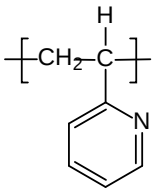


(56) The interfacial free energy characterizing the interface high molecular weight versions of polystyrene (PS) and poly(2-vinyl pyridine) (P2VP) is$3x10^{-3}$ J/m^2 at 180 °C. Assume the volume per repeat unit is the same for both polymers (165Å^3, and use this as your reference volume. Also assume that the statistical segment length for both polymers is 6.7 Å.

PS



P2VP



a What isχ for the PS/P2VP system at 180 °C?

b Suppose PS and P2VP polymers with equal degrees of polymerization are mixed with one another. What range of molecular weights will form mixtures that are completely miscible with one another at all compositions at 180 °C?

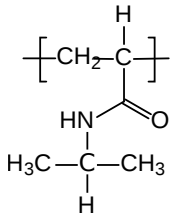
c Suppose monodisperse PS and P2VP samples with degrees of polymerization of 1000 are mixed with one another. What is the solubility of the PS in the P2VP at 180 °C? (hint: for a symmetric system like this, where both degrees of polymerization are equal to one another, the common tangent construction is satisfied at the point where the first derivative of the free energy of mixing with respect to volume fraction is equal to zero).

(57) Use the data from the PS/P2VP system from the previous page to answer the following questions:

a A diblock copolymer is synthesized that has a PS block with a degree of polymerization of 500 attached to a 2nd block with a degree of polymerization of 500. Calculate the periodicity of the lamellar structure that is formed by this material at equilibrium.

b Describe in qualitative terms the equilibrium geometries of the structures formed when the P2VP block length is fixed, but the PS block degree of polymerization is steadily increased.

(58) Poly(n-isopropylacrylamide) (chemical structure shown below) exhibits a lower crtical solution temperature in water. It is fully soluble in water at temperatures below 32°C, but the phase diagram with water has a two phase region for temperatures above 32°C.



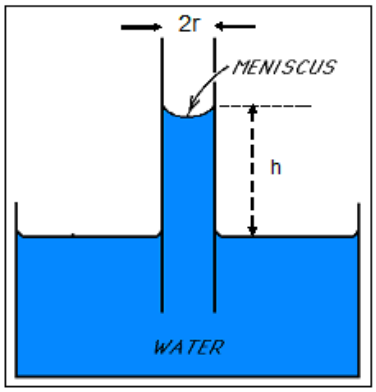
a The temperature dependence ofχ for the Poly(n-isopropylacrylamide)/water system can be written in the following form:$\chi = A + BT$, where$A$ and$B$ are constants and$T$ is the temperature in°C. What is the sign (positive or negative) of$B$ in this equation?

b Is the entropy of mixing for this system positive or negative? Briefly explain your answer.

c Calculate the osmotic pressure of an aqueous solution of high molecular weight poly(n-isopropylacrylamide) with a polymer volume fraction of 0.2.

100.9 Capillarity

(59) A hollow glass tube is immersed in water as shown in the illustration below:



a Calculate the internal tube diameter that gives a capillary rise of 10 cm, assuming that water completely wets the glass (contact angle of zero).

b Calculate the capillary rise if the same glass tube from part a is coated with a thin layer of polystyrene, changing the contact angle to 90° .

(60) Solve the following problems:

a Calculate the force required to push a cylindrical fiber with a radius of$1\mu m$ into water, if the contact angle the fiber makes with the water is$130°$.

b What stress can be supported if you create an array of these fibers on a square lattice, with a spacing between fiber centers of$4\mu m$? (Assume the density of the fibers is equal to the density of water, so you don't need to worry about forces due to buoyancy).

c How much weight could be supported by an object like this with a total area of$1cm^2$?

101 331 Computational Exercises

(61) In binary blends where the two components have very different molar volumes (often the case with polymers), the following form of the free energy expression is commonly used:

$$\frac{v_0 G_v}{k_B T} = \frac{\phi_b ln \phi_b}{N_b} + \frac{(1-\phi_b)\ln(1-\phi_b)}{N_a} + \chi \phi_b (1-\phi_b)$$

The expressions for the chemical potentials for the A and B molecules are given by the following expression:

$$\frac{\mu_a}{k_B T} = \ln(1-\phi_b) + \phi_b (1 - N_a/N_b) + \chi N_a \phi_b^2$$

$$\frac{\mu_b}{k_B T} = \ln(\phi_b) + (1-\phi_b)(1 - N_b/N_a) + \chi N_b (1-\phi_b)^2$$

Here G_v is the free energy per volume, $\phi_b$is the volume fraction of B in the system, and N_a and $N_b$are proportional to the molecular volumes of the two components. The quantity v_0 is a reference volume, so the molecular volume of the 'A' molecule is $v_0 N_a$ and the molecular volume of the 'B' molecule is $v_0 N_b$. Adapt the MATLAB script below to do the following:

```
global na nb chi % these values also get used in function definitions
set(0,'defaultaxesfontsize',16)
set(0,'defaultlinelinewidth',2)
phi=linspace(0.001,0.999,1000);
na=150;
nb=100;
chi=0.02;
% fv is the expression for the free energy of mixing
fv=@(phi) phi.*log(phi)/nb+(1-phi).*log(1-phi)./na+chi.*phi.*(1-phi);

%mua and  mub are the chemical potentials of A and B
mua=@(phi) log(1-phi)+phi*(1-na/nb)+chi*na*phi.^2;
mub=@(phi) log(phi)+(1-phi)*(1-nb/na)+chi*nb*(1-phi)^2;

% now write the function that is equal to zero when the A and B chemical
% potentials are equal to one another for phi=phi(1) and phi=phi(2)

ftosolve=@(phi) [mua(phi(1))-mua(phi(2));
    mub(phi(1))-mub(phi(2))];

plot(phi,fv(phi));
xlabel('\phi_{b}')
ylabel('G_{v}V_{0}/RT')

% start with a guess for the equilibrium volume fractions
phiguess = [0.2; 0.8];      % Make a starting guess at the equilibrium compositions
[phicalc,fval] = fsolve(ftosolve,phiguess);  % Call solver

% now we add the tangent line
slope=(fv(phicalc(2))-fv(phicalc(1)))/(phicalc(2)-phicalc(1));
intercept=fv(phicalc(1))-slope*phicalc(1);
tangentline=intercept+slope.*phi;
hold on
plot(phi,tangentline,'r')
hold off
title(['\chi=' num2str(chi) ', N_a=' num2str(na), ', N_b=' num2str(nb)], 'fontsize', 11)

% now save the plot as a jpg file
saveas(gcf,'commontangent.jpg');

% this saves the file a .eps file, used to embed the figure into the
% solution set - students can comment out this next line if they don't wan
% the .svg file
print(gcf,'-dsvg','commontangent.svg')
```

a Plot the free energy expression $\chi = 0.02$, $N_a = 200$, and $N_b = 100$, along with its common tangent construction, and report the values of the equilibrium compositions of the two phases. Include the plot with your solution.

b Repeat the previous calculation for $\chi = 0.025$ and $\chi = 0.05$ (you may need to adjust the initial guess for the volume fractions).

c Decrease χ to the critical value, χ_{crit}, which can be calculated analytically. Include a plot of G_v as a function of ϕ_b for value of χ.

(62) Extend work from the previous problem to obtain a MATLAB code that generates the full phase diagram (with χ on the vertical axis and ϕ_b on the horizontal axis) for a system with $N_a = 200$ and $N_b = 100$. Show both the bimodal and spinodal curves, and turn in a printout of your actual MATLAB code.

Hint: You'll want to solve the equations for a variety of χ values above the critical value of χ, which you can calculate analytically. You can use a 'for' loop in MATLAB to do this. Calculation of the spinodal curve is pretty easy, because there's an analytic expression for that. When you're developing your initial guesses for

the bimodal compositions to pass to the 'fsolve' command, keep in mind that the bimodal compositions must lie outside the spinodal compositions. This can help you develop sensible guesses to use for this quantity.

102 Appendix 1: MATLAB scripts

102.1 Symbolic Math

The following script uses MATLAB's symbolic math capability to evaluate the integrals in Eq.92.34.

```
% symbolic_visco.m
syms tau w go t real positive % specify symbolic variables and insist that they be real and
  positive
g=go*exp(-t/tau); % define exponential  relaxation function
storage=w*int(g*sin(w*t),t,0,inf); % evaluate the ingetral for G'
loss=w*int(g*cos(w*t),t,0,inf); % evaluate the integral for G"
pretty(storage) % visualize the expressions in a simpler form
pretty(loss)
```

The 'pretty' command returns the solutions in a format that is relatively easy to interpret, and is shown in Figure102.1. Note that the first expression (forG') needs to be manipulated a little bit to give the form shown in Eq.92.34.

```
                go
go - -----------
                2  2
           tau  w  + 1

  go tau w
-----------
   2  2
tau  w  + 1
```

Figure 102.1: Output generated by symbolicvisco.m.

102.2 Solution to a Nonlinear Equation

The following script returns$\phi_s = 0.2885$ as the solution to Eq.97.9 for$\chi = 1.05$.

```
chi=1.05;  % this is the value of chi we want to use
mus=@(phis)log(phis)+1-phis+chi*(1-phis)^2; % solvent chemical potential expression
phiguess=0.2;  % initial guess for phiguess
phisval=fsolve(mus,phiguess)  % returns the value of phis for which mus=0, with an initial
  guess for phis of phiguess
```

103 Appendix 2: Data Tables

103.1 Dynamic Moduli

103.1.1 Poly(2-vinyl pyridine)

130 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-0.8	4.48	4.25
-0.6	4.55	4.25
-0.4	4.6	4.3
-0.2	4.66	4.28
0	4.71	4.35
0.2	4.76	4.39
0.4	4.81	4.43
0.6	4.87	4.46
0.8	4.92	4.49
1	4.97	4.5
1.2	5.02	4.49
1.4	5.05	4.44
1.6	5.08	4.36
1.8	5.12	4.23
2	5.13	4.3

140 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-1	3.6	4.01
-0.8	3.88	4.13
-0.6	4.12	4.2
-0.4	4.28	4.24
-0.2	4.4	4.25
0	4.48	4.26
0.2	4.54	4.25
0.4	4.6	4.27
0.6	4.65	4.3
0.8	4.71	4.33
1	4.76	4.36
1.2	4.81	4.4
1.4	4.87	4.44
1.6	4.91	4.44
1.8	4.97	4.45
2	5.02	4.44

160 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-1	2.1	2.83
-0.8	2.21	3.03
-0.6	2.36	3.22
-0.4	2.58	3.42
-0.2	2.86	3.6
0	3.19	3.79
0.2	3.5	3.96
0.4	3.8	4.1
0.6	4.05	4.19
0.8	4.24	4.24
1	4.37	4.25
1.2	4.46	4.25
1.4	4.53	4.25
1.6	4.6	4.26
1.8	4.65	4.26
2	4.7	4.29

180 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-0.4	1.72	2.53
-0.2	1.79	2.73
0	1.9	2.92
0.2	2.06	3.12
0.4	2.31	3.32
0.6	2.59	3.51
0.8	2.92	3.7
1	3.26	3.88
1.2	3.59	4.05
1.4	3.9	4.19
1.6	4.16	4.27
1.8	4.35	4.31
2	4.47	4.3

200 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
0.4	1.71	2.63
0.6	1.81	2.82
0.8	1.97	3.02
1	2.18	3.21
1.2	2.45	3.41
1.4	2.77	3.6
1.6	3.11	3.79
1.8	3.46	3.97
2	3.79	4.11

103.1.2 Poly(t-butyl acrylate)

75 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-1	5.01	4.37
-0.83	5.03	4.39
-0.67	5.05	4.42
-0.5	5.07	4.46
-0.33	5.09	4.51
-0.17	5.11	4.58
0	5.13	4.66
0.17	5.16	4.74
0.33	5.19	4.83
0.5	5.23	4.92
0.67	5.26	5.02
0.83	5.31	5.11
1	5.36	5.21
1.17	5.41	5.31
1.33	5.47	5.41
1.5	5.53	5.51
1.67	5.6	5.61
1.83	5.68	5.72
2	5.76	5.82

85 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-1	4.88	4.39
-0.83	4.91	4.38
-0.67	4.94	4.37
-0.5	4.97	4.37
-0.33	4.99	4.37
-0.17	5.01	4.38
0	5.03	4.4
0.17	5.06	4.44
0.33	5.08	4.48
0.5	5.1	4.54
0.67	5.12	4.61
0.83	5.14	4.69
1	5.17	4.77
1.17	5.2	4.86
1.33	5.24	4.95
1.5	5.28	5.05
1.67	5.33	5.14
1.83	5.38	5.24
2	5.43	5.34

100 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-1	4.6	4.41
-0.83	4.66	4.42
-0.67	4.72	4.42
-0.5	4.78	4.41
-0.33	4.82	4.41
-0.17	4.86	4.4
0	4.9	4.39
0.17	4.93	4.38
0.33	4.95	4.38
0.5	4.98	4.38
0.67	5	4.39
0.83	5.03	4.4
1	5.05	4.43
1.17	5.06	4.47
1.33	5.09	4.52
1.5	5.11	4.59
1.67	5.13	4.66
1.83	5.16	4.74
2	5.19	4.83

120 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-1	3.79	4.15
-0.83	4.01	4.25
-0.67	4.20	4.32
-0.5	4.35	4.37
-0.33	4.46	4.40
-0.17	4.56	4.41
0	4.63	4.42
0.17	4.70	4.42
0.33	4.75	4.42
0.5	4.80	4.42
0.67	4.84	4.41
0.83	4.88	4.41
1	4.91	4.40
1.17	4.94	4.39
1.33	4.97	4.39
1.5	5.00	4.40
1.67	5.02	4.41
1.83	5.04	4.44
2	5.06	4.47

140 °C

$\log(\omega)$ (s^{-1})	$\log(G')$ (kPa)	$\log(G'')$ (kPa)
-1	2.72	3.57
-0.83	2.98	3.73
-0.67	3.24	3.88
-0.5	3.50	4.02
-0.33	3.75	4.14
-0.17	3.98	4.25
0	4.17	4.32
0.17	4.33	4.38
0.33	4.45	4.41
0.5	4.55	4.42
0.67	4.63	4.43
0.83	4.70	4.44
1	4.75	4.44
1.17	4.80	4.44
1.33	4.85	4.43
1.5	4.89	4.42
1.67	4.92	4.42
1.83	4.95	4.41
2	4.98	4.41

104 Other Resources

104.1 Polymer Synthesis

The Macrogalleria web site is an excellent online source describing the reaction mechanisms for most of the polymerization reactions that we have considered:

http://www.pslc.ws/macrog/level4.htm

This site includes reaction mechanisms for free radical, anionic anionic, cationic and step growth polymerizations.

Part VIII
351-1: Quantum Mechanics

105 Why Do We Need Quantum Mechanics?

105.1 Classical (Planetary) Picture of the Hydrogen Atom

Consider a classical analysis of the hydrogen atom, where the electron can be viewed as a charge $-e$ that is orbiting around a proton of charge e.
The Coulomb attraction force F_{coul} between the proton and the electron is given by the following formula:

$$F_{coul} = -\frac{1}{4\pi\varepsilon_o}\frac{e^2}{r^2} \tag{105.1}$$

Suppose that the electron is moving with a velocity of v in a circular orbit with a radius of r. The electrons are being accelerated toward the center of the orbit by v^2/r. In the classical picture the relationship between r and v is determined by Newton's second law relating the Coulomb force from Eq. 105.1 to the acceleration:

$$F = m_e a = m_e\frac{v^2}{r} = \frac{1}{4\pi\varepsilon_o}\frac{e^2}{r^2} \tag{105.2}$$

Solving this equation for r results in the following expression for the radius:

$$r = \frac{e^2}{4\pi\varepsilon_o m_e}\frac{1}{v^2} \tag{105.3}$$

Classically, v is continuous suggesting that r can be anything. This implies that the Hydrogen atom does not have a fixed size, which doesn't really make sense.

105.2 The Bohr Atom

How can this issue of an indeterminate size of the hydrogen atom be fixed? Suppose certain orbits are preferred, and that only orbits meeting the following condition are allowed (we'll justify this assumption later):

$$m_e v r = h/2\pi \tag{105.4}$$

Here m_e is the electron mass and h is Planck's constant . If we use Eq. 105.4 to eliminate v from Eq. 105.3 and then solve for r we obtain:

$$r = \frac{4\pi\varepsilon_o}{m_e e^2}\hbar^2 \equiv a_o \tag{105.5}$$

Where a_o is the Bohr radius. If we generalize so that the allowed radii are $r_n = n^2 a_o$, where $n = 1, 2, 3...$ as shown in Figure we can account for key features of hydrogen spectra.

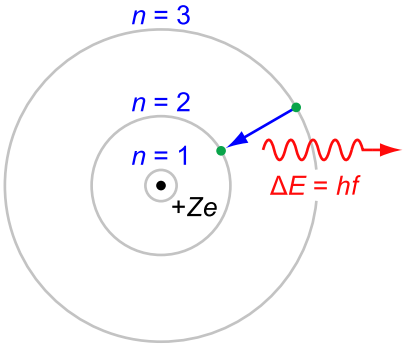


Figure 105.1: The Bohr atom [2].

105.3 Units and Magnitude of h

$\hbar = h/2\pi \; h = 2\pi\hbar$
Let $a_o \simeq 0.5\mathring{A} = 0.5 \times 10^{-10} m$.

$$\hbar^2 = \frac{h^2}{(2\pi)^2} = \frac{m_e e^2}{4\pi\varepsilon_0} r_0$$

$$h^2 = \frac{\pi \left(9.109 \times 10^{-31} kg\right)}{8.85 \times 10^{-12} \frac{C^2}{Jm}} \left(1.6 \times 10^{-19} C\right)^2 \left(0.5 \times 10^{-10} m\right) \tag{105.6}$$

Checking units:

$$\frac{\text{Jm}}{\text{C}^2} \left(\text{kg}\right) \left(\text{C}^2\right) \left(\text{m}\right) = \text{Jm}^2\text{kg} = \text{J} \left(\frac{\text{kgm}^2}{\text{s}^2}\right) \text{s}^2 \tag{105.7}$$

$$h = 6.43 \times 10^{-34} \text{J} \cdot \text{s}$$

The Bohr Radius is actually 0.529Å.

- Note that Rutherford established that the diameter of the electron cloud as: $d_{e^- cloud} \approx 10^{-10} m$.
- h had been showing up in other contexts, as we'll see.

From classical electrodynamics, orbits should **not** be stable because accelerating point charges radiate energy. This is described by the Larmor Formula:

$$P_{RAD} = \frac{dE_{RAD}}{dt} = \frac{e^2 a^2}{6\pi\varepsilon_0 c^3} \tag{105.8}$$

Where $E_{RAD}$is the radiated energy, P_{RAD} is the radiated power, c is the speed of light and a is the acceleration. Note:

$$P_{RAD} = \frac{dE_{RAD}}{dt} = \frac{e^2 a^2}{6\pi\varepsilon_o c^3} > 0$$

For the Planetary Atom,

$$E_{ATOM} = KE + PE = \frac{1}{2} m_e v^2 - \frac{e^2}{4\pi\varepsilon_0 r} \tag{105.9}$$

But, recall that

$$v^2 = \frac{e^2}{4\pi\varepsilon_0 m_e r} \tag{105.10}$$

$$E_{ATOM} = \frac{e^2}{8\pi\varepsilon_0 r} - \frac{e^2}{4\pi\varepsilon_0 r} = -\frac{e^2}{8\pi\varepsilon_0 r} < 0 \tag{105.11}$$

Conservation of Energy:

$$E_{ATOM} + E_{RAD} = Constant$$

$$\frac{dE_{ATOM}}{dt} + \frac{dE_{RAD}}{dt} = 0 \Rightarrow \frac{dE_{ATOM}}{dt} = -\frac{dE_{RAD}}{dt} \tag{105.12}$$

Since

$$\frac{dE_{RAD}}{dt} > 0 \Rightarrow \frac{dE_{ATOM}}{dt} < 0$$

E_{ATOM} becomes more negative with time and r decreases as the electron radiates.
As the orbiting electron radiates, it spirals into the nucleus.
But, we know that atoms are stable and do not radiate.
The Planetary Model fails catastrophically! (See homework #1).
We **need** quantum mechanics to understand atomic (and solid state) physics!

106 Origins of Quantum Theory

Prior to 1900, most experiments were well-explained or consistent with the beliefs that:

(1) Matter consists of particles obeying Newton's Laws.

(2) Radiation consists of waves obeying Maxwell's Laws.

These laws were incredibly successful at explaining numerous phenomena, however new concepts were needed to explain new (and old) experiments:

(1) Relativistic Mechanics (Einstein)

(2) Quantization of Observed Physical Quantities (e.g. E)

(3) Particle Properties of Radiation

(4) Wave Properties of Matter

Quantization is a consequence of attempts to reconcile the particle and wave picture

106.1 Black Body Radiation

A Blackbody is a perfect emitter/ absorber of radiation.
In 1879, J. Stefan found the **empirical** relationship between total emissive power R and T:

$$R(T) = E\sigma T^4 \tag{106.1}$$

Stefan-Boltzmann Law (Boltzmann derived this in 1884.)
E: emissivity = 1 for a blackbody
σ: $5.67 \times 10^{-8} \frac{W}{m^2 k^4}$ Stefan's Constant
Spectral Density:
$R(T) = \int_o^\infty d\lambda R(\lambda, T)$

Wien's Displacement Law: Max of $R(\lambda, T)$ occurs when $\lambda_{\max} T = b$, where $b = 2.898 \times 10^{-3} \mathrm{m} \cdot \mathrm{K}$.
(This is useful for temperature measurement, but we also want to know distance.)
Inside the cavity, energy is distributed according to spectral distribution function $\rho(\lambda, T)$, where $R(\lambda, T) = \frac{c}{4}\rho(\lambda, T)$.
From (Classical) Thermodynamics, Wein derived:

$$\rho(\lambda, T) = \lambda^{-5} F(\lambda, T)$$

$F(\lambda, T)$ cannot be derived from Thermo, but Wein proposed $F(\lambda, T) \sim e^{-\propto/\lambda T}$ (okay for short λ).
Lord Rayleigh and J. Jeans used electromagnetism to show that:

$$\rho(\lambda, T) = \frac{8\pi}{\lambda^4} < ESpectralDensity :>_\lambda$$

$\langle E \rangle_\lambda$ is the average energy of mode with wavelength λ (normal for long λ).
(How can we calculate this?)
From Classical Stat Mech, assuming $P_{\sigma e} \sim e^{-E/kT}$:
$< E >= \frac{\int_0^\infty dE E \exp(-\beta E)}{\int_0^\infty dE \exp(-\beta E)}$
$(\beta = \frac{1}{kT}$

$$k = 1.38 \times 10^{-23} \backslash mathrm$$

Boltzmann's Constant)
$< E >= -\frac{d}{d\beta} \ln\left[\int_0^\infty dE \exp(-\beta E)\right] \frac{d}{dx} f(x) = \frac{f'(x)}{f(x)}$
$< E >= -\frac{d}{d\beta} \ln\left(\frac{1}{\beta}\right) = -\beta\left(-\frac{1}{\beta^2}\right) = \frac{1}{\beta} = kT$
$\therefore \rho(\lambda, T) = \frac{8\pi}{\lambda^4} kT \Rightarrow f(\lambda, T) = 8\pi k \lambda T$
At long λ, this approaches experimental values.
But $\rho(\lambda, T) \to \infty$ as $\lambda \to o$
"Ultraviolet Catastrophe"
1900 - Planck's Quantum Theory
Revolutionary Hypothesis: Radiation emitted by microscopic oscillators of frequency V that can only take on discrete values

$$nE_o, n = 0, 1, 2, ...$$

(How does this affect the distribution function?)

$$< E >= \frac{\sum_{n=0}^{\infty} nE_0 \exp(-\beta n E_o)}{\sum_{n=0} \exp(-\beta n E_o)}$$

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$$

$$< E >= -\frac{d}{d\beta}\left[\ln \sum_{n=0}^{\infty} \exp(-\beta n E_o)\right]$$

$$< E >= -\frac{d}{d\beta} \ln\left(\frac{1}{1 - \exp(-\beta E_o)}\right)$$

$$\frac{d}{dx} \ln f(x) = \frac{f'(x)}{f(x)}$$

$$< E >= \frac{1 - \exp(-\beta E_o)(-1)}{[1 - \exp(-\beta E_o)]^2} [-E_o \exp(-\beta E_o)]$$

$$< E >= \frac{E_o}{\exp(\beta E_o) - 1}$$

Therefore,

$$\rho(\lambda, T) = \frac{8\pi}{\lambda^4} \frac{E_o}{\exp(E_o/kT - 1)}$$

For Planck's Formula to satisfy Wien's Law, where $\rho(\lambda, T) = \lambda^{-5} f(\lambda, T)$:
$E_o \propto \frac{1}{\lambda}$
$E_o = hV = \frac{hc}{\lambda}$, where h is Planck's Constant.
$h = 6.626 \times 10^{-34} J \cdot s$
So $\rho(\lambda, T) = \frac{8\pi hc}{\lambda^5} \frac{1}{exp(\frac{hc}{\lambda kT}) - 1}$ (consider extremes).
This was a successful interpolation of (1) Wien's Exponential Law and (2) the Rayleigh-Jeans Law.

(1) For short λ, $\rho(\lambda, T) \propto \exp\left(-\frac{hc}{\lambda kT}\right)$ (goes to zero for short λ)

(2) For long λ, $\exp\left(\frac{hc}{\lambda kT}\right) \simeq 1 + \frac{hc}{\lambda kT}$ (Taylor series)
$\lambda \gg hc/kT$
$\rho(\lambda, T) \simeq \frac{8\pi hc}{\lambda^5} \frac{1}{1 + \frac{hc}{\lambda kT} - 1} - \frac{8\pi}{\lambda^4} kT$

(3) Can calculate h using $\lambda_{MAX} T = Constant$

Photoelectric Effect (evidence for quantum of radiation) Electrons emitted from metal surface when exposed to radiation (Hertz, Hallwachs - 1888, J.J. Thomson - 1899).

- $E_{\max} \alpha \nu$, not intensity (Leonard, 1902)
- $\frac{dN_e}{dt} \propto$ intensity (Leonard, 1902)
- No current for $\nu < \nu_{TH}$ (Quant confirmed by Milikan)

1905 - Einstein explained that $E_{ph} = h\nu$ (theoretically).

- Radiation consists of Quanta (photons) of energy $h\nu$.
- Energy is transferred in discrete amounts.
- Einstein received the Nobel Prize for this work.

Electrons are bound to the surface with energy ϕ, the work function.
By conservation of energy, $E_{\max} = h\nu - \phi$.
$h\nu_{TH} = \phi$ (was a prediction)
$E_{\max}$ has no dependence on intensity.
(HW Aside)
Work Function - work done on electron moving through potential
Suppose $F_x = -eE_x = +e\frac{dV}{dx}$
$\int F_x dx = e \int \frac{dV}{dx} dx$
Work = $eV = 1.6 \times 10^{-19} c \cdot J/c = 1.6 \times 10^{-19} J$ (an energy)
eV is the potential energy of electon in potential V.

Compton Scattering (He mentions Quantum Theory.) (1923)
Do photons behave as particles kinematically?
(Photoemission is intermediate case - no outgoing photon.)

- The electron carries momentum from incident photon.
- New photon energy is smaller.
- Photon and electron are treated as relativistic particles:

  $E = \gamma mc^2$, $\rho = \gamma mv$, where $\gamma \equiv 1/\sqrt{1 - \frac{v^2}{c^2}}$

Results:

- $\triangle\lambda = \lambda_1 - \lambda_o = 2\lambda_c$
- $\sin^2\left(\frac{\theta}{2}\right) = \lambda c\,(1 - \cos\theta)$, where $\lambda_c = \frac{h}{m_e c}$
- Compton Wavelength, confirmed by experiments⇒radiation behaves as a wave and a particle.

Matter Waves 1923 - De Broglie proposed that if matter behaved as waves, then $\lambda = h/p$.
For light:
$E = hV = \frac{hc}{\lambda}$ and $E = pc$
$\Rightarrow \lambda = h/p$

- not usually observed becauseρ is large/ h is small
- need something very light, not too fast

1927 - Davisson and Germer confirm with electrons.
Consider acceleration of e^-through potential V.
$E = \frac{1}{2}mv^2 = \frac{m^2v^2}{2m} = \frac{\rho^2}{2m}$
$\rho = \sqrt{2mE}$ use $E = cV$
$\lambda = \frac{h}{\rho} = \frac{h}{\sqrt{2mE}} = \frac{h}{\sqrt{2meV}}$
$\lambda \simeq \frac{12.3\text{Å}}{\sqrt{V(volts)}}$ e.g. V = 150 volts gives $\lambda \sim 1$Å
Constructive Interference of Electrons occurs when:
$d\sin\theta = n\lambda$, where $n = 1, 2, 3...$
(saw clear path after annealing oxidized Ni crystal after vacuum failure)
The same h shows up to explain particles.
Helium atom diffraction: surface science
Neutron: diffraction in solids
Mass is ∼1000x bigger. (How much smaller is λ?)
→Need low energy particles.

Matter behaves like waves and particles.

Classical physics does not describe:

(1) Quantization

(2) Wave - Particle Duality

1925 - Heisenberg, Born and Jordan develop Matrix Mechanics
1925 - Schrödinger develops Wave Mechanics (Matter Waves)
1930 - Dirac Formulation: More General + Relativity →Spin, Magnetism

107 Origins of Schrödinger Equation

Consider the Two-Split Experiment for Electrons. (What do you expect?) Each electron is detected when it hits screen in fixed location.

$\rightarrow$localized particle

What about the statistical behavior?

Over time, a pattern builds up.

"Interference" Pattern builds up even if only 1 electron is in transit at any given time.

Interpretation:

- Electrons pass through both slits.
- Position is not determined until it reaches the detector.

Proof: $I \neq I_1 + I_2$

(We need to give up the possibility of knowing $x\,(t)$at all times and adopt a probabilistic description of particle behavior.)

Introduce the wave function $\Psi\,(x,y,z,t)$ as a probability amplitude, i.e...

$$P\,(x,y,z,t) \propto |\Psi\,(x,y,z,t)|^2$$

P is the probability of finding the particle at x, y, z within volume at time t.

If slit #1 is open, $P_1 \propto |\Psi_1|^2$.

If slit #2 is open, $P_2 \propto |\Psi_2|^2$.

If both are open, $\Psi = \Psi_1 + \Psi_2$.

$\rightarrow P \propto |\Psi_1 + \Psi_2|^2 \neq |\Psi_1|^2 + |\Psi_2|^2$ (This says little about form of Ψ.)

(Observation of large numbers of particles provides a sense of what a single particle is likely to do.)

1926 - Born formulated the Wave Function.

Assume a large number of identical, independent systems that are identically prepared and subjected to the same external force.

$$P\,(\vec{r},t)\,d\vec{r} = |\Psi\,(\vec{r},t)|^2\,d\vec{r} = \Psi^*\,(\vec{r},t)\,\Psi\,(\vec{r},t)\;\,d\vec{r}$$

This is an inherently statistical description of matter. (It's probabilistic, but statistically deterministic.)

The probability of finding the particle somewhere is unity.

$$\therefore \int_{ALL\,SPACE} d\vec{r}\;|\Psi\,(\vec{r},t)|^2 = 1$$

$\Psi\,(\vec{r},t)$ must be normalizable, cannot diverge.

$\Rightarrow$provides constraints for solving for Ψ

Ψ obeys Principle of Superposition. Assume: $\Psi = c_1\Psi_1 + c_2\Psi_2 c_1$,$c_2$ complex constants

$\Psi_1 = |\Psi_1|\,e^{i\propto 1}\Psi_2 = |\Psi_2|\,e^{i\propto 2}$

$e^{ix} = \cos x + i\sin x$

$\left|\Psi^2\right| = \Psi^*\Psi = \left(c_1^*\Psi_1^* + c_2^*\Psi_2^*\right)\left(c_1\Psi_1 + c_2\Psi_2\right)$

$\left|\Psi^2\right| = |c_1\Psi_1|^2 + |c_2\Psi_2|^2 + c_1^*c_2\Psi_1^*\Psi_2 + c_2^*c_1\Psi_2^*\Psi_1$

$\left|\Psi^2\right| = |c_1\Psi_1|^2 + |c_2\Psi_2|^2 + 2Re\left\{c_1c_2^*\,|\Psi_1|\,|\Psi_2|\,exp\left[i\left(\propto_1 - \propto_2\right)\right]\right\}$

$2Re\left\{c_1c_2^*\,|\Psi_1|\,|\Psi_2|\,exp\left[i\left(\propto_1 - \propto_2\right)\right]\right\}$- Interference term $\neq 0$ in general.

Probability amplitudes **interfere**.

Note: Any complex number can be written: $A^* = |A|\,c^{i\phi}$.

$A^* + A = \left|A\right|e^{-i\theta} + \left|A\right|e^{io}$

$A^* + A = \left|A\right|\left(cos\theta - isin\theta\right) + \left|A\right|\left(cos\theta + isin\theta\right)$

$A^* + A = 2\left|A\right|cos\left(\theta\right) = 2Re\left\{A\right\}$

Build Ψ with Physical Characteristics Following De Broglie, assume particle has definite momentum $p = \frac{h}{\lambda}$, $E = hv$.
And define:

(1) Angular Frequency $\omega = 2\pi v$

(2) Wave Number $R = 2\pi/\lambda$

(3) Reduced Planck Constant $\hbar = h/2\pi$

Then, $E = \hbar w$ and $p = \hbar\kappa$.
For photons, we know $\omega = 2\pi v = \frac{2\pi}{\lambda} \cdot c = \kappa c$.
Note that: (1) $\omega \propto \kappa$ and (2) ω/κ is a speed.
For free particles, we will see that $\omega \propto \kappa^2$.
- What is the speed of a free particle wave?

- $F(x - vt)$ represents a waveform traveling along x with speed v.
- For a free particle of mass m in moving along $+x$ axis,

$$\vec{P} = P_x \hat{x}, P_x > 0, \kappa = P_x/\hbar$$

Represent particle with definite momentum as **a Plane Wave**.
Since $\kappa = P/\hbar$, $\omega = E/\hbar$, $\Psi(x,t) = Aexp\left[\frac{i}{\hbar}\left(P_x X - E(P_x)t\right)\right]$.
Note:

(1) $-i\hbar\frac{d}{dx}\Psi = Px\Psi$

(2) $i\hbar\frac{d}{dt}\Psi = E\Psi$

And in 3-D,
$\Psi(\vec{r},t) = Aexp[i(\vec{\kappa}\cdot\vec{r} - \omega(\kappa)t)]$
$\Psi(\vec{r},t) = Aexp\left[\ddot{\psi}\hbar(\vec{p}\cdot\vec{r} - E(p)t)\right]$
We have:

(1) $-i\hbar\nabla\Psi = \vec{p}\Psi$; $-i\hbar$ = pop ; ∇: gradient

(2) $i\hbar\frac{\delta}{\delta t}\Psi = E\Psi$; $i\hbar\frac{d}{dt} = E_{op}$

Postulate: The differential operators give $\vec{p}$ and $\vec{E}$ even in the presence of a potential. (Particles may not be free.)
Classically,$E_{tot} = \frac{p^2}{2m} + V(\vec{r},t)$.
1926: Schrödinger proposes that $E_{op}\Psi(\vec{r},t) = \left[\frac{p_{op}^2}{2m} + V(\vec{r},t)\right]\Psi(/\vec{r},t)$.
From the above postulate, it follows that $i\hbar\frac{d}{dt}\Psi(\vec{r},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r},t)\right]\Psi(\vec{r},t)$.
This is the time dependent **Schrödinger Equation**.
It:

(1) Describes non-relativistic quantum mechanics.

(2) Is not derived, but is physically plausible (can be related to Hamiltonian classical mechanics).

(3) Correctly predicts experiments to date.

Define:
Hamiltonian Operator: $H = -\frac{\hbar^2}{2m}\nabla^2 + V$
The Schrödinger Equation can then be written:

$$i\hbar\frac{d}{dt}\Psi(\vec{r},t) = H\Psi(\vec{r},t)$$

In MSE 351-1, we will solve this equation for:

(1) Free Particles ($V = 0$)

(2) Potential Wells and Barriers in 3-D (tunneling): $E < V$ and $E > V$

(3) Coulomb Potential in 3-D (periodic table) ($V \sim -\frac{1}{r}$)

(4) Periodic Potentials: $V \sim sinx$ (as in solids $\rightarrow$electronic "band" structure)

Tasks for Students

(1) Learn to solve Schrödinger equation for these few cases.

(2) Calculate probability distributions, energies and dispersion relationships.

(3) Provide physical interpretations.

(from S+W 3.3)
We will consider only time-independent potentials: $V(t)$.
(Time-dependent potentials require Perturbation Theory. This is needed for transitions and optical properties in particular.)

Separation of Variables - attempted solution $\Psi(\vec{r},t) = \psi(\vec{r})\,\omega(t) = \psi\omega$

$-\frac{\hbar^2}{2m}\omega\nabla^2\psi + V\psi\omega = i\hbar\psi\frac{d\omega}{dt}$ (divide by $\psi\omega$)

(1) $-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi$ (Time-dependent Schrödinger Equation)

(2) $i\hbar\frac{dw}{dt} = E\omega \Rightarrow \omega = exp\left(-i\frac{E}{\hbar}t\right)$

A solution for time independent potentials is therefore: $\Psi(\vec{r},t) = \psi(\vec{r})e^{-i\frac{E}{\hbar t}}$, where $\psi(\vec{r})$ is the solution of (1).
Apply to free particle in 1-D:

$\nabla^2 \rightarrow \frac{d^2}{dx^2}$, $V(x) = 0$

$\frac{-\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi(x)\psi'' = \frac{-2mE}{\hbar^2}\psi$, $\kappa = \frac{\sqrt{2mE}}{\hbar^2}$

$\rightarrow \psi(x) = Ae^{i\kappa x} + Be^{-i\kappa x}A$, B constants: $\kappa = \frac{\sqrt{2mE}}{\hbar}$ (see below)

$\Psi(x,t) = Ae^{i\left(\kappa x + \frac{Et}{\hbar}\right)} + Be^{-i\left(\kappa x + \frac{Et}{\hbar}\right)}$

right-moving: $Ae^{i\left(\kappa x + \frac{Et}{\hbar}\right)}$

left-moving: $Be^{-i\left(\kappa x + \frac{Et}{\hbar}\right)}$

These are plane waves (De Broglie Matter Waves):

$$E = \frac{p^2}{2m} = \frac{(h/\lambda)^2}{2m} = \frac{\hbar^2\kappa^2}{2m} = \hbar\omega\left(\omega \propto \kappa^2\right)$$

Consider the right-moving wave discussed earlier:

$$\Psi_\kappa(x,t) = e^{i(\kappa x - wt)}\kappa = \frac{\sqrt{2mE}}{\hbar}w = \frac{\hbar}{2m}\kappa^2$$

If Ψ_κ represents 1 electron, then: $|\Psi_\kappa(x,t)|^2 = 1$ $\rightarrow$constant for **all** x, t.
Interpretation: A particle with exact momentum is completely delocalized in space.
Normalization Condition: $\int_{-\infty}^{\infty} dx\,|\Psi_\kappa(x,t)|^2 = \int_{-\infty}^{\infty} dx = \infty - (-\infty)$?!
$\Rightarrow \Psi_\kappa(x,t)$ cannot represent wave function of a single particle.
Particles cannot have "perfectly defined" or exact momenta.
Approach: Represent particle as a superposition of plane waves that produces a localized, normalizable wave packet.

Wave Packets in 1-D (S+W 2.3, in part)

$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int d\kappa \phi(\kappa) \Psi_\kappa(x,t)$(Fourier Transform)

$\Psi(x,t)$ - no index

Normalizable if $\phi(\kappa) \to 0$ as $\kappa \to \pm\infty$.

Find $\phi(R)$:$\Psi(x,o) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\kappa \phi(\kappa) e^{i\kappa x} = Ae^{i\kappa_o x}$for $\frac{-d}{2} \le x \le \frac{d}{2}$

Approach: Multiply both sides by $e^{-i\kappa x'}$ and integrate.

Substitute $\kappa = \kappa^1$:

$$\phi(\kappa) = \sqrt{\frac{2}{\pi}} A \frac{sin\left[(\kappa_o - \kappa)\frac{\ell}{2}\right]}{\kappa_o - \kappa}$$

Interpretation:

Particle is locatlized and moves with speed $\backsim V = \frac{p_o}{m} = \frac{\hbar\kappa_o}{m}$ (we'll prove).

But there is a spread in momenta about κ_o:

For $\kappa \sim \kappa_o$, new peak, $(\kappa_o - \kappa)\frac{\ell}{2} \approx \pi$ This defines width of $\phi(\kappa) \to \triangle\kappa = \frac{2\pi}{\ell}$ (argument runs over $\sim \pi$).

$\triangle p = \hbar\triangle\kappa \approx \hbar\frac{2\pi}{\ell} = \frac{h}{\ell}$

$\triangle p \triangle x \simeq \frac{h}{\ell} \cdot \ell \simeq h$

$\triangle p \triangle x \sim h$ is the Heisenberg Uncertainty Principle that can be derived using matrix mechanics.

Comments:

(1) Here, the relationship between uncertainties follows directly from Fourier Transforms.

(2) Analagous relationships exist for $\omega, t(E,t)$.

Velocity of the Wavepacket Assume $\phi(\kappa)$ is peaked about κ_o, consider κ values $\sim\kappa_o$.

$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\kappa \phi(\kappa) e^{i(\kappa x - \omega t)}$, but $\omega(\kappa) \simeq W_o + \left.\frac{d\omega}{d\kappa}\right|_{\kappa_o} (\kappa - \kappa_o)$(Taylor Series Expansion)

$\Psi(x,t) \simeq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\kappa \phi(\kappa) exp\left[i\left(\kappa x - \left(\omega_o + \left.\frac{d\omega}{d\kappa}\right|_{\kappa_o}(\kappa - \kappa_o) + ...\right)t\right)\right]$(Bring terms in front that do not depend on κ.)

$\omega = \frac{\hbar\kappa^2}{2m}$, $\left.\frac{d\omega}{d\kappa}\right|_{\kappa_o} = \frac{\hbar\kappa_o}{m} \to V_g = \frac{p_o}{m}$

The center of the wavepacket travels at the group velocity $V_g = \left.\frac{d\omega}{d\kappa}\right|_{\kappa_o}$.

Individual components travel at phase velocity $\frac{\omega}{\kappa}$, which may cause wavepacket to spread out with time.

Proof: $\Psi_\kappa(x,t) = e^{i(\kappa x - \omega t)} = e^{i\phi}$

ϕ is a phase, and the equation defines a plane of constant phased (e.g. peak or trough).

$\phi = \kappa x - \omega t = constant$

$x = \frac{\phi}{\kappa} + \frac{\omega}{\kappa} t$

$\left.\frac{dx}{dt}\right|_{\phi=constant} = +\frac{\omega}{\kappa} = V_p$

Re-examine the Two-Split Experiment:

- Attempt to measure which slit the electron passes through.
- Use a microscope that can resolve the slits $\to$ implies that $\lambda_\gamma < \delta$.

(1) $\triangle P_e \simeq P_\gamma = \frac{h}{\lambda\gamma} > \frac{h}{d}$ $(\lambda_\gamma \le d)$

Note that $\triangle P_e d > h$.

(2) $P_e = \frac{h}{\lambda_e}$

$\frac{\triangle P_e}{P_e} = \frac{h}{d}\frac{\lambda_e}{h} = \frac{\lambda_e}{d} = \frac{\triangle x}{L}$ (from above)

$\Rightarrow \triangle x \simeq L\frac{\lambda_e}{d}$

The uncertainty in position is comparable to the distance between interference maxima.

$\Rightarrow$ No pattern is observed.

Why not? The act of measurement has "collapsed the wave function."

108 Time-Independent Potentials: Wells and Barriers

Introduction Plane Waves: Solutions to Schrödinger Equation for Free Particle ($V = 0$)

But,

This formulation usefully explains the Two-Slit Experiment.

To build up to hydrogen atom (sow of Schrödinger Equation for e-bound to proton), we start with a simpler case.

Recall that for time-independent potentials, in 1-D, solutions have the form:

$$\Psi(x,t) = \psi(x)\, e^{-i\frac{E}{\hbar}t}, \text{ where } \left[\frac{-\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi(x) = E\psi(x)$$

$$\left[\frac{-\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi(x) = E\psi(x)$$

$$\int_{-\infty}^{\infty} \psi^*(x)\psi(x)dx = 1$$

Note that $\left|\Psi(x,t)\right|^2 = \left|\psi(x)\right|^2$ are independent of time.

We call these stationary states of definite energy.

Because superposition holds, the general solution is: $\Psi(x,t) = \sum_{n=0}^{\infty} C_n \psi_n(x)\, e^{-i\frac{E_n}{\hbar}t}$.

C_n: complex constant

E_n: definite energy

$\left|\Psi(x,t)\right|^2$ is time dependent: does not have definite E.

1-D Time-Independent Potentials: Wells & Barriers Recall time-independent Schrödinger Equation in 1-D:

$$H\psi = E\psi, \text{ where } H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)$$

Solutions of Schrödinger Equation with constant potential have the form:

$$\Psi(x,t) = \psi(x)\, e^{-1\frac{E_n}{\hbar}t}, \text{ where } \left[\frac{-\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi(x) = E\psi(x)$$

Note that $\left|\Psi(x,t)\right|^2 = \left|\psi(x)\right|^2$ is time-dependent.

Such **stationary states** have definite energy (the region of potential sampled by electron is constant).

The Schrödinger Equation is a linear equation, so Principle of Superposition holds.

Let $\Psi_n(x,t) = \psi_n(x)\, e^{-i\frac{E_n}{\hbar}t}$ be the infinite.

Set of Solutions Since superposition holds, the general solution is $\Psi(x,t) = \sum_{n=0}^{\infty} C_n f_n(x)\, e^{-i\frac{E_n}{\hbar}t}$.

C_n : complex constrants

E_n: definite energies

Note: $\left|\Psi(x,t)\right|^2$ is time-dependent and does not have a definite energy.

Infinite Square Well Potential Solve Schrödinger Equation for Stationary States:

Outside Box: $V = \infty \to \psi(x) = 0$

Boundary:

(1) $\psi(0) = 0$

(2) $\psi(L) = 0$

Inside Box: $E\psi(x) = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}$ $(V = 0)$

Soln: $\psi(x) = Asin\kappa x + Bcos\kappa x, \kappa = \frac{\sqrt{2mE}}{\hbar}$

Boundary conditions provide restrictions:

(1) $\psi(0) = 0 \Rightarrow B = 0 \Rightarrow \psi(x) = Asin\kappa x$

(2) $\psi(L) = Asin\kappa l = 0 \Rightarrow \kappa = \frac{\pi}{L} \cdot n, n = 1, 2, 3...$ (discrete values)

Now, E, ω are related to κ:
$\kappa = \frac{\sqrt{2mE}}{\hbar}$, $E = \frac{\hbar^2\kappa^2}{2m}$(Particle is "free" inside box.)
$E = \frac{\hbar^2\pi^2}{2m}\frac{n^2}{L^2} n = 1, 2, 3...$
Energy levels are quantized.
Observations:

(1) $\psi_1(x)$ is symmetric; $\psi_2(x)$ is antisymmetric.

(2) E is quantized.

(3) $\kappa \propto \frac{1}{L}$ $\left(E \propto \frac{1}{L^2}\right)$ is consistent with uncertainty principle. (The increase in ground state energy is new).

Determine A by normalization:
(Recall that this is physics, not just math).

$$\int_{-\infty}^{\infty} dx\, \psi^*(x)\, \psi(x) = |A|^2 \int_o^L dx\, sin^{-2}\left(\frac{\pi n}{L}x\right) = 1$$

Change variables: $y = \frac{\pi n}{L}x$, $dy = \frac{\pi n}{L}dx$
$= |A|^2 \frac{L}{2} = 1 \Rightarrow A = \sqrt{\frac{2}{L}}$
$\psi_n(x) = \sqrt{\frac{2}{L}} sin\frac{\pi n}{L}x E_n = \frac{\pi^2\hbar^2}{2m}\frac{n^2}{L^2}$
Properties of $\psi_n(x)$:

(1) Orthogonal: $\int dx \psi_m^*(x)\, \psi_n(x) = 0$ for $m \neq n$

(2) Complete: Any solution $F(x)$ that satisfies the boundary conditions can be written as:
$F(x) = \sum_n C_n \psi_n(x)$

If particle is in particular state ψ_m, $|\Psi(x,1)|^2 = (\psi_m)^2$ is independent of time.
The Ψ_m are stationary states.
If particle is in superposition $|\Psi(x,t)|^2$ is time dependent $\rightarrow$ does not have definite E.
If we know $F(x)$,we can calculate C_n:
If we know $\Psi(x,0)$ at $t = 0$, we can now calculate $\Psi(x,t)$ (evolution of system).
General Solution: $\Psi(x,t) = \sum_n C_n \psi_n(x)\, e^{i\frac{E}{\hbar}t}$
If we know initial state of system $\Psi(x,0)$, then $\Psi(x,0) = \sum_n C_n \psi_n(x)$, $C_n = \int dx \psi_n^*(x)\, \Psi(x,0)$.
Put $C_n$back into general solution.

Finite Square Well Potential Look for bound, stationary states.
(example:$E < V_o$ "trapped")
Apply boundary conditions:

(1) $\psi(\pm\infty) = 0$ $\psi(x)$ must be normalizable

(2) $\psi(x)$ continuous at $x = \pm\frac{L}{2}$

(3) $\frac{d\psi(x)}{dx}$ continuous at $x = \pm\frac{L}{2}$ (finite K.E.); This is guaranteed by Schrödinger Equation for these boundary conditions.

Outside of the Well

$$\left(x < -\frac{L}{2}, x > \frac{L}{2}\right) \quad -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = (E - V_1)\,\psi E < V_1 \frac{\partial^2\psi}{\partial\lambda^2} = \frac{2m(V_1-E)}{\hbar^2}\psi$$

$$\text{Left: } \psi^{\ell}(x) = Ae^{-\gamma\lambda} + Be^{\gamma x} = Be^{\gamma\lambda}\gamma = \frac{2m(V_1-E)}{\hbar}$$

$$\text{Right: } \psi^{r}(x) = Fe^{-\gamma\lambda} + Ge^{\gamma\lambda} = Fe^{-\gamma\lambda}$$

(1) $\psi(\pm\infty) = 0 \Rightarrow A = 0, G = 0$

Inside the Well:

$\frac{-\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi$

$\psi(x) = Csin\kappa x + Dcos\kappa x\kappa = \sqrt{\frac{2mE}{\hbar}}$

(There will be symmetric and antisymmetric states.)

Sketch Solutions:

There are only a finite number of bound states (Soln with $E < V_1$).

Consider symmetric states $\psi(x) \sim cos(\kappa_n x)$.

Apply Boundary Conditions (at $\frac{L}{2}$).

(2) Continuity: $Dcos\left(\frac{\kappa L}{2}\right) = Fe^{-\gamma\frac{L}{2}}$

(3) Continuity of Derivative: $\frac{d\psi}{dx}$

$-\kappa Dsin\left(\frac{\kappa L}{2}\right) = -\gamma Fe^{-\gamma\frac{L}{R}}$

Divide (3) by (2) to get: $ktan\left(\frac{kL}{2}\right) = \gamma$, where $k = \sqrt{\frac{2mE}{\hbar}}\gamma = \sqrt{\frac{2m(V_1-E)}{\hbar}}$.

$\kappa tan\left(\frac{\kappa L}{2}\right) = \gamma$

$\sqrt{\frac{2mE}{\hbar}}tan\left(\frac{L}{2}\sqrt{\frac{2mE}{\hbar}}\right) = \sqrt{\frac{2m(V_1-E)}{\hbar}}$

$\sqrt{E}tan\left[\left(\frac{2m}{\hbar^2}E\frac{L^2}{\psi}\right)^{\frac{1}{2}}\right] = \sqrt{V_1 - E}$

We can solve this graphically:

- There must be one bound state given $E < V_1$.
- The same approach can be applied to antisymmetric states.

In the limit of a deep well, $V_1 \to \infty$.

$$tan\left[\left(\frac{2m}{\hbar^2}t^2\frac{L^2}{\psi}\right)^{\frac{1}{2}}\right] \to \infty$$

$$\left(\frac{2m}{\hbar^2}E^2\frac{L^2}{2}\right)^{\frac{1}{2}} = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}...$$

$E = \frac{\pi^1\hbar^2}{2mL^2}\frac{(2n+1)^2}{n^2}$ $\frac{n=0,1,2...}{n=1,3,5...}$just as for infinite potential well

Scattering from a Potential Barrier Consider electron incident on a potential step:

(1) $\psi_1(x) = Ae^{i\kappa_1 x} + Be^{-i\kappa_1 x}$ $\kappa_1 = \sqrt{\frac{-2mE-V_1}{\hbar}} = \sqrt{\frac{2mE}{\hbar}}$

(2) $\psi_2(x) = ce^{i\kappa_2 x} + De^{-i\kappa_2 x}$ $\kappa_2 = \sqrt{\frac{2m(E-V_2)}{\hbar}}$

What is the probability that the electron can be found in regions (1) and (2), respectively?

Interpretation:

Wave with amplitude A is partially reflected ($B > 0$) and partially transmitted (C). $D = 0$

Boundary Conditions:

(1) $\psi_1\Big|_{x=0} = \psi_2\Big|_{x=0} \Rightarrow A + B + C$

(2) $\frac{d\psi_1}{dx}\Big|_{x=0} = \frac{d\psi_2}{dx}\Big|_{x=0} \Rightarrow i\kappa_1 (A - B) = i\kappa_2 (C)\; A - B = \frac{\kappa_2}{\kappa_1} C$

$$\text{Add (1) and (2)} \Rightarrow 2A = C\left(1 + \tfrac{\kappa_2}{\kappa_1}\right)\left[\tfrac{C}{A} = \tfrac{2\kappa_1}{\kappa_1+\kappa_2}\right]$$

(1) $A + B = C$, $\frac{B}{A} = \frac{C}{A} - 1$

$\frac{B}{A} = \frac{2\kappa_1}{\kappa_1+\kappa_2} - 1 = \frac{\kappa_1-\kappa_2}{\kappa_1+\kappa_2}$

Consider two cases:

a $E > V_2$, so κ_2 is real.

- Oscillatory solution in Region 2
- $\left(\left|\frac{B}{A}\right| > 0\right)$ Partial reflection (does not occur classically)
  Note: if we used wave packets instead of a single plane wave, reflected amplitudes would be different because the κ's are different.

b $E < V_2$, so κ_2 is imaginary,

- Solution decays exponentially in Region 2.
- $\left|\frac{C}{A}\right| > 0$ Finite probability of finding electron at $x > 0$
- "Tunneling" is a phenomenon that is classically "forbidden."

Homework: Consider narrow barrier.
$E < V_2$
$\kappa_1 = \kappa_3$, real
κ_2 is imaginary
There is a finite probability that the electron will tunnel into Region 3.

109 3-D Potentials

Time-independent Schrödinger Equation:

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r}) + V\psi(\vec{r}) = E\psi(\vec{r}) \quad \Psi(\vec{r},t) = \Psi(r)\, e^{-i\frac{E}{\hbar}t}$$

($\psi(\vec{r})$ are stationary states of definite energy.)
Consider particle in 3-D Box:
$V(x,y,z) = 0$ if $0 < x,y,z < L$; otherwise, it's equal to ∞.
Inside: $-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r}) = E\psi(\vec{r})$ $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$
Solve by separation of variables:
$\psi(\vec{r}) = \psi(x)\,\psi(y)\,\psi(z)$
$-\frac{\hbar^2}{2m}\left[\psi(y)\,\psi(z)\,\frac{\partial^2\psi(x)}{\partial x^2} + \psi(x)\,\psi(z)\,\frac{\partial^2\psi(y)}{\partial y^2} + \psi(x)\,\psi(y)\,\frac{\partial^2\psi(z)}{\partial z^2}\right] = E\psi(x)\,\psi(y)\,\psi(z)$
Divide both sides by $\psi(x)\,\psi(y)\,\psi(z)$.

(1) $\frac{-\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2} = E_x\psi(x)$

(2) $\frac{-\hbar^2}{2m}\frac{\partial^2\psi(y)}{\partial y^2} = E_y\psi(y)$

(3) $\frac{-\hbar^2}{2m}\frac{\partial^2\psi(z)}{\partial z^2} = E_z\psi(z)$

Note: $E_x + E_y + E_z = E$; These are the equations for the 1-D particle in a box.

From Lecture #4:

$$\psi_{nx}(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n_x \pi}{L} x\right) \; E_{nx} = nx^2 \frac{\pi^2 \hbar^2}{2mL^2}$$

$$\therefore \psi_{n_x n_y n_z}(\vec{r}) = \psi_{n_x}(x)\, \psi_{n_y}(y)\, \psi_{n_z}(z) \text{ (reminder: "n" is a quantum \#)}$$

$$E_{n_x n_y n_z} = \frac{\pi^2 \hbar^2}{2mL^2}\left(n_x^2 + n_y^2 + n_z^2\right)$$

The ground state energy $E_{111} = 3E_1$, where $E_1 = \frac{\pi^2 \hbar^2}{2mL^2}$

First excited states:

$$E_{211} = E_{121} = E_{112} = \left(2^2 + 1^2 + 1^2\right) E_1 = GE_1$$

$E_{211} = E_{121} = E_{112}$ is a 3-fold "degeneracy" in energy .

Spherically Symmetric Potentials: $V(\vec{r}) = V(r), r = \left|\vec{r}\right|$

Spherical Coordinates: $x = r \sin\theta \cos\phi$

$$y = r \sin\theta \sin\phi$$

$$z = r \cos\theta$$

$$x^2 + y^2 + z^2 = r^2$$

$$d\vec{r} = r^2 \sin\theta dr d\theta d\phi$$

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial}{\partial\theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial\theta^2}$$

Solve Schrödinger Equation by separation of variables in spherical coordinates:

$$\psi(\vec{r}) = R(r)\, y(\theta, \phi)$$

$$\frac{-\hbar^2}{2m}\left[Y\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial^2}\right) R + R\frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial}{\partial\theta}\right) Y + R\frac{1}{r^2 \sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right] + VRY = ERY$$

Multiply by $\frac{r^2}{RY}$.

(1) Radial Equation: $\frac{1}{R}\frac{\partial}{\partial r}\left(r^2 \frac{\partial R}{\partial r}\right) - \frac{2m}{\hbar^2} r^2 (V - E) = \ell(\ell + 1)$ (We will solve this second.)

(2) Angular Equation: (Multiply by $Y \sin^2\theta$.)

$$\sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial Y}{\partial\theta}\right) + \frac{\partial^2 Y}{\partial\phi^2} = -\ell(\ell + 1)\sin^2\theta Y$$

Solve by separation of variables:

Subsitute: $Y(\theta, \phi) = H(\theta)\Phi(\phi)$

Divide by $H\Phi$.

(3) $\frac{1}{H}\sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial H}{\gamma\theta}\right) + \ell(\ell + 1)\sin^2\theta - m^2 = 0$

(4) $\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2} = -m^2 \Rightarrow \Phi = e^{im\phi}$

In our coordinate system, $\phi \leftrightarrow \phi + 2\pi$, so $\Phi(\phi + 2\pi) = \Phi(\phi)$.

This implies that $m = 0, \pm 1, \pm 2$; m is the Magnetic Quantum Number.

Now solve (3):

$$\sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial H}{\partial\theta}\right) + \left[\ell(\ell + 1)\sin^2\theta - m^2\right] H = 0$$

Soln: $H(\theta) = AP_\ell^m(\cos\theta)$ A: Normalization Constant

$P_\ell^m(x)$ Associated Legendre Function

$$P_\ell^m(x) = \left(1 - x^2\right)^{|m|/2}\left(\frac{d}{dx}\right)^{|m|} P_\ell(x)$$

$P_\ell(x)$ Legendre Polynomial (for integer ℓ)

$$P_\ell(x) = \frac{1}{2^\ell \ell!}\left(\frac{d}{dx}\right)^\ell \left(x^2 - 1\right)^\ell$$

$\ell = 0, 1, 2...$ Azimuthal / Orbital Quantum Number

Rodrigues Formula $P_1^o = \cos\theta$

$P_1^1 = \sin\theta$

$P_2^1 = 3\sin\theta\cos\theta$

$P_2^2 = 3\sin^2\theta$

Magnetic Quantum Number: $|m| \leq \ell$

$-\ell, -\ell+1, ...0, \ell-1, \ell$

Because $P_\ell(x)$is a polynomial of degree ℓ, $P_\ell^m(x)$is non-zero only for $|m| \leq \ell$.

Spherical Harmonies $Y_\ell^m(\theta,\phi) = AP_\ell^m(\cos\theta)\, e^{im\phi}$ A= Normalization Constant

They are normalized such that:

$\int d\eta \left|Y_\ell^m\right|^2 = 1$ where $\int d\eta = \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi$

E.g. $Y:(\theta,\phi) = -\frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta e^{i\phi}$

$Y_2^{-1}(\theta,\phi) = \frac{1}{2}\sqrt{\frac{t5\sin\theta}{2\pi\cos\theta e^{-i\phi}}}$

They are also orthogonal:

$\int d\eta Y_\ell^{m*} Y_{\ell^1}^{m^1} = \delta_{\ell\ell^1}\delta_{mm^1}$ =1 if $\ell = \ell^1$ and $m = m^1$; otherwise, it is equal to 0.

(On the unit sphere, any square integrable function can be expanded as a linear combination of spherical harmonies.)

Radial Equation:

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - \frac{2mr^2}{\hbar^2}(V-E)R = \ell(\ell+1)R$$

Let $v = rR$, where "v" is the radial wavefunction:

$R = \frac{v}{r}$, $\frac{dR}{dr} = \frac{-v}{r^2} + \frac{1}{r}\frac{dv}{dr}$

$\frac{dv}{dr}\left(r^2\frac{dR}{dr}\right) = \frac{d}{dr}\left(-v + r\frac{dv}{dr}\right)$

$\frac{dv}{dr}\left(r^2\frac{dR}{dr}\right) = -\frac{dv}{dr} + \frac{dv}{dr} + r\frac{dv^2}{dr^2} = r\frac{d^2v}{dr^2}$

We then have $r\frac{d^2v}{dr^2} - \frac{2mr}{\hbar^2}(V-E)v = \ell(\ell+1)\frac{v}{r}$.

Multiply by $\frac{-\hbar^2}{2mr}$:

$\frac{-\hbar^2}{2m}\frac{d^2v}{dr^2} + \left[V + \frac{\hbar^2}{2m}\frac{1}{r^2}\ell(\ell+1)\right]v = Ev$

$\frac{\hbar^2}{2m}\frac{1}{r^2}\ell(\ell+1)$ is the Centrifugal Term.

Same form as 1-D Schrödinger Equation, but with extra term that acts like a repulsive potential. (Relate to stability; radius of ψ.)

In summary, for **any** spherically symmetric potential, we have stationary states:

$\Psi(\vec{r},t) = \psi(\vec{r})\, e^{-iE\hbar/t}$, where $\psi(\vec{r}) = R(r)\, Y_\ell^m(\theta,\phi)$

θ, ϕ are spherical harmonies; $\ell = 0, 1, 2...$; $|m| \leq \ell$

and $v = rR$, satisfying the Radial Equation

Normalization $\int d\vec{r}\,|\psi(\vec{r})|^2 = 1$

$\int d\vec{r}\,|\psi(\vec{r})|^2 = \int_0^\infty dr \underbrace{r^2|R|^2}_{|v|^2}\int d\eta \left|Y_\ell^m\right|^2$

$\int d\eta \left|Y_\ell^m\right|^2 = 1$

$\therefore \int_0^\infty dr\,|v|^2 = 1$, so$v \to 0$ as $r \to \infty$

110 The Hydrogen Atom

Task: Solve Radial Equation for $V(r) = \frac{-e^2}{4\pi\varepsilon_0}\frac{1}{r}$.

Define $k = \frac{\sqrt{2m(-E)}}{\hbar}$ for $E < 0$ (Bound State):

$$\frac{1}{\kappa^2}\frac{d^2v}{dr^2}+\left[\frac{e^2}{4\pi E_0}\frac{2m}{\hbar^2\kappa^2}\frac{1}{r}-\frac{1}{\kappa^2}\frac{\ell(\ell+1)}{r^2}\right]v=v$$

Define $p=kr$ (Dimensionless)

$$\frac{d^2v}{dp^2}+\left[\frac{e^2}{4\pi E_0}\frac{2m}{\hbar^2\kappa}\frac{1}{p}-\frac{\ell(\ell+1)}{P^2}\right]v=v\,\frac{e^2}{4\pi E_0}\frac{2m}{\hbar^2\kappa}=p_0$$

$$\frac{d^2v}{dp^2}=\left(1-\frac{p_0}{p}+\frac{\ell(\ell+1)}{p^2}\right)v$$

Considering asymptotic behavior of v, we find:

$$v(p)=p^{\ell+1}e^{-p}v(p)$$

Power Series solution gives:

$V(p)=\sum_{p=0}^{v}a_jp^j$ AND $z(j\max+\ell+1)-p_0=0$This second equation is from the Recursion Relation.

Impose $a_{jmax+1}=0$ so the series will terminate.

Define $n=j\max+\ell+1$ as principle quantum number (int),

then $p_0=zn=\frac{e^2}{4\pi\varepsilon_0}\frac{2m}{\hbar^2\kappa}$ but $\kappa=\frac{\sqrt{2m(-E)}}{\hbar}$

Apparently, $\left[\left[E_n=-\left(\frac{e^2}{v\pi\varepsilon_0}\right)^2\frac{m}{2\hbar^2}\cdot\frac{1}{n^2}=E_1\cdot\frac{1}{n^2}\right]\right]E_1=-13.61eV$.

What values of the orbital quantum number ℓ are allowed for a given n ?

$$n=1,j\max=0\to\ell=0$$

$$n=2,j\max=0\to\ell=1$$

$$j\max=1\to\ell=0$$

$$[[\ell\le n-1]]\kappa_n=\frac{\sqrt{2m(-E_n)}}{\hbar}=\frac{e^2}{4\pi\varepsilon_0}\frac{m}{\hbar^2n}=\frac{1}{an}$$

$a\equiv\frac{4\pi\varepsilon_0}{me^2}\hbar^2$ Bohr Radius

$a=0.529$Å

Find Radial Wavefunctions Ground State: $n=1\to j\max=0\ \ell=0$

$$a_0=A\ a_1,a_2,a_3...=0$$

$$V(\ell)=A$$

$$\Rightarrow v(p)=Ape^{-1}\ (\ell=0)$$

$v(r)=\frac{A}{a}re^{-r/a}$,where $\kappa_n=\frac{1}{na}=\frac{1}{a}$ for $n=1$.

Normalization: $\int_0^\infty|v(r)|^2dr=\frac{|A|^2}{a^2}\int_0^\infty drr^2e^{-2r/a}=1$

$$\int_0^\infty|v(r)|^2dr=2\cdot\frac{|A|^2}{a^2}\left(\frac{a}{2}\right)^3=1$$

$$\text{so }\frac{|A|}{a}=2a^{-3/2}$$

Note: $\int_0^\infty drr^ne^{-2r/a}=n!\left(\frac{a}{2}^{n+1}\right)$

$$\underset{n\ell}{R_{10}}(r)=\frac{v(r)}{r}=2a^{-2/3}e^{-r/a}$$

$$\overset{\psi_{100}}{n\ell m}\left(r,\vec{\theta},\phi\right)=R_{10}(r)\,Y_0^0(\theta,\phi)=\frac{1}{\sqrt{\pi a^3}}e^{-r/a}$$

To generate other Radial Wavefunctions, note that:

$$a_{j+1}=\frac{2(j+\ell-1-n)}{(j+1)(j+2\ell+2)}a_j$$

(In addition, the series must terminate, so we imposed: $a_{j\max+1}=0$ for $V(p)=\sum_{j=0}^{\infty}a_jp^i$).

Excited States where $\psi_{n\ell m}=R_{n\ell}Y_\ell^m$

$$\ell\le n-1\ |m|\le\ell$$

$$E_2=\frac{E_1}{(z)^2}=\frac{-13.6eV}{4}=-3.4eV$$

For $\ell=0:R_{20}(r)=\frac{1}{\sqrt{2a^3}}\left(1-\frac{r}{2a}\right)e^{-r/2a}$ # of radial nodes

For $\ell=1:R_{21}(r)=\frac{1}{2Ya^3}\frac{r}{a}e^{-r/2a}=n-\ell-1$

$$\psi_{200}=R_{20}(r)\,Y_0^0(\theta,\phi)=R_{20}(r)\cdot\frac{1}{\sqrt{4\pi}}$$

$$\psi_{210} = R_{21}(r)\, Y_1^0(\theta,\phi) = R_{21}(r)\sqrt{\tfrac{3}{4\pi}}\cos\theta$$
$$\psi_{21\pm1} = R_{21}(r)\, Y_1^{\pm1}(\theta,\phi) = R_{21}(r) \mp \sqrt{\tfrac{3}{8\pi}}\sin\theta e^{\pm i\phi}$$

$\ell = 0$: Spherically Symmetric Wavefunctions
$\ell = 1$: Not Spherically Symmetric
Look at Slice parallel to z axis.
The Schrödinger Equation in three-dimension:

Lecture 6: Slide Notes Questions to answer today:

(1) What is the relationship between $R_{n\ell_1} Y_\ell^m$ and the probability density? (visualize)

(2) What determines the energy levels and spectrum of the hydrogen atom?

(3) What are the consequences of electron spin on atomic structure?

For Discussion:

- Radial Wavefunctions:
  - What are the right and left columns?
  - For $\ell = 0$, what is the trend with n ?
  - For given n, what is the trend with ℓ ?
- Spherical Harmonics:
  - What are the trends with ℓ, m ?
- Combined:
  - What are the trends in the number of nodes?
  - Why is $E_n \alpha -\frac{1}{n^2}$

Hydrogen Spectrum When electron undergoes transition from excited state to lower energy state, energy can be conserved through emission of a photon with energy E_γ.

$$E_\gamma = E_i - E_f = -\left(\tfrac{e^2}{4\pi\varepsilon_0}\right)^2 \tfrac{m}{2\hbar}\left(\tfrac{1}{n_i^2} - \tfrac{1}{n_f^2}\right) E_r = \tfrac{nc}{X}$$
$$\Rightarrow \tfrac{1}{\lambda} = \left(\tfrac{e^2}{4\pi\varepsilon_0}\right)^2 \tfrac{m}{4\pi\hbar^3 c}\left(\tfrac{1}{nr^2} - \tfrac{1}{nc^2}\right)$$

$\left(\tfrac{e^2}{4\pi\varepsilon_0}\right)^2 \tfrac{m}{4\pi\hbar^3 c} \to$ Rydberg Constant $1.10 \times 10^7 m^{-1}$
$n_f = 1 \to$Lyman Series (UV)
$n_f = 2 \to$Balmer Series (VIS)
$n_f = 3 \to$Daschen Series (IR)

The states of the hydrogen atom can be specified by quantum numbers:

(1) n : Principle Quantum # specifies the energy level $\quad n = 1, 2, 3...$

(2) ℓ : Orbital Quantum # $\ell \leq n - 1$

(3) m : Magnetic Quantum # $|m_\ell| \leq \ell$

(4) m_s: Spin $m_s = \pm\frac{1}{2}$

From spectroscopy, we have the nomenclature:
$\ell = 0$ "s" shap
$\ell = 1$ "p"principal
$\ell = 2$ "d"diffuse
$\ell = 3$ "f"fundamental
$\ell = 4$ "g"

Physical Interpretation of ℓ, m_ℓ ℓ and m are related to the orbital angular momentum of the electron.
Classically,

$$L_x = yp_z - zp_y$$
$$L_y = zp_x - xp_z$$
$$L_z = xp_y - yp_x$$

We have previously defined quantum mechanical operators for p_x , p_y, p_z, so we can write:

$$L_x = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \quad L_y = \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \quad L_z\frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

These operators do not commute, i.e.,

$$\left[L_x,\, L_y\right] = L_xL_y - L_yL_x \neq 0$$

They are incompatile observables - they cannot both be known with arbitrary certainty at the same time, just like x and p.
(In fact, $[x, p_x] \neq 0) \rightarrow \triangle x \triangle p = \frac{\hbar}{2}$
It can be shown that both $L_z$and L^2 are compatible observables, with:

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

(See Griffiths.)
For our purposes, it will be sufficient to note that:
ℓ is a measure of $L^2L^2\psi = \hbar^2\ell\left(\ell+1\right)\psi$
m is a measure of $L_zL_z\psi = \hbar m\psi$
(The forms of these operators are given in Griffiths, p. 168 and 169.)

Spin

- **Intrinsic** angular momentum of point particle
- Purely quantum mechanical: quantum #'s s, m_s
- **Important** practically$\rightarrow$ magnetism, optical transitions
- Profound consequences for structure of matter

Classically,
Spin can be composed into orbital momenta. This is not possible for point particles.
For electrons, protons and neutrons, $|spin| = \frac{1}{2}$.
For photons, $|spin| = 1$.

Classes of Identical Particles Suppose we represent two identical particles by the wave function $\psi(\vec{r_1}, \vec{r_2})$:
Define:

- Bosons: $\psi(\vec{r_1}, \vec{r_2}) = \psi(\vec{r_2}, \vec{r_1})$ Symmetric Integer Spin
- Fermions: $\psi(\vec{r_1}, \vec{r_2}) = -\psi(\vec{r_2}, \vec{r_1})$ Antisymmetric Half-Integer Spin

Let's express $\psi(\vec{r_1}, \vec{r_2})$ in terms of wave functions for two states a and b: $\psi_a(\vec{r_1}), \psi_b(\vec{r_2})$.

Bosons: $\psi(\vec{r_1}, \vec{r_2}) = A\left[\psi_a\left(\vec{r_1}\right)\psi_b\left(\vec{r_2}\right) + \psi_b\left(\vec{r_1}\right)\psi_a\left(\vec{r_2}\right)\right]$

- Symmetric under exchange of $\vec{r_1}$ and $\vec{r_2}$
- If both particles are in state a, $\psi\left(\vec{r_1}, \vec{r_2}\right) = 2A\psi_a\left(\vec{r_1}\right)\psi_a\left(\vec{r_2}\right)$.
- Bosons can occupy the same state.

Fermions: $\psi(\vec{r_1}, \vec{r_2}) = A\left[\psi_a(\vec{r_1})\,\psi_b(\vec{r_2}) - \psi_b(\vec{r_1})\,\psi_a(\vec{r_2})\right]$

- Antisymmetric under exchange of $\vec{r_1}, \vec{r_2}$
- Both particles **cannot** be in the same state: e.g. if both "in" a, then $\psi(\vec{r_1}, \vec{r_2}) = 0$.
- Pauli Exclusion Principle

The entire wave function must be antisymmetric with respect to exchange. This has implications for bonding (allowed configurations of electons + energy of configurations).
We now have 4 quantum numbers to describe electron wavefunctions in spherically symmetric potentials:
Now, consider the rest of the Periodic Table, beginning with He.
Schrödinger Equation becomes:
For neutral atoms of atomic Hz,

$$H = \sum_{j=1}^{z}\left[-\frac{\hbar^2}{2m}\nabla_j^2 - \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{ze^2}{r_j}\right] + \frac{1}{2}\left(\frac{1}{4\pi\varepsilon_0}\right)\sum_{j\neq\kappa}^{z}\frac{e^2}{|\vec{r_j}-\vec{r}_\kappa|}$$

Note: This cannot be solved exactly for $z > 1$! The $e^- - e^-$ repulsion term is the problem.
If we ignore this interpretation,
$E_n = -13.6\frac{z^2}{n^2} = -54.4eV$ for $n = 1$
But the measured value is $-24.6eV$
$\therefore e^- - e^-$ interaction is very important to electronic structure.
(If we were to ignore interactions) How many degenerate states would we have for each v?
$\ell \leq n-1\ |m_\ell| \leq \ell\ m_s = \pm\frac{1}{2}$
$2\ell + 1$ values of m_ℓ for each ℓ, 2 values of m_s
$2 \times \sum_{\ell=0}^{n-1}(2\ell+1) = 2n^2$ (arithmetic series)
Here, $m_s = 2$ and $2\ell + 1 = m_\ell$.
Recall that Pauli Exclusion forbids electrons for being in the same state.

n	#
1	2
2	8
3	18
4	32

This is not the same structure as the periodic table, however. $e^- - e^-$ interaction changes energies and influences state filling.

111 Bonding

- Basic Concepts: Mechanics of Bonding
- Bond Types: Qualitative Description
- Quantum Mechanical Description of Covalent Bond

Characteristics of a Bond:

- Negative Potential Energy
- Stable: **Energy is minimized** (E_0).
  - $F_{TOT} = F_R + F_A = 0$ at r_0
  - R refers to repulsive; A refers to attractive.
  - Alt: $F = \left.\frac{-dE}{dr}\right|_{r_0} = 0$ r_0 : bond length

- $E_0 = \int_\infty^{r_0} dr F_{TOT} = \int_\infty^{r_0} dr F_R(r) + \int_\infty^{r_0} dr F_A(r)$
  $E_0 = \int_\infty^{r_0} dr F_{TOT} = E_R + E_A$
- For example, E_R is the Coulomb Attraction and E_A is the Coulomb Repulsion.

Quantitatively...

Example Potential: $E(r) = \frac{A}{r^n} - \frac{B}{r^m} \; > m$
Bond Length r_0 typically $1.5 - 3\mathring{A}$
Bond Energy E_0 typically $0.1 - 10eV$
At finite T, $\langle r \rangle = r_0$

Influence on Mechical Properties:

(1) Elastic Modulus = $\frac{Stress}{Strain}\left(\frac{\Delta F}{\triangle x}\right)$

a $ElasticModulus \propto \left.\frac{dF}{dr}\right|_{r=r_0} \propto \left.\frac{d^2E}{dr^2}\right|_{r=r_0}$ (Curvature of Potential)

(2) Thermal Expansion:

a $\langle r \rangle$ increases with temperature:

i. Smaller if bonds are strong (Deep Wells)

ii. Depends on $\frac{d^3E}{dr^3}$ (Anharmonic Terms)

(3) Strong Bonds Increase T_m

Bond Types

(1) **Ionic**

- Energy is minimized by electron transfer from electropositive to electronegative atom.
- Atoms bound by Coulomb Attraction:
  - E.g. Na $1s^2 2s^2 2p^6 3s^1 \rightarrow$ Ne
  - E.g. Cl $1s^2 2s^2 2p^6 3s^2 3p^5 \rightarrow$ Ar
  - $E(r) = \frac{A}{r^n} - \frac{e^2}{4\pi\varepsilon_0 r}$ for 1 pair
  - Summing:
    - $-6 \times \frac{e^2}{4\pi\varepsilon_0 a} \quad a = r_0$
    - $+12 \times \frac{e^2}{4\pi\varepsilon_0 a\sqrt{2}}$
    - etc.
  - Infinite Sum $\Rightarrow E(r) = \frac{A}{r^n} - \frac{Me^2}{4\pi\varepsilon_0 r}$

Structure	Madelung Constant
CsCl	1.7627
NaCl	1.7476
Zincblende	1.6381

$Etyp$ Few eV

(2) **Van Der Walls**

Fluctuating dipole on one atom/molecule induces dipole on neighbor⇒Attraction.

e.g. liquid He, molecular crystals

Potential well described by Lennard-Jones Potential:

$$E(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

For a crystal, $E(r) = 2\varepsilon\left[A_{12}\left(\frac{\sigma}{r}\right)^{12} - A_6\left(\frac{\sigma}{r}\right)^{6}\right]$, where A_{12}, A_6 depend on crystal structure.

VDW bonds are weak: $\sim 0.01eV$.

(3) **Covalent Bonds**

- Strong bonds arising from **shared** electrons
- Highly directional (unlike ionic, metallic)
- Less dense packing

Feynman's Coupled Mode Approach to Hydrogen Molecular Ion

a $H\psi = i\hbar\frac{\partial\psi}{\partial t}$, where $H = -\frac{\hbar^2}{2m}\nabla^2 + V$
H = Hamiltonian Operator.
We attempt a general solution.

b $\Psi\left(\vec{r}, t\right) = \sum_j w_j(t)\,\psi_j\left(\vec{r}\right)$ $\quad$ $\psi_j\left(\vec{r}\right)$: Stationary State

Q: How does probability of being in ψ_j vary with time?

A: According to w_j , substitute (b) into (a) to test solution:

$$\sum_j w_j H\psi_j = i\hbar\sum_j \psi_j\frac{dw_j}{dt}$$

Multiply both sides by ψ_k^* and integrate over all $\vec{r}$:

$\sum_j w_j \int \psi_k^* H\psi_j d\vec{r} = i\hbar\sum_j \frac{dw_j}{dt}\int d\vec{r}\,\psi_k^*\psi_j = i\hbar\frac{dw_k}{dt}$

$d\vec{r}\,\psi_k^*\psi_j = \delta_{kj}$

Define $H_{kj} = \int d\vec{r}\,\psi_k^* H\psi_j$.

Then, $i\hbar\frac{dw_k}{dt} = \sum_j H_{kj}w_j$

Consider for 2 Stationary States

a **Example #1**: $1e^-$, $2p$ very far apart
$i\hbar\frac{dw_1}{dt} = H_{11}w_1 + H_{12}w_2$
$i\hbar\frac{dw_2}{dt} = H_{21}w_1 + H_{22}w_2$

b For "Uncoupled" States, $H_{12} = H_{21} = 0$
$\Rightarrow i\hbar\frac{dw_1}{dt} = H_{11}w_1$, $w_1(t) = \kappa_1 e^{-iH_{11}t/\hbar}$
Probability of being in $\psi_1 \propto \left|w_1(t)\right|^2 = \left|\kappa_1\right|^2 = Constant$
(Physically, there is no chance of sharing electron between ψ_1, ψ_2)
If d is small, electron can tunnel $(H_{12} \neq 0)$.
Associate tunneling with weak coupling - increases exponentially as d decreases
$H_{12} = H_{21} = -A$
Assume: $H_{11} = H_{22} = E_0$ and $H_{12} = H_{21} = -A$
- E_0 is energy in absence of coupling

c Attempt Solutions:

$$i\hbar\frac{dw_1}{dt} = E_0 w_1 - A w_2 \; w_1(t) = \kappa_1 e^{-iEt/\hbar}$$
$$i\hbar\frac{dw_2}{dt} = -A w_1 + E_0 w_2 \; w_2(t) = \kappa_2 e^{-iEt/h}$$
$$\kappa_1 E = E_0\kappa_1 - A\kappa_2 \; 0 = (E_0 - E)\kappa_1 - A\kappa_2$$
$$\kappa_2 E = -A\kappa_1 + E_0\kappa_2 0 = -A\kappa_1 + (E_0 - E)\kappa_2$$

	Attempt Solutions:
$i\hbar\frac{dw_1}{dt} = E_0 w_1 - A w_2$	$w_1(t) = \kappa_1 e^{-iEt/\hbar}$
$i\hbar\frac{dw_2}{dt} = -A w_1 + E_0 w_2$	$w_2(t) = \kappa_2 e^{-iEt/h}$
$\kappa_1 E = E_0\kappa_1 - A\kappa_2$	$0 = (E_0 - E)\kappa_1 - A\kappa_2$
$\kappa_2 E = -A\kappa_1 + E_0\kappa_2$	$0 = -A\kappa_1 + (E_0 - E)\kappa_2$

A solution exists if and only if:

(1) $$\begin{vmatrix} E_0 - E & -A \\ -A & E_0 - E \end{vmatrix} = 0$$

a $(E_0 - E)^2 = A^2$ –> $E = E_0 \pm A$
Sharing ⇒ Energy Lowering ⇒Bonding

The Hydrogen Molecule Let $\psi\left(\vec{r_1}, \vec{r_2}\right)$.

Describe $2e^-$, $2p$.

For identical fermions, we must have $\psi\left(\vec{r_1}, \vec{r_2}\right) = -\psi\left(\vec{r_2}, \vec{r_1}\right)$.

Consider spatially antisymmetric ψ (parallel spins).

Zero probability of electron being between protons.
→ Large repulsion

But if e^- spins are antiparallel, spatial wavefunction must be symmetric for $\psi\left(\vec{r_1}, \vec{r_2}\right) = -\psi\left(\vec{r_2}, \vec{r_1}\right)$.

Probability of finding e^- between protons **lowers** the energy.

Spins	Total Spin	Spatial Symmetry	Energy	Name
Anti-parallel	$\frac{1}{2} + \left(-\frac{1}{2}\right) = 0 \; ; m_s = 0$	Symmetric	Lower; "Bonding"	Singlet, $s = 0$
	$\left(\frac{1}{2}, -\frac{1}{2}\right) - \left(-\frac{1}{2}, \frac{1}{2}\right)$	Symmetric	Lower; "Bonding"	Singlet, $s = 0$
Parallel	$m = 1; \frac{1}{2} + \frac{1}{2} = 1$	Antisymmetric	Higher; "Anti-Bonding"	Triplet, $s = 1$
	$m = 0; \left(\frac{1}{2}, -\frac{1}{2}\right) + \left(-\frac{1}{2}, \frac{1}{2}\right) = 1$	Antisymmetric	Higher; "Anti-Bonding"	Triplet, $s = 1$
	$m = 0; \frac{1}{2} + \left(-\frac{1}{2}\right) = 0$	Antisymmetric	Higher; "Anti-Bonding"	Triplet, $s = 1$
	$m = -1; -\frac{1}{2} + \left(-\frac{1}{2}\right) = -1$	Antisymmetric	Higher; "Anti-Bonding"	Triplet, $s = 1$

(2) **Metallic Bonding**

- Valence electrons delocalized throughout lattice.
- Electrons are like free electrons (**Free Electron Theory of Metals**).

Consequences

(1) High electrical conductivity
(2) High thermal conductivity
(3) High ductility & malleability

112 Supplement (Bridge to Lecture 9.1)

The distribution of particles in allowed states is given by **the density of states times (x) probability of occupation.**

States	
Distinguishable Particles	$N_n = d_n e^{-(\propto t\beta E_n)} = d_n F_0(E)$
Fermions	$N_n = \frac{dn}{e^{\propto+\beta E_n}+1} = dnF_{FD}(E)$
Bosons	$N_n = \frac{dn}{e^{\propto+\beta E_n}-1} = dnF_{\beta D}(E)$

Define $\beta \equiv \frac{1}{\kappa T} \propto\equiv -M \cdot \beta$
N_n: Most probable occupation number (number within energy window dE)
d_n: Degeneracy of energy level $\rightarrow$ Density of states $\frac{dN}{dE}$, $\frac{dN}{d\kappa}$
In calculations of total particle number, total energy, and quantities that depend on them, we replace discrete sums with integrals.

$$N = \sum_{n=1}^{\infty} N_n \rightarrow \int dE \frac{dn}{dE} F(E) = \int dE z(E) F(E) \text{ OR}$$

$$N = \sum_{n=1}^{\infty} N_n \rightarrow \int d\kappa \frac{dn}{d\kappa} F(\kappa) = \int dk z(\kappa) F(\kappa)$$

$z(E)$ and $z(\kappa)$ refer to the density of states.
We next learn how to calculate the density of states for electrons and phonons (and photons).

113 Modeling Metals

(S+W 6.1-6.3) (Griffiths 5.3)
Assume particles experience no forces except at impenetrable walls (non-interacting).

- $V(x,y,z) = 0, 0 < x,y,z < L$
- $V(x,y,z) = \infty$ otherwise

Sowing the 3-D Schrödinger Equation, the allowed energies of the stationary states are:

$$E_{n_x n_y n_z} = \frac{\hbar^2}{2m}\frac{\pi^2}{L^2}\left(n_x^2 + n_y^2 + n_z^2\right) n = 1, 2, 3...$$

$$E_{n_x n_y n_z} = \frac{\hbar^2}{2m}\left(k_x^2 + k_y^2 + k_z^2\right) = \frac{\hbar^2 k^2}{2m}, \text{ where } k_{x,y,z} = \frac{\pi}{L} n_{x,y,z}.$$

114 Free Electron Gas

k values associated with these states look like a grid of points in a plot of k_x, k_y, k_z

- Each point represents a stationary state
- Each state/ cube takes up volume in k-space: $\left(\frac{\pi}{L}\right)^3 = \frac{\pi^3}{V}$
- At $T = 0$, all Bosons or classical particles would be in lowest energy state: ψ_{111}

- Electrons, being Fermions, obey the Pauli Exclusion Principle:
  $\rightarrow$ Only 2 per state $(M_s \pm 1/2)$.
  N electrons fill up one octant of sphere of radius k_f:

$$\frac{1}{8}\left(\frac{4}{3}\pi k_F^3\right) = \frac{N}{2}\left(\frac{\pi^3}{V}\right) \Rightarrow k_F = \left(3\pi^2\frac{N}{V}\right)^{1/3}$$

  At $T = 0$, all states inside this **Fermi surface** are occupied, while all outside are empty

Energy of highest occupied state
Fermi Energy, $E_F = \frac{\hbar^2}{2m}k_F^2 = \frac{\hbar^2}{2m}\left(3\pi^2\frac{N}{V}\right)^{2/3}$
Density of States - number of states at a specific energy

- Between E and E+dE

First find number N of states between k and $k + dk$

- In k-space, a shell of thickness dk contains volume $\frac{1}{8}\left(4\pi k^2\right)dk$.
- $dN = \frac{\frac{1}{8}\left(4\pi k^2\right)dk}{\frac{\pi}{V}\frac{1}{2}} = \frac{V}{\pi^2}k^2dk$
- $\frac{dN}{dk} = \frac{V}{\pi^2}k^2$

Write in terms of energy:
$E = \frac{\hbar^2k^2}{2m} \Rightarrow dE = \frac{\hbar^2k}{m}dk$
$\frac{dN}{dk} = \frac{V}{\pi^2}k^2$ note: $\frac{dE}{dk} = \frac{\hbar^2}{m}k\left(E = \frac{\hbar^2}{2m}k^2\right)$
$\frac{dN}{dE} = \frac{V}{\pi^2}k\frac{m}{\hbar^2} = \frac{V}{\pi^2}\sqrt{\frac{2mE}{\hbar^3}}m = \frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{E}$
$\therefore g(E) = \frac{dN}{dE} = \frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}E^{1/2}$
At non-zero temperature: thermal excitations move some electrons to states at energies above EF.
Described by the **Fermi-Dirac distribution function:**
$f(E) = \frac{1}{e^{(E-\mu)/kT}+1}$
We first consider the simple case of $T = 0$
$F(E) = 1, E < \nu(o)$
$F(E) = 0, E > \nu(o)$
$\nu(o) = E_f$
$E_{Tot} = \int_0^\infty Eg(E)f(E)dE$
$E_{Tot} = \int_0^{E_F}\frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{1/2}E^{3/2}dE = \frac{V}{5\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}E_F^{5/2}$
$N = \int_0^\infty g(E)f(E)dE = \int_0^{E_F}\frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}E^{1/2}dE = \frac{V}{3\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}E_F^{3/2}$
Since $\langle E\rangle = \frac{E_{Tot}}{N}$, dividing the two expression above yields:
$\langle E\rangle = \frac{3}{5}E_F$
$E_F = \frac{\hbar^2}{2m}\left(3\pi^2\frac{N}{V}\right)^{2/3} = \frac{\hbar^2}{2m}\kappa_F^2$
Alternatively, $E_{TOT} = \frac{\hbar^2\kappa_F^5}{10\pi^2m}\cdot V$ (as one could also find by integrating $d\omega/d\kappa$)
What is the average energy per electron?
$k_F = 3\pi^2\frac{N}{V}$
For classical particles, $\langle E\rangle = 0$ circa $T = 0$, $P = 0$ also:
Here $P = -\frac{dE_{Tot}}{dV} = -\frac{d}{dV}[\frac{3}{5}N\frac{\hbar^2}{2m}\left(3\pi^2\frac{N}{V}\right)^{2/3}] = \left(3\pi^2\right)^{2/3}\frac{\hbar^2}{5m}\left(\frac{N}{V}\right)^{5/3}$
(A positive change in volume decreases the energy.)
Therefore, $P = \left(3\pi^2\right)^{2/3}\frac{\hbar^2}{5m}\left(\frac{N}{V}\right)^{5/3}$
The pressure is non-zero even at $T = 0$ due to the Pauli Exclusion Principle.

$$B = -V\frac{\partial P}{\partial V}$$

$$B = -V\frac{\partial P}{\partial V} = -V\frac{\partial}{\partial V}[\left(3\pi^2\right)^{2/3}\frac{\hbar^2}{5m}\left(\frac{N}{V}\right)^{5/3}] = \left(3\pi^2\right)^{2/3}\frac{\hbar^2}{3m}\left(\frac{N}{V}\right)^{5/3} = \frac{2}{3}\frac{N}{V}E_F$$

What is the specific heat?

$$C_V = \frac{d}{dT}\left\langle E\right\rangle = \frac{1}{N}\frac{d}{T}\int_0^\infty dE\, E\cdot z\left(E\right)\, F\left(E\right)$$

This integral is difficult to evaluate for $T > 0$ (See A+M Chapter 2 for Sommerfeld Expansion to get this.)

$$C_V = \frac{\pi^2}{2}\frac{\kappa_B^2 T}{E_F}$$

We can estimate the electronic contribution to specific heat by considering how many electrons will be affected by thermal excitations.
Electrons within $\sim \kappa_B T$ below E_F will be excited to states within $\sim \kappa_B T$ above E_F.

$$N \backsimeq \left.\frac{dN}{dE}\right|_{E=E_F} \times \kappa_B T = z\left(E_F\right)\kappa_B T$$

The additional energy stored in electronic degrees of freedom is $N \times \kappa_B T$ or

$$\Delta E_{TOT} \simeq z\left(E_\kappa\right)\cdot\left(\kappa_B T\right)^2$$

$$\therefore C_V = \frac{d\left\langle E\right\rangle}{dT} \simeq \frac{2z(E_F)}{N}\kappa_B^2 T$$

Note: Show $z\left(E_f\right) = \frac{3}{2}\frac{N}{E_F}$.

So $C_V \backsimeq \frac{3\kappa_B^2 T}{E_F}$, which is close to $\frac{\pi^2}{2}\frac{\kappa_B^2 T}{E_F}$.

For a classical "ideal" gas, $C_V = \frac{3}{2}\kappa_B$.

Electrons make a negligible contribution to C_V at $300K$ because $\kappa_B T \ll E_F$ (little excitation to higher states)

Experimentally, $C_V = \gamma T + AT^3$.

For a metal:

- Low Temperatures: Linear Dependence
- High Temperatures: Cubic Dependence

Low temperature specific heat is consistent with expectations of free electron model:

$$C_V = \left(\frac{\pi^2}{2}\frac{\kappa_B^2}{E_F}\right)T$$

This model fails at higher temperatures, where lattice vibrations, or phonons, become more important.
Next Steps: Apply quantum statistical mechanics to phonons.

115 Phonons: Quantized Lattice Vibrations / Elastic Waves

(1) Transverse: Displacement perpendicular to direction of propagation (like light)

(2) Longitudinal: Displacement parallel to direction of propagation

We will describe two models for the phonon density of states and calculate N, E_{TOT} , and C_v.

Einstein Model: Assume N atoms of lattice oscillate with frequency ω_0.

$z(\omega) = \frac{dn}{d\omega} = 3 \times N\delta(\omega - \omega_0)$; 3 refers to "3 polarizations"

$E = \hbar\omega$, as with photons

Also, as phonons are bosons for which N is not conserved, $\propto = 0$, and

$$N_\omega = \frac{dN}{e^{\hbar\omega/\kappa_B T} - 1} \text{ or } F(\omega) = \frac{1}{e^{\hbar\omega/\kappa T} - 1}$$

$$E_{TOT} = \int_0^\infty d\omega z(\omega) F(\omega) \cdot E(\omega)$$

$$E_{TOT} = \int_0^\infty d\omega \frac{3N\delta(\omega - \omega_0) \cdot \hbar\omega}{e^{\hbar\omega/\kappa T} - 1} = \frac{3N\hbar\omega_0}{e^{\hbar\omega_0/\kappa T} - 1}$$

$$C_v = \frac{d}{dT}\langle E \rangle = 3\hbar\omega_0 \frac{(-1)}{\left(e^{\hbar\omega/\kappa T} - 1\right)^2} \left(-\frac{\hbar\omega_0}{\kappa T^2}\right) e^{\hbar\omega_0/\kappa T}$$

$$C_v = 3\kappa_B \left(\frac{\hbar\omega_0}{\kappa T}\right)^2 \frac{e^{\hbar\omega_0/\kappa T}}{\left(e^{\hbar\omega_0/\kappa T} - 1\right)^2}$$

Rewrite with $\theta_E = \frac{\hbar\omega_0}{\kappa_B}$, so

$$C_v = 3\kappa_B \left(\frac{\theta_E}{T}\right)^2 \frac{\exp(\theta_E/T)}{[\exp(\theta_E/T) - 1]^2}$$

At high temperatures,

$$C_v \simeq 3\kappa_B \left(\frac{\theta_E}{T}\right)^2 \cdot \frac{1}{\left[1 + \frac{\theta_E}{T} - 1\right]^2} \simeq 3\kappa_B$$

This is in agreement with high temperature data, but this does not give $C_v \propto T^3$.

Debye Model:

$$\frac{dn}{d\kappa} = \frac{3 \times \frac{1}{8} 4\pi\kappa^2}{\pi^3/v} = \frac{3}{2}\frac{V}{\pi^2}\kappa^2$$

Assume: $\kappa = \frac{\omega}{\nu}$ and $d\kappa = \frac{d\omega}{\nu}$

$\frac{dn}{d\omega} = \frac{3}{2}\frac{V}{\pi^2}\frac{\omega^2}{\nu^3}$, where ν is the speed of sound.

Assume: There are a finite number of modes $3N$, as there are N oscillators.

How many modes are there?

$$3N = \int_0^{\omega_D} \frac{3V\omega^2}{2\pi^2\nu^3} d\omega = \frac{V\omega_D^3}{2\pi^2\nu^3} \Rightarrow \frac{\omega_D}{\nu} = \left(6\pi^2\frac{N}{V}\right)^{1/3}$$

Where ω_D is a cut-off frequency:

$$E_{TOT} = \int_0^{\omega_D} d\omega \frac{3V}{2\pi^2}\frac{\omega^2}{\nu^3} \cdot \frac{1}{e^{\hbar\omega/\kappa T} - 1} \cdot \hbar\omega$$

$$z(\omega) = \frac{3V}{2\pi^2}\frac{\omega^2}{\nu^3}$$

$$F(\omega) = \frac{1}{e^{\hbar\omega/\kappa T} - 1}$$

$$E = \hbar\omega$$

$$C_v = \frac{d\langle E \rangle}{dT} = \frac{1}{N}\frac{3V}{2\pi^2}\frac{\hbar}{\nu^3}\int_0^{\omega_D} d\omega\omega^3 \frac{(-1)}{\left(e^{\hbar\omega/\kappa T} - 1\right)^2}\left(\frac{-\hbar\omega}{\kappa_B T^2}\right) e^{\hbar\omega/\kappa T}$$

$$C_v = \frac{V}{N}\frac{3}{2\pi^2}\frac{\hbar^2}{\nu^3\kappa T^2}\int_0^{\omega_D} d\omega \frac{\omega^4 e^{\hbar\omega/\kappa T}}{\left(e^{\hbar\omega/\kappa T} - 1\right)^2}$$

Let $x = \frac{\hbar\omega}{\kappa_B T}$, $\omega = \frac{\kappa T}{\hbar}x$, $d\omega = \frac{\kappa T}{\hbar}dx$, $x_D = \frac{\hbar\omega_D}{\kappa_B T} = \frac{\theta_D}{T}$

θ_D is the Debye Temperature.

$$\theta_D = \frac{\hbar\omega_D}{\kappa_B} = \frac{\hbar\nu}{\kappa_B}\left(6\pi^2\frac{N}{V}\right)^{1/3}$$

$$C_v = 9\kappa_B \left(\frac{T}{\theta_D}\right)^3 \int_0^{x_D} dx \frac{x^4 e^x}{(e^x - 1)^2}$$

For $T \gg \theta_D$, $x \to 0$, and $C_v \to 3\kappa_B$ Classical Result

(See figures.)

For $T \ll \theta_D$, $x_D \to \infty$

$$C_v = 9\kappa_B \left(\frac{T}{\theta_D}\right)^3 \int_0^\infty \frac{x^4 e^\kappa}{(e^x-1)^2} dx \int_0^\infty \frac{x^4 e^\kappa}{(e^x-1)^2} dx = 4\pi^4/15$$

$$C_v = \frac{12\pi^4}{\hbar} \kappa_B \left(\frac{T}{\theta_D}\right)^3$$

$\Rightarrow$Debye Model gives $C_v \propto T^3$ at low T

Phonons (quantized lattice vibrations) act like waves with momenta $p = \hbar\kappa$ (crystal momentum).
We will need to consider energy and momentum conservation in interactions with other particles.

Modifications to the Free Electron Theory of Metals

(1) **Work Function**: finite potential barrier to electron escape (recall the photoelectric effect)
Note: this doesn't affect the energy levels much.

(2) **Thermionic Emission**: Thermal energy can also induce electron emission (high temps).
From Free Electron Theory, $E = \frac{1}{2m}\left(Px^2 + Py^2 + Pz^2\right)$
Assume surface normal parallel to x,
electrons would be emitted classically if: $\frac{Px^2}{2m} > E_f + \phi$.
Quantum mechanics: finite probability of reflecBution from potential step
Flux: $\frac{p_x}{m}\frac{N(p_x)}{V} dp_x$ is the number/area arriving at surface with momentum between P_x and $P_x + dP_x$.

Emission Current Density:

$$J = \int_{\sqrt{2m(E_f+\phi)}}^\infty \left[1 - R\left(p_x\right)\right] \frac{p_x}{m} \frac{eN\left(p_x\right)}{V} dp_x$$

$p_x > \sqrt{2m\left(E_f + \phi\right)}$ is the minimum P_x needed to get past step.

$$N\left(p_x, p_y, p_z\right) dp_x dp_y dp_z = V(\frac{2}{h})^3 exp\left[\frac{E_F - (p_x^2 + p_y^2 + p_z^2)/2m}{k_B T}\right] dp_x dp_y dp_z$$

$$N\left(p_x\right) dp_x = V(\frac{2}{h})^3 exp\left[\frac{E_F - (p_x^2/2m)}{k_B T}\right] dp_x \int_0^\infty exp\left[\frac{-p_y^2}{2mk_B T}\right] dp_y \int_0^\infty exp\left[\frac{-p_z^2}{2mk_B T}\right] dp_z$$

$$N\left(p_x\right) dp_x = V(\frac{4\pi m k_B T}{h^3}) exp\left[\frac{E_F - (p_x^2/2m)}{k_B T}\right] dp_x$$

$$J = (\frac{4\pi mek_B T}{h^3})\left[1 - R\right] exp\left[\frac{E_F}{k_B T}\right] \int_{\sqrt{2m(E_f+\phi)}}^\infty p_x exp\left[\frac{-p_x^2}{2mk_B T}\right] dp_x$$

We use the following to solve:

(1) Free Electron Theory: $p_x = \frac{h}{2L} N_x$

a $R\left(P_x\right) = constant$ (approximately)

b Maxwell-Boltzmann Tail of Fermi-Dirac Distribution matters most:
$J = A_0 (1-R) T^2 e^{-\phi/k_B T}$, $A_0 = \frac{4\pi e m k_B^2}{h^3}$
Current depends exponentially on ϕ, T.

(2) **Schottky Effect**: Thermionic Emission modified by the image force and the electric field.

a Image force modifies the potential.
b Voltage is often applied.

$\therefore$ The new barrier height is given by maximum of $V(x) = V_1 + V_2$
$J = A(1-R) T^2 e^{-\phi eff/\kappa_B T} = A_0 (1-R) T^2 exp\left[-\left(\phi - e\sqrt{\frac{e\varepsilon}{4\pi\varepsilon_0}}\right)/\kappa_B T\right]$
This indicates that $\ln(J) \propto \sqrt{\varepsilon}$

(3) **Field Emission**: At high fields, lower/more narrow barrier leads to significant tunneling.
If we assume a rectangular barrier, then:

$$T = \frac{1}{1 + \frac{V_0^2 sinh^2 2K_2 a}{4E(V_0 - E)}}$$

$$K_2 = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

$$T \propto exp \frac{-4a\sqrt{2m(V_0 - E)}}{\hbar}$$

$$T \propto exp \frac{-2\phi\sqrt{2m\phi_{eff}}}{\hbar e\varepsilon}$$

$$J \sim \exp\left(-\frac{\sqrt{2m\phi eff}}{\hbar} xf\right)$$

$$J \sim \exp\left[-\frac{\sqrt{2m\phi eff}}{e\hbar}\frac{\phi}{\varepsilon}\right]$$

Tunneling increases exponentially with field ε, relatively insensitive to temperature.
All of these processes are important in electronic devices.

116 Band Theory of Solids

116.1 Approach:

- Build qualitative picture of electrons scattering from periodic bumps in potential.
- Introduce Bloch Waves as appropriate solution to Schrödinger Equation in periodic potential.
- Use Kronig-Penney Model to develop band structure.
- Return to coupled modes description to connect with atomic picture.

116.2 Ziman Model:

Free electron waves scatter from dips/ bumps in the potential associated with ions:

$$\psi_R \sim e^{ikx}\ \psi_L \sim e^{-ikx} \psi_\pm = \frac{1}{\sqrt{2}}\left(\psi_R \pm \psi_L\right)$$

These two combinations produce standing waves:

- $\psi_+ \propto e^{ikx} + e^{-ikx} \sim \cos kx$
- $\psi_- \propto e^{ikx} - e^{-ikx} \sim \sin kx$

Consider λ_{e-} such that $n\lambda = 2a$ (strong interference):

$$n\frac{2\pi}{k} = 2a \Rightarrow k = n \cdot {}^{\pi}/_{a}$$

Then $\psi_+^* \psi_+ \propto \cos^2\left(\frac{\pi}{a}x\right)$ →For $n = 1$, max e ions, low energy

$\psi_-^* \psi_- \propto \sin^2\left(\frac{\pi x}{a}\right)$ →minimum e ions, higher energy

For free electrons, $E = \frac{\hbar^2 k^2}{2m}$

Here we have $E = \frac{\hbar^2 k^2}{2m} \pm V_0$ ←Due to ion potentials and scattering when $k \sim \pm n{}^{\pi}/_{a}$

Strong scattering modifies free electron:E vs. k

To make this picture more quantitative, we introduce:

- Correct ψ: Bloch Theorem
- Model $V(x)$: Kronig-Penney Model

116.3 What is the actual ψ?

Consider $-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi = E\psi$ where $V(x) = V(x+a)$.

Bloch proposed that $\psi(x) = u_k(x)\, e^{ikx}$ where $u(x) = u(x+a)$.

(Actually, in 3-D, $\vec{r} \to \vec{r}+\vec{a}$).

This is a "Bloch" wave containing:

- Local, atomic character, $u(x)$
- Plane wave modification e^{ikx}, free electron character (reasonable, also deep)

116.4 "Proof in 1-D":

Model infinite array of atoms by a ring with spacing a and N total atoms, i.e.,

$V(x) = V(x+na)$ where n is an integer

$\psi(x+a) = c\psi(x)$for periodic system where $c = e^{i\alpha}$ (complex constant)

$\psi(x+a) = e^{i\kappa a}\psi(x)$ (suppose)

$\psi(x+Na) = e^{iN\kappa a}\psi(x) = \psi(x)$

$\Rightarrow e^{iN\kappa a} = 1$, or $N\kappa a = 2\pi n$

So $\kappa = \frac{2\pi}{Na} \cdot n$, where $n = 0, \pm 1, \pm 2...$

Quantized values of κ arise from (periodic) boundary conditions, $\kappa \leftrightarrow$wavenumber,

so $\psi(x) = u_\kappa(x)\, e^{i\frac{2\pi}{Na}nx}$ and $u_k(x) = u_k(x+a)$

This 1-D Bloch Wavefunction meets conditions established by potential.

116.5 Kronig-Penney Model:

Inside Wells:

$$\frac{d^2\psi}{dx^2} = -\alpha^2\psi$$

$$,\alpha = \frac{\sqrt{2mE}}{\hbar}$$

Outside Wells:

$$\frac{d^2\psi}{dx^2} = \beta^2\psi$$

$$,\beta = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$$

Impose Bloch Condition on ψ:

$$\psi\left(a < x < a + b\right) = \psi\left(-b < x < 0\right)\ e^{ik(a+b)}$$

We also require the usual continuity of ψ, $\frac{d\psi}{dx}$.
These conditions lead to the following equation:

$$\frac{\beta^2 - \alpha^2}{2\alpha\beta}\ \sinh\ \beta b\ \sin\alpha a + \cosh\ \beta b\ \cos\alpha a = \cos\ k\left(a + b\right)$$

To simplify, we let $b \to 0$, while $V_0 \to \infty$ (δ-function barriers)
with

$$P = \lim_{b\to 0; V_0 \to \infty} \left(\frac{mba}{\hbar^2} V_0\right)$$

as a measure of the barrier (or scattering strength)

$$\Rightarrow P\frac{\sin\alpha a}{\alpha a} + \cos\ \alpha a = \cos\ ka$$

Note that $|\cos ka| \leq 1$.
For what values of α (and $E = \frac{\hbar^2\alpha^2}{2m}$) is $-1 < LHS < 1$?

(1) Plot $P\frac{\sin\alpha a}{\alpha a} + \cos\alpha a$ versus αa .

(2) Solutions exist when the function is between -1 and 1 (first figure).

(3) Plot $E = \frac{\hbar^2}{2m}\alpha^2$ versus ka. (There will be two values of αa that give the same ka.)

(Go to figures.)

116.6 Reduced Zone Scheme:

- Lattice is periodic $\to k$-space is periodic
- E versus k diagram can be reduced to 1st Brillouin Zone from $k = 0$ to π/a.
- $k \pm 2\pi/a \leftrightarrow k$

116.7 More Generally:

- E versus $\vec{\kappa}$: Dispersion depends on direction in the crystal.
- Potential is more complex, especially for compounds.

116.8 Feynman Model:

Interactions between atoms lead to energy level splittings:

- 2 atoms: 2 degenerate levels $\Rightarrow$bonding/ anti-bonding states, 2 levels
- N atoms: N energy levels (per original level)

(See Figure).

116.8.1 Consider 1-D Array:

Assume only nearest neighbor coupling:

$$i\hbar\frac{dw_j}{dt} = E_1 w_j - Aw_{j-1} - Aw_{j+1}$$

$$i\hbar\frac{dw_j}{dt} = E_1 w(x_j) - Aw(x_{j-1}) - Aw(x_{j+1})$$

$$i\hbar\frac{dw_j}{dt} = E_1 w(x_j) - Aw(x_j - a) - Aw(x_j + a)$$

A: coupling coefficient

E_1 : energy in absense of coupling

As before, attempt solution of the form: $w_j = \kappa_j e^{-iEt/\hbar}$

$$E\kappa_j = E_1\kappa_j - A\left(\kappa_{j-1} + \kappa_{j+1}\right)$$

Assume that κ is a function of x , since

$$E\kappa\left(x_j\right) = E_1\kappa\left(x_j\right) - A\left[\kappa\left(x_j + a\right) + \kappa\left(x_j - a\right)\right]$$

Try solution of the form $\kappa\left(x_j\right) = e^{ikx_j}$

$$Ee^{ikx_j} = E_1 e^{ikx_j} - A\left[e^{ik\left(x_j - a\right)} + e^{ik\left(x_j + a\right)}\right]$$

Divide both sides by e^{ikx_j} to get:

$$E = E_1 - A\left[e^{ika} + e^{-ika}\right] = E_1 - 2A\cos ka$$

$$E = E_1 - 2A\cos ka$$

This is an allowed energy band.

Larger A $\rightarrow$Wider Band

Similarly, we expect $E = E_2 - 2B\cos ka$

The generation and analysis of E vs. k is rather straightforward. **(**Physical Interpretation)

116.8.2 In 3-D:

$$i\hbar\frac{dw(x,y,z,t)}{dt} = E_1 w(x_j, y_l, z_m)$$

$$-A_x[w(x_j - a, y_l, z_m) + w(x_j + a, y_l, z_m)]$$

$$-A_y[w(x_j, y_l - b, z_m) + w(x_j, y_l + b, z_m)]$$

$$-A_z[w(x_j, y_l, z_m - c) + w(x_j, y_l, z_m + c)]$$

Try $w\left(x, y, z, t\right) = e^{-iEt/\hbar}\exp\left[i\left(k_x x + k_y y + k_z z\right)\right]$

$$E = E_1 - 2A_x\cos ka - 2A_y\cos k_y b - 2A_z\cos k_z c$$

$$E_{\min} = E_1 - 2\left(A_x + A_y + A_z\right)$$

$$E_{\max} = E_1 + 2\left(A_x + A_y + A_z\right)$$

117 Implications of Band Theory

- Periodic potential modifies parabolic E vs. κ.

  Diagram of Free Electron

- Consider motion analagous to that of free electron.

Classically, in electric field, $F = e\varepsilon = ma = m\frac{dv}{dt}$.

In an Efield, $e\varepsilon = m\frac{dv}{dt}$

Quantum mechanically, we consider electron as wave packet with $v_g = \frac{\partial\omega}{\partial k} = \frac{1}{\hbar}\frac{\partial E}{\partial k}$ as $E = \hbar\omega$.

In a semi-classical picture, then:

$$e\varepsilon = m\frac{dv_g}{dt}$$

But what is $\frac{dv_g}{dt}$?

$$\frac{dv_g}{dt} = \frac{1}{\hbar}\frac{d}{dt}\frac{\partial E}{\partial k} = \frac{1}{\hbar}\frac{\partial^2 E}{\partial k^2}\frac{dk}{dt}$$

But $\frac{dk}{dt} = \frac{1}{\hbar}\frac{dp}{dt} = \frac{1}{\hbar}e\varepsilon\left(F = \frac{dp}{dt}\right)$

(Alternatively, $dE = Fdx$ and $dE = Fv_g dt$.)

So $\frac{dv_g}{dt} = \frac{1}{\hbar^2}\cdot\frac{\partial^2 E}{\partial k^2}e\varepsilon$

Rewriting, $e\varepsilon = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2}\right)^{-1} \frac{dv_g}{dt}$

But we stated that $e\varepsilon = m\frac{dv_g}{dt}$, implying that electron moves as if it had a mass $\hbar^2 \left(\frac{\partial^2 E}{\partial k^2}\right)^{-1}$.

Define effective mass: $m^* = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2}\right)^{-1}$

Check for free electron: $E = \frac{\hbar^2}{2m}k^2$

$\frac{\partial E}{\partial k} = \frac{\hbar^2}{m}k$, $\frac{\partial^2 E}{\partial k^2} = \frac{\hbar^2}{m} \Rightarrow m = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2}\right)^{-1}$

How does this result compare with the Free Electron Model?

To provide a qualitative answer, consider the Feynman Model:

$E = E_1 - 2A\cos ka$

$\frac{\partial E}{\partial k} = 2Aa \sin ka$

$\frac{\partial^2 E}{\partial k^2} = 2Aa^2 \cos ka$

$v_g = \frac{1}{\hbar}\frac{\partial E}{\partial k} = \frac{2Aa}{\hbar}\sin ka$

$m^* = \frac{\hbar^2}{2Aa^2}\sec ka$

Notes:

- m^*depends on k
- Negative around top of band where $k = \pi/a$
- At $k = 0$, $m^* \propto \frac{1}{A}$
- More overlap $\rightarrow$ Wider bands, smaller m^*

Current Flow for Electrons in a Crystal: Free Electrons:

$I = e \cdot \frac{N}{L} \cdot v_g$

$\frac{dI}{dt} = e\frac{dv_g}{dt}\frac{N}{L} = \frac{e^2\varepsilon}{m}\frac{N}{L}$

Not all electrons in crystal can be accelerated.

$N = \int_0^\infty f(E)\, g(E)\, dE = \int_{-\infty}^\infty f(k)\, g(k)\, dk$

From homework, $g(k) = \frac{2L}{\pi}$, for particle in box, so for QM electrons:

$\frac{dI}{dt} = \frac{e}{L}\int_{-\infty}^\infty \frac{dv_g(k)}{dt} f(k)\, g(k)\, dk$

$\frac{dI}{dt} = \frac{e}{L}\frac{e\varepsilon}{\hbar^2}\int_{-\infty}^\infty \frac{d^2E}{dk^2} f(k)\, \frac{L}{\pi} dk$

$\frac{dI}{dt} = \frac{2e^2\varepsilon}{\pi\hbar^2}\int_0^{k_F} \frac{\partial^2 E}{\partial k^2} dk$ for $T = 0$ since $f(k) = 0$ for $k > k_F$

$\frac{dI}{dt} = \frac{2e^2\varepsilon}{\hbar^2\pi}\left(\frac{\partial E}{\partial k}\right)\Big|_{k=k_F}$

But $\frac{dI}{dt} = \frac{e^2E}{m}\frac{N}{L}$

$\Rightarrow N_{Eff} = \frac{2mL}{\pi\hbar^2}\left(\frac{\partial E}{\partial k}\right)_{k=k_F}$ (Those electrons that can be anticipated.)

Observations:

(1) $\left(\frac{\partial E}{\partial k}\right)\Big|_{k=k_F}$ is largest when E_F is in the middle of the band (Half-filled).

$\Rightarrow N_{Eff}$ maximized, high conductivity

(2) $\left(\frac{\partial E}{\partial k}\right)\Big|_{k=k_F} \rightarrow 0$ as $E_F \rightarrow$Band Gap (Filled Band)

$N_{Eff} \rightarrow 0$, Conductivity $\rightarrow 0$ (at$T = 0$)

(3) Number of Electrons per atom will influence the filling of bands, and therfore the type of material.

For Chain of N atoms with P.B.C. (Periodic Boundary Conditions):

- $2N$ states/band (including spin)
- Atoms with even (odd) numbers of electrons will have filled (half-filled) bands.
- Some materials with even numbers are metals, however.
  E.g., Be $(4e^-)$and Mg $(12e^-)$

We need to consider $> 1D$
$\Rightarrow$ use Feynman Model to develop understanding

Feynman Model in 2-D

- $E = E_1 - 2A_x \cos k_x a - 2A_y \cos k_y b$
- Bandgaps occur at $k_x = \pm\pi/a$, $k_y = \pm\pi/b$.
  Example: Choose $a = b$, $A_x = A_y = 0.25$, $E_1 = 1$, $E_g = 0.5$
  - Observe bands "along" $k_y (k_x = 0)$:
    $E_{k_y} = \cos k_y a$
  - Observe bands "along" k_{xy} $\left(k_x = k_y\right)$:
    $E_{k_{xy}} = 2\cos k_{xy} a$

E_F	**Conduction**
0.15	Electrons, Band #1
1.0	Holes along k_y, Electrons along k_{xy}
1.5	Electrons in Band #2, Holes in Band #1

For a single, free electron $\frac{dI}{dt} = \frac{e \cdot e\varepsilon}{m} = e\frac{dv_g}{dt}$.

For nearly free electrons in a solid, $\frac{dI_{Tot}}{dt} = e^2\varepsilon \sum_{occ.states} \frac{1}{m_i^*}$ where $m_i^* = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2}\right)^{-1}\Bigg|_{k=k_i}$.

For a full band: $\frac{dI}{dt} = e^2\varepsilon \sum_i \frac{1}{m_i^*} = 0$.

If certain electron j is missing at top of band:

$$\frac{dI_h}{dt} = e^2\varepsilon \sum_{i\neq j} \frac{1}{m_i^*} = e^2\varepsilon \sum_{i\neq j} \frac{1}{m_i^*} + e^2\varepsilon \frac{1}{m_j^*} - e^2\varepsilon \frac{1}{m_j^*}$$

Note: $e^2\varepsilon \sum_{i\neq j} \frac{1}{m_i^*} + e^2\varepsilon \frac{1}{m_j^*} = 0$

So $\frac{dI_h}{dt} = -\frac{e^2\varepsilon}{m_j^*}$ But $m_j^* < 0$ around top

$\Rightarrow \frac{dI_h}{dt} = \frac{e^2\varepsilon}{\left|m_j^*\right|}$ Remaining electrons act collectively as particle with positive charge.

"Holes" move in opposite direction with opposite charge and therefore make same contribution to current as the electrons.

- Can have **both**, e.g., semi-metals

At finite temperatures, occupation changes according to F-D Distribution, which broadens.
Materials with narrow gaps can become conducting semiconductors.
(Figure #1, Lecture 13)

118 Semiconductors

Classes of Band Structures: (see slides) [1]

(1) Metals: E_F within one band

(2) Semi-Metal: E_F within two overlapping bands

(3) Insulators: E_F within large gap ($>\sim 2eV$)

(4) Semiconductors: E_F within small gap

SemiConductor Sub-Classes: Direct or Indirect Gap [2] (See Examples of GaAs, AlAs)
Maxima and minima depend on both structure and composition [3]

Energies of Electrons Electrons move as free particles with effective mass; near band bottom; taking $E_V = 0$ and $C\kappa - 0$ for reference.

$$E_C \simeq E_g + \frac{\hbar^2\kappa^2}{2m_n^*} \; E_V \simeq -\frac{\hbar^2\kappa^2}{2m_p^*}$$

Negative m^*, + charge $\longleftrightarrow$ Positive m^*, - charge

Intrinsic Semiconductors At finite $T > 0$, electrons occupy conduction band, leaving holes in the valence band.

$$n = p = n_i$$

n = density of free electrons
p = density of free holes
n_i= intrinisic carrier concentration

$$n = \int g_c(E) f(E) dE \text{ conduction band}$$
$$p = \int g_v(E) [1 - f(E)] dE \text{ valence band}$$

From Lecture #9, the D.O.S. (density of states) for free electrons:

$$g(E) = \frac{dn}{dE} = \frac{1}{2\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{3/2}\sqrt{E}$$

For bottom of conduction and top of valence band:

$$g_c(E) = \frac{1}{2\pi^2}\left(\frac{2m_c^*}{\hbar^2}\right)^{3/2}\sqrt{E - E_c} \; g_v(E) = \frac{1}{2\pi^2}\left(\frac{2m_v^*}{\hbar^2}\right)^{3/2}\sqrt{E_V - E}$$

These states become occupied according to $f(E)$:

$$f(E) = \frac{1}{e^{(E-E_F)/k_BT}+1} \cong e^{-(E-E_F)/k_BT} \text{For } E_F \text{ in gap and } E - E_F \gg \kappa T$$

(Maxwell-Boltzmann)

Examine product $g_c(E) f(E)$ graphically to inform the integration:

(1) $g_c(E)$ is zero in gap.

(2) $f(E)$ is zero at top of band $\rightarrow$ only bottom will matter.

$$\text{Hence, } n \approx \frac{(2m_c^*)^{3/2}}{2\pi^2\hbar^3}\int_{E_c}^{\infty}(E - E_c)^{1/2} e^{-(E-E_F)/\kappa_BT} dE$$
$$\text{Let } x = \frac{E-E_c}{k_BT} \; dE = k_BT dx \; \sqrt{E - E_c} = x^{1/2}\sqrt{k_BT}$$
$$n = \frac{(2m_c^*)^{3/2}}{2\pi^2\hbar^3}(k_BT)^{3/2} e^{-(E_C-E_F)/k_BT}\int_0^{\infty} x^{1/2} e^{-x} dx$$
$$n = N_C e^{-(E_C-E_F)/k_BT}, \text{ where } N_C = 2\left(\frac{m_c^* k_BT}{2\pi\hbar^2}\right)^{3/2}$$
$$p = N_V e^{-(E_F-E_V)/k_BT} \; N_V = 2\left(\frac{m_v^* k_BT}{2\pi\hbar^2}\right)^{3/2}$$

- As E_F increases, n increases and p decreases.

- Depends on m^*

  $\Rightarrow$ develop expression for carrier concentration for intrinsic semiconductors

Need expression for E:

$$n = p = n_i \text{ Let } E_F = E_{F,i}$$
$$N_C e^{-(E_C - E_{F,i})/k_B T} = N_V e^{-(E_{F,i} - E_V)/k_B T}$$
$$e^{2E_{F,i}/k_B T} = \frac{N_V}{N_c} e^{(E_C + E_V)/k_B T}$$
$$\frac{2E_{F,i}}{k_B T} = \frac{E_C + E_V}{k_B T} + \frac{3}{2} \ln \frac{m_v^*}{m_c^*}$$
$$E_{F,i} = \frac{1}{2}(E_V + E_C) + \frac{3}{4} k_B T \ln \frac{m_v^*}{m_c^*}$$
$$E_{F,i} = E_V + \frac{E_G}{2} + \frac{3}{4} k_B T \ln \frac{m_v^*}{m_c^*}$$

Fermi level is approximately mid-gap for intrinsic semiconductors.

$$n = p = n_i \Rightarrow np = n_i^2$$
$$n_i^2 = N_C e^{-(E_C - E_F)/k_B T} N_V e^{-(E_F - E_V)/k_B T}$$
$$n_i^2 = N_C N_V e^{-(E_C - E_V)/k_B T} \; E_C - E_V = E_G$$
$$\therefore n_i = \sqrt{N_C N_V} e^{-E_G/2k_B T}$$

At given temperature, smaller gap implies more carriers and a higher intrinsic carrier concentration.

Extrinsic Semiconductors Impurities (dopants) added to control n, p

"**n -type**": negative charge carriers

- Add atom with extra valence electron: donor
- e.g. P, As, Sb (Group V) to Si, Ge (IV).

Donor Level Energy (estimate) Electron + ionized donor is like H atom but in material with $E \neq \varepsilon_0$

$$E_C - E_D = \left(\frac{e^2}{4\pi\varepsilon}\right)^2 \frac{m^*}{2\hbar^2} = \left[\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \frac{m_0}{2\hbar^2}\right] \frac{m^*}{m_0} \left(\frac{\varepsilon_0}{\varepsilon}\right)^2$$
$$= \frac{m^*}{m_0} \left(\frac{\varepsilon_0}{\varepsilon}\right)^2 (13.6eV)$$

Estimate for Si: $\varepsilon \simeq 12\varepsilon_0$ $m^* = 0.5 m_0$

$\Rightarrow E_C - E_D \simeq 0.05eV$ Donor Binding Energy

$\Rightarrow E_C - E_D \simeq 2k_B T$ around $T = 300K$

($k_B T \simeq \frac{1}{2}(E_i - E_D)$ around $T = 300K$)

$\Rightarrow$ Most donors are ionized $n \simeq N_D$ (for $N_D \gg n_i$)

We showed that $n = N_C e^{-(E_C - E_{F,i})/k_B T}$

In thermal equilibrium, for $N_D \gg n_i$, but not too large:

$$n = N_C e^{-(E_C - E_F)/k_B T} \approx N_D \; (E_F > E_{F,i})$$

By rewriting in terms of E_i, N_i, one can show, however, that:

$$n = n_i e^{(E_F - E_{F,i})/k_B T}$$
$$p = n_i e^{(E_{F,i} - E_F)/k_B T}$$

$\Rightarrow np = n_i^2$ can increase the number of electrons, but there will be fewer holes.

(Figure 3-16)

Law of Mass Action "**p-type**": positive charge carriers

- Add atom with 1 less valence electron: acceptor
- e.g. Add B, Al, Ga (Group III) to Si, Ge (Group IV)
- $p \simeq N_A$ for $N_A \gg n_i$ at room temperature

Observations

- At very low temperatures, not all donors will be ionized. (Ionization Regime)
- At very high temperatures, n_i becomes appreciable due to thermal activation across gap. (Ionization Regime)

- More generally, both donors and acceptors may be present. **The number of free carriers is given by**:

  (1) Space-charge neutrality: $p + N_D^+ = n + N_A^-$ (in equilibrium)

  (2) Law of Mass Action: $np = n_i^2$

  At room temperatures, $N_D^+ \simeq N_D$, $N_D^- \simeq N_A$
  If, e.g., $N_D > N_A$, then $n - p = N_D - N_A$
  Compensation: We have a reduced net electron concentration.

119 Current Flow in Semiconductors

(S+W, Chapter 8 and A+M, Chapter 1)

Equilibrium

- Carriers move quickly $(\sim \nu_F)$, but no **net** motion of charge.
- Random scattering from phonons and impurities: this randomizes momentum.
  - ℓ: mean free path between collisions
  - $\tau = \frac{\ell}{\nu_F}$ average time between collisions (inverse rate)

In Applied Field $\vec{\varepsilon}$ Electron feels a force $\vec{F} = \frac{d\vec{p_\varepsilon}}{dt} = -e\vec{\varepsilon}$, accelerates, but collisions randomize $\vec{p}$.

- i.e. $\Delta\vec{p}_{col} = -\left\langle \vec{p} \right\rangle$

For an average collision, $\frac{d\vec{p}_{col}}{dt} = \frac{-\left\langle \vec{p} \right\rangle}{\tau}$.

In steady state, $\frac{d\vec{p_e}}{dt} + \frac{d\vec{p}_{col}}{dt} = 0$

$$\Rightarrow -e\vec{\varepsilon} = \frac{\left\langle \vec{p} \right\rangle}{\tau}, \nu$$

$$-e\tau\vec{\varepsilon} = m_n^* \left\langle \vec{\nu} \right\rangle$$

$\langle \nu \rangle = -\frac{e\tau}{m_n^*}\vec{\varepsilon}$ where $\left\langle \vec{\nu} \right\rangle \equiv \nu_d$ Drift Velocity

$\left\langle \vec{\nu} \right\rangle = -\frac{e\tau}{m_n^*}\vec{\varepsilon}$ where $\left\langle \vec{\nu} \right\rangle \equiv \vec{\nu}_d$ Drift Velocity

$\vec{J} = -e \cdot n \cdot \vec{\nu}_d = \frac{ne^2\tau}{m_n^*}\vec{\varepsilon}$, so we have Ohm's Law

$\vec{J} = \sigma\vec{E}$, where $\sigma = \frac{ne^2\tau}{m_n^*} = ne\mu_n$ is the conductivity.

(Note that σ is a tensor $\overleftarrow{\sigma}$.)

Also, note that $\vec{\nu}_d = -\mu n\vec{\varepsilon}$ (simple interpretation).

$\vec{J} = ne\mu n\vec{\varepsilon}$ for electrons.

Overall, the current due to drift is:

$\vec{J} = e\left(n\mu_n + p\mu_p\right)\vec{\varepsilon}$ when $\vec{\nu}_d \propto \vec{\varepsilon}$.

You may be more familiar with $V = IR$, where $R = p\frac{\ell}{A} = \frac{1}{\sigma}\frac{\ell}{A}$.

- ℓ refers to length; A refers to the cross-sectional area

Observations:
Conductivity is proportional to τ, inversely proportional to m^*.
What are the origins of scattering?

Scattering Mechanisms

(1) Phonons (Lattice Scattering)
$\tau_{Ph} \propto T^{-3/2}$

(2) Impurities/ Defects
$\tau_{Imp} \propto T^{3/2}$ (less time near imp)

(3) Charge Carriers
(can neglect for moderately doped semiconductors)

Scattering Rate:
$$R_{Tot} = R_{Ph} + R_{Imp}$$
$$\frac{1}{\tau} = \frac{1}{\tau_{Ph}} + \frac{1}{\tau_{Imp}}$$
$$\left(\frac{1}{\mu} = \frac{1}{\mu_{Ph}} + \frac{1}{\mu_{Imp}}\right)$$

Observations:

- Doping increases σ by increasing n, p, but impurity scattering also increases.
- At high fields, additional scattering mechanisms lead to non-linear changes in $\vec{v}\,d$ with $\vec{\varepsilon}$.
  - Si, Ge, velocity saturation (Optical Phonon Scattering)
  - GaAs NDR for e^- (Inter-valley Scattering)

Hall Effect (skipped in 2012)
Experiment used to determine n, p, μ_n, μ_p.
$$\vec{F} = e\left(\vec{\varepsilon} + \vec{v}_x \vec{B}\right) \text{ for holes.}$$
Assume $\vec{v} = v_x\hat{x}\vec{B} = B_z\hat{z}$.
$$F_y = e\left(\varepsilon_y - v_x B_z\right)$$
Holes are deflected in $-\hat{y}$ direction until field builds up to oppose the flow.
$\Longrightarrow \varepsilon_y = v_x B_z$ in steady state (**Hall Effect**)
(The Hall voltage is $\varepsilon_y w$) $\{J_x = e \cdot p \cdot v_{dx}\}$
$\varepsilon_y = \frac{J_x}{ep} B_z = R_H J_x B_z$ Note: $R_H = \frac{1}{ep}$
$p = \frac{J_x B_z}{e\varepsilon_y} = \frac{\frac{I_x}{wt} \cdot B_z}{qV/w} = \frac{I_x B_z}{qtV_H}$
Also: $R = \frac{V_x}{I_x} = \frac{1}{\sigma} \cdot \frac{L}{\omega t}$
$\Longrightarrow \sigma = \frac{I_x}{V_x} \cdot \frac{L}{wt} = pe\mu_p$
$\Rightarrow \mu_p = \frac{\sigma}{ep} = R_H\sigma$
We can therefore determine the majority carrier type and mobility.

Additional Transport Processes: Diffusion A gradient in free carrier concentration leads to diffusion. (See figure)

Fick's First Law: (1-D) $\phi = -D\frac{dc}{dx}$

- ϕ refers to flux $\frac{1}{cm^2-s}$
- $-D$ is the diffusion coefficient $^{cm^2}/_s$
- $\frac{dc}{dx}$ refers to the concentration cm^{-3}

These are the diffusion currents:

- $J_n = -(-e)\,D_n\frac{dn}{dx} = eD_n\frac{dn}{dx}$
- $J_p = -(e)\,D_p\frac{dp}{dx} = -eD_p\frac{dp}{dx}$

In the presence of $\vec{\varepsilon} \neq 0$ and $\frac{dp}{dx}, \frac{dn}{dx} \neq 0$:

$$J_n(x) = e\mu_n n(x)\,\varepsilon(x)\,(drift) + eD_n\frac{dn(x)}{dx}(diffusion)$$
$$J_p(x) = e\mu_p p(x)\,\varepsilon(x)\,(drift) + eD_p\frac{dp(x)}{dx}(diffusion)$$
$$J_{Tot}(x) = J_n(x) + J_p(x)$$

Whereas majority carriers dominate current flow due to drift ($n \gg p$ or $p \gg n$), minority carrier diffusion currents can be substantial ($\frac{dp}{dx}$, $\frac{dn}{dx}$).

Consider energy band diagram in presence of applied field. (See att version.)

The energy of an electron is $E(x) = -eV(x)$ and $\varepsilon(x) = -\frac{dV}{dx}$

So $\varepsilon(x) = -\frac{d}{dx}\left(\frac{E_i(x)}{-e}\right) = \frac{1}{e}\frac{dE_i}{dx}$

Electrons drift "downhill."

Holes drift "uphill."

Variations in **concentration** also reduce fields. Consider holes:

$J_p(x) = e\mu_p p(x)\,\varepsilon(x) - eD_p\frac{dp(x)}{dx} = 0$ No current when $\frac{dp}{dx} \neq 0$

$\Longrightarrow \varepsilon(x) = \frac{D_p}{\mu_p}\frac{1}{p(x)}\frac{dp}{dx}$ {Q: What $p(x)$ would produce this? See homework.}

$p = n_i e^{(E_i-E_F)/\kappa T} \Rightarrow \frac{dp}{dx} = n_i e^{(E_i-E_F)/\kappa T}\cdot\frac{1}{\kappa T}\left(\frac{dE_i}{dx} - \frac{dE_F}{dx}\right)$

Note that $e\varepsilon$ refers to $\frac{dE_i}{dx}$ and $\frac{dE_F}{dx}$ is zero in the above equation.

$\varepsilon(x) = \frac{D_p}{\mu_p}\frac{1}{\kappa_B T}\cdot e\varepsilon$

$\therefore \frac{D_p}{\mu_p} = \frac{\kappa_B T}{e}$ Einstein Relation (also holds for electrons)

Concentration Gradients → "Built-in" Fields

Inside a semiconductor, $\varepsilon(x) = \frac{1}{e}\frac{dE_i}{dx}$

Slope in bands indicates field

Fields result from:

(1) Applied voltage: $\varepsilon = {}^{V}/_{L}$

(2) Carrier concentration gradients

Consider holes in the absense of current:

$J_p(x) = e\mu_p p(x)\,\varepsilon(x) - eD_p\frac{dp}{dx} = 0$ $\frac{dp}{dx} \neq 0$

$\varepsilon(x) = \frac{D_p}{\mu_p}\frac{1}{p(x)}\frac{dp}{dx} \Longrightarrow \varepsilon \neq 0$

We can relate drift to diffusion:

$p = n_i e^{(E_i-E_F)/\kappa T} \Rightarrow \frac{dp}{dx} = n_i e^{(E_i-E_F)/\kappa T}\frac{1}{\kappa T}\left(\frac{dE_i}{dx} - \frac{dE_F}{dx}\right)$

Note: $\frac{dE_i}{dx} = e\varepsilon$ and $\frac{dE_F}{dx} = 0$

Substituting above,

$\varepsilon(x) = \frac{D_p}{\mu_p}\frac{1}{\kappa_B T}\frac{dE_i}{dx} = \frac{D_p}{\mu_p}\frac{1}{\kappa_B T}e\varepsilon$

So $\left[\frac{D_p}{\mu_p} = \frac{\kappa_B T}{e}\right]$ Einsteen Relation (also for e^-)

This assumes equilibrium.

HW Problem #8.4 (in class, 2010, 2011)

(1) Consider donor doping profile:
$n \simeq N_d = N_0 e^{-ax}$ $(Nd \gg n_i)$

a What is the field?
In equilibrium, $J(x) = 0$ and $e\mu_n n(x)\varepsilon(x) + eD_n \frac{dn(x)}{dx} = 0$
$\varepsilon(x) = -\frac{D_n}{\mu_n}\frac{dn}{dx}\frac{1}{n(x)}$ Use Einstein Relation.
$\varepsilon(x) = -\frac{\kappa T}{e}\frac{\left(-aN_0 e^{-ax}\right)}{N_0 e^{-ax}} = \frac{\kappa T}{e}a$
Field depends on termperature and rate of dopant decrease (juction abruptness).

b $a = 1\mu m^{-1} \qquad T = 300\kappa$
$\varepsilon = \frac{\kappa T}{a} = \frac{\left(1.38\times10^{-23} J/\kappa\right)(300\kappa)(1eV)}{e1.602\times eo^{-19} J}\frac{1}{10^{-4} cm}$
$= 258 V/cm$

Additional Homework Solutions:

(1) Parts (a) and (b):

a $F(E_A) = \frac{1}{e^{+(E_A - E_F)/\kappa T}+1}$

b $(-v)\, n = N_c e^{-(E_i - E_F)/\kappa_B T}$
$p = N_v e^{-(E_F - E_V)/\kappa_B T}$
$p + N_D^+ = n + N_A^- N_D = 0$

(2) $\sigma_n(T) = n(T_i)\, e\mu_n(T)$
$= n(T)\frac{e^2}{m_n^*}\tau(T)$
$\frac{\sigma(2T_1)}{\sigma(T_1)} = \frac{n(2T_1)}{n(T_1)} \cdot \frac{T(2T_1)}{T(T_1)}$
$\frac{1}{T} = \frac{1}{T_{PH}} + \frac{1}{T_{IM}}$
$n \propto T^2$
$T_{imp} = AT^{3/2}$
$T_{PH} = BT^{-3/2}$
Also: $T_{imp}(T_1) = T_{IM}(T_2)$ (Connects A with B)

$$\frac{I}{A} = J = \sigma\varepsilon = \sigma \cdot \frac{V}{L}$$
$$I = \sigma \cdot A \cdot \frac{V}{L}$$
$$I = N_d$$
$$\nu_d = \mu\varepsilon$$
$$\sigma = n \cdot e \cdot \mu_n \ (n \text{ is an estimate})$$
$$\sigma \simeq N_d \cdot e\mu_n$$
$$t = \frac{d}{\nu}$$
$$t = \frac{i}{\nu_d} = \frac{L}{\mu\varepsilon} = \frac{LL}{\mu V} = \frac{L^2}{\mu V}$$

(1) $J_{sat} = ne \cdot \vec{\nu}_{sat}$
For $\nu \propto \varepsilon$, $J = \sigma\varepsilon$
For high fields, $\nu \to \nu_{sat}$
$J_{sat} = n \cdot e \cdot \nu_{sat}$
$J = \frac{I}{A}; \sigma = n \cdot e \cdot \mu$
n is approximate; μ should be looked up

a $T(T_1) = \frac{T_{I1} T_{P1}}{T_{I1} + T_{PI}}$

$T_F = AT^{3/2}$, $T_{PH} = BT^{-3/2}$, $AT_1^{3/2} = BT_1^{-3/2}$

$\Rightarrow B = AT_1^3$

$\frac{1}{T_1(T_1)} = \frac{1}{AT_1^{3/2}} + \frac{1}{AT_1^3 \cdot T_1^{-3/2}} \frac{1}{AT^{3/2}} \left(\frac{1}{2^{3/2}} + \frac{1}{2^{-3/2}} \right)$

$T_1 = \frac{A}{2} T_1^{3/2} T_2 = \frac{A}{2} T_2^{3/2} = \frac{A}{2} \cdot 2^{3/2} T_1^{3/2} = A\sqrt{2} T_1^{3/2}$

$T(T_2) = AT_1^{3/2}$

$\frac{T(T_2)}{T(T_1)} = \frac{\sqrt{2}}{2}$

$\sigma = 4 \cdot \frac{\sqrt{2}}{2} = 2\sqrt{2}$

$\left(2^{-3/2} + 2^{3/2} \right)^{-1} = 0.314$

120 351-1 Problems

(1) Newton's Laws Can be derived from Hamilton's equations.

a Identify the Hamiltonian from conservation of energy using only momentum p (for Kinetic Energy) and position x (for potential Energy):

$$H = KE + PE$$

Use the Hamiltonian for a particle in a 1-D quadratic potential like a mass on a spring. What is KE, PE?

b Hamiltion's equations are

$\frac{d\rho}{dt} = -\frac{\partial H}{\partial x}$
$\frac{dx}{dt} = \frac{\partial H}{\partial \rho}$

Show these give Newton's laws of motion for the mass on a spring.

c Derive the 1-D differential equation of motion from Hamilton equation. For a particle of total energy E and spring constant k,$x = 0$ at $t = 0$, what is the equation of motion$x(t)$. what is $p(t)$. show the region of Phase Space (p vs$x$) that describes the particle throughout its motion.

d Use the equipartition theorem (where <A> is time average of A, and i is index for each spatial dimension):

$$\left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle = \left\langle x_i \frac{\partial H}{\partial x_i} \right\rangle = k_B T$$

to derive the relationship between thermal velocity and temperature $\frac{1}{2} m v_{th}^2 = \frac{3}{2} k_B T$

e Derive the Dulong-Petit law for atoms in a 3-D potential from the equipartition theorem:

$$\langle E \rangle = 3 k_B T$$

f What is the heat capacity $C = dE/dT$ in this case?

(2) Problem 1.1 from Solymar and Walsh:

A 10 mm cube of germanium passes a current of 6.4 mA when 10 mV is applied between two of its parallel faces. Calculate the resistivity of the sample. Assuming that the charge carriers are electrons that have a mobility of 0.39 $m^2 V^{-1} s^{-1}$, calculate the density of carriers. What is their collision time if the electron's effective mass in germanium is 0.12 m_e where m_e is the free electron mass?

(3) Give a one line description of each of these experiments and their significance to modern physics: Photoelectric Effect, Compton Effect, Black Body Radiation, Rutherford Backscattering (Bohr model), Electron Diffraction, Atomic spectra.

Classical: https://www.youtube.com/watch?v=yXsHflXB7QM

Bohr Atom: https://www.youtube.com/watch?v=ydPzEZTd-98

Wave Particle Duality - photoelectric effect: https://www.youtube.com/watch?v=frNLtEm1glg

Schrödinger waves: https://www.youtube.com/watch?v=C8XGIYz1PCw

Probability interpretation: https://www.youtube.com/watch?v=p7xIKoBdViY

Compton Effect: https://www.youtube.com/watch?v=0Y648TNGAIo

(4) Problem 2.1 from Solymar and Walsh.

Find the de Broglie wavelength of the following particles, ignoring relativistic effects:

(i) an electron in a semiconductor having average thermal velocity at $T = 300K$ and an effective mass of $m_e^* = a m_o$,

(ii) a helium atom having thermal energy at $T = 300K$

(iii) an α particle (He^4 nucleus) of kinetic energy 10 MeV.

Hint: See question 1. For a gas of non-interacting particles, $\frac{1}{2} m v_{th}^2 = \frac{3}{2} k_B T$.

(5) A particle of mass, m, is confined to a 1-D region $0 \leq x \leq L$. In class, we derived the following stationary state wavefunctions and energies for this 1-D infinite square well potential:

$$\psi_n(x) = Asin\left(\frac{\pi n}{L}s\right)$$

a Normalize the wave function to find the value of A.

b Find the Energy of these stationary states

Assuming that the initial normalized wavefunction of this particle at$t = 0$ is:

$$\Psi(x, t = 0) = \sqrt{\frac{8}{5L}}\left[1 + cos\left(\frac{\pi}{L}x\right)\right] sin\left(\frac{\pi}{L}x\right)$$

a Derive an expression for the wavefunction $\Psi(x, T)$, at all later times $t > 0$.

b Show that the probability of finding the particle in the left half of the box (i.e., in the region $0 \leq x \leq L/2$) at time $t > 0$ is: $\frac{1}{2} + \frac{16}{15\pi}cos\left(\frac{3\pi^2 ht}{2mL^2}\right)$

(6) Assuming that Ψ(x,t) is a solution of the 1-D Schrödinger Equation, the current density is defined as:

$$J(x) = -\frac{i\hbar e}{2m}\left[\Psi^*\frac{\partial\Psi}{\partial x} - \Psi\frac{\partial\Psi^*}{\partial x}\right]$$

In this problem, consider the potential barrier of height (V) and width (d) that is depicted in Fig. 3.3 of Solymar and Walsh. Assume that the electron energy (E) is less than V.

a By applying suitable boundary conditions and your knowledge of quantum mechanics, develop a system of equations that could be solved to determine the transmitted current through the barrier (J_t) in terms of the incident current on the barrier (J_i).

b By solving your system of equations from part (a), show that:

$$\frac{J_t}{J_t} = \frac{1}{1 + \frac{V^2}{4(V-E)E}sinh^2(\gamma d)}$$

where $\gamma = \frac{\sqrt{2m(V-E)}}{\hbar}$ and $sinh(\gamma d) = \frac{1}{2}(e^{\gamma d} - e^{-\gamma d})$

c In the limit where $\gamma d \gg 1$, show that: $\frac{J_t}{J_t} = \frac{16(V-E)E}{V^2}e^{-2\gamma d}$

d The exponential dependence of the "tunneling" current on distance is utilized for atomic resolution imaging of conductive surfaces with the scanning tunneling microscope (STM). Conservatively assume that the STM can detect changes in the tunneling current of 1%. Under typical tunneling conditions (e.g., $V - E$ 1eV, d ~ 10 Å), estimate the vertical spatial resolution of the STM. Hint: The answer can be expressed in picometers (10^{-12} m)!

(7) Derive the solution of the 2-D particle in a box (particle is constrained in both x, and y directions; $V \to \infty$ at the boundaries).

a Solve for the energies and wave functions of the ground and first excited states.

b What is the degeneracy of the first excited state?

c Plot the probability distributions 2 ψ of the first excited states as surface plots.

d Optional: plot probability distributions of the ground state, first excited states, and 2nd excited states, and comment on the evolution.

(8) Consider the infinite spherical well: $V(r) = 0$ if $r \leq a$, $V(r) = \infty$ if $r > a$.

a For $l = 0$, determine the allowed energies (E_n).

b For $l = 0$, show that the corresponding wavefunctions are:

$$\Psi_n(r) = \frac{1}{\sqrt{2\pi a}} \frac{sin(n\pi r/a)}{r}$$

In class, we worked through the Schrödinger equation in spherical coordinates for spherically symmetric potentials by breaking the solution into radial and angular functions. You should be able to solve the radial equation for the conditions given here.

(9) Consider the infinite spherical well: $V(r) = 0$ if $r \leq a$, $V(r) = V_o$ if $r > a$.

a For $l = 0$, show that there is no bound state if: $V_o a^2 < \frac{\pi^2 \hbar^2}{8m}$. This can be shown without resorting to numerical computation.

b Given $V_o = \frac{2\hbar\pi^2}{ma^2}$, find the energies of the two bound states by graphing in MATLAB or Excel. You can also use MATLAB to check your solution by solving the transcendental equation directly.

This problem is analogous to the finite square well problem solved in 3.8 of Solymar and Walsh, but in the spherical coordinate system. Through the appropriate application of boundary conditions, you should arrive at a transcendental equation whose argument can be analyzed to establish the condition for the existence of bound states.

(10) Suppose that the nucleus of a hydrogen atom is located at a distance d from a two-dimensional infinite potential wall which, of course, tends to distort the hydrogen atom. As d approaches zero, determine the following items:

a The ground state wavefunction.

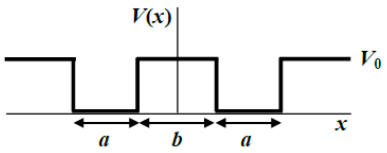
b The degeneracy of the first excited state (ignore degeneracy due to spin).

c The wavelength of light that is emitted upon transition between the first excited state and the ground state (express your answer in nanometers).

With the exception of part (c), this problem does not involve mathematical calculation. The proper choice of coordinate system can make the relationship between these solutions and the usual hydrogen atom solutions clear. Bear in mind that the ground state refers to the lowest energy state that exists.

(11) Consider a double finite potential well in one dimension. Suppose that the depth V_0 and the width a are fixed such that the following equation is obeyed:

$$\beta^2 = \frac{V_0 m a^2}{2\hbar^2} = 4\pi^2$$



(12) Qualitatively sketch the ground state wavefunction $\psi_1$and the first excited state wavefunction ψ_2 for: (i) b =0, (ii) $b = a/2$, and (iii) $b \gg a$.

a For $b = 0$, show that the ground state energy (E_1) is given by: $E_1 = \frac{\hbar^2 x_1^2}{2ma^2}$, where x_1 is the solution of the following equation: $tan x_1 = \sqrt{\left(\frac{2\beta}{x_1}\right)^2 - 1}$. CalculateE_1 in units of $\frac{\hbar}{ma^2}$.

b For$b = 0$, show that the first excited state energy (E_2) is given by: $E_2 = \frac{\hbar^2 x_2^2}{2ma^2}$ where x_2 is the solution of the following equation: $-cot x_2 = \sqrt{\left(\frac{2\beta}{x_2}\right)^2 - 1}$. Calculate E_2 in units of $\frac{\hbar^2}{ma^2}$.

c For $b \gg a$, estimate E_1 and E_2 in units of $\frac{\hbar^2}{ma^2}$.

d Use the MATLAB code derived from Garcia, R., Zozulya, A. & Stickney, J. MATLAB codes for teaching quantum physics: Part 1. *arXiv* **physics.ed-ph**, (2007). http://arxiv.org/pdf/0704.1622.pdf
Consult the original publication for background on the code, which is reproduced below.

i. What is the Heaviside function? How is it used in this code (for what purpose)?

ii. Using the MATLAB code, generate the plots you sketched in part (a) (ground and first excited states for b=0, b=a/2, and b>>a). You should have three graphs with two curves on each. Label the graphs and the wave functions.

iii. Generate 2 more plots with intermediate barrier widths, and sketch the trends in E_1 and E_2 as a function of barrier width (you should have 5 different widths).

Note: in the code, you can change the potential profile quite easily (e.g. the quadratic harmonic oscillator potential or single square well). Parameters can be varied to develop insight into how the wavefunctions vary with the potentials.

e Provide a physical explanation for the variation of E_1 with b that you observed in part (e).

f The double well is a primitive one dimensional model for the potential experienced by an electron in a diatomic molecule (the two wells represent the attractive force of the nuclei). If the nuclei are free to move, they will adopt the configuration of minimum energy. In view of your conclusions in (b), does the electron in the ground state (E_1) tend to draw the nuclei together or push them apart? What about E_2? Provide a physical reason for these behaviors, considering your answer to the previous question.

```
% ***************************************************************
% Program 4: Find several lowest eigenmodes V(x) and
% eigenenergies E of 1D Schrodinger equation
% -1/2*hbar^2/m(d2/dx2)V(x) + U(x)V(x) = EV(x)
% for arbitrary potentials U(x) %
   ***************************************************************
% Parameters for solving problem in the interval -L < x < L
% PARAMETERS:
L = 5;% Interval Length
N = 1000; % No of points
x = linspace(-L,L,N)';% Coordinate vector
dx = x(2) -x(1); % Coordinate step

% POTENTIAL, choose one or make your own
U = 1/2*100*x.^(2); % quadratic harmonic oscillator potential
%U = 1/2*x.^(4); % quartic potential

% Finite square well of width 2w and depth given
%w = L/50;
%U = -500*(heaviside(x+w)-heaviside(x-w));

% Two finite square wells of width 2w and distance 2a apart
%w = L/50; a=3*w;
%U = -200*(heaviside(x+w-a) -heaviside(x-w-a) ...
% + heaviside(x+w+a) -heaviside(x-w+a));

% Three-point finite-difference representation of Laplacian
% using sparse matrices, where you save memory by only
% storing non-zero matrix elements
e = ones(N,1); Lap = spdiags([e -2*e e],[-1 0 1],N,N)/dx^2;

% Total Hamiltonian hbar = 1; m = 1;
% constants for Hamiltonian H = -1/2*(hbar^2/m)*Lap + spdiags(U,0,N,N);

% Find lowest nmodes eigenvectors and eigenvalues of sparse matrix
nmodes = 3; options.disp = 0;
[V,E] = eigs(H,nmodes,'sa',options);% find eigs
[E,ind] = sort(diag(E));% convert E to vector and sort low to high
V = V(:,ind); % rearrange corresponding eigenvectors

% Generate plot of lowest energy eigenvectors V(x) and U(x)
Usc = U*max(abs(V(:)))/max(abs(U)); % rescale U for plotting
plot(x,V,x,Usc,'- k'); % plot V(x) and rescaled U(x)

% Add legend showing Energy of plotted V(x)
```

```
lgnd_str = [repmat('E = ',nmodes,1),num2str(E)];
legend(lgnd_str) % place lengend string on plot
```

(13) Consider a cesium chloride crystal where the potential energy per formula unit is:

$$E(r) = \frac{A}{r^n} - \frac{Me^2}{4\pi\epsilon_0 r}$$

where A is a constant, $n > 1$, r is interionic distance, and M is the Madelung constant.

a Express the binding energy (E_0) in terms of M, n, r_0, and fundamental constants. *Hint: first express A in terms of these constants by considering the equilibrium condition.*

b The cesium chloride crystal structure consists of cations located on a simple cubic lattice (lattice constant = a) with an anion located at the center of the cube. What is the volume per formula unit (v) in terms of r?

c From thermodynamics, the bulk modulus (B) is known to be: $B = v\frac{\partial}{\partial v}\left(\frac{\partial E}{\partial v}\right)$ .Show that B_o is of the form $B_0 = \frac{C}{r_o}\frac{d^2E}{dr^2}|_{r=r_o}$, and find C. *Hint: use your result from the previous problem to rewrite the derivatives in terms of r.*

d Using the result from part (c), express the equilibrium bulk modulus (B_0) in terms of M, n, r_0, and fundamental constants.

e The experimentally determined values of B_0 and r_0 for CsCl are 19.8 GPa and 3.571 Å respectively. Calculate E_0 for CsCl in eV. Note: M = 1.7627 for CsCl.

Part IX
351-2: Solid State Physics

121 Catalog Description (351-1,2)

Quantum mechanics; applications to materials and engineering. Band structures and cohesive energy; thermal behavior; electrical conduction; semiconductors; amorphous semiconductors; magnetic behavior of materials; liquid crystals. Lectures, laboratory, problem solving. Prerequisites: GEN ENG 205 4 or equivalent; PHYSICS 135 2,3.

122 Course Outcomes

123 351-2: Solid State Physics

At the conclusion of 351-2 students will be able to:

(1) Apply the basic concepts of semiconductor physics to semiconductor devices – p-n junctions, solar cells, light-emitting diodes, lasers, and transistors – and use this for understanding/predicting device characteristics.

(2) Describe the nature of lattice vibrations in solids and calculate their effect on materials properties.

(3) Apply an understanding of the basic polarization mechanisms to predict and/or understand dielectric properties.

(4) Articulate the basic principles of magnetic materials and be predict basic magnetic behaviors.

(5) Articulate the basic principles of superconductivity.

(6) Understand and interpret data from a wide range of electrical, thermal, optical, and magnetic probes of materials properties.

124 Principles of Semiconductor Devices

Recall that the conductivity of semiconductors is given by

$$\sigma = ne\mu + pe\mu \tag{124.1}$$

where n is the number of electrons, e is the electronic charge, μ is the mobility in $\mathrm{cm}^2/\mathrm{V/s}$. The following trends for conductivity versus temperature are noted for metals, semiconductors, and insulators:

(1) Metals: n and p are constant with temperature. Mobility is related to temperature as $\mu \propto T^{-a}$, where T is temperature and a is a constant. Conductivity decreases with increasing temperature. The resistivity ρ ($\rho = 1/\sigma$) can be written as a sum of contributing factors using Matthiessen's rule as

$$\rho = \rho_0 + \rho(T)$$

where ρ_0 is a constant and $\rho(T)$ is the temperature dependent resistivity. A typical carrier concentration for a metal is $10^{23}\mathrm{cm}^{-3}$. Metals do not have a gap between the conduction and valence bands.

(2) Semiconductors: n and p are not constant with temperature but are thermally activated. Mobility is related to temperature as $\mu \propto T^{-a}$. The typical range of carrier concentrations for semiconductors is $10^{14} - 10^{20}\mathrm{cm}^{-3}$. The range of bandgaps for semiconductors is typically $0.1 - 3.0\,\mathrm{eV}$.

(3) Insulators: n and p are much lower than they are in metals and semiconductors. A typical carrier concentration for an insulator is $< 10^{7}\mathrm{cm}^{-3}$. The conductivity is generally a function of temperature. The bandgap for insulators is $> 3.0\,\mathrm{eV}$.

124.1 Law of Mass Action

The equilibrium concentration of electrons and holes can be determined by treating them as chemical species. At equilibrium,

$$[np] \rightleftharpoons [n] + [p] \tag{124.2}$$

The rate constant is given by

$$\frac{[n][p]}{[np]} = K(T) \tag{124.3}$$

where

$$K(T) = [F(T)]^2 \exp[-\Delta E/kT] \tag{124.4}$$

Consider the intrinsic case, that is, when the semiconductor is not doped with chemical impurities. For this case,

$$[n] = [p] = [n_i] = A \exp[-\Delta E/2kT] \tag{124.5}$$

where n_i is the intrinsic carrier concentration. For an intrinsic semiconductor, the conductivity is

$$\sigma = 2n_i e\mu \cong \sigma_0 \exp[-\Delta E/2kT] \tag{124.6}$$

A $\log \sigma$ versus $1/T$ plot gives a straight line as shown in Fig. 124.1a.

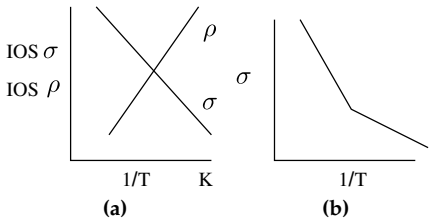


Figure 124.1: (a) Log versus $1/T$ plot of conductivity and resistivity in semiconductors. (b) Log of conductivity versus $1/T$ plot for a semiconductor spanning the regimes of extrinsic and intrinsic conduction.

At lower temperatures where $\Delta E_g > kT$, the conductivity is dominated by carriers contributed from ionized impurities. For temperatures where $\Delta E_g < kT$, thermal energy is sufficient to excite electrons from the valence to conduction band. In this regime, the intrinsic carriers dominate the conduction. The two regimes of conduction are illustrated in Fig. 124.1b.

124.2 Chemistry and Bonding

From the periodic table, we know the valence, atomic weight, and atomic size. A portion of the periodic table with groups IIIA-VIA is shown in Fig. 124.2.

On the left side (group IIIA) are metals. The right hand side (group VIA) contains nonmetals. In between these two groups are elemental semiconductors and elements that form compound semiconductors. At the bottom of the table (In, Sn, and Sb) are more metallic elements. Group IVA consists of the covalent semiconductors Si, Ge, and gray Sn.

Electronic bands are made from an assembly of atoms with individual quantum states. The spectroscopic notation for quantum states is given by

$$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 ... \tag{124.7}$$

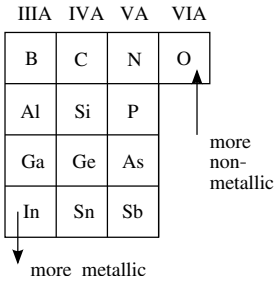


Figure 124.2: Portion of the periodic table showing elements in typical elemental and compound semiconductors.

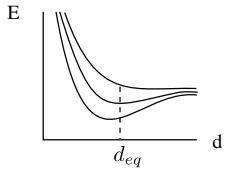


Figure 124.3: Potential energy versus distance between atoms. The equilibrium distance, d_{eq} corresponds to the minimum in the energy curve. The different energies corresponding to d_{eq} form discrete levels in an electronic band.

By the Pauli exclusion principle, no two electrons can share the same quantum state. This results in a formation of electronic bands when atoms are brought together in a periodic lattice. The atomic bonding energy as a function of distance between atoms is shown in Fig. 124.3. In this figure, d_{eq} is the equilibrium distance between atoms with the lowest potential energy and is equal to the lattice constant. The energies correponding to d_{eq} form the allowed energies in electronic bands. A diagram showing the energy levels in electronic bands is shown in Fig. 124.4.

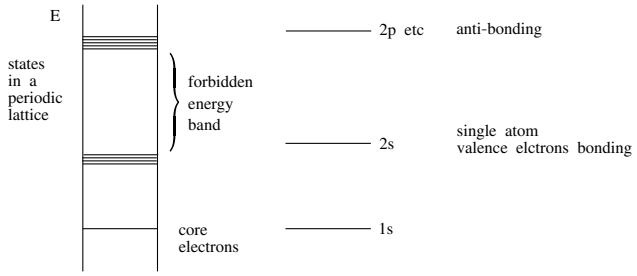


Figure 124.4: Energy diagram showing the electronic bands formed in a crystalline solid.

Core electrons do not contribute to conduction. Conduction of charge carriers occurs in only in partially filled bands. A diagram showing the filling of energy states in the valence band is given in Fig. 124.5. Note that two electrons occupy each energy state having different values of the spin quantum number, that is, "spin up" or "spin down."

Recall that semiconductors have band gap energies of $0.1 - 3.0\,\text{eV}$ that separate the conduction from the valence band. An electron can be promoted from the valence to conduction band if it is supplied with energy greater than or equal to the band gap energy. The source of this energy may be thermal, optical, or electrical in nature. A simple two-level band model for a semiconductor showing the conduction and valence band levels is shown in Fig. 124.6a Also shown (Fig. 124.6b) is a diagram illustrating the process of promoting an electron from the valence to conduction band. When an electron is excited to the conduction band, it leaves behind a "hole" in the valence band, which represents an unfilled energy state. As discussed earlier with regards to the conductivity in semiconductors, both electrons and holes contribute to conduction (see Eqn. 124.1).

124.3 Reciprocal Lattice and the Brillouin Zone

- Reciprocal space: k-space, momentum space

$$\boxed{\vec{p} = \hbar\vec{k}}$$

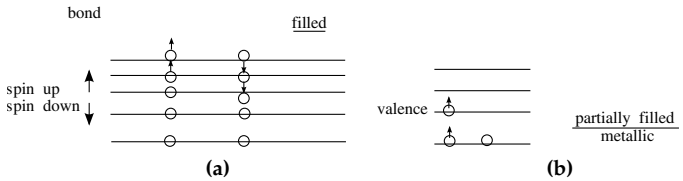


Figure 124.5: (a) Diagram of energy states in a filled valence band. Up arrows indicate the "spin up" state and down arrows indicate the "spin down" state.. (b) Filling of states in a partially filled band. Since the band is partially filled, the electrons can contribute to conduction.

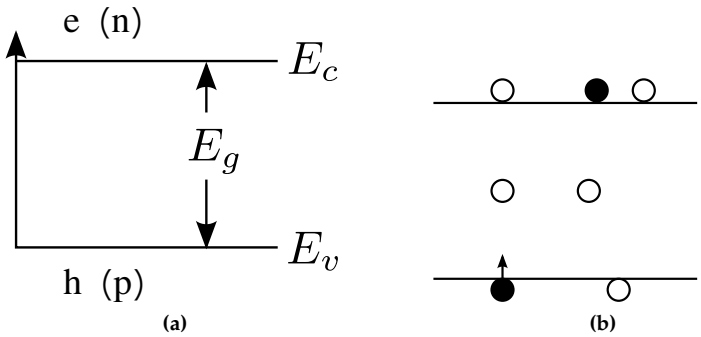
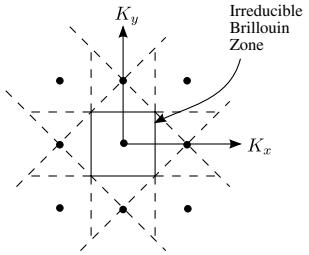


Figure 124.6: (a) Energy band diagram for a simple two-level system. (b) Diagram illustrating the generation of an electron and hole pair.

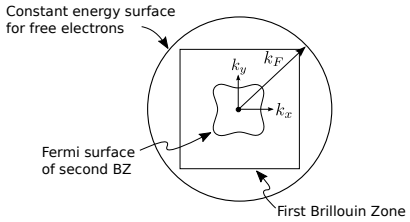
- $\vec{k}$ is the reciprocal lattice vector with units $1/\text{cm}$

Reciprocal 2D Square Lattice

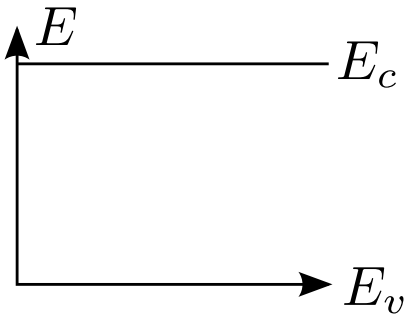


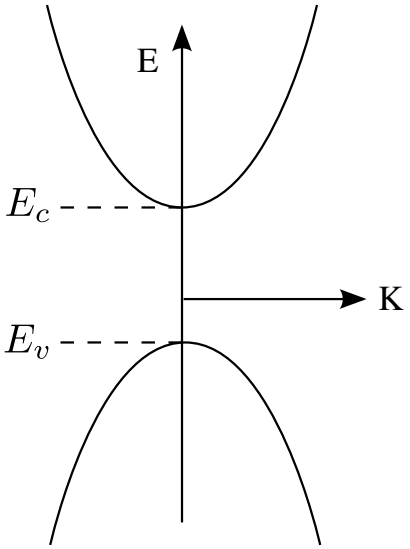
124.4 Nearly Free Electron Model

- Fermi wavevector: $\boxed{k_F = (3\pi^2 n)^{1/3}}$
- Fermi energy: $\boxed{E_F = \frac{\hbar^2 k_F^2}{2m}}$
- Conductivity is proportional to the Fermi surface area, S_F. $\boxed{\sigma \propto S_F}$



124.5 Two Level Model

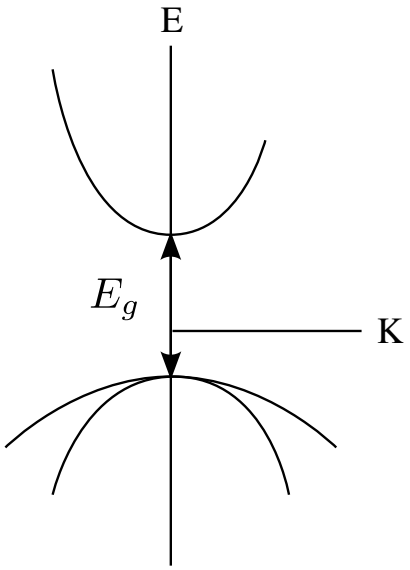


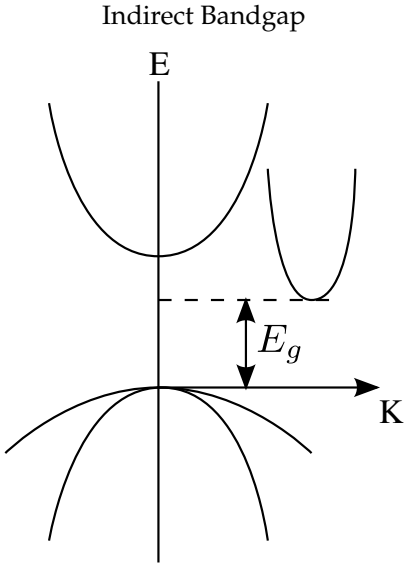


- Parabolic bands: $E = \hbar^2 k^2 / 2m^*$. $\boxed{m^* \equiv \text{effective mass}}$
- Origin? Solution to the Schrodinger Equation, $H\Psi = E\Psi$ for free electrons

125 Semiconductor Band Diagrams

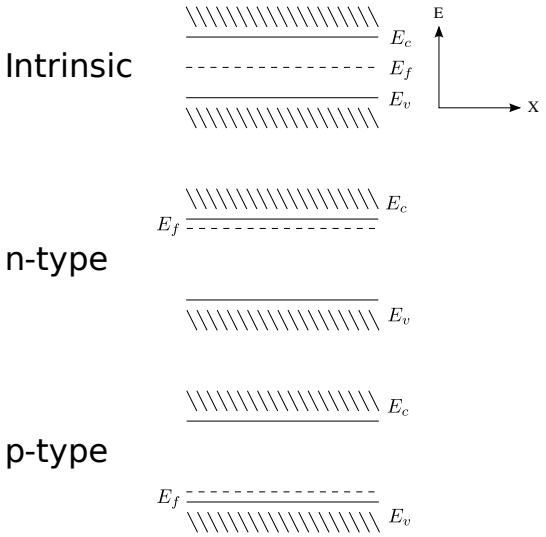
Direct Bandgap



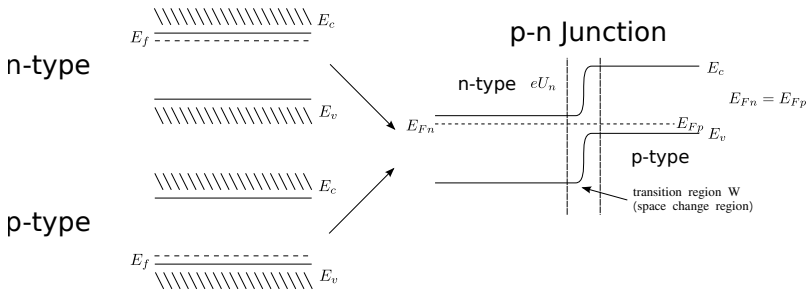


- Direct bandgap semiconductors: III-V's
- Indirect bandgap semiconductors: Si, Ge

125.1 P- and N-Type Semiconductors



125.2 P-N Junctions



125.3 Boltzmann Statistics: Review

For E_v as the reference energy ($E_V = 0$),

$$n = N_s \exp\left[-\frac{E_c - E_F}{kT}\right] \tag{125.1}$$

$$p = N_s \exp\left[-\frac{E_F - E_V}{kT}\right] = N_S \exp\left(\frac{-E_F}{kT}\right) \tag{125.2}$$

For fully ionized acceptors, $p \cong N_A$. For fully ionized donors, $n \cong N_D$.

N_s = Density of states

The Fermi Level for electrons and holes are calculated from Eq.125.1 and 125.2, respectively:

$$\begin{aligned} \text{Electrons:} \quad & E_F = E_c - kT\ln(N_S/n) \\ \text{Holes:} \quad & E_F = E_V + kT\ln(N_S/p) \end{aligned}$$

Note: these equations are true for non-generate semiconductors.

125.4 P-N Junction Equilibrium

Poisson's equation - describes potential distribution $\phi(x)$:

$$\frac{d^2\phi(x)}{dx^2} = -\frac{\rho(x)}{\epsilon}$$

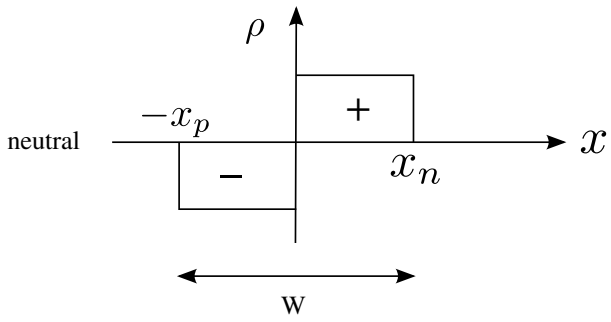
$\phi(x) \equiv$ Potential
$\epsilon \equiv$ Dielectric Constant of Material
$\rho(x) \equiv$ Charge Density

Assume that the donors and acceptors are fully ionized

$$\begin{aligned} \text{p-type:} \quad & N_A \rightarrow N_A^- + p^+ \\ \text{n-type:} \quad & N_D \rightarrow N_D^+ + e^- \end{aligned}$$

Note: the space charge region is positive on the n-side and negative on the p-side.

125.5 Charge Profile of the p-n Junction



- Total charge: $\rho(x) = e\left[p(x) + N_D(x) - n(x) - N_A(x)\right]$
- Space charge width: $w = x_n + x_p$
- Charge balance: $N_A x_p = N_D x_n$
- Note that $x = 0$ at the junction interface.

125.6 Calculation of the Electric Field

$$\frac{d^2\phi}{dx^2} = -\frac{\rho}{\epsilon} \quad \frac{d\phi}{dx} = -\mathcal{E}$$

$$\Rightarrow \frac{d\mathcal{E}}{dx} = \frac{\rho}{\epsilon}$$

$$\mathcal{E}(x) = \int \frac{\rho}{\epsilon} dx$$

Separate the integral for the n- and p-regions:

$$\mathcal{E}(x) = \underbrace{\int_{-x_p}^{0} \frac{\rho(x)}{\epsilon} dx}_{\text{p-region}} + \underbrace{\int_{0}^{x_n} \frac{\rho(x)}{\epsilon} dx}_{\text{n-region}}$$

Important assumptions made:

(1) $\rho \approx eN_D$ in the n region, and $\rho \approx -eN_A$ in the p region.

(2) ρ is constant in the separate n and p regions.

Boundary conditions:

(1) $\mathcal{E}(x = -x_p) = \mathcal{E}(x = x_n) = 0$

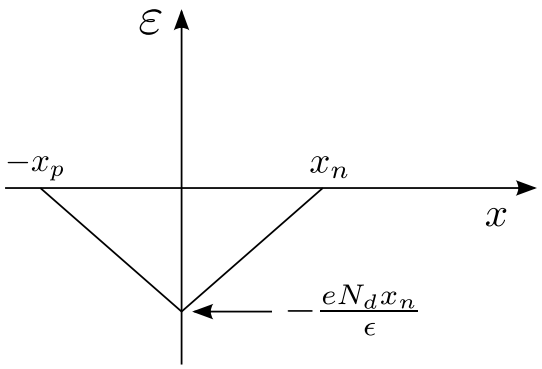
(2) $\mathcal{E}$ is continuous at the junction

Resulting electric field across the junction:

$$\mathcal{E}(x) = \begin{cases} -\frac{eN_A}{\epsilon}(x + x_p), & -x_p \leq x \leq 0 \\ \frac{eN_D}{\epsilon}(x - x_n), & 0 \leq x \leq x_n \end{cases}$$

$$\mathcal{E}(x) = \begin{cases} -\frac{eN_A}{\epsilon}(x + x_p), & -x_p \leq x \leq 0 \\ \frac{eN_D}{\epsilon}(x - x_n), & 0 \leq x \leq x_n \end{cases}$$

Electric Field in the Junction



125.7 Calculation of the Electric Potential

Recall

$$\phi = -\int \mathcal{E} dx$$

Boundary conditions:

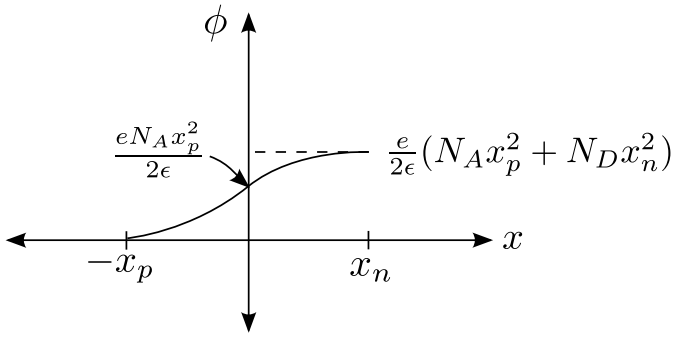
(1) ϕ is continuous at the junction boundary

(2) $\phi = 0$ at $x = -x_p$ (chosen arbitrarily)

Resulting potential across the junction:

$$\phi(x) = \begin{cases} \frac{eN_A}{\epsilon}\left(\frac{1}{2}x^2 + x \cdot x_p + \frac{1}{2}x_p^2\right), & -x_p \le x \le 0 \\ \frac{eN_D}{\epsilon}\left(\frac{N_A}{2N_D}x_p^2 + x \cdot x_n - \frac{1}{2}x^2\right), & 0 \le x \le x_n \end{cases}$$

Electric Potential in the Junction



125.8 Junction Capacitance

- Space charge width:

$$w = x_p + x_n = \left[\frac{2\epsilon\phi_0}{e(N_A + N_D)}\right]^{1/2}\left[\left(\frac{N_A}{N_D}\right)^{1/2} + \left(\frac{N_D}{N_A}\right)^{1/2}\right]$$

$\phi_0 \equiv$ Built-in potential

- Note that, from charge balance, $N_A x_p = N_D x_n$
- Junction capacitance:

$$\boxed{C = \frac{\epsilon}{w}} \quad \text{F/cm}^2$$

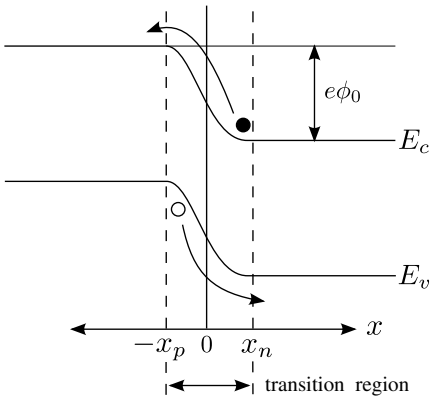
- One-sided abrupt junction: heavily doped p or n
- Consider the case of a p^+-n junction. For $N_A \gg N_D$,

$$\boxed{w = \left(\frac{2\epsilon\phi_0}{eN_D}\right)^{1/2}}$$

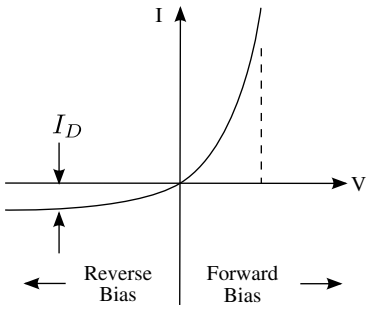
$$\boxed{C = \left(\frac{e\epsilon N_D}{2\phi_0}\right)^{1/2}}$$

125.9 Rectification

Band Diagram



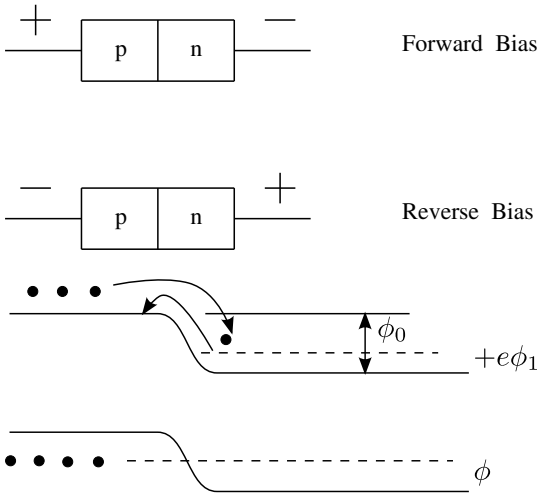
Current-Voltage Characteristic



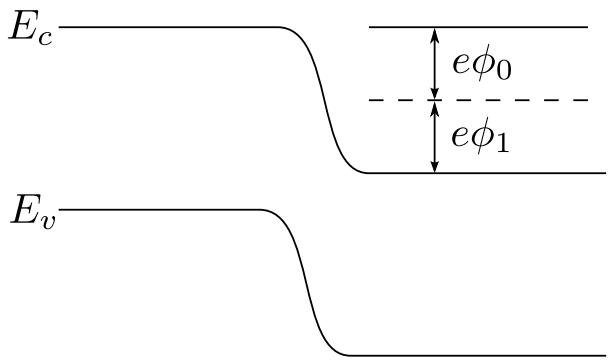
- Larger E_g, larger ϕ_0
- $I_D = I_{0D} \exp(-e\phi_0/kT)$

	E_g	n_i ($\text{cm}^{-3} \times 10^{14}$)
Ge	0.66	0.24
Si	1.08	0.00015
GaAs	1.4	-
GaP	2.25	-
GaN	3.4	-

- Under a forward bias $+\phi_1$, the built in potential is reduced by ϕ_1 and there is a higher probability of electrons going over the barrier.
- Barrier height under forward bias: $e(\phi_0 - \phi)$
- Current increases exponentially with forward bias voltage.



- Under a reverse bias $+\phi_1$, the built in potential is increased by ϕ_1 and there is a lower probability of electrons going over the barrier.



125.10 Junction Capacitance

Definition of capacitance:

$$C \equiv \frac{\partial Q}{\partial \phi}$$

For a one-sided abrupt junction, recall that $Q = eN_D x_n$, where

$$x_n = \left[\frac{2\epsilon(\phi_0 + \phi_1)N_A}{eN_D(N_D + N_A)}\right]^{1/2}$$

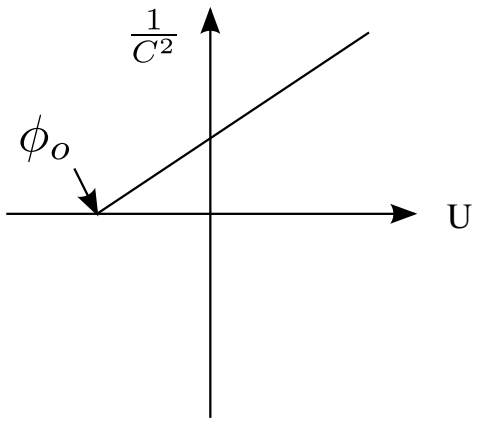
The total charge Q for a one sided junction is then

$$Q = \left[2e\epsilon(\phi_0 + \phi_1)\frac{N_A N_D}{N_A + N_D}\right]^{1/2}$$

Calculate the capacitance as a function of bias voltage ϕ_1:

$$\boxed{C = \frac{\partial Q}{\partial \phi_1} = \left[\frac{e\epsilon}{2(\phi_0 + \phi_1)}\frac{N_A N_D}{N_A + N_D}\right]^{1/2}}$$

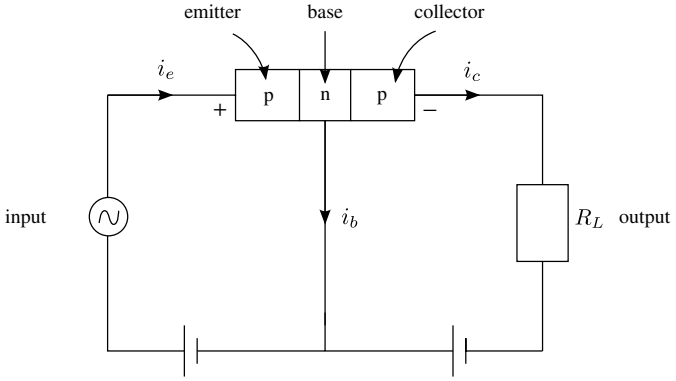
A plot of $1/C^2$ versus ϕ_1 is called a Mott Schottky plot and yields N_A or N_D.



126 Transistors: Semiconductor Amplifiers

- Three terminal device

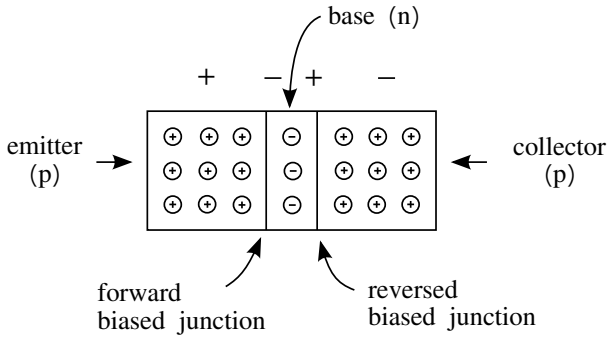
- Bipolar Junction Transistor (BJT): consists of two p-n junctions back-to-back
- The three seminconductor regions are the emitter, base, and collector
- pnp and npn devices



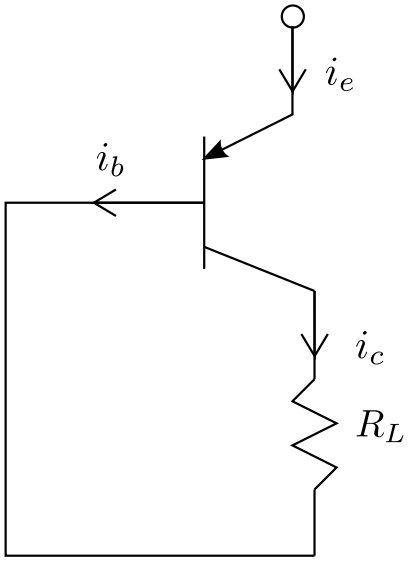
126.1 pnp Device

- Emitter-Base junction is forward biased, and the Collector-Base junction is reverse biased
- Little recombination in the Base region
- Holes are injected into the Base from the Emitter and collected at the Collector.

Device Diagram

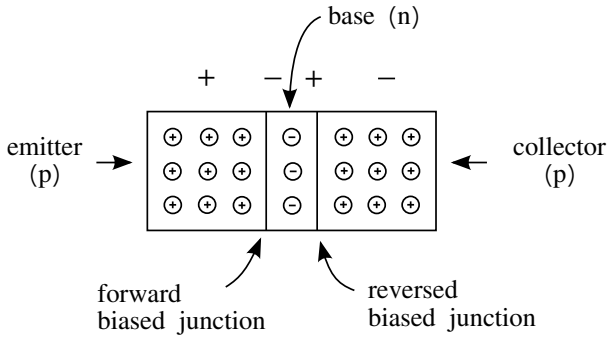


Circuit Diagram

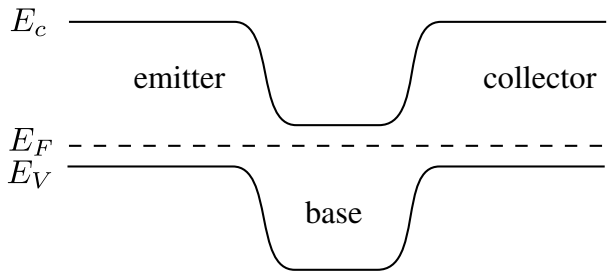


126.2 pnp Device

Device Diagram



Band Diagram

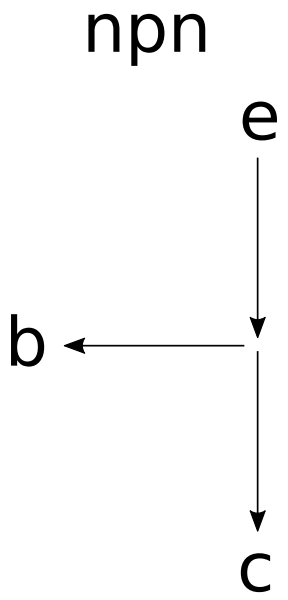


126.3 Amplifier Gain

Voltage across the junctions:

$$i_c\phi_r \gg i_e\phi_f \quad \frac{\phi_r}{\phi_f} \approx 10\,\text{V/V}$$

Nodal analysis at the base for pnp and npn devices:



$$\boxed{i_c \approx i_e} \quad \text{and} \quad \boxed{i_c = \alpha i_e}$$

From nodal analysis at the base, $i_c = i_e - i_b$.

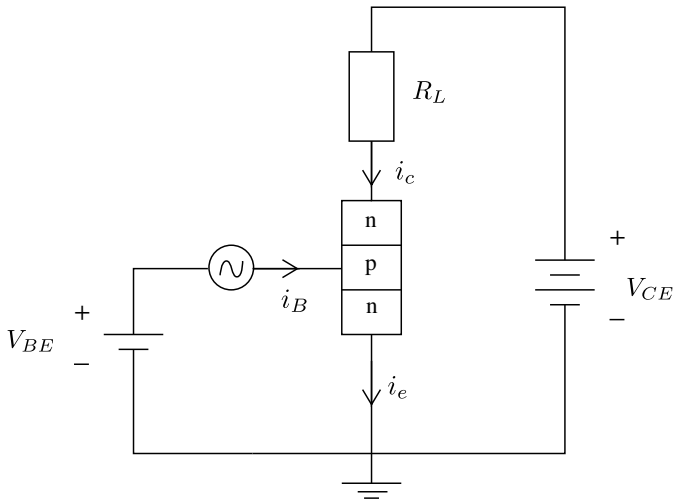
$$\Rightarrow i_c = \frac{\alpha i_b}{1-\alpha} = h_{fe} i_b \quad \text{with} \quad \boxed{h_{fe} = \frac{\alpha}{1-\alpha}}$$

h_{fe} is the current gain parameter. $\alpha \sim 0.9$.

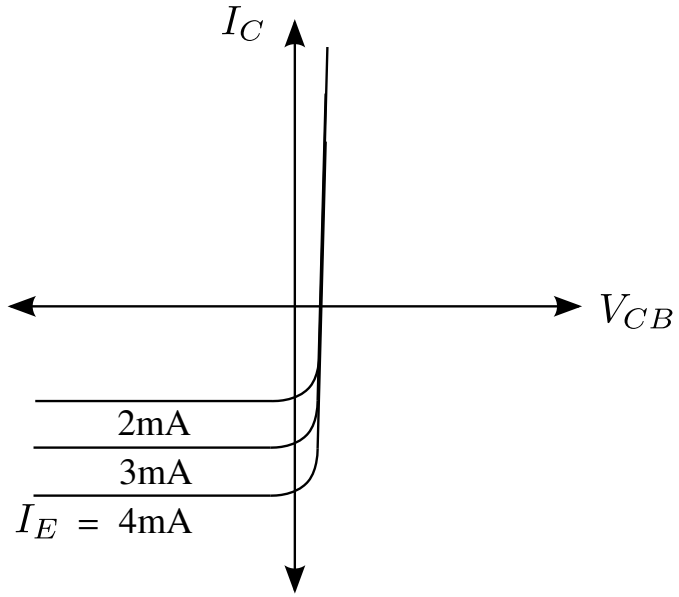
126.4 Circuit Configurations

- Circuit configurations include the common-base, common-collector, and common-emitter.
- The term following "common" indicates which terminal is common to the input and output.

Common-Emitter Diagram

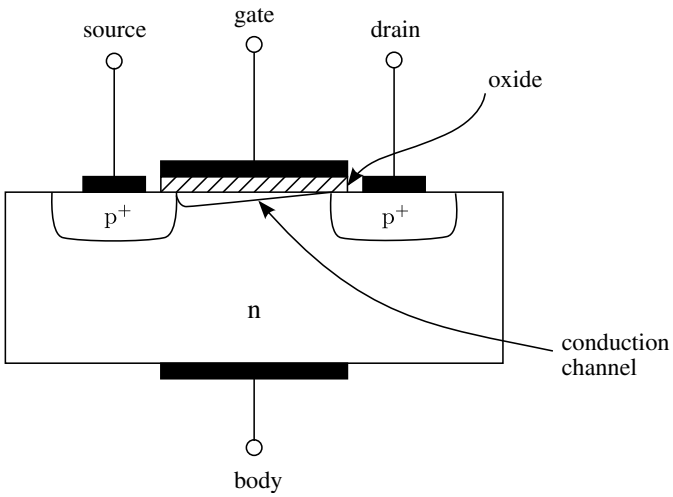


Common-Base I-V Characteristic



126.5 MOSFETs

- MOSFET: Metal Oxide Semiconductor Field Effect Transistor
- An applied electric field induces a channel between the source and drain through which current conducts.



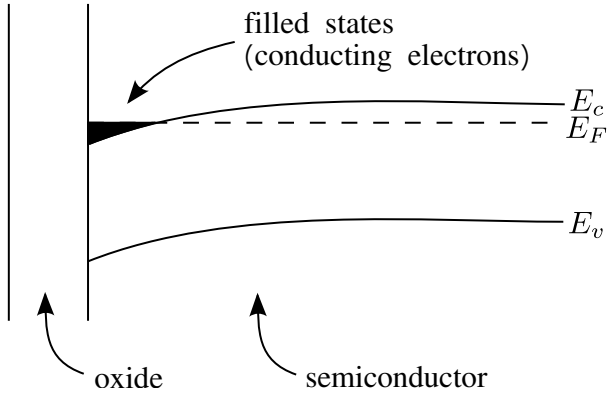
126.6 Oxide-Semiconductor Interface

Changing the gate bias causes the following:

(1) Band bending at the oxide-semiconductor interface.

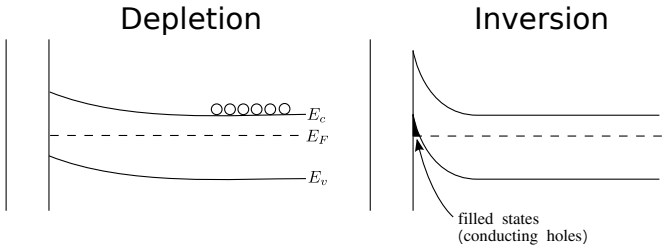
(2) Accumulation of charge at the interface

(3) An increase in channel conductance.



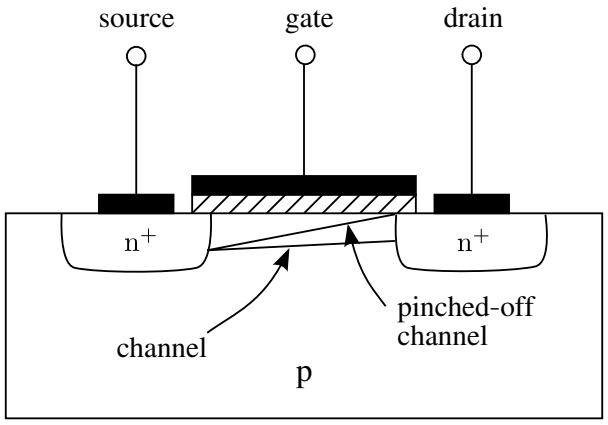
126.7 Depletion and Inversion

- Band bending in the opposite sense leads to depletion of charge carriers.
- Inversion occurs when the Fermi level of an n (p) type semiconductor intercepts the valence (conduction) band due to band bending.
- The channel in a MOSFET is an inversion layer.

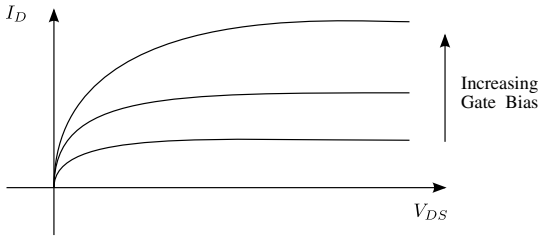


126.8 Channel Pinch-Off

- A threshold voltage V_T is required for channel inversion.
- For a positive gate voltage V_G, the potential across the oxide-semiconductor junction is $V_G - V_T$.
- For a positive voltage between the drain and source (V_{DS}), the potential near the drain is $V_G - V_T - V_{DS}$.
- For $V_{DS} \geq V_G - V_T$, the channel is depleted near the drain, resulting in "pinch-off."



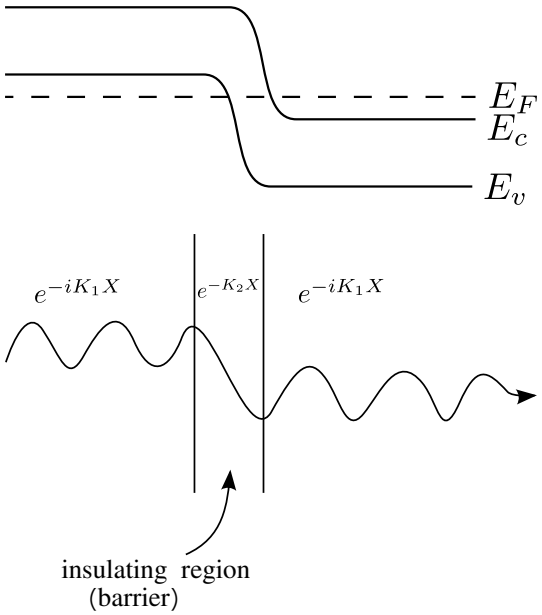
- In saturation mode ($V_{DS} \geq V_G - V_T$), $I_D \propto V_{GS}^2$
- A small change in V_{DS} causes a large change in I_D.



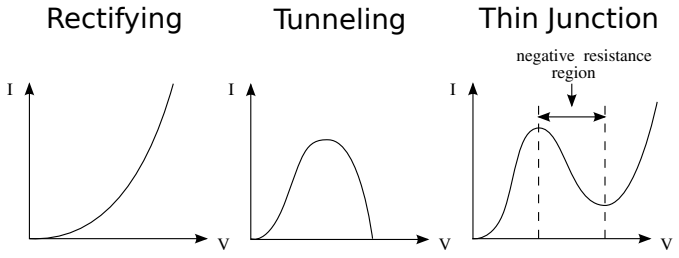
$G =$ Gate
$S =$ Source
$D =$ Drain

126.9 Tunnel Diodes

- Consists of a heavily doped p^+-n^+ structure
- Tunneling current: $I \propto N_{eff} e^{-x/a}$
- Electrons tunnel from filled to empty states across the junction



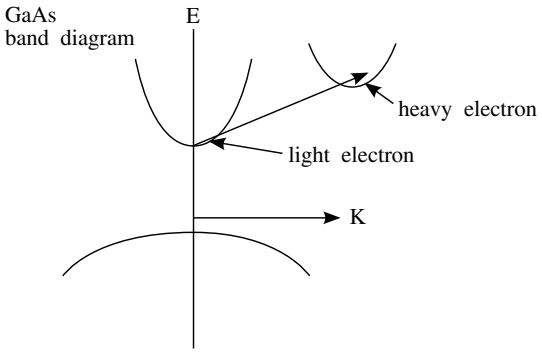
- Electrons can go over the barrier or through the barrier (tunneling)



126.10 Gunn Effect

- Excitation of electrons under a high field to a higher energy conduction band with larger effective mass

$$\sigma = ne\mu \quad \text{and} \quad \mu = \frac{e\tau}{m^*}, \quad m^* = \left(\frac{1}{\hbar^2} \underbrace{\frac{\partial^2 E}{\partial k^2}}_{\text{band curvature}} \right)^{-1}$$

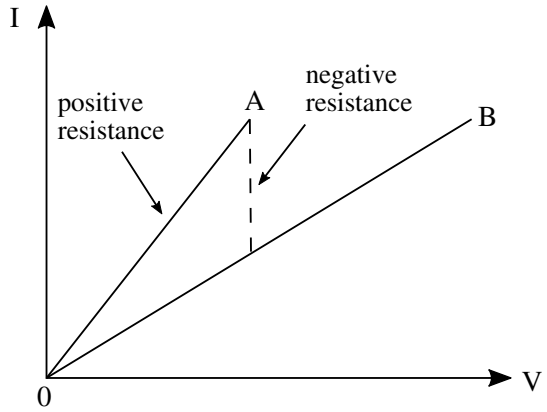


126.11 Gunn Diodes

- Negative resistance due to increased effective mass

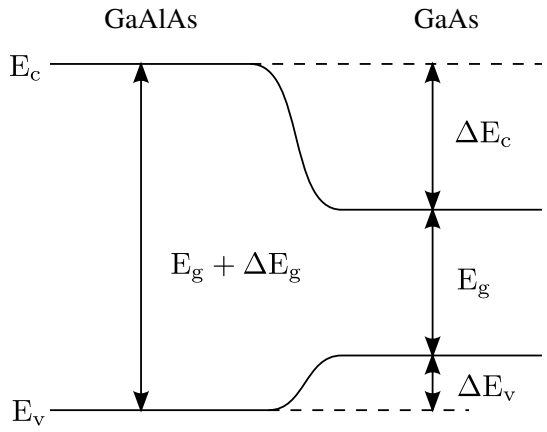
$$J = \sigma\mathcal{E} = \frac{ne^2\tau}{m^*}\mathcal{E}$$

- Transfer of electrons from one valley to another (Transferred Electron Device, TED)



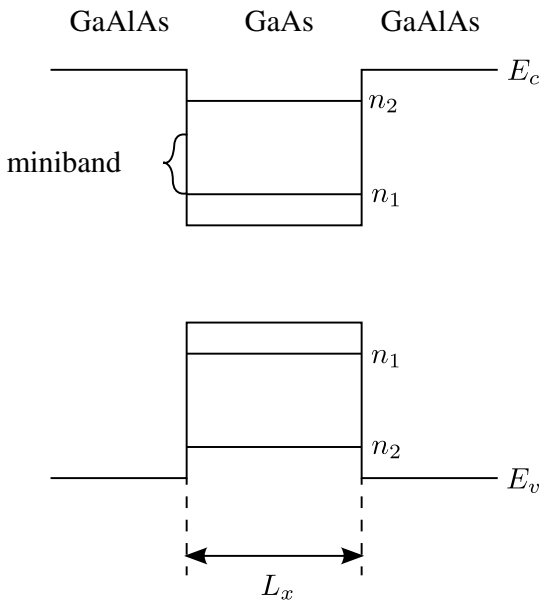
127 Heterojunctions and Quantum Wells

- Composed of semiconductor super-lattices (alternating layers of different semiconductors)
- Multilayers grown by Molecular Beam Epitaxy (MBE) and Metal-Organic Chemical Vapor Deposition (MOCVD)
- Used in the construction of lasers, detectors, and modulators
- Utilizes band offset, can be used to confine carriers



127.1 Quantum Wells

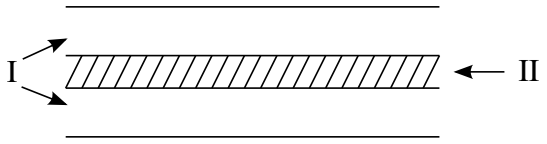
- From the particle in a box solution, minibands are formed within the well



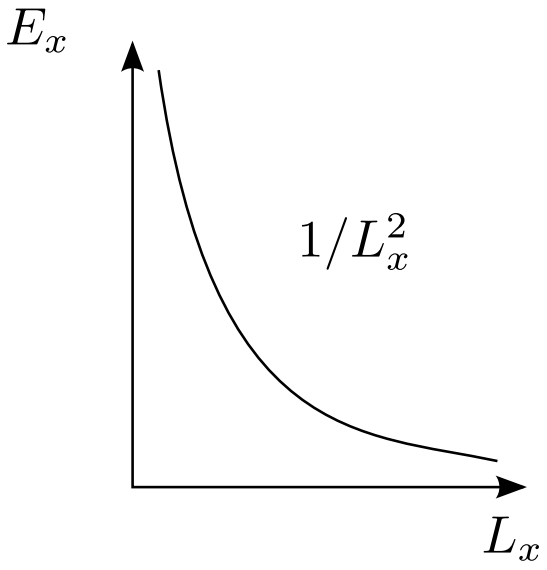
$$H\Psi = E\Psi; \quad \Psi = A\exp(-k_n x)$$

$$\boxed{k_n = \frac{n\pi}{L_x}}$$

$$\boxed{E_x = \frac{\hbar^2 k^2}{2m^*} = \frac{\hbar^2}{2m^*}\left(\frac{\pi^2 n_x^2}{L_x^2}\right)}$$

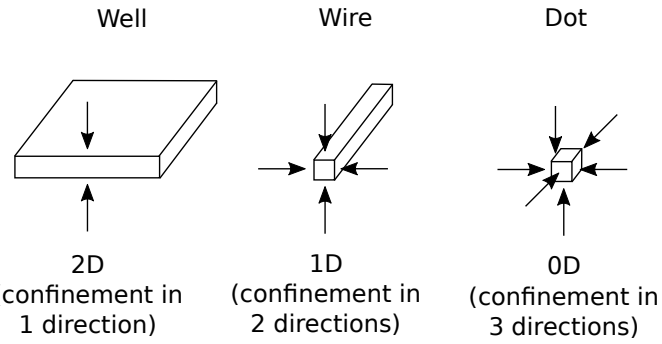


- Smaller L_x, larger energy of miniband
- L_x can be tuned to engineer the band gap



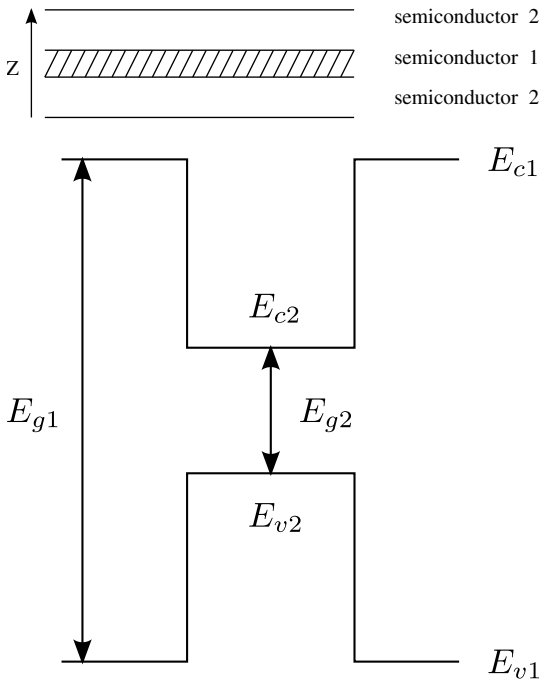
127.2 Heterostructures and Heterojunctions

- Devices that use heterojunctions and heterostructures: lasers, modulation-doped field effect transistors (MODFETs)
- Types of heterostructures: quantum wells (2D), wires (1D), and dots (0D)



127.3 Layered Structures: Quantum Wells

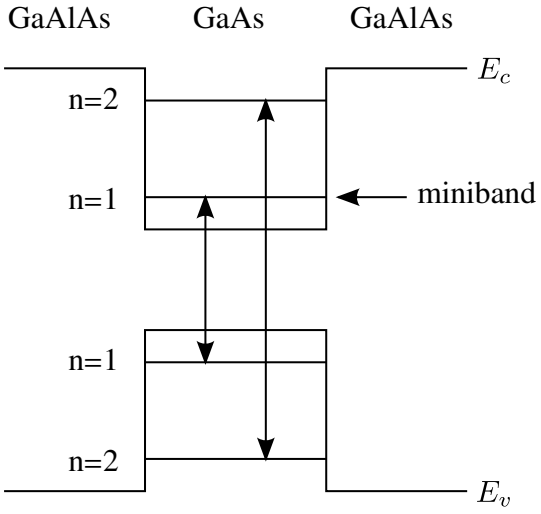
- "Cladding" layer can confine carriers due to difference in bandgap and light due to difference in refractive index.
- Examples: GaAs/GaAlAs, InP/InGaAsP, GaN/InGaN
- Semiconductor 1 is "lattice matched" (same lattice constant as substrate)
- Semiconductor 2 is the strained layer (different lattice constant)



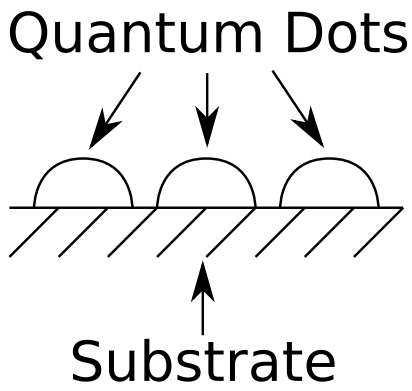
127.4 Transitions between Minibands

- Allowed initial and final states are governed by "selection rules."
- Will have conduction in minibands
- Energy of states:

$$E = \frac{\hbar^2 k^2}{2m^*} = \frac{\hbar^2}{2m^*}\left(\frac{n\pi}{L_x}\right)^2$$



127.5 Quantum Dots

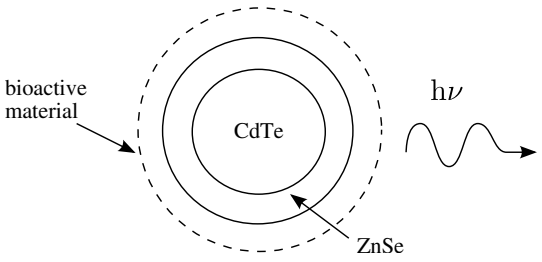


- Electrons are confined in all dimensions to "dots."
- Quantum Dots: clusters of atoms 3-10 nm in diameter.
- QDs can be created by colloidal chemical synthesis or island growth epitaxy.
- Modeled after the hydrogen atom with radius R. Quantized energies:

$$E_i = \frac{\hbar^2}{2m_i}\left(\frac{\alpha_i}{R}\right)^2$$

$\alpha_i \equiv$ quantum number

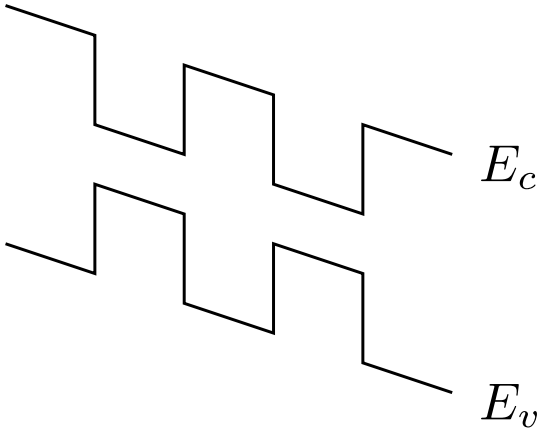
127.6 Quantum Dot Example: Biosensor



- Solution synthesized colloidal quantum dots (CQD)
- Fluorescence energy depends on surface and size of dot
- Binding of molecules to the surface of the QD quenches the photoluminescent intensity.

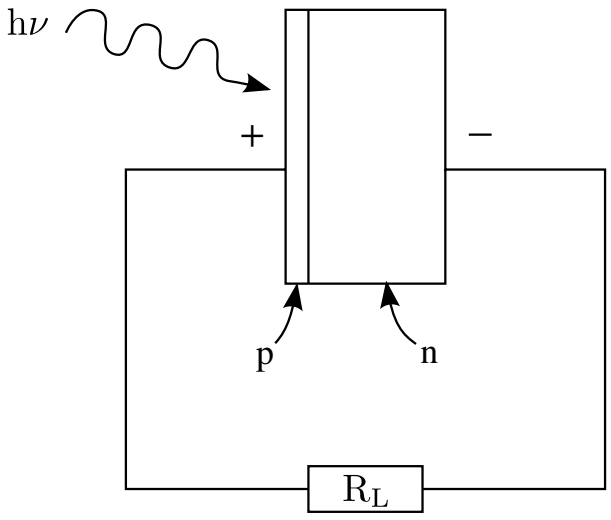
127.7 Quantum Cascade Devices

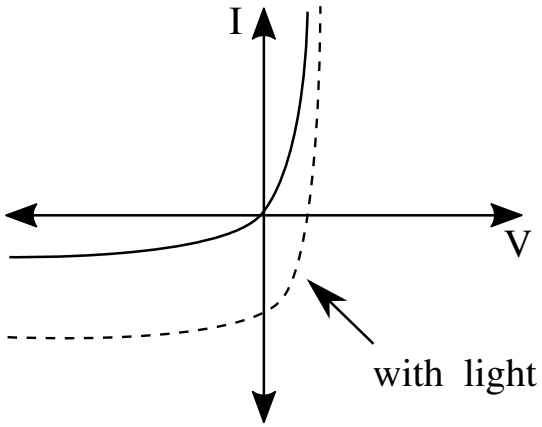
- Consist of multiple quantum wells
- Band bending occurs with bias

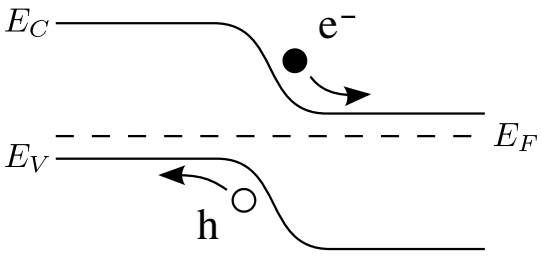


128 Optoelectronic Devices: Photodetectors and Solar Cells

- Light incident on a p-n junction generates electron-hole pairs which are separated by the built in potential.
- Separated carriers contribute to the photocurrent.







128.1 I-V Characteristics

- Diode equation: $I = I_0 \left(e^{qV/kT} - 1 \right)$
- Short circuit current: $\boxed{I_{sc} = Aq(L_e + L_h)G}$

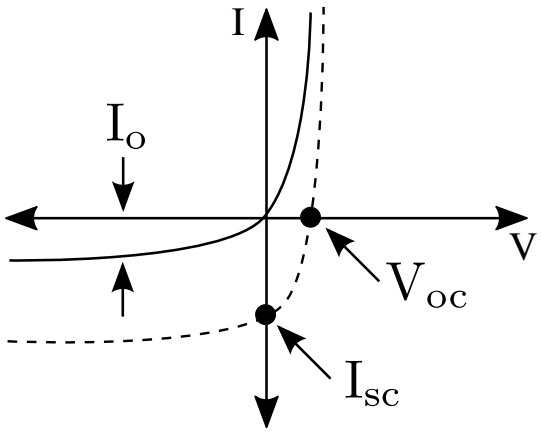
$L_e \equiv$ Electron diffusion length
$L_h \equiv$ Hole diffusion length
$G \equiv$ Carrier generation rate (depends on light intensity)

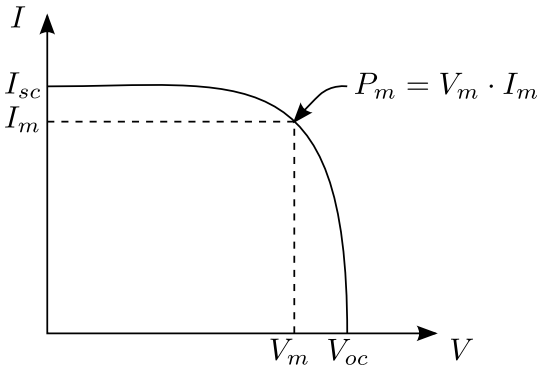
- Current under illumination: $I = I_0 \left(e^{qV/kT} - 1 \right) - I_{sc}$
- Open circuit voltage: $\boxed{V(I = 0) = V_{oc} = \frac{kT}{q} \ln \left(\frac{I_{sc}}{I_0} + 1 \right)}$

128.2 Power Generation

- Quadrant IV is the power quadrant.
- Theoretical maximum power: $P_{max,theory} = V_{oc} I_{sc}$
- Maximum obtainable power: $\boxed{P_{max} = ff(I_{sc} V_{oc})}$

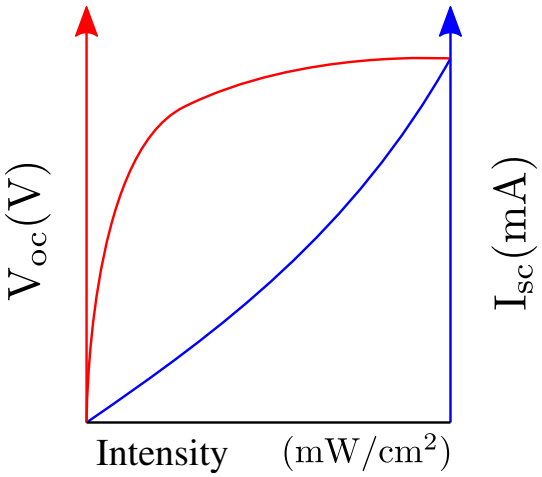
Fill Factor: $ff \equiv \frac{I_m V_m}{I_{sc} V_{oc}}$

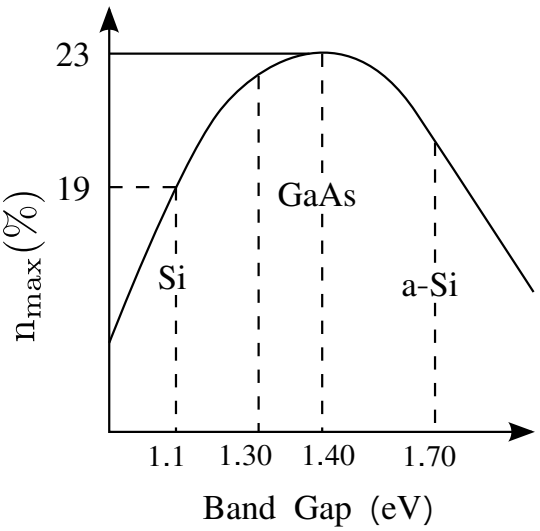




128.2.1 Solar Cell Efficiency

- Efficiency is governed by the fill factor and bandgap
- Want larger fill factors and absorption that matches the solar spectrum



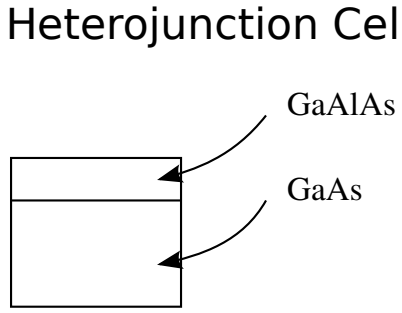


128.2.2 Other Types of Solar Cells

(1) Multi-junction solar cells: GaAlAs/GaAs. More efficient absorption.

(2) Thin films: amorphous silicon (a-Si), CdTe. Inexpensive.

(3) Metal-semiconductor solar cells

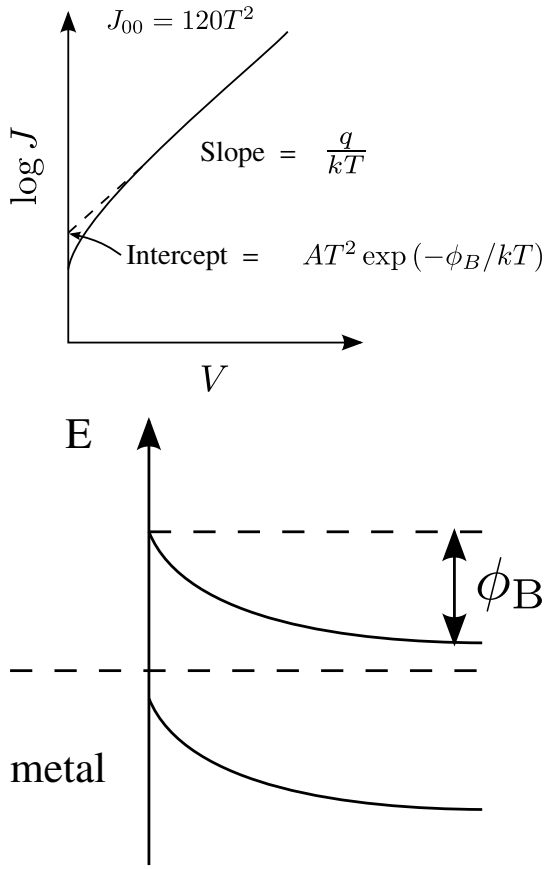


128.2.3 Metal-Semiconductor Solar Cells

- Energy barrier ϕ_B at metal-semiconductor junction
- Current density, J:

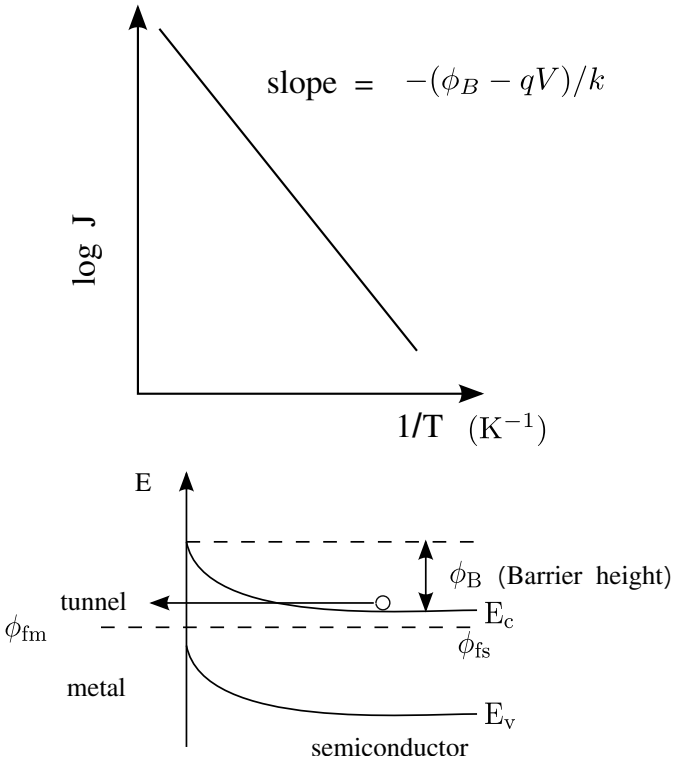
$$J = J_{00} \exp\left(\frac{-\phi_B}{kT}\right)\left[\exp\left(\frac{qV}{kT}\right) - 1\right]$$

$$\boxed{J \equiv \frac{\text{current}}{\text{area}} = \frac{I}{A}}$$

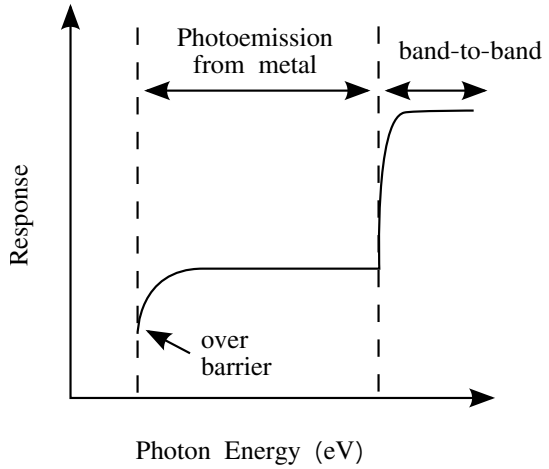


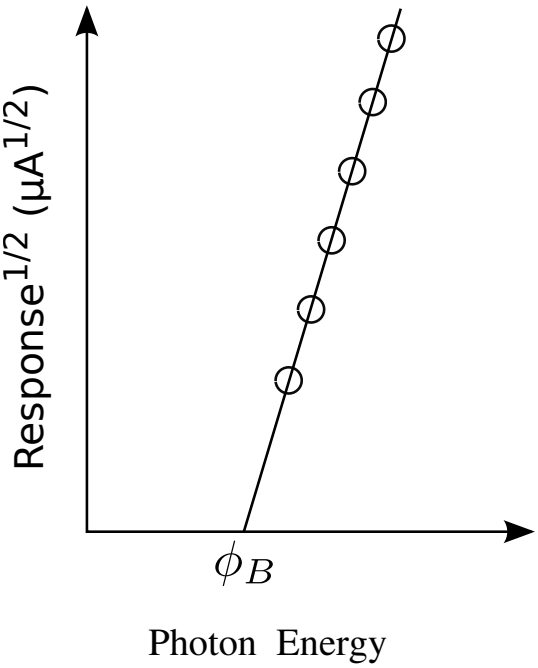
128.2.4 Schottky Barrier and Photo Effects

- Barrier height is obtained from $\log(J)$ versus $1/T$ plot
- Electrons can tunnel across the barrier



- Photo-response is characterized by two regions: metal photo-emission, and band-to-band excitation.
- Photo-response $\propto \sqrt{\phi_B}$

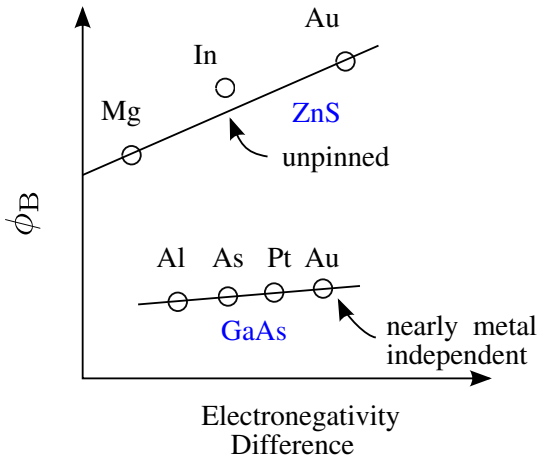


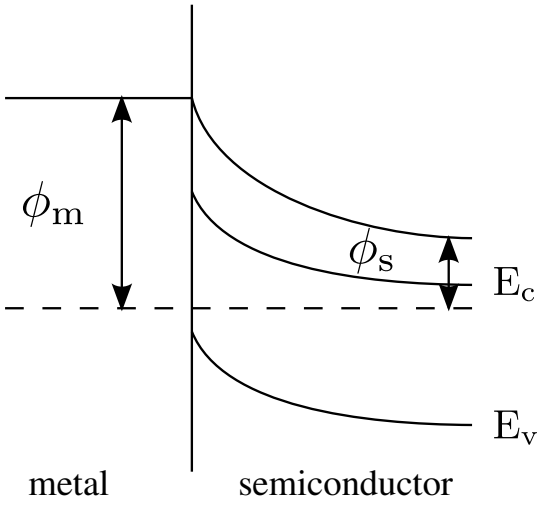


128.2.5 Dependence of ϕ_B on Work Function

$\phi \equiv$ energy required to take an electron from the Fermi Level to the Vacuum Level

$$\phi_B = \phi_m - \phi_s$$



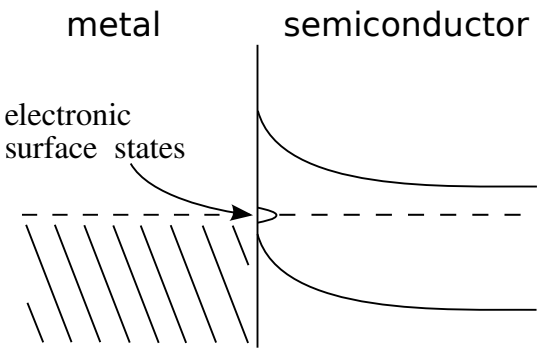


128.2.6 Pinned Surfaces

- Surface states can influence the amount of band bending at a metal/semiconductor interface
- For "pinned" surfaces, the barrier height is nearly independent of the metal work function.
- Total amount of band bending is equal to the band gap energy:

$$\phi_{B(n)} + \phi_{B(p)} = E_g$$

- Example, InP (n): $\phi_{B(n)} \sim 0.5$ eV $\quad \phi_{B(p)} \sim 0.8\,\text{eV} \quad E_g \sim 1.27\,\text{eV}$
- For making an Ohmic contact to an n-type semiconductor, $\phi_B = \phi_m - \phi_s < 0$

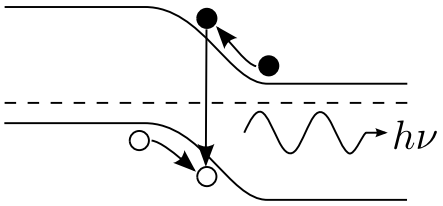


128.2.7 Light Emitting Diodes (LEDs)

- Forward biasing a p-n junction results in electron-hole recombination at the junction
- Recombination results in light emission. The wavelength depends on E_g

$$h\nu = E_g$$

p-n Junction under Forward Bias



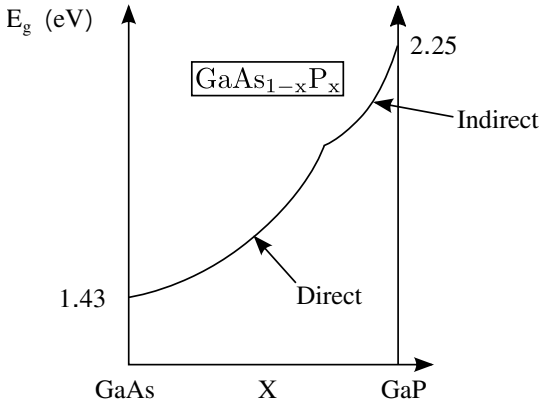
128.2.8 Light Emitting Diodes (LEDs)

Material	E_g at 300K (eV)	Comments
GaP	2.25	
GaAsP	1.7-2.25	Alloy
GaAs	1.43	
InGaAsP	0.8-1.4	
ZnSe	2.58	Blue
GaN	3.4	Ultraviolet
AlGaInN	4.2	

128.2.9 Solid Solution Alloys

- The bandgap can be engineered by alloying
- Vegard's law can be used to approximate the bandgap

$$E_g \approx xE_{g1} + (1-x)E_{g2}$$



128.2.10 LED Efficiency

- LED efficiency is the ratio of the radiative recombination probability (W_R) to the total recombination probability (W_{total})

$$\boxed{\eta = \frac{W_R}{W_{total}} = \frac{W_R}{W_R + W_{NR}}}$$

- W_{NR}: the probability of non-radiative recombination
- Non-radiative processes limit the efficiency
- Recombination probability is related to the lifetime (τ)

$$W \propto \frac{1}{\tau} \ [\mathrm{s}^{-1}]$$

$$\boxed{\frac{1}{\tau_{total}} = \frac{1}{\tau_R} + \frac{1}{\tau_{NR}}}$$

- Efficiency in terms of probability (W) and lifetime (τ)

$$\eta = \frac{W_R}{W_{total}} = \frac{W_R}{W_R + W_{NR}}$$

$$\eta = \frac{1/\tau_R}{1/\tau_{total}} = \frac{1}{1 + \tau_R/\tau_{NR}}$$

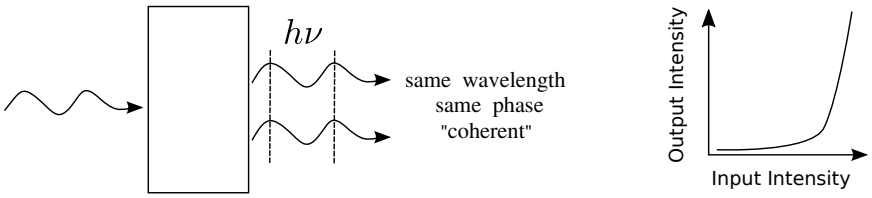
- For $\tau_R \gg \tau_{NR}$, τ_{NR} determines the radiation efficiency $\therefore$ want as few non-radiative defects (dislocations, impurities,etc.) as possible

$$\boxed{\eta = \frac{\tau_{NR}}{\tau_R}}$$

- GaN example: 10^7 dislocations per cm^2

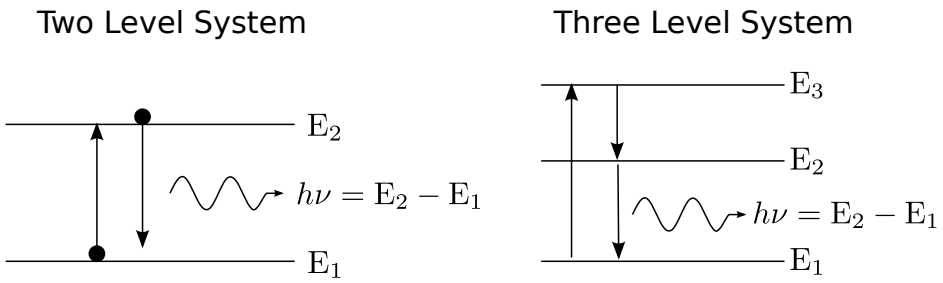
129 Lasers

- LASER: Light Amplification through Stimulated Emission of Radiation
- Laser light has narrow divergence and consists of photons of the same frequency and phase
- Photons of a certain frequency and phase stimulate emission of photons with the same frequency and phase



129.1 Emission Rate and Laser Intensity

- Calculate emission rate and laser intensity by accounting for both emission and absorption processes using Boltzmann Statistics
- Calculate the transition rate between states and Einstein A and B coefficients
- Consider 2 and 3 level systems



129.2 Two Level System

- At equilibrium, the population of levels is determined by the Boltzmann distribution

$$N_2 \propto Ne^{-E_2/kT} \quad \text{or} \quad \frac{N_2}{N} = \frac{e^{-E_2/kT}}{Z}$$

$$Z \equiv \text{sum over states}$$

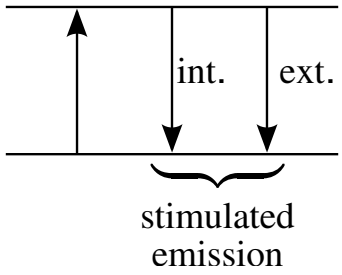
- Ratio of populations in a two level system:

$$\frac{N_2}{N_1} = e^{-(E_2-E_1)/kT}$$

At equilibrium, fewer electrons are in level 2

Triggers of Stimulated Emission

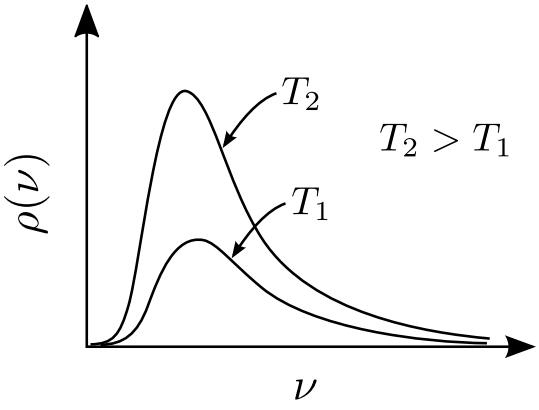
- Internal or external radiation can cause the (stimulated) transition to occur
- Internal sources are described by "black body" radiation.



129.3 Planck Distribution Law

- Number of photons inside the cavity is determined by Planck's Distribution Law
- Spectral density of photons of frequency ν in linewidth $d\nu$

$$\rho(\nu)d\nu = \underbrace{\frac{8\pi n^3 h\nu^3}{c^3}}_{\text{Density of states}} \underbrace{\frac{1}{\exp(h\nu/kT)-1}}_{\text{Planck distribution}} d\nu$$



129.4 Transition Rates

- Transition rate R_{ij}: rate at which electrons transition from state i to state j.
- Absorption ("up") rate R_{12}:

$$R_{12} = N_1 B_{12} \rho(\nu_{21}) d\nu$$

$$N_1 \equiv \text{Number of photons in lower state}$$

$$\rho(\nu_{21}) \equiv \text{number of photons available having transition energy} h\nu$$

- Emission ("down") rate R_{21}:

$$R_{21} = N_2 [\underbrace{A_{21}}_{\text{Spont.}} + \underbrace{B_{21}\rho(\nu_{21})}_{\text{Stim.}}] d\nu$$

- A_{21} and B_{21} are the Einstein A and B coefficients for spontaneous and stimulated transitions, respectively, between states 2 and 1.

129.5 Two Level System

- At equilibrium (steady state):

$$R_{12} = R_{21}$$

$$N_1 B_{12} \rho(\nu_{21}) d\nu = N_2 \left[A_{21} + B_{21}\rho(\nu_{21})\right] d\nu$$

- Substitute in $\frac{N_2}{N_1} = e^{-(E_2 - E_1)/kT}$, simplifying and solving for $\rho(\nu_{21})$:

$$\rho(\nu_{21}) d\nu = \frac{A_{21} d\nu}{B_{12} \exp\left[(h\nu_{21})/kT\right] - B_{21}}$$

- Note: same form as Planck Distribution Law with $B_{12} = B_{21}$
- In a two level system, the stimulated emission and absorption rates are equal, i.e., $B_{12} = B_{21}$
- Comparing to the Planck Distribution Law,

$$A_{21} = B_{21} \frac{8\pi n^3 h\nu^3}{c^3}$$

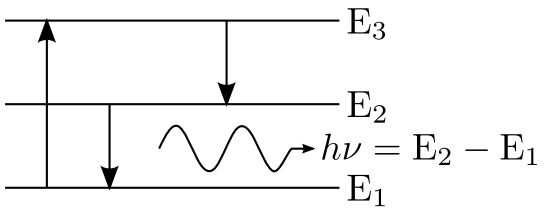
- The coefficient of spontaneous emission A_{21} is related to the stimulated emission rate

Two Level System - Summary

- Number of photons inside the cavity is determined by Planck's Distribution Law
- The stimulated emission and absorption rates are equal
- The coefficient of spontaneous emission A_{21} is related to the stimulated emission rate
- A population inversion is required for lasing

129.6 Three Level System

- A three level system consists of a ground state (E_1), excited state (E_2), and intermediate state (E_3).
- Electrons are pumped from E_1 to E_3 by a pump source with energy $\geq h\nu_{31}$; N_2 population is unaffected
- More electrons are in the upper state E_3 than the lower (metastable) state $E_2 \Rightarrow$ a population inversion is achieved

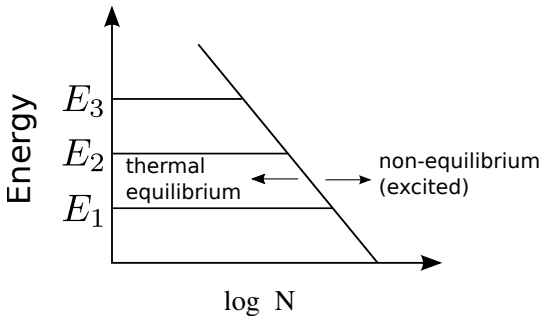


Carrier Population

- The carrier population is given by the Boltzmann distribution

$$\begin{aligned} N(E) &= N_0 \exp(-E/kT) \\ \log N &= -E/kT + \log N_0 \end{aligned}$$

- Lower energy states have higher carrier populations at equilibrium



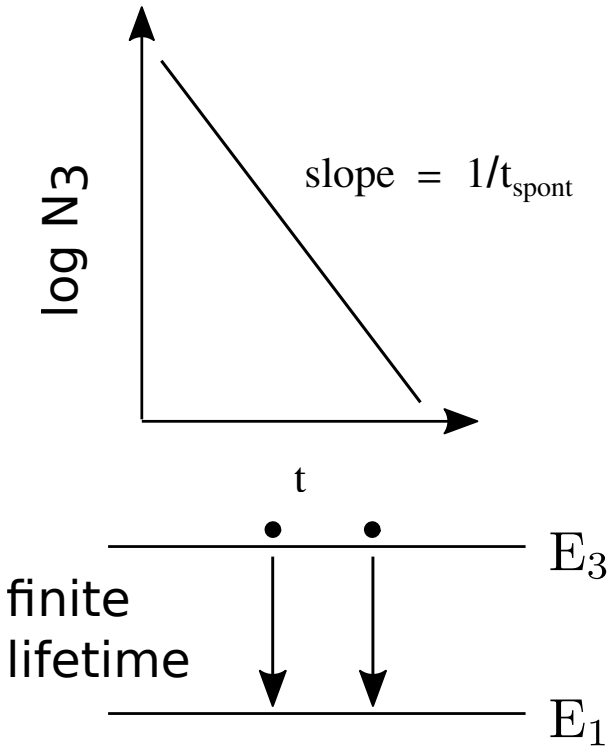
Spontaneous Emission

- A_{31}: spontaneous emission rate; measure of the spontaneous depopulation
- Assuming first order kinetics,

$$\frac{dN_3}{dt} = -A_{31}N_3$$

$$A_{31} = \frac{1}{t_{spont}}$$

- Solution for initial population $N_{3,i}$: $N_3(t) = N_{3,i} \exp(-A_{31}t)$



Populations under High Pumping

- Under high pumping at steady state, the populations in levels 1 and 3 are

$$N_1^* = N_3^* = \frac{N_1 + N_3}{2}$$

- Equal probabilities of emission and absorption: material is "transparent"
- From energy balance,

$$\underbrace{B_{23}N_2\rho(\nu)}_{\text{Stim. absorption}} = \underbrace{A_{32}N_3}_{\text{Spont. emission}} + \underbrace{B_{32}N_3\rho(\nu)}_{\text{Stim. emission}}$$

Laser Gain

- The net number of photons gained per second per unit volume is

$$NW_{12} = (N_2 - N_1)W_{12}$$

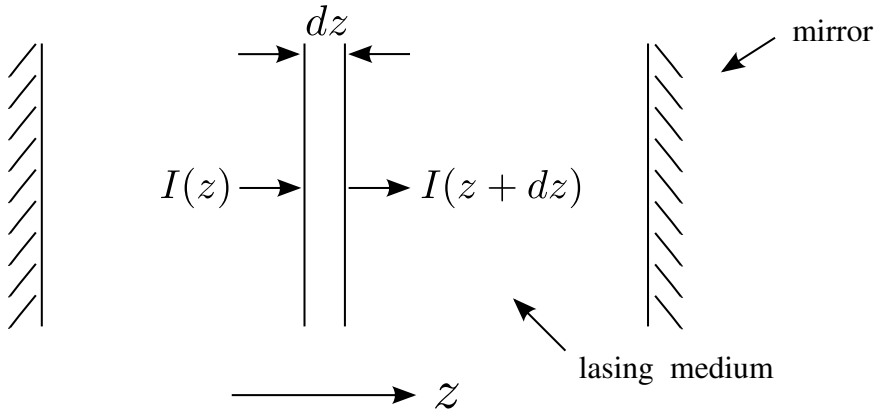
$$W_{12} \equiv \text{Transition rate from state 1 to 2}$$

- For $N > 0$: population inversion, material can act as an amplifier
- For $N = 0$: the material is transparent
- For $N < 0$: light is attenuated

For amplification, a population inversion must exist.

Derivation of Laser Gain Factor

- Stimulated-emission photons travel in the same direction (z) as the incident photons and further stimulate emission
- $\Rightarrow$ photon flux and intensity increase in z direction



- Increase in photon intensity through a volume of thickness dz:

$$dI = (N_2 - N_1)W_{12}h\nu_{12}dz$$

$$\therefore \frac{dI}{dz} = (N_2 - N_1)W_{12}h\nu_{12}$$

- Transition probability is defined as

$$W_{12} = \frac{I}{h\nu}\frac{c^2 g(\nu)}{8\pi n^2 \nu^2 t_{spont}}$$

- $g(\nu)$ is the lineshape factor
- Laser gain factor $\gamma(\nu)$ is defined by the equation

$$\frac{dI}{dz} = \gamma(\nu)I$$

$$\therefore \boxed{\gamma(\nu) = (N_2 - N_1)\frac{c^2 g(\nu)}{8\pi n^2 \nu^2 t_{spont}}}$$

- The change in laser intensity with distance z is

$$\frac{dI}{dz} = \gamma(\nu)I$$

- Solution for $I(z)$:

$$\boxed{I(z) = I(0)\exp\left[\gamma(\nu)z\right]}$$

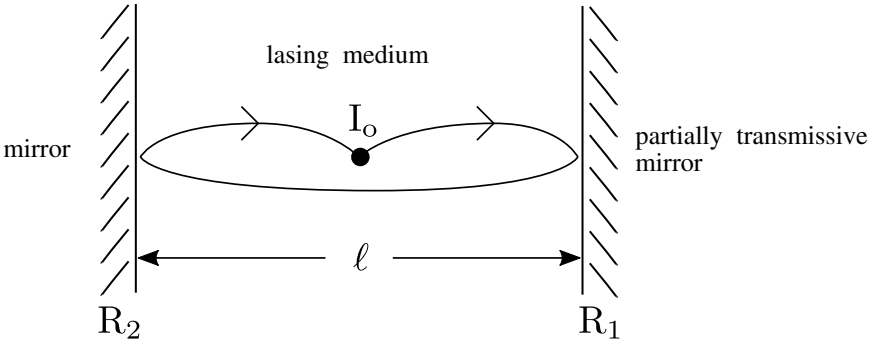
- Recall laser gain factor from previous slide:

$$\boxed{\gamma(\nu) = (N_2 - N_1)\frac{c^2 g(\nu)}{8\pi n^2 \nu^2 t_{spont}}}$$

- For amplification, $I(z)/I(0) > 1$ is required $\therefore N_2 > N_1$. A population version is required for amplification
- Laser threshold: operation at which $I(z)/I(0) = 1$

129.7 Lasing Modes

- Lasing frequencies are chosen by confining the stimulated-emission photons inside a cavity called a Fabry Perot resonator
- Standing waves of certain discrete frequencies are established inside the cavity



- Laser intensity after a round trip in the cavity:

$$I = I_0 R_1 R_2 \exp\left[(\gamma - \alpha)2l\right]$$

$\alpha \equiv$ absorption and scattering losses
$\gamma \equiv$ laser gain factor

- Criteria for net gain: gain $\geq$ sum of losses

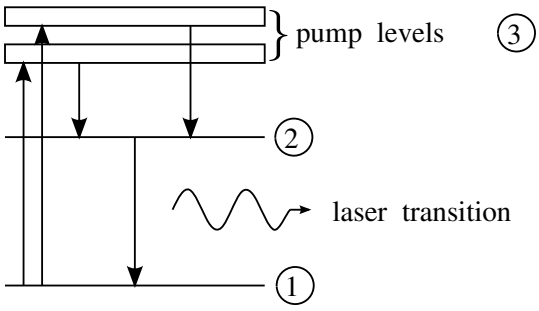
$$\boxed{R_1 R_2 \exp\left[(\gamma - \alpha)2l\right] \geq 1}$$

- Modes that are in phase after a round trip are stabilized

129.8 Laser Examples

129.8.1 Ruby

- Transition metal ion (Cr) doped alumina (Al_2O_3)
- 3 level system - first demonstrated laser
- Optically pumped by a "discharge lamp"



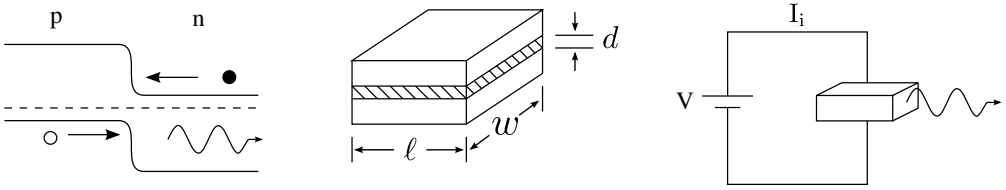
129.8.2 Others Examples

Semiconductor:

- Emission from carrier recombination in a forward biased p-n junction
- Heterostructures form both p-n junction and waveguiding region (cavity)

Others:

- Ceramic: Nd^{3+} :YAG, 4-level system
- Glass fiber laser Er^{3+} :SiO_2 Erbium Doped Fiber Amplifier (EDFA), basis of modern optical communication



$d \equiv$ active region thickness

129.9 Threshold Current Density

- Number of electrons injected into lasing region per unit volume: N_e
- Recombination rate:

$$\frac{N_e l w d}{t_{rec}} = \frac{I_i \eta}{e}$$

$$\begin{aligned} \eta &\equiv \text{Efficency} \\ I_i &\equiv \text{Injection current} \\ e &\equiv \text{Fundamental charge} \end{aligned}$$

- Recall the gain coefficient, $\gamma(\nu)$

$$\gamma(\nu) = (N_3 - N_2)\frac{c^2 g(\nu)}{8\pi n^2 \nu^2 t_{spont}}$$

$$\boxed{t_{rec} \cong t_{spont}}$$

$$N_e = N_3 - N_2$$

- Substituting in $N_3 - N_2 = N_e = \frac{I_i \eta}{e}\frac{t_{rec}}{lwd}$ and $t_{spont} \cong t_{rec}$,

$$\gamma(\nu) = \frac{c^2 g(\nu)\eta}{8\pi n^2 \nu^2 e l w d} I_i \ (\text{cm}^{-1})$$

Relates gain coefficient to I_i

- The linewidth factor is $\boxed{g(\nu) \cong 1/\Delta\nu}$
- Let the mirror reflectivity be $R_1 R_2 = R$. Recall the lasing threshold condition

$$R\exp(\gamma-\alpha)l' = 1$$

$$(\gamma-\alpha)l' + \ln R = 0$$

- Define threthold current density: $j_{th} = I_{i,th}/wl$
- Substituting in for γ and solving for j_{th}:

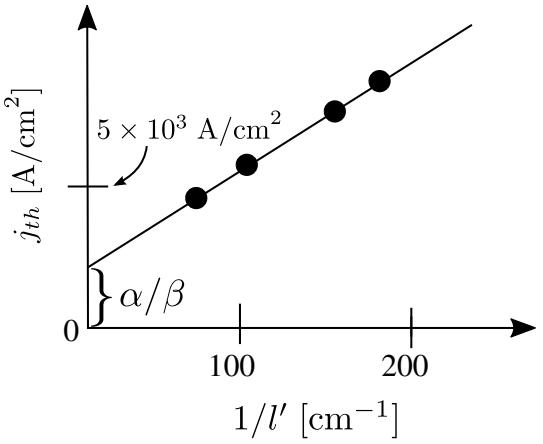
$$j_{th} = \underbrace{\frac{8\pi n^2\nu^2 ed}{c^2\eta g(\nu)}}_{\beta^{-1}}\left(\alpha - \frac{1}{l'}\ln R\right) \quad [\mathrm{A/cm^2}]$$

- Lasing threshold current density:

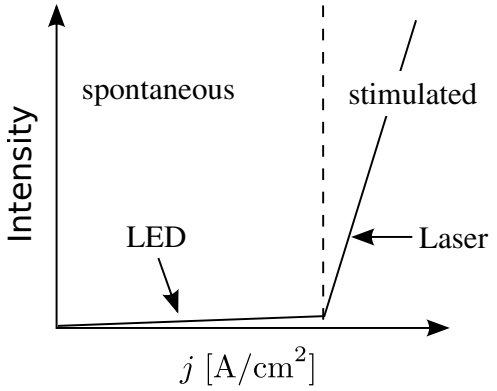
$$j_{th} = \frac{\alpha}{\beta} - \frac{1}{\beta l'}\ln R \;\; (R<1)$$

Lasing Threshold Condition:

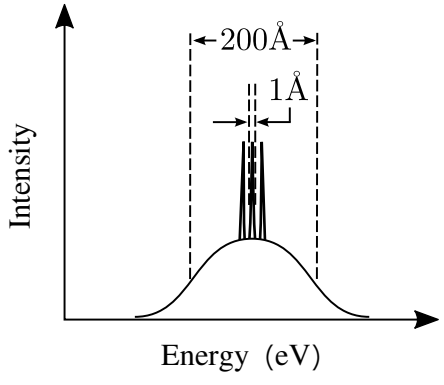
$$j_{th} = \frac{\alpha}{\beta} - \frac{1}{\beta l'}\ln R \;\; (R<1)$$



Lasing Threshold



Lasing Spectrum



129.10 Comparison of Emission Types

Emission	Intensity	Linewidth	Directional	Coherent
Spontaneous	Weak	Broad	No	No
Stimulated (Super-Radiant)	Intense	Narrow	Not Necessarily	No
Stimulated Lasing	Intense	Very Narrow	Yes	Yes

129.11 Cavities and Modes

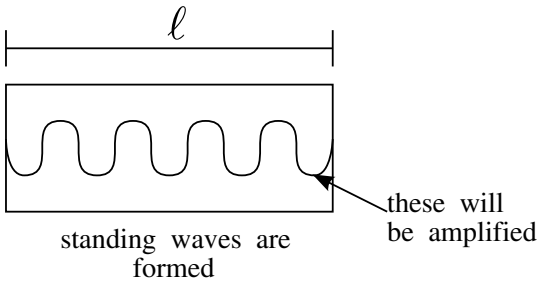
- Inside a cavity, standing waves are formed
- Light frequencies that are in phase after a round trip through the cavity will lase. This condition is given by

Integer number of wavelengths = Optical path length

$$m\lambda = 2nl$$

$$m\frac{\lambda}{2n} = l \quad m = 1, 2, 3...$$

- n is the refractive index of the cavity and l is the cavity length

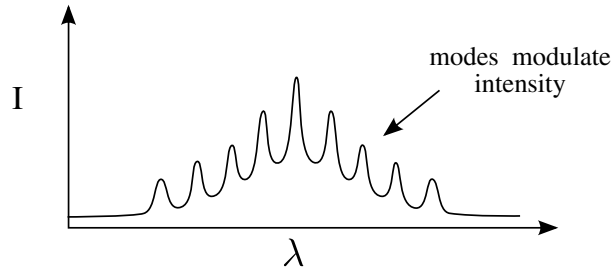


129.11.1 Longitudinal Laser Modes

- Modes whose condition for lasing depends on the length of the cavity are called longitudinal modes
- Wavelength spacing between modes:

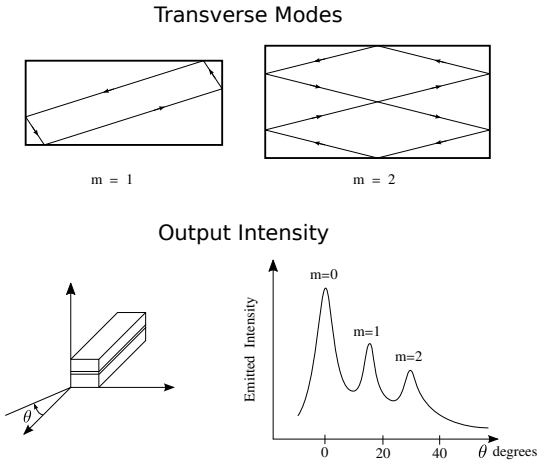
$$\Delta\lambda = \frac{\lambda^2}{2l\left(n - \lambda\frac{dn}{d\lambda}\right)}$$

- $dn/d\lambda$ is the dispersion of the refractive index



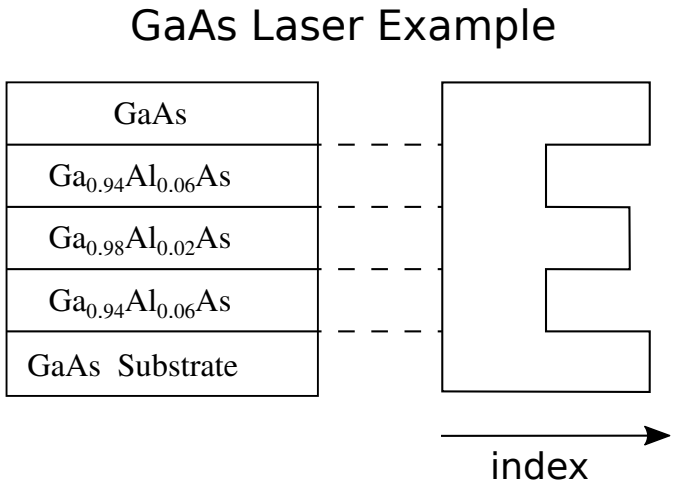
129.11.2 Transverse Laser Modes

- If the other facets are flat, then transverse modes are supported in the cavity.
- The effect of transverse modes is visible in the angular dependence of the laser output intensity



129.12 Semiconductor Lasers

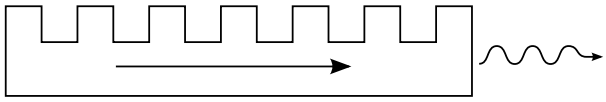
- Light is confined to regions of high refractive index
- Semiconductor heterostructure is designed to confine light in the junction



129.13 Photonic Bandgap Materials

- Light interacts with periodic structures with periodicity on the order of the wavelength
- Photonic crystals: structures with spatial periodic variation in the refractive index in 1, 2, or 3 dimensions. The periodicity is on the order of optical wavelengths.
- Photonic crystals can be engineered to slow down the propagation of optical pluses (“slow light”)
- The speed (v_g) of a pulse is described by the group index (n_g)

$$v_g = c/n_g \quad \text{as } n_g \to \infty, \quad v_g \to 0$$

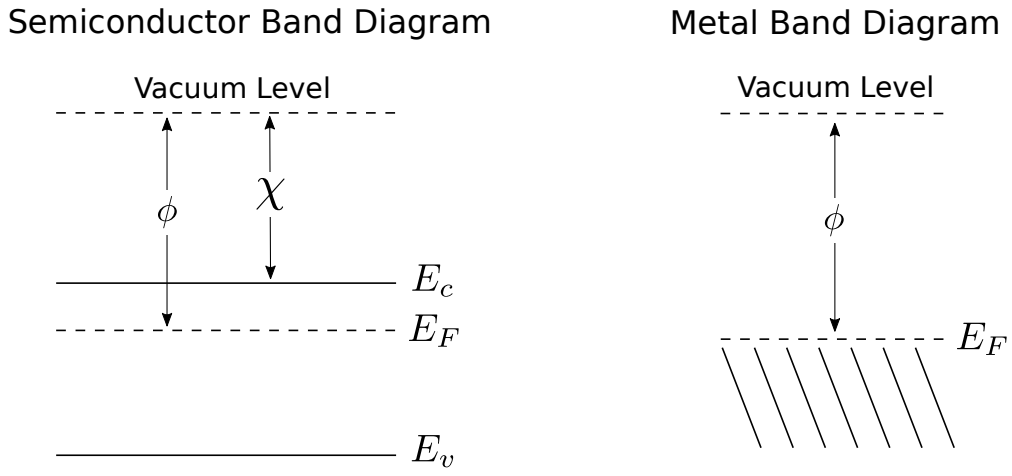


130 Band Diagrams

- Energy band levels are represented in real space
- Energy levels of interest: Conduction band edge, valence band edge, and Fermi Level
- Energies of interest: work function and electron affinity

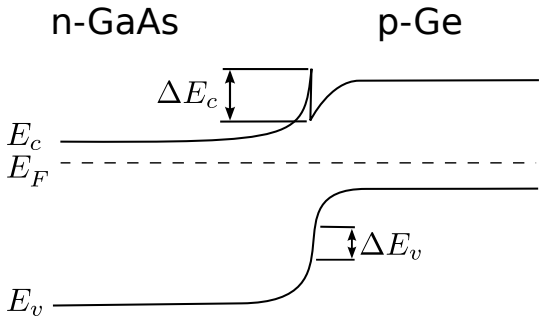
ϕ ≡Work Function. Energy required to move an electron from the fermi level to the vacuum level.
χ ≡Electron Affinity. Energy required to move an electron from the conduction band edge to the vacuum level.
E_c ≡Conduction band energy.
E_v ≡Valence band energy.
E_F ≡Fermi level.

130.1 Band Diagrams



130.2 Heterojunctions and the Anderson Model

- Heterojunction: metallurgical junction between two different types of semiconductors
- Anderson Model: junction between different semiconductors results in discontinuities in the energy bands.
- Example: n-GaAs ($E_g = 1.45$ eV) and p-Ge ($E_g = 0.7$ eV)



130.3 Band Bending at p-n Junctions

Band diagrams for isolated p- and n-type semiconductors are shown in Fig. 130.1below.
When brought into contact without application of an external bias field, thermal equilibrium requires that the Fermi levels equilibrate, i.e.

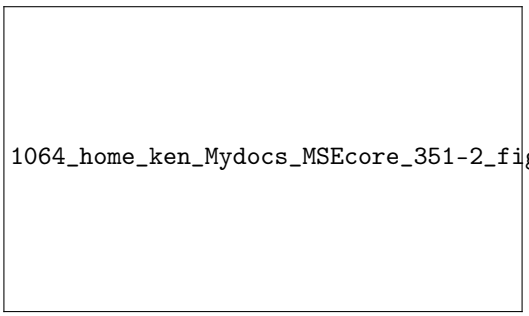


Figure 130.1: Band diagrams of isolated n-GaAs and p-Ge.

$$E_{F,n} = E_{F,p} \tag{130.1}$$

This results in a bending of the conduction and valence bands as shown in Fig. 130.2. With the Fermi levels

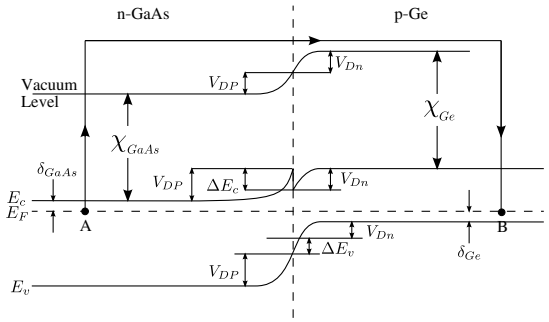


Figure 130.2: Band diagram of a n-GaAs/p-Ge heterojunction.

equilibrated, the following expression is true considering the path A-B in Fig. 130.2:

$$E_{F,n} - E_{F,p} = 0 = (\chi_{GaAs} + \delta_{GaAs} + qV_{D,n} + qV_{D,p}) - (\chi_{Ge} + E_{g,Ge} - \delta_{Ge}) \tag{130.2}$$

Rearranging Eqn. 130.2,

$$q(V_{D,n} + V_{D,p}) = \underbrace{(\chi_{Ge} - \delta_{Ge} + E_{g,Ge})}_{p} - \underbrace{(\chi_{GaAs} + \delta_{GaAs})}_{n} \tag{130.3}$$

Substituting in known values,

$$\begin{aligned} q(V_{D,n} + V_{D,p}) &= (4.13 - 0.1 + 0.7) - (4.07 + 0.1) \\ &= 0.56\,\text{eV} < E_{g,Ge} \end{aligned} \tag{130.4}$$

where

$$\frac{V_{D,n}}{V_{D,p}} = \frac{N_A \epsilon_{Ge}}{N_D \epsilon_{GaAs}} \tag{130.5}$$

130.3.1 Calculate the Conduction Band Discontinuity

Let the Fermi level be zero on the energy scale.

$$\delta_{GaAs} + qV_{D,n} = (E_{g,Ge} - \delta_{Ge}) - qV_{D,p} + \Delta E_c$$

$$\Delta E_c = \delta_{GaAs} + qV_{D,n} - (E_{g,Ge} - \delta_{Ge}) + qV_{D,p}$$

On substituting in Eqn. 130.3, we get

$$\Delta E_c = \chi_{Ge} - \chi_{GaAs} = +0.6$$

The positive value for ΔE_c indicates a spike in the conduction band.

130.3.2 Calculate the Valence Band Discontinuity

The valence band discontinuity ΔE_v is calculated in a similar manner. In this case, a negative value for ΔE_v indicates a spike in the valence band. ΔE_v is calculated as follows:

$$\Delta E_v = \Delta E_g - \Delta\chi$$

$$\Delta E_v = (E_{g,GaAs} - E_{g,Ge}) - (\chi_{Ge} - \chi_{GaAs})$$

Substituting in known values,

$$\Delta E_v = (1.45 - 0.7) - (4.13 - 4.07)$$

$$\Delta E_v = +0.69$$

Since ΔE_v is positive, there is no spike in the valence band.

130.3.3 Example: p-GaAs/n-Ge

The isolated band diagrams for p-GaAs and n-Ge are shown in Figs. 130.3a and 130.3b. Here, we calculate the

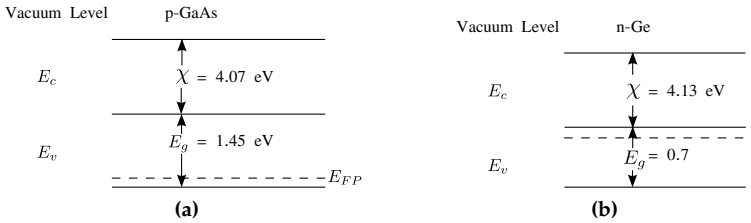


Figure 130.3: Isolated band diagrams of (a) p-GaAs and (b) n-Ge.

conduction and valence band discontinuities using the approaches outlined in Sections 130.3.1 and 130.3.2. For the conduction band discontinuity,

$$q(V_{D,n} + V_{D,p}) = \underbrace{(\chi_{GaAs} + E_{g,GaAs} - \delta_{GaAs})}_{p} - \underbrace{(\chi_{Ge} + \delta_{Ge})}_{n}$$

$$= (4.07 + 1.45 - 0.1) - (4.13 + 0.1)$$

$$= 1.19 < E_{g,GaAs}$$

For calculating the valence band discontinuity,

$$\Delta E_v = (E_{g,Ge} - E_{g,GaAs}) - (\chi_{GaAs} - \chi_{Ge})$$

$$\Delta E_v = -0.69 \Rightarrow \text{ spike} \tag{130.6}$$

The resulting band diagram is shown in Fig. 130.4.

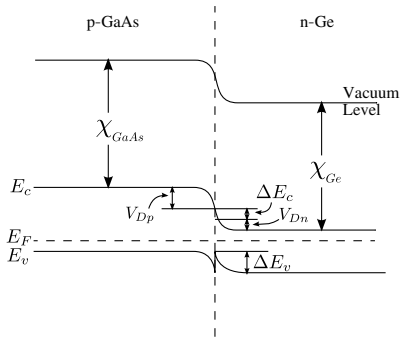


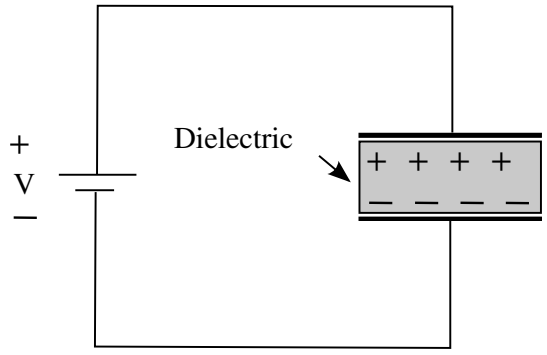
Figure 130.4: Band diagram of a p-GaAs/n-Ge heterojunction as predicted using the approach in Section130.3.3. Note that there is a spike in the valence band which impedes the conduction of holes.

131 Dielectric Materials

- Dielectrics are insulating materials which include glasses, ceramics, polymers, and ferroelectrics
- Dielectric constant (ϵ) describes how well a material stores charge
- ϵ is related to capacitance (C)

Parallel Plate Capacitance

$$\begin{aligned} C &= \frac{\epsilon A}{d} \\ d &\equiv \text{distance between plates} \\ A &\equiv \text{plate area} \end{aligned}$$



131.1 Macroscopic Dielectric Theory

- From Maxwell's equations,

$$D = \epsilon \mathcal{E} \quad [\text{C/m}^2]$$

$$\begin{aligned} D &\equiv \text{Electric flux density} \\ \epsilon &\equiv \text{Dielectric constant} \\ \mathcal{E} &\equiv \text{Electric field} \end{aligned}$$

- Note: $\epsilon = \epsilon_r \epsilon_0$

$$\begin{aligned} \epsilon_r &\equiv \text{Relative dielectric constant (unitless)} \\ \epsilon_0 &\equiv \text{Permittivity of free space [F/m]} \end{aligned}$$

- Total charge stored in a capacitor in vacuum:

$$Q = \epsilon_0 \frac{V}{d} \quad [C]$$

with

$$\begin{aligned} V &\equiv \text{Voltage [V]} \\ d &\equiv \text{Distance [m]} \end{aligned}$$

- Putting in a dielectric,

$$Q' = \epsilon_0 \epsilon_r \frac{V}{d} \quad \text{and} \quad C = Q'/V$$

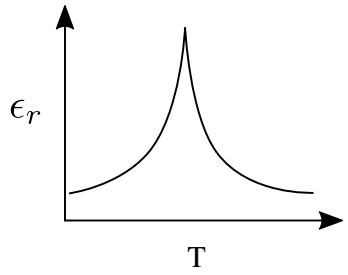
- An increase in charge is stored in the capacitor with insertion of the dielectric
- D (electric flux density) is equal to the surface charge
- P (polarization density) is the additional surface charge

$$\begin{aligned} P &\equiv D - \epsilon_0 \mathcal{E} \\ P &\equiv (\epsilon_r \epsilon_0 - \epsilon_0)\mathcal{E} = \epsilon_0 \underbrace{(\epsilon_r - 1)}_{\chi} \mathcal{E} \\ P &\equiv \epsilon_0 \chi \mathcal{E} \end{aligned}$$

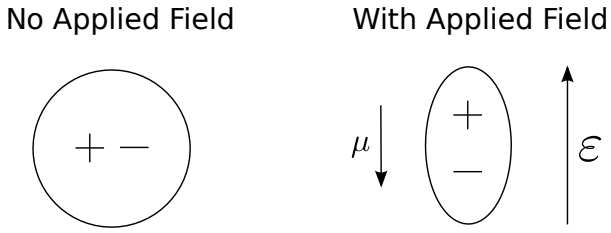
- $\chi = (\epsilon_r - 1)$ is the dielectric susceptibility

131.2 Relation of ϵ to the Microscopic Structure

- Knowledge of the microscopic structure is needed to explain dielectric phenomena. Example: ferroelectric $\epsilon_r(T)$



- In the microscopic approach, we consider 1) Atomic behavior and 2) Deformation of the atomic charge cloud (orbitals) when a field is applied.



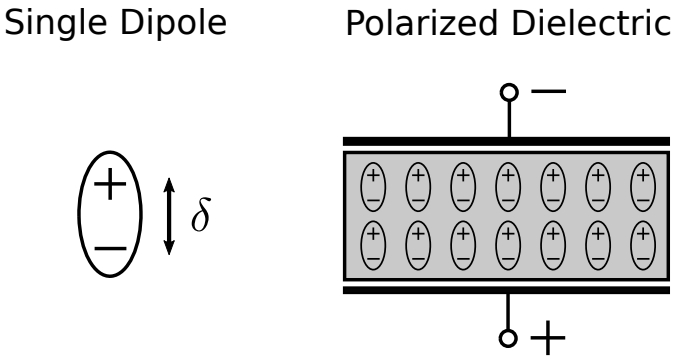
131.2.1 Relation of Macroscopic to Microscopic

- Induced dipole moment due to an applied field: $\mu = q\delta$
- For N_m molecules per unit volume, the polarization density is $P = N_m q\delta$ or

$$\boxed{P = N_m \mu}$$

- For small fields, $\mu = \alpha \mathcal{E}'$ and $P = N_m \alpha \mathcal{E}'$

$\mathcal{E}'$ $\equiv$ Local electric field
α $\equiv$ Polarizability of species (atoms, molecules)



131.3 Contributions to the Polarizability

(1) Electronic: Deformation of the electronic cloud (orbitals) due to optical fields.

(2) Molecular: Bonds between atoms can stretch, bend, and rotate

(3) Orientational: Polymers, liquids, and gases can be reoriented in an applied electric field. Example: poled polymers, useful for electro-optic devices.

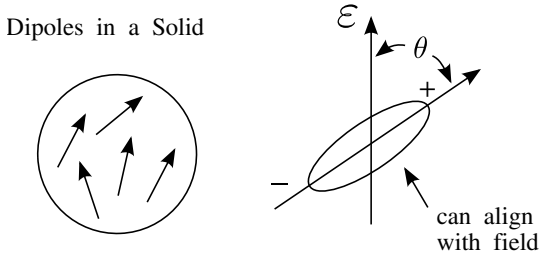
131.4 Polarization in Solids

- In an applied field, some molecules will be aligned, others will not be completely aligned. The orientation energy is:

$$E = -\mu \mathcal{E} \cos\theta$$

$\theta \equiv$ Angle between μ and $\mathcal{E}$

- The number of dipoles with a certain E will depend on the Boltzmann factor $\exp(-E/kT)$. Increasing temperature will favor more random alignment.



131.5 Calculation of the Average Dipole Moment

- Average Dipole Moment ($\langle\mu\rangle$)

$$\langle\mu\rangle = \frac{\text{weighted average energy for each orientation}}{\text{total energy for all orientations}}$$

- Consider the 3-dimensional case. In a sphere with radius r, the total number of molecules with energy E in a volume element $2\pi r^2 \sin\theta d\theta$ is

$$\#\text{ of dipoles with energy } E = 2\pi A \sin\theta d\theta \exp(-E/kT)$$

$$A = \text{Constant including radius}$$

- Recall the orientation energy is given by $E = -\mu\mathcal{E}\cos\theta$. Substituting in for E,

$$\#\text{ of dipoles with energy } E = 2\pi A \sin\theta d\theta \exp\left(\frac{\mu\mathcal{E}\cos\theta}{kT}\right)$$

- The total energy for all orientations is then

$$\int_0^\pi A \exp\left(\frac{\mu\mathcal{E}\cos\theta}{kT}\right) 2\pi \sin\theta d\theta$$

- The weighted average for each orientation is

$$\int_0^\pi A \exp\left(\frac{\mu\mathcal{E}\cos\theta}{kT}\right) (\mu\cos\theta)\, 2\pi \sin\theta d\theta$$

- The average dipole moment is then

$$\langle\mu\rangle = \frac{\int_0^\pi A \exp\left(\frac{\mu\mathcal{E}\cos\theta}{kT}\right) (\mu\cos\theta)\, 2\pi \sin\theta d\theta}{\int_0^\pi A \exp\left(\frac{\mu\mathcal{E}\cos\theta}{kT}\right) 2\pi \sin\theta d\theta}$$

- Integrating,

$$\frac{\langle\mu\rangle}{\mu} = L(a) = \coth a - \frac{1}{a}$$

$$\text{with } a = \frac{\mu\mathcal{E}}{kT}$$

- For small a, $\coth a \approx \frac{1}{a} + \frac{a}{3} + \cdots$

$$\therefore \frac{\langle\mu\rangle}{\mu} = \frac{a}{3} \quad \text{for small a}$$

$$\boxed{\langle\mu\rangle = \frac{\mu^2\mathcal{E}}{3kT}} \quad \text{or} \quad \boxed{\frac{\langle\mu\rangle}{\mu} = \frac{\mu\mathcal{E}}{3kT}}$$

131.6 Polarizability of Solids

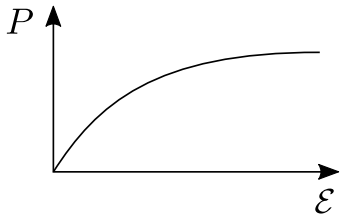
- The polarizability of the solid is

$$\boxed{P = N_m \langle \mu \rangle}$$

- P is proportional to the field $\mathcal{E}$ and inversely proportional to the temperature T

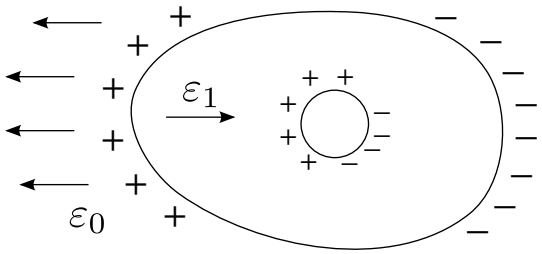
$$P = N_m \frac{\mu^2 \mathcal{E}}{3kT}$$

- The definition for the electric susceptibility is $\chi = \frac{\partial P}{\partial \mathcal{E}} \Rightarrow \chi \propto \frac{1}{T}$



131.7 Dielectric Constant for a Solid

- Under an applied field $\mathcal{E}_0$, induced dipoles in the solid create an opposing field $\mathcal{E}_1$. The field inside the material is $\boxed{\mathcal{E} = \mathcal{E}_0 + \mathcal{E}_1}$.



- For a crystal, the polarizability will depend on the interaction of dipoles in the lattice

$$\boxed{P = \sum_j N_j P_j = \sum_j N_j \alpha_j \mathcal{E}_{loc,j}}$$

- $\mathcal{E}_{loc,j}$ is the local field of atom j due to interactions with other dipoles. For cubic materials,

$$\mathcal{E}_{loc} = \mathcal{E} + \frac{P}{3\epsilon_0}$$

- The polarizability is then

$$P = \left(\sum_j N_j \alpha_j \right) \left(\mathcal{E} + \frac{P}{3\epsilon_0} \right)$$

131.8 Claussius Mossotti Relation

- From the relation for the macroscopic electric susceptibility (see slide 131.1), $\chi = P/\epsilon_0 \mathcal{E}$

$$\chi = \left(\sum_j N_j \alpha_j \right) \frac{\mathcal{E} + \frac{P}{3\epsilon_0}}{\epsilon_0 \mathcal{E}}$$

- Recall that $\chi = \epsilon_r - 1$ and $P = \chi \epsilon_0 \mathcal{E} = \epsilon_0 \mathcal{E}(\epsilon_r - 1)$

$$\epsilon_r - 1 = \left(\sum_j N_j \alpha_j\right) \frac{\mathcal{E} + \frac{1}{3\epsilon_0}\epsilon_0 \mathcal{E}(\epsilon_r - 1)}{\epsilon_0 \mathcal{E}}$$

Simplifying and rearranging gives the Claussius Mossotti Relation:

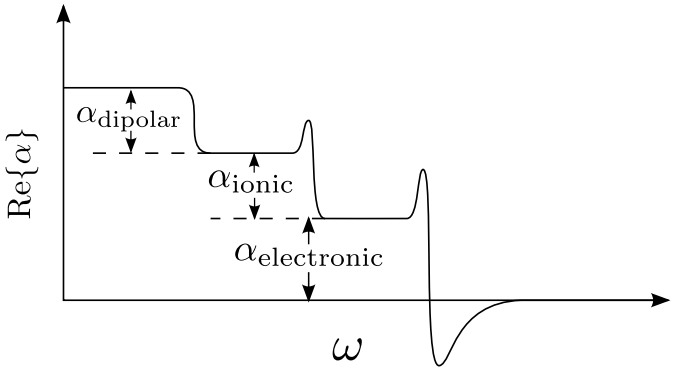
$$\boxed{\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{1}{3\epsilon_0}\sum_j N_j \alpha_j} \quad \text{(SI Units)}$$

- The Claussius Mossotti relation relates the dielectric constant to the atomic polarizability

131.9 Frequency Dependence of the Polarizability

- Dipolar, ionic, and electronic contributions to the polarizability

$$\alpha(\omega) = \sum \alpha_i(\omega)$$



- Classical approach: treat electrons as harmonic oscillators
- Apply Hooke's Law

$$e\mathcal{E}_{loc} = Cx = m\omega_0^2 x$$

$$C = m\omega_0^2 \equiv \text{Force constant}$$

$$\mathcal{E}_{loc} = \frac{m\omega_0^2 x}{e}$$

- Note: ω_0 is the natural or resonance frequency
- The electronic contribution to the polarizability is $P = N_m \alpha_{el} \mathcal{E}_{loc}$. Recall that $P = N_m q\delta = Nex$

$$\alpha_{el} = \frac{P}{N\mathcal{E}_{loc}} = \frac{Nex}{N\mathcal{E}_{loc}} = \frac{e^2}{m\omega_0^2}$$

- For an applied field with frequency ω, the equation of motion for the electron response is

$$m\frac{d^2x}{dt^2} + m\omega_0^2 x = -e\mathcal{E}_{loc}\sin\omega t$$

- The solution is $x = x_0 \sin\omega t$. Substituting the solution into the equation of motion,

$$m(-\omega^2 + \omega_0^2)x_0 = -e\mathcal{E}_{loc}$$

- Solving for x_0

$$\boxed{x_0 = \frac{-e\mathcal{E}_{loc}}{m(\omega_0^2 - \omega^2)}}$$

- The electronic dipole moment is

$$\mu_{elec} = -ex_0 = \frac{e^2\mathcal{E}_{loc}}{m(\omega_0^2 - \omega^2)}$$

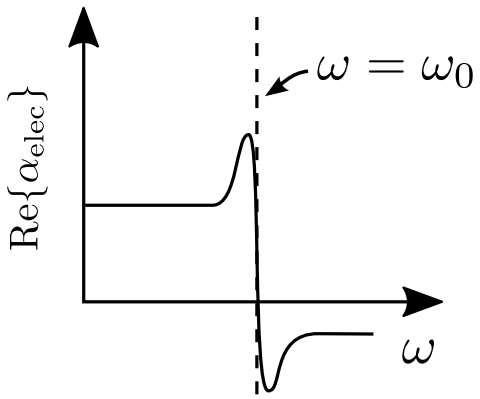
- The electronic polarizability is

$$\alpha_{elec} = \frac{\mu_{elec}}{\mathcal{E}_{loc}} = \frac{e^2}{m(\omega_0^2 - \omega^2)}$$

- α_{elec} becomes large at resonances where $\omega_0^2 - \omega^2 \to 0$
- Calculate the atomic polarizability:

$$\mu_0 = -ex_0 = \frac{e^2\mathcal{E}_{loc}}{m(\omega_0^2 - \omega^2)}$$

$$\alpha_{elec} = \frac{\mu_0}{\mathcal{E}_{loc}} = \frac{e^2}{m(\omega_0^2 - \omega^2)}$$



131.10 Quantum Theory of Polarizability

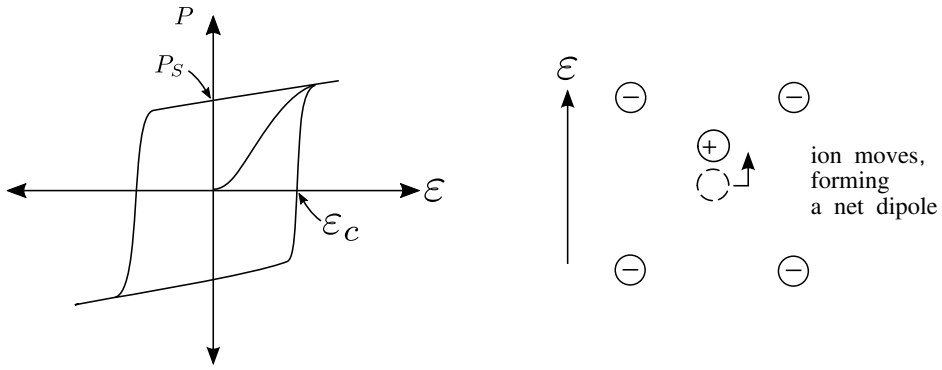
$$\alpha_{elec} = \frac{e^2}{m}\sum_j \frac{f_{ij}}{\omega_0^2 - \omega^2}$$

$f_{ij} \equiv$ Oscillator strength for a transition from state i to j

$$P_{elec} = N\alpha_{elec}\mathcal{E}$$

131.11 Advanced Dielectrics: Ferroelectrics

- Ferroelectrics: materials that are non-centrosymmetric and have a resulting net polarization.
- The polarization is a function of the applied field and displays hysteresis.

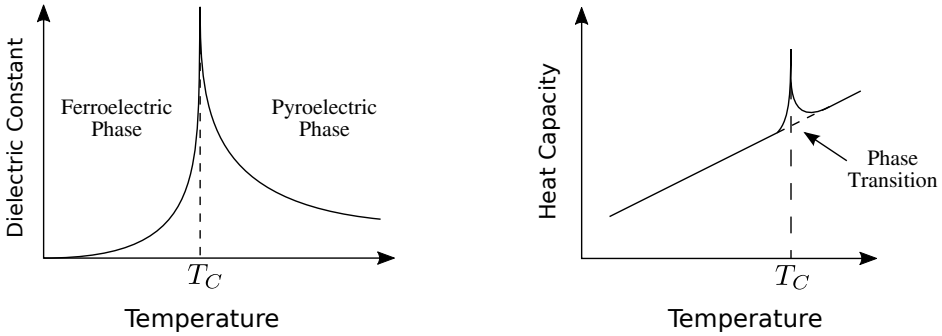


$P_S \equiv$ Spontaneous polarization
$\mathcal{E}_C \equiv$ Coercive field

132 Phase Transitions

- The transition from a ferroelectric (net polarization) to pyroelectric (no net polarization) structure occurs at the Curie Temperature (T_C)
- Dielectric constant:

$$\epsilon = \frac{\xi}{T - T_C} \quad \xi \equiv \text{Curie constant}$$



132.1 Lattice Instabilities

- Recall the Clausius-Mossotti relation:

$$\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{1}{3\epsilon_0} N\alpha$$

- Solving for ϵ_r,

$$\epsilon_r - 1 = \frac{N\alpha}{3\epsilon_0}(\epsilon_r + 2) = \frac{N\alpha\epsilon_r}{3\epsilon_0} + \frac{2}{3}\frac{N\alpha}{\epsilon}$$

$$\epsilon_r \left(1 - \frac{N\alpha}{3\epsilon_0}\right) = \frac{2N\alpha}{3\epsilon_0} + 1$$

$$\boxed{\epsilon_r = \frac{2N\alpha/\epsilon_0 + 3}{3 - N\alpha/\epsilon_0}}$$

- Singularity (polarization catastrophe) at the condition $N\alpha/\epsilon_0 = 3$

132.2 Curie Weiss Law

- What happens close to the singularity? Consider a small deviation $2s$

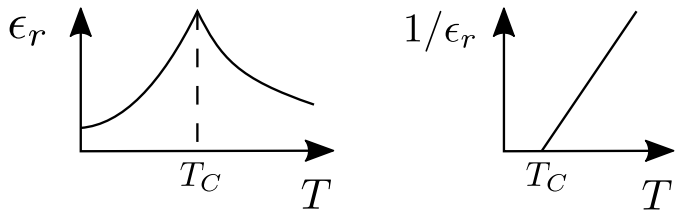
$$\frac{N\alpha}{\epsilon_0} \cong 3 - 2s$$

- Rewriting the equation for ϵ_r,

$$\epsilon_r = \frac{2N\alpha/\epsilon_0 + 3}{3 - N\alpha/\epsilon_0 - 2s} \propto \frac{1}{s} \quad \text{near the singularity}$$

- Near the critical temperature, suppose $s \propto T$ such that $s \cong (T - T_C)/\xi$

$$\therefore \epsilon_r \cong \frac{\xi}{(T - T_C)} \quad \text{or} \quad \frac{1}{\epsilon_r} \cong \frac{T - T_C}{\xi}$$



132.3 Ferroelectric Phase Transitions: Landau Theory

- Consider the Helmholtz free energy $\hat{F}$ for a ferroelectric material as a function of polarization (P), temperature (T), and applied field ($\mathcal{E}$).
- Taking a Taylor expansion in the order parameter s about $s = 0$,

$$\hat{F}(P, T, \mathcal{E}) = -\mathcal{E} \cdot P + s_0 + \frac{1}{2}s_2P^2 + \frac{1}{4}s_4P^4 + \frac{1}{6}s_6P^6 + \cdots$$

- Note that there are no odd powers if the unpolarized crystal has a center of inversion.
- At thermal equilibrium, the minimum in $\hat{F}$ with respect to P is

$$\frac{\partial \hat{F}}{\partial P} = 0 = -E + s_2P + s_4P^3 + s_6P^5$$

- Thermal equilibrium condition:

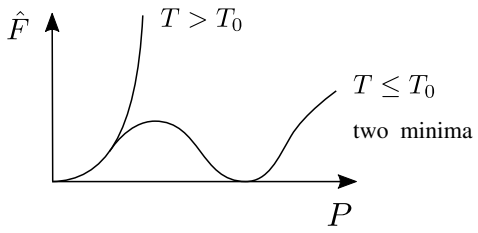
$$\frac{\partial \hat{F}}{\partial P} = 0 = -E + s_2P + s_4P^3 + s_6P^5$$

- For a ferroelectric state transition, the coefficient of the first order P term must pass through zero at some temperature T_0 for no applied field.

$$\therefore \text{assume } s_2 = \gamma(T - T_0), \quad \text{with } \gamma \text{ a positive constant}$$

- s_2 can be positive or negative:
- $s_2 < 0$: lattice is "soft" close to the instability

- $s_2 > 0$: unpolarized lattice is unstable



- Rewriting the equilibrium condition,

$$\frac{\partial \hat{F}}{\partial P} = -\mathcal{E} + \gamma(T - T_0)P + s_4 P^3 = 0$$

- Second order transition: no volume change, smooth transition.
- For $\mathcal{E} = 0$ and $P = P_S$ at equilibrium,

$$\gamma(T - T_0)P_S + s_4 P_S^3 = 0$$

$$P_S \left[\gamma(T - T_0) + s_4 P_S^2\right] = 0$$

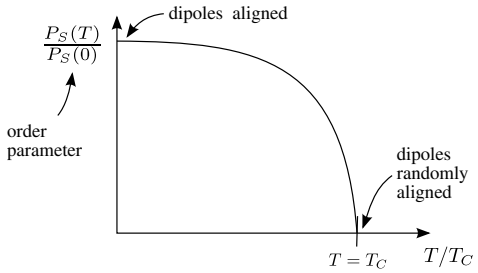
- Solutions: $P_S = 0$, or
- For $T \geq T_0$: $P_S^2 = (\gamma/s_4)(T_0 - T) \Rightarrow P_S$ is imaginary; the only real root is $P_S = 0$.
- For $T < T_0$:

$$\boxed{|P_S| = (\gamma/s_4)^{1/2} (T_0 - T)^{1/2}}$$

- Second order transition for $T < T_0$:

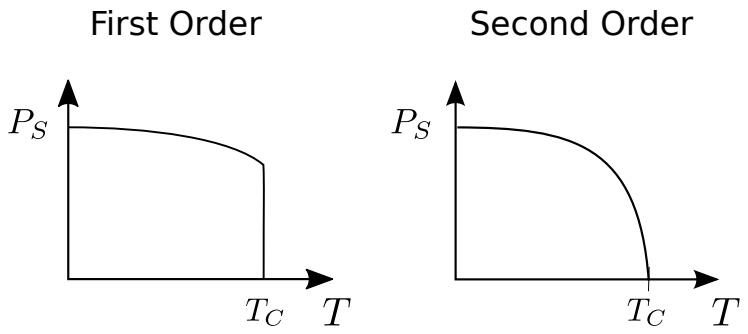
$$P_S| = (\gamma/s_4)^{1/2} (T_0 - T)^{1/2}$$

- The phase transition is second order since the polarization goes continuously to zero



132.3.1 First versus Second Order Transitions

- The change in P_s is discontinuous for first order transitions and continuous for second order transitions
- First order transition results from a structure change. Example: cubic (pyroelectric) → tetragonal (ferroelectric)



- Recall the Helmholtz free energy:

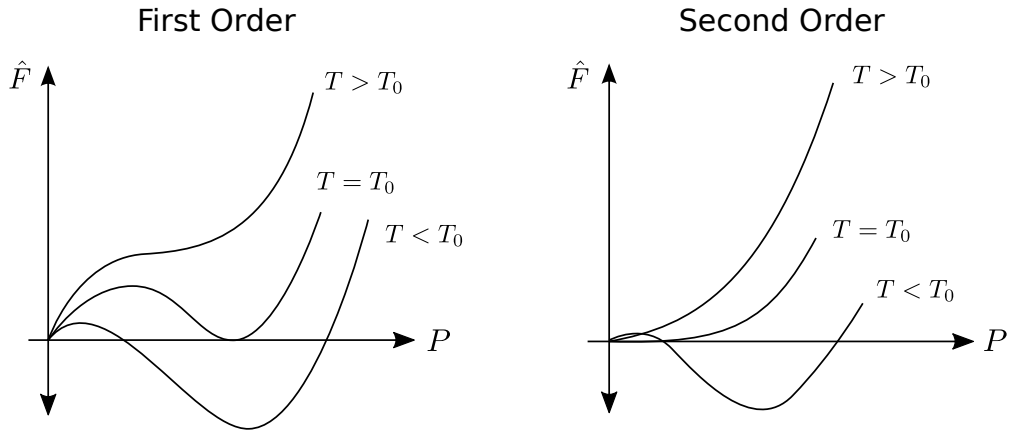
$$\hat{F} = -\mathcal{E} \cdot P + s_0 + \frac{1}{2}s_2P^2 + \frac{1}{4}s_4P^4 + \frac{1}{6}s_6P^6 + \cdots$$

- For first order transitions, s_4 is negative and s_6 cannot be neglected
- Equilibrium condition

$$\frac{\partial \hat{F}}{\partial P} = \gamma(T - T_0)P_S - |s_4|P_S^3 + s_6P_S^5 = 0$$

- Solutions: $P_S = 0$ or

$$\gamma(T - T_0) - |s_4|P_S^2 + s_6P_S^4 = 0$$

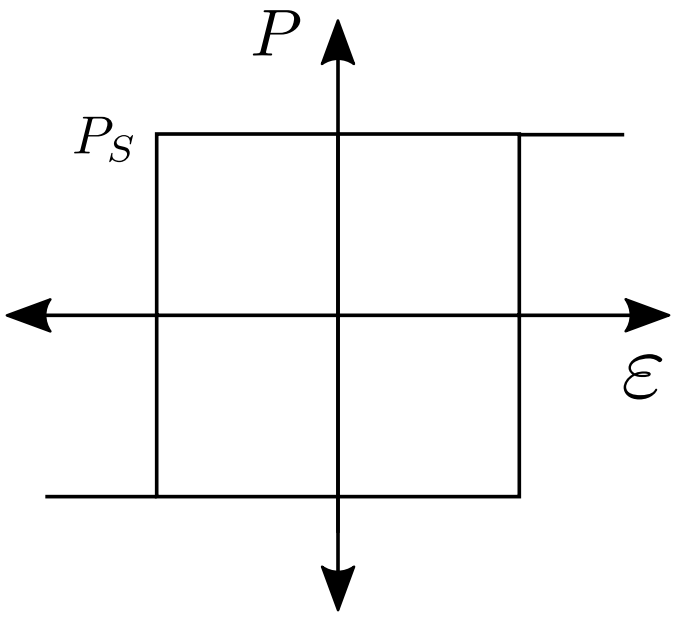


132.3.2 Ferroelectric Example: $BaTiO_3$

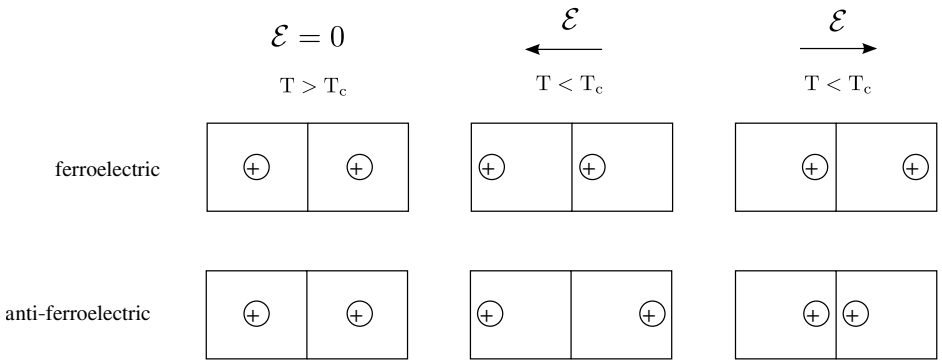
- Spontaneous polarization: $P_s = N_m q \delta = 26\mu C/cm^2 \approx 3 \times 10^{-1} C \cdot m^{-2}$
- Dipole moment of a unit cell: $\mu = P_s \cdot$ volume. For a lattice constant $a \approx 0.4\,nm$

$$\mu = (3 \times 10^{-1} C \cdot m^{-2}) \times 64 \times 10^{-30} m^3 \approx 2 \times 10^{-29} C \cdot m$$

Material	T_C (K)
$BaTiO_3$	408
$BaTiO_3$	1480



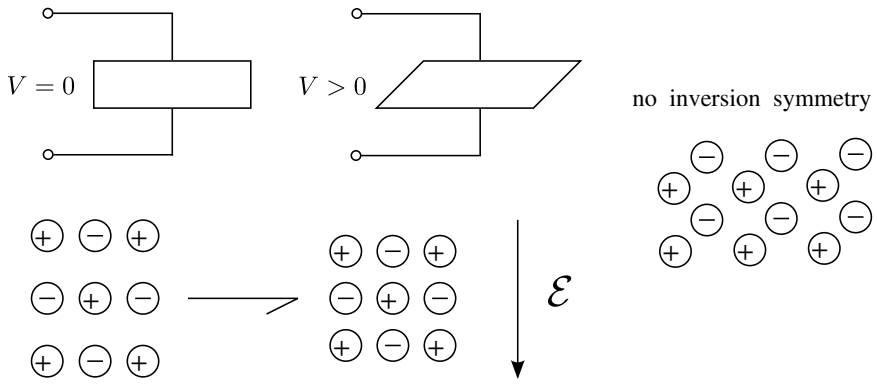
132.4 Other Instabilities



- Pyroelectrics: heat causes distortion and electrical signal, "pyroelectric detectors"

132.5 Piezoelectrics

- Piezoelectric materials: an applied electric field causes a mechanical deformation
- Piezoelectric materials lack inversion symmetry



- Stress and displacement:

$$\begin{aligned} \text{Stress:} \quad T &= CS - e\mathcal{E} \\ \text{Displacement:} \quad D &= \epsilon\mathcal{E} + eS \end{aligned}$$

$$\begin{aligned} C &\equiv \text{Elastic constant} \\ e &\equiv \text{Piezoelectric coefficient} \\ \mathcal{E} &\equiv \text{Electric field} \\ S &\equiv \text{Strain} \end{aligned}$$

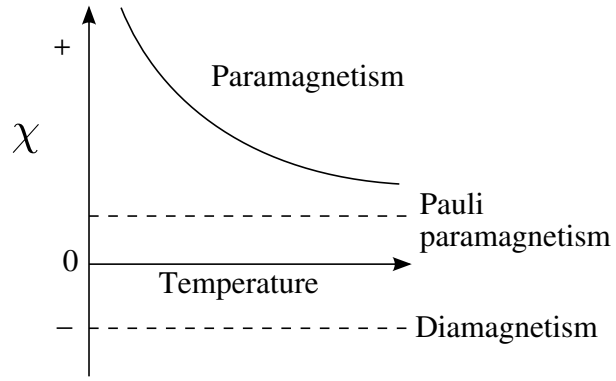
- Note: when $\mathcal{E} = 0$, $D \neq 0$ for non-zero e and S.
- Hooke's Law for $\mathcal{E} = 0$.
- Applications: electro-mechanical transducers, microphones, micro-motors, micro-electro-mechanical structures (MEMS), nano-electro-mechanical structures (NEMS).

133 Diamagnetism and Paramagnetism

- Current and magnetism related through Maxwell's Equations
- Magnetic susceptibility:

$$\boxed{\chi = \frac{\mu_0 M}{B}}$$

$$\begin{aligned} M &\equiv \text{Magnetization} \\ B &\equiv \text{Magnetic field intensity} \\ \mu_0 &\equiv \text{Permeability of free space} \end{aligned}$$

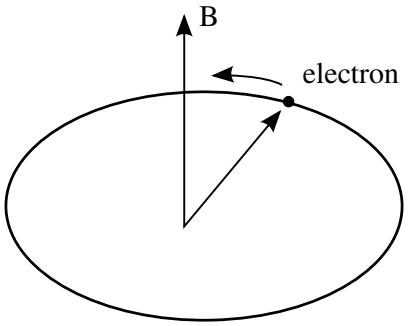


133.0.1 Diamagnetism

- In a magnetic field, induced current field opposite to applied current
- Electron precesses around field axis with frequency

$$\omega = \frac{eB}{2m}[\text{sec}^{-1}]$$

- Magnetic field will induce net electric current around the nucleus.



- Consider multi-electron atom. The "electron current" for Z electrons is

$$\begin{aligned} I &= (\text{charge}) \times (\text{revolutions per second}) = -Ze\frac{\omega}{2\pi} \\ I &= -Ze\left(\frac{1}{2\pi}\frac{eB}{2m}\right) = -\frac{Ze^2B}{4\pi m} \end{aligned}$$

- Magnetic moment:

$$\mu \equiv (\text{current}) \times (\text{area of loop}) = I \times A$$

- For a loop area $A = \pi\rho^2$,

$$\mu = -\frac{Ze^2B}{4m}\langle\rho^2\rangle$$

- $\langle\rho^2\rangle$ is the mean square of the perpendicular distance from the field axis through charge

$$\langle\rho^2\rangle = \langle x^2\rangle + \langle y^2\rangle$$

- The atomic radius is $\langle r^2\rangle = \langle x^2\rangle + \langle y^2\rangle + \langle z^2\rangle$

$$\therefore \langle\rho^2\rangle = \frac{2}{3}\langle r^2\rangle$$

- Diamagnetic susceptibility of electrons: $M = N\mu$

$$\text{Langevin Result: } \boxed{\chi = \frac{N\mu_0\mu}{B} = -\frac{\mu_0 NZe^2}{6m}\langle r^2\rangle}$$

(note negative sign)

- Note: need to calculate $\langle r^2\rangle$ from quantum mechanics

133.1 Paramagnetism

- Positive χ occurs for
- atoms, molecules, and lattice defects with an odd number of electrons
- free atoms with partially filled inner shells
- metals
- Magnetic moment μ:

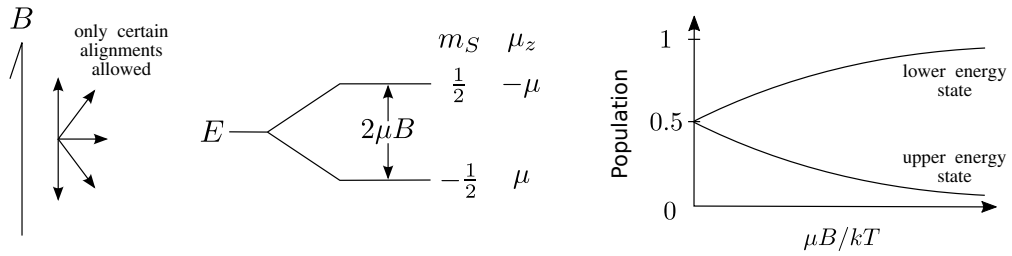
$$\mu \equiv \gamma\hbar J = -g\mu_B J$$

$\gamma \equiv$ gyromagnetic ratio
$\hbar J \equiv$ angular momentum
$g \equiv$ g factor ($\sim$2.00 for electron spin)
$\mu_B \equiv$ Bohr magneton (atomic unit)
$J \equiv$ angular quantum number for electrons

- From quantum mechanics, only certain alignments of μ with B are allowed
- Consider single electron case. Energy U of an electron in a magnetic field:

$$\boxed{U = -\vec{\mu} \cdot \vec{B} = m_J g \mu_B B}$$

- For single electron atom with no orbital momentum, $m_J = m_S = \pm 1/2$



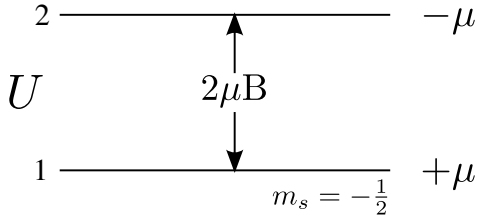
- For a two spin orientation case, $m_J = \pm\frac{1}{2}$, $g = 2$, and $U = \pm\mu_B B$
- Probability of occupying state i:

$$\boxed{P_i = \frac{\exp(-E_i/\tau)}{Z}}$$

$Z \equiv$ sum over all states (normalization)
$\tau \equiv kT$

- For N total atoms,

$$\frac{N_1}{N} = \frac{\exp(\mu B/\tau)}{\exp(\mu B/\tau) + \exp(-\mu B/\tau)}$$
$$\frac{N_2}{N} = \frac{\exp(-\mu B/\tau)}{\exp(\mu B/\tau) + \exp(-\mu B/\tau)}$$



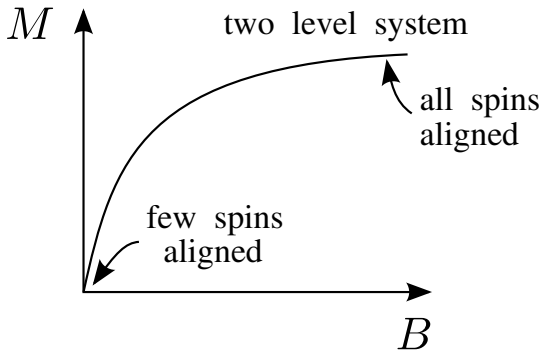
- Magnetization $M \equiv$ (net spin density $\|$B) $\times$ (magnetic moment per electron)

$$M = (N_1 - N_2)\mu = N\mu \frac{e^x - e^{-x}}{e^x + e^{-x}} = N\mu \tanh x$$

$$x \equiv \mu B/kT$$

- For $x \ll 1$: high T, small field, no interaction.

$$\tanh x \approx x \Rightarrow M \cong N\mu(\mu B/kT)$$



- For multiple electron atoms, the azimuthal quantum number m_J takes values of $J, J-1, \ldots, -J$
- For an atom with quantum number J, there are $2J+1$ energy levels
- Curie-Brillouin Law:

$$M = NgJ\mu_B B_J(x)$$

$$x \equiv gJ\mu_B B/kT$$

- Brillouin function (obtained from $\langle\mu\rangle$ over all quantum configurations)

$$B_J(x)\frac{2J+1}{2J}\coth\left[\frac{(2J+1)x}{2J}\right] - \frac{1}{2J}\coth\frac{x}{2J}$$

- For $x \ll 1$

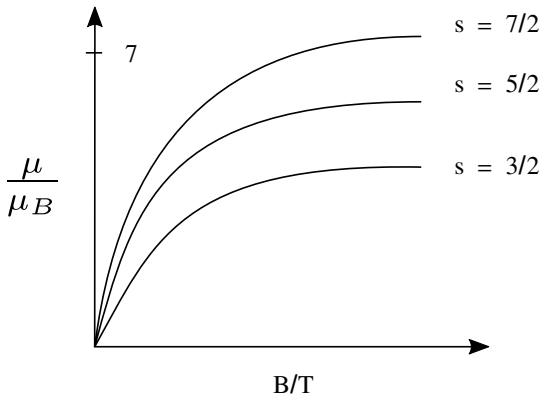
$$\boxed{\coth x \approx \frac{1}{x} + \frac{x}{3}}$$

133.1.1 Calculation of Susceptibility

For small fields and high temperatures,

$$\chi = \frac{M}{B} \cong \frac{NJ(J+1)g^2\mu_B^2}{3kT} = \frac{Np^2\mu_B^2}{3kT} = C/T$$

$$\begin{aligned} p^2 &= g^2 J(J+1) \Rightarrow p = g\left[J(J+1)\right]^{1/2} \\ C &\equiv \text{Curie constant} \end{aligned}$$



133.1.2 Calculation of Total Angular Momentum J

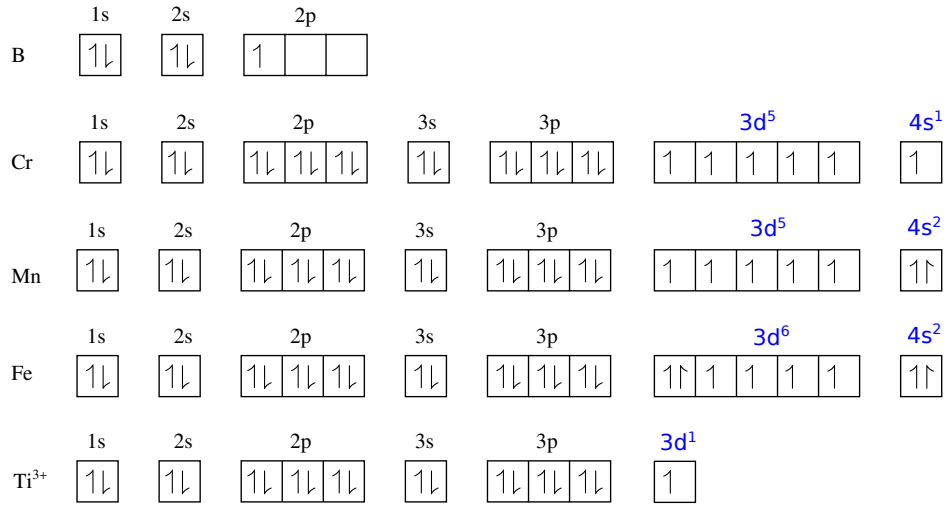
- Atoms with filled shells have no magnetic moment. Example:

$$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10}$$

- The spins arrange themselves so as to give the maximum possible S consistent with Hund's rule
- Pauli Principle: no two electrons in the same system can have the same quantum numbers n, l, m_l, and m_s

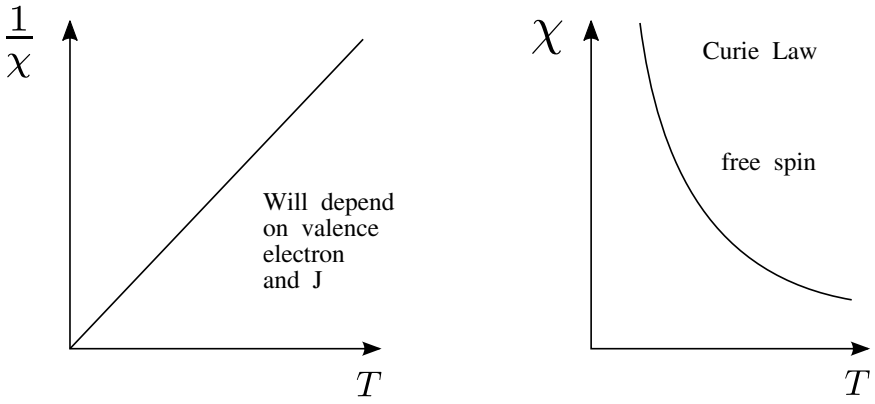
133.1.3 Spectroscopic Notation

Example Orbital Configurations



133.1.4 Paramagnetic Susceptibility

Curie Law for paramagnetic susceptibility: $\chi = C/T$



133.1.5 Calculation of J

- Rules for calculating J:

Value of J	Condition
$\lvert L - S \rvert$	When the shell is less than half full
$\lvert L + S \rvert$	When the shell is more than half full
S	When the shell is half full

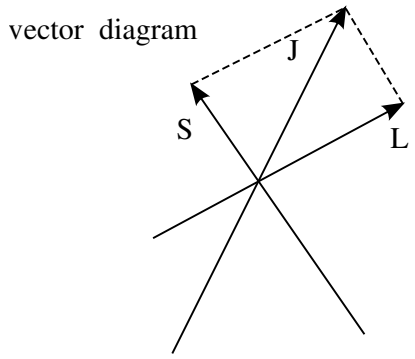
- Values for L and S:

$$L = \sum m_l \quad \text{and} \quad S = \sum m_s$$

- L is the orbital angular momentum

133.1.6 Spin Orbit Interactions

- For multi-electron atoms,
- Orbital angular momentum: $L = \sum_i L_i$
- Total spin angular momentum: $S = \sum_i S_i$
- Total angular momentum: $\hbar J$, where $J = L + S$
- Treat J, L, and S as vectors. L and S precess around J; J remains constant ("conserved")
- Note: for iron metal group, $J = S$; only the spin contributes. Orbital momentum is "quenched."



133.1.7 Effective Magnetic Number

- Effective magnetic number:

$$P = g\left[J(J+1)\right]^{1/2}$$

- Example: Ti^{3+}

$\boxed{J \sim S}$ For transition metals, "orbital angular momentum quenched"

$$3d^1 \quad \uparrow \quad \Rightarrow \quad S=1/2$$

$$P = 2\left[S(S+1)\right]^{1/2}$$

$$P = 2\left[\frac{1}{2}\cdot\left(\frac{3}{2}\right)\right]^{1/2} = 1.73$$

- Example: Fe^{2+}

$$3d^6 \quad \uparrow\uparrow\uparrow\uparrow\uparrow\downarrow \quad \Rightarrow \quad S=2$$

$$P = 2\left[2\,(3)\right]^{1/2} = 4.90$$

133.1.8 Paramagnetic Properties of Metals

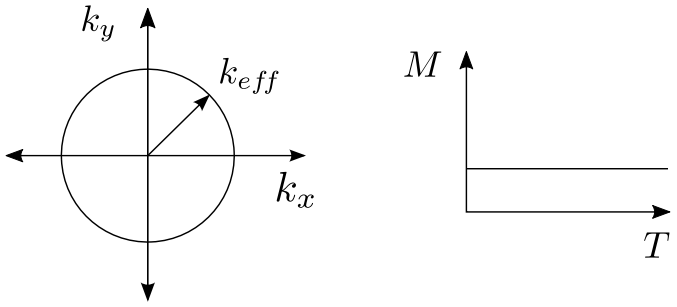
- For paramagnetic solids with N magnetic ions

$$M = \frac{N\mu^2 B}{kT_F}$$

- For metals, not all spins are accessible

$$M \cong \frac{N_{eff}\mu^2 B}{kT} \Rightarrow N_{eff} \cong \frac{NkT}{kT_F}$$

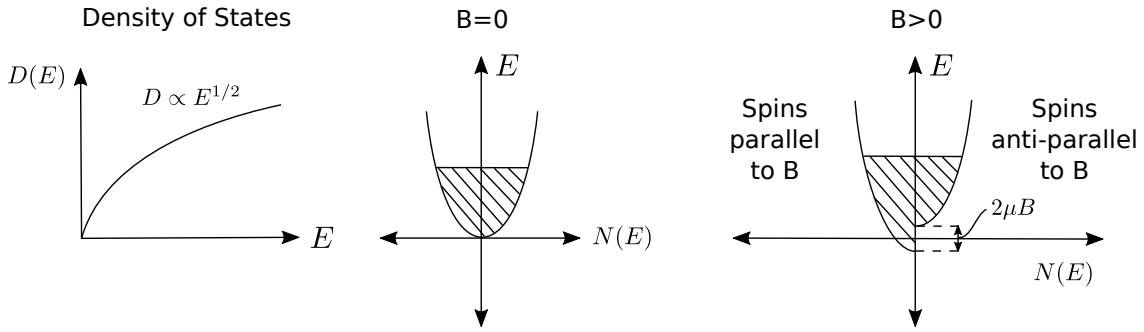
$$\boxed{M \cong \frac{N\mu^2 B}{kT_F}}$$ Temperature independent



133.1.9 Band Model

- Total magnetization: $\boxed{M = (N_+ - N_-)\mu}$
- Total number of electrons with "up" spins (N_+)

$$N_+ = \frac{1}{2}\int_{-\mu_B}^{E_F} dE \underbrace{f(E)}_{\text{Fermi function}} \underbrace{D(E+\mu B)}_{\text{Density of states}}$$



- Taking a Taylor expansion of the density of states for N_+,

$$\begin{aligned} N_+ &\approx \frac{1}{2}\int_0^{E_F} dE f(E)D(E) + \frac{1}{2}\int_0^{E_F} dE f(E)\frac{dD(E)}{dE}\mu B \\ &\approx \frac{1}{2}\int_0^{E_F} dE f(E)D(E) + \frac{1}{2}\int_0^{E_F} f(E)dD(E)\mu B \\ &\approx \boxed{\frac{1}{2}\int_0^{E_F} dE f(E)D(E) + \frac{1}{2}D(E_F)\mu B} \end{aligned}$$

- Repeat the procedure for N_-
- Taylor series expansion of the density of states for N_-

$$\begin{aligned} N_- &= \frac{1}{2}\int_{\mu_B}^{E_F} dE f(E) D(E-\mu_b) \\ &\approx \boxed{\frac{1}{2}\int_0^{E_F} d(E) f(E) D(E) - \frac{1}{2} D(E_F)\mu B} \end{aligned}$$

- Total magnetization:

$$M = (N_+ - N_-)\mu = D(E_F)\mu^2 B$$

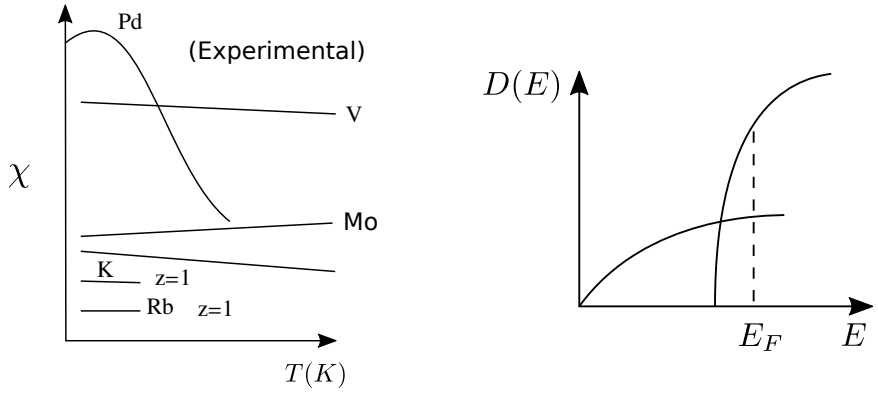
- For a metal,

$$D(E_F) = \frac{3N}{2E_F} = \frac{3N}{2kT_F}$$

$$\boxed{M = \frac{3N\mu^2 B}{2kT_F}} \quad \boxed{\chi = \frac{\partial M}{\partial B} = \frac{3N\mu^2}{2kT_F}} \quad \text{Independent of T!}$$

133.1.10 Multivalent Effects

- High density of states near Fermi level.
- d-band contribution to χ



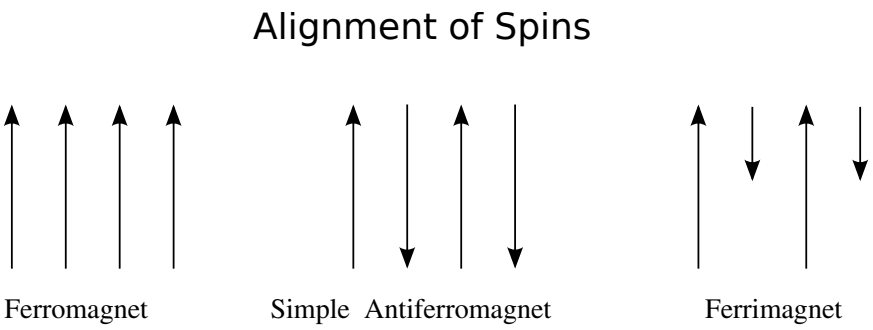
134 Ferromagnetism

- Magnetic ordering in solids is due to interaction of spins
- Below a transition temperature, an internal field tends to line up spins. This is called the "exchange field" B_E

$$H\Psi = E\Psi \quad \text{with} \quad \text{H=K+}V_{ex}$$

$V_{ex} \equiv$ Exchange Potential

- Mean field approximation: $\boxed{B_E = \lambda M}$ where $\lambda \equiv$ Weiss constant



134.1 Ferromagnetic Phase Transition

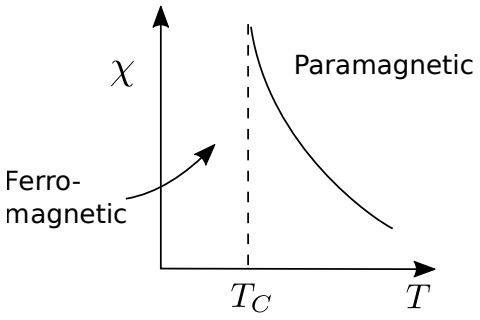
- Above the transition temperature, ferromagnetism gives way to paramagnetism

$$T > T_C \quad \text{Paramagnetism}$$

$$T < T_C \quad \text{Ferromagnetism}$$

- Magnetization: $M = \chi_p(B_A + B_E)$

$$\begin{aligned} \chi_P &\equiv \text{Paramagnetic susceptibility} \\ B_A &\equiv \text{Applied field} \end{aligned}$$



134.2 Molecular Field

- For paramagnetism, we had $\chi_p = C/T$
- "Molecular" or "exchange" field for ferromagnetism: $B_E = \lambda M$

$$M = \chi_P(B_A + \lambda M) \Rightarrow M = \frac{C}{T}(B_A + \lambda M)$$

- Solving for M,

$$M\left(1 - \frac{C\lambda}{T}\right) = \frac{CB_A}{T}$$

$$M(T - C\lambda) = CB_A$$

$$\chi = \frac{M}{B_A} = \frac{C}{(T - C\lambda)}$$

- Taking $C\lambda = T_C$,

$$\boxed{\chi = \frac{C}{T - T_C}} \quad \text{for } T > T_C$$

Weiss constant: $\boxed{\lambda = T_C / C}$

- Ferromagnetic behavior:

$$\boxed{\chi \propto \frac{1}{(T - T_C)^{1.33}}} \quad \text{Scaling theory}$$

- Calculation of field parameter:

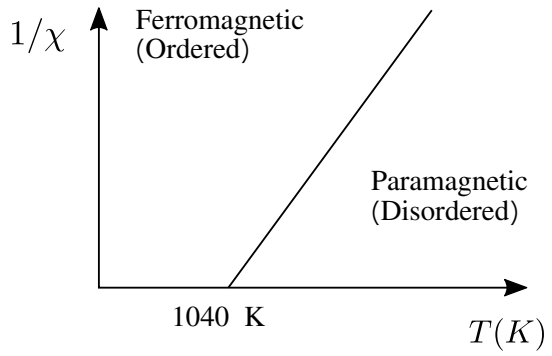
$$\lambda = \frac{T_C}{C} = T_C \underbrace{\frac{3k}{Ng^2 S(S+1)\mu_b^2}}_{\text{Curie Constant}}$$

- For iron, $T_C \cong 1000K$, $g = 2$, $S = 1$, $(3d^6 4s^2 \rightarrow 3d^8)$, $\lambda = 5000$. For a saturation magnetization of $M_S = 1700$ Gauss,

$$B_E = \lambda M_S$$

$$B_E = 5000 \times 1700 \approx 10^7 \text{ Gauss} = 10^3 \text{ Tesla}$$

- Highest superconducting magnet: 40 Tesla



Material	T_C (K)
Fe (bcc)	1043
Co	1388
Ni	627

Note: fcc iron is not magnetic

134.2.1 Prediction of T_C

- Heisenberg model. For atoms i and j with spins S_i and S_j, the ordering energy U is

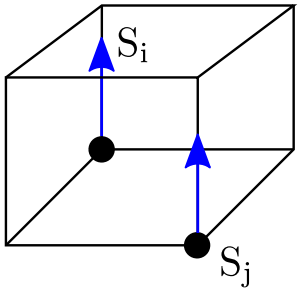
$$U = -2J_{ex}\vec{S}_i \cdot \vec{S}_j$$

$J_{ex} \equiv$ Exchange energy integral

- From mean field theory,

$$J_{ex} = \frac{3kT_C}{2ZS(S+1)} \quad \text{or} \quad T_C = \frac{2ZS(S+1)J_{ex}}{3k}$$

$Z \equiv$ Number of neighbors



134.2.2 Temperature Dependence of $M(T)$ for Ferromagnetism

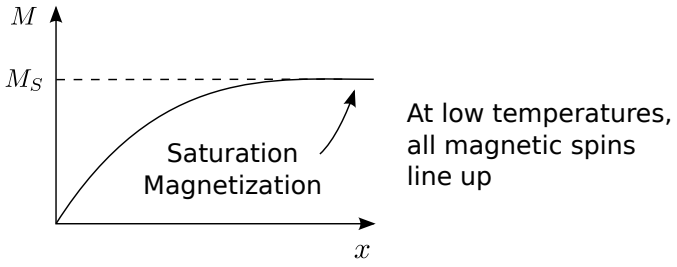
- Consider spin 1/2 system. Just as for paramagnetism, $M = N\langle\mu\rangle$

$$M = N\mu \tanh\left(\frac{\mu B}{kT}\right)$$

Where $B = B_A + B_E$

- For a ferromagnet, $B \approx B_E = \lambda M$, where $B_E \equiv$ Molecular field.

$$\therefore M = N\mu \tanh\left(\frac{\mu\lambda M}{kT}\right) = N\mu \tanh(x) \quad \text{with} \quad x \equiv \frac{\mu\lambda M}{kT}$$



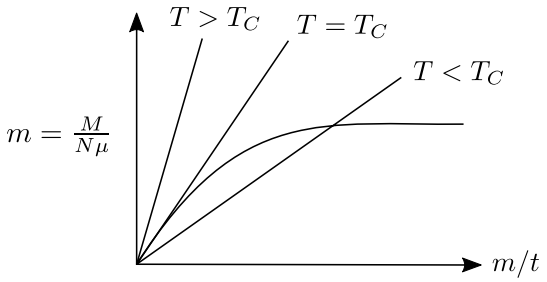
134.2.3 Ferromagnetic-Paramegnetic Transition

- Define reduced magnetization m and reduced temperature t

$$m \equiv \frac{M}{N\mu} \quad t \equiv \frac{kT}{N\mu^2\lambda}$$

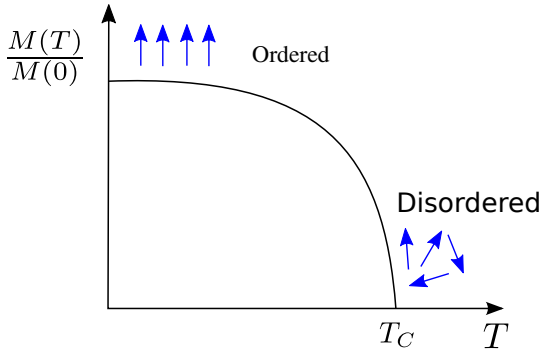
$$\therefore \underbrace{m}_{LHS} = \underbrace{\tanh(m/t)}_{RHS}$$

- Solve the transcendental equation graphically by plotting both the left hand side (LHS) and right hand side (RHS) versus m/t



134.2.4 Ferromagnetic-Paramagnetic Transition

- From the Landau model, as T increases, $M(T)$ decreases
- Second order phase transition: continuous change in $M(T)/M(0)$ versus T to T_C



134.2.5 Ferromagnetic-Paramagnetic Transition

- At low temperatures,

$$\tanh(m/t) = \tanh\xi \cong 1 - 2e^{-2\xi}$$

- Define change in magnetization with temperature, $\Delta M \equiv M(0) - M(T)$. For large ξ (low T),

$$\tanh\xi \cong 1 - 2e^{-2\xi} \approx 1$$

$$\therefore \Delta M = 2N\mu e^{-2\xi} = 2N\mu e^{-2\mu\lambda M/kT}$$

- Noting that $M = N\mu$,

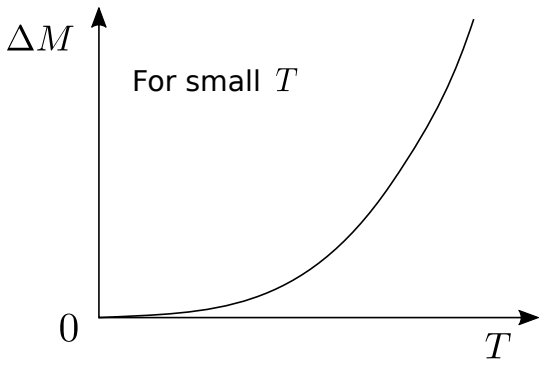
$$\Delta M = 2N\mu e^{-2\mu^2\lambda N/kT}$$

$$\boxed{\Delta M = 2N\mu\exp(-2\lambda N\mu^2/kT) = 2N\mu\exp(-2T_C/T)}$$

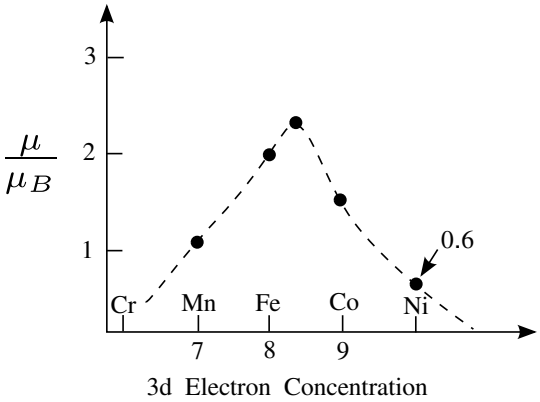
134.2.6 Ferromagnetic-Paramagnetic Transition

Change in magnetization with respect to $T = 0$ value versus temperature

$$\boxed{\Delta M = 2N\mu\exp(-2\lambda N\mu^2/kT) = 2N\mu\exp(-2T_C/T)}$$

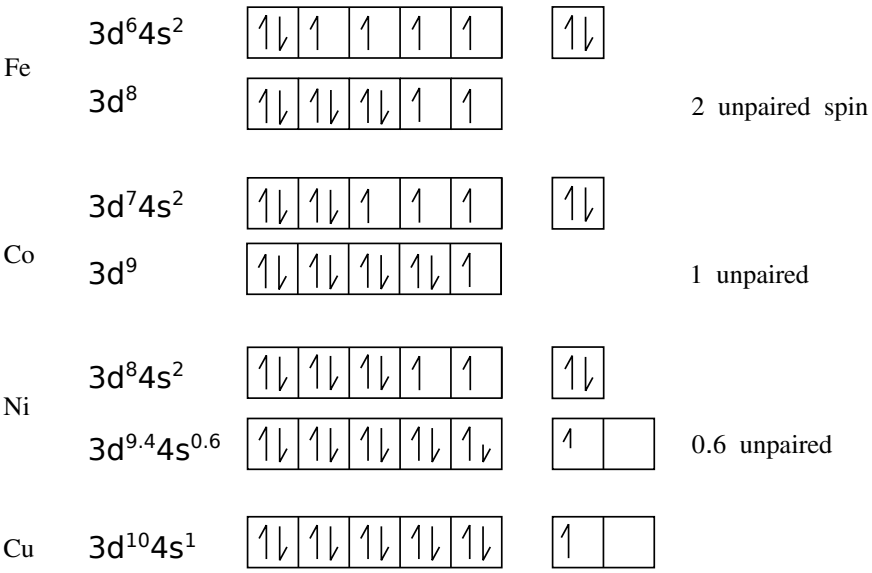


134.2.7 Ferromagnetism of Alloys



134.2.8 Transition Metals

d Orbital Configurations of Transition Metals

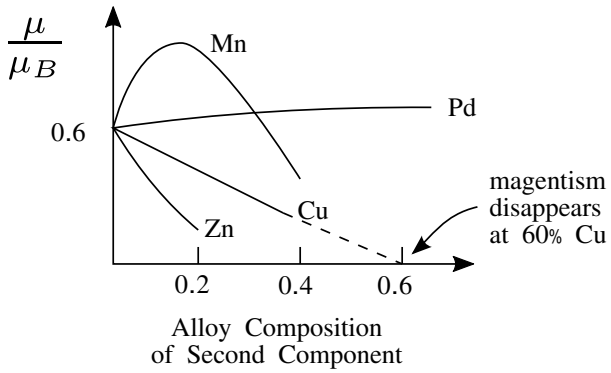


Others: Mn $3d^54s^2$

Pd $4d^{10}$

134.2.9 Ni Alloys

- Atomic moment of nickel transition metal alloys changes with composition
- For nickel alloyed with copper, magnetization disappears at 60% Cu



134.2.10 Band Model

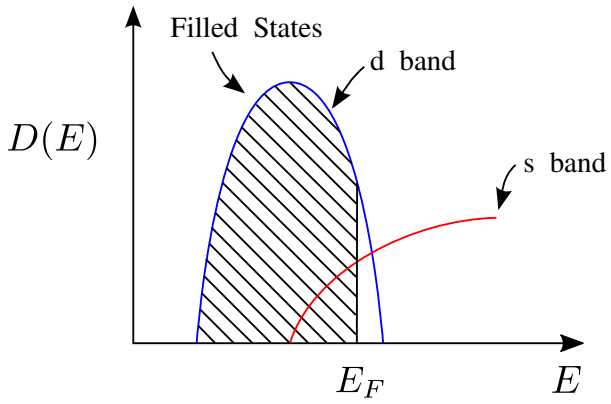
- For nickel, 3*d* bands and 4*s* bands are partially occupied

$$3d^{10} \longrightarrow 3d^{9.4}4s^{0.6}$$

5 up
4.4 down

Paired spins
(Pauli paramagnetism)

- Net spin for Ni is $0.6\mu_B$. Only the *d* bands contribute in this case



- Magneton numbers
- Fe: $2.2\mu_B$
- Co: $1.7\mu_B$
- For iron,

$$3d^8 \longrightarrow 3d^{7.8}4s^{0.2}$$

- Can copper be ferromagnetic?

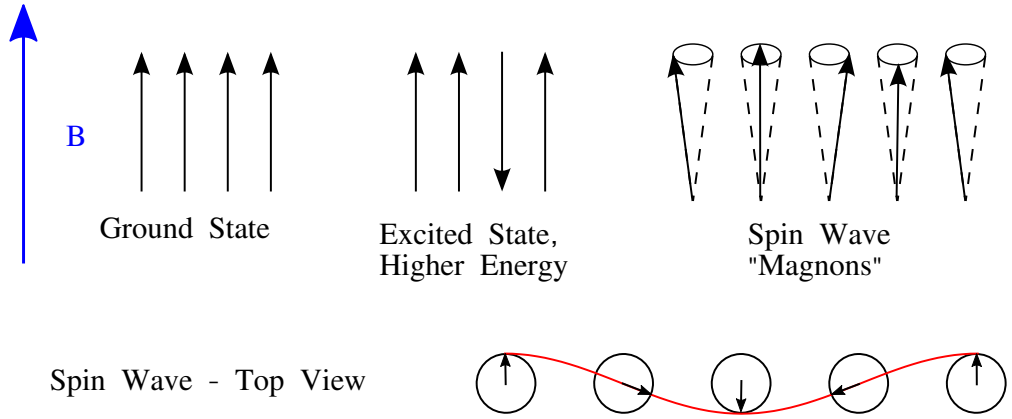
$$3d^{10}4s^1 \longrightarrow 3d^9 4s^2 \quad \text{(yes)}$$

superconducting properties?

YBa_2CuO

134.2.11 Spin Waves

- Spin waves: low energy excitations
- Lower energy excited state is obtained by having spins arranged at an angle θ with respect to neighboring spins ("spin waves" or "magnons")
- Precession angle depends on angle of neighbor

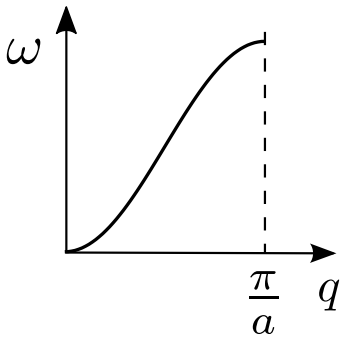


134.2.12 Magnon Dispersion

- Magnons are particles and have an ω versus q dispersion relation

$$\omega = \frac{4J_{ex}S^2}{\hbar}\sin^2\frac{aq}{2}$$

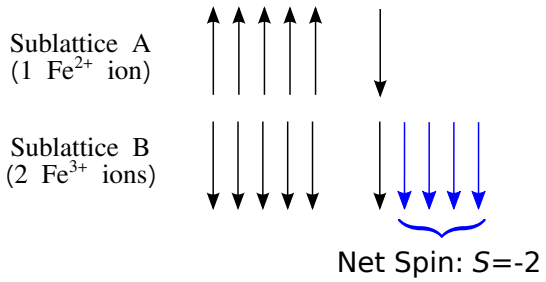
J_{ex} $\equiv$ Exchange energy term
S $\equiv$ Total spin angular momentum
a $\equiv$ Distance between atoms with magnetic moments
q $\equiv$ Wavevector



134.3 Ferrimagnetic Order in Magnetic Oxides

- Iron oxide (Fe_3O_4) consists of two sublattices:

$$Fe_3O_4 \rightarrow FeO \cdot Fe_2O_3$$



Sublattice	Fe Oxid. State	e^- Config.	S	$gJ\ (\mu_B)$
(A) FeO	2+	$3d^6$	2	4
(B)Fe_3O_4	3+	$3d^5$	$\frac{5}{2}$	5

- Total expected magnetic moment per formula unit: $2 \times 5 + 4 = 14\,\mu_B$
- Measured magnetic moment per formula unit: $4.1\,\mu_B$ (due to ferrimagnetic ordering)

134.3.1 Magnetic Oxides

- Calculation of saturation magnetization for $NiOFe_2O_3$ (nickel ferrite)
- Unit cell: inverse spinel with 8 Ni^{2+} ions (B sites) and 16 Fe^{3+} ions (A sites)

Ion	$g\sqrt{J(J+1)}(\mu_B)$
Ni^{2+}	2
Fe^{3+}	5

- The Fe^{3+} ions have anti-parallel magnetic moments

$$\therefore \text{ net } 2\mu_B \text{ per formula unit due to } Ni^{2+}$$

- For 8 Ni^{2+} ions per formula unit each with $2\mu_B$ magnetic moment,

$$M_S = \frac{8 \times 2\mu_B}{\text{cell volume}} = \frac{(16)(9.27 \times 10^{-24})}{(8.37 \times 10^{-10})^3} = 2.5 \times 10^5\ \text{A} \cdot \text{m}^{-1}$$

- Calculation of saturation magnetization for Fe_3O_4 (magnetite)
- Unit cell: inverse spinel with 8 Fe^{3+} ions on tetrahedral sites, 8 Fe^{3+} ions on octahedral sites, and 8 Fe^{2+} ions on octahedral sites.

Ion	Site	$g\sqrt{J(J+1)}(\mu_B)$
Fe^{3+}	Tetrahedral	5
Fe^{3+}	Octahedral	5
Fe^{2+}	Octahedral	4

- Fe^{3+} ions on tetrahedral sites have magnetic moments aligned anti-parallel to those of Fe^{3+}on the octahedral sites

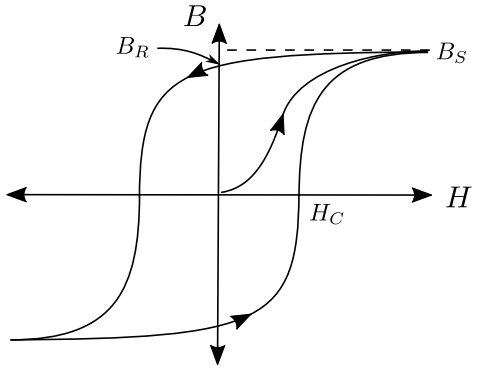
$$\therefore \text{ net } 2\mu_B \text{ per formula unit due to } Fe^{2+}$$

- For 8 Fe^{2+} ions per formula unit each with $4\mu_B$ magnetic moment,

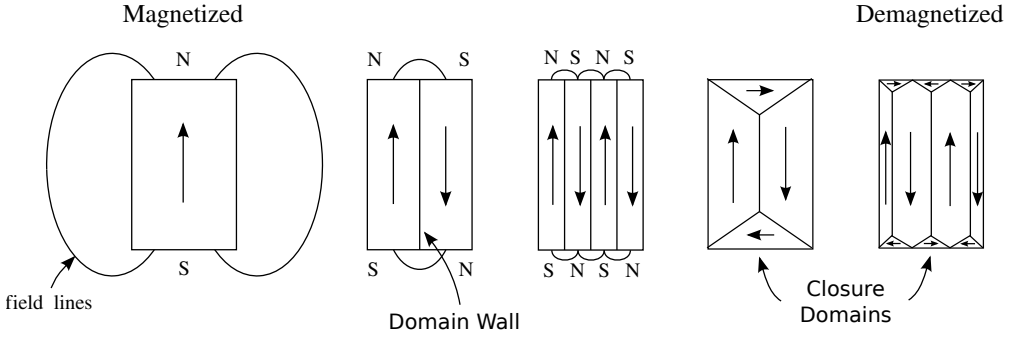
$$M_S = \frac{8 \times 4\mu_B}{\text{cell volume}} = \frac{(32)(9.27 \times 10^{-24})}{(8.40 \times 10^{-10})^3} = 5.0 \times 10^5\ \text{A} \cdot \text{m}^{-1}$$

134.3.2 Magnetization and Hysteresis

- In ferromagnetic materials, the magnetization is a nonlinear function of the applied magnetic field.
- Important terms describing the hysteretic behavior:
- H_C: coercive field. The applied magnetic field for which the flux density B disappears
- B_S: saturation flux density. The maximum flux density measured when all of the atomic moments are aligned with the applied magnetic field.
- B_R: remnant flux density. The magnetization remaining in the material after reaching the saturation flux density and removing the field.

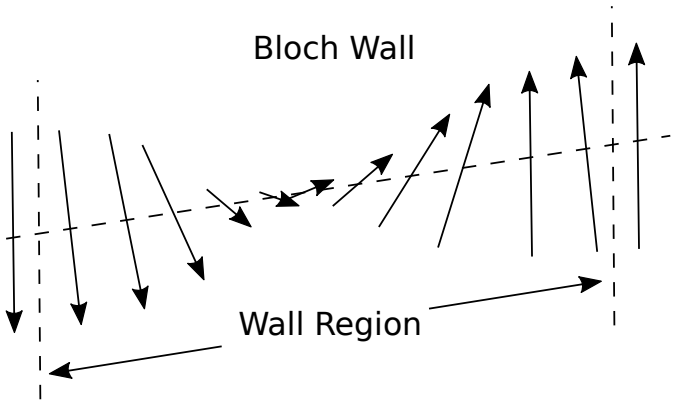


- Ferromagnetic domain: a region in a ferromagnetic material over which all magnetic moments are aligned
- Domains exist in in the demagnetized state in order to minimize the large magnetostatic energy associated with single domains



134.4 Domains and Walls

- Domain wall: region between adjacent domains over which the direction of the local magnetic moment changes
- It takes energy to form and to move domain walls



134.4.1 Energy of Bloch Domain Walls

- Energy U (Heisenberg Model):

$$U = -2J_{ex}\vec{S}_i \cdot \vec{S}_j$$

- Consider the spin vectors by a classical model, i.e. $\vec{S}_i \cdot \vec{S}_j = |S|^2 \cos\phi$. Take a Taylor series expansion of $\cos\phi$ about $\phi = 0$

$$U = -2J_{ex}|S|^2 \cos\phi \approx -2J_{ex}|S|^2 \left(1 - \frac{1}{2}\phi^2\right)$$

- The angle-dependent exchange energy between two adjacent spins at an angle ϕ is then

$$w_{ex} = J_{ex}|S|^2\phi^2$$

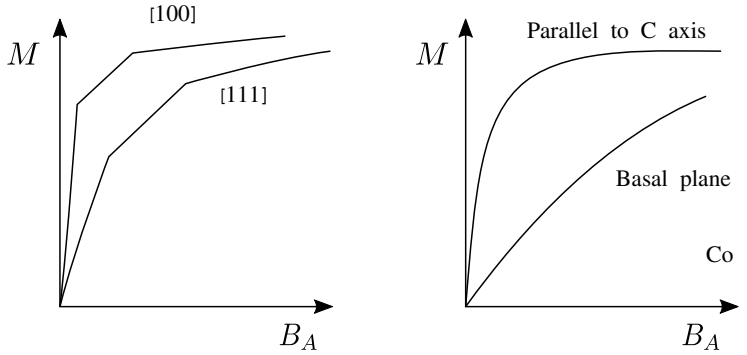
- For a rotation of the magnetic moment by π radians in N steps, $w_{ex} = J_{ex}|S|^2(\pi/N)^2$. For a Bloch wall of N spins, the total energy is then

$$Nw_{ex} = J_{ex}|S|^2\frac{\pi^2}{N}$$

- Anisotropy energy limits the Bloch wall size since the spins are arranged away from the direction of easy magnetization. (The more that the spins point the "wrong way," the higher the energy).

134.5 Anisotropy of Magnetization

- Alignment of magnetic moments is more favorable for certain crystalline directions, resulting in anisotropic energy term
- Under an applied field, domains rotate to align with the field
- The hysteretic response is dependent on the direction of the applied field with respect to the crystalline direction
- "Soft" direction: crystalline direction with low coercivity (H_C)
- "Hard" direction: crystalline direction with high coercivity (H_C)

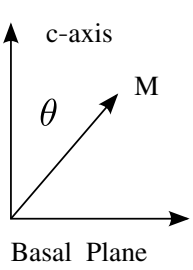


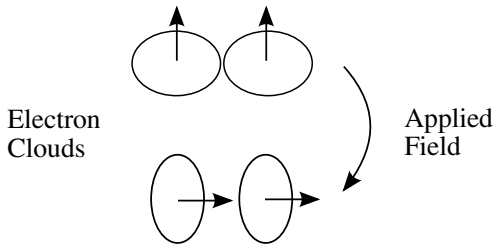
134.5.1 Anisotropy of Magnetization

- Rotation of domains changes the overlap of the electron clouds of neighboring atoms, resulting in different exchange energy J_{ex}
- Anisotropy energy U_k

$$U_k = k_1' \sin^2\theta + k_2' \sin^4\theta$$

- Easy axis depends on values of k_1' and k_2'

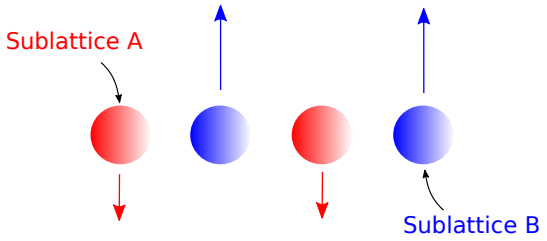




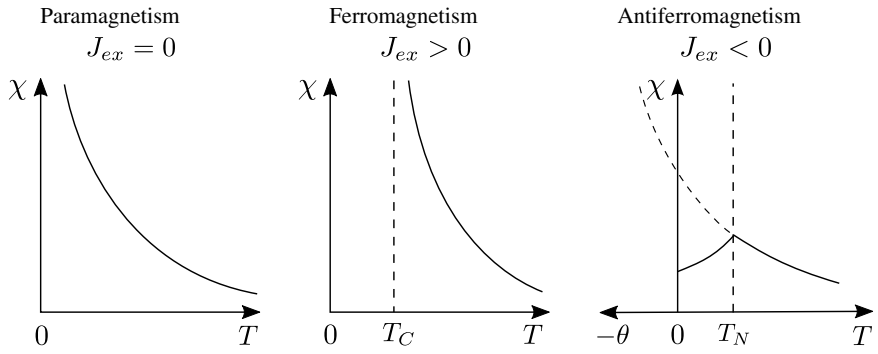
134.6 Ferrimagnetic Ordering - Exchange Terms

- Exchange term for neighboring Fe atoms is negative for atoms on different sublattices and positive for atoms on the same sublattice.
- Different sign of exchange term *and* different number of Fe atoms per unit volume on each sublattice account for ferrimagnetic ordering in Fe_3O_4
- Exchange terms for A and B sublattices:

$$J_{AA} > 0 \quad J_{AB} < 0 \quad J_{BB} > 0$$



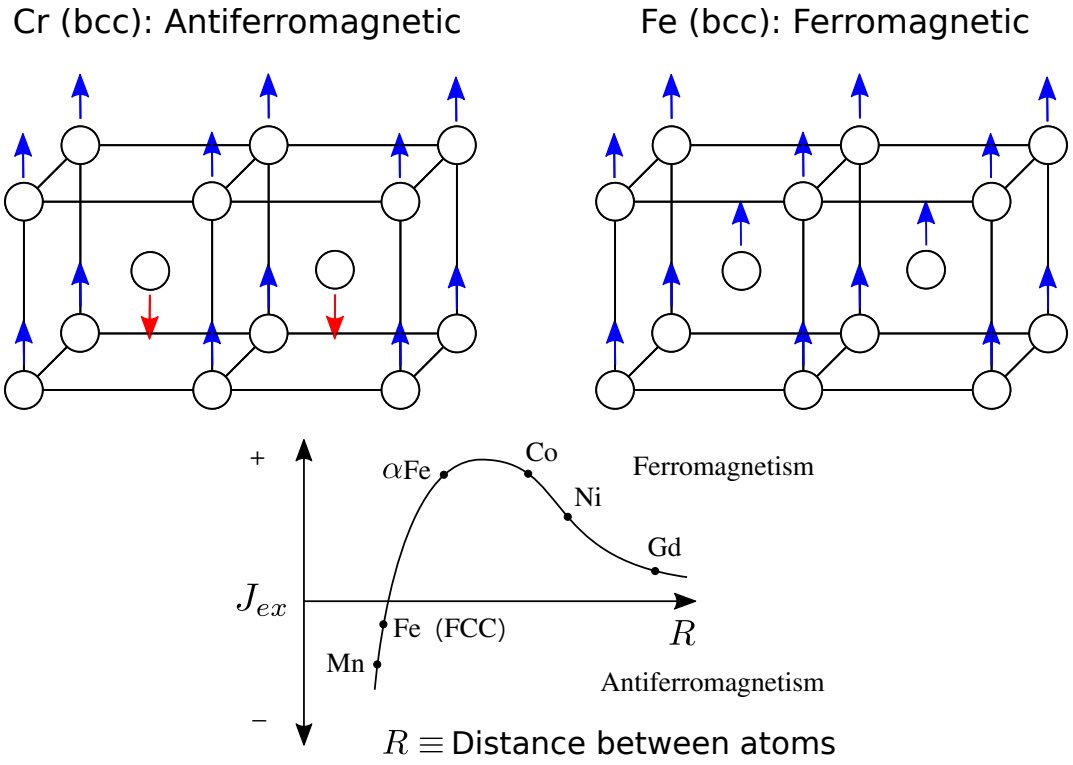
134.6.1 Exchange Terms and Susceptibility



$$\chi = \frac{C}{T} \qquad \chi = \frac{C}{T - T_C} \qquad \chi = \frac{C}{T + T_N}$$

$T_C \equiv$ Curie temperature
$T_N \equiv$ Neel temperature

134.6.2 Structure Dependence of J_{ex}



135 Electro-optic and Nonlinear Optical Materials

- Nonlinear optical materials have a nonlinear relation between the polarization density (P) and electric field ($\mathcal{E}$)
- Expand electric susceptibility in higher order terms

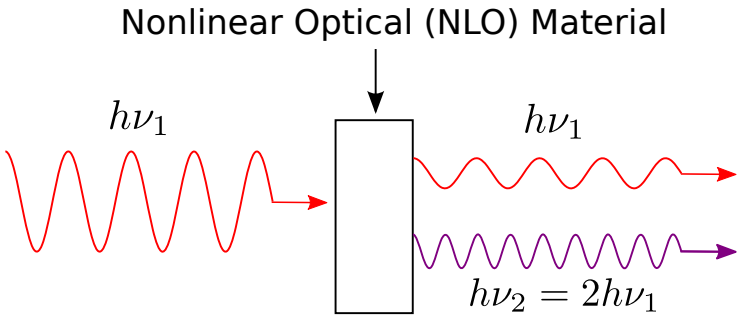
$$P = \epsilon_0 \left[\chi^{(1)}\mathcal{E} + \chi^{(2)}\mathcal{E}^2 + \chi^{(3)}\mathcal{E}^3\right]$$

$\chi^{(1)} \equiv$ Linear susceptibility
$\chi^{(2)} \equiv$ Quadratic susceptibility
$\chi^{(3)} \equiv$ Cubic susceptibility

- $\chi^{(2)}$ is nonzero in non-centrosymmetric materials. Examples of $\chi^{(2)}$ effects: linear (Pockels) electro-optic effect, second harmonic generation (frequency doubling)
- The lowest order nonlinearity for cubic materials is $\chi^{(3)}$. Examples of $\chi^{(3)}$ effects: quadratic (Kerr) electro-optic effect, third harmonic generation (frequency tripling)

135.1 Frequency Doubling

- In frequency doubling, two photons with energy $h\nu_1$ combine to form one photon with energy $h\nu_2$. The total energy is conserved.
- Sending a beam with photon energy $h\nu_1$ into a nonlinear optical (NLO) material results in two output beams with energies $h\nu_1$ and $h\nu_2$.
- Examples:
- Nd:YAG laser with 1.10μm (infrared) radiation $\xrightarrow{NLO}$ 0.55μm (green)
- Blue laser by frequency doubling GaAs lasers



135.2 Nonlinearity in Refractive Index

- Velocity of light in matter: $v = c/n$

$$\begin{aligned} c &\equiv \text{Speed of light in free space } (3\times10^8 \text{ m/s}) \\ n &\equiv \text{Refractive index} \end{aligned}$$

- Refractive index (n) in nonlinear material:

$$n^2 = 1 + \chi^{(1)} + \chi^{(2)} + \chi^{(3)} + \cdots$$

- Dependence of refractive index on light intensity:

$$n = n_0 + n_2 I$$

$$\begin{aligned} n_0 &\equiv \text{Linear and } \chi = \epsilon_r - 1 \\ n_2 &\equiv \text{Nonlinear index} \\ I &\equiv \text{Intensity} \end{aligned}$$

- Dielectric constant: $\epsilon_r \cong n^2$

135.3 Electro-Optic Modulators

- The optical properties of non-centrosymmetric crystals are modified by applying an external electric field.

$$\Delta\left(\frac{1}{\epsilon_r}\right) = \Delta\left(\frac{1}{n^2}\right) = r\mathcal{E}$$

- Expanding $n = n(\mathcal{E})$ in a Taylor series expansion about $\mathcal{E} = 0$, the following expression is obtained

$$\Delta n = -\frac{1}{2} n^3 r \mathcal{E}$$

$r \equiv$ Electro-optic coefficient [pm/V]

- Example material: $LiNbO_3$, $r = 31$ pm/V, $n = 2.29$. Fields of $10^6 - 10^7$ V/m are easily obtained in micro-photonic devices (1-10 V across $\sim 1\mu$m)
- Other materials: electro-optic polymers ($r = 20 - 50$ pm/V), $BaTiO_3$ ($r = 300$ pm/V)

135.4 Optical Memory Devices

- Photorefractivity: change in refractive index by light. Light empties trap states, creating a space charge region that locally changes the refractive index.
- Volume holography:
- Two plane waves (object and reference beams) incident upon a photosensitive material. A standing wave is formed which is preserved in the material as an interference pattern
- Material contains intensity *and* phase information and the object can be reconstructed
- The difference between hologram and a photograph is that phase information is preserved in a hologram

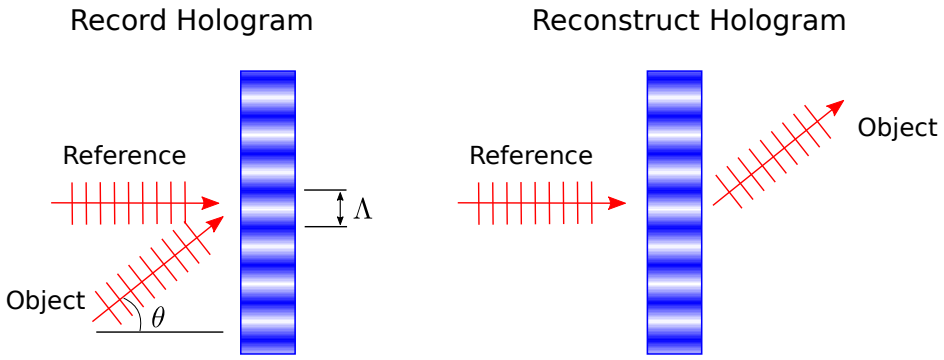
135.5 Volume Holography

- Pattern in material creates a dielectric grating or spatial modulation in the refractive index and dielectric constant

$$\begin{aligned} \epsilon_r &= \epsilon_{r0} + \epsilon_{r1} \cos(2ky \sin\theta) \\ k &\equiv \text{Magnitude of light wavevector} \end{aligned}$$

- Period of dielectric grating:

$$\Lambda = \frac{2\pi}{2k \sin\theta} = \frac{\lambda}{2n \sin\theta} \quad \text{with} \quad k = 2\pi n/\lambda$$



135.6 Photorefractive Crystals

- Electro-optic and photoconductive - organic and inorganic materials
- Mechanism for photorefraction:

(1) Generation of electrons and holes

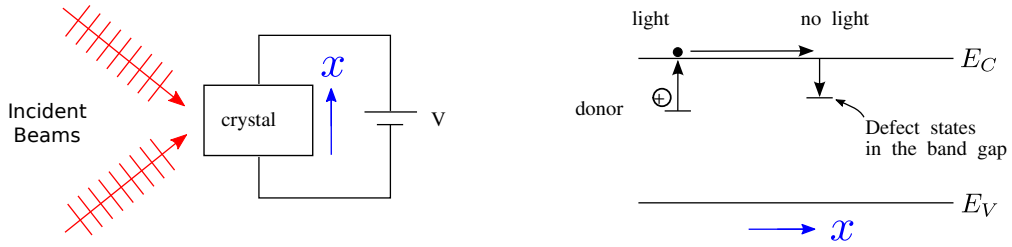
(2) Transport charges through the material

(3) Trap the charges

(4) Local field of trapped charges changes the refractive index n

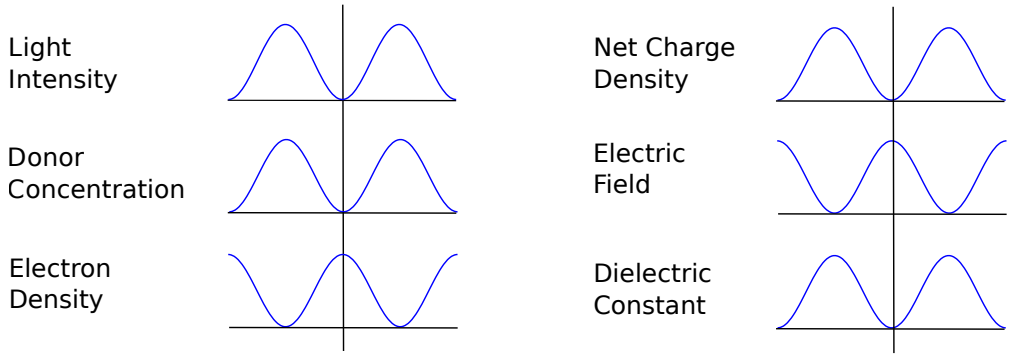
- Spatially varying charge distribution results in spatially varying field. Field changes n through electro-optic coefficient r; larger r, larger photorefractive effect.

$$\Delta n = -\frac{1}{2} n^3 r \mathcal{E}$$

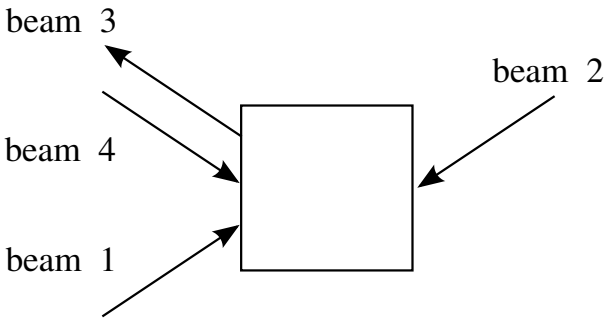


135.7 Photorefractive Crystals

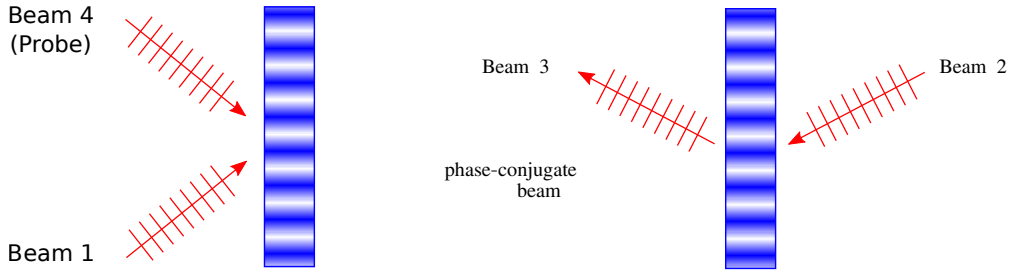
Steps in creation of periodic dielectric constant



135.8 Phase Conjugation - All Optical Switching



- Beams 1 and 2 are pump beams
- Beam 4 is a probe beam
- Beam 3 is the phase conjugate beam
- Diffraction grating is created by beams 1 and 4 through interference and the photorefractive effect
- Beam 2 is diffracted by the grating, forming beam 3



135.9 Acousto-Optic Modulators

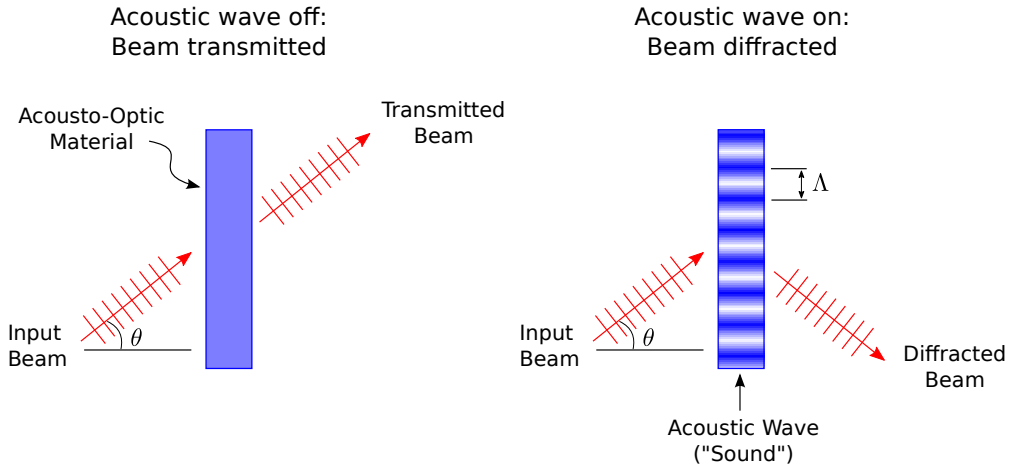
- Photoelastic effects: strain causes a change in the refractive index

$$\Delta\left(\frac{1}{\epsilon_r}\right) = pS \quad \text{and} \quad \Delta\frac{1}{n^2} = pS$$

$S \equiv$ Strain
$p \equiv$ Photoelastic constant
$n \equiv$ Refractive index

- Acousto-optic effect is used in Bragg cells to create a grating with period Λ by launching an acoustic wave. Light is diffracted from the grating.

$$\Lambda = \frac{\lambda}{2n\sin\theta}$$

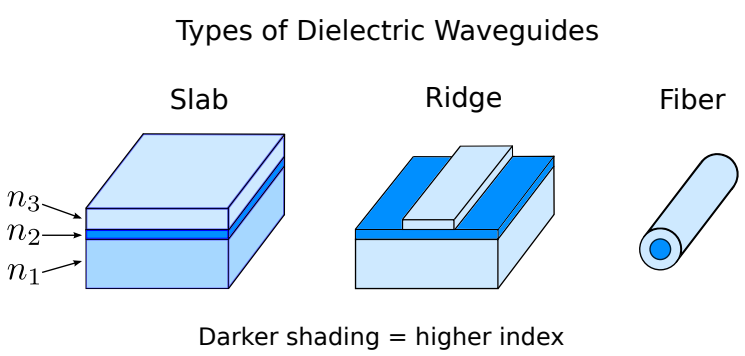


135.10 Integrated Optics

- Integrated optics: passive and active optical elements incorporated onto a wafer of material for local manipulation of light.
- Passive elements: elements that do not require power, such as waveguides, couplers, and filters
- Active elements: elements that require power, such as switches, modulators, directional couplers, amplifiers, lasers
- Photonics: analogue of electronics. Photonics encompasses the control of photons.

135.11 Dielectric Waveguides

- In integrated optics, dielectrics are used to guide light in waveguides. Waveguides act as "optical wires"
- Light is confined to the region of highest refractive index
- Why dielectric materials instead of metals? Ease of fabrication and lower losses



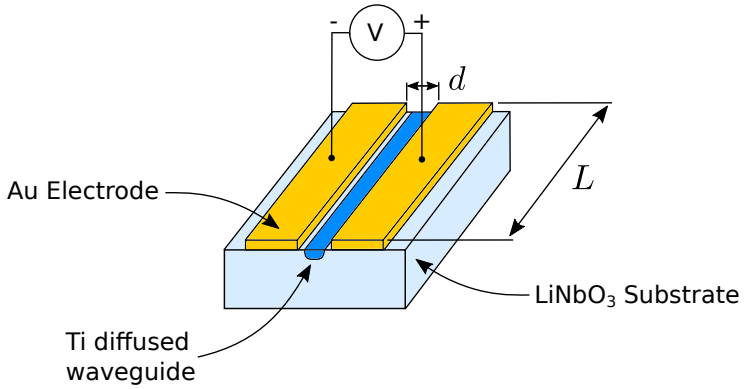
135.12 Example: Phase Shifter

- An electro-optic dielectric waveguide confines and guides light
- An electric field is applied to electrodes surrounding the waveguide. The refractive index of the dielectric waveguide is then modified by the electro-optic effect

$$\Delta n = -\frac{1}{2} n^3 r \mathcal{E} = -\frac{1}{2} n^3 r \frac{V}{d}$$

- Phase delay of optical wave due to the electro-optic effect:

$$\Delta\phi = \frac{2\pi}{\lambda} \Delta n L$$



135.13 Example: $LiNbO_3$ Phase Shifter

- Electric field of 10^6 V/m (10 V across 10 μm) yields $\Delta n = 1.86 \times 10^{-4}$ for $LiNbO_3$
- For a phase shift $\Delta\phi = \pi$ with operation at $\lambda = 1.5\,\mu$m,

$$\Delta\phi = \pi = \frac{2\pi}{\lambda} L \Delta n$$

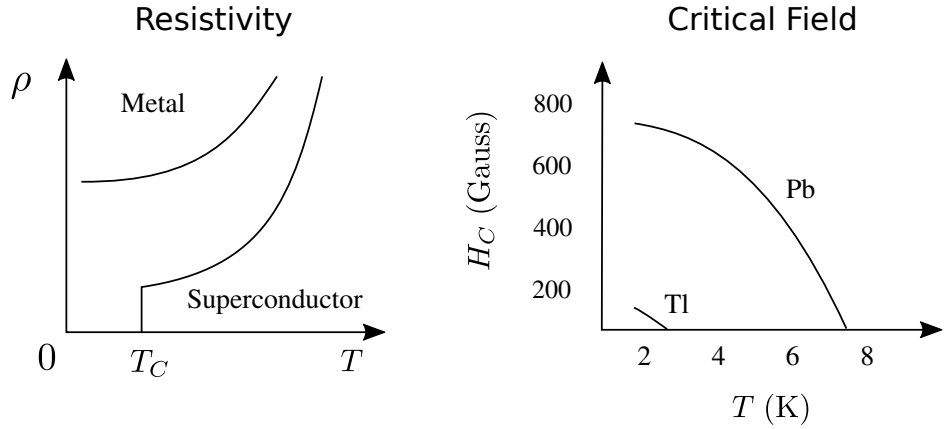
$$L = 0.4\,\text{cm}$$

- Long interaction lengths are needed; materials with higher electro-optic coefficients are needed for more compact devices
- The voltage required for a π phase shift is the half-wave voltage V_π

$$V_\pi = \frac{\lambda d}{n^3 r L}$$

136 Superconductivity

- Superconductivity: a phenomena in which a material undergoes a transition to zero resistivity with zero magnetic induction (Meissner effect) below a critical temperature T_C.
- At temperatures below T_C application of a critical magnetic field H_C will cause a transition to the normal state

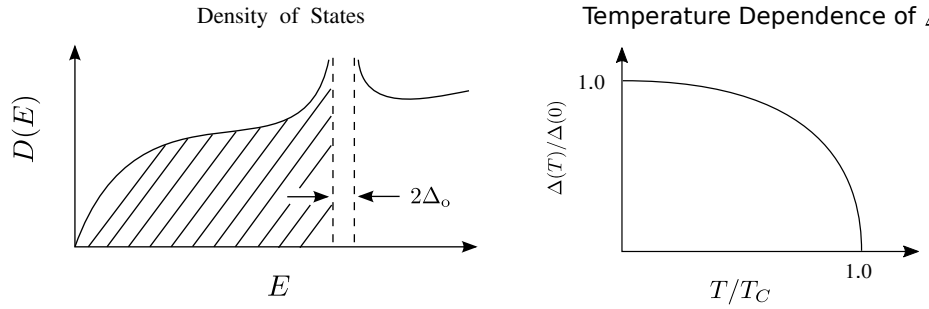


136.1 BCS Theory of Superconductivity

- BCS Theory: quantum mechanical theory of superconductivity formulated by Bardeen, Cooper, and Schreiffer in 1957 (Nobel Prize in 1972).
- Electron-phonon coupling leads to the formation of electron pairs (Cooper pairs)
- BCS Theory accounts for observed properties of superconductivity, including the energy gap, Meissner effect, critical temperature, and quantization of magnetic flux through a superconducting ring

136.2 Density of States and Energy Gap

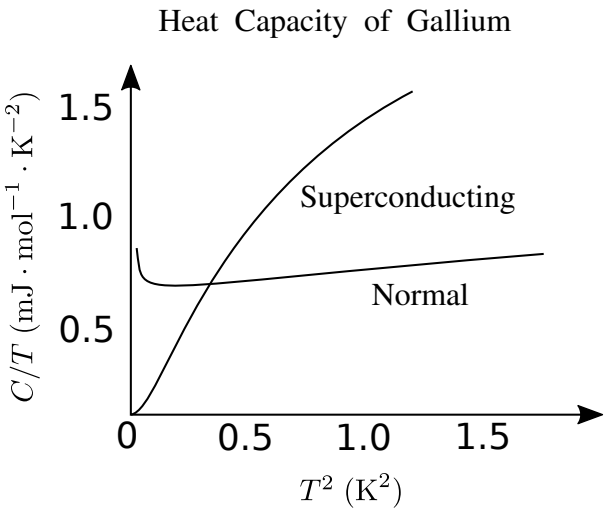
- The energy gap in superconductors is caused by electron-electron interactions rather than electron-lattice interactions for dielectrics
- The energy gap is maximum at 0 K and decreases to zero at the transition temperature.



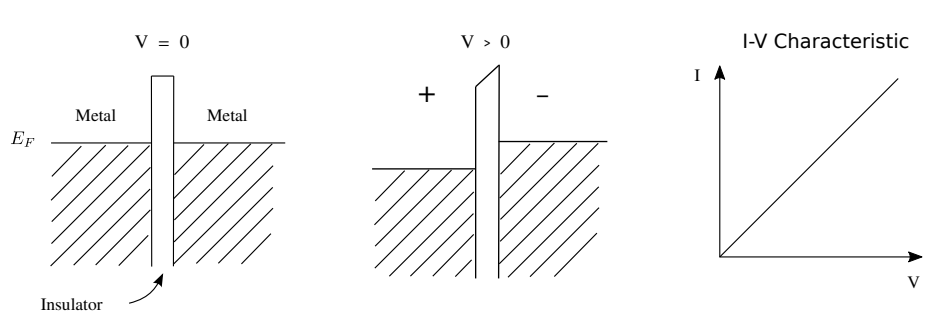
136.3 Heat Capacity of Superconductors

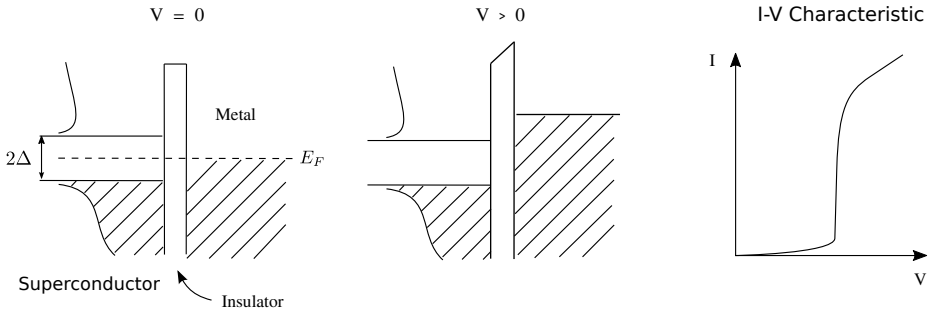
- The entropy of superconductors decreases upon cooling below T_C, resulting in electronic ordering.
- In the superconducting state, for $T \ll T_C$,

$$H_C = H_0 \left[1 - 1.07(T/T_C)^2\right]$$



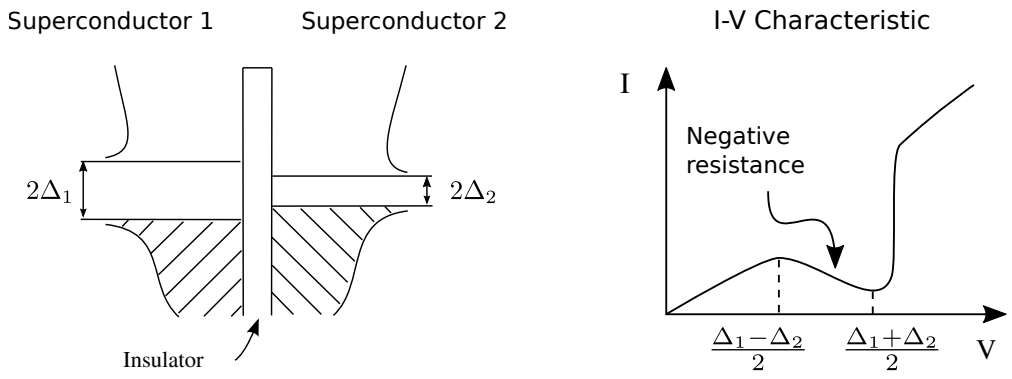
136.4 Tunneling in Superconductor Junctions





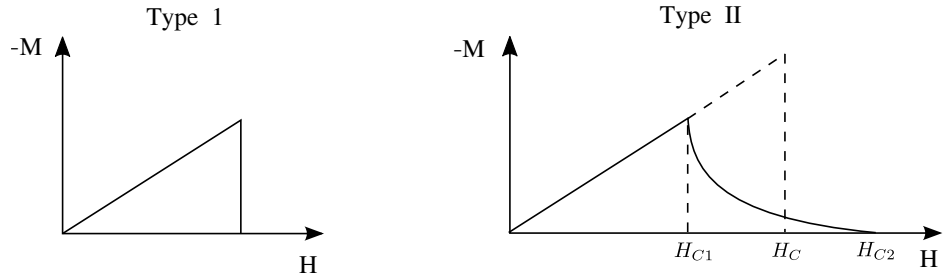
136.5 Semiconductor-Insulator-Semiconductor Junction

- Tunneling across a thin oxide junction between two superconductors with energy gaps $2\Delta_1$ and $2\Delta_2$
- Nobel Prize in physics, Ivan Giaever, 1973



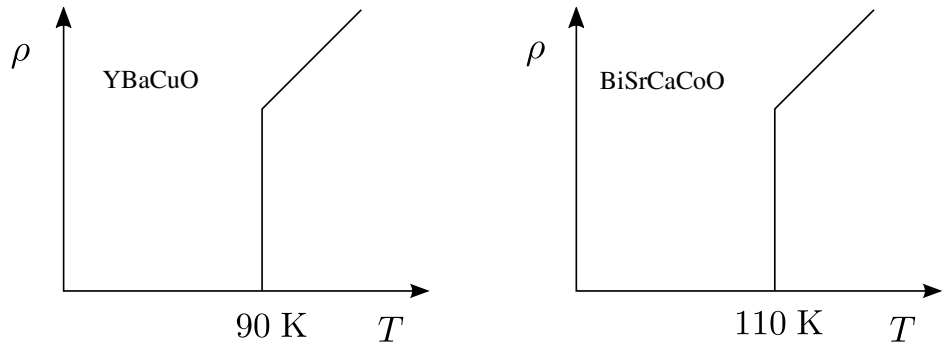
136.6 Type I and Type II Superconductors

- Type I: superconductor is diamagnetic in an applied magnetic field up to the critical field H_c
- Type II: superconductor exhibits two magnetic behaviors. Up to a lower critical field H_{C1} the superconductor is diamagnetic, and between H_{C1} and an upper critical field H_{C2}, the flux density is not equal to zero ("incomplete" Meissner effect).
- Type II superconductors tend to be alloys or transition metals with high room temperature resistivity
- Type II superconductors are used in high field magnets, magnetic resonance imaging (MRI) applications



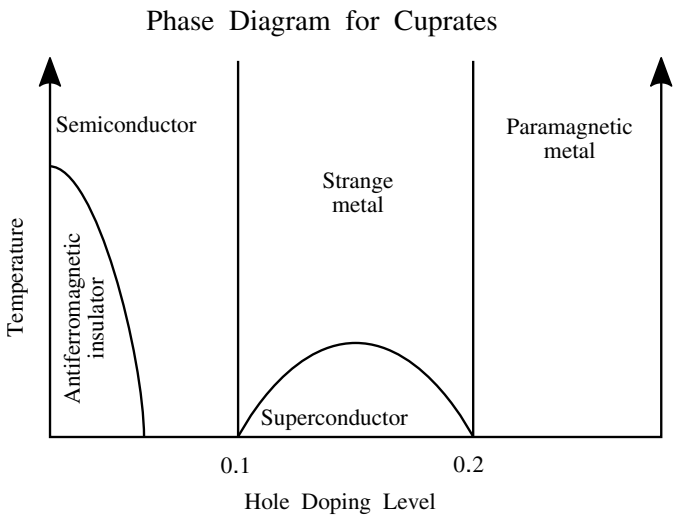
136.7 High T_C Superconductors

- Highest values for T_C have been reported in cuprates
- 2 layer cuprate compounds support 2D conduction



136.8 Cuprate Superconductors

- Theory: pairing is involved. Nature of pairing? Spin Waves?
- $2\Delta(0) \approx 6kT_C$
- Future applications: current carrying wires, Maglev trains



137 351-2 Problems

(1) An abrupt Si p-n junction has$N_a = 10^{18}cm^{-3}$ on one side and$N_d = 10^{15}cm^{-3}$on the other.

a Calculate the Fermi level position at 300K on both sides.

b Draw an equilibrium band diagram for the junction.

c Determine the contact potentialΦ_o for this junction.

(2) A silicon$p^+ - n$ junction$10^{-2}cm^2$ in area had$N_d = 10^{15}cm^{-3}$ doping on the n-side. Calculate the junction capacitance with a reverse bias of 10V.

(3) For metallic aluminum, calculate:

a The valence electron density.

b The radius of the Fermi spherek_F.

c Fermi energy in eV.

(4) From the Schrodinger equation for a quantum well, show that the wave vector is equal to$n\pi/L$ where L is the well width.

(5) Calculate the energy of light emitted from a 10 nm wide AlGaAs/GaAs quantum well structure that is photoexcited with 2.5 eV laser light.

(6) What is the luminescent energy for a CdSe quantum dot with a 2 nm radius.

(7) For a MOSFET device briefly describe how the three types of device work: a) enhancement mode b) depletion mode c) inversion mode.

(8) Calculate the capacitance of an MOS capacitor with a 10 nm thickHfO_2 dielectric oxide. What is the ratio of capacitances forC_{HfO_2}/C_{SiO_2}. The relatie dielectric constant for$HfO_2$is 25.

(9) Problem 9.9 in Solymar and Walsh

(10) Problem 9.14 in Solymar and Walsh

(11) Problem 9.16 in Solymar and Walsh

(12) Problem 12.10 in Solymar and Walsh

(13) Consider a quantum cascade laser (QCL) made from GaAs and GaAlAs. What well thickness is needed for laser emission at 3 microns?

(14) Derive the expression for the average value of the dipole moment. Show that it is given by:
$< \mu >= \mu[\coth a - \frac{1}{a}]$

(15) The saturation polarizationP_s of$PbTiO_3$, a ferroelectric, is0.8 coulombs/m^2. The lattice constant is 4.1$\dot{A}$. Calculate the dipole moment of unit cell.

(16) Calculate the polarization P of one liter of argon gas at 273 K and 1 atm. The diameter of an argon atom is 0.3 nm.

(17) Consider the frequency dependence of the atomic polarizability. The polarizability and its frequency dependence can be modeled as a damped harmonic oscillator. Derive the expression forα in this case.
The expression is given by:
$m\frac{dx}{dt^2} + b\frac{dx}{dt} + \omega_0^2 x = -e\varepsilon_{loc}\sin\omega t$
Plotα vs.ω for this case.

(18) Problem 4.6 in Solymar and Walsh.

(19) Problem 4.7 in Solymar and Walsh.

(20) Problem 4.8 in Solymar and Walsh.

(21) Problem 4.9 in Solymar and Walsh.

(22) Calculate the magnetic susceptibility of metallic copper. How does it compare to the measured value of -1.0?

(23) Calculate the effective magneton number p forMn^{2+}, Co^{2+}. Show work.

(24) Consider Mn doped GaP. There are$10^{20}Mn^{2+}$ ions.

(25) What is the electron configurationMn^{2+} in spectroscopic notation.

a Calculate its magnetic moment at saturation in Bohr magnetons.

b Calculate its magnetic susceptibility.

(26) For metallic Co, which has a Curie temperature of 1388 K, calculate the Weiss constantλ. Calculate the exchange constant in meV.

Part X
361: Crystallography and Diffraction

138 Symmetry of Crystals

A crystal can be defined as a solid which is composed of an atomic array that is periodic in three dimensions. Crystal systems may exhibit various kinds of symmetry, or types of repetitiveness. The specific type of symmetry is determined by a number of operations, performed on the crystal, which bring it into self coincidence.

138.1 Types of Symmetry

Symmetry elements can be of the following types:

- Rotational axis
- Mirror plane
- Inversion center
- Translation center

Combinations of these symmetry elements are also possible. The plane object pictured in Figure 138.1 demonstrates rotational symmetry, as a 120° turn about its center brings it into self coincidence. In addition, the object has three mirror planes, and the object's reflection across these planes also brings it into self coincidence.

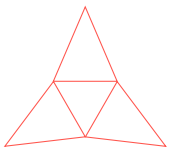


Figure 138.1: Rotational symmetry

An example of inversion center symmetry is pictured in Figure 138.2. When a line is drawn through the center of this plane object, corresponding points of the object are equidistant from the center line. It will come into coincidence with itself if all points on the object are inverted through the inversion center which is also on this center line. This object also exhibits rotational symmetry, as a 180turn about its center brings it into self coincidence.

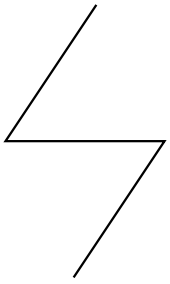


Figure 138.2: Inversion center and rotational symmetry

The object in Figure 138.3 demonstrates translation center symmetry or a repetitiveness in space. This figure may be interpreted as a flat representation of a diatomic molecule, which is being translated in equal steps both to the right and to the left.

Figure 138.3: Translation center symmetry

138.1.1 Proper Rotation Axes

Proper rotation axes may occur in both two dimensions and three dimensions. Figure 138.4 shows the types of proper rotation axes as well as the symbols which will be used to represent each type. Referring back to Figure 138.1, the object pictured has 3-fold rotational symmetry about an axis directed out of the page, as a 360/3 or 120rotation allows for self coincidence. In other words, an object has n-fold symmetry about an axis if a $360/n$ rotation about that axis brings it into self coincidence.

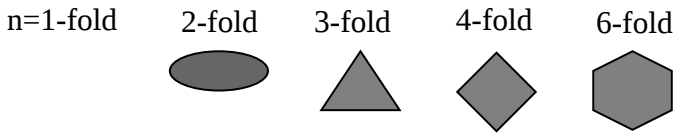


Figure 138.4: Types of rotational Axes

All objects exhibit 1-fold symmetry, or are brought into self coincidence by a full 360 turn. This case is trivial but important nonetheless. Furthermore, other types of rotational axes (i.e. 5, 7 or greater) are not permitted due to translational symmetry restrictions. For instance, you cannot fill space by translating a pattern with an object that has a 5-fold rotation axis. (See Figure 138.5).

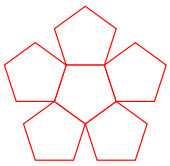


Figure 138.5: 5-fold axis

138.1.2 Improper Rotation Axes

An improper rotation axis repeats a left-handed object from a right-handed object and vice-versa. This movement of the object combines rotation and inversion in a single hybrid operation (rotoinversion), and is only possible in three dimensions. Just as there is a set number of proper rotation axes, there is a set number of improper rotation axes as shown in Figure 138.6.

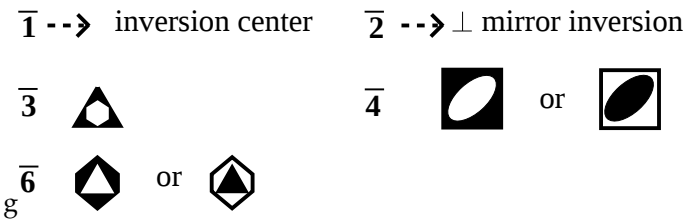


Figure 138.6: Rotoinversion or Improper Axes

Figure 138.7, from Vainhstein's *Modern Crystallography* Vol I., depicts a perfectly asymmetrical three dimensional object undergoing all of the symmetry operations that have been mentioned thus far. Notice that the $\bar{1}$ is exactly

like inversion center symmetry, and that the $\bar{2}$ operation is exactly like mirror plane symmetry with the dashed circle as the mirror plane.

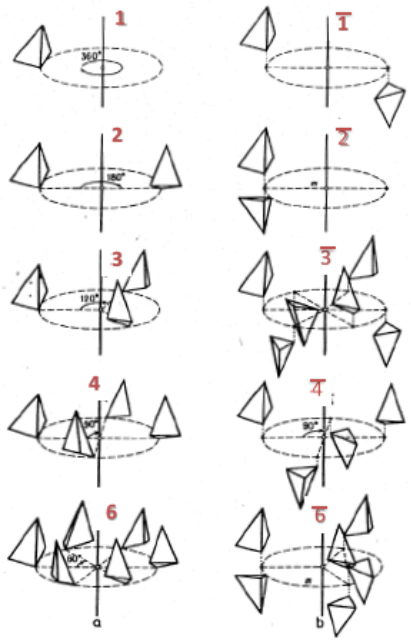


Figure 138.7: Proper Rotation Axes (2D & 3D), Improper Rotoinversion Axes (3D only)

138.2 Projections of Symmetry Elements

There are simpler ways of representing symmetry operations, namely through a system of stereographic projections. The top half of Figure138.8 depicts a view of a body with its symmetry axis pointing out of the page and the circle representing its perimeter. The symbol in the center indicates the 2-fold rotation symmetry operation, and the two solid circles depict the 180 rotation of the body. The darkened circle in the $\bar{2}$ symmetry operation projection represents a mirror plane.

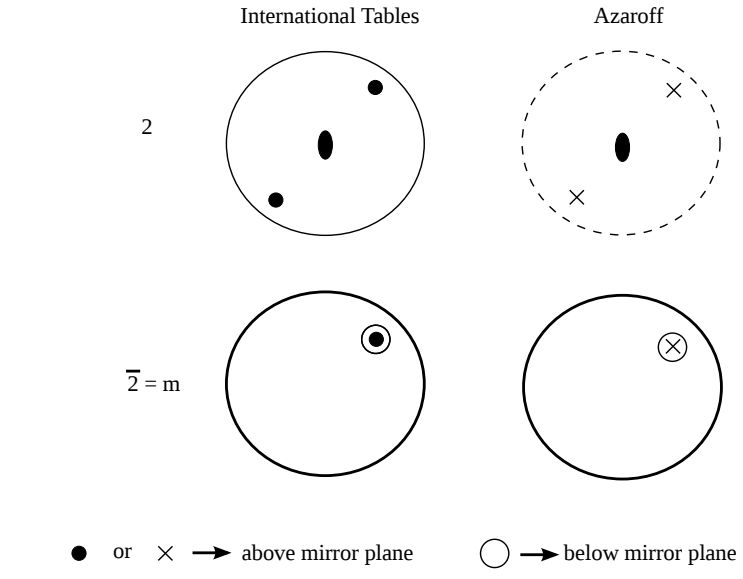


Figure 138.8: Stereographic representation of symmetry operations

Figure 138.9: Three bar symmetry operation

Arrangement of symmetry equivalent objects as an effect of a rotoinversion axis The following steps breakdown the $\bar{3}$ operations used to generate the object in Figure 138.10:

- Start at 1
- Rotate 120°
- Invert 2
- Rotate 120°
- Invert 3, etc. until you return to 1

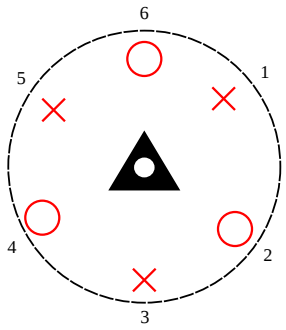
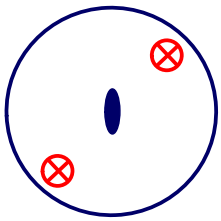


Figure 138.10: $\bar{3}$ symmetry example

Certain symmetry elements can be combined to form new patterns, such as the product of a 2-fold and inversion center shown in 138.11. This produces a 2-fold rotation axis which is perpendicular to a mirror plane as indicated by the darkened circle around the object.
Just like an algebraic operation, if A is one symmetry operation and B is another, they may be grouped through some operator to produce a third symmetry element, or C.

$$A \cdot B = C$$



$$2 \cdot \bar{1} = \frac{2}{m} \neq \bar{2}$$

Figure 138.11: Product of 2-fold and inversion center

For the 2mm symmetry operation shown in 138.12, the mirror plane lies parallel to the rotation axis. There is also a third mirror plane that arises from this combination, as shown by the second line drawn through the center of the dashed circle.

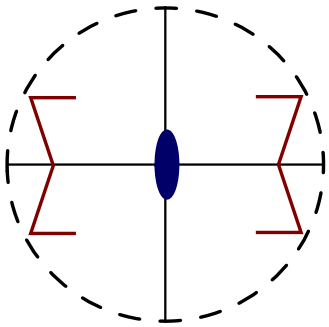


Figure 138.12: Stereographic projection of 2*mm* symmetry

Figure138.13 illustrates an object with symmetry that combines two rotation axes, or a 3-fold rotation axis perpendicular to a 2-fold rotation axis. Note that it does not matter whether the 3-fold or 2-fold operation is performed first, but the 3 2 symmetry is not equivalent to the 2 3 symmetry.

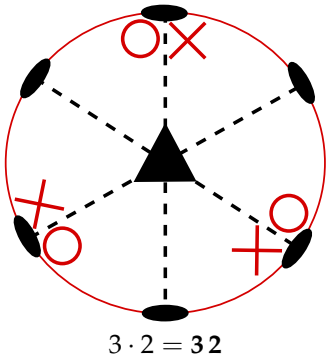


Figure 138.13: 3-fold perpendicular to a 2-fold rotation axis

Quartz Crystal

Note the point symmetry of a quartz crystal, as pictured in the figure below.

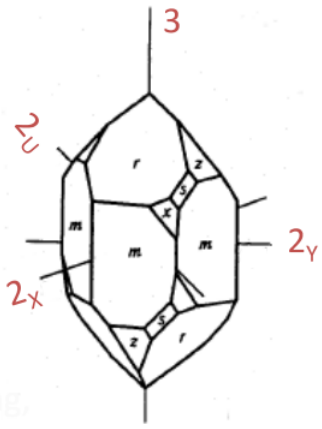


Figure 138.14: Quartz crystal

The perfect shape (or habit) is brought into self coincidence by (or is invariant to) the following symmetry operations:

$$g_0 = e, = 1$$

$$g_1 = 3,\ g_2 = 3^2$$

$$g_3 = 2_X,\ g_4 = 2_Y,\ g_5 = 2_U$$

Note that successive symmetry operations also bring the object into self coincidence.

$$g_1 \cdot g_1 = g_1^2 = g_2,\ g_1 \cdot g_4 = g_3$$

Note that $g_1$and $g_2$are the inverse of each other and that g_3, $g_4$and $g_5$are their own inverse operation. For instance, if you perform the $g_1$operation followed by the $g_2$operation, the object is at its starting position.

$$g_1 \cdot g_2 = g_0 \rightarrow g_1 = g_2^{-1}\ g_3 \cdot g_3 = g_3^2 = g_0 \rightarrow g_3 = g_3^{-1}$$

This is the **3 2** or $\mathbf{D_3}$ point symmetry group.

138.2.1 Group Theory Axioms

(1) A group G consists of an **operator** "multiplication" and a set of **elements** $g_i \in G$, such that the group is **closed** under "multiplication." $g_i \cdot g_j = g_k \in G$

(2) Multiplication is **associative:** $g_i \times (g_j \cdot g_k) = (g_i \cdot g_k) = (g_i \cdot g_j) \cdot g_k$

(3) There exists an identity element $e \in G$ such that $e \cdot g_i = g_i$

(4) For any g_i there is an inverse $g_i^{-1} \in G$ such that $g_i \cdot g_i^{-1} = e$

The completion of the following table is left as an exercise for the reader.

	g_0	g_1	g_2	g_3	g_4	g_5
g_0	g_0					
g_1		g_2	g_0			
g_2	g_2					
g_3						
g_4					g_0	
g_5						

Table 138.1: Multiplication of symmetry groups

The following demonstrates the multiplication for the **3 2** point symmetry group:

$$g_0 = e, g_1 = 3, g_2 = 3^2, g_3 = 2_x, g_4 = 2_Y, g_5 = 2_U$$

Note that the Cyclic group $C_3 = \{0, 1, 2\}$under modulo-3 addition is directly analogous to the 3-fold rotation axes point group.

There are only 32 point groups that can be generated by combinations of the 1,2,3,4,6,$\bar{1}$,m,$\bar{3}$,$\bar{4}$,$\bar{6}$symmetry operators, whose stereographic projections are shown in Figure 138.15. Furthermore, every crystal has a set of symmetry elements that is one of these 32 point groups or Crystal Classes. A crystal's highest degree of symmetry is used to describe its structure. Crystal classes are further separated into Crystal Systems based on minimal symmetry requirements. (See Hammond Table 3.1, Azaroff Table 1-2)

Crystal system	Minimal symmetry†	Crystal classes
Triclinic	1(or $\bar{1}$)	$1, \bar{1}$
Monoclinic	One 2 (or $\bar{2}$)	$2, m, \frac{2}{m}$
Orthorhombic	Three 2's (or $\bar{2}$)	$222, 2mm, mmm$
Tetragonal	One 4 (or $\bar{4}$)	$4, \bar{4}, \frac{4}{m}, 422, 4mm, \bar{4}2m, \frac{4}{m}\frac{2}{m}\frac{2}{m}$
Cubic(isometric)	Four 3's (or $\bar{3}$)	$23, 432, \frac{2}{m}\bar{3}, \bar{4}3m, \frac{4}{m}\bar{3}\frac{2}{m}$
Hexagonal	One 3 or 6 (or $\bar{3}$or $\bar{6}$)	$3, \bar{3}, 32, 3m, \bar{3}\frac{2}{m}, 6, \bar{6}, 622, 6mm, \bar{6}m2, \frac{6}{m}, \frac{6}{m}\frac{2}{m}\frac{2}{m}$

† The symmetry elements shown in this column are the least symmetry that a crystal can have and still belong to the corresponding crystal system. In this sense, this column shows the diagnostic symmetry elements of each system.

Table 138.2: Azaroff Table 1-2; Crystal systems and crystal classes

Figure 138.15: Stereographic projections of 32 point groups

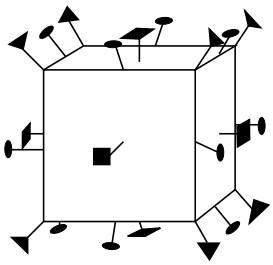


Figure 138.16: Rotation symmetry axes of a cube

Note that different crystal habits can have the same symmetry, as shown in Figure 138.17.

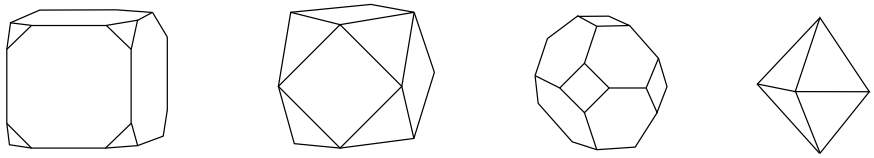


Figure 138.17: Crystal habits (Schartz & Cohen Fig. 1-6)

138.2.2 The 5 Point Groups belonging to a Cubic Crystal System

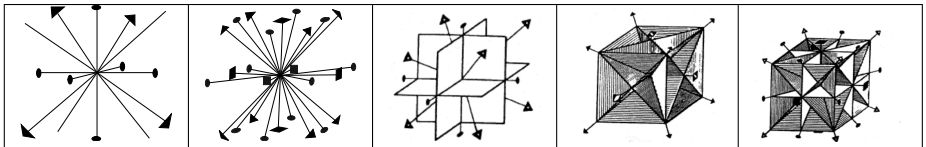


Figure 138.18: 5 Point groups - cubic crystal system

Note that a cubic crystal does not necessarily have 4-fold axes. (See Vainshtein "Modern Cryst." Vol 1 Fig 2.45-2.47)

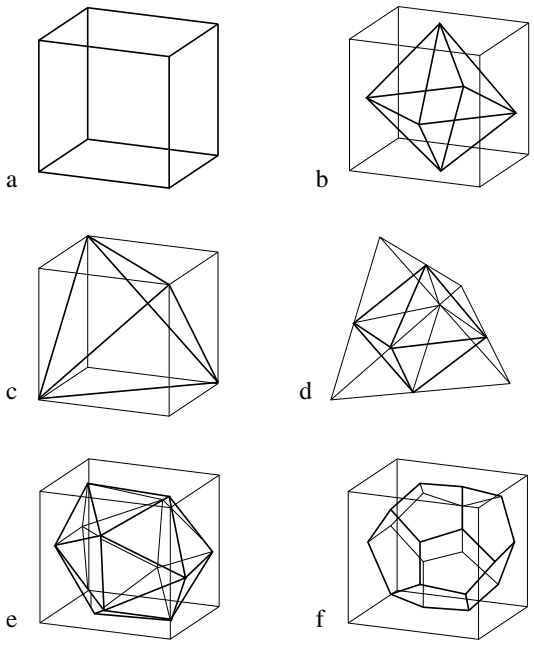


Figure 138.19: Crystal habits

Note the symmetry of the octahedron in Figure 138.20 with vertices at midpoints of square faces of cube. A regular octahedron has $m\bar{3}m$ point group symmetry, also known as $\frac{4}{m}\bar{3}\frac{2}{m}$. Figure 138.21 shows the projection of the octahedron on the [111] plane, with a $\bar{3}$ improper rotation axis.

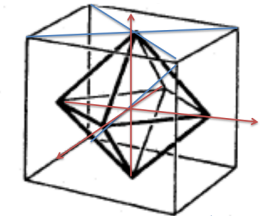


Figure 138.20: Octahedron

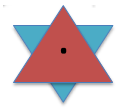
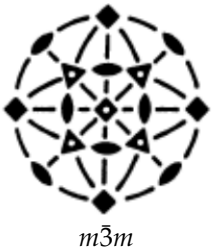


Figure 138.21: [111] projection



$m\bar{3}m$

Figure 138.22: Stereographic projection of octahedron symmetry

139 Crystal Lattices

139.1 Indexing within a crystal lattice

The unit cell of a crystal may be described by the following three parameters:
Crystal axes: a, b, c
Length: a, b, c
Angles: α, β, γ
Crystal axes, illustrated in Figure 139.1, share an origin at one corner of the unit cell. The angle between these axes, or α,β and γ, are based on the symmetry of the crystal. For instance, only the cubic crystal system has $\alpha = \beta = \gamma = 90$. The c-axis usually points to the highest symmetry element. Finally, the lengths of the axes are determined by the natural repeat lengths of the unit cell and based on the crystal habit.

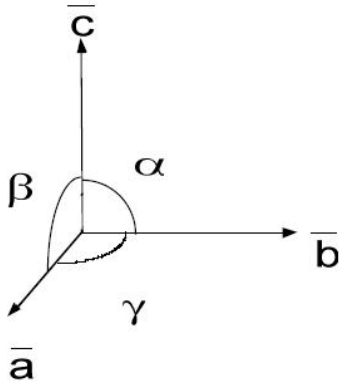


Figure 139.1: Crystal axes

The crystal axes may be used to specify planes within the crystal lattice structure using Miller indices. The following procedure is used to determine the Miller indices of a plane.

(1) Determine intercepts on crystal axes

(2) Take reciprocals of the intercepts

(3) Remove common factors

(4) Put in ( ), no commas

All parallel planes have the same (hkl) Miller Indices. If a plane crosses through the origin, as shown in the left axes in Figure 139.2, it can be translated to determine its intercepts. By convention, a bar over the number is used to indicate a negative indice.

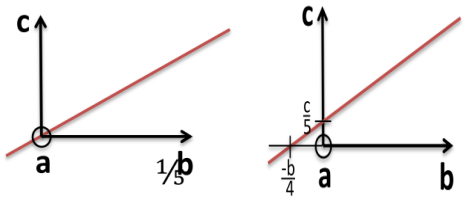


Figure 139.2: Determining Miller indices

a, b ,c Intercepts: $\infty, \frac{-1}{4}, \frac{1}{5}$

Reciprocals: 0, -4, 5

Miller Indices: $(0\ \bar{4}5)$

Determine the hkl indices of the plane shown in the figure below:

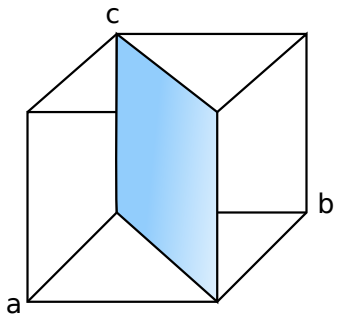


Figure 139.3: Plane in cubic unit cell

Note: Faces of cube are of the same ((hkl)) form ((100))

139.1.1 Hexagonal Prism

Indices of hexagonal prism faces are not of the same (hkl) form. Instead, a new index i is introduced to identify planes, in which $i = -(h + k)$ or $a_1 + a_2 = -a_3$. Thus, there are only three independent indices. The four-index $(hkil)$Bravais-Miller system more easily defines planes within a hexagonally structured crystal. Figure 139.4 shows indexing within a hexagonal structure using the three Miller indices, whereas Figure139.5 uses the a_1, a_2, a_3, c axes in the Bravais-Miller scheme.

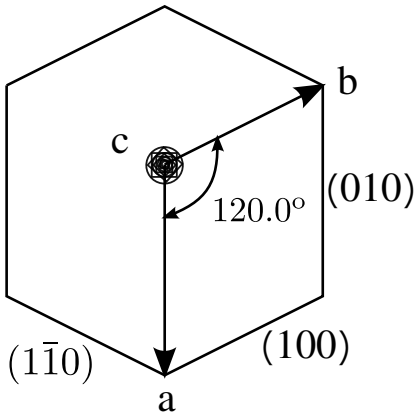


Figure 139.4: Using Miller indices (hkl)

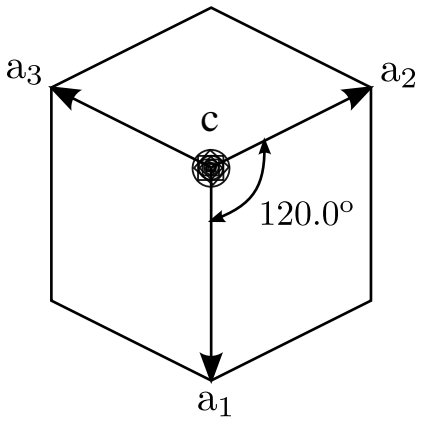


Figure 139.5: Using Bravais-Miller indices (hkil)

By using the Bravais-Miller scheme, the symmetry equivalent faces are all of the same form (1100).

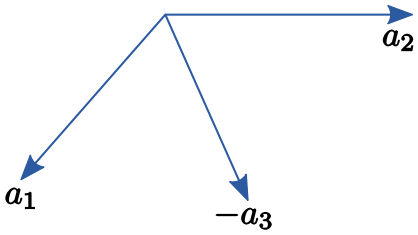


Figure 139.6: 3 indices

139.1.2 Form

Planes (or faces) of a form are related by crystal symmetry and generally have the same spacing but different indices. For instance, the six square faces of a cube (100), (010), (001),($\bar{1}$00), (0$\bar{1}$0), and (00$\bar{1}$), are of the form ((100)) or {100}. These planes are all related by the 4-fold rotation axes perpendicular to the cube faces. Note that a form is indicated by the indices of one plane enclosed in braces, or {hkl}.

Example: Octahedron
Faces are of the {111} plane form

What Crystal System?

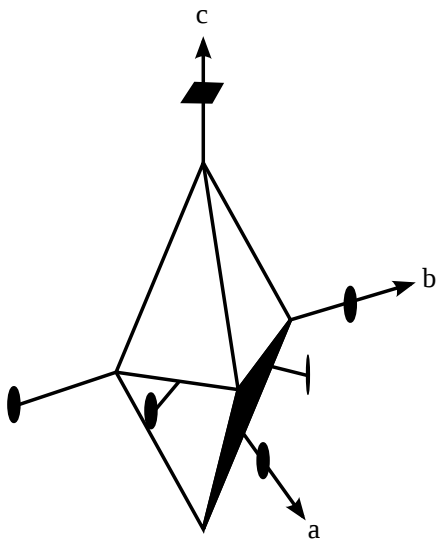


Figure 139.7: Crystal system (From Azaroff Fig. 1-24)

139.1.3 Direction Indices

Direction indices define a direction with respect to the unit cell axes - **a**, **b** and **c**.

$$[uvw] = u\mathbf{a} + v\mathbf{b} + w\mathbf{c}$$

139.8 demonstrates direction indices with respect to unit cell axes.

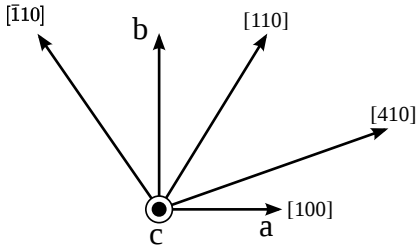


Figure 139.8: Direction Indices

139.1.4 Zone

Planes of a zone intersect at a line [uvw] known as a zone axis, to which they are all parallel. All (hkl) planes that contain this line belong to the same zone.
For instance:

- Let σ_{hkl}be the normal vector for the plane (hkl).
- Then $\sigma_{hkl} \cdot [uvw] = 0$ if hkl plane belongs to [uvw] zone
- For cubic unit cell $\sigma_{hkl} = c[hkl] \rightarrow hu + kv + lw = 0$
- [010] is the zone axis for (001), (100), (102), $(10\bar{2})$

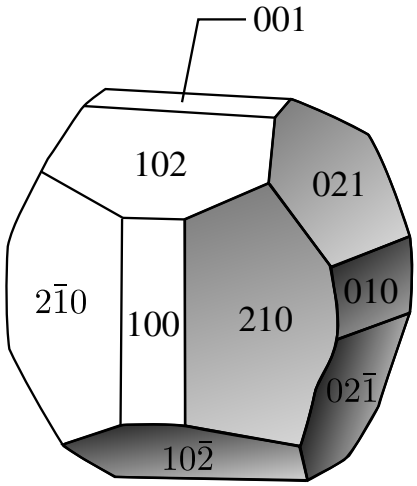


Figure 139.9: Zone axis of [010]

139.2 Lattices

A crystal lattice is a periodic array of objects, i.e. 7's, atoms, molecules ions. Each object has identical surroundings. Periodicity is generated by translation symmetry operators t_1, t_2, t_3.
The entire infinite array of objects can be described by giving the structure within the unit cell and by giving t_1, t_2, t_3.

1-D	t_1	**(a)**
2-D	t_1, t_2	**(a,b)**
3-D	t_1, t_2, t_3	**(a,b,c)**

Table 139.1: Translation symmetry operations

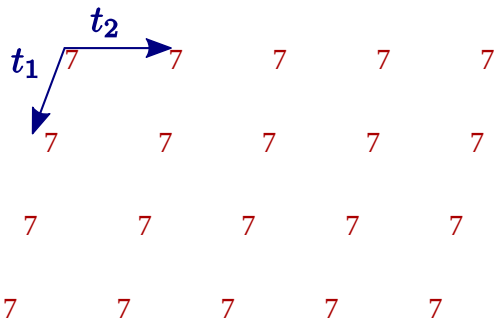


Figure 139.10: Array of 7's

Lattice points are an imaginary set of points which make up the vertices of a unit cell and are translated to make an array pertaining to crystal's structure. Each lattice point has identical surroundings.

2-D →Plane Lattice
3-D →Space Lattice

There are five possible plane lattices, or ways to arrange points in two dimensions such that each has identical surroundings. The five plane lattices, shown in Figure 139.11, include the oblique p-lattice, rectangular p-lattice, rectangular c-lattice, square p-lattice, and hexagonal p-lattice.

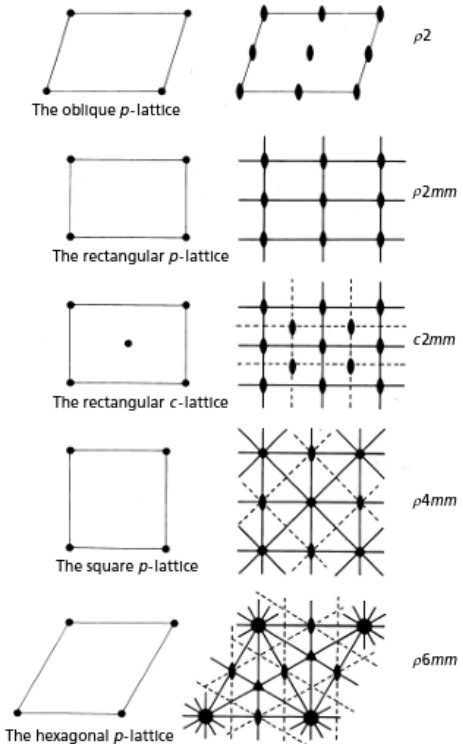


Figure 139.11: Five plane lattices (Hammond Fig. 2.4)

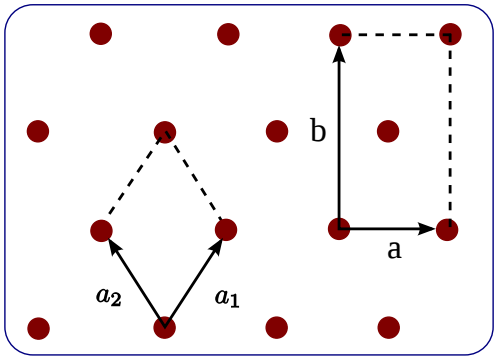


Figure 139.12: Parallelogram and square lattices

Azaroff	International Tables
Parallelogram	Oblique p
Rectangular	Rectangular P
Diamond	Rectangular C
Triangular	Hexagonal P
Square	Square P

Table 139.2: Two-dimensional lattices

The five types of plane lattices are grouped into four axial systems.

Axial System	Min. Sym.		
Oblique	1	$a \neq b$	$\gamma = 90$
Rectangular	m	$a \neq b$	$\gamma = 90$
Hexagonal	3	$a = b$	$\gamma = 120$
Square	4	$a = b$	$\gamma = 90$

Table 139.3: Four Axial Systems

Note: Only 1,2,3,4,6 are consistent with translation symmetry.

139.2.1 Unit cell

A unit cell is defined as a small region of space chosen to represent a lattice. It is defined by cell edges **a, b**. Translation of cell by **a, b** generates lattice.

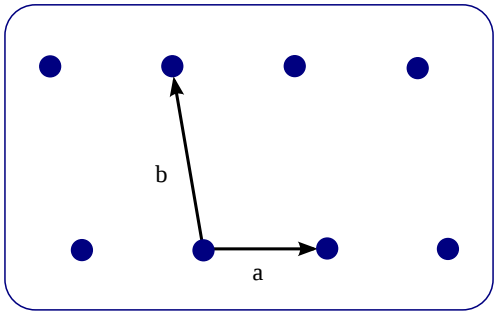


Figure 139.13: Unit cell edges

A primitive unit cell has one lattice point per cell and uses the smallest possible area to represent the lattice. A non-primitive unit cell, such as the rectangular-c unit cell in two dimensions, may contain one or more lattice points within the unit cell and is conventionally and conveniently chosen to emphasize the symmetry of the lattice.

139.2.2 Plane Groups

Two dimensional point symmetry is combined with two dimensional translation symmetry to create aplane group, as shown in Figure 139.14.

$$A \cdot t = B$$

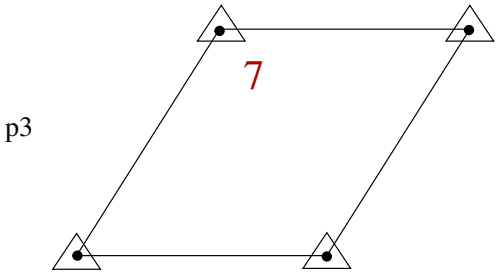


Figure 139.14: 3-fold and Hexagonal

There are 17 such plane groups within which any two dimensional periodic pattern may be classified. The shorthand names of these plane groups, as shown in Figure139.15 below, are generated by the following scheme:

the "1" represents no symmetry, the integers indicate the type of rotational symmetry, the "m" indicates the presence of a mirror plane, and the "g" stands for glide symmetry. Glide symmetry results from the combination of point group symmetry with the lattice. The first image in Figure 139.16 demonstrates a familiar mirror plane symmetry, with R's reflected across the solid lines, while the second image shows reflection with one R translated half a lattice spacing down. The dashed lines in the second image indicate a special type of symmetry known as a reflection glide or a glide line of symmetry. The glide lines give this pattern a translational form of symmetry along the lattice rather than a simple reflection. In crystal structures, this shift in glide symmetry would only be observable in an electron microscope which allowing for sight of distances of the order of 0.5–2 Å (50–200 pm). Lastly, note that 3-fold symmetry axes do not exist in oblique, rectangular or square lattices.

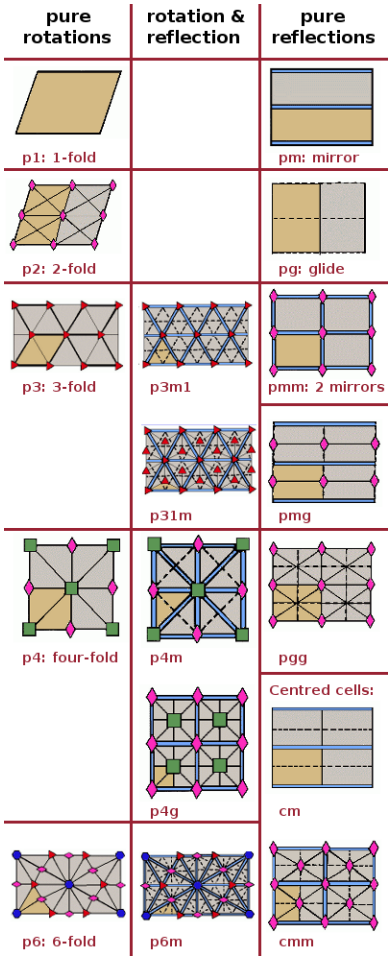


Figure 139.15: 17 Plane Groups

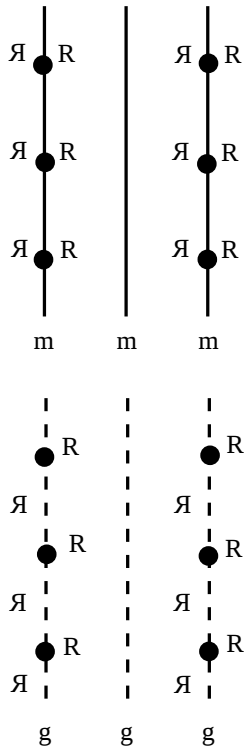


Figure 139.16: Mirror Plane and Glide Symmetry

139.2.3 Space Lattices or Bravais Lattices

Points in 3-D space can be arranged such that each has identical surroundings in 14 ways. These are divided into7 Crystal Systems.

Crystal System	Lattice Types		
Triclinic	P	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma$
Monoclinic	P,C	$a \neq b \neq c$	$\alpha = \gamma = 90 \neq \beta$
Orthorhombic	P,I,F,C	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90$
Tetragonal	P,I	$a = b \neq c$	$\alpha = \beta = \gamma = 90$
Trigonal	P,R	$a = b = c$	$\alpha = \beta = \gamma \neq 90$
Hexagonal	P	$a = b \neq c$	$\alpha = \beta = 90; \gamma = 120$
Cubic	P,I,F	$a = b = c$	$\alpha = \beta = \gamma = 90$

P: Primitive (1/cell) **I**:Body Centered (2/cell) **C**: Base Centered (2/cell)
F: Face Centered (4/cell) Rhombohedral pattern in Hex. Cell (3/cell)

Table 139.4: 7 Crystal Systems

139.2.4 Positions Within the Unit Cell

The origin of cell is arbitrary. By convention, it is the highest symmetry point. The position of any point in cell is given by fractional translations along a,b,&c, written with no parentheses. Table 139.5 notes the positions of lattice points using this scheme.
Translation to an identical point is achieved by the following: $\mathbf{t} = n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}$, where n is an integer.

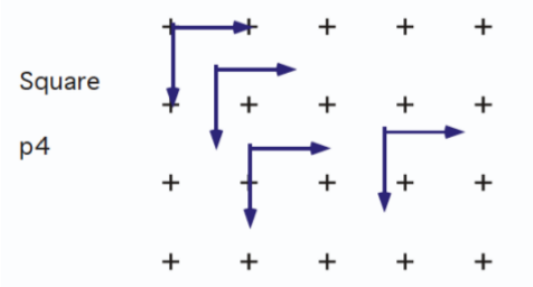


Figure 139.17: Translation along a square lattice

Body Centered	I	$000;\frac{1}{2}\frac{1}{2}\frac{1}{2}$
C-Centered	C	$000;\frac{1}{2}\frac{1}{2}0$
Face Centered 000+fct	F	$000;\ \frac{1}{2}\frac{1}{2}0;\frac{1}{2}0\frac{1}{2};0\frac{1}{2}\frac{1}{2}$

Table 139.5: Lattice Points

Hexagonal Close Packed Structure HCP

- E.g.; Zn, Mg, Be

The HCP structure has two atoms per unit cell (or two atoms per lattice point). The 000 and $\frac{1}{3}\frac{2}{3}\frac{1}{2}$ points do not have equivalent surroundings.

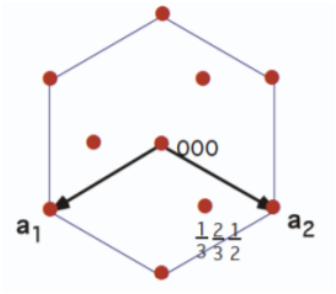


Figure 139.18: HCP

Rhombohedral (Hexagonal) For a Rhombohedral Primitive Cell:

- 1 lattice point per unit cell at 000

For a Hexagonal R Non-Primitive Unit Cell:

- 3 lattice points per unit cell at 000, $\frac{2}{3}\frac{1}{3}\frac{1}{3};\frac{1}{3}\frac{2}{3}\frac{2}{3}$

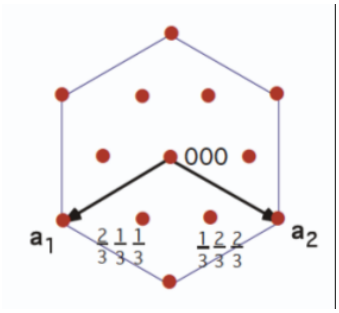


Figure 139.19: Hexagonal R Non-Primitive Unit Cell

Counting Lattice Points: 1 corner point per unit cell at 000, since 8 corners shared with 8 cells.

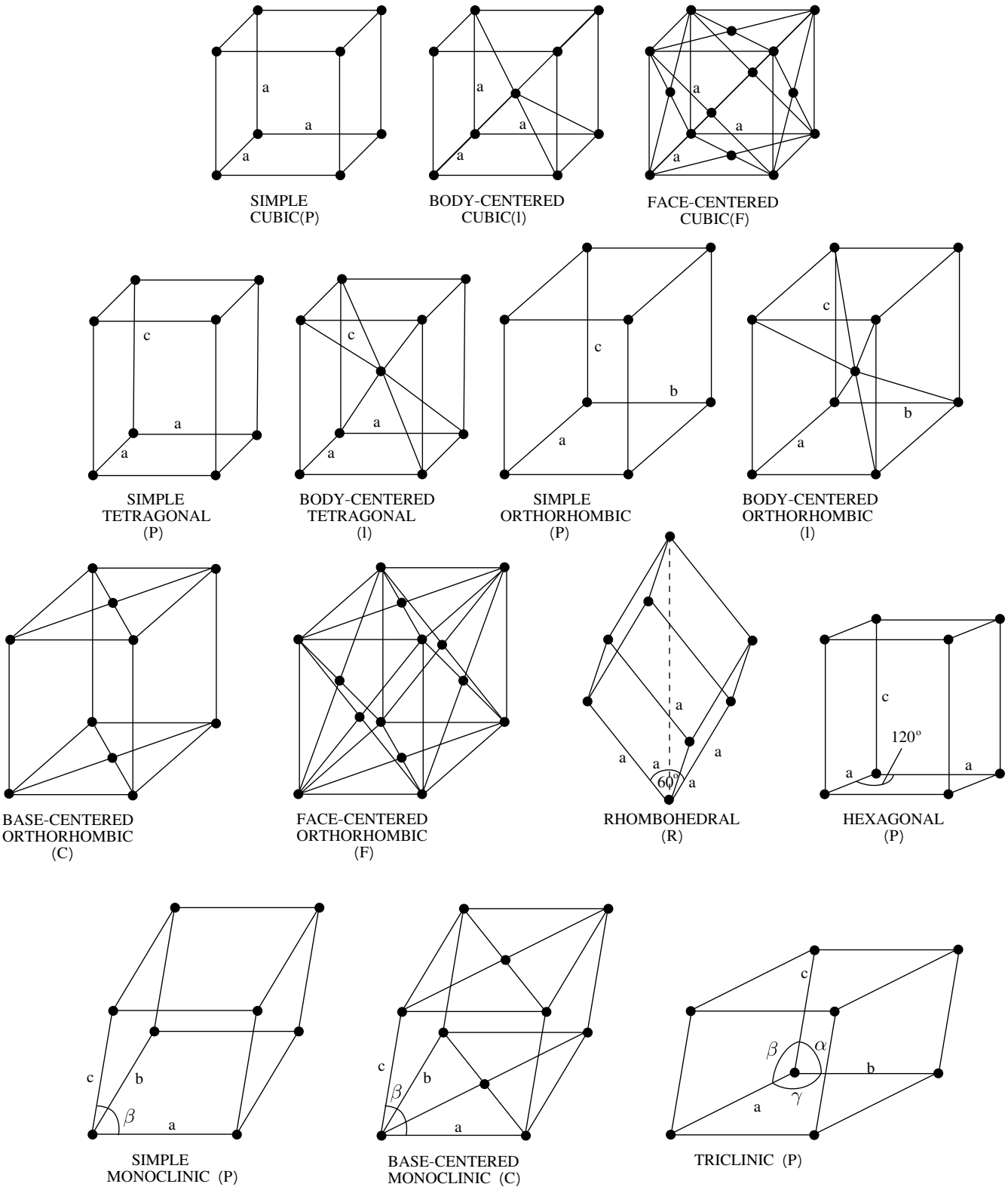


Figure 139.20: 14 Bravais Lattices

140 Stereographic Projections

Recall the regular octahedron illustrated in Figure 140.1 exhibiting various types of symmetry.

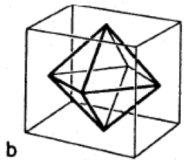


Figure 140.1: Octahedron - Cubic Crystal System

How did we position these Symmetry Rotation Axes and Mirror Planes in this 2D projection?

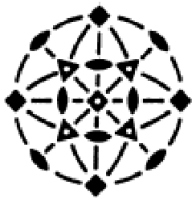


Figure 140.2: Octahedron - Stereographic Projection

Stereographic projections are a two-dimensional form of representing three-dimensional crystal structure. Namely, these are graphic representation of the orientation of planes. Stereographic projections allow for an easy visualization of crystallographic features. Looking at Figure 140.3, the normal to the (hkl) plane of the unit cell at the center of the sphere, (σ_{hkl}), intersects the reference sphere at the hkl pole.
Consider that all hkl planes pass through the center of the reference sphere.

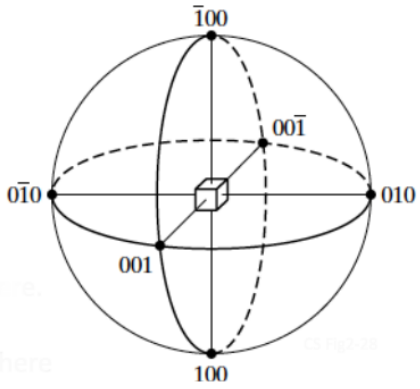


Figure 140.3: Reference Sphere

A plane can be represented by a trace made by the plane through the reference sphere.
Traces for planes passing through the center of reference sphere are Great Circles, or Longitudinal Lines (or Meridians).
Traces for planes perpendicular to diameter of reference sphere, but not passing through center are Small Circles or Latitudes (except those at the equator).

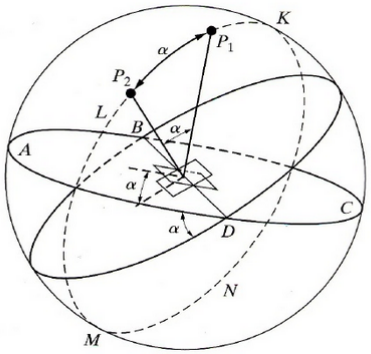


Figure 140.4: Reference Sphere with Traces

Rather than measuring angles between planes on the surface of the sphere, we employ an equi-angular stereographic projection. Note: geographers use equal area projection.

Angle between two planes

=Angle between their Great Circles (Angle-True)

=Angle between their Normals

=Angle along Great Circle connecting their Poles

The plane of the crystal being described is located at the center of the reference sphere, labeled "C" in Figure 140.5. The following terms describe useful parameters to define in the stereographic projection method:

Projection plane Placed normal to the end of any chosen diameter. See point "A" in Figure 140.5.

Point of projection Other end of chosen diameter. "B"
Note the projections of the Great Circles.

Basic Circle N′ E′ S′ W′ = Projection of Great Circle N E S W from plane N E S W with normal A B.
Steps for representing the angle for a plane C on a projection:

(1) Plane C →Normal CP

(2) Normal CP →Pole P

(3) Pole P →Stereographic Projection P′

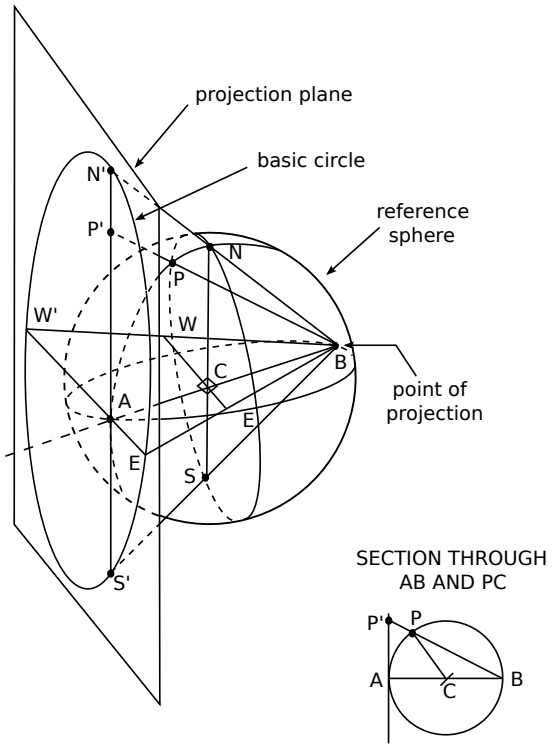
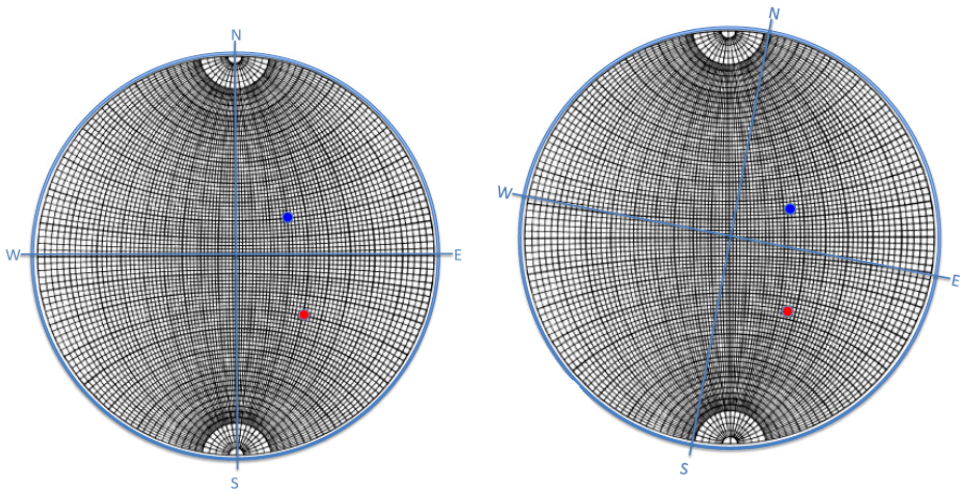


Figure 140.5: Projections from a reference sphere

140.0.1 Wulff Net

A Wulff Net is used to make angle measurements between planes using their surface normals that produce poles on the reference sphere (σ_{hkl}). Starting from the center of the Wulff Net and moving laterally outward, the angles increase from 0to 90. Angles between poles are measured along great circles (longitude lines, basic circle, or equator). The horizontal arcs are small circles.
To find the angle between two planes, rotate the projection around until both poles lie on same great circles, as shown in Figure 140.6.



Pole P1 at 20N, 30E
Pole P2 at 30S, 40E
Rotate →50°is the distance between the planes

Figure 140.6: Wulff Net

Standard Projections in a Cubic System Figure 140.7 below illustrates stereographic projections of cubic system planes on (001) (left) and (011) (right).

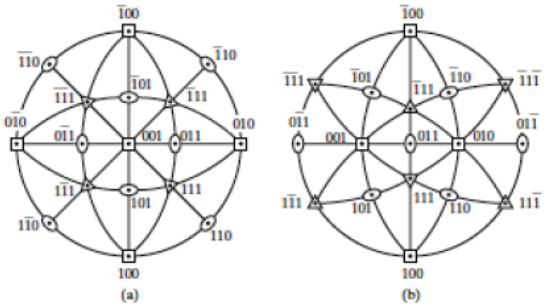


Figure 140.7: Standard projections of cubic crystals, from C&S textbook

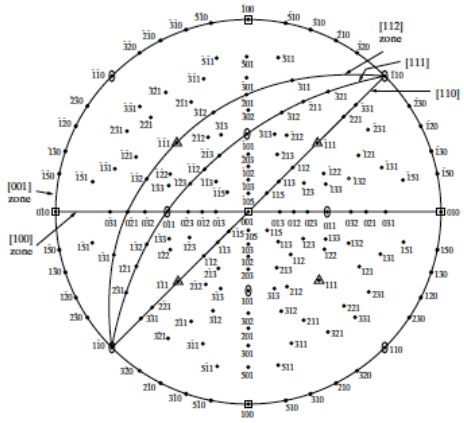


Figure 140.8: Standard (001) projection of cubic crystal, from C&S textbook

As shown in figure above, the [uvw] zone has (hkl) planes whose normals σ_{hkl} are perpendicular to **[uvw]**.These normals lie in a plane whose intersection with the reference sphere forms a trace.
Note: See C&S Table 2-4 for listing of cubic inter-planar angles.

140.1 Representing (hkl) Planes with Vectors

When representing planes within a crystallographic plane, the vector direction is perpendicular to the (hkl) plane. The vector magnitude is determined by using interplanar spacing, also known as d-spacing (d_{hkl}).

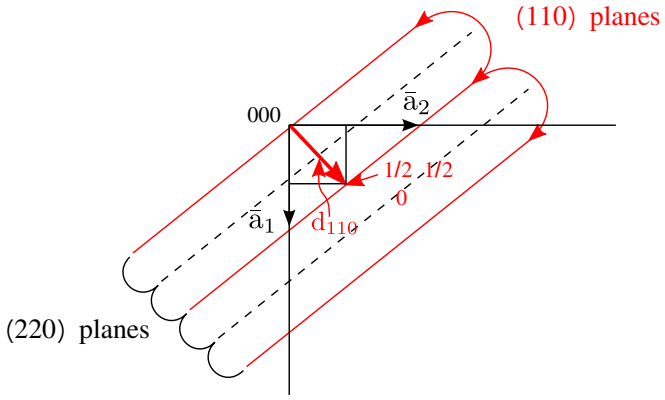


Figure 140.9: Cubic

$$d_{hkl} = \frac{a}{\sqrt{h^2+k^2+l^2}}$$

$$d_{110} = \frac{a}{\sqrt{2}}$$

$$d_{220} = \frac{a}{2\sqrt{2}}$$

Define vector ,$\bar{d}_{hkl}$ to represent cubic (hkl) planes:

$\bar{d}_{hkl} \perp (hkl)$ planes

Magnitude: $d_{hkl} = \frac{a}{\sqrt{(h^2+k^2+l^2}}$

Direction: $[hkl] = h\bar{a}_1 + k\bar{a}_2 + l\bar{a}_3$

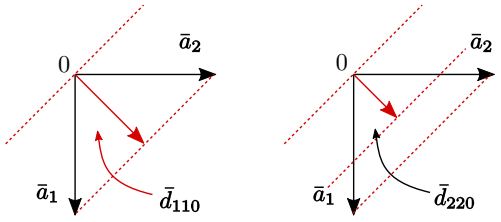


Figure 140.10: Defining$\bar{d}_{hkl}$

Unfortunately, this scheme only works for cubic.
Try orthorhombic, $\alpha = \beta = \gamma = 90$

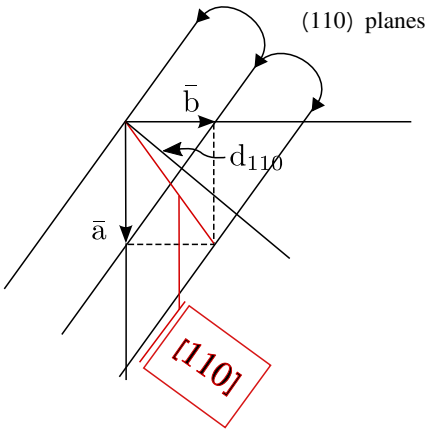


Figure 140.11: Orthorhombic Example

For the orthorhombic case, the [110] vector is not perpendicular to the (110) planes, therefore the plane spacing is not constant. We need a vector scheme that works for all axial systems, not just cubic. This can be achieved by using the reciprocal lattice concept.

Concept of Reciprocal Lattice Any (hkl) set of diffraction planes will be represented by a point whose coordinate is hkl, such that:

(1) The vector r^*_{hkl} which starts at the origin 000 and ends at hkl is normal to the (hkl) plane.

(2) The distance from the origin to hkl $|r^*_{hkl}| = 1/d_{hkl}$where $d_{hkl}$is the interplanar spacing

Consider a monoclinic crystal with $a = 4\mathring{A}, b = 3\mathring{A}, c = 6\mathring{A}$,and $\gamma = 70$
Show reciprocal lattice points 000,100, and 0101 in relation to direct space axes **a** and **b**.

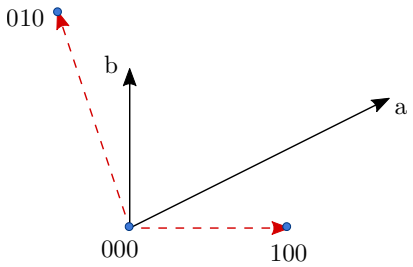


Figure 140.12: Reciprocal lattice points for a monoclinic crystal

Construct a hk0 grid corresponding to the (hkl) planes belonging to the [001] zone.

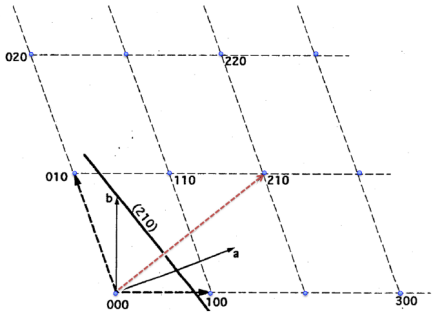


Figure 140.13: hk0 grid

Note: $r^*_{210} \perp (210)$plane

141 Reciprocal Lattice

The reciprocal lattice is a vector representation of direction of planes and their inter-planar spacing. It is important for understanding the geometry of diffraction patterns.

Consider a primitive lattice in direct space, whose unit cell is defined by $\bar{a}$,$\bar{b}$,$\bar{c}$ $(a, b, c, \alpha, \beta, \gamma)$.

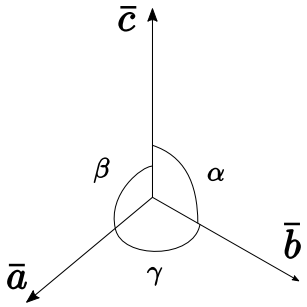


Figure 141.1: Unit cell defined by $\bar{a}$,$\bar{b}$,$\bar{c}$ axes

The reciprocal axes, $\bar{a}^*, \bar{b}^*, \bar{c}^*$, are illustrated in Figure 141.2.

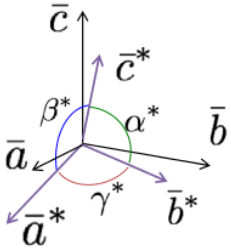


Figure 141.2: Reciprocal axes

The following equations define the reciprocal axes:

$$\bar{a}^* = \frac{\bar{b}\times\bar{c}}{\bar{a}\cdot(\bar{b}\times\bar{c})}$$

$$\bar{b}^* = \frac{\bar{c}\times\bar{a}}{\bar{b}\cdot(\bar{c}\times\bar{a})}$$

$$\bar{c}^* = \frac{\bar{a}\times\bar{b}}{\bar{c}\cdot(\bar{a}\times\bar{b})}$$

Note:

- $\bar{a}^* \perp bc$ plane $\bar{b}^* \perp ac$plane $\bar{c}^* \perp ac$ plane
- $\bar{a}\cdot(\bar{b}\times\bar{c}) = \bar{b}\cdot(\bar{c}\times\bar{a}) = \bar{c}\cdot(\bar{a}\times\bar{b}) = V$ =Volume of unit cell
- $\bar{a}^*\cdot\bar{a} = 1, \bar{b}^*\cdot\bar{b} = 1, \bar{c}^*\cdot\bar{c} = 1$ Reciprocal Relationship

Special Case:
For $\alpha = \beta = 90$(i.e. all but triclinic or rhombohedral)

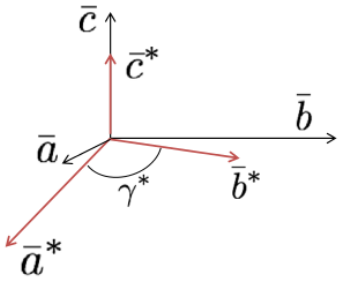


Figure 141.3: Special Case ($\alpha = \beta = 90$)

$$\bar{c}^* \| \bar{c}$$

$$\gamma^* = 180 - \gamma$$

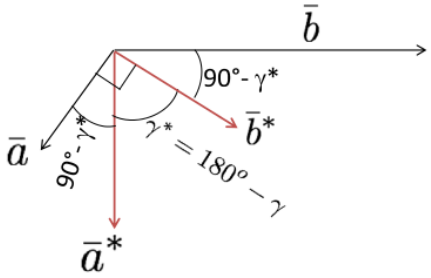


Figure 141.4: Special Case

$$\gamma = 90 - \gamma^* + 90 - \gamma^* + \gamma^*$$

$$\gamma = 180 - \gamma^*$$

For $\alpha = \beta = \gamma = 90$ →Cartesian axial system
(Cubic, tetragonal, orthorhombic)
$\bar{a}^* \| \bar{a}$, $\bar{b}^* \| \bar{b}$, $\bar{c}^* \| \bar{c}$
And...

$$|\bar{b} \times \bar{c}| = bc$$

$$|\bar{a} \cdot \bar{b} \times \bar{c}| = abc = V$$

$$|\bar{a}^*| = |\frac{\bar{b} \times \bar{c}}{\bar{a} \cdot \bar{b} \times \bar{c}}| = \frac{bc}{abc} = \frac{1}{a}$$

$$a^* = \frac{1}{a}, b^* = \frac{1}{b}, c^* = \frac{1}{c}$$

...hence "reciprocal."

Reciprocal Lattice Vector The reciprocal lattice vector extends in reciprocal space from origin 000 to rectangular lattice point and is perpendicular to the (hkl) plane. When the vectors of many planes are mapped out, these generate a reciprocal lattice. These vectors have units of 1/length, so a larger vector indicates smaller spacing between planes.

$$r^*_{hkl} = ha* + kb* + lc*$$

Note: Hammond - $r^*_{hkl} = d^*_{hkl}$ Cullity - $r^*_{hkl} = H_{hkl}$

The following examples show applications of reciprocal lattice vectors in determining inter-planar relationships.

(1) Find inter-planar angle ϕ between $(h_1k_1l_1)$and $(h_2k_2l_2)$

i.e.find angle between reciprocal lattice vectors r^*_1& r^*_2.

$$cos(\phi) = \frac{\bar{r}^*_1 \cdot \bar{r}^*_2}{r^*_1 r^*_2} \; r^*_1 = |\bar{r}^*_1| = \sqrt{\bar{r}^*_1 \cdot \bar{r}^*_2}$$

$$\bar{r}^*_1 \cdot \bar{r}^*_2 = (h_1\bar{a}* + k_1\bar{b}* + l_1\bar{c}*) \cdot (h_2\bar{a}* + k_2\bar{b}* + l_2\bar{c}*)$$
$$= h_1h_2a*^2 + k_1k_2b*^2 + l_1l_2c*^2 + (h_1k_2 + k_1h_2)\bar{a}* \cdot \bar{b}* + (h_1l_2 + l_1h_2)\bar{a}* \cdot \bar{c}* + (k_1l_2 + l_1k_2)\bar{b}* \cdot \bar{c}*$$

Note: $\bar{a}* \cdot \bar{b}* = a*b*cos(\gamma*)$,etc.

Example:

Find angle between (111) and (110) in cube?

$$cos(\phi) = \frac{\bar{r}^*_{111} \cdot \bar{r}^*_{110}}{r^*_{111} r^*_{110}} = \frac{(1+1+0)a^{*2}}{\sqrt{3}a^* \sqrt{2}a^*} = \frac{2}{\sqrt{6}} \Rightarrow \phi = 35.26$$

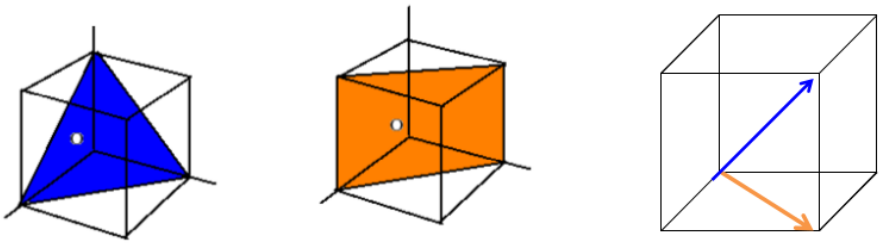


Figure 141.5: Angle between (111) and (110) in cube

(2) Find the zone axis for two intersecting planes

$(h_1k_1l_1)$and $(h_2l_2k_2)$

$$\bar{z} = [uvw] = u\bar{a} + v\bar{b} + w\bar{c}$$

Since $\bar{r}_{hkl}* \perp (hkl)$plane, $\bar{r}_{hkl}* \perp \bar{z}$ for any (hkl) belonging to zone $\bar{z}$

$$\bar{z} \cdot \bar{r_1}* = uh_1 + vk_1 + wl_1 = 0$$

$$\bar{z} \cdot \bar{r_2}* = uh_2 + vk_2 + wl_2 = 0$$

Solve simultaneous equations for u, v, w

$$u = k_1l_2 - l_1k_2$$

$$v = l_1h_2 - h_1l_2$$

$$w = h_1k_2 - k_1h_2$$

2 equations and 3 unknowns →uvw determined to within a scale factor

Example:
What is the zone axis (or edge formed by) for $(\bar{1}10)$ and (010)?

$$u = (1)(0) - (0)(1) = 0$$

$$v = (0)(0) - (-1)(0) = 0$$

$$w = (-1)(1) - (1)(0) = -1$$

$$\bar{z} = [uvw] = [00\bar{1}] \text{or } [001] = \bar{c} - axis$$

Equivalent to doing vector cross product:

$$\bar{r}^*_{\bar{1}10} \times \bar{r}^*_{010}$$

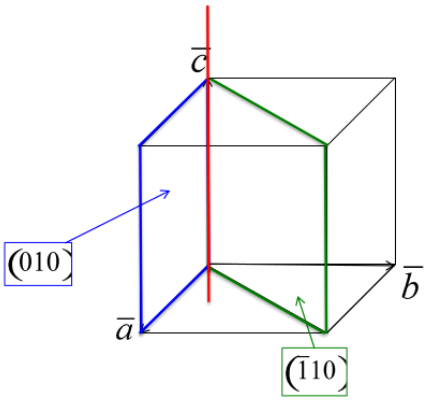


Figure 141.6: Zone axis for $(\bar{1}10)$ &(010)

(3) Find d-spacing d_{hkl}

Use $|\bar{r}^*_{hkl}| = r^*_{hkl} = \frac{1}{d_{hkl}}$
d_{hkl} =⊥distance between 2 || planes with intercept differences = $\frac{a}{h}, \frac{b}{k}, \frac{c}{l}$

$r^*_{hkl} = \sqrt{\bar{r}^*_{hkl} \cdot \bar{r}^*_{hkl}}$
$r^{*2}_{hkl} = h^2a*^2 + k^2b*^2 + l^2c + 2hk\bar{a}* \cdot \bar{b}* + 2hl\bar{a}* \cdot \bar{c}* + 2kl\bar{b}* \cdot \bar{c}*$

For $\alpha = \beta = \gamma = 90$, $2hk\bar{a}* \cdot \bar{b}* + 2hl\bar{a}* \cdot \bar{c}* + 2kl\bar{b}* \cdot \bar{c}* = 0$

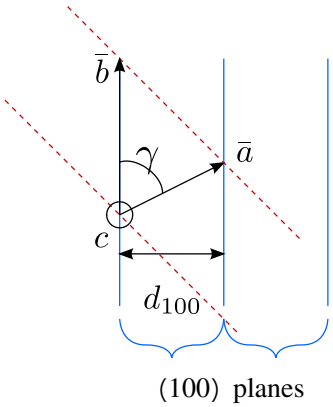


Figure 141.7: d_{100} spacing

Note: $\bar{a}* \cdot \bar{b}* = a* b* cos(\gamma*)$

For Orthorhombic - $d_{hkl} = \frac{1}{r_{hkl}*} = \frac{1}{(\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2})^{1/2}}$

For Cubic - $d_{hkl} = \frac{a}{(h^2+k^2+l^2)}$

Monoclinic
$(\alpha = \beta = 90)$ $a = 4\mathring{A}$, $b = 3\mathring{A}$, $c = 6\mathring{A}$, $\gamma = 70$

$r_{100}* = a* = \frac{1}{d_{100}} \neq \frac{1}{4\mathring{A}}$ $\quad d_{100} = a*^{-1}$ $\quad a* =? \neq \frac{1}{a}$ since $\gamma \neq 90$

Use $\bar{a}* \cdot \bar{a} = 1 \rightarrow a* acos(\frac{\pi}{2} - cos\gamma) = 1$
$a* = \frac{1}{asin(\gamma)} = \frac{0.250}{0.94} = 0.266\mathring{A}^{-1}$ $\quad \therefore d_{100} = asin(\gamma) = 3.759\mathring{A} \neq 4\mathring{A}$

Likewise, $b* = \frac{1}{bsin(\gamma)} = \frac{0.333}{0.94} = 0.355\mathring{A}^{-1}$ $\quad \therefore d_{010} = bsin(\gamma) = 2.819\mathring{A}.$

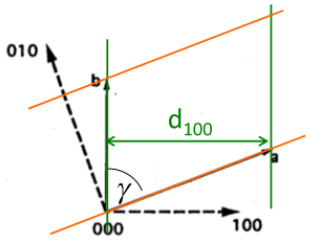


Figure 141.8: Monoclinic d_{100}spacing

Note: For monoclinic, $\frac{a*}{b*} = \frac{b}{a}$ $\quad c* = \frac{1}{c} = \frac{1}{6}\mathring{A}^{-1}$ since $\bar{c} \parallel \bar{c}*$

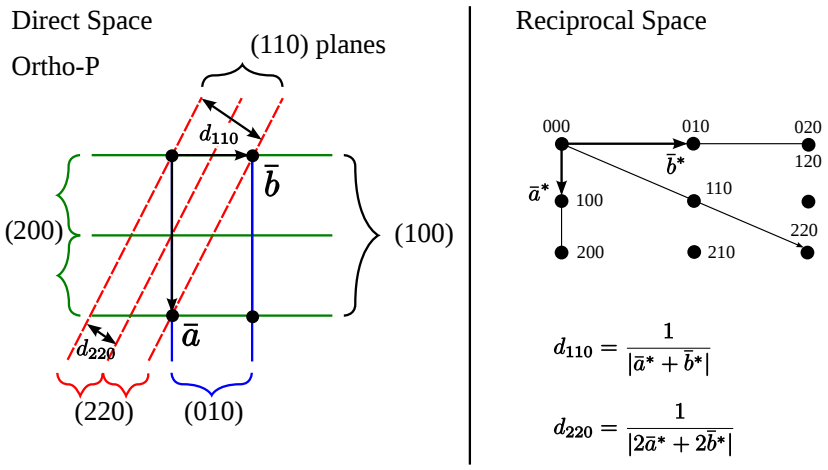


Figure 141.9: Direct vs Reciprocal Space for Ortho-P

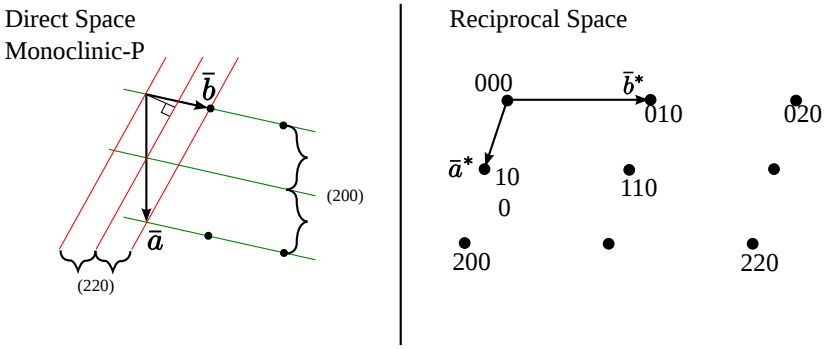


Figure 141.10: Direct vs Reciprocal Space for Monoclinic-P

(4) Cell axis transformations (primitive →non-primitive)
e.g. Rhombohedral ($\alpha = \beta = \gamma = 60$) →FCC

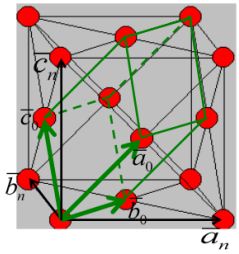


Figure 141.11: Cell axis transformation

Let old axes be $\bar{a}_0, \bar{b}_0, \bar{c}_0$
The new axes $\bar{a}_n, \bar{b}_n, \bar{c}_n$are related to the old by:

$$\bar{a}_n = u_1\bar{a}_0 + v_1\bar{b}_0 + w_1\bar{c}_0$$

$$\bar{b}_n = u_2\bar{a}_0 + v_2\bar{b}_0 + w_2\bar{c}_0$$

$$\bar{c}_n = u_3\bar{a}_0 + v_3\bar{b}_0 + w_2\bar{c}_0$$

or:

$$\begin{pmatrix} \bar{a}_n \\ \bar{b}_n \\ \bar{c}_n \end{pmatrix} = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \begin{pmatrix} \bar{a}_0 \\ \bar{b}_0 \\ \bar{c}_0 \end{pmatrix}$$

Matrix of transformation

For rhombohedral to FCC example:

$\bar{a}_n = \bar{a}_0 + \bar{b}_0 - \bar{c}_0$

$\bar{b}_n = -\bar{a}_0 + \bar{b}_0 + \bar{c}_0$

$\bar{c}_n = \bar{a}_0 - \bar{b}_0 + \bar{c}_0$

Transformation Matrix: $\begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$

Modulus of Transformation:

Determinant $\begin{vmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{vmatrix} = 4 \rightarrow V_n = 4V_0$ (There is four times the volume of the rhombohedral unit cell in the FCC unit cell)

For transforming $(h_0k_0l_0) \longleftrightarrow (h_nk_nl_n)$

Since the interplanar spacing and direction are the same, the planes have the same reciprocal lattice vector, $\bar{r}^*_{h_0k_0l_0} = \bar{r}^*_{h_nk_nl_n}$.

$$(h_0\bar{a}^*_0 + k_0\bar{b}^*_0 + l_0\bar{c}^*_0 = (h_n\bar{a}^*_n + k_0\bar{b}^*_n + l_n\bar{c}^*_n)$$

take dot product with $\bar{a}_n = u_1\bar{a}_0 + v_1\bar{b}_0 + w_1\bar{c}_0$

$$(u_1\bar{a}_0 + v_1\bar{b}_0 + w_1\bar{c}_0) \cdot (h_n\bar{a}^*_n + k_0\bar{b}^*_n + l_n\bar{c}^*_n) = h_n$$

$$\begin{array}{l} h_n = u_1h_0 + v_1k_0 + w_1l_0 \\ k_n = u_2h_0 + v_2k_0 + w_2l_0 \\ l_n = u_3h_0 + v_3k_0 + w_3l_0 \end{array} \rightarrow \begin{pmatrix} h_n \\ k_n \\ l_n \end{pmatrix} = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \begin{pmatrix} h_0 \\ k_0 \\ l_0 \end{pmatrix}$$

Same matrix of transformation used for axes and indices.

Old Rhombohedral →New FCC example

$$\begin{pmatrix} h_n \\ k_n \\ l_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} h_0 \\ k_0 \\ l_0 \end{pmatrix}$$

$$\begin{array}{c} h_n = h_0 + k_0 - l_0 \\ k_n = -h_0 + k_0 + l_0 \\ l_n = h_0 - k_0 + l_0 \end{array}$$

Note that $h_n + k_n = 2k_0$, or an even integer and $k_n + l_n = 2l_0$, also an even integer.
Therefore,h_n, k_n, l_nmust all be even or all odd "unmixed."

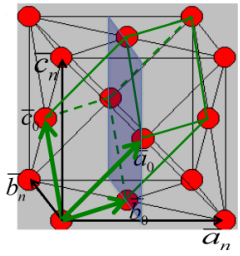


Figure 141.12: Rhombohedral to FCC

Transformation of the (110) plane in the rhombohedral structure to its (hkl) in the face-centered cubic structure.

$$(110)_R = (???)_{FCC} \begin{pmatrix} h_n \\ k_n \\ l_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = (200)$$

In Figure 141.13 below, a non-primitive orthorhombic-c unit cell is illustrated, with the c axis pointing out of the page. The transformation of its lattice in real space to reciprocal space is shown in Figure 141.14. Notice that the a and b axes only reach halfway to the first lattice point in their respective directions. The reason for this is due to the organization of lattice planes in the **a** and **b** directions of the real space. For instance, going vertically in the **a** direction of Figure 141.13, the first set of planes appears halfway between the lattice spacing. The planes that follow in this **a** direction appear at the same spacing. Thus, this set of (200) planes within the Orthorhombic-c unit cell, can represent the fundamental period of planes in the **a** direction. The same analysis can be done in the **b** direction, yielding the (020) fundamental period of planes. From this, the first point in the reciprocal space appears at 200 in the $a*$direction and 020 in the $b*$ direction. Note that in the **c** direction, which is not shown, the first plane is found at (001), due to the base-centered structure.

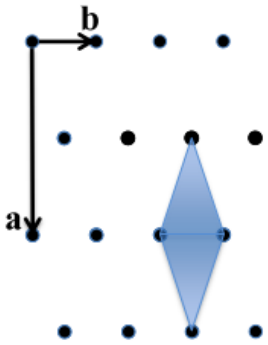


Figure 141.13: Orthorhombic-C

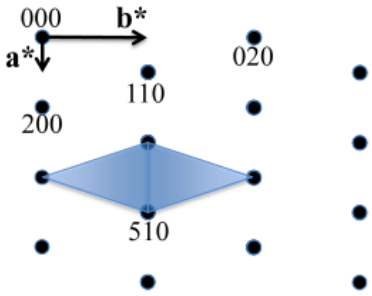


Figure 141.14: Reciprocal Lattice l=0 layer

Note that the original, real lattice and the reciprocal lattice have a similar shape, with the aspect ratio appearing to switch in the reciprocal lattice. Furthermore, they are both base-centered and have a similar primitive unit cell, which is monoclinic. The missing base-centered reciprocal lattice points (e.g. 010, 100, 210, etc.) are due to the use of a non-primitive unit cell in real space for the transformation. Had the primitive cell been used, all allowed reciprocal lattice points, corresponding to allowed (hkl) Bragg diffraction planes, would be present.
A two dimensional projection of the FCC unit cell on the **a** and **b** axes is illustrated in Figure 141.15 below. The solid circles represent the lattice points in the plane of the paper, and the circles with the $\frac{1}{2}$ label represent lattice points half a lattice spacing above the plane. The dashed lines indicate the set of (110) planes within this unit cell.
The transformation from the FCC to its reciprocal lattice generates a BCC structure as shown in Figure 141.16, and vice versa. Note that the (100) planes are missing, since the fundamental period of planes is (200) in the **a** direction. This reasoning can be applied to other missing reciprocal space points, using the fundamental period of planes in the real FCC lattice.

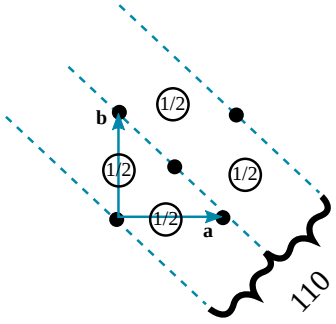


Figure 141.15: FCC

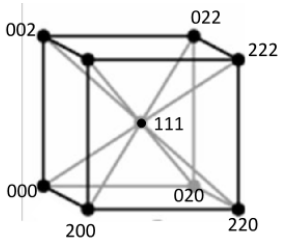


Figure 141.16: Reciprocal Lattice - hkl unmixed (all even or all odd)

142 Representative Crystal Structures

Crystal Structure Examples The combination of the previously discussed point and translation symmetry elements with the fourteen Bravais lattices give rise to 230 possible three-dimensional space groups. For these three-dimensional crystal structures, an identical atom or a fixed arrangement of atoms (ions, molecules) is found at each Bravais lattice point.

The following examples show a one atom per lattice point structure. Another example which is not shown below is the simple cubic structure.
Ex.: BCC Metals $\alpha - Fe$,Cr, Mo, V, etc.

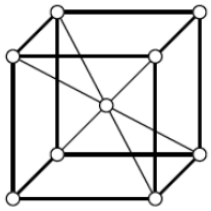


Figure 142.1: BCC with one atom per lattice point

Ex.: FCC Metals $\gamma-$Fe, CU, Pb, Ni, etc.

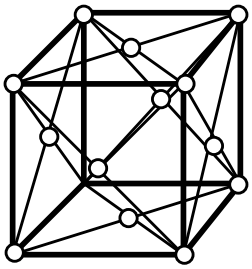


Figure 142.2: FCC with one atom per lattice point

The hexagonal close packed structure, unlike the examples in Figures 142.1 and 142.2, has two atoms per lattice point. These two atoms are considered to be the motif, or repeating object within the HCP lattice. Ex.: HCP (hexagonal close-packed)

Figure 142.3 shows four different representations of a hexagonal close-packed structure with two atoms per unit cell - one atom at 000, and one atom at $\frac{2}{3}\frac{1}{3}\frac{1}{2}$or $\frac{1}{3}\frac{2}{3}\frac{1}{2}$. The nine atoms marked with an X in Figure 142.3(b) correspond ot the nine atoms illustrated in Figure 142.3(a). The association of two atoms with a single lattice point is represent by the dashed lines connecting the pairs. Furthermore, (c) and (d) show more of a hexagonal representation of the HCP atomic structure with the stacking of atomic planes in a periodic arrangement. Some elements with this structure include zinc, magnesium, and beryllium.

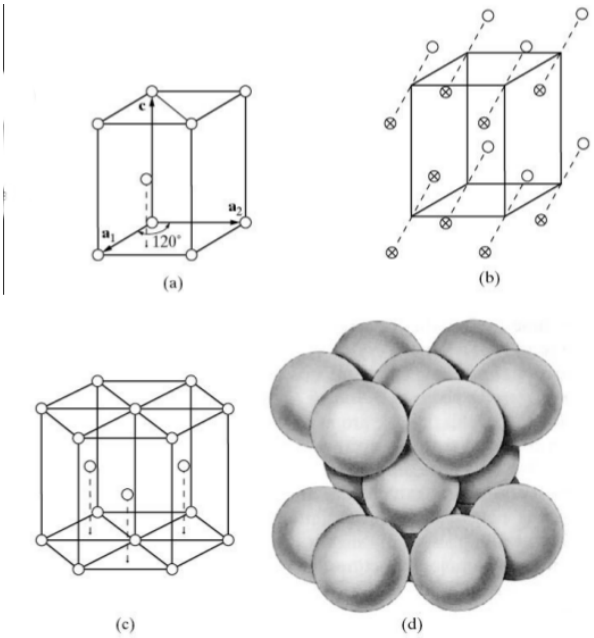


Figure 142.3: HCP structure

The FCC structure is as close-packed as the HCP structure. Though its relation to the HCP structure is not immediately apparent, Figure 142.4 demonstrates the stacking of FCC (111) planes in a hexagonal pattern, similar to the (0002) HCP planes stacking. One distinction between the two is that the HCP layer stacking occurs in an ABAB pattern in the $\bar{c}$ or [001] direction, and the layer stacking for FCC occurs in an ABCABC pattern in the [111] direction. This is illustrated in Figures 142.5 and 142.7 below.

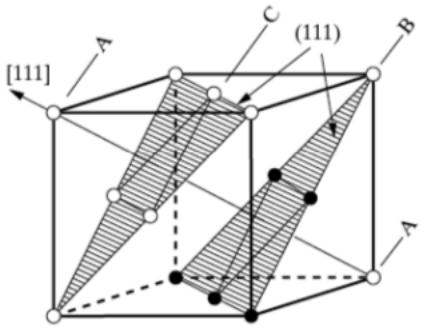


Figure 142.4: Face-centered cubic planes

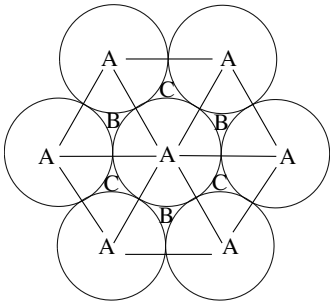


Figure 142.5: Stacking of (111) planes

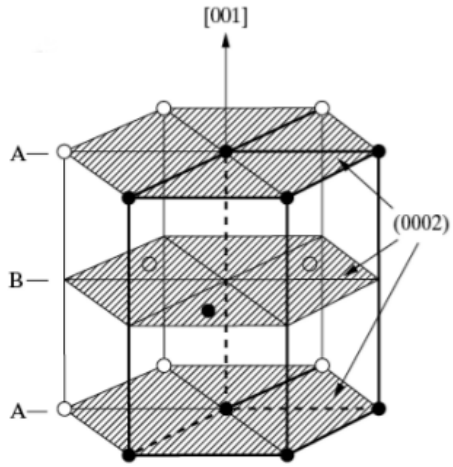


Figure 142.6: Hexagonal close-packed planes

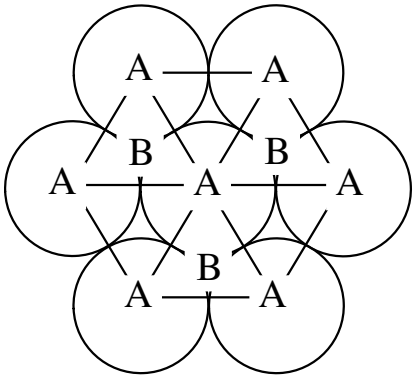


Figure 142.7: Stacking of (0002) planes

The following are other examples of crystal structures with two atoms per lattice point. The first example, in Figure 142.8, is a diamond-like structure that is seen in other atoms besides carbon, such as silicon and germanium. In the examples following Figure 142.8, compounds consisting of two unlike atoms follow unique characteristics of arrangement. For instance, if a compound's structure consists of atoms A and B, the structure of A atoms must possess the same symmetry elements of the crystal as a whole. Namely, a symmetry operation characteristic of the entire crystal, when performed on an a given A atom, must bring it into coincidence with another A atom. Furthermore, body-centering, face-centering, and base-centering translations, if present, should begin and end on like atoms. An example of this is detailed above figure 142.11.

Diamond Structure

ex. C, Si, Ge

4C at 000 + face-center translations (fct)

4C at $\frac{1}{4}\frac{1}{4}\frac{1}{4} + fct$

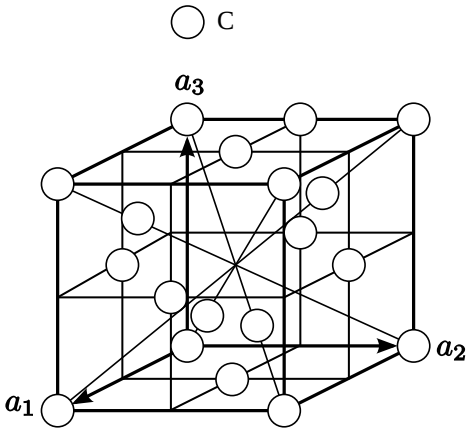


Figure 142.8: Diamond

Zinc Blend

The face-centered cubic zinc blend structure can be likened to the diamond structure in Figure 142.8 above, but the atoms at 000+fct (S) are now different from those at the second set of positions - $\frac{1}{4}\frac{1}{4}\frac{1}{4} + fct$ (Zn).

Ex. GaAs, ZnS

4S at 000+fct

4Zn at $\frac{1}{4}\frac{1}{4}\frac{1}{4} + fct$

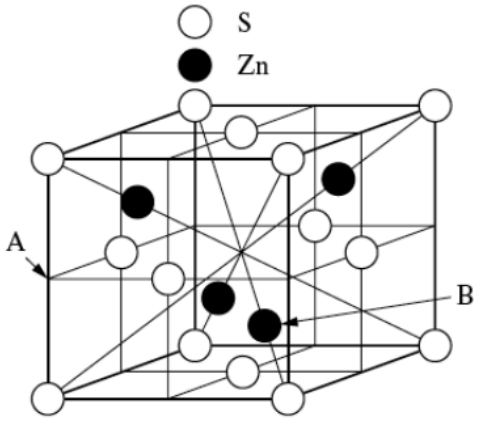


Figure 142.9: Zinc Blend

The CsCl structure illustrated in Figure 142.10 below is common to compounds such as CsBr, NiAl, ordered $\beta - brass$, and ordered CuPd. Though it appears to be a body-centered cubic structure at first glance, the body centering translation of $\frac{1}{2}\frac{1}{2}\frac{1}{2}$ connects two unlike atoms, rather than the required self-coincidence of a single atomic element. Therefore, the structure is simple cubic, and one may think of the Cs atom at 000 and the Cl^-atom at $\frac{1}{2}\frac{1}{2}\frac{1}{2}$ as being related to the single lattice point at 000.

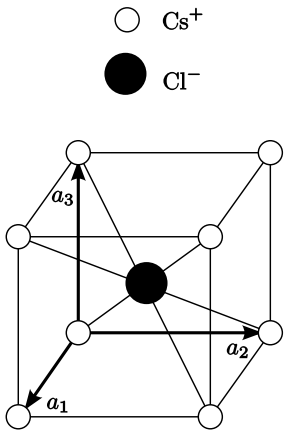


Figure 142.10: CsCl structure

In the NaCl structure of Figure 142.11 below, the sodium atoms are face-centered. When face centering translations are applied to the chlorine atom at $(\frac{1}{2}\frac{1}{2}\frac{1}{2})$, all chlorine atom positions are mapped. In addition, a 90rotation about the four fold rotation axis at [010] brings the chlorine atoms into self coincidence.

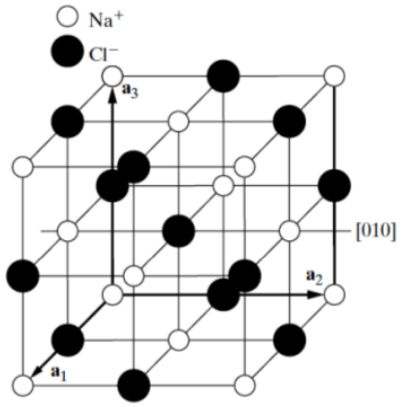


Figure 142.11: NaCl (common to KCl, CaSe, PbTe,etc.)

Voids in FCC

The FCC structure has two types of voids, namely octahedral and tetrahedral voids, which naturally arise from the arrangement of atoms in its crystal lattice.

Octahedral voids are located at $\frac{1}{2}\frac{1}{2}\frac{1}{2}$+fct in the NaCl structure. There is one located at the center of the unit cell, and one at the center of each of the unit cell edges. Since there are twelve edges and each edge is shared with four other unit cells, this yields three edge octahedral voids per unit cell in addition to the one at the center. Thus, there are a total of four octahedral voids per unit cell. The space labeled "A" in Figure 142.12 is an octahedral hole in the structure, which is surrounded by eight sulfur atoms. In Figure 142.11, note that $Cl^-$is NaCl octahedrally surrounded by Na^+.

The Zn atom at $\frac{3}{4}\frac{3}{4}\frac{1}{4}$, marked "B" in Figure142.12, is surrounded by four S atoms at the corners of a tetrahedron. This is known as atetrahedral void in the structure, as the Zn atom is surrounded by a tetrahedral structure of sulfur atoms. In the zinc blend structure as well as other FCC structured crystals, there are a total of eight octahedral voids per unit cell. For the zinc blend structure, however,only half of these tetrahedral sites are occupied by Zn atoms, which is illustrated more clearly in Figure 142.9.

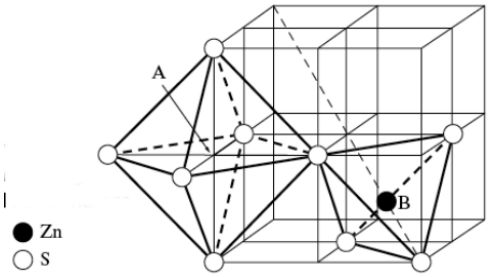


Figure 142.12: Portion of the zinc blend structure

Atom Sizes and Coordination When picturing atoms within a crystal structure, atoms are represented by a simplified hard sphere structure, with atomic spheres in direct contact with one another. This number is specific to a given crystal structure. The radius of a given atomic element is generally considered to be constant regardless of conditions such as phase or presence in a solution. The size of atom may also be deduced from its coordination number (CN), or number of nearest neighbors specific to its crystal structure. The Cs and Cl atoms in the simple cubic structure, for instance, have a coordination number of 8 as shown below.

Ex.: CsCl
Structure - Cubic P (not BCC)
Cs^+at 000 →CN=8
Cl^-at $\frac{1}{2}\frac{1}{2}\frac{1}{2}$ → CN=8

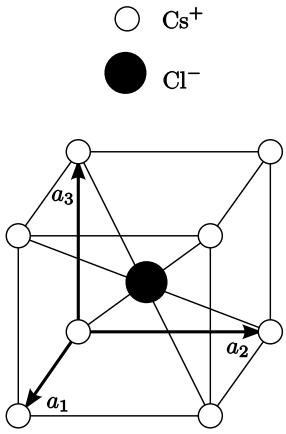


Figure 142.13: CsCl

Using the hard sphere model for atoms, the placement of atoms in the unit cell of a given structure, as well as known atomic radii, the lattice constant of a crystal structure may be determined.
The following procedure demonstrates the evaluation lattice constant "a" in CsCl using known ionic radii:

$$r_{Cs^+} = 1.69\mathring{A} \quad r_{Cl^-} = 1.81\mathring{A}$$
$$\text{Body diagonal of cube} = \sqrt{3}a = 2r_{Cs^+} + 2r_{Cl^-}$$
$$a = \frac{2}{\sqrt{3}}(r_{Cs^+} + r_{Cl^-}) = 4.04\mathring{A}$$

Note that the nearest neighbors experience an attractive force due to their opposite charges and that the second nearest neighbors experience a repulsive force due to their like charges. Therefore, if the two ionic sizes become more disparate than CsCl, then the packing changes to an NaCl structure.

NaCl structure:
4 Na^+at 000 + fct CN=6 (octahedral surroundings)
4 Cl^-at $\frac{1}{2}\frac{1}{2}\frac{1}{2}$+ fct CN=6 (octahedral surroundings)
fct ≡face-center translations
$\{(000), (\frac{1}{2}\frac{1}{2}0), (\frac{1}{2}0\frac{1}{2}), (0\frac{1}{2}\frac{1}{2})\}$
$r_{Na^+} = 0.95\mathring{A} \quad r_{Cl^-} = 1.81\mathring{A}$

If the NaCl structure were like "CsCl" structure,
then $a = \frac{2}{\sqrt{3}}(r_{Na^+} + r_{Cl^-}) = 3.19\mathring{A}$
Therefore, the NaCl CN=8 structure would not bond since the Cl^-- Cl^-distance $a = 3.19\mathring{A} < 2r_{Cl^-} = 3.62\mathring{A}$
In other words, for the CN=8 CsCl structure, the distance of the required lattice parameter would be greater than the available distance of two radii of the larger ions that make up the lattice parameter "a". Thus, there is a critical or minimum radius ratio for which cation–anion contact is established in the CsCl structure.
$a > 2r_>$, where $r_>$ =radius of larger ion & $r_<$ =radius of smaller ion

$$\therefore \textit{Limit } a = 2r_> \rightarrow \frac{2}{\sqrt{3}}(r_< + r_>) = 2r_>$$

$$\text{Radius Ratio} \equiv \frac{r_<}{r_>} = \sqrt{3} - 1 = 0.732$$

This radius ratio rule is different for different crystal structures, as shown in Table 142.1 below. Deviation from the radius ratio rule indicates a shift from ionic to covalent bonding.
Ex. Zinc blend structure: $r_{Zn^{2+}} = 0.6\mathring{A} \quad r_{S^{2-}} = 1.80\mathring{A} \rightarrow \frac{r_<}{r_>} < 0.33$

Crystal Structure	CN	Radius Ratio Rule	Example
CsCl	8	$0.732 < \frac{r_<}{r_>} < 1$	CsCl, CsBr
NaCl	6	$0.414 < \frac{r_<}{r_>} < 0.732$	KCl, NaCl
Wurtzite or Zinc Blend	4	$\frac{r_<}{r_>} < 0.414$	ZnS, GaAs

Table 142.1: Radius Ratio Rules

Figure 142.14: Close-packed structure

What about HCP, FCC close-packed structures? For the HCP structures, the coordination number is 12. An ionic crystal cannot have this coordination number because there would be an equal number of anions and cations surrounding each ion.

143 Introduction to Diffraction

The experimental technique of X-ray diffraction in crystals, since early applications at the start of the 20th century, has been an important method for mapping crystal structures. The following section introduces the basic, key concepts of X-ray diffraction such as the nature of X-rays and Bragg's Law.

X-ray An X-ray is an electromagnetic wave similar to light, but of a much shorter wavelength. As illustrated in Figure 143.1 below, a beam of X-rays of a single wavelength traveling in the same direction (i.e. $\vec{k}$) has an electric field ($\bar{E}$) perpendicular to its propagation direction and a magnetic field ($\bar{H}$) perpendicular to its electric field. A plane-polarized wave is one in which the electric field is confined to the a plane, such as the x-y plane, as the wave travels along the x direction, labeled the Poynting vector in Figure 143.1. The magnetic field is not considered in the following applications.

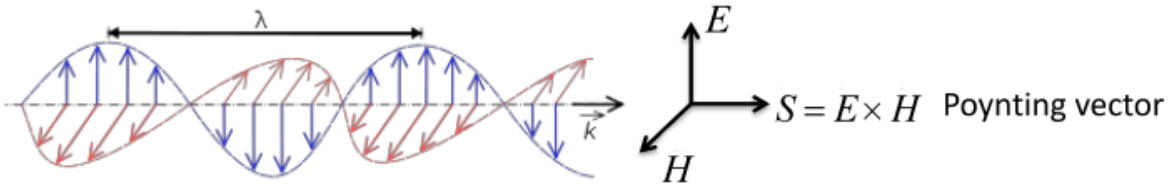


Figure 143.1: X-ray

The following equation models the movement of the electric field as a transverse plane wave traveling in the x-direction, with an amplitude of $\bar{E}_0$ in the and a time varying $\bar{E}$ in the y direction:

$$E = E_0 cos[\tfrac{2\pi}{\gamma}(x - vt)]$$ γ = wavelength,v=velocity traveling in x-direction

Let $\bar{E}_0 = E_0\hat{j}$

After a time, t, has elapsed the wave has been phase shifted to the right by $x' = vt$.

Interference Interference occurs when two or more coherent, or same wavelength, waves are superimposed. If two interfering waves are in phase, then constructive interference occurs, as illustrated in Figure 143.2. Waves out of phase, as shown in Figure 143.3, may produce a resultant wave wit

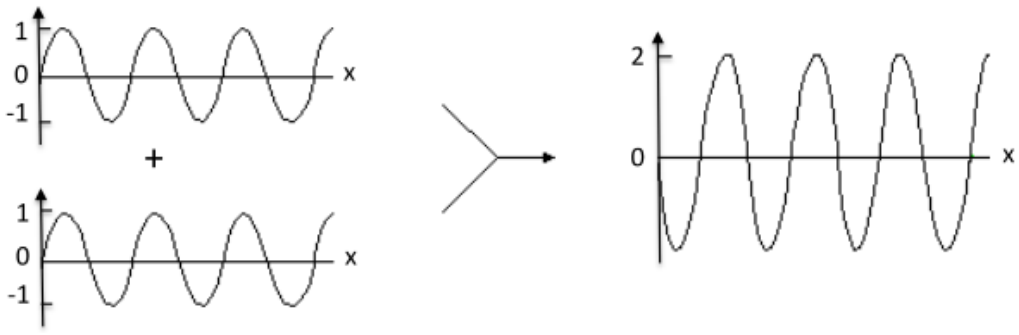


Figure 143.2: Constructive interference

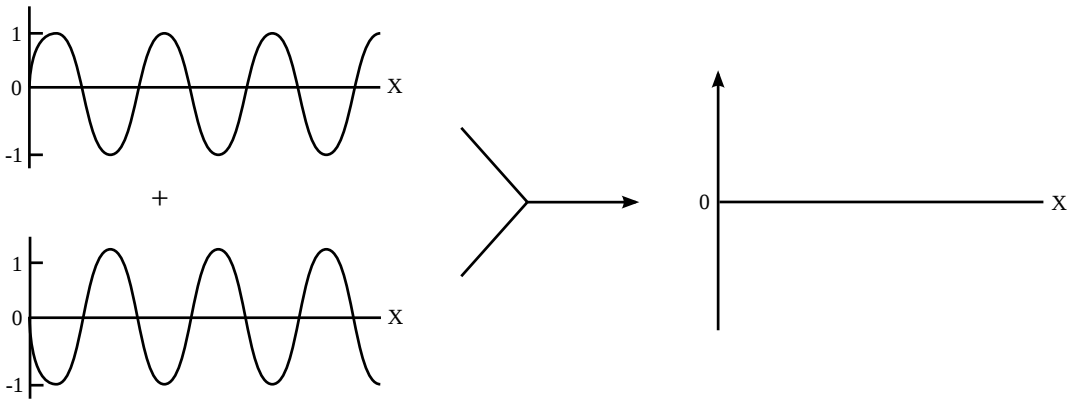


Figure 143.3: Destructive interference

(1800) Young's two-slit experiment
Young's two slit experiment, as illustrated in Figure 143.4 below, is a classic example of the interference of light waves and resulting diffraction patterns.

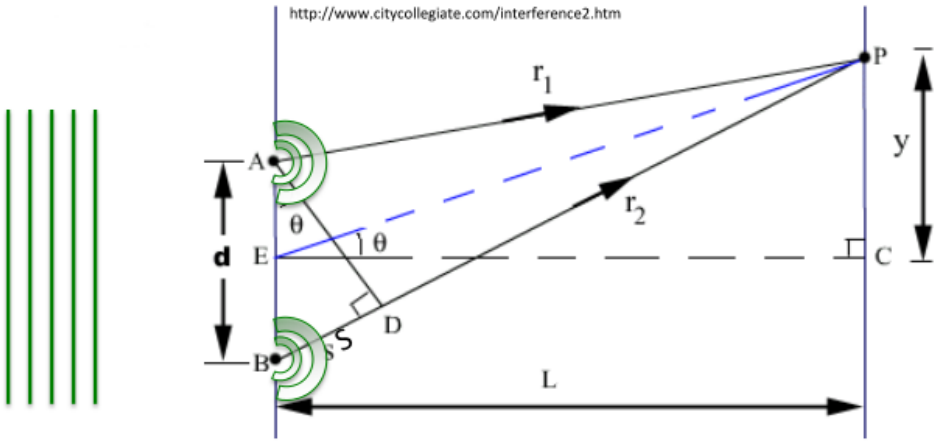


Figure 143.4: Young's two-slit experiment

In Young's experiment, a plane light wave of wavelength γ traveling to the right with an "in-phase" wave front parallel to the grating produces two coherent circular waves emanating from the 2 slits at A and B. The two waves will add up in phase at point P on a screen, if the path length difference, $s = BD$, is a multiple of wavelength γ.

$$s = dsin\theta = n\gamma \rightarrow \text{interference maxima at y}$$

To produce interference fringe pattern, of dark destructive interference and bright constructive interference spots,

$$d \approx \gamma$$

If $d < \gamma$, $sin\theta = \frac{n\gamma}{d} > 1 \rightarrow$ impossible

If $d >> \gamma$, $sin\theta = \frac{\gamma}{d} \rightarrow$ too small to separate n=1 max. from direct beam

X-ray Diffraction History

In 1912, Max von Laue, a young Assistant Professor in Munich, was inspired by PhD thesis work of Ewald who was working in Sommerfeld Lab at the University of Munich. Ewald's thesis concerned solving for a mathematical physics relationship of the scattering of waves from a periodic array of dipole oscillators. Laue realized that Ewald's discovery directly related to scattering of waves from a crystal's periodic structure, and collaborated with Friedrich and Rongten's (discoverer of X-rays) student Knipping to explore this idea. Their experiment demonstrated an X-ray diffraction pattern from a crystal, and proved that X-rays were waves with wavelength of the order of a d-spacing within a crystal.

Later in 1912, W.L. Bragg, excited by the Laue discovery, worked with his father to develop a simpler picture to explain X-ray diffraction.

How does X-ray diffraction work? First, consider reflection from a plane of atoms, as represented by a horizontal line in Figure 143.5. Note that the atoms do not have to evenly spaced within the plane.

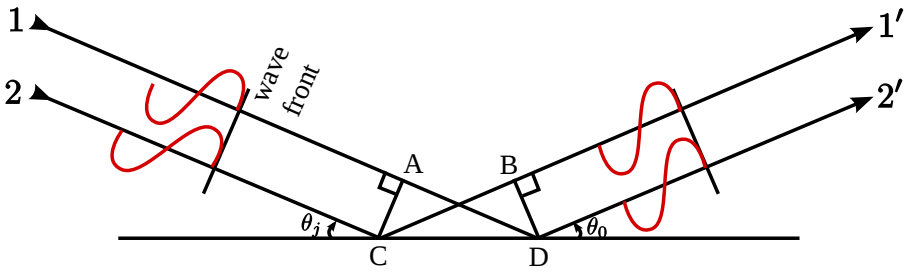


Figure 143.5: Reflection from a plane of atoms

A plane-wave is represented by rays 1 & 2 in-phase at AC in Figure 143.5, and also reflecting in phase . The path length difference of rays 1-1′ and rays 2-2′ are equal when $AD = BC$, and by trigonometry, when the incident angle is equal to the outgoing angle.

Since path lengths will be equal for all rays under the specular reflection condition, the outgoing angle will equal the angle of incidence. This is known as the Law of Reflection.

$$AD = CB \rightarrow \theta_i = \theta_o$$

Law of Reflection

Scattering from two planes

Second, consider the reflection from a set of equally spaced atom planes as illustrated in 143.6.

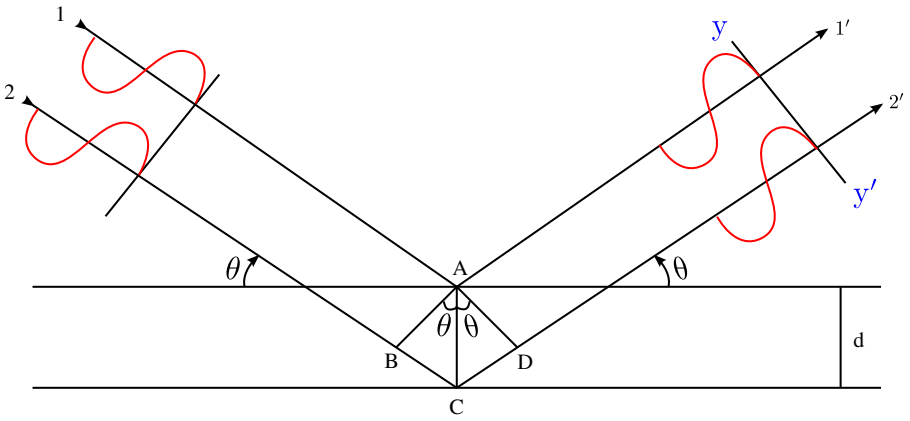


Figure 143.6: Reflection from a set of equally spaced atom planes

In this case, scattering occurs in both planes. Here, we would like to know when the scattering from both planes will result in constructive interference. If the path difference, δ=BCD of rays 1 and 2 is equal to the length equal to an integer multiple of the wavelength of the X-rays ($n\gamma$,where n is an integer) then the scattered rays 1′ & 2′ are in-phase at yy.′ If the angle of incidence were tilted down slightly, then the waves would no longer be in phase, and the conditions would no longer allow for constructive interference. Applying the geometric relationship below yield's Bragg's law. This states the essential condition necessary for diffraction to occur.

$$BC = CD = dsin\theta$$

Bragg's Law: $n\gamma = 2dsin\theta$

n- order of relection, number of wavelengths in the path difference between rays scattered by adjacent planes

If $\theta$is slightly different from the Bragg condition, each successive ray 1′,2′,3′,...will be slightly phase shifted with respect to previous ray as illustrated in Figure 143.7. The amplitudes of the waves would sum to zero since every plane will have another plane at some depth which will scatter out of phase by 180.

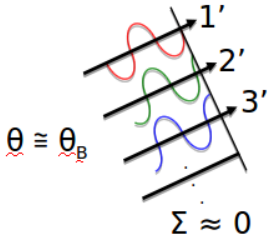


Figure 143.7: Successive phase shifts in reflected rays

What about the "n" in $n\lambda = 2dsin\theta$?

The "n" value , or order of reflection, may take on any integral value as long as sinθ does not exceed one in the Bragg's law equation. For fixed values of d and λ, various n values (n=1,2,3...) correspond to different angles of incidence ($\theta = \theta_1, \theta_2, \theta_3...$) at which diffraction may occur.
For a first-order reflection (n=1) the path length and phase difference of the scattered X-rays 1′ and 2′ in Figure 143.6 would differ by one wavelength. For a second-order reflection (n=2) the path length and phase difference of the scattered rays would differ by two wavelengths. Thus, the rays scattered by all the planes of atoms are completely in phase, leading to constructive interference.
1st order n=1 $\rightarrow \delta = \lambda$ between consecutive planes spaced by "d"
2nd order n = 2 $\rightarrow \delta = 2\lambda$between consecutive planes spaced by "d"
Another way of looking at the reflection order:

$$\lambda = 2\frac{d}{n}sin\theta \quad \frac{d}{n} = \frac{d_{hkl}}{n} = d_{nhnknl}$$

where hkl corresponds to the 1st order or fundamental reflection planes. Thus, the Bragg equation may be rearranged in the following form, in which the "n" term is dropped and the d_{hkl} does not need to be the lowest set of indices for a given plane (i.e. d_{hkl}=110,220,etc.):

$$\lambda = 2d_{hkl}sin\theta \text{or } \lambda = 2dsin\theta_{hkl}$$

If the diffracted intensity is plotted as a function of sinθ, peaks corresponding to crystallographic planes would be visible as shown in Figure 143.8.

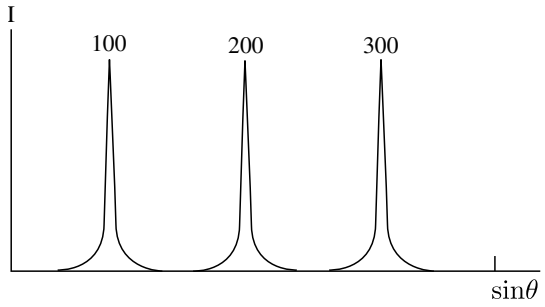


Figure 143.8: Diffracted intensity

Now, consider diffraction from NaCl (111), for which a two-dimensional representation is shown in Figure 143.9 below.

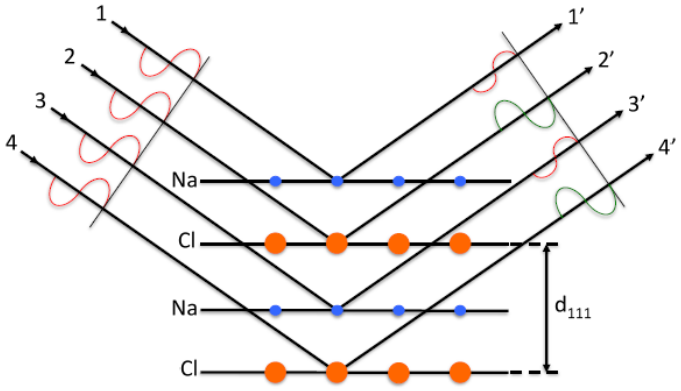


Figure 143.9: NaCl diffraction

The (111) diffraction conditition ($\theta = \theta_{111}$) allows for constructive interference from the 1′ and 3′ waves and the 2′ and 4′ waves , as their path length difference measures exactly one wavelength. Note that rays 1′&3′ from $Na^+$are in phase with each other, but 180out of phase with rays 2′&4′ from Cl^-.
If Na and Cl were equal in scattering strength, then this would lead to complete destructive interference and 0 intensity. Due to the number of electrons in the two atomic species, which are the primary source of scattering in

atoms, the scattering strength is not equal and therefore a small diffraction intensity peak will be observed for $\theta = \theta_{111}$ as shown in Figure 143.10.

For $\theta = \theta_{222}$, the overall path length becomes longer and all rays 1',2',3',4' from all planes $Na^+$and $Cl^-$are in phase. This leads to constructive interference and higher intensity peaks shown in Figure 143.10. Note that this is a simplified model, which solely accounts for the path length difference effect in scattering from different planes which are not identical to each other. Other effects will be discussed later in the text.

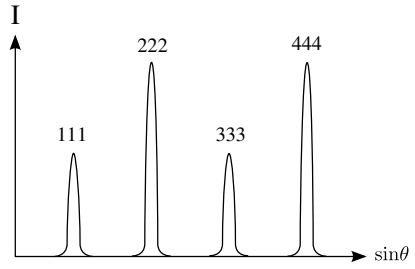


Figure 143.10: NaCl diffraction peaks

Observing the Bragg Angles tells us about unit cell dimensions, namely the d-spacing between planes which can be used to derive the lattice constant. Furthermore, observing relative intensities as shown in the NaCl structure for different θ_{hkl} tells us about the structure within the unit cell. Namely, that there were a pair of ions ($Na^+$and Cl^-) associated with an FCC lattice point.

On the other hand, in a simple FCC metal (111),(333), etc., would not have been weak.

Hint: Look for weak and absent reflections

Absent (or forbidden) Reflections Figure below shows a two dimensional projection side view of (100) planes in a simple FCC metal. The white circles represent atoms that are half of a lattice spacing out of the plane.

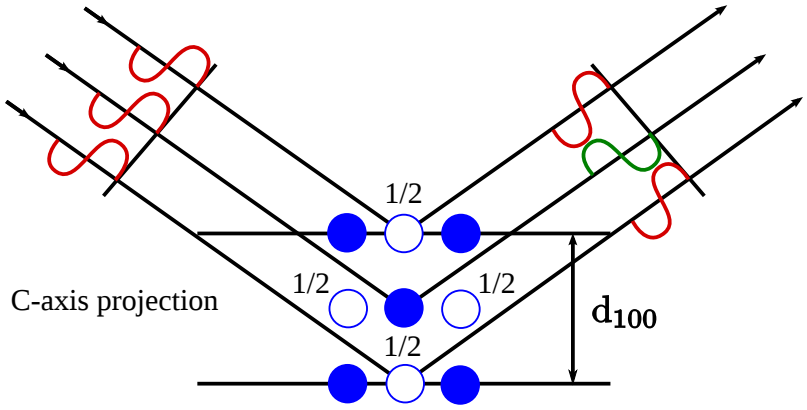


Figure 143.11: FCC (100)

The plane of atoms half way between the planes will scatter at 180out of phase, canceling out the waves reflecting at the 100 planes. Therefore, the $\theta = \theta_{100}$ Bragg diffraction condition leads to destructive interference, and the (100) reflection is absent or not allowed. For simple cubic, the atoms halfway between the planes would not be present and th

However all (200) planes scatter in phase and the (200) reflection would be observed. This reinforces the fact that in an FCC structure, the following conditions are true:

- (200), (400), (600),etc. allowed.
- (100), (300), (500),etc. forbidden
- ∴hkl all even or hkl all odd (unmixed) for FCC

How would Bragg diffraction show that a crystal is cubic? i.e., not orthorhombic
Realize that for the cubic structure the following relationship between the d-spacing of the 110 and the 100 planes is true: $\frac{sin\theta_{100}}{sin\theta_{110}} = \frac{d_{110}}{d_{100}} = \frac{1}{\sqrt{2}}$.
By determining the reflection angle for the (100) and (110) planes and then taking the ratio of the sines, the cubic structure can be confirmed if it is equal to $\frac{1}{\sqrt{2}}$.
This same procedure may be used to determine if a crystal structre is FCC. Since (110) and (100) are forbidden for FCC, then (220) and (200) can be used instead.

144 Absorption/Emission

144.1 Spectrometers

The following three X-ray sources will be discussed in this text:

- X-ray tube e.g., Cu target
- Radioactive Source e.g., Fe^{55}
- Synchrotron e.g., e^-storage ring

An X-ray tube must contain three vital components, namely a source of electrons, a high accelerating voltage to move these electrons, and a metal target on which these highly-accelerated electrons will collide. Figure 144.1(a) below illustrates the cross section of an X-ray filament tube. This consists of an evacuated glass casing that insulates the anode at one end from the cathode at another. The cathode is a tungsten filament and the anode is a water-cooled block of copper which also contains the desired target metal at one end. A lead of the high-voltage transformer, detailed in Figure 144.1(b) is connected to the filament and the other is connected to ground. The target is grounded by te cooling water connection. In operation, the filament is heated with a current and emits electrons that are drawn to the target. Around the filament is a metal focusing cup which is kept at a high negative voltage close to that of the filament and tends to focus the electrons on a focal spot on the target. X-rays are emitted from this focal spot in all directions and escabe through windown in the tube usually made from beryllium.

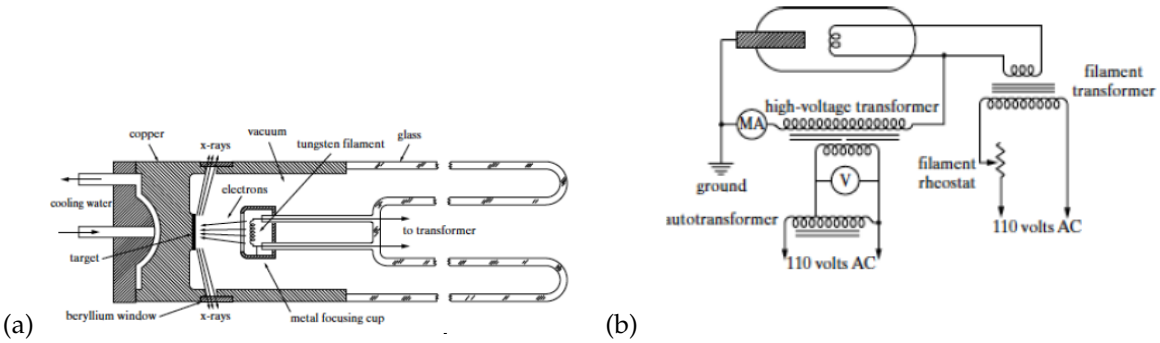


Figure 144.1: X-ray tube

Figure 144.2, below illustrates an X-ray radiation spectrum, plotting X-ray intensity as a function of energy.

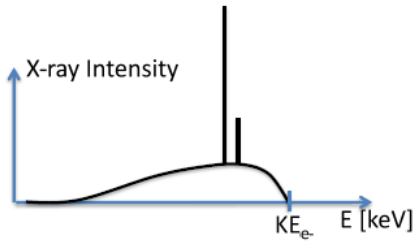


Figure 144.2: X-ray radiation spectrum

144.1.1 Bremsstrahlung Radiation

Bremsstrahlung radiation is a continuous X-ray spectrum from X-ray tube, which is the smooth curve in the intensity plot in Figure 144.2. This occurs as a result of the de-acceleration, or braking of electrons in the target. The accelerating (or de-accelerating) of any charge q, an electron in this case, generates electromagnetic radiation.

Figure 144.3 illustrates a charge q, with acceleration in the $\bar{a}$ direction. If you look in the α direction relative to its acceleration vector, you will have electromagnetic radiation at point P whose distance from charged particle q is indicated by $\bar{r}$. The electric field $\bar{E}$ of this radiation is pointing in the transverse direction to the radius $\bar{r}$. The following formula yields the magnitude of the electric field generated by the accelerated charge:

$$|\bar{E}| = \frac{qasin\alpha}{4\pi\epsilon_0 c^2 r}$$

The following relationships for the electric field also hold true:

- $\bar{E} \perp \bar{r}$ (transverse wave)
- $\bar{E} \perp (\bar{a} \times \bar{r})$

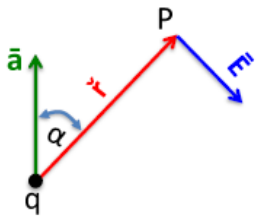


Figure 144.3: Radiation from an accelerated charge

In the case of X-ray tube Bremsstrahlung radiation, the electrons accelerate (negatively) from the tungsten filament cathode to the grounded anode. Electrons accelerating in the negative $\bar{z}$ direction as illustrated in Figure 144.4 create an azimuthally symmetric dipole radiation field. In Figure 144.4, the angle θ is the complement of the angle α in Figure 144.3. Note that the magnitude of $\bar{r}$ does not change, and θ is the factor that affects the intensity of the generated $\bar{E}$ field. Because the intensity is proportional to the magnitude of the electric field squared, it is also proportional to $sin^2\alpha$ or $cos^2\theta$. This is consistent with what is observed in Figure 144.4 and the three-dimensional representation in Figure 144.5, since the intensity at a minimum when when θ is 90 and at a maximum when θ goes to 0.

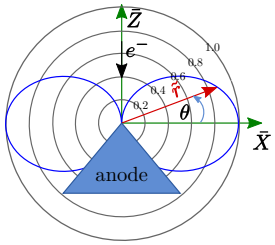


Figure 144.4: Intensity of $\bar{E}$ field with varying θ

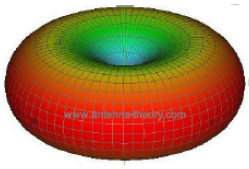


Figure 144.5: Azimuthally symmetric radiation field

$$sin\alpha = cos\theta$$

$$I \propto |\bar{E}| \propto cos^2\theta$$

Max intensity $\perp \hat{z}$

Synchrotron Radiation Synchrotron radiation occurs when electrons are accelerated in a curved path. The electron experiences centripetal acceleration in the radial direction, producing the dipole radiation pattern is of the same toroid shape as illustrated in Figures 144.4 and 144.5. The minimum intensity is observed in the radial direction and maximized perpendicular. However, there is a high relativistic effect due to the electron's speed, such that the dipole radiation field is distorted in the lab frame of reference creating an extremely narrow cone of radiation in the forward direction of the electron's tangential acceleration.

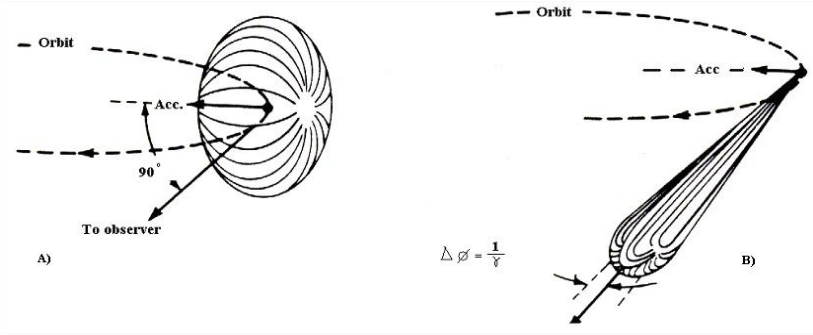


Figure 144.6: (A) Relativistic effect; (B) Extremely narrow forward directed cone of radiation

Recall that Bremsstrahlung radiation is white, or polychromatic, radiation resulting from the deacceleration of electrons following inelastic collisions with target atoms. The kinetic energy analysis of this process is as follows:

The maximum energy loss is the kinetic energy (KE) of the electron at the vacuum target interface. This is the case in which the electron loses all of its energy from one inelastic collision.

$$\Delta E_{Max} = KE_{e^-} = \frac{1}{2}mv^2 = hv = \frac{hc}{\lambda}$$

The maximum energy of the photon, by the de Broglie relationship $\frac{hc}{\lambda}$, is directly related to the wavelength of the X-ray tube. Note that this is a maximum case, and the range of energies will be anywhere from this maximum value to zero.This maximum kinetic energy is therefore directly related to the set potential of the X-ray tube, as shown below.

KE [keV] = V [kV]

Typically, If X-ray tube high voltage, HV = 50kV $\rightarrow$Max photon energy = 50keV

The maximum photon energy in the continuous Bremsstrahlung spectrum will show 50kV at its maximum along with lower values.

de Broglie wavelength relationship: $\lambda = \frac{hc}{E}\lambda[\mathring{A}] = \frac{12.40}{E[keV]}\rightarrow$short wavelength limit $\lambda_{SWL}[\mathring{A}] = \frac{12.40}{V[kV]}$

Therefore, in the Bremsstrahlung spectrum there will be a continuum of wavelengths showing an intensity pattern as well as a short wavelength limit (SWL) corrresponding to the maximum available energy. This is due to the fact that the wavelength and energy are inversely proportional.
Figure 144.7 below shows a radiation spectrum from a copper X-ray tube, for various set potentials (25kV, 20kV, 15kV, etc.). The smooth curve indicates the continuous or Bremsstrahlung radiation curve, with the short wavelength limit (SWL) corresponding to the leftmost point on a given curve. The spikes are indicative of characteristic radiation, but

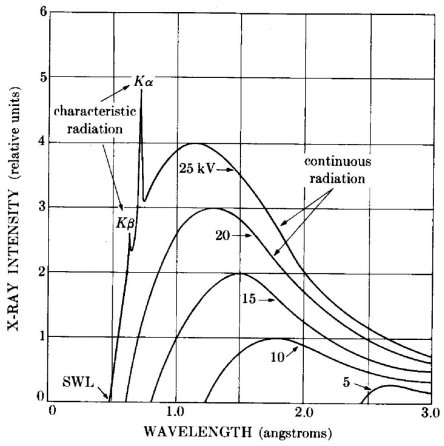


Figure 144.7: Bremsstrahlung radiation spectrum
Note: for V=20kV $\rightarrow\lambda_{SWL} = 0.62\mathring{A}$

The integrated intensity, or area under the curves in Figure 144.7 is proportional to the voltage of the X-ray cubed multiplied by the the atomic number of the target and the current throught the filament. Typically the the voltage will be increased rather than the current, as this increases the intensity by three orders of magnitude.

$$I \propto V^3 Z_{target} i$$

The following is a comparison of the spectra from different synchrotons on a log scale. Notice the Bremsstrahlung spectrum located near the bottom of the plot.

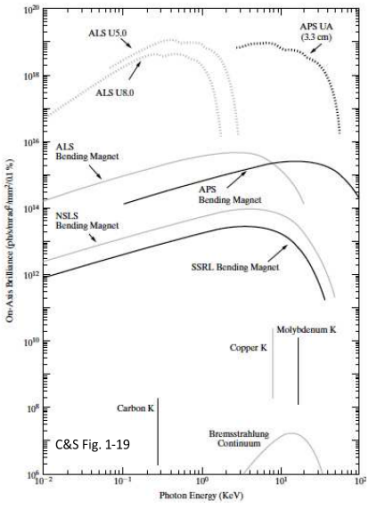


Figure 144.8: On-Axis Brilliance vs Photon Energy (KeV)

144.1.2 Characteristic Spectrum

Raising the voltage on an X-ray tube above a certain critical value for a given target metal yields sharp, narrow intensity peaks, or fluorescent lines superimposed on the continuous radiation spectrum. The characteristic spectrum is comprised of several of these fluorescent lines (i.e. K,L,M,etc.) that are specific to the target in use. In order for fluorescent lines from the target to occur, the electron beam must have enough kinetic energy to overcome the binding energy of an inner core electron and knock it out of the target. The transition of a higher energy electron to fill this hole leads to the production of an X-ray which is characteristic of the given target element. Figure 144.10 illustrates these fluorescent peaks on a continuous background radiation spectrum for a copper target sample at 30 kV.

For fluorescence: $KE_{e^-} = eV \geq W_K (\text{or} \geq W_L)$

where KE_{e^-} is electron energy & W_k *or* W_L is the binding energy of target electrons.

Typically, the K fluorescence lines comprised of $K\alpha_1$, $K\alpha_2$ and $K\beta_1$ are useful in X-ray diffraction. The $K\alpha_1$ and $K\alpha_2$ lines, referred to as the $K\alpha$ doublet, often have very close wavelengths and are treated as a single line. In addition, the subscript in the $K\beta_1$ is usually dropped. The K indicates the principal quantum number of an electron which corresponds to its energy level in the electron cloud. A useful diagram of the electron energy states for copper is shown in Figure 144.10, with different principal energy levels (n=1=K and n=2=L) followed by an integer number that indicates the sub-shell within a given energy level (i.e. L1, L2, etc.). When an electron in the K shell is knocked out , the atom is left in a high-energy state. The target metal atom becomes grounded when a higher energy electron, such as one from the L shell or the M shell, drops down to fill the K electron-hole. Filling a K electron with an L-shell electron is more probable than filling it with M-shell electron, therefore yielding a higher $K\alpha$ (corresponding to the L shell transition) intensity peak relative to the $K\beta$ (corresponding to the M shell transition) intensity peak. Note that the K radiation has a definite wavelength for a given target atom and that it is impossible to excite one K fluorescence line without also exciting the other.

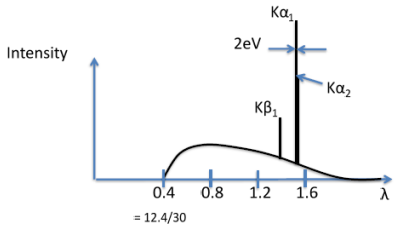


Figure 144.9: Cu characteristic spectrum at 30 kV. Note that the short wavelength limit is at 0.4 due to this 30 kV condition.

In Figure 144.9 above, note that the intensity of the $K\alpha$ peak is one hundred times stronger than the background radiation curve. Furthermore, the $K\alpha$ doublet differ in intensity by a factor of two, and the $K\alpha_1$and $K\beta_1$ differ by approximately a factor of five.

$$\frac{I_{K\alpha}}{I_{Cont}} \approx 100$$

$$\frac{I_{K\alpha_1}}{I_{K\alpha_2}} = 2 \quad \frac{I_{K\alpha_1}}{I_{K\beta_1}} \approxeq 5$$

The arrows in Figure 144.10 below demonstrate the $K\alpha_1 K\alpha_2$and $K\beta_1$ transitions, with their lengths representing the difference in the energies of the two states. This energy difference also equals the energy of the outgoing fluorescent photon. Note that an arrow does not go from the 1s to the 2s electron energy state. Because the electron originally in the 1s state had an angular momentum of 1, the angular momentum of the photon produced must equal 1 by its conservation law. The photon's angular momentum is derived from the difference in the angular momentum of the two electron energy states before and after the transition. The 1s and 2s states do not differ in angular momentum by 1, thus the transition is not favored.

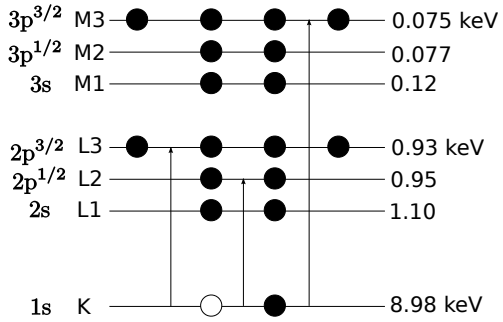


Figure 144.10: Copper energy transitions

Note that in Figure 144.10 there are twice the number of electrons in the L3 state than there are in the L2 state. Due to this, the $K\alpha_1$electron transition from the L3 energy level to K energy level is twice as likely to occur as the $K\alpha_2$ transition from the L2 energy state to the K energy state. This is consistent with the double intensity of the $K\alpha_1$ fluorescence peak relative to the $K\alpha_2$ fluorescence peak shown above.
The following calculations determine the characteristic $K\alpha$ and $K\beta$ fluorescence energies for a copper target and their corresponding wavelengths:

$$E_\gamma \qquad \lambda$$

$$E_{K\alpha_1} = W_K - W_{L_3} = 8.05keV, 1.540\mathring{A}$$

$$E_{K\alpha_2} = W_K - W_{L_2} = 8.03keV, 1.544\mathring{A}$$

$$E_{K\beta_1} = W_K - W_{M_3} = 8.19keV, 1.392\mathring{A}$$

$$(\lambda = \frac{hc}{E} \rightarrow \lambda[\mathring{A}] = \frac{12.4}{E[keV]})$$

144.1.3 Auger Effect

The Auger effect is a non-radiative process that competes with the radiative X-ray emission process to fill an inner electron hole created by the accelerated-electron bombardment of the target. Rather than producing a photon, this process ejects a valence energy level electron known as the Auger electron. An energy level diagram for Auger electron emission from a copper atom is illustrated in Figure 144.11. This process involves three discrete electron energy levels: the first is the K energy level of the electron hole, the second is the L1 or higher energy level from which an electron drops down to fill the K-hole, and the last is the L3 valence energy level from which the Auger electron is emitted. The energy released from the transition of the L1 level electron the lower energy level of the K-hole is transmitted to the L3 valence shell Auger electron, allowing for its ejection from the atom. The kinetic energy of the Auger electron, therefore, is equal to the difference between its binding energy and the energy difference between the L1 and K energy levels involved in the filling of the electron hole. Note that the creation of the K hole leads an L1 electron to drop down and fill the hole as shown by the arrow indicating the movement of the hole from the K to L1 energy level. Following this, the energy is transmitted to the L3 valence electron which becomes the ejected Auger electron, as shown by its arrow pointing outside of the electron's energy levels. The energy of this ejected L3 electron is equal to the binding energy difference between the K and L1 energy levels and the binding energy of the L3 that must be overcome for its emission from the copper atom.

Example:
Copper K L_1L_2Auger electron emission follows creation of k-hole

- K-hole is created in copper atom via accelerated electron bombardment
- L1 electron drops down to fill K hole, creating useful energy
- Useful transition energy is transferred to L3 electron, overcomes the binding energy and ejects the L3 electron from the atom

Energy analysis, where E_e is the energy of the Auger electron:
$$E_e = W_K - W_{L1} - W_{L3}$$

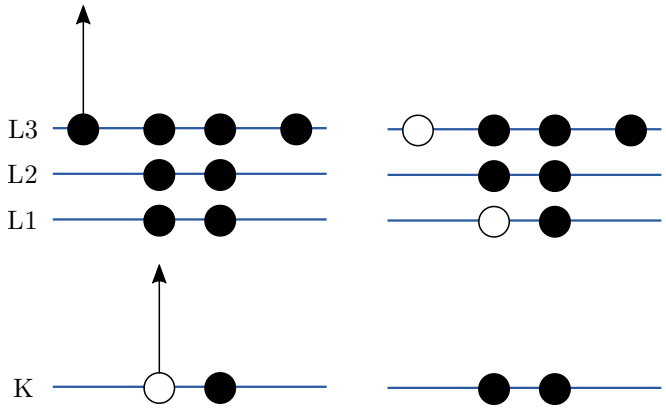


Figure 144.11: Cu K L_1L_3Auger

$$E_e = 8.98 - 1.10 - 0.93 = 6.95keV$$

Recall that the Auger effect and radiative fluorescence are two process that compete to fill a K-electron hole. If the fluorescence yield, or probability of filling a K-electron hole via X-ray emission, is ω_k, then the probability of an filling an electron hole via the Auger effect is $1 - \omega_k$. This Auger emission probability is very high for a low atomic number. For germanium of atomic number 32, there is an equal probability of filling an electron hole by either the nonradiative or radiative processes. Elements with an atomic number lower than germanium (i.e. Z<32) will favor the Auger effect over a radiative process.

Fluorescence yield: ω_k= probability for K hole being filled by radiative process

K Auger Probability = 1-ω_k

The "K" line in Figure 144.12 below shows this relationship, in which a higher atomic number corresponds to a higher fluorescence yield probability for filling a K electron hole. The L-hole electron fluorescence probability is also shown in this diagram as the line labeled "L."

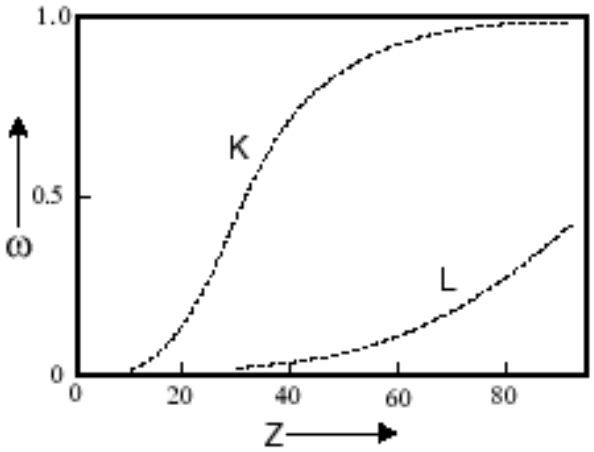


Figure 144.12: Fluorescence yield vs atomic number

Note that because electrons cannot through air very easily like X-rays, the Auger spectrometer is located inside a vacuum chamber. Furthermore, it is much more surface sensitive than X-ray spectrometers.

144.1.4 Radioactive Source

Another form of radioactive emission is the Fe^{55} isotope of iron, which exhibits a high probability of undergoing electron capture. In this process, a proton in the nucleus combines with the 1s electron, creating a neutron and a neutrino. The atomic species is now manganese with an electron hole in its K energy level. Following this, the manganese emists a $K\alpha_1$ radiation wave in filling the K hole with an L3 level electron as shown in Figure 144.13. The observed spectrum will be characteristic of manganese.

$$p + e^- \rightarrow n + u \text{ (neutron + neutrino)}$$

$$Fe^{55} + e^- \rightarrow (Mn^{55}) \text{ with K hole}$$

$$\swarrow\downarrow\searrow$$

$$(Mn^{55}) \text{ with } L_3\text{hole + } K\alpha_1\text{fluorescence}$$

Example:
Fe^{55}Electron Capture

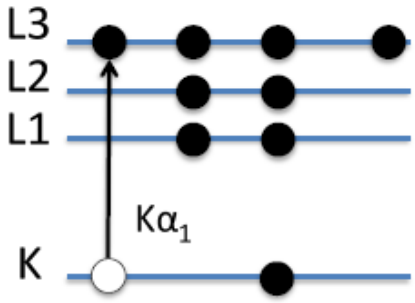


Figure 144.13: $K\alpha$ fluorescence in manganese following Fe^{55} electron capture

$$E_\gamma = W_K - W_{L3} = 5.90keV \leftarrow K\alpha_1$$

The Cd^{109} isotope of cadmium undergoes a similar process, whereby the electron capture leads to its conversion into Ag^{109} and characteristic fluorescence lines of Ag may be observed on the radiation spectrum.

$$Cd^{109} + e^- \rightarrow (Ag^{109})$$

144.2 X-ray Absorption

When an X-ray comes into contact with a material, part of it is transmitted an part of it is absorbed. The attenuation of the incident X-ray beams intensity has been experimentally shown to be directly related to the traversed distance of the X-ray through the material. Referring to the diagram of Figure 144.14, if an X-ray of incident intensity I_o comes into contact with a slab of material of known composition and passes through the slab through a distance x, the transmitted intensity I_x will have a fractional intensity of the incident beam. This difference is equal to the adsorbed intensity, I_x, which is proportional to the distance traversed.

$$I_{absorbed} = I_o - I_x$$

I_o= incident intensity

I_x= transmitted intensity

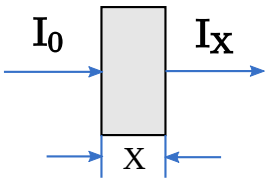


Figure 144.14: Drop in X-ray intensity over distance x through material

The constant of proportionality relating the absorbed intensity to the distance traversed isμ, or the linear absorption coefficient. The plot in Figure 144.1 below shows this exponential relationship, in which the incident intensity I_0 decays at a rate of μ times the distance x through the material.

Differential form: $\frac{-dI}{I} = \mu dx$,μ= linear absorption coefficient

$$\int_{I_0}^{I_x} \frac{dI}{I} = -\int_0^x \mu dx \rightarrow ln(I_x) - ln(I_0) = -\mu x$$

Exponential Relationship:

$$I_x = I_0 e^{-\mu x}$$

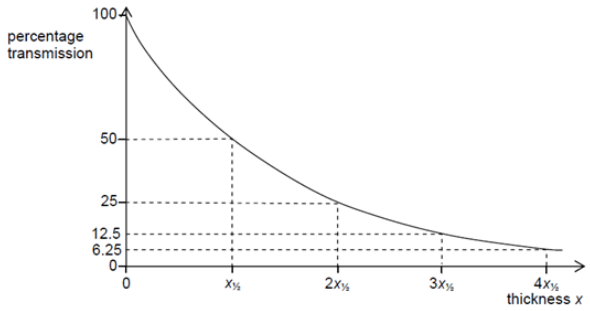


Figure 144.15: Transmitted intensity vs thickness traversed

The linear absorption coefficient is directly proportional to the density of the material, thus the quantity $\frac{\mu}{\rho}$, called the mass absorption coefficient, is constant for a given material regardless of its state of matter. The $\frac{\mu}{\rho}$ value for a given atom is normally listed in tables, rather than listing variousμ values for various compounds or elements and accounting for structure or density factors. Furthermore, the mass absorption coefficient is a function of the atomic number of the element as well as the energy of the X-ray. The energy dependence is due to the fact that there is a critical energy above which the X-rays are able to overcome the binding energy of the electrons in the atoms of the material and emit a photoelectron. Thus, there is a much higher probability of X-ray absoprtion above this critical energy leading to a higher μ value. Furthermore, there is a certain distance at which the initial intensity $I_0$is divided by a factor of e.This distance is equal to $\frac{1}{\mu}$, which is referred to as the absorption length.

$$\mu \propto \rho \rightarrow \frac{\mu}{\rho} \text{ constant of material, independent of state (solid,liquid,gas)}$$

$$\frac{\mu}{\rho} = f(Z, E_\gamma) \equiv \text{mass absorption coefficient}$$

$$\mu^{-1} = \text{ absorption length}$$

For more complicated structures such as chemical compunds or mixtures, the mass absorption coefficient of the substance is the weighted average of the mass absorption coefficient of its constituent elements. A general formula for this is shown below. The mass absorption coefficient for the mixture of air is calculated in Table 144.1, using tabulated mass absorption coefficients of nitrogen, oxygen, and argon along with their respective weight percents in air.

$$\frac{\mu}{\rho} = \sum_{i=1}^{n} w_i (\frac{\mu}{\rho})_i \; w_i = \text{weight fraction}$$

Mass absorption coefficient for variable composition

Example:

Air=76% N_2 +23% O_2+1.3% Ar

i	w_1	$(\frac{\mu}{\rho})\ \frac{cm^2}{g}$	$w_i \times (\frac{\mu}{\rho})_i$
N_2	0.76	7.14	5.53
O_2	0.23	11.0	2.53
Ar	0.013	120.0	1.56
		(C&S Appendix 8)	$9.52 cm^2/g = (\frac{\mu}{\rho})_{Air}$

@ STP $\rho_{Air} = 0.0013 g/cm^3$,$\mu_{Air} = 0.0124 cm^{-1}$,$\frac{1}{\mu} = 81 cm$

Table 144.1: X-ray absorption in Air at E_γ=8.05 keV (Cu $K\alpha$)

From these calculations, it is determined that a copper $K\alpha$ (8.05 keV) X-ray beam is attenuated to $\frac{1}{e} = 0.37 = \frac{I_x}{I_0}$ after traveling through 81 cm of air.
Figures 144.16 and 144.17 below splot the linear absorption coefficients and X-ray transmission for beryllium as a function of the photon energy. Recall that a beryllium window is used in an X-ray filament tube.

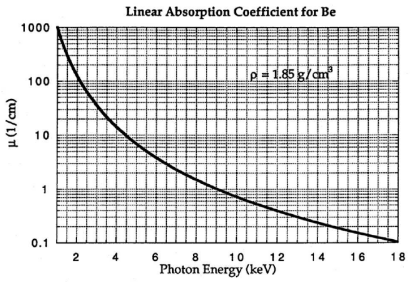


Figure 144.16: Beryllium linear absorption coefficient

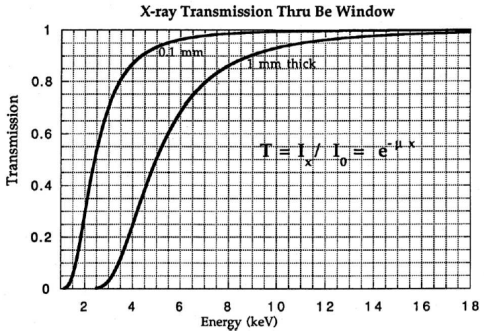


Figure 144.17: X-ray Transmission

X-ray Transmission Filters As previously shown, the spectrum from an X-ray tube yields several fluorescent lines, such as the characteristic $K\alpha$ and the $K\beta$. Some applications require that an X-ray beam be monochromatic, such that the $K\beta$ fluorescence is significantly reduced in comparison to the $K\alpha$ fluorescence. This more monochromatic spectrum may be achieved by using a metal foil filter that preferentially reduces the intensity of the $K\beta$ in the X-ray spectrum compared to the $K\alpha$. Note that absorption filters cannot be used to remove the unwanted $K\alpha_2$ component from the $K\alpha$ radiation. Figure 144.18 below shows a basic setup of a nickel foil filter used in a copper X-ray tube.

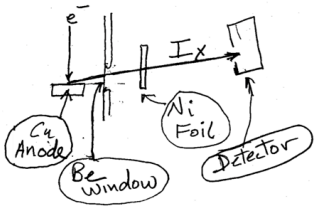


Figure 144.18: Ni foil filter and Cu X-ray tube

Transmission filters make use of the X-ray absorption edge of a particular element. In the second plot of Figure 144.19 below, the absorption edge for nickel can be observed. This discontinuity corresponds to the energy required to eject an inner core electron from an atom of the material. At energies just below this edge the $\frac{\mu}{\rho}$ mass absorption coefficient is significantly less than the mass absorption coefficient at energies slightly above the edge. This translates into higher transmission for x-rays with energies below the absorption edge line on the energy axis. By this process the intial x-ray spectrum, plotted as I_o at the top of the figure, undergoes significant

reduction in the intensity of the $K\beta$ line as shown in the final plot of transmitted intensity, I_x. Note that the high energy background radiation is also significantly reduced in the final plot.

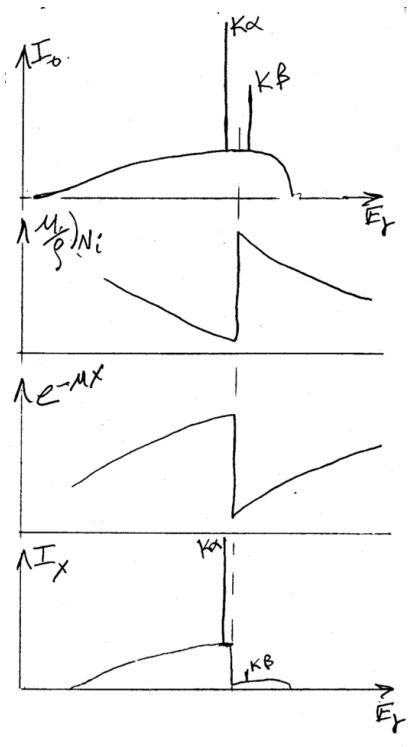


Figure 144.19: Ni transmission filter

Recall the X-ray transmission intensity formula, $I_x = I_0 e^{-\mu x}$.
For an 8 mil Ni Filter, (1mil=0.001"=25.4μm) the intensity of the $K\alpha$ fluorescence is five hundred times greater than the $K\beta$ fluorescence. Recall that without a filter, the $K\alpha$ transmitted intensity was five times greater than the$K\beta$ intensity; thus, the nickel filter only transmits $\frac{1}{100}$ of the original $K\beta$ X-ray intensity.
8 mil Ni Filter:
$x = 20\mu m \ \frac{I(K\alpha)}{I(K\beta)} = 500$
In general, a filter with an atomic number that is one less than the target metal in the X-ray tube is used and the thickness X is chosen such that $\frac{I_{K\alpha}}{I_{K\beta}} = \frac{1}{500}$. This Z-1 relationship is utilized so that the absorption edge lies just above the desired $K\alpha$ and below the $K\beta$ to be filtered. The mass-absorption law $I_x = I_0 e^{-(\mu/\rho)\rho x}$ may used to calculate the thickness of the filter.
Generally, use a $Z-1$ filter:
Cu→Ni Filter
Ag→Pd Filter
Mo→Zr ←(Z-2)
A thicker filter will better suppress the $K\beta$ component, but this will also result in unwanted suppression of the $K\alpha$ component. Thus, it is useful to follow the $\frac{1}{500}$ intensity ratio.

Figure 144.20 below illustrates a general absorption spectra for an atom, by plotting the log of the mass absorption coefficient against increasing energy. Note the presence of the absorption edge previously discussed, which is labeled as the the "K-edge." In addition to the K-edge, other discontinuity spikes corresponding to the L electron shell binding energies are present. The location of the K-edge, and its associated K-branch, further to the right on the energy axis is due to the fact that this inner electron shell has the highest binding energy in a given atomic element.
The overall absorption, labeled $\left(\frac{\mu}{\rho}\right)_{tot}$ is equal to the addition of the absorption from the different electron energy level branches, i.e. $\left(\frac{\mu}{\rho}\right)_K$, $\left(\frac{\mu}{\rho}\right)_{L_1}$,etc. Note that a single increase in the energy unit or in the atomic number have

a cube effect on the mass absorption coefficient. For instance, if the X-ray energy is increased from 10keV to 20keV, the x-ray absorption will decrease by one eighth. Furthermore, if the atomic species is increased by two from, say, nickel to beryllium, then eight times the absorption occurs. Since the mass absorption coefficient $(\frac{\mu}{\rho})$ is inversely proportional to the cube of the energy, it follows that it the wavelength of the X-ray cubed is dirrectly propotional to the mass absoption coefficient. These relationships are summarized in Figure 144.20 below.

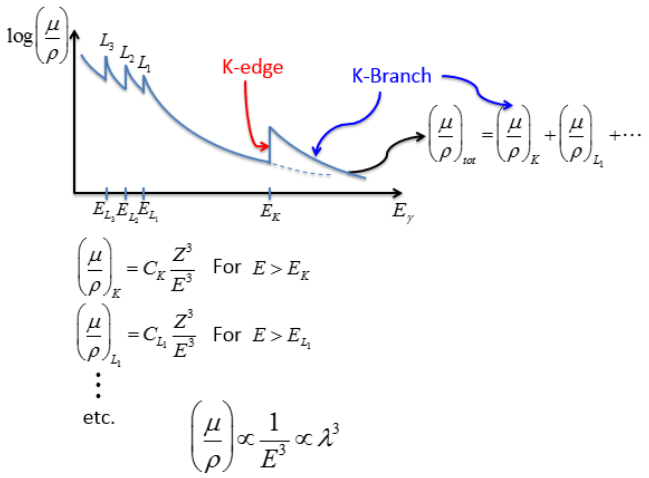


Figure 144.20: X-ray absorption spectra

Photoelectric effect X-ray absorption is dominated by the photoelectric effect, as intially discovered by Einstein. Photoelectric absorption of polarized x-rays occurs when an x-ray photon is absorbed and a 1s or K electron from the atom's inner core is emitted. This process is illustrated in Figure 144.21 below.

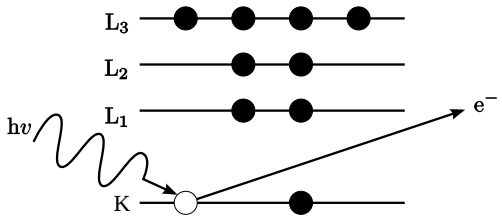


Figure 144.21: Photoelectric effect

In order for photoelectric absorption to occur, the photon's energy must be greater than the binding energy of the K-shell electron. The energy of this emitted photoelectron is given by the difference of the incident photon's energy, $h\nu$ and the binding energy overcome in the electrons emission. This is followed by an X-ray emission, allowing for the atom to return to its ground state.

$$E_{e^-} = h\nu - W_K$$

In addition to conservation of energy, the angular momentum in this process is conserved. That is, the incident photon's intrinsic angular momentum of one and the 1s electrons angulare momentum of one are combined into a dipole angular distribution in the emitted photoelectron. That is, there is a high probability that the emitted electron will be on along the z-axis perpendicular to the incident direction of the photon, and a minimal probability of its emission in the x and y direction.

Dipole selection rule for electronic transition:$\Delta l = \pm 1$

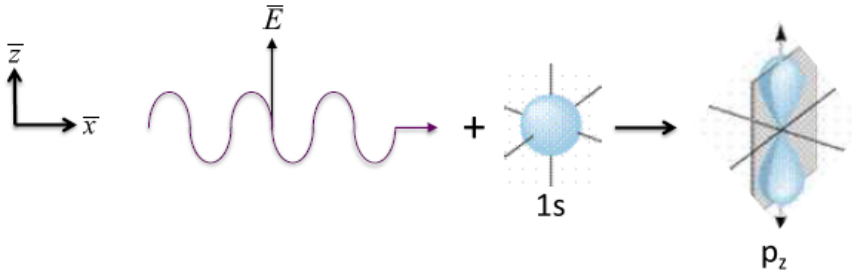


Figure 144.22: Conservation of angular momentum

144.2.1 X ray scattered by an electron

Classical thomson scattering is a wave phenomena in which an non-polarized X-ray beam collides with an electron, causing an acceleration of the electron due to the X-ray's electric field. The effect of the X-ray beam's magnetic field on the electron's acceleration is negligible. This idea is illustrated in Figure 144.23 below.

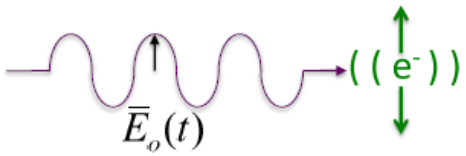


Figure 144.23: Thomson Scattering

The oscillating $\bar{E}$ field of the incoming X-ray is given by the equation: $\bar{E}_0(t) = \bar{E}_0 cos[2\pi(x - vt)/\lambda]$.
The force on this electron is given by its charge multiplied by the magnitude of the incident electric field.

$$\bar{F} = -e\bar{E} = m_e\bar{a}$$

Therefore, the acceleration of the electron is given by the following equation:

$$\bar{a}(t) = \frac{-e\bar{E}_0(t)}{m_e}$$

This oscillating electron emits electromagnetic radiation in all directions, or a scattered wave. The wavelength of this scattered wave λ_s is equal to the wavelength of the incident X-ray wave λ_0.Therefore, this type of scattering is coherent, or elastic.

Oscillating E-field $\bar{E}_0(t) = \bar{E}_0 cos[2\pi(x - vt)/\lambda]$

$$\downarrow$$

Oscillating Charge Particle (e^-)

$$\downarrow$$

E-M Radiation (oscillating fields) $\lambda_s = \lambda_0$

Theory for Radiation Generating by Accelerated Charge The electric field generated by an accelerated charge q at a given position $\bar{R}$ away from the charge is illustrated in Figure 144.24 below. The acceleration vector is indicated as $\bar{a}$ in the vertical direction, with the $\bar{R}$ position being at angle α relative to the charge q's accleration vector.

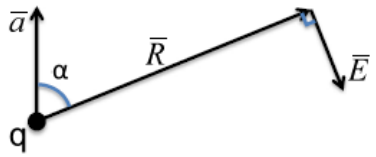


Figure 144.24: E-field at $\bar{R}$

(1) Direction of the electric field: $\bar{E} \perp \bar{R}$ and in-plane of $\bar{R}$ and $\bar{a}$

(2) Magnitude of the electric field: $|\bar{E}| = \frac{1}{4\pi\epsilon_0}\frac{qasin\alpha}{c^2R} = \frac{1}{4\pi\epsilon_0}\frac{e^2}{m_ec^2}\frac{E_0 sin\alpha}{R}$ Note that the electric field gets weaker as you increase the distance $\bar{R}$ from the charged particle.
$|\bar{E}| = r_e\frac{E_0 sin\alpha}{R}$ $r_e = 2.818 \times 10^{-5}\mathring{A}$ Classical e^-radius
$r_e = \frac{1}{4\pi\epsilon_0}\frac{e^2}{m_ec^2}$

The classical thomson scattering equation, as shown below, gives the intensity of the scattered x-rays from the electron as the square of the magnitude of the electric field generated by the oscillating electron. Note that the intensity is a maximum in the direction perpendicular to the acceleration $\alpha = 90$.

$$I_e \propto |\bar{E}|^2 = E_0^2\frac{r_e^2}{R^2}sin^2\alpha$$

$$\frac{I_e}{I_0} = \frac{r_e^2}{R^2}sin^2\alpha$$

$I_e \equiv$Intensity of scattered X-rays from e^-at $\bar{R}$
$I_0 =$Intensity of incident X-rays
$I = \frac{c}{4\pi}|\bar{E}|^2 \equiv [energy \cdot area^{-1} \cdot time^{-1}]$

Why ignore Thomson scattering from the (charged) nucleus?
Relative to the scattering of electrons, the scattering intensity from protons in the nucleus of an atom is negligible, as shown in the following ratio:
$\frac{I_p}{I_e} = (\frac{m_e}{m_p})^2 = 3 \times 10^{-7}$

145 Scattering

145.1 X-ray Scattering & Polarization

Consider an incident X-ray traveling in $\hat{x}$direction and scattered from an electron "e^-" whose position is at some arbitrary origin "O" as shown in Figure 145.1 below. The incident X-ray, $\hat{s_0}$, has an electric field that is polarized in some direction on the y-z plane. We can therefore represent this electric field $\bar{E}_0$ of the incident X-ray in two vector components - one in the y-direction $(\bar{E}_{0y})$ and one in the z-direction $(\bar{E}_{0z})$. Note that the angle 2θ is the complement of the angle α used in the thomson scattering equation and illustrated in Figure 144.24.

$$\bar{E}_0 = E_{oy}\hat{y} + E_{oz}\hat{z}$$

Electric field of incident X-ray traveling in the $\hat{x}$ direction

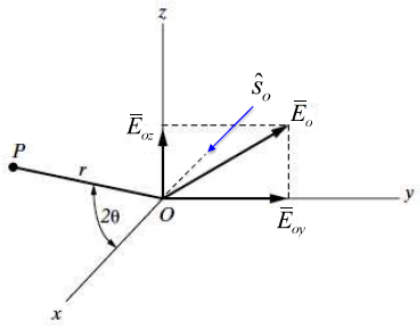


Figure 145.1: Scattered X-ray

For an unpolarized incident x-ray from an x-ray tube, the electric field is taken to be evenly distributed in the $\hat{y}$ direction and in the $\hat{z}$ direction, thus each vector component is equated to half of the total electric field. Recall that the square of the electric field vectors measures the amplitude of the wave, which is proprtional to the intensity of the wave. It follows that the intensity in the $\hat{y}$ direction is equal to the intensity in the $\hat{z}$ direction and half of the total intensity.

$$E_{oy}^2 = E_{oz}^2 = \frac{1}{2}E_o^2 \text{ i.e. } I_{oy} = I_{oz} = \frac{1}{2}I_o$$

Note that the $\hat{z}$ component of the electric field is what accelerates the electron located at the origin in Figure 145.1. The following procedure to determines the scattered intensity I_e from a given electron at the position P in Figure 145.1 above, which is a distance$\bar{r}$ away from the electron in the x-z plane:

If the electric field is in the $\hat{y}$ direction and perpendicular to the scattering plane (the scattering plane being x-z in Figure 145.1):

$\bar{E}_0 = E_o\hat{y}$ (σpolarization case, $\bar{E}_0 \perp$scattering plane)

Here, $\alpha = \frac{\pi}{2}$, thus yielding maximum intensity as expected by the $sin^2\alpha$ term in the classical thomson scattering equation.

$$I_e^{\perp} = I_o\frac{r_e^2}{R^2}$$

If the electric field is in the $\hat{z}$ direction and parallel to the x-z scattering plane:

$\bar{E}_0 = E_o\hat{z}$ (πpolarization case, $\bar{E}_0 \parallel$scattering plane)

Here, $\alpha = \frac{\pi}{2} - 2\theta$, yielding an intensity which varies depending on the 2θ term.

$$I_e^{\parallel} = I_0\frac{r_e^2}{R^2}cos^2 2\theta$$

Figure 145.2 below plots this $cos^2 2\theta$ relationship with a varying 2θ value that is present in the intensity function for the π polarization case.

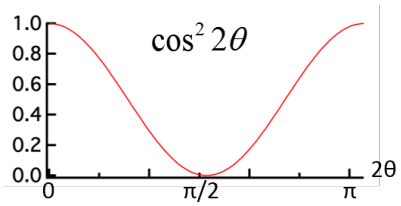


Figure 145.2: Scattered intensity -πpolarization case

For unpolarized e.g. characteristic radiation from an X-ray tube, half of the X-rays are σ polarized and half are π polarized. To obtain the total scattered intensity at P, sum the intensities of the σ polarized and π polarized components.

$$I_e = \frac{1}{2}I_e^{\perp} + \frac{1}{2}I_e^{\parallel} = I_0\frac{r_e^2}{R^2}\left(\frac{1+cos^2 2\theta}{2}\right)$$

↑polarization factor

Note that this relationship is independent of λ and dependent on $\frac{1}{R^2}$.

At $2\theta = 0$ (forward scattering) and R=1cm

$\frac{I_e}{I_0} = \frac{r_e^2}{R^2} = (2.818 \times 10^{-13})^2 = 7.94 \times 10^{-26}$ ← very weak, a small fraction of the incident beam's intensity

At $2\theta = \frac{\pi}{2}$

$\frac{I_e}{I_0} = \frac{1}{2}(\frac{I_e}{I_0})_{2\theta=0}$only σpolarization X-rays are scattered in $2\theta = \frac{\pi}{2}$ direction.

145.1.1 Barkla

British researcher Charles Barkla performed a double scattering experiment to prove that scattering at $\frac{\pi}{2}$ produces a polarized beam. In this experiment, as shown in Figure 145.3 below, unpolarized light interacts with a charged particle at O, and a scattered beam OP is polarized in the $\hat{y}$ direction. The polarized beam is incident on the particle at P, and the scattered beam from P may be detected at a point Q in the x'y' plane. The intensity of this scattered beam at Q is measured as a function of the ϕ direction angle, and varies as $cos^2\phi$ as shown in Figure 145.4. The intensity pattern of this final scattered beam matches the intensity pattern observed when the incident ray scattered is polarized, thus proving that $\frac{\pi}{2}$ scattering produces a polarized beam. Essentially, in this experiment we have an x-ray polarizer at O and a x-ray polarimeter at Q.

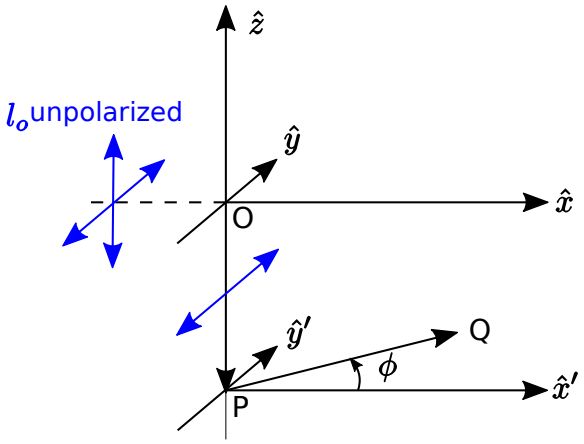


Figure 145.3: Barkla's experiment

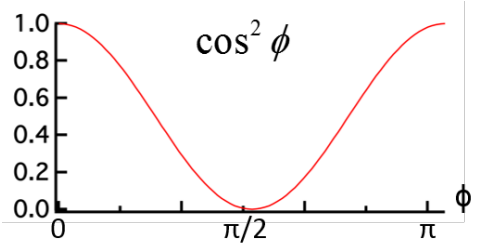


Figure 145.4: Scattered intensity at Q

145.2 Thomson Scattering by two electrons

Recall the intensity of a scattered wave from a free electron:

$I_e = I_0 \frac{r_e^2}{R^2}\left(\frac{1+cos^2 2\theta}{2}\right)$ unpolarized X-ray beam

$r_e = 2.818 \times 10^{-13} cm$

Note: λindependent & $\frac{1}{R^2}$dependent

As shown in Figure 145.5 below, we can define the following parameters:
$\bar{S}_0$ ≡incident direction unit vector
$\bar{S}$ ≡scattered direction unit vector
2θ =scattering angle

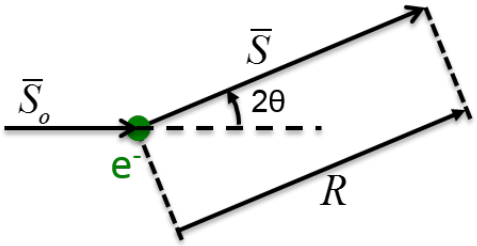


Figure 145.5: Define $\bar{S}_0, \bar{S}, 2\theta$

Now consider the scattering from 2 free electrons, one at origin and one at $\bar{r}$, as shown in Figure 145.6 below.

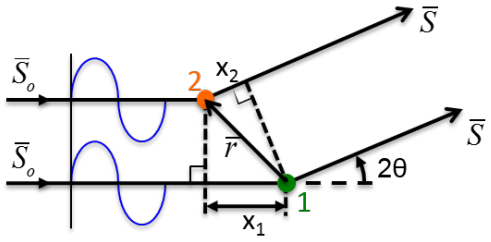


Figure 145.6: Scattering from two free electrons

The two scattered rays will have a phase difference $\delta = (x_1 + x_2)\frac{2\pi}{\lambda}$
Where $x_1 = \bar{r} \cdot \bar{S}_0$, and $x_2 = \bar{r} \cdot \bar{S}$
Equivalently, $\delta = -\frac{2\pi}{\lambda}(\bar{S} - \bar{S}_0)\bar{r}$
Note that in the forward scattering direction, as shown in Figure 145.7 below, the angle 2θ is 0, so the scattered direction is equal to incident direction and there is no phase difference.

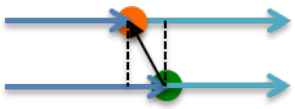


Figure 145.7: Forward scattering direction

The scattering waves from each electron at a detector located a distance R in $\bar{S}$ direction are represented by the following equations:

$$E_1(R,t) = E_1 cos\left[\frac{2}{\pi}R - \omega t\right] = E_1 cos\phi$$

$$E_2(R,t) = E_2 cos\left[\frac{2\pi R}{\lambda} + \delta - \omega t\right] = E_2 cos(\phi + \delta)$$

Note that the second electron's scattered wave has a phase shift δ added into its cosine function. Since the detector is at a distance which is much greater than the distance between the two electrons, the amplitude of the electric fields scattered from both electrons are taken to be equal. Furthermore, the cosine factors of the equations may be condensed using Euler's formula.

$$E_2 = E_1 \leftarrow R \gg r$$

$$\text{use } e^{i\phi} = cos\phi + isin\phi$$

$$e^{i(\phi+\delta)} = e^{i\phi}e^{i\delta}$$

At the detector, the total field is equal to the sum of the scattered field amplitudes from each electron. Note that Euler's formula is used in place of the cosine terms in the first line below. The intensity is determined by dotting the total field, labeled ϵ below, by its complex conjugate ϵ^*. Recall that the scattered intensity from one electron was simply E^2. Now, this E^2 term includes a cosine term that accounts for the fact that the electrons are at two different positions and a certain 2θ angle that can lead to scattering waves at different phases. Notice that if $\delta = 0, 2\pi$ or any even multiple of π then, the scattered waves are in phase and four times the intensity is observed. On the other hand, if δ is an odd multiple of π, then the scattered waves are out of phase, and no intensity is detected. Of course, there are instances in between, such as $\delta = \frac{\pi}{2}$, in which the intensity is in between the two extremes due to interference effects in the scattered waves.

$$\epsilon(R,t) = E_1 e^{i\phi} + E_2 e^{i\phi}e^{i\delta} = E_1 e^{i\phi}(1 + e^{i\delta})$$

$$I \propto \epsilon^2 = |\epsilon|^2 = \epsilon \cdot \epsilon^* = E_1^2(1 + e^{i\delta})(1 + e^{-i\delta}) = E_1^2(2 + e^{i\delta} + e^{-i\delta}) = 2E_1^2(1 + cos\delta)$$

$$\epsilon^2 = 4E_1^2 \text{ for } \delta = 0, 2\pi, ...$$

$$\epsilon^2 = 0 \text{ for } \delta = \pi, 3\pi, ...$$

$$\epsilon^2 = 2E_1^2 \text{for } \delta = \pm\tfrac{\pi}{2}, ...$$

$$\rightarrow x_1 + x_2 = -\bar{r} \cdot (\bar{S} - \bar{S}_o) = n\lambda$$

145.2.1 Review of Thomson scattering from one and two electrons

Thomson scattering from one free electron

$$I_e = I_0 \frac{r_e^2}{R^2}\left(\frac{1 + cos^2 2\theta}{2}\right)$$

Thomson scattering from 2 free electrons separated by $\bar{r}$, as shown in Figure 145.8 below. The phase of scattered wave 2 is relative to scattered wave 1.

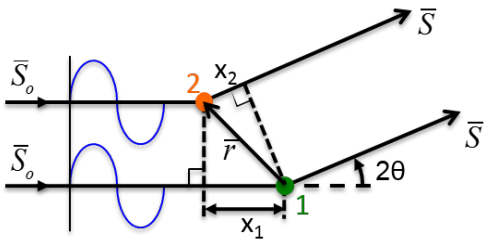


Figure 145.8: Scattering from two free electrons

$$\delta_2 = (x_1 + x_2)\frac{2\pi}{\lambda} = \frac{-2\pi}{\lambda}(\bar{S} - \bar{S}_0)\bar{r}$$

The total field is the sum of the field generated by each of the two electrons.

$$\bar{\epsilon}_{Total} = \bar{\epsilon}_1 + \bar{\epsilon}_2 = e^{i\phi}(\bar{E}_1 e^{i\delta_1} + \bar{E}_2 e^{i\delta_2})$$

where $\phi = \frac{2\pi R}{\lambda} - \omega t$

The phase difference of scattered wave 1 is zero:$\delta_1 = 0$

Where R is the distance of the detector and r is the distance between the electrons, $R \gg r \rightarrow E_2 = E_1 = E_e$

In general, for N free electrons, the total electric field generated is the sum of the scattered fields of all of the electrons. This is based on the superposition principle, in which the sum of the individual forces yields an equivalent total force. The detector "sees" the sum of all of these individual electric force fields, having equal amplitude but individual phase factors resulting from varying positions and 2θ angles. Recall that only in the forward direction will this phase difference equal zero.

$$\bar{\epsilon}_{tot} = \sum_{n=1}^{N} \bar{\epsilon}_n = \bar{E}_e e^{i\phi} \sum_n e^{i\delta_n}$$

Thus, the total intensity is equal to the square of this total electric field, where in the equation below, the first terms in front of the summation equal the scattering from one electron. This is abbreviated as I_e as seen in the second line below.

$$I_{tot} \propto |\epsilon_{tot}|^2 = E_e^2 |\sum_{n=1}^{n} e^{i\delta_n}|^2 = \frac{r_e^2}{R^2}\left(\frac{1 + cos^2 2\theta}{2}\right) |\sum_{n=1}^{N} e^{i\delta_n}|^2$$

$$I_{tot} = I_e |\sum_{n=1}^{N} e^{i\delta_n}|^2$$

If the scattered direction is zero, then the total intensity will be the intensity of one electron times N^2 , where N is the total number of electrons. The intensity of the scattered wave is N^2 times, rather than N times, the intensity of an individual electron, because of the interference of coherently scattered waves (waves of equal wavelength). Furthermore, if the 2θ angle is not zero, then the intensity is less than N^2.

$$2\theta = 0 \rightarrow \frac{I_{tot}}{I_e} = N^2 2\theta \neq 0 \rightarrow \leq N^2$$

Note if the scattered waves had slightly different wavelengths $\lambda_n > \lambda_0$ this is called modified or incoherent scattering. The waves will not interfere and the intensity will be affected by a factor of N as opposed to N^2 in classical Thomson scattering. $\rightarrow (I_{tot})_{mod} = N(I_e)_{mod}$

145.2.2 X-ray Scattering from atom with Z electrons in spherically symmetrical distributions

According to quantum mechanics, the electrons involved in scattering are not at discrete points in space, but have a probability distribution that describes their location around an atom. This probability of finding an electron in a given shell is given by a wave function and may be pictorially represented by an electron-cloud rather than set electron locations, as illustrated Figure 145.9 below. The electron density in a given atom, therefore, is the sum of the individual probability densities of its Z amount of electrons. This the sum is equal to the sum of each electron's wave functions multiplied by the wave function's complex conjugate, as shown below.

$$\rho_{atom} = \sum_{1}^{Z} \rho_e = \sum \psi_e \psi_e^*$$

H-like wave function↑

$$\rho_e = |\psi_e|^2$$

Defining Atomic Scattering Factor In Figure 145.9 below, an incident x-ray beam, represented by the blue arrow pointing to the right and positioned to the left of the red electron cloud, is incident on some differential volume multiplied by the electron density to yield a differential charge value.

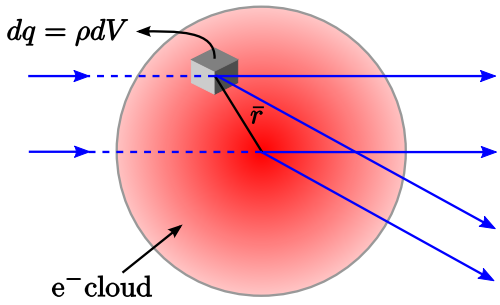


Figure 145.9: Scattering from e^- cloud

The atomic scattering factor, f, is defined as the ratio of the magnitude of the scattered field from an atom as compared to the scattered field from one single electron according to Thomson scattering.

$$f = \frac{E_{atom}}{E_{e^-}}$$

The atomic scattering factor is calculated by the following integral which sums the scattering contributions from each differential volume while accounting for phase differences ($e^{i\delta} = e^{i\frac{2\pi}{\lambda}(\bar{S}-\bar{S}_0)\bar{r}}$)over the entire atomic volume.

$$f = \int_{atom} \rho(\bar{r}) e^{i\frac{2\pi}{\lambda}(\bar{S}-\bar{S}_0)\bar{r}} dV$$

Recall that for N electrons, the magnitude of the electric field scattered in the forward direction should equal N, since there is no phase shift between the scattered waves. This is consistent with the formula below, as the $e^{i\delta}$ term, where δ is the phase shift, goes to one, and $f = Z$. The atomic scattering factor, therefore, is dictated by the θ angle describing the phase shift between scattered waves of multiple electrons.

Essentially, as waves scattered by the atom's electrons become more out of phase (θ increases), the atomic scattering factor decreases (f decreases) due to the partially destructive interference in scattered waves. In addition to the θangle, the atomic scattering factor also depends on the wavelength. At a fixed θ value, for instance, a decrease in wavelength will make path differences larger relative to the wavelength, thus causing greater interference between scattered waves.

Calculating the Atomic Scattering Factor Figure 145.10 below illustrates an incident x-ray beam $\bar{S}_0$, interacting with an atom O and scattering in the $\bar{S}$ direction. The following formula may be used to calculate the amplitude of the scattering for one electron:

$$f_e = \int \rho_e e^{i\frac{2\pi}{\lambda}(\bar{S}-\bar{S}_0)\bar{r}} dV$$

Here, ρ_e is the charge density for the electron, and the exponential term is what relates the direction of the incident beam $\bar{S}_0$ to that of the scattered beam $\bar{S}$.

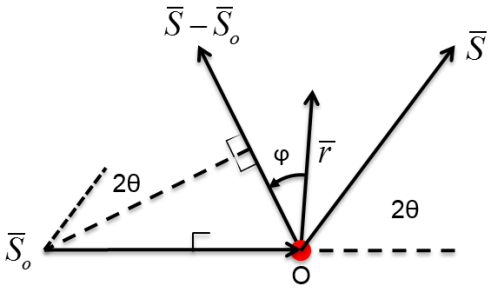


Figure 145.10: Atomic Scattering

$$(\bar{S}-\bar{S}_0)\bar{r} = |\bar{S}-\bar{S}_0||\bar{r}|cos\phi = 2sin\theta r cos\phi$$

Note: $\int_0^\pi e^{iqrcos\phi} sin\phi d\phi$

Let $u = qrcos\phi$ $\frac{du}{dq} = -qrsin\phi$

$\phi = 0 \rightarrow u = qr$

$\phi = \pi \rightarrow u = -qr$

$$\Rightarrow \int_0^\pi e^{iqrcos\phi} sin\phi d\phi = -\frac{1}{qr}\int_{-qr}^{qr} e^{iu} du = -\frac{1}{qr}\left[\int_{-qr}^{qr} cosudu + i\int_{-qr}^{qr} sinudu\right] = -\frac{2sin(qr)}{qr}$$

Let $q = 4\pi sin\theta/\lambda$ Azaroff $q \equiv k$

$$f_e = 2\pi \int_{r=0}^{\infty}\int_{\phi=0}^{\pi} \rho(r) e^{iqrcos\phi} r^2 sin\phi d\phi dr$$

If $\rho(\bar{r}) = \rho(r)$ (spherical symmetric)

$$f_e = 4\pi \int_{r=0}^{\infty} \rho(r) r^2 \frac{sinqr}{qr} dr$$

for spherically symmetrical charge distribution

Example: Lithium

Suppose that the individual electron densities for a neutral lithium atom, of atomic number 3 and electron configuration$(1s^2 2s^1)$, are given by the following hydrogen-like expression:

$$\rho_e = \frac{e^{(-2r/a)}}{\pi a^2}$$

Where for one K electron$a_K = 0.20\mathring{A}$ and for one L electron $a_L = 1.60\mathring{A}$

The scattering factor for each electron may be calculated from the following equation:

$$f_e = 4\pi \int_0^\infty \frac{e^{-(2r/a)}}{\pi a^2} r^2 \frac{sinqr}{qr} dr = \frac{1}{\left[1+\left(\frac{2\pi a sin\theta}{\lambda}\right)^2\right]^2}$$

Furthermore, the total scattering factor for lithium is obtained by summing the scattering factors of the two electrons in its K shell and the one electron in its L shell. $f_{Li} = 2f_{eK} + f_{eL}$

Note that the electron scattering factors for lithium are decreasing a function of $\frac{sin\theta}{\lambda} = \frac{1}{2d}$ in Figure 145.13 below.

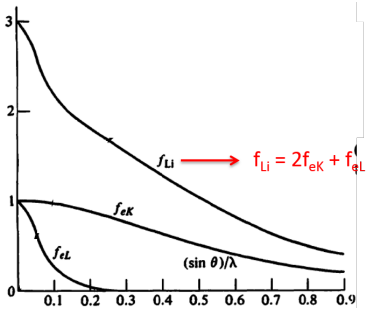


Figure 145.11: Li scattering factor

Figure 145.12

145.3 Compton Scattering

Compton scattering, also known as modified or incoherent scattering is the effect in which scattered waves have slightly different wavelengths than their incident x-ray wavelengths. This occurs when x-rays encounter loosely bound electrons, and may be explained by considering the incident beam to be comprised of a stream of photons of energy $\hbar\omega_o$, as illustrated in Figure 145.13 below. When one of these photons strikes a loosely bound electron that is initally at rest, energy and momentum are transferred from the photon to the electron. Due to this transfer of energy and momentum, the scattered x-ray photon has lower energy and therefore a larger wavelength. Note that another important characteristic of Compton scattering is that the phase of the scattered wave has no direct relationship to the phase of the incident beam, thus making this scattering incoherent and preventing interference effects.

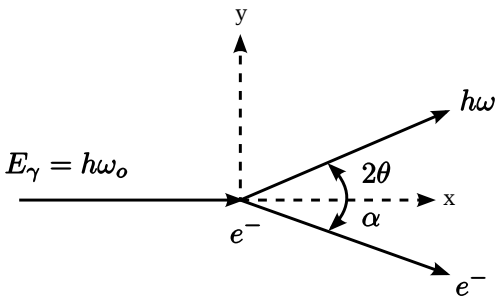


Figure 145.13: Compton Scattering

The following equations combine the principles of conservation of energy and momentum in the photon-electron collision to yield a function of the change in wavelength of the x-ray beam with respect to the scattering angle, 2θ. Note that the theory of Compton scattering expresses the energy and momentum of the photons and electrons in relativistic values. These relativistic values are necessary because the photons scattered are massless and because the energy transferred to the electron is measured relative to its rest energy.

Conservation of energy:
$\hbar\omega_0 = \hbar\omega + \frac{1}{2}m_e v_e^2$ assuming initial e^- at rest

Conservation of momentum
$\frac{\hbar\omega_0}{c} = \frac{\hbar w}{c}cos2\theta + m_e v_e cos\alpha$ ←x direction
$0 = \frac{\hbar\omega}{c}sin2\theta - m_e v_e sin\alpha$ ←y direction

Eliminate α and v_e $\lambda = \frac{2\pi c}{\omega}$ $\hbar = \frac{h}{2\pi}$

$$\lambda = \lambda_0 = \Delta\lambda = \frac{h}{m_e c}(1 - cos2\theta)$$

$$\Delta\lambda(\mathring{A}) = 0.0243(1 - cos2\theta)$$

Note that the change in wavelength in the forward direction is 0 and the change in wavelength in the backward direction $2\theta = 180$ is about $0.05\mathring{A}$.

$$\Delta\lambda = 0 \text{ at } 2\theta = 0$$
$$\Delta\lambda = 0.0486\mathring{A} \text{ at } 2\theta = 180$$

The change in wavelength may be related to the change in energy by the following equations:

$$\lambda - \lambda_0 = \Delta\lambda = \frac{h}{m_e c}(1 - cos2\theta) \qquad \Delta E = \frac{E_\gamma^2(1-cos2\theta)}{m_e c^2 + E_\gamma(1-cos2\theta)}$$

where $m_e c^2 = 511keV$ (rest energy of electron)

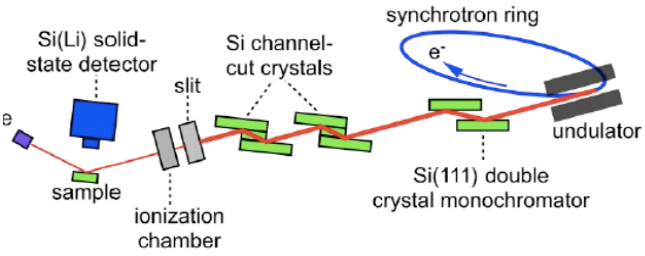


Figure 145.14

Ex: If $2\theta = 90$ $E_\gamma = 18.5KeV$ $\Delta E = 0.65KeV$

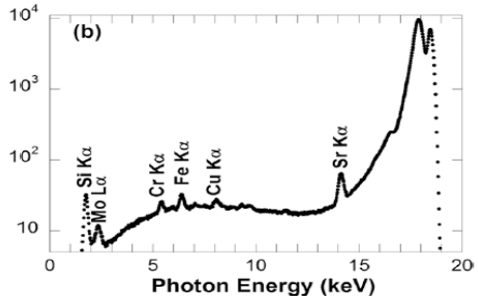


Figure 145.15: Photon Energy

Note that compton scattering will not occur if the change in energy is less than the binding energy for the electron.$\Delta E < E_B$

For hydrogen, a one electron atom, the total intensity according to the Thomson scattering equation is equal to the summed intensities of the modified (Compton) and unmodified (Bragg) scattered wave.

$$I_e = I_{unmod} + I_{mod}$$

$$I_e = I_0(\frac{r_e}{R})^2 sin^2\alpha \rightarrow Thomson$$

$$I_{unmod} = Coherent \rightarrow Bragg \qquad I_{mod} = Incoherent \rightarrow Compton$$

$$1 = \frac{I_{unmod}}{I_e} + \frac{I_{mod}}{I_e} \text{ in electron units}$$
$$1 = f_e^2 + (I_{eu})_{mod} \Rightarrow (I_{eu})_{mod} = 1 - f_e^2$$

The modified, or Compton scattering, and the unmodified, or Thomson scattering, vary in opposite ways with respect to $\frac{sin\theta}{\lambda}$, as shown in Figure 145.16 below.

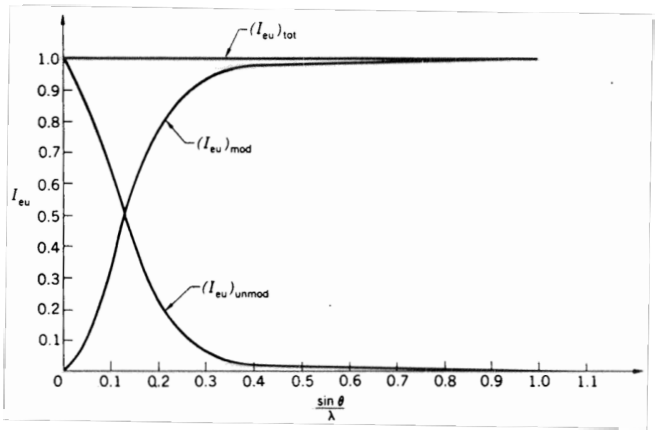


Figure 145.16: Hydrogen scattered intensities

The scattering intensity for atoms with atomic numbers greater than one, such as lithium (Z=3), varies from the simple hydrogen model shown above. Again, the modified scattering intensity for each electron is given by $1 - f_e^2$. Recall that no interference effects can occur for incoherent scattering due to wavelength variations after collisions; thus, the total incoherent scattering intensity is simply given by the sum of the incoherent scattering intensities of the individual electrons.

$$(I_{eu})_{mod} = 2(1 - f_e^2)_K + (1 - f_e^2)_L = Z - \sum_n (f_e)_n^2$$

Furthermore, the unmodified scattering intensity is equal to the square of the atomic scattering factor, since $f = \frac{E_{atom}}{E_e}$, and the intensity is the square of the magnitude of the electric field. Recall that for coherently scattered rays, the most efficient scattering is in the forward direction, where $f = Z$. Since the unmodified intensity varies as f^2, its magnitude will be less than or equal to Z^2. Furthermore, the modified scattering intensity will be less than or equal to the atomic number of a given element. For lighter elements and for a low wavelength, the modified scattering intensity will be significant relative to the total scattering intensity.

$$(I_{eu})_{unmod} = f^2 \leq Z^2$$

$$(I_{eu})_{mod} \leq Z \leftarrow \text{significant for low Z and low } \lambda$$

Figure 145.17 below demonstrates the relationship between scattering intensities and $\frac{sin\theta}{\lambda}$ for atoms with $Z > 1$.

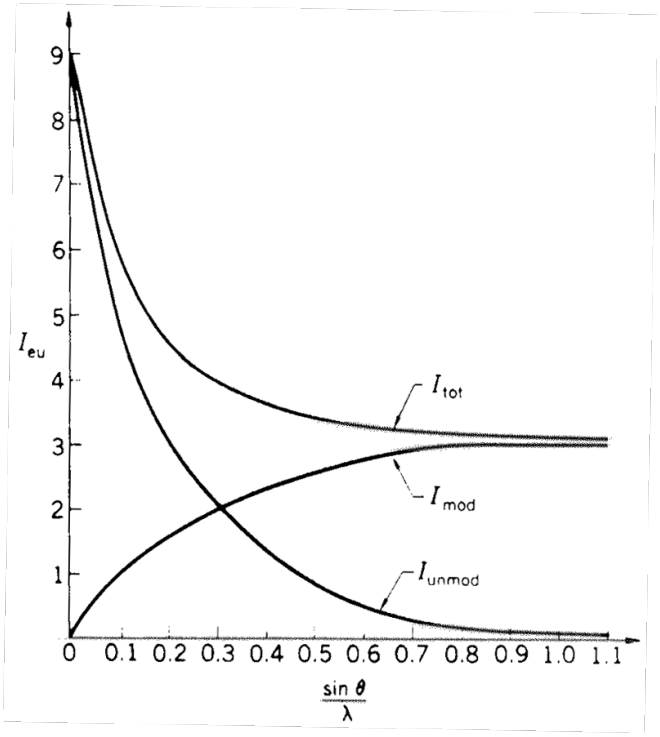


Figure 145.17: Scattering intensities for elements of Z>1

Anomalous Dispersion - λdependence of f Anomalous dispersion is a phenomenon that occurs when the x-ray's photon energy is close to the energy of the absorption edge for a given element. In other words, the frequency of the incoming photon is close to the natural bound frequency for an electron within the atom, causing a resonance effect and a stronger scattering intensity. When this effect is considered, the following wavelength-dependent formula gives the atomic scattering factor:

$$f = f_o\left(\frac{sin\theta}{\lambda}\right) + \Delta f'(\lambda) + i\Delta f''(\lambda)$$

$f_o \rightarrow$Atomic sattering factor

$$f_o(\bar{q}) = \int_{atom} \rho(\bar{r}) e^{i\bar{q}\cdot\bar{r}} d\bar{r}$$

scattering factor is fourier transform of atom's electron distribution, and vice versa

$$q = \frac{2\pi}{\lambda}(\bar{S} - \bar{S}_o)$$

$$q = \frac{4\pi sin\theta}{\lambda}$$

The $f_o$term is what was previously taken to be the atomic scattering factor, and the $\Delta f'$ and $i\Delta f''$ correction terms account for the anomalous dispersion effect. The latter two terms are dependent on the element and the wavelength of the incident x-ray.

If the frequency of the incident x-ray beam is much greater than the natural resonance frequency of the electron (i.e. if the energy of the x-ray is much greater than the binding energy for a K shell electron) the scattered wave will be π out of phase with respect to the incident wave, similar to the π phase lag in Thomson scattering by free electrons. However, if the energy of the incident beam is close to the electron's binding energy, the scattered wave will deviate from this π phase shift with respect to the incident wave. Therefore, the in-phase rays scattered from the K electrons destructively interfere with the out of phase x-rays scattered from outer electrons.

Hönl's Correction Factors As shown in Figure 145.18 below, the Δf real component is 180out of phase with the normallly scattered radiation, f_o, and affects the amplitude of the wave. Furthermore, the $\Delta f''$ imaginary component is 90 out of phase with respect to f_o and affects the phase of the scattered wave.

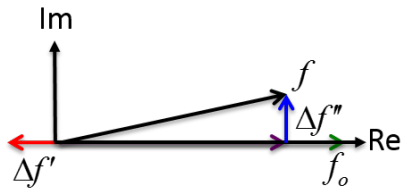


Figure 145.18: Honl's correction factors

Figure 145.19 below demonstrates the effect of increasing energy on the $\Delta f'$ and $\Delta f''$ correction factors. Note that the $\Delta f''$ imaginary component becomes highly positive slightly above the absorption edge. Conversely, the real $\Delta f'$ component becomes largely negative close to the absorption edge. This proximity to the absorption edge indicates that the frequency of the incident ray is close to the electron's natural frequency. Since the absorption edge is characteristic of a given element, the correction factors in $f = f_o + \Delta f' + i\Delta f''$ can give chemical sensitivity to the x-ray scattering process. These $f_o, \Delta f', \Delta f''$ terms are tabulated in the Internation Tables of Crystalllography Vol. IV. In addition, since the intensity is proportional to $ff*$ - the scattering factor times its complex conjugate - the scattered intensity may be manipulated by using a synchrotron radiation source with a tunable E_γ, or wavelength of the incident x-ray beam.

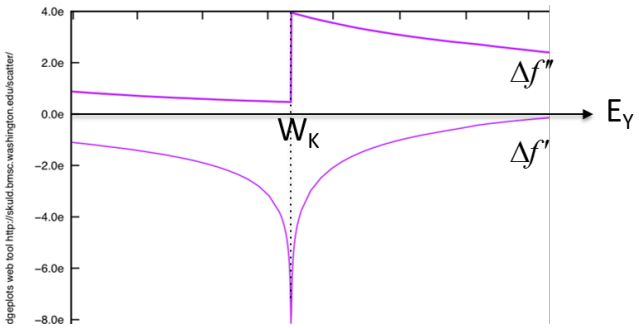


Figure 145.19: Correction factors as a function of incident x-ray energy

Evidently, if the incident x-ray is not close to the absorption edge for an element, i.e. ,E_K or E_L etc., then the corection factors will be negligible with respect to f_o. In 361, we will ignore anomalous dispersion effects unless specifically stated.

145.3.1 X-ray Scattering by Many Atoms

For one atom scattered E-field amplitude is
$E_a = fE_e = fE_o \frac{r_e}{R} sin\alpha sin\alpha = 1$ for σ-polarized case; $sin\alpha = cos2\theta$ for π-polarized case
Elastic Coherent Scattering $\rightarrow f$
$r_e = 2.818 \times 10^{-13} cm$
$f = f_o \left(\frac{sin\theta}{\lambda}\right) \rightarrow \Delta f' = \Delta f'' = 0$
$f_o(0) =$
For unpolarized E_o $E_{ox} = E_{oz} = \frac{E_o}{\sqrt{2}}$

146 Kinematic Theory

Consider the scattering from a collection of N atoms, as shown in Figure 146.1 below. Note the incident direction $\bar{S}$ and scattering direction $\bar{S}$.

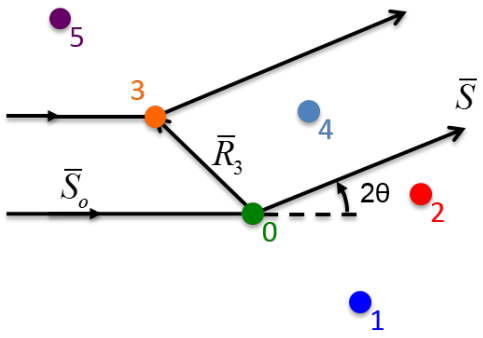


Figure 146.1: Scattering from collection of atoms

If we consider the scattering from atoms "0" and "3," their scattered wave vectors will be parallel and of the same wavelength, but the scattered waves will be out of phase with each other due to a separation between the two atoms, indicated by $\bar{R}_3$. The difference between the incident vector and the scattered wave vector dotted with the distance between the two atoms yields this extra path length, as shown in the following expression:

$$\text{Extra path length } \bar{R}_m \cdot (\bar{S} - \bar{S}_o)$$

The path length difference normalized by the wavelength gives the scattering vector.

$$\text{Scattering vector } \bar{Q} = \frac{\bar{S} - \bar{S}_o}{\lambda}$$

Multiplying the scattering vector by 2π and taking its dot product with $\bar{R}_m$ yields the path difference will be in terms of the phase angle. If this phase angle is an integer multiple of 2π, then perfect constructive interference occurs. The following equation sums the total scattered field by N atoms, in which the exponential terms make up a phase factor that keeps track of the phase difference for all of the atoms in a sample:

$$\frac{E_{Total}}{E_e} = \sum_{m=0}^{N-1} f_m e^{2\pi i \bar{Q} \cdot \bar{R}_m}$$

Note that f_m is the amplitude of scattering by an individual atom, which will be the same for all atoms of the same type and distinct for different atoms. The summed fields are normalized by the scattering amplitude that is expected from one electron according to classical Thomson scattering- $\frac{E_{Total}}{E_e}$.

$f_1 = f_2$: If all atoms are same type, otherwise different

The scattered intensity in electron units, or the scattered intensity of N atoms as compared to the scattered intensity of a single electron is given by the following equation:

$$I_{(eu)} = |\frac{E_{Total}}{E_e}|^2 = \left[\sum_{m=0}^{N-1} f_m e^{2\pi i \bar{Q} \cdot \bar{R}_m}\right] \left[\sum_{m=0}^{N-1} f_m e^{-2\pi i \bar{Q} \cdot \bar{R}_m}\right]$$

Here, the normalized total electric field equation is multiplied by its complex conjugate to yield the intensity as given by $|E|^2$.
In the following step, the $\bar{R}_m$ is subdivided into $\bar{R}_m$ and $\bar{R}_n$ allowing for a grouping into a double summation. Here the $f_m$and f_n factors account for possibilities of interference of scattered waves between two atoms.

$$= \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f_m f_n e^{2\pi \bar{Q} \cdot (\bar{R}_m - \bar{R}_n)}$$

$$I_{eu} = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f_m f_n e^{2\pi i \bar{Q} \cdot \bar{r}_{mn}}$$

In the following vector diagram for the $\bar{R}_m - \bar{R}_n$ term, the origin is conveniently chosen to be at the location of the green "0" atom. The difference of $\bar{R}_3$ and $\bar{R}_4$ gives the ineratomic distance between atoms 3 and 4, labeled $\bar{r}_{34}$.

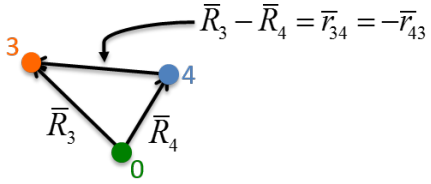


Figure 146.2: $\bar{R}_m - \bar{R}_n$

Scattered intensity for N atoms

Now, the $\bar{r}_{mn}$interatomic spacing vector replaces the $\bar{R}_m - \bar{R}_n$ term previously used.

$$\frac{I_{Total}}{I_e} = I_{eu} = \sum_{n=0}^{N-1} f_m f_n e^{2\pi i \bar{Q}\cdot\bar{r}_{mn}}$$

From this double summation, there are N^2 terms that are of the type $f_m \cdot f_n$
There are N terms of the type: f_m^2 m=0,1,2...N-1
This term is like the scattering intensity from a single isolated atom, in which no interference effects occure. Here $\bar{r}_{mn}$is 0, and the exponential term goes to 1.
There are $\frac{N(N-1)}{2}$ of the type: $f_m f_n (e^{2\pi i \bar{Q}\cdot\bar{r}_{mn}} + e^{-2\pi i \bar{Q}\cdot\bar{r}_{mn}}) = 2 f_m f_n cos(2\pi i \bar{Q}\cdot\bar{r}_{mn})$
These are like cross terms, where $f_m \neq f_n$.

This is the general equation for kinematical scattering from any type of atomic collection - i.e. gas, solid, liquid, crystal, amorphous, etc. The following conditions are specific to to gases, liquids and crystals:

- gases & liquids ←require proper averaging of $\bar{Q}\cdot\bar{r}_{mn}$ (angle averaging) by small angle x-ray scattering
- crystals are periodic ←can use symmetry to reduce N^2terms

Consider an X-ray plane wave $(\lambda, \bar{S}_0)$ incident on a small crystal with a detector for coherently scattered X-rays at R and direction $\bar{S}$, as seen in Figure146.3 below.

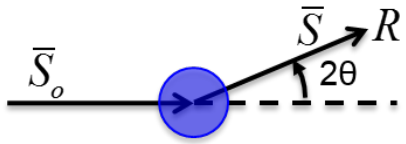


Figure 146.3: X-ray plane wave incident on crystal

The scattering vector determination is illustrated in Figure 146.4 below.

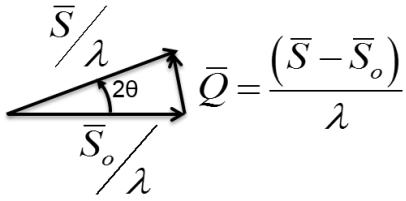


Figure 146.4: Scattering vector

The crystal contains $M = M_1 M_2 M_3$unit cells, in which $M_1 = \bar{a}, M_2 = \bar{b}, M_3 = \bar{c}$. Each unit cell contains N atoms, which may or may not be of different types. Therefore, there are MN total atoms in the unit cell.

Each atom in the unit cell is located by a vector $\bar{r}_n$, as shown in Figure 146.5 below. The atom is illustrated as a black dot in the unit cell block.

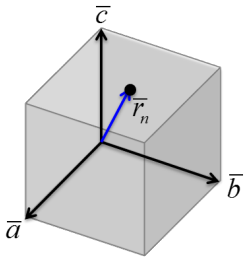


Figure 146.5: Atom located by $\bar{r}_n$

Assuming the crystal is parallelepiped, any unit cell's origin may be located by $m_1\bar{a} + m_2\bar{b} + m_3\bar{c}$, where $0 \leq m_i \leq M_i - 1$ is an integer. This is given by translational symmetry of unit cells.
Therefore, any atom within entire crystal is located by the following position vector:

$$\bar{R}^n_{m_1 m_2 m_3} = \underset{\text{u.c. location}}{m_1\bar{a} + m_2\bar{b} + m_3\bar{c}} + \underset{\text{atom location in u.c.}}{\bar{r}_n}$$

The first three indices locate the specific unit cell containing the atom, using a translational symmetry operator. The $\bar{r}_n$term indicates the atom's position within that specific unit cell.
Figure 146.6 below illustrates the use of these four indices to locate an atom within a crystal. Note that the location of the origin is selected and arbitrary.

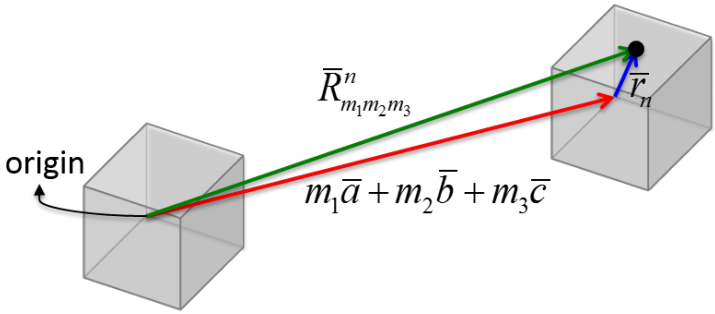


Figure 146.6: Location of atom within entire crystal ($\bar{R}^n_{m_1 m_2 m_3}$)

The scattered electric field from the entire crystal relative to the scattered field form a single electron is given by the following summation:

$$\frac{E_{Total}}{E_e} = \sum_{m_1=0}^{M_1-1}\sum_{m_2=0}^{M_2-1}\sum_{m_3=0}^{M_3-1}\sum_{n=0}^{N-1} f_n e^{2\pi i \bar{Q}\cdot\bar{R}^n_{m_1 m_2 m_3}}$$

$$= \sum_{m_1=0}^{M_1-1}\sum_{m_2=0}^{M_2-1}\sum_{m_3=0}^{M_3-1}\sum_{n=0}^{N-1} f_n e^{2\pi i \bar{Q}\cdot(m_1\bar{a}+m_2\bar{b}+m_3\bar{c})} e^{2\pi \bar{Q}\cdot\bar{r}_n}$$

The four indices yield four summations, where only the last summation term is related to the exponential phase factor.
The structure factor, given by the following summation, is the amplitude of the scattered electric field from a single unit cell:

$$F = \sum_{n=0}^{N-1} f_n e^{2\pi i \bar{Q}\cdot\bar{r}_n}$$

This structure factor accounts for internal interference effects from N atoms in the unit cell. Adding the structure factor to the calculation of the total scattered field yields the following triple summation:

$$\frac{E_{Total}}{E_e} = F \sum_{m_1=0}^{M_1-1} e^{2\pi i \bar{Q}\cdot m_1 \bar{a}} \sum_{m_2=0}^{M_2-1} e^{2\pi i \bar{Q}\cdot m_2 \bar{b}} \sum_{m_3=0}^{M_3-1} e^{2\pi i \bar{Q}\cdot m_3 \bar{c}}$$

Consider the following simplification:

$$\sum_{m_1=0}^{M_1-1} e^{2\pi i \bar{Q}\cdot m_1 \bar{a}} = \sum_{m=0}^{M-1} x^m$$

Geometric series $\rightarrow$ converges

$$= \frac{x^M - 1}{x - 1} = \frac{e^{2\pi i \bar{Q}\cdot \bar{a} M_1} - 1}{e^{2\pi i \bar{Q}\cdot \bar{a}} - 1} = \frac{e^{2i\eta M} - 1}{e^{2i\eta} - 1}$$

$$= \frac{e^{i\eta M}(e^{i\eta M} - e^{-i\eta M})}{e^{i\eta}(e^{i\eta} - e^{-i\eta})} = \frac{e^{i\eta M} sin\eta M}{e^{i\eta} sin\eta}$$

$$= \frac{sin(\pi \bar{Q}\cdot \bar{a} M_1) e^{\pi i \bar{Q}\cdot \bar{a} M_1}}{sin(\pi \bar{Q}\cdot \bar{a}) e^{\pi \bar{Q}\cdot \bar{a}}} = \frac{sin(\pi \bar{Q}\cdot \bar{a} M_1)}{sin(\pi \bar{Q}\cdot \bar{a})} e^{\pi i \bar{Q}\cdot \bar{a}(M_1 - 1)}$$

In one dimension, the total scattering intensity for M unit cells relative to the scattering intensity from one intensity is given by the following:

$$\bar{I}_{eu} = \left| \frac{E_{total}}{E_e} \right|^2 = |F|^2 \frac{sin^2(\pi \bar{Q}\cdot \bar{a} M)}{sin^2(\pi \bar{Q}\cdot \bar{a})}$$

1D Interference Function

The scattering from a single unit cell is given by the structure factor squared, and the second term accounts for interference from M repeated unit cells.
In three dimensions, each M factor is multiplied by its complex conjugate

$$\bar{I}_{eu} = \left| \frac{E_{total}}{E_e} \right|^2 = FF^* \left| \frac{e^{2\pi i \bar{Q}\cdot \bar{a} M_1} - 1}{e^{2\pi i \bar{Q}\cdot \bar{a}} - 1} \cdot \frac{e^{2\pi i \bar{Q}\cdot \bar{b} M_2} - 1}{e^{2\pi i \bar{Q}\cdot \bar{b}} - 1} \cdot \frac{e^{2\pi i \bar{Q}\cdot \bar{c} M_3} - 1}{e^{2\pi i \bar{Q}\cdot \bar{c}} - 1} \right|^2$$

yielding the following three dimensional interference function. This is a rigorous theory for explaining x-ray scattering from a 3-D periodic lattice. Note that the three terms account for repeated unit cells along three axes.

$$\bar{I}_{eu} = |F|^2 \frac{sin^2(M_1 \pi \bar{Q}\cdot \bar{a})}{sin^2(\pi \bar{Q}\cdot \bar{a})} \cdot \frac{sin^2(M_2 \pi \bar{Q}\cdot \bar{b})}{sin^2(\pi \bar{Q}\cdot \bar{b})} \cdot \frac{sin^2(M_3 \pi \bar{Q}\cdot \bar{c})}{sin^2(\pi \bar{Q}\cdot c)}$$

3D Interference Function

Considering the 1D Interference Function

$\frac{sin^2 Mx}{sin^2 x} = M^2$ when $x = \pm h\pi$ $h = 0, 1, 2, ...$

Both the numerator and denominator vary, with the numerator having M times more oscillations than the denominator. In addition, both have periodic zeros at the same integer values of π, this causes the function to go to M^2 at these values. Physically, this condition is when $\bar{Q} \cdot \bar{a} = \pm h$

Recall that $\bar{Q} \cdot \bar{a}$ is extra length in units of $\lambda$for scattering by two atoms separated by $\bar{a}$. This makes sense with previous considerations of Bragg's law, in which an integer multiple wavelength difference in paths yields a diffraction peak.

In Figure 146.7 below, $sinx$, $sin^2 x$, and $\frac{1}{5} sin^2 x$ are plotted against x. The green $sin^2 x$ curve represents the denominator, and the red curve represents the $sin^2 Mx$ numerator where M=5. Note that the red curve has five times the oscillations of the green curve and the two have three common zeros in this range of x.

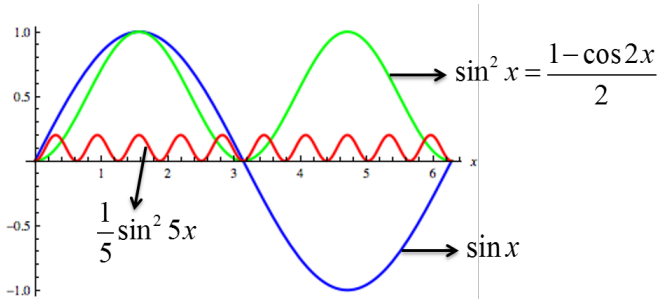


Figure 146.7: 1D interference oscillations

In the one dimensional interference function, the red function in Figure 146.7 is divided by the green function to yield the plot in Figure 146.8.

Note that the one dimensional interference function is multiplied by a $\frac{1}{M^2}$ factor to yield the following normalized function that has a maximum of one:

$$I(x) = \frac{sin^2 Mx}{M^2 sin^2 x}$$

Figure146.8 illustrates the unity amplitude diffraction peaks that occur when the numerator and denominator have common zeros, at $x = \pm h\pi$ with $h = 0, 1, 2$, etc. There is a zeroth order peak that occurs when all waves are scattered in phase and in the forward direction. In addition, there are always $M - 2$ subsidiary peaks between any two diffraction peaks, with the other two peaks hidden at half of the width of the large diffraction peaks at either side. These subsidiary peaks are due to the fact that the numerator, $sin^2 Mx$, gives zero intensity, except when x is an integer multiple of π. The width of these peaks is also due to the $\frac{\pi}{m}$ periodicity of the numerator.

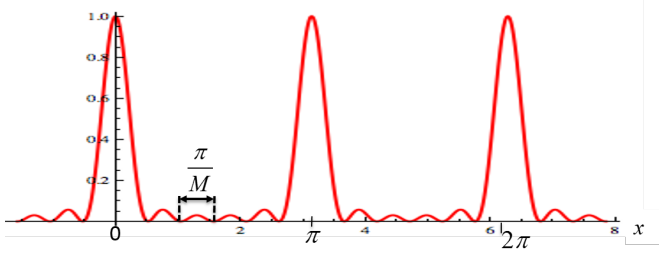


Figure 146.8: Normalized 1D interference function, where M=6

What is x in terms of the angle θ?

The quantity x is the scattering vector Q multiplied by π , and projected in the direction of the lattice constant. If the lattice constant and the scattering vector are in the same direction , this simplifies to the following expression:

$$x = \pi\bar{Q} \cdot \bar{a} = \pi Qa = \pi\frac{|\bar{S}-\bar{S}_0|}{\lambda}a = \pi\frac{2sin\theta}{\lambda}a$$

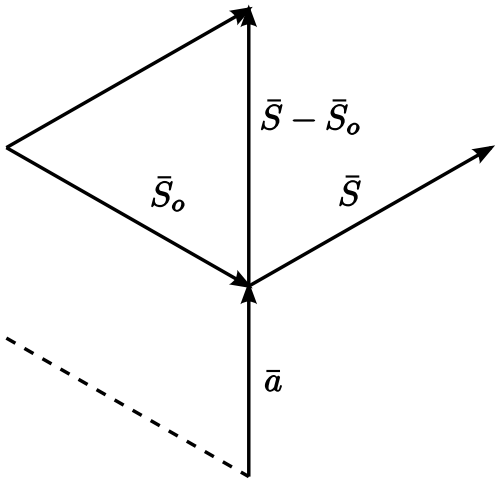


Figure 146.9: Physical relationship between x-ray directions, and $\bar{a}$ - used in describing "x"

$$\text{At } x = \frac{\pi}{M} = \frac{2\pi sin\theta_{min}}{\lambda}a \rightarrow \frac{sin\theta_{min}}{\lambda} = \frac{1}{2Ma} \rightarrow \theta_{min} = sin^{-1}\left[\frac{\lambda}{2Ma}\right]$$

Here, θ_{min} represents the first appearing minimum of θ in the periodic intensity pattern.

For a very large M, $sin\theta$can be set equal to θ, and the following expression can be used to determine the full width, full max diffraction peak width as illustrated in Figure 146.10. Note that this is simply twice the value of θ_{min}.

$$\Delta\theta = 2\theta_{min} = \frac{\lambda}{Ma} \text{ for h=0}$$

Therefore by measuring the width of the entire diffraction peak, the value M may be determined, which indicates the number of unit cells in the direction of $\bar{a}$. Multiplying M by a yields the thickness of the crystal in the direction of a. Notice that $\Delta\theta$ and Ma are inversely proportional, therefore as the width of a crystal increases the $\Delta\theta$, or width of the peak, gets smaller.

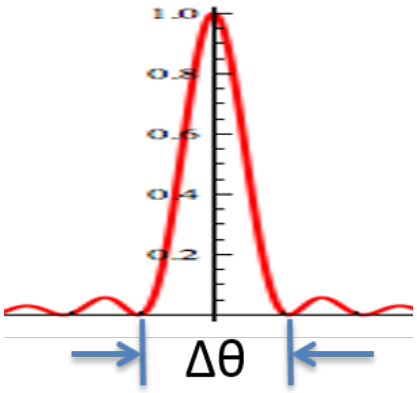


Figure 146.10: $\Delta\theta$- Full width full max

Note the following correlation with Bragg's Law:

$\lambda = 2dsin\theta$, in this case $\lambda = 2asin\theta$

$\frac{2asin\theta}{\lambda} = 1 \rightarrow \bar{Q} \cdot \bar{a} = 1 = h$

Furthermore, for all other peaks ($h \neq 0$)

$$\Delta\theta = \frac{\lambda}{Macos\theta}$$

For all peaks, therefore, the width is inversely proportional to the size of the crystal. Therefore, by looking at the width of the peak, one can see the size of the crystal along the $\bar{a}$ direction.

- for *large* M, Area $=M^2\Delta\theta = \frac{M\lambda}{acos\theta}$

As M approaches ∞, the intensity as a function of x approaches a delta function as shown in Figure 146.11.

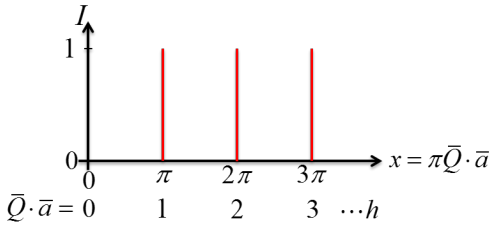


Figure 146.11: Delta functions, where $\bar{Q}$ is the scattering vector and $\bar{a}$ is the scattering section

147 Diffraction

Recall the the following 3D interference function for crystal, where M gives the number of unit cells:
$\bar{I}_{eu} = |F|^2 \frac{sin^2(M_1\pi\bar{Q}\cdot\bar{a})}{sin^2(\pi\bar{Q}\cdot\bar{a})} \cdot \frac{sin^2(M_2\pi\bar{Q}\cdot\bar{b})}{sin^2(\pi\bar{Q}\cdot\bar{b})} \cdot \frac{sin^2(M_3\pi\bar{Q}\cdot\bar{c})}{sin^2(\pi\bar{Q}\cdot c)}$, $M^2 = (M_1M_2M_3)^2$ M = # of unit cells

147.1 Laue condition for diffraction

In three dimensions, the diffraction condition is satisfied and the intensity reaches a maximum when all three fractional factors after the $|F^2|$ term in the interference function simultaneously go to zero, that is when the following three conditions are met:
$\bar{Q}\cdot\bar{a} = h, \bar{Q}\cdot\bar{b} = k, \bar{Q}\cdot\bar{c} = l$ whereh,k,l are integers
This is known as the Laue condition for diffraction. Note that theh,k,l indices are the same indices used in describing the planes of a crystal in real space.
The three scalar quantities describing $\bar{Q}$ in the Laue condition are equivalent to the following vector equation, which also describes the reciprocal lattice vector that points perpendicular to the *hkl* planes in real space with a length of $\frac{1}{d_{hkl}}$:
$\bar{Q} = h\bar{a}^* + k\bar{b}^* + l\bar{c}^* = \bar{r}^*_{hkl}$
Consider the scalar equation $|\bar{Q}| = |\bar{r}^*_{hkl}|$. Recall that $|\bar{Q}|$ is equivalent to $\frac{2sin\theta}{\lambda}$ and that $|\bar{r}^*_{hkl}|$ is equivalent to the inverse of the interplanar spacing, which leads to Bragg's Law:
Bragg's Law: $\frac{2sin\theta}{\lambda} = \frac{1}{d_{hkl}} \rightarrow d_{hkl} = \frac{d}{n}$

Therefore, the h,k,l diffraction peak maxima occurs at:

$$\bar{Q} = \frac{\bar{S}}{\lambda} - \frac{\bar{S}_0}{\lambda} = \bar{r}^*_{hkl}$$

$\bar{Q}$- Scattering vector, $\frac{\bar{S}}{\lambda}$- Scattered wave vector,$\frac{\bar{S}_0}{\lambda}$- Incident wave vector ,$\bar{r}^*_{hkl}$- Reciprocal lattice vector

This relationship is illustrated in Figure 147.1 below. The scattering vector may be seen as the instrument being used in the diffraction experiment, since the incident direction, the wavelength, the location of the crystal, and the orientation of the detector as indicated by 2θ may be selected before performing the experiment. The $\bar{r}^*_{hkl}$ describes the sample, which is a crystal consisting of a discrete set of planes in discrete directions indexed by r_{hkl} in real space and $\bar{r}^*_{hkl}$ in reciprocal space. Therefore, as described by the Laue condition, when the instrument $\bar{Q}$ coincides with an $\bar{r}^*_{hkl}$ vector that is part of a set of discrete vectors pertaining to the crystal, a diffraction peak will appear.

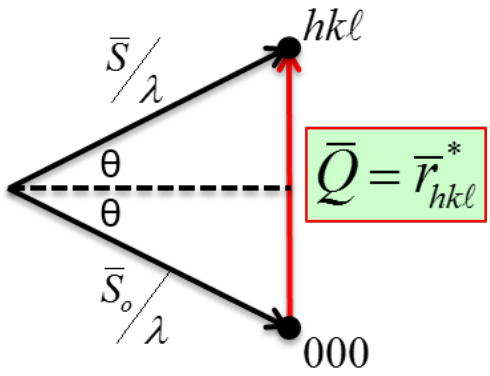


Figure 147.1: Diffraction peak condition

Vector Representation in Reciprocal Space In order to picture a diffraction experiment in reciprocal space, recall that the crystal planes as referenced by *hkl* Miller indices are transformed into *hkl* reciprocal lattice points in reciprocal space.

Figure147.2 illustrates the 012 reflection in the $h = 0$ layer of a cubic unit cell. In this example incident beam I_o is normalized by the wavelength into $\frac{\bar{S}_o}{\lambda}$, and has a fixed direction and wavelength. The beam interacts with the sample at point C, and the scattered wave is in the direction of $\frac{\bar{S}}{\lambda}$. The incident wave vector and the scattered wave vector form a locus of points that is known as the Ewald sphere. A cross section of the Ewald sphere with radius $\frac{1}{\lambda}$ is illustrated in Figure 147.2 .

Ex. 012 reflection, cubic h=0 layer

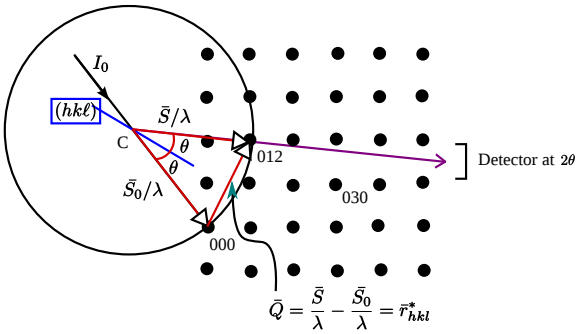


Figure 147.2: Ewald Sphere

Ewald Sphere (Radius=$\frac{1}{\lambda}$) If incident direction $(\bar{S}_0)$ is fixed and λfixed

(1) Sphere does not move.

(2) $\frac{\bar{S}_0}{\lambda}$ always points from center (C) to 000.

(3) Rotate crystal, represented by reciprocal lattice, about 000.

(4) $\frac{\bar{S}}{\lambda}$ points form C to any points on sphere.

(5) Diffraction occurs when an (*hkl*) reciprocal lattice point coincides with the surface of the Ewald sphere.

(6) For X-ray wavelengths multiple reflections are rare. This probability increases with higher energies and a larger Ewald sphere.

147.2 Single Crystal Diffractometer

Consider a silicon wafer

How would you use a $\theta - 2\theta$diffractometer and a monochromatic beam to determine if its face is (111) or (001)?

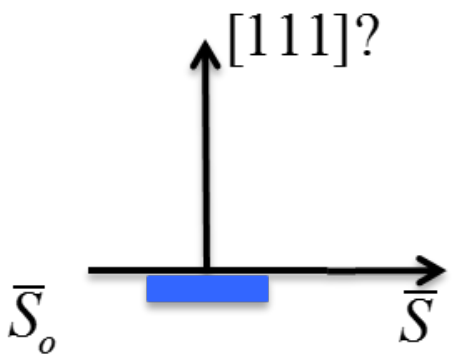


Figure 147.3: Si wafer

(1) Rotate a detector to $2\theta = 0$

(2) Rotate a sample to $\omega = 0$, i.e., surface $\|$to beam

In reciprocal space, move $\bar{r}^*_{hhh}$ into diffractometer plane, i.e., rotate crystal $\rightarrow \bar{r}^*_{hhh} \perp \bar{S}_0 \rightarrow \bar{r}^*_{hhh} \perp 2\theta$axis

(3) Do coupled $\theta - 2\theta$ scan

Notice that the incident wave vector and the scattered wave vector coincide with each other. ω and 2θ will be rotated, so that ω is always half of 2θ. This rotation is continued until the condition in the third illustration of Figure 147.4 is reached. Note that the Ewald sphere is fixed, and the reciprocal lattice points are now angled indicating rotation of the crystal.

(4) Looking at the diffraction peaks, the silicon wafer face is found to be (111), since a (001) face would yield diffraction peaks of the order $(00h)$.

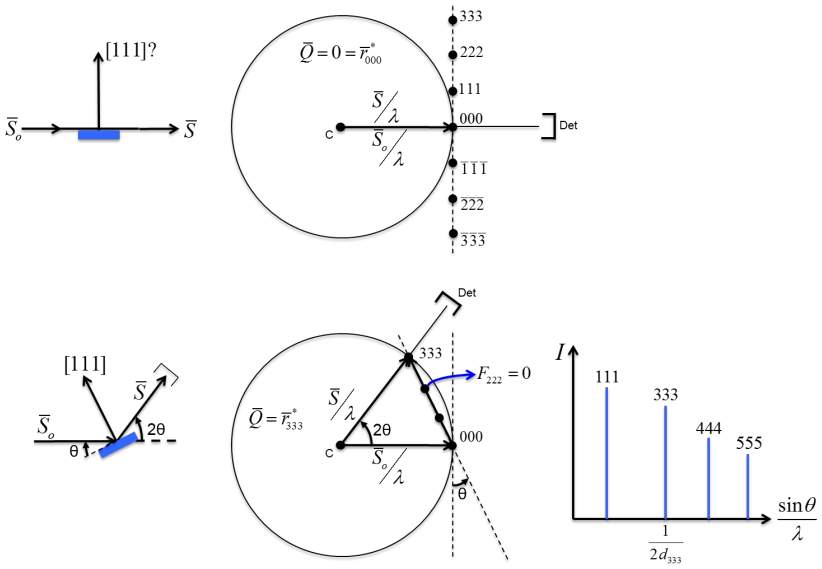


Figure 147.4: $\theta - 2\theta$ scan of Si wafer

147.3 Diffraction Methods

147.1 summarizes the laue method, rotating crystal method, and powder method based on the variability of λ and θ. The rotating crystal method, discussed in the previous section, uses a monochromatic beam of fixed wavelength and has a variable θ for scanning diffraction peaks in series. The laue method has a fixed θ and a variable wavelength since the incident beam is continuous rather than monochromatic and allows for various diffraction peaks to appear at once. The powder method has a monochromatic beam of fixed wavelength and a variable θ since the crystallites in the beam have random orientations.

Diffraction Methods	λ	θ
Laue Method	Variable	Fixed
Rotating Crystal Method	Fixed	Variable
Powder Method	Fixed	Variable

Table 147.1: Diffraction methods

147.3.1 Laue Method - white beam / single crystal

For the Laue method, there is an incident white beam consisting of a continuum of wavelengths which interacts with a crystal as shown in Figure 147.5 below.The crystal consists of a set of planes, indexed in reciprocal space by the $\bar{r}^*_{hkl}$ vector perpendicular to the planes in real space. Satisfying Bragg's law based on these crystal planes creates a Laue spot on the film located downstrean from the sample. This film location is used in the transmission laue experiment setup.

Transmission Laue

$\theta < 2\theta < 90$

$0 < \theta < 45$

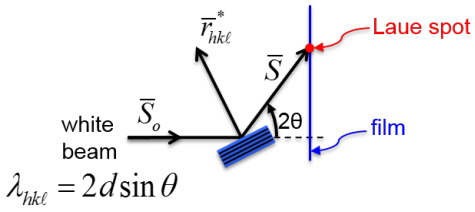


Figure 147.5: Transmission Laue

In the back reflection geometry, the beam comes through a hole in the film and hits the crystal planes in the single crystal sample as illustrated in Figure 147.6. The crystal selects a particular wavelength that matches up to a particular reflection of planes and spacings such that the scattered monochromatic beam that satisfies Bragg's law. There are multiple wavelengths that are harmonics or integer multiples of each other that make up a Laue spot i.e. 220,440,880,etc. However, the diffraction intensity gets weaker for higher harmonics due to many different factors including the atomic scattering factor.

Back Reflection Laue

$90 < 2\theta < 180$

$45 < \theta < 90$

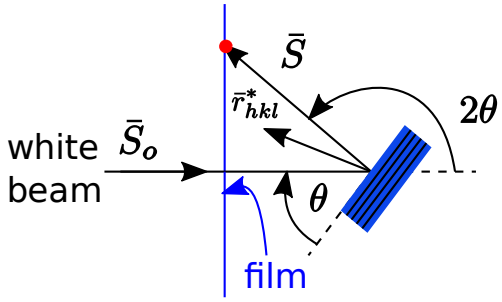


Figure 147.6: Back reflection Laue

Reciprocal lattice treatment of Laue method In reciprocal space, there is a continuum of Ewald spheres due to the continuum of wavelengths in the incident white beam. Recall that the Ewald sphere radius is equal to $\frac{1}{\lambda}$. All Ewald spheres touch the origin, as illustrated in Figure 147.7. Note that this illustration shows cross sections of the Ewald spheres , and a projection of the orthorhombic P

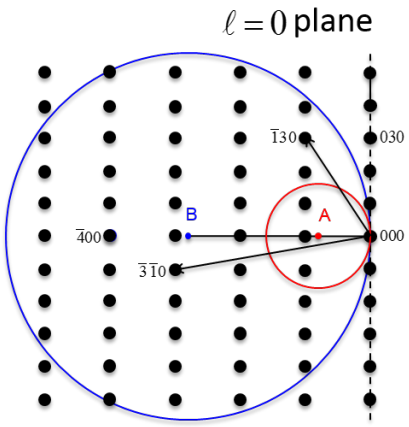


Figure 147.7: Reciprocal lattice of Laue method

The incident wave vector is in the direction of $\bar{S}_o$, and its length is continuum over some range due to the continuum of wavelengths in the incident beam. In the case of Figure 147.8 , the crystal is oriented in a fixed direction such that the [100] direction is coincident with the incident wave direction. The wavelength range of this beam is from the short wavelength limit set by the accelerating voltage to a rough upper limit of $2\mathring{A}$ due to absorption effects.

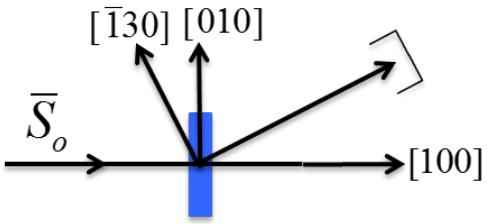


Figure 147.8: Incident wave vector

$\lambda_{SWL} < \lambda < 2\mathring{A}$

Looking back at Figure 147.7, the smaller sphere corresponds to the upper wavelength limit of lower energy and the larger sphere corresponds to the short wavelength limit of higher energy. Note that the radius of the small sphere is AO and the radius of the larger sphere is BO. The two radii correspond to the limits of the incident wave vector $\bar{S}_0$, ranging in length from AO to BO.

- $AO < \frac{\bar{S}_0}{\lambda} < BO$
- $AO = \frac{1}{2\mathring{A}} = 0.5\mathring{A}^{-1}$
- $BO = \frac{1}{\lambda_{SWL}} \approx 5\mathring{A}^{-1}$

Consider the $(\bar{1}30)$ Bragg reflection whose vector diagram is illustrated in Figure 147.9. Note that the length of the normalized incident wave vector is equal to the length of the normalized scattering vector due to their equal wavelengths. This $(\bar{1}30)$ reflection is allowed since its vector is between the minimum and maximum Ewald spheres AO and BO in Figure 147.7. Note that the two vectors form an isosceles triangle and that the scattering vector $\bar{Q}$ is coincident with the reciprocal lattice vector $\bar{r}^*_{\bar{1}30}$. Since the 2θ angle is less than 90, this is the transmission laue condition.

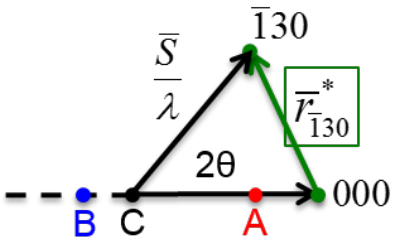


Figure 147.9: $(\bar{1}30)$ reflection

$2\theta < 90 \rightarrow$Transmission Laue

Now consider the $(\bar{3}\bar{1}0)$ reflection illustrated in Figure 147.10 , in which $2\theta$is greater than 90. This is a back reflection laue condition. Note that the incident and scattered wave vector form two legs of an isosceles triangle with their vertex at some point betwen the smallest and largest Ewald spheres for the experiment.

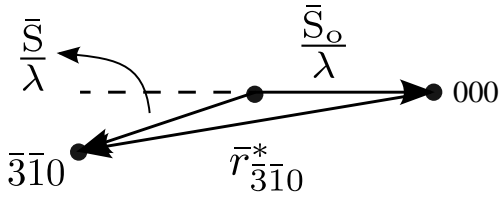


Figure 147.10: $\bar{3}\bar{1}0$ reflection

$2\theta > 90 \rightarrow$Back Reflection Laue

In general, the reciprocal lattice points outside the smallest Ewald sphere of radius $\frac{1}{2\bar{A}}$ and inside the largest sphere of radius $\frac{1}{\lambda_{SWL}}$ may produce Bragg reflections.

Recall that at the Bragg peak, the intensity relative to the intensity of one electron scatterer is derived from the three dimensional interference function. In one dimension, the following equation yields this relative intensity:

$$I_{(eu)} = FF * M^2$$

Here, M represents the number of unit cells that are diffracted, where in three dimensions M_1, M_2 and M_3 define the number of unit cells in the three distinct axes. The structure factor times its complex conjugate, F^2 , affects the relative intensity of the Bragg peak and is given by summation of the individual atomic scattering factors of the atoms in the unit cell and their respective phase factors.

$$F = \sum_{n=0}^{N-1} f_n e^{2\pi i \bar{Q}\cdot\bar{r}_n}$$

Structure Factor for unit cell with N atoms per unit cell

Recall that the individual atomic scattering factor has a form factor and correction factors where $f_n = (f_0(\frac{sin\theta}{\lambda}) + \Delta f' + i\Delta f'')_n$ in general. We simplify this by assuming that $f_n = (f_0)_n$ for atoms at positions $\bar{r}_n$ where $\bar{r}_n = x_n\bar{a} + y_n\bar{b} + z_n\bar{c}$.

At the Bragg peak the scattering vector coincides with a distinct $\bar{r}^*_{hkl}$ vector within the crystal. $\bar{Q} = \bar{r}^*_{hkl} = h\bar{a}^* + k\bar{b}^* + l\bar{c}^*$. This $\bar{r}^*_{hkl}$ may replace the scattering vector in the structure factor equation to yield the following simplified relationship for a structure factor at a Bragg peak:

$$F_{hkl} = \sum_{n=0}^{N-1} f_n e^{2\pi i(hx_n + ky_n + lz_n)}$$

Structure Factor at hkl Bragg Peak

147.4 Structure Factor Examples

147.4.1 Primitive Unit Cell (P)

The simplest structure factor example is a primitive unit cell with one atom per lattice point. This atom is placed at the origin for convenience. In this case, F^2 is the same for all reflections and does not depend on the h, k, l.
1 atom/lattice point $\rightarrow$N=1 $x_n = y_n = z_n = 0$

$$F_{hkl} = f\left(\frac{sin\theta}{\lambda}\right) = f\left(\frac{1}{2d_{hkl}}\right)$$

$$FF^* = f^2$$

for all hkl

147.4.2 Body-Centered (I) Bravais Lattice

For the body centered cubic, tetragonal, and orthorhombic Bravais lattices, there are a total of two atoms of the same kind with one atom per lattice point. A two dimensional projection is illustrated in Figure 147.11.
1 atom/lattice point $\rightarrow$N=1
The atoms are located at positions r_1 and r_2 :
$r_1 = 000, r_2 = \frac{1}{2}\frac{1}{2}\frac{1}{2}$

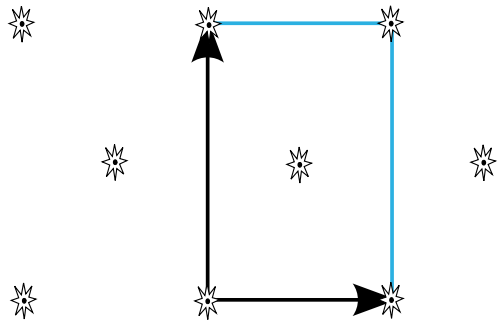


Figure 147.11: Body-centered (I) Bravais lattice

The following is the structure factor calculation for the body-centered bravais lattice:

$$F_{hkl} = f\left(1 + e^{2\pi i\left(\frac{h}{2}+\frac{k}{2}+\frac{l}{2}\right)}\right) = f(1 + e^{\pi i(h+k+l)})$$

If the exponential is raised to an even power, then the structure factor is equal to $2f$ and the two atoms are scattering perfectly in phase.

$$F_{hkl} = f(1+1) = 2f$$

for all $h + k + l = 2n$ (even sum)

If the exponential is raised to an odd power, then the structure factor is zero, and the two atoms are scattering perfectly out of phase.

$$F_{hkl} = f(1-1) = 0$$

for all $h + k + l = 2n + 1$ (odd sum)

Note: $e^{2\pi i n} = 1$

- $e^{\pi i} = -1^n = e^{3\pi i} = e^{5\pi} ...etc.$
- $2n + 1 = 1, 3, 5, ...$

The l index has no effect on the body-centered strucutre factor - for instance, 111, 112, and 113 all have the same $2f$ structure factor.

147.4.3 Base Center (C) Bravais Lattice

The base centered lattices also has one atom per lattice point and two atoms in total.
1 atom\lattice point $\rightarrow$N=2
The atoms are located at the following positions:
$r_1 = 000, r_2 = \frac{1}{2}\frac{1}{2}0$
Plugging these positions into the structure factor equation yields the following:

$$F_{hkl} = f(1 + e^{\pi i(h+k)})$$

Again, there are two cases, one in which the two atoms are scattering perfectly in phase and one in which they are scattering perfectly out of phase.

$$F_{hkl} = 2f \text{ for } h + k = 2n$$

$$F_{hkl} = 0 \text{ for } h + k = 2n + 1$$

147.4.4 Face Center (F) Bravais Lattice

For the face centered lattice, there is again one atom per lattice point, yielding four atoms total in the nonprimitive unit cell.
1 atom\lattice point $\rightarrow$N=4
The atoms are located at the following positions with the unit cell:
$r_n = \left\{000, \frac{1}{2}\frac{1}{2}0, 0\frac{1}{2}\frac{1}{2}, \frac{1}{2}0\frac{1}{2}\right\}$

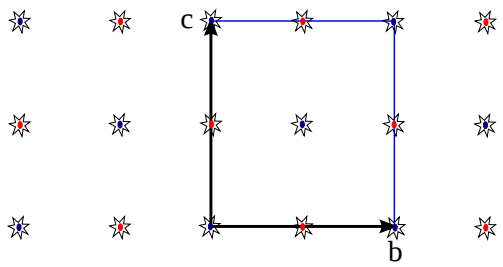


Figure 147.12: Face-center (F) Bravais lattice

The face center lattice positions yield the following structure factor:

$$F_{hkl} = f(1 + e^{\pi i(h+k)} + e^{\pi i(k+l)} + e^{\pi i(h+l)})$$

There are two distinct cases for the face center structure. That is, when the hkl are unmixed (all even or all odd), then all four atoms scatter waves perfecly in phase and when the hkl are mixed (even and odd) then the four atoms scatter waves that destructively interfere.

$$F_{hkl} = 4f, \text{ if } h, k, l \text{ are unmixed}$$

$$F_{hkl} = 0, \text{ if } h, k, l \text{ are mixed}$$

For instance, the 022 planes would yield perfectly constructive interference and the 011 planes would yield perfectly destructive interference.

$$F_{022} = 4f\left(\frac{1}{2d_{022}}\right), F_{011} = 0$$

147.4.5 Hexagonal Closed Pack (HCP)

For the hexagonal closed pack structure, there are two atoms of the same type per lattice point. Recall that the unit cell is hexagonal primitive, and there are two atoms in total. One atom is positioned at the origin 000 and the other atom is in the $\frac{1}{3}\frac{2}{3}\frac{1}{2}$ position.
Hexagonal-P Bravais Lattice
Ex. Zn, Ti, Mg

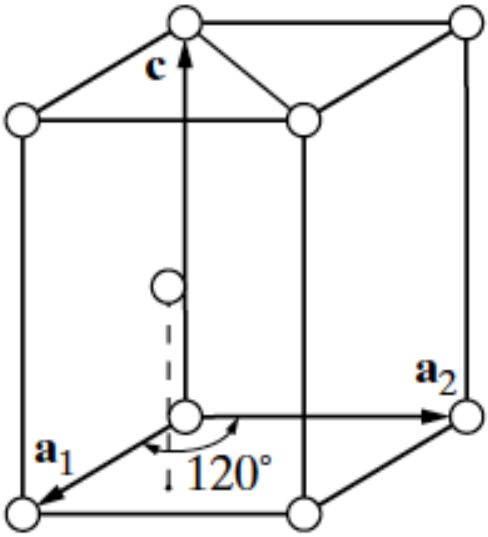


Figure 147.13: HCP structure

Using the hexagonal closed pack atom locations yields the following structure factor:
$$F_{hkl} = f\left[1 + e^{2\pi i\left(\frac{h}{3}+\frac{2k}{3}+\frac{l}{2}\right)}\right] = f(1 + e^{2\pi iq}), \text{ where } q = \frac{h+2k}{3} + \frac{l}{2}$$
The structure factor times its complex conjugate eliminates the imaginary terms and simplifies the expression into the following:
$$FF* = f^2\left(2 + e^{2\pi iq} + e^{-2\pi iq}\right) = 2f^2(1 + cos2\pi q)$$
Because there are two atoms in the unit cell, we expect two extreme conditions - one in which the two scatter perfectly out of phase and one in which the two scatter perfectly out of phase. When the parameter q is an integer, $cos2\pi q$ will equal 1, producing a very strong reflection due to constructive interference. If q is equal to an odd integer, then $2\pi q$ will equal -1, yielding a total intensity of zero.
Squaring the structure factor yields the following expression:
$$|F|^2 = 4f^2cos^2\pi q = 4f^2cos^2\left[\pi\left(\frac{h+2k}{3}\right) + \frac{l}{2}\right]$$

(1) $l = 2n$ (even) AND $h + 2k = 3m$
$\Rightarrow q = (m + n) \in Z \rightarrow cos^2\pi q = (\pm 1)^2 = 1$
$|F|^2 = 4f^2$ e.g., (002), (112), ... very strong (in-phase scattering)
$l = 2n$ AND $h + 2k = 3m \pm 1 \Rightarrow q = (m + n) \pm \frac{1}{3} = m' \pm \frac{1}{3}$
$$cos\pi q = cos\pi m' cos\tfrac{\pi}{3} \pm sin\pi m' sin\tfrac{\pi}{3} = cos\tfrac{\pi}{3} = \tfrac{1}{2}$$
$|F|^2 = f^2$ e.g., (102), (200), (100)...weak reflection

(2) $l = 2n + 1$ (odd) AND $h + 2k = 3m$
$\Rightarrow q = n + \frac{1}{2} + m = m' + \frac{1}{2} \rightarrow cos\pi q = cosm'\pi cos\frac{\pi}{2} - sinm'\pi sin\frac{\pi}{2} = 0$
$|F|^2 = 0$ e.g., (001), (111), ... forbidden reflection
$l = 2n + 1$ AND $h + 2k = 3m \pm 1 \Rightarrow q = n + \frac{1}{2} + m \pm \frac{1}{3} = m' \pm \frac{1}{6}$
$|F|^2 = 3f^2$ e.g., (103), (013), (101)... strong reflection

Therefore, there are four types of diffraction peak intensities for this hexagonal closed pack structure, ranging from forbidden reflections to very strong reflections.

147.4.6 More than one atom type per unit cell

We will now analyze the case of more than one atom type per unit cell such as the rock-salt lithium fluoride structure.

LiF→NaCl structure FCC

The following indices locate the lithium the lithium and fluorine atoms within the face centered cubic unit cell:

$$Li^+ \text{at } 000 + fct$$
$$F^- \text{at } \tfrac{1}{2}00 + fct$$
$$fct = \left\{000, \tfrac{1}{2}\tfrac{1}{2}0, 0\tfrac{1}{2}\tfrac{1}{2}, \tfrac{1}{2}0\tfrac{1}{2}\right\}$$

The structure factor is given by the following equation:

$$F_{hkl} = \sum_{n=1}^{8} f_n e^{2\pi i(hx_n+ky_n+lz_n)}$$
$$F_{hkl} = f_{Li^+}\left(1 + e^{\pi i(h+k)} + e^{\pi i(k+l)} + e^{\pi i(h+l)}\right) + f_{F^-}e^{2\pi i\frac{h}{2}}\left(1 + e^{\pi i(h+k)} + e^{\pi i(k+l)} + e^{\pi i(h+l)}\right)$$
$$F_{hkl} =$$

- $4(f_{Li^+} + f_{F^-})$ for h, k, l all even →(200), (222), etc.
- $4(f_{Li^+} - f_{F^-})$ for h, k, l all odd →(111), (311), etc.
- 0 for h, k, l mixed →(210), (001), etc.

Recall, in the lab that you saw: strong, weak, & forbidden FCC.
Kinematical scattering ignores multiple scattering due to low probabilities of re-scattering after a wave has scattered from one atom.

Note that the Laue method is useful for finding the symmetry of crystal orientations, but not useful for finding the lattice constant due to a variable wavelength. Furthermore, the Laue spot position is is independent of the length of $r*$and only dependent on direction.

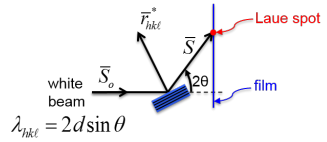


Figure 147.14: Laue spot position

147.5 Width of Diffraction Peaks (single crystal)

Reciprocal lattice points have dimensions which are in proportional to the size of a crystal in a given direction.

$$\propto \frac{1}{aM_1}, \frac{1}{bM_2}, \frac{1}{cM_3}$$

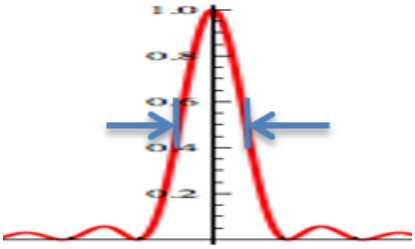


Figure 147.15: Diffraction peak

Recall the 3-D Interference Function whose M_1 component is given by the following:

$$\frac{sin^2(M_1\pi\bar{Q}\cdot\bar{a})}{sin^2(\pi\bar{Q}\cdot\bar{a})}$$

The numerator in the function determines the oscillation and width of the diffraction pattern, whereas the denominator causes the Bragg peaks to occur. When M_1 M_2 and M_3 are small numbers, the three quotients increase, thus visibly broadening the peaks in the diffraction pattern. Since the area under the peak is constant, this also results in a decreased peak height. The reciprocal lattice points which were previously assumed infintesimal, now have a breadth related to the width of the Bragg peaks as shown in the two dimensional projection of Figure 147.16. Note that the size of the crystal may be different in the a and b directions.

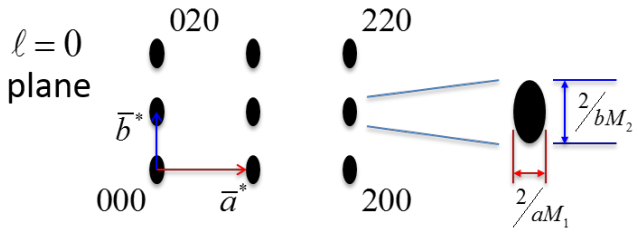


Figure 147.16: Widths of diffraction peaks

For a $\theta - 2\theta$ scan along the $\bar{r}^*_{100}$ direction, the diffraction peak width is proportional to the length of reciprocal lattice spot in $\bar{r}^*_{100}$ direction or $\frac{2}{aM_1}$. The full width half max $\Delta\theta$ length is estimated by Scherrer's Formula

$$\Delta\theta FWHM = \epsilon_{1/2}(radians) = \left(\frac{ln2}{\pi}\right)^{1/2} \frac{\lambda}{D_{hkl}cos\theta} = \frac{0.47\lambda}{D_{hkl}cos\theta}$$

Scherrer's Formula

where $D_{hkl}$is the crystal size (or X-ray coherence length) in the $\bar{r}^*_{hkl}$ direction. N
e.g., $D_{100} = aM_1$
If crystallites, such as those in a powder sample, are on the order of a tenth of a micron, then the width of the pic will be one milliradian.$D \approx 1000\mathring{A} \rightarrow \epsilon_{1/2} \approx 1mrad = 0.06$
Therefore when measuring a diffraction pattern, you first locate the peaks which give information about the lattice constant and symmetry of the crystal. The shape of the peak is an indicator of the crystal size in different direction.
Each Bragg peak maximum is given by $I^{(eu)}_{hkl} = |F_{hkl}|^2 M^2$, where F_{hkl} is affected by atomic positions in the unit cell.
Each Peak Width →crystal size
In addition, crystal size may be determined from a transverse or longitudinal scan.
D_{hkl} from $\theta - 2\theta$ scan (longitudinal scan) $D_{(hkl)\perp}$from ωscan (transverse scan)

What else can affect the height and width of the diffraction peaks?

- Angular divergence of incident beam - if an incident beam is very divergent, the
- Crystal defects - affects the width of the

To smooth over these two effects, we use angle integrated peak intensities to find relative values of $|F|^2$ (then compare to calculated $|F|^2$ from model).

Transverse scan $(\omega - scan)$ **≡rock crystal through peak with** 2θ **fixed** In a transverse scan, the sample is always at the center of the Ewald sphere. This experiment setup is illustrated in Figure 147.17 below. The incident beam wave vector of length $\frac{1}{\lambda}$ enters the sample and always hits the origin (000) of the reciprocal lattice on the Ewald sphere. For the 2θ fixed scan, in which the angle of the sample ω is varied, the reciprocal lattice points rotate about 000, as illustrated by the black dashed lines on the left side of Figure 147.17. Since 2θ is restricted to the width of the slit, the Ewald sphere intersection with lattice point at 010 will produce a line of points that does not capture the entire integrated intensity of the reciprocal lattice ellipsoid. Recall that this is a two dimensional projection, and that the line of points depicted by a horizontal red line across the blue ellipse in Figure 147.17, is actually a plane intersection of the 010 reciprocal lattice ellipsoid. If we remove the slit and do the same scan

across ω, the entire intensity is captured, as illustrated by the red ellipse in Figure 147.17. This is an angle integrated scan.

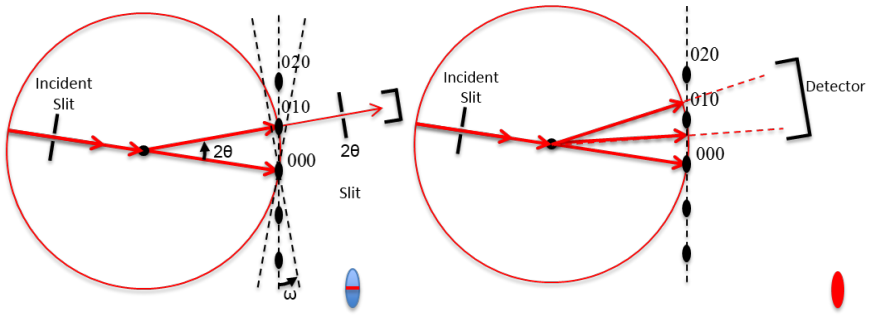


Figure 147.17: Transverse and Longitudinal Scan

The integrated counts or the energy in a peak is given by the following equation:

$$I_{hkl} = \frac{I_0}{\omega'} r_e^2 F^2 \frac{\lambda^3}{V_{uc}^2} \left(\frac{1 + cos^2 2\theta}{2 sin 2\theta} \right) \delta V$$

$\omega' = \frac{d\omega}{dt}$ ≡angular velocity of crystal rotation, a step rotational scan is now used
$\omega' = \frac{\Delta\omega}{\Delta t} \equiv \frac{angular\ step\ size}{time\ per\ step}$ →step rotational scan, a faster scan yields fewer counts
V_{uc} =unit cell volume $V_{uc} = (V_{uc}^*)^{-1}$
$\delta V = M_1 M_2 M_3$ =Effective volume of irradiated sample
$\left(\frac{1+cos^2 2\theta}{2} \right)$ =Polarization Factor for unpolarized beam
$(sin 2\theta)^{-1}$ =Lorentz Factor for single crystal diffraction with constant ω', θ ∝time taken to sweep Ewald sphere through a relative point (reciprocal lattice point)

Note that the tangential velocity of relative point through Ewald sphere is given by $\omega' \bar{r}_{hkl}^*$.
In Figure 147.18, notice that the 200 and 400 diffraction conditions are scanned at different speeds due to the geometry of the Ewald sphere . The intensity for the 200 lattice point will be greater because the Ewald sphere travels and scans more quickly through this point yielding a smaller intensity relative to the 400 lattice point. This is another reason why $I_{200} > I_{400}$ for LiF in lab

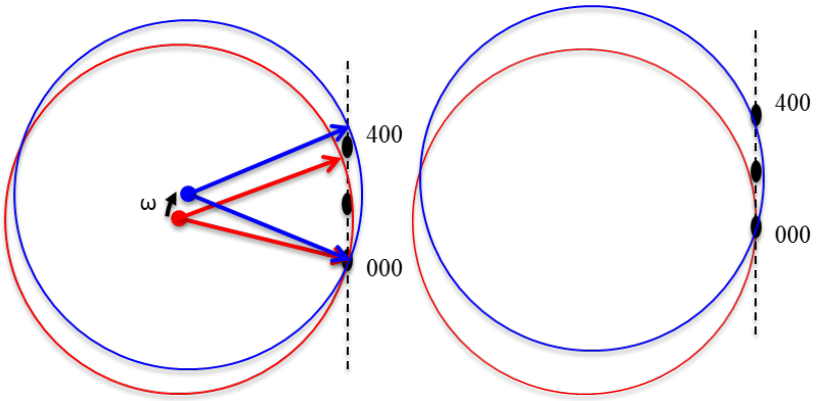


Figure 147.18: $\bar{r}_{200}^*$ vs $\bar{r}_{400}^*$ scan

147.6 Polycrystalline Aggregate (Powder sample)

In the powder method, the crystal is reduced to a fine powder of thousands of randomly oriented small grains to be placed in a beam of monochromatic x-rays. In reciprocal space, this random orientation results in a continuum

of concentric spherical shells with a range of r^*_{hkl}. The illustration in Figure 147.19 below illustrates a reciprocal lattice vector range from r^*_{200} to r^*_{400}.

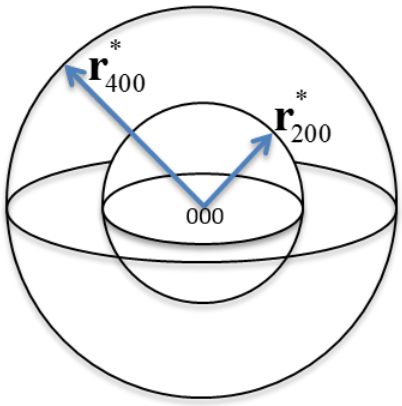


Figure 147.19: Ewald spheres for powder sample

The surface density of *hkl* relative points is equal to the number of crystallites divided by the surface are a of the spherical shell.

- Surface density$\propto \frac{N}{4\pi r^{*2}_{hkl}}$

We assume that since the powder sample is isotropic, each sphere has uniform surface density. If a texture sample is used, a preferred direction may exist and the density will be non-uniform.

Ewald construction for one r^*_{hkl} with uniform density The Ewald sphere for one r^*_{hkl} with uniform density is illustrated in Figure 147.20 below. This blue Ewald sphere has a definite center and radius, and the monochromatic incident beam $\frac{\bar{s}_0}{\lambda}$ passes through its center to hit the origin of the sample's reciprocal lattice sphere at 000. The intersection of the Ewald sphere and the sample's sphere for one r^*_{hkl}, marked by the red dashed circle, represents the entire set of allowed scattered wave vectors $\frac{\bar{s}}{\lambda}$ that will emanate from the center of the Ewald sphere. This circular locus of points will produce a diffraction cone of scattered wave vectors beginning from the origin of the Ewald sphere and ending at the circular intersection of the two spheres. In a small slit powder diffractometer, only a segment of the circle is captured as the slit scans around 2θ. Placing a film downstream from the sample would show the entire Bragg ring.

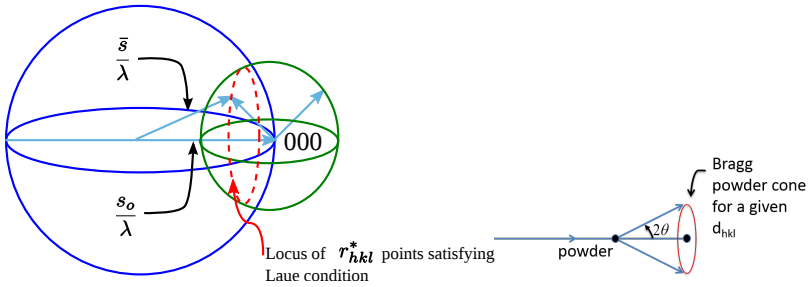


Figure 147.20: Ewald construction for powder method

This is is satisying the Laue condition in which the scattering vector is equal to a specific reciprocal lattice vector of the sample.

$$\frac{\bar{s}-\bar{s_0}}{\lambda} = \bar{r}^*_{hkl}$$

Calculated Diffracted Intensity (Power)

$$P = I_0 \frac{cos\theta}{2} r_e^2 F^2 \frac{\lambda^3}{V_{uc}^2} \left(\frac{1+cos^2 2\theta}{2sin2\theta} \right) N\delta V$$

This is for entire Bragg Ring of circumference $C = 2\pi R(sin2\theta)$, $cos\theta \propto \frac{locus\,circle}{great\,circle\,of\,r^*}$

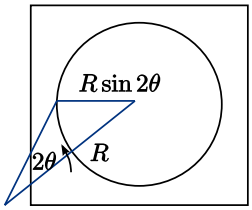


Figure 147.21: Spherical film of radius R

A horizontal diffractometer with 2θdetector slit at R and vertical slit height l records a fraction of the Bragg ring $\frac{l}{2\pi R sin2\theta}$.

Multiplicity Correction The case of multiplicity occurs when different combinations of hkl have the same same r^*_{hkl} length in reciprocal space. Since these all contribute to the same spherical shell, then the overall intensity of the diffraction condition will be greater.

E.g. Cubic:
$$r^*_{100} = r^*_{010} = r^*_{001} = r^*_{\bar{1}00} = r^*_{0\bar{1}0} = r^*_{00\bar{1}} = \frac{1}{a}$$
$$\therefore m_{100} = 6, m_{h00} = 6$$
Note: C&S $m \to p$

Powder Diffraction Intensity Adding in the multiplicity correction factor, the measure of the power for a diffraction condition of a given hkl is given by the following:

$$P'_{hkl} = \frac{I_0 l}{16\pi R} r_e^2 \frac{\lambda^3}{V_{uc}^2} F_{hkl}^2 m_{hkl} \left(\frac{1 + cos^2 2\theta}{sin2\theta sin\theta}\right) Vs$$

WhereV_s =effective sample volume $N\delta V$ and $LP = \frac{1+cos^2 2\theta}{sin2\theta sin\theta}$ =Lorentz-Polarization factor for powder diffraction
Note: $sin2\theta = 2sin\theta cos\theta$
Why do we use a symmetric reflection geometry for powder diffraction, if the random orientation of the crystals in the powder allow for various sampling angles and diffraction conditions at once?
(i.e. $\omega = \theta = \frac{2\theta}{2}$ or $\theta - 2\theta$scan)

Absorption Effects Recall that the linear absorption coefficient is represented by $\mu$and that the absorption length of a material is given by $\frac{1}{\mu}$. A low Z material like water or a hydrocarbon has a longer absorption length $\frac{1}{\mu}$than a solid metal.
Consider a large planar sample with thickness t in reflection geometry as illustrated in Figure 147.22. The incident beam with area A across enters at an angle α, and scatters from a differential volume at some depth z. The 2θdetector is lined up to pick up the scattered intensity at the angle β. In general, α and β do not have to be equal.

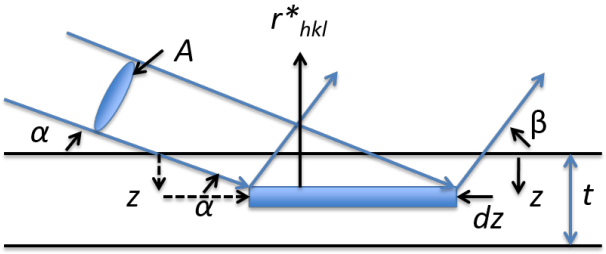


Figure 147.22: Planar sample in reflection geometry

the beam area produces a projected area of the slab that is A/sinα
Therefore the volume element is given by$dV = \frac{A}{sin\alpha}dz$

The diffracted intensity at depth z is given by the following:

$$dI_z = C\frac{A}{sin\alpha}dzI_0e^{-uz/sin\alpha}e^{-uz/sin\beta}$$

There is an absorption effect that attenuates the beam as it enters and exits sample, which is accounted for by the following two transmission factors:

$e^{-uz/sin\alpha}$- in transmission factor

$e^{-uz/sin\beta}$- out transmission factor

To calculate the total intensity, we integrate over the total thickness:

$$I = \int_0^t dI_z = I_0C\frac{A}{\mu}\left(1 + \frac{sin\alpha}{sin\beta}\right)^{-1}$$

Assuming the thickness is much greater than the attenuation length, or $\mu t \gg 1$ - for a typical oxide material this thickness is about 0.5 mm.
$\int_0^t dI_z = \int_0^\infty dI_z$ "t" drops out by assuming $\mu t \gg 1$ or $t \to \infty$

For symmetric reflection $\alpha = \beta = \theta \to I = I_0C\frac{A}{2\mu}\frac{A}{2\mu}$ =effective volume = V_{eff}

Note that the effective volume,V_{eff} , is constant and independent of $\theta$for $\alpha = \beta = \theta$ As you increase the angle in this geometry, you are increasing the depth but decreasing the effective area. Therefore, it is convenient to use symmetric reflection geometry for a powder diffraction experiment, in order to be able to avoid varying absorption effects from a varying effective volume.

We could also do diffractometry in symmetric transmission mode as illustrated in Figure 147.23.
We want as much material in front of the beam to yield a high scattering intensity, but we also want a small enough thickness to avoid too much absorption. There is an optimal thickness in this transmission geometry, $\frac{1}{\mu}$, that yields the maximum scattering compromised correctly with the absorption. Recall that $\frac{1}{\mu}$ is the absorption length.
Effective Volume: $\frac{At}{cos\theta}e^{\frac{-\mu t}{cos\theta}}$
Here,V_{eff} depends on θ.

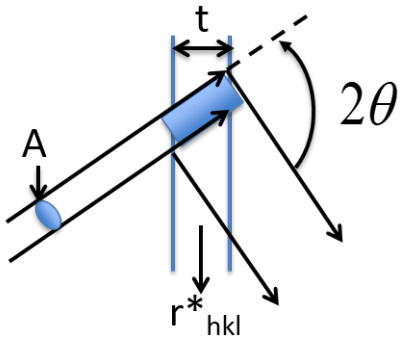


Figure 147.23: Transmission geometry

Summarize: Powder diffractometer in symmetric reflection geometry The area under the diffraction peak is proportional to:

$$I_{hkl} = |F_{hkl}|^2 m_{hkl} \left(\frac{1 + cos^2 2\theta}{sin^2 \theta cos\theta} \right) e^{-2M}$$

F_{hkl}- structure factor hkl and $sin\theta/\lambda$

m_{hkl}- multiplicity (hkl)

$\left(\frac{1+cos^2 2\theta}{sin^2\theta cos\theta}\right)$- Lorentz polarization factor, θ

e^{-2M}- temperature dependent Debye-Waller factor $sin\theta/\lambda$accounting for atomic vibrations

Our objective is to measure the integrated intensities $I'_{hkl}s$ and use the structure factor $F'_{hkl}s$ to determine the crystal structure.
At extreme angles, the Lorentz polarization factor has a considerable effect, as shown in Figure 147.24.

$$LP = \left(\frac{1+cos^2 2\theta}{sin^2\theta cos\theta}\right)^{\leftarrow Polarization\ correction}_{\leftarrow Geometrical\ correction}$$

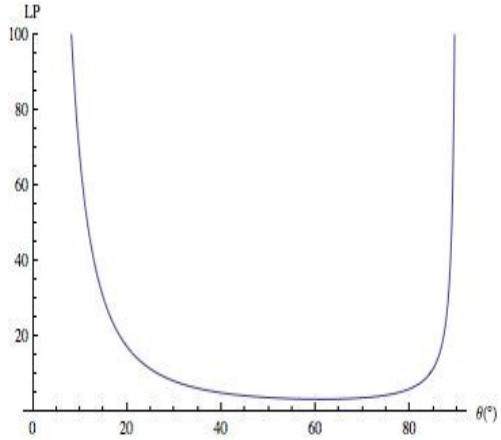


Figure 147.24: Theta vs. LP

Multiplicity Factor m_{hkl} : number of (hkl) planes in the $\{hkl\}$ family for a given crystal system
Example:

0kk

022	$0\bar{2}\bar{2}$
202	$\bar{2}0\bar{2}$
220	$\bar{2}\bar{2}0$
$02\bar{2}$	$0\bar{2}2$
$20\bar{2}$	$\bar{2}02$
$2\bar{2}0$	$\bar{2}20$

Cubic $\rightarrow m_{0kk} = 12$, since all 12 $|r^*|'s$ are equal
triclinic $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$

$r^*_{022} \neq r^*_{220}$ triclinic $m_{hkl} = 2$
$r^*_{022} = r^*_{02\bar{2}}$ Since only $r^*_{\bar{h}\bar{k}\bar{l}} = r^*_{hkl}$
See Cullity and Stock Appendix 11

Cubic:	$\frac{hkl}{48*}$	$\frac{hhl}{24}$	$\frac{0kl}{24^*}$	$\frac{0kk}{12}$	$\frac{hhh}{8}$	$\frac{00l}{6}$	
Hexagonal and Rhombohedral:	$\frac{hk\cdot l}{24^*}$	$\frac{hh\cdot l}{12^*}$	$\frac{0k\cdot l}{12^*}$	$\frac{hk\cdot 0}{12^*}$	$\frac{hh\cdot 0}{6}$	$\frac{0k\cdot 0}{6}$	$\frac{00\cdot l}{2}$
Tetragonal:	$\frac{hkl}{16^*}$	$\frac{hhl}{8}$	$\frac{0kl}{8}$	$\frac{hk0}{8^*}$	$\frac{hh0}{4}$	$\frac{0k0}{4}$	$\frac{00l}{2}$
Orthorhombic:	$\frac{hkl}{8}$	$\frac{0kl}{4}$	$\frac{h0l}{4}$	$\frac{hk0}{4}$	$\frac{h00}{2}$	$\frac{0k0}{2}$	$\frac{00l}{2}$
Monoclinic:	$\frac{hkl}{4}$	$\frac{hol}{2}$	$\frac{0k0}{2}$				
Triclinic:	$\frac{hkl}{2}$						

Table 147.2: Mutliplicity factors for powder photographs

147.6.1 Indexing Powder Diffraction Patterns

We can determine r^*_{hkl} lengths (not directions) from experimentally measured $\theta's$ and a known λ using the following relationships:

$$\frac{2sin\theta_{hkl}}{\lambda} = r^*_{hkl} = (\bar{r} * \cdot \bar{r} *)^{\frac{1}{2}}$$

$$= [h^2a*^2 + k^2b*^2 + l^2c*^2 + 2hka^*b*cos\gamma* + 2klb*c*cos\alpha* + 2hla^*c*cos\beta*]^{\frac{1}{2}}$$

First index the series of peaks (i.e. determine $(hkl)_n$), then determine the lattice constants of the Bravais lattice : $a, b, c, \alpha, \beta, \gamma$
Initially check for cubic, the simplest case, where$a = b = c$, $\alpha = \beta = \gamma = 90$.

Cubic: $\underbrace{r*^2}_{measured} = \underbrace{a*^2 [h^2 + k^2 + l^2]}_{unknown}$ subject to Bravais lattice conditions.
See Cullity and Stock handout, Appendix 9 - also shown below in Table 147.3.

$h^2 + k^2 + l^2$	hkl	P	F	I	Diamond
1	100	x			
2	110	x		x	
3	111	x	x		x
4	200	x	x	x	
5	210	x			
6	211	x		x	
7					
8	220	x	x	x	x
9	300,221	x,x			
10	310	x		x	
11	311	x	x		x
12	222	x	x	x	
13	320	x			
14	321	x		x	
15					
16	400	x	x	x	x

Table 147.3: Determining unit cells from $h^2 + k^2 + l^2$

Therefore, we can check for pattern in the sequence of measured $r*^2$values, which are proportional to $h^2 + k^2 + l^2$, and relate these to the corresponding unit cells using Table 147.3.

If $\{r*_i^2\} =$
$C\{1, 2, 3, 4, 5, 6, 8, ...\}$ →Cubic-P

$C\{1, \frac{4}{3}, \frac{8}{3}, \frac{11}{3}, 4, ...\} \rightarrow$Cubic-F
$C\{1, 2, 3, 4, 5, 6, 7, 8\} \rightarrow$Cubic-I
$C\{1, \frac{8}{3}, \frac{11}{3}, \frac{16}{3}, \frac{19}{3}, ...\} \rightarrow$Diamond

Note that we need to measure $r* = 2sin\vartheta/\lambda$ for at least the first seven peaks to distinguish between Cubic-P and Cubic-I.

Tetragonal and hexagonal ("uniaxial")

Tetragonal: $r^{*2}_{hkl} = a^{*2}(h^2 + k^2) + c^{*2}l^2 = a^{*2}\left[(h^2 + k^2) + \left(\frac{c^*}{a^*}\right)^2 l^2\right]$

Hexagonal: $cos\gamma* = \frac{1}{2}$

$r^{*2}_{hkl} = a^{*2}(h^2 + k^2 + hk) + c^{*2}l^2 = a^{*2}\left[(h^2 + k^2 + hk) + \left(\frac{c^*}{a^*}\right)^2 l^2\right]$

$\left(\frac{c^*}{a^*}\right)^2 = \left(\frac{a}{c}\right)^2$ is unknown

is 1st peak ?:

100 or 001 (P)

110 or 002 (Tetragonal I)

100 or 002 (HCP)

$sin^2\theta = \frac{\lambda^2}{4} r^{*2}_{hkl} = A(h^2 + k^2) + Cl^2 \leftarrow Tetragonal$

$A(h^2 + k^2 + hk) + Cl^2 \leftarrow Hexagonal$

Example 1 $\lambda = 1.542\mathring{A}$

Line No.	$sin^2\theta$	$\frac{(sin^2\theta)_n}{(sin^2\theta)_1}$		$h^2 + k^2 + l^2$	hkl
1	0.0603	1	3/3	3	111
2	0.1610	2.67	8/3	8	220
3	0.221	3.67	11/3	11	311
4	0.322	5.34	16/3	16	400
5	0.383	6.35	19/3	19	331
6	0.484	8.03	24/3	24	422
7	0.545	9.04	27/3	27	333
8	0.645	10.7	32/3	32	440

Table 147.4: Indexing Powder Patterns, Example 1

Ratios $\rightarrow$Cubic

$\frac{(sin^2\theta)_n}{(sin^2\theta)_1} = \frac{(h^2+k^2+l^2)_n}{(h^2+k^2+l^2)_1}$

unmixed hkl: Cubic F

$h + k + l \neq 4n \pm 2$: diamond cubic

What is "a"?

$r^*_{400} = 16a*^2 = \frac{16}{a^2} = \frac{4sin^2\theta_{400}}{\lambda^2}$

$a = \frac{4\lambda}{2sin\theta_{400}} = \frac{2\times 1.542\mathring{A}}{(0.322)^{1/2}} = 5.434\mathring{A}$

Silicon a=5.431$\mathring{A}$

Example 2 $\lambda = 1.542\mathring{A}$

Line No.	$sin^2\theta$	$(sin^2\theta)_n\ (sin^2\theta)_1$	$nA+mC$	$(sin^2)_{calc}$	hkl
1	0.0806	1.00	A+C	0.0806	101
2	0.0975	1.21	4C	0.0976	002
3	0.1122	1.39	2A	0.1124	110
4	0.210	2.61	2A+4C	0.210	112
5	0.226	2.80	4A	0.225	200
6	1.274	3.40	A+9C	0.276	103
7	0.305	3.78	5A+C	0.306	211
8	0.321	3.98	4A+4C	0.323	202

Table 147.5: Indexing powder patterns, example 2

Ratios →not cubic
Note Relations:
(4)=(2)+(3)
(5)=2(3)
(7)=(5)+(1)
(8)=4(1)

Assume Tetragonal:
$sin^2\theta = A(h^2+k^2)+Cl^2$

Tetragonal P→(1) = 100 or 001
No peak at 2(1) →not P
2(1) =110, if (1) = 100
Tetragonal I: (1) =110,101, or 002

Try (1)=101 → $sin^2\theta = A+C$
$sin^2\theta = A(h^2+k^2)_n + C(l^2)_n$

hk	00 10 11 20 21 22 30 31 32 ... l 0 1 2 3 4 ...
h^2+k^2	0 1 4 5 8 9 10 13 ... l^2 0 1 4 9 16 ...

If $(1) = 101 = A+C \rightarrow (2) = 002 = 4C$ *
If $(1) = 002 = 4C \rightarrow (2) = 101 = A+C$ } Think about it
If $(1) = 110 = 2A \rightarrow (2) = 101 = A+C$ }
* Try:
$(2) \rightarrow 4C = 0.0975$ } $A = 0.0562, C = 0.0244$
$(1) \rightarrow A+C = 0.083$ }
Note $(3) = 2A \rightarrow 110 = A(1^2+1^2)+C(0^2)$

Check to see if each $(sin^2\theta)_n = A(h^2+k^2)_n + Cl_n^2$ for A = 0.0652 and C = 0.0244
Conclusion: Tetragonal I

Lattice constant. a=? c=?
$(2) \rightarrow 2c* = \frac{2}{c} = \frac{2sin\theta_{002}}{\lambda} \Rightarrow c = \frac{\lambda}{sin\theta_{002}} = \frac{1.542\mathring{A}}{(0.0975)^{\frac{1}{2}}} = 4.94\mathring{A}$
$(3) \rightarrow \sqrt{2}a* = \frac{\sqrt{2}}{a} = \frac{2sin\theta_{110}}{\lambda} \Rightarrow a = \frac{\lambda}{\sqrt{2}sin\theta_{110}} = 3.62\mathring{A}$ **Indium**

Procedure for hexagonal similar to tetragonal except:
$(sin^2\theta)_n = A(h^2+k^2+hk)_n + l^2C$
where $h^2+k^2+hk = \{0,1,3,4,7,9,12,13,...\}$
Indexing difficulty increases as crystal symmetry decreases: need computer
Note: When $K\alpha_1$ and $K\alpha_2$are unresolved (at lower 2θ), use avg $\lambda_{K\alpha} = 2/3\lambda_{K\alpha 1} + 1/3\lambda_{K\alpha_2} = 1.5418\mathring{A}$
At higher 2θ,$K\alpha_1$and $K\alpha_2$may be resolved, use correct $\lambda_{K\alpha 1}$ and $\lambda_{K\alpha 2}$, not $\lambda_{K\alpha}$

147.6.2 Quantitative Analysis

Recall: $I_{hkl} = C|F_{hkl}|^2 \frac{m_{hkl}}{V_C^2} LP \frac{A}{2\mu}$, ignore e^{-2M}

Consider a 2 phase powder with:
v_A =volume fraction of phase A, $1 - v_A = v_B$ =volume fraction of phase B
Phase A diffraction peaks

$$I_A = C \left[|F_{hkl}|^2 \frac{m_{hkl}}{V_C^2} LP \right]_A v_A \frac{A}{2\mu} = C K_A v_A \frac{A}{2\mu}$$

$K_A$varies with *hkl*

μvaries with v_A

Phase B diffraction peaks

$$I_B = C K_B (1 - v_A) \frac{A}{2\mu}$$

$\mu$is the same

Hence,

$$\frac{I_A}{I_B} = \frac{K_A}{K_B} \frac{v_A}{(1 - v_A)}$$

To determine volume fraction for mixture of 2 powder crystalline phases A&B

$$\underbrace{\frac{I_A}{I_B}}_{\text{measure}} = \overbrace{\frac{K_A}{K_B}}^{calculate} \underbrace{\frac{v_A}{(1 - v_A)}}_{\text{determine } v_A}$$

The relative intensities of peaks for each phase should be checked to ensure an ideally random powder sample. **Texture** will cause errors in quantitative analysis.

In the analysis of the integrated area under the diffraction peaks, the Debye-Waller factor, e^{-2M_A};e^{-2M_B} , may be included if known. The two different phases are made up of atoms in two different lattices with different vibrational amplitudes. The vibrational amplitude M is given by the following:

$$M = \frac{2\pi^2 \langle u_{hkl}^2 \rangle}{d_{hkl}^2}$$

$\sqrt{\langle u_{hkl}^2 \rangle}$= rms vibrational amplitude in hkl direction, typically isotropic and ~ 0.1$\mathring{A}$

An increase in temperature increases the vibrational amplitude of atoms.

147.7 Rotating Crystal Method

In therotating crystal method, a single crystal is mounted with an axis that is normal to a monochromatic beam. A cylindrical film encloses the single crystal, which records a diffraction pattern as the crystal is rotated. This setup is illustrated in Figure 147.25. Whereas older films required chemical exposure to develop, x-ray sensitive image plates are used today and diffraction patterns are read out by a laser.

In this procedure, the crystal is rotated smoothly about a chosen axis that coincides with the axis of the film cylinder and different planes cause diffraction peaks at points in the rotation. Diffraction peaks occur when a set of lattice planes make the correct Bragg angle for reflection. In the setup illustrated in Figure 147.25, prior to inserting the crystal in the camera, the c-axis of the crystal's unit cell was determined and the crystal is then rotated about this axis. Due to this rotation about the c-axis, the spots appear in layers with heights according to constant l index values. All of the center spots are $hk0$.

The rotating crystal method helps determine the structure of a crystal in an organized manner, and is useful for methods where a single crystal can be attained as well as when the$c$ and c^* axes of a crystal's unit cell coincide with each other.

Rotating Crystal Method

(1) Monochromatic X-ray beam diffraction from single crystal surrounded by film cylinder.

(2) The crystal continually rotated about the axis that coincides with film cylinder axis

(3) If rotation axis = unit cell axis (e.g., $\vec{c}$) then reflection pattern (spots) on film form circles at heights corresponding to constant (e.g., l)

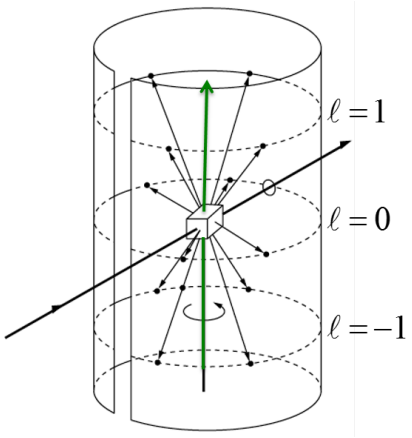


Figure 147.25: Rotating crystal method

Figure 147.26 below illustrates a developed rotating-crystal film of a quartz crystal. The extra streaks in the around the center of the pattern are due to Bremsstrahlung radiation not removed by the filter.

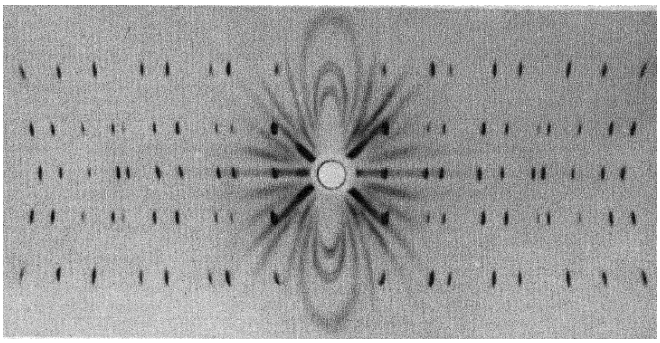
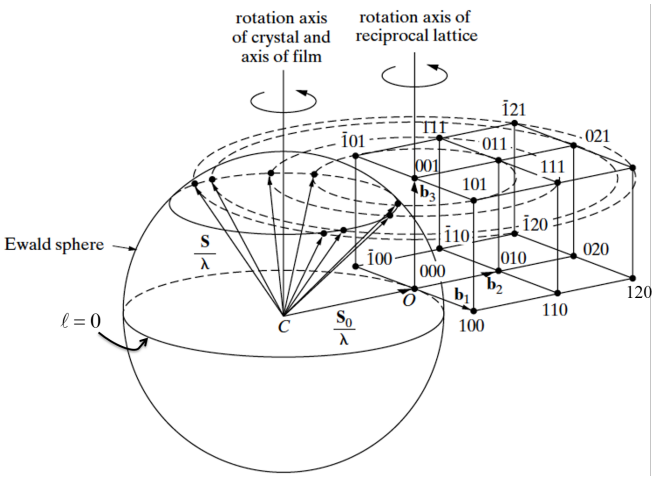


Figure 147.26: Rotating-crystal pattern of a quartz crystal (hexagonal) rotated about its c axis. Filtered copper radiation. (Courtesy of B.E. Warren.)

147.7.1 Reciprocal Lattice Treatment of Rotating-Crystal Method

How do we picture the rotating-crystal method in reciprocal space? Visualizing the method in reciprocal space, the Ewald sphere is the instrument, and the reciprocal lattice structure is the crystal sample. An orthorhombic crystal crystal centered at the Ewald sphere may be rotated about the c-axis, as illustrated in Figure 147.27 Recall that the incident wave vector always points from the center of the Ewald sphere to the origin of the reciprocal lattice. If the crystal is rotating about the c-axis, then reciprocal lattice rotates about its $\bar{c}^*$ or $\bar{b}_3$axis. A diffraction peak will appear when a reciprocal lattice point touches the Ewald sphere. In this camera, all diffracted beams will be detected and recorded except those that travel close to the cylinder's central axis and miss the film.

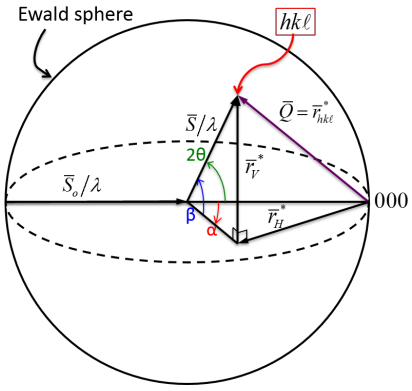


$\bar{b}_1 = \bar{a}^*, \bar{b}_2 = \bar{b}^*, \bar{b}_3 = \bar{c}^*$

Figure 147.27: Rotating crystal method in reciprocal space

- The reciprocal lattice rotates about $\bar{r}^*_{100} = \bar{c}^*$as the crystal rotates about $\bar{c}$.
- Spots appear on the film when reciprocal lattice points coincide with the Ewald sphere.
- Recall the Laue condition, when $\frac{\bar{S}}{\lambda} - \frac{\bar{S}_0}{\lambda} = \bar{Q} = \bar{r}^*_{hkl}$

We will divide the scattering vector into horizontal and vertical components, $\bar{r}^*_H$ and $\bar{r}^*_V$.The α and β angles determine the in plane and out of plane components of $\bar{r}^*_{hkl}$.



$\bar{r}^* = \bar{r}^*_H + \bar{r}^*_V$
Horizontal and Vertical components
α =In-Plane Angle
β =Elevation Angle

Chicago = $88W, 42N$

Figure 147.28: Ewald sphere geometry

After the scattered diffracted beam places various spots on the film, the film is unrolled and the specific x and y coordinates of these characteristic spots are noted. Using these coordinates and the radius of the cylinder, α and β may be determined using the following equations:

$$(1)\ tan\beta - y/R$$

$$(2)\alpha = x/R$$

$$(3)\ |\vec{r}_V^*| = \left(\frac{1}{\lambda}\right) sin\beta$$

$$(4)\ |\vec{r}_H^*| = |h\vec{a}* + k\vec{b}*|^2 = \frac{1}{\lambda^2}[2(1 - cos\alpha cos\beta) - sin^2\beta]$$

Example: $\lambda = 1.542\mathring{A}$,$a^* = 0.25\mathring{A}^{-1}$, $b^* = 0.15\mathring{A}^{-1}$, $c^* = 0.36\mathring{A}^{-1}$, and $R = 50$ mm

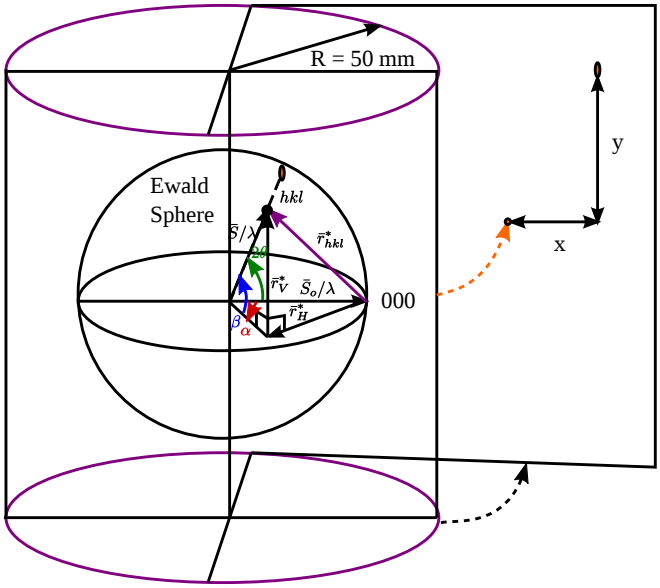


Figure 147.29: Convert x,y coordinates of film spot to α, β

147.7.2 Ewald Construction for Rotating Crystal Method

In the rotating crystal method, the diffraction spot positions (x, y coordinates) are experimentally determined and the film cylinder radius is known. The y coordinate gives the layer line, such as $l = 0, l = \pm 1$,etc. and the x coordinate indicates where this spot is located on the layer line. Using these values, we can determine the in-plane (α) and elevation (β) angles, as well as information about the scattering vector which is coincident with the reciprocal lattice vector for a diffraction spot. Geometric relationships using the Ewald sphere are pictured in Figure 147.20 below.

Define angles:

$$\alpha = x/R \rightarrow \text{in-plane angle}$$

$$tan\beta = y/R \rightarrow \text{elevation angle}$$

$$\text{Separate } \bar{r}_{hkl}^* = \bar{r}_H^* + \bar{r}_V^*$$

$$|\vec{r}_V^*| = \left|\frac{\bar{S}}{\lambda}\right| sin\beta \rightarrow \lambda r_V^* = sin\beta$$

Note $cos2\theta = cos\alpha cos\beta$

$$\lambda^2 (r_H^*)^2 = 2(1 - cos\alpha cos\beta) - sin^2\beta$$

Bragg's law below

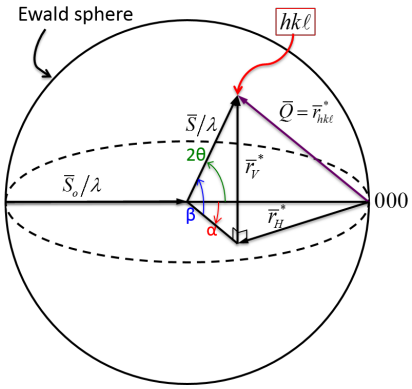


Figure 147.30: Ewald sphere for rotating crystal method

$$(r^*_{hkl})^2 = (r^*_H)^2 + (r^*_V)^2 = \frac{4sin^2\vartheta}{\lambda^2}$$

A. For hexagonal ,tetragonal, cubic, orthorhombic, etc., where $\bar{c}$ coincides with $\bar{c}^*$, measuring the y coordinate allows us to calculate the elevation angle, and the vertical component of the scattering vector

$y \to \beta \to r^*_V \Rightarrow r^*_{00l} \to c$ determined directly

By remounting the crystal to rotate about $\bar{a}$or $\bar{b}$, we can determine a and b.

B. Indexing a reflection in layer line

For a given layer line, β is constant.

$\bar{r}^*_H = h\bar{a}^* + k\bar{b}^*$for l =constant layer line

$$\lambda^2 (r^*_H)^2 = [h^2(a^*)^2 + k^2(b^*)^2 + 2hka^*b^*cos\gamma^*]\lambda^2$$
$$= 2(1 - cos\alpha cos\beta) - sin^2\beta$$
$$\alpha = x/R \; tan\beta = y/R$$
$$\therefore x, y \to r^*_H, r^*_V$$

Bernal Chart The Bernal chart is a graphical method that aids in the conversion the x and y coordinates of diffraction spots to $\lambda r^*_H, \lambda r^*_V$. This graphical aid can be superimposed on the film roll, as shown in Figure 147.31 below.

Graphical method for converting $x, y \to \lambda r^*_H, \lambda r^*_V$

Horizontal Lines →constant l for $\bar{c}$ axis rotating

λr^*_H lines equally incremented from 0 to 2.0. The maximum is when $2\theta = \pi$

$$\lambda r^*_H = 0 \to x = 0 \to \alpha = 0 \to hk = 00$$
$$\lambda r^*_H = 2 \to x = \pi R \to \alpha = \pi \to r^* = r^*_H = \frac{2}{\lambda} \text{ for } l = 0$$

Start with $l = 0$ layer

- Assume $a^* = b^*, \gamma^* = 90 \to \frac{1}{a}(h^2 + k^2)^{\frac{1}{2}}$
- Assume $a^* \neq b^*, \gamma^* = 90 \to r^*_H = \left(\frac{h^2}{a^2} + \frac{k^2}{b^2}\right)^{1/2}$

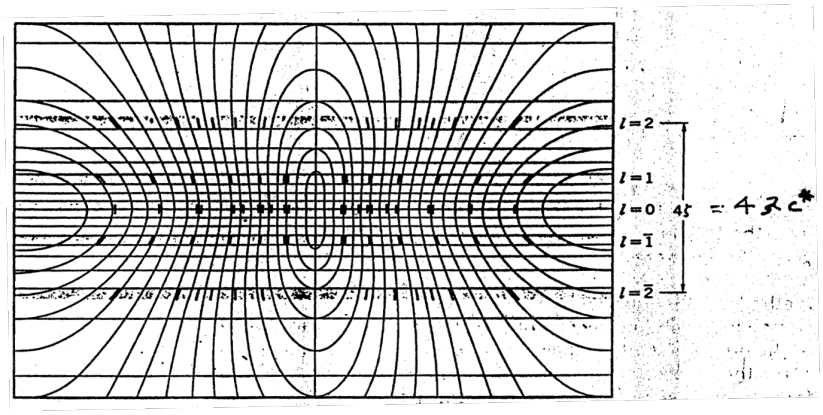


Figure 147.31: Schematic view of film overlain on a Bernal chart. Note that the X-ray diffraction spots lie along rows of constant ξandζ. (Only a few possible reflections are indicated)

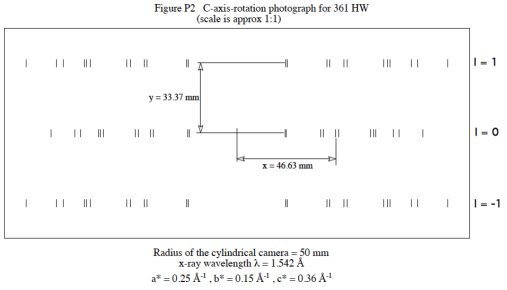


Figure 147.32: C-axis rotation photograph

Once the lattice constants and orthorhombic unit cell have been determined, index and determine Lattice Type: P, I, C, F.

Measure X and Y from the film
R = 50 mm , $\lambda = 1.542\mathring{A}$
For $l = 0$ layer, $\rightarrow \beta = 0 \rightarrow r_H^* = \frac{1}{\lambda}\sqrt{2(1 - cos\alpha)}$

X [mm]	$\alpha = X/R$	α[]	$r_H^*[\mathring{A}^{-1}]$	hkl
22.4	0.488	25.67	0.288	110
23.3	0.466	26.70	0.299	020
39.6	0.792	45.38	0.500	200
40.8	0.816	46.75	0.515	130
46.6	0.932	53.40	0.583	2220
48.1	0.962	55.12	0.600	040
63.1	1.262	72.31	0.765	310
64.6	1.292	74.03	0.781	240
65.5	1.31	75.06	0.790	150
74	1.48	84.80	0.875	330
76.7	1.534	87.89	0.900	060
88	1.76	100.84	1.000	400

r_H^*defines radius of circle in the $l = 0$ reciprocal space plane

Table 147.6: Conversion of X to hkl

Figure 147.33 below demonstrates the graphical solution for layer $l = 0$.
Plot hk0 lattice: $a^* = 0.25\mathring{A}^{-1}$ & $b^* = 0.15\mathring{A}^{-1}$
Plot circles with radii r_H^*, shown in Figure 147.33 below.
The blue squares represent the reciprocal lattice points corresponding to the planes that caused diffraction to occur.

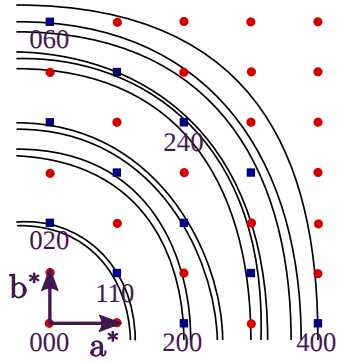


Figure 147.33: Graphical Solution for $r_H^* = h^2a^{*2} + k^2b^{*2}$

Repeat this procedure for $l = 1$ layer $\rightarrow \beta = tan^{-1}(y/R)$
$r_H^* = \frac{1}{\lambda}\sqrt{2(1 - cos\alpha cos\beta - sin^2\beta}$

148 Temperature Effects

The atoms in a crystal vibrate about equilibrium lattice positions, even at 0K and the vibrational amplitudes increase as temperature increases. With increasing atomic kinetic energies, the unit cell expands and the lattice constant becomes greater. This temperature effect, also known as the Debye-Waller effect, affects the area under diffraction peaks even at room temperature. It is important to account for this effect since the area under diffraction peaks gives information about the Bravais lattice of the crystal.
Figure 148.1 illustrates a simple model of interatomic spacing in a crystal as a function of energy, in which the red dots indicate the average interatomic distance. Note that as temperature and energy increase, the width of the potential well widens to the right, towards a greater separation distance. An increase in temperature leads to a sampling of higher energies in the potential well and a lattice expansion as indicated by the red dots in the plot. The width bounded by the two walls of the potential also increases, allowing the atoms to take greater path lengths and have greater vibrational amplitudes.

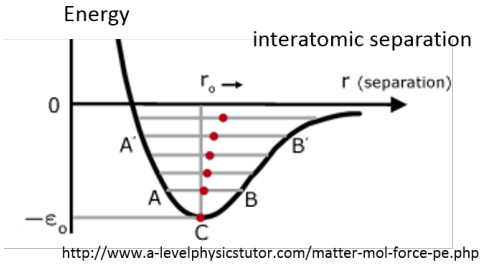


Figure 148.1: Interatomic separation as a function of energy

Describe a crystal in terms of primitive unit crystal. Considering a crystal of M primitive unit cells, the $\bar{R}'_m$ will replace the former lattice point position vector, which now includes an extra factor - a time dependent displacement vector for each atom.

- 1 atom per lattice point at origin
- Let $M = M_1M_2M_3$ =# of unit cells
- Let $m = \{m_1m_2m_3\}$= unit crystal index
- Let $\bar{R}_m$ =lattice point position vector. Locates origin of the m^{th}unit cell
- Let $\bar{R}'_m(t) = \bar{R}_m + \underbrace{\bar{\Delta}_m}(t)$
instantaneous displacement changing in time

Atoms can vibrate longitudinally or transversely, as shown in the figures below.

Figure 148.2: Longitudinal Wave

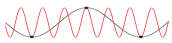


Figure 148.3: Transverse Wave

The summation of the coherently scattered wave intensities in electron units from M atoms for a crystal with one atom per unit cell is given by the following:

$$I_{(eu)}(\bar{Q}) = \left[\sum_m^M f_m e^{2\pi i \bar{Q}\cdot\bar{R}_m}\right]\left[\sum_m^M f_m e^{-2\pi i \bar{Q}\cdot\bar{R}_m}\right] = \sum_m^M\sum_n^M f_m f_n e^{2\pi i \bar{Q}\cdot(\bar{R}_m-\bar{R}_n)}$$

Assuming all of the atoms are the same, the f_m and f_n terms may be taken out of the summation. Adding in the time dependent Debye-Waller factors yields the following:

$$I_{(eu)}(\bar{Q}) = f^2\sum_m\sum_n e^{2\pi i \bar{Q}\cdot(\bar{R}_m-\bar{R}_n)}e^{2\pi i \bar{Q}\cdot(\bar{\Delta}_m-\bar{\Delta}_n)}$$

Recall that $\bar{Q}$ is the scattering vector that is coincident with r^*_{hkl} under the Bragg diffraction condition.

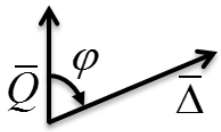


Figure 148.4: Scattering vector and Δ

$$\bar{Q} \cdot \bar{\Delta}_m = \underset{\hookrightarrow Q}{\frac{2sin\theta}{\lambda}} \Delta_m cos\phi_m$$

The vector u_m is the projection of the delta vector along the scattering vector's direction.
$u_m = \Delta_m cos\theta_m$

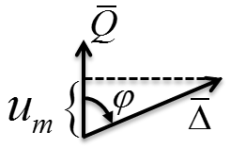


Figure 148.5: u_m

Note that $u_m = 0$ for longitudinal wave along diffraction plane.

Using $2\pi Q = \frac{4\pi sin\theta}{\lambda} = q$, the intensity equation is the following:
$I_{(eu)}(\bar{Q}) = f^2 \sum_m e^{2\pi i \bar{Q} \cdot (\bar{R}_m - \bar{R}_n)} e^{iq(u_m - u_n)}$

Note that $\Delta_m; u_m$are instantaneous vibrations that occur in the phonon lifetime which is on the order of ~ 10^{-14} seconds , and the intensity measurement time is > 10^{-3}seconds. Therefore, a time average of these instantaneous intensities as affected by vibrational effects is used.
The function is integrated over a period of time T that is long relative to the fluctuations and is normalized by dividing by T.

$$\langle f(t) \rangle = \frac{1}{T} \int_0^T f(t) dt$$

The time averaged intensity is indicated by bracket notation:

$$\langle I_{(eu)} \rangle = f^2 \sum_m \sum_n e^{2\pi i \bar{Q} \cdot (\bar{R}_m - \bar{R}_n)} \langle e^{iq(u_m - u_n)} \rangle$$

Simplifying,

$$e^{ix} = 1 + ix - \frac{x^2}{2} - \frac{ix^3}{6} + ...$$

$$\langle e^{iq(u_m - u_n)} \rangle = 1 + iq\langle u_m - u_n \rangle - \frac{q^2}{2} \langle (u_m - u_n)^2 \rangle - ...$$

$$\approx 1 - \frac{q^2}{2} \langle (u_m - u_n)^2 \rangle = e^{\frac{-q^2}{2} \langle (u_m - u_n)^2 \rangle} = e^{\frac{-q^2}{2} \langle u_m^2 + u_n^2 - 2u_m u_n \rangle}$$

$$\langle u_m^2 \rangle = \langle u_n^2 \rangle = \langle u^2 \rangle$$

$$\Rightarrow$$

$$\langle e^{iq(u_m-u_n)}\rangle \cong e^{-q^2\langle u^2-u_m u_n\rangle}$$

Let

$$\langle u_m u_n\rangle = \underbrace{\alpha_{mn}}_{\text{correlation function}}\langle u^2\rangle$$

which accounts for the correlation of vibrational motion between the atoms, and may range from 0 to 1.

$$\langle e^{iq(u_m-u_n)}\rangle = e^{-q^2\langle u^2\rangle(1-\alpha_{mn})}$$

and

$$\langle I_{(eu)}\rangle = \sum_m\sum_n f^2 e^{2\pi i\bar{Q}\cdot(\bar{R}_m-\bar{R}_n)}\underbrace{e^{q^2\langle u^2\rangle\alpha_{mn}}}_{=1+q^2\langle u^2\rangle\alpha_{mn}\text{ for }q^2\langle u^2\rangle\alpha_{mn}\ll 1}$$

$$= e^{-q^2\langle u^2\rangle}\underbrace{\sum_m\sum_n f^2 e^{2\pi i\bar{Q}\cdot(\bar{R}_m-\bar{R}_n)}}_{[CR]} + \underbrace{q^2\langle u^2\rangle e^{-q^2\langle u^2\rangle}}_{xe^{-x}}\sum_m\sum_n f^2\alpha_{mn}e^{2\pi i\bar{Q}\cdot(\bar{R}_m-\bar{R}_n)}$$

$$[CR] = \text{Crystal Reflection } = f^2[\text{3-D interference function}]$$

With higher 2θ, the diffraction peak intensities decrease due to the damping term $e^{-q^2\langle u^2\rangle}$at the front.

$$xe^{-x} = x(1-x+\frac{x^2}{2}-....) \approx x-x^2 \approx 1-e^{-x} = x-\frac{x^2}{2}+\frac{x^3}{6}...$$

$$\Rightarrow \langle I_{(eu)}\rangle = e^{-q^2\langle u^2\rangle}[CR] + \left(1-e^{-q^2\langle u^2\rangle}\right)\sum_n f^2\alpha_{mn}e^{2\pi i\bar{Q}\cdot(\bar{R}_m-\bar{R}_n)}$$

$$q^2\langle u^2\rangle = (2\pi Q)^2\langle u^2\rangle = 2B\left(\frac{sin\theta}{\lambda}\right)^2 = \frac{BQ^2}{2} = 2M$$

The new term, B, is sensitive to the temperature and vibrational amplitudes. Using B, the intensity can be written as follows:

Thermal diffuse factor increases with Q

$$\Rightarrow \langle I_{(eu)}\rangle = \underbrace{e^{\frac{-BQ^2}{2}}}_{\text{attenuated sharp diffraction peaks}}[CR] + \overbrace{\left(1-e^{-\frac{BQ^2}{2}}\right)}\underbrace{\sum_m\sum_n f^2\alpha_{mn}e^{2\pi i\bar{Q}\cdot(\bar{R}_m-\bar{R}_n)}}_{\text{broadened peaks centered at }\bar{Q}=\bar{r}^*_{hkl}}$$

Since the correlation function is equal to 1 at an atom and only has a considerable effect at a small region around an atom, the diffraction peaks in the second summation are broadened. Recall that a smaller number of unit cells produces broader peaks than a relatively larger amount of unit cells.

Why time average?

Because measurement time for intensity (1 sec) $\gg$phonon lifetime (10^{-14}sec)

Why do odd powers of $\langle u^n\rangle = 0$?

We assume about u = 0

The average time displacement for odd powers averages to zero, but for even powers, the area under the curve is π.This explains why even powers are present in the time average and odd powers are not.

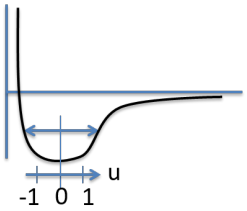


Figure 148.6: ~ Harmonic Oscillation

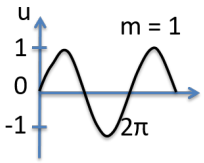


Figure 148.7: $\int_0^{2\pi} sin(mx)dx = 0$

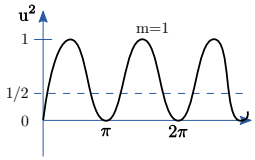


Figure 148.8: Area = $\int_0^{2\pi} sin^2(mx)dx = \pi$

$$Average = \frac{Area}{2\pi} = \frac{1}{2}$$

148.1 Debye Theory of Lattice Dynamics/heat capacity — Corrected by Waller

The quantity B, which describes the vibration of the atoms, is described by the following equation:

$$B = \frac{6h^2T}{m_A k\theta_D^2} Q\left(\frac{\theta_D}{T}\right) = 8\pi^2\langle u^2\rangle$$

m_A =atomic mass = $A \times 1.660 * 10^{-24} g/amu$

h =Planck's constant = $6.628 * 10^{-27} erg - sec$

$T = K$ absolute temp

k =Boltzmann's constant = $1.380 * 10^{-16} erg/K$

θ_D =Debye temperature (related to stiffness of crystal) e.g. $\theta_D^{Ge} = 360K, \theta_D^{Al} = 394K, \theta_D^{Si} = 625K$

$Q(\theta_D/T) \approx 1$ for $T > \theta_D/2$ declines rapidly for $T < \theta_D/2$

$$\text{Unvarying constants: } \frac{6h^2}{m_A k} = 1.151 \times 10^{-12} cm^2 = 1.151 \times 10^4 \mathring{A}^2/A$$
$$B[\mathring{A}^2] = 1.151 \times 10^4 \frac{T[K]}{A[amu]\theta_D^2[K]} Q\left(\frac{\theta_D}{T}\right)$$

Therefore, the mean square vibrational amplitude (assume Q=1) is given by the following:

$$\langle u^2 \rangle = \frac{B}{8\pi^2} = 145.8 \frac{T}{A\theta_D^2}[\mathring{A}^2]$$

- For $Si^{28.1}(\langle u^2 \rangle)^{\frac{1}{2}} = 0.063\mathring{A}$ at R.T. $(T = 300K)$
- For $Ge^{72.6}(\langle u^2 \rangle)^{\frac{1}{2}} = 0.068\mathring{A}$ at R.T. $(T = 300K)$
- For $Al^{27}(\langle u^2 \rangle)^{\frac{1}{2}} = 0.10\mathring{A}$ at R.T. $(T = 300K)$

Debye-Waller Factor $= e^{-2M} = e^{-BQ/2} = e^{-4\pi^2 \langle u_{hkl}^2 \rangle / d_{hkl}^2}$

If we look at the Debye-Waller factor in terms of real components, $u_{hkl}$and d_{hkl}, we see a comparison between the vibrational displacement and the d-spacing within a crystal.

For silicon, the (220) peak intensity is diminished by 4% as shown below. If we move to greater peaks, such as (880) with a smaller d spacing, the attenuation significantly affects the peak intensity.

- For Si (220), $d_{220} = 1.92\mathring{A}$ $e^{-2M} = e^{-0.0425} = 0.958$ at R.T.
- For Si (880), $d_{880} = 0.48\mathring{A}$ $e^{-2M} = (e^{-0.0425})^{16} = e^{-0.0680} = 0.506$ at R.T.

Note that the (880) d-spacing is 1/4 of the (220) d-spacing. Since the d_{hkl} is squared in the Debye Waller factor, this causes the peak intensity of the (880) to be $\frac{1}{d_{hkl}^2} = \frac{1}{16}$ of the (220) intensity.

$$\frac{e^{-2M_{h_2l_2k_2}}}{e^{-2M_{h_1k_1l_1}}} = (e^{-2M_{h_1k_1l_1}})^{\overbrace{(h_2^2+k_2^2+l_2^2)}^{\text{if cubic}}/(h_1^2+k_1^2+l_2^2)} = (0.0958)^{16} = 0.503$$

$$\langle I_{eu} \rangle = e^{-BQ^2/2}[CR] + \underbrace{(1 - e^{-BQ^2/2}) \sum_m \sum_n f^2 \alpha_{mn} e^{2\pi i \bar{Q} \cdot (\bar{R}_m - \bar{R}_n)}}$$

Note : Original Debye Theory α_{mn} ={ 1, m=n; 0, otherwise $= (1 - e^{-BQ^2/2}) f^2 M$

If we only account for the 3-D interference function and divided by f^2, then the intensity will always have an intensity of M^2 , as shown by the dashed curves in Figure. However, the Debye-Waller factor attenuates the peaks in an exponential fashion with an increasing Q, as shown by the three peaks with amplitudes following the $e^{-BQ^2/2}$ curve. Finally, the TDS increases the background radiation and widens the peaks, but this effect is exaggerated in the figure below.

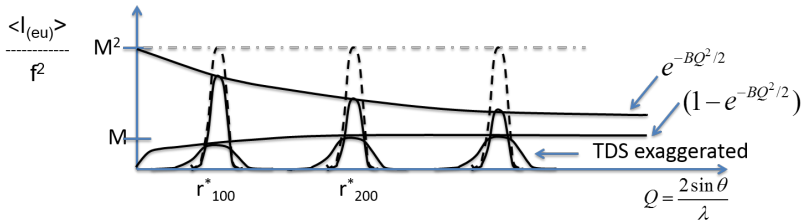


Figure 148.9: Intensity as a function of Q

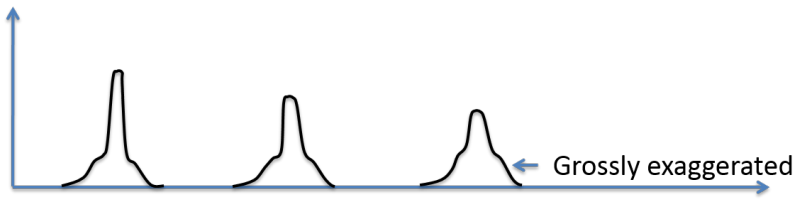


Figure 148.10: Intensity with temperature effects

Symmetrical Reflection Powder Diffraction

$$P' = \frac{I_0 l}{16\pi R} r_e^2 \frac{\lambda^3}{V_c^2} m_{hkl} |F_{hkl}|^2 e^{-2M} (LP) \frac{A}{2\mu} = C(LP) m_{hkl} |F_{hkl}|^2 e^{-2M}$$

The intensities are multiplied by the Debye-Waller factor - $e^{-2M} = e^{-BQ^2/2}$, therefore as Q increases, the peak intensities diminish.
M depends on:

- crystal type (Si, Ge, etc.)
- direction r^*_{hkl}(anistropic crystals) - network of "springs" different in different directions, such as graphite
- d-spacing ~ $1/r^*_{hkl}$
- Temp. T
- Crystal bonding (covalent, molecular, etc.)

"B" can be measured by comparing integrated intensity of hkl to 2h 2k 2l reflection, Measuring B is like measuring bonding properties and temperature effects.

For h k l $P'_1 = C m_{hkl} e^{-BQ_1^2/2} |F_{hkl}|^2 (LP)_1$

For 2h 2k 2l $P'_2 = C m_{hkl} e^{-BQ_2^2/2} |F_{2h2k2l}|^2 (LP)_2$

$\frac{P'_1}{P'_2} = \frac{(LP)_1}{(LP)_2} \frac{|F_{hkl}|^2}{|F_{2h2k2l}|^2} e^{2BQ_1^2/2}$ **Since** $Q_2 = 2Q_1$ $Q_2^2 = 4Q_1^2$
All of the above are known, except B
$B = 8\pi^2 \langle u^2 \rangle$ where, u is the vibrational amplitude around a given atom.

149 Thin Film X-ray Scattering

Recall the 1-D interference function:
$q = 2\pi Q = \frac{4\pi sin\theta}{\lambda}$
$\theta = \frac{1}{2}$ of the scattering angle 2θ
An incident wave vector and a scattered wave vector, as well as the scattering vector are illustrated in Figure below.

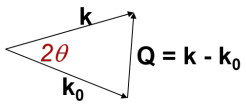


Figure 149.1: Scattering Angle

1-D interference function If there are a series of scattering points that are linearly periodic, then the following relationship describes the scattering intensity from these points.

$$I \propto f^2 \frac{sin^2(Nqd/2)}{sin^2(qd/2)}$$

Note that the function peaks at N^2, as shown in Figure 149.2. The f^2 indicates the scattering pattern by one scatterer. The Bragg peak half width full max is shown in the period of the fringes, which is numerically the period of the numerator, given by $\Delta q = 2\pi/t$.

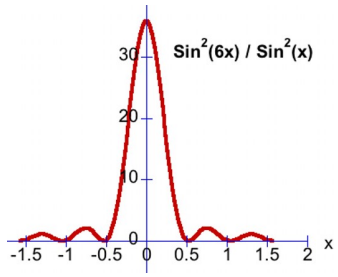


Figure 149.2: 1-D interference

- Peak intensity $\propto N^2$
- Denominator$\rightarrow 0 \rightarrow$Bragg peaks
- Numerator $\rightarrow 0 \rightarrow$Bragg peak width and I
- Crystal size (thickness) = $Nd = t$
- Bragg peak HWFM = Fringe period = period of the numerator = $\Delta q = 2\pi/t$
- Bragg peak area $\propto N$

HRXRD of single crystal Ferro-Electric Ultra-Thin Film In the following experiment, a $PbTiO_3$ film is grown on a $SrTiO_3$substrate, where the two lattices nearly line up and create a structure very similar to a single crystal,

$q = \frac{4\pi sin\theta}{\lambda}$

At 001 Bragg peak:

$c = \frac{2\pi}{q_{peak}}$

$SrTiO_3$cubic:

$a = c = 3.905\mathring{A}$

$PbTiO_3$tetragonal:

$a = 3.905\mathring{A}, c = 4.135\mathring{A}$

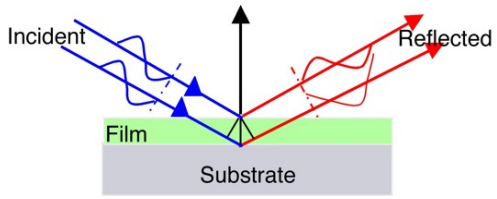


Figure 149.3: Thin film

The following figure shows the reflectivity from this thin film in a $\theta - 2\theta$ in reciprocal space up to to the (001) strontium titanate plane. The strontium titanate is like a semi-infinite crystal, therefore yielding a very narrow. The broad $PbTiO_3$ peak is twice the width of the subsidiary peaks

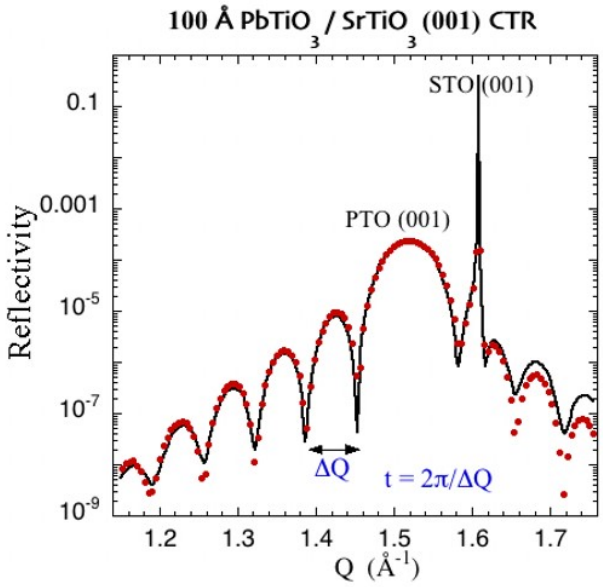


Figure 149.4: Reflectivity of $SrTiO_3$ -$PbTiO_3$

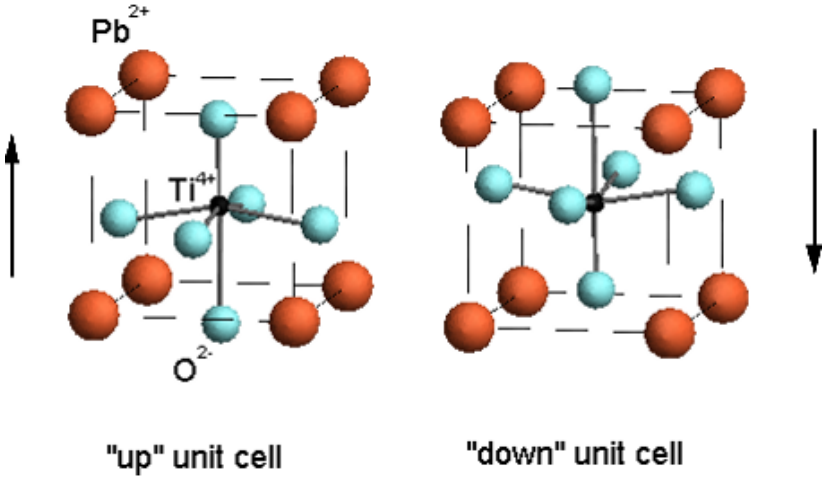


Figure 149.5: $PbTiO_3$Structure

perroelectric and ferroelectric
$PE \Leftrightarrow FE$ transition temperature 490
Tetragonal structure: c/a = 1.063
Metal organic chemical vapor deposition grown films on $SrTiO_3(001)$:
single c-domain at $t \leq 70$ nm
Polarization of as-grown c-domain film?
The structures are comprised of anions and cations, and there is displacement of the positive anions relative to the negative cations. Overall, there is an electric dipole moment in a given unit cell. There are two energy equivalent states, “up” or “down,” as shown in Figure 149.6.

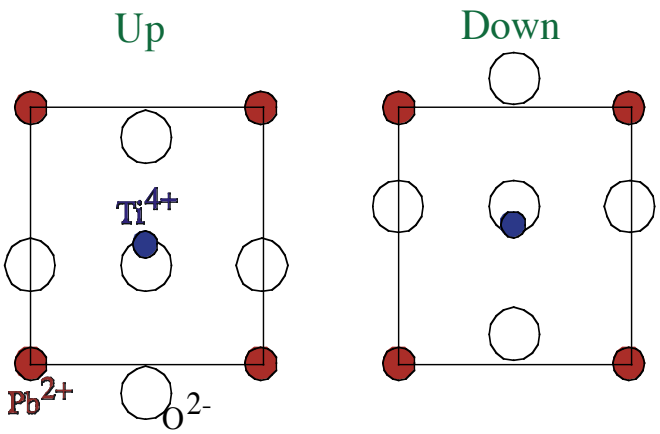


Figure 149.6: $PbTiO_3$ b-axis projection

In determining the structure factor, the exact positions of the atoms must be known.

Friedel's Law If X-ray Absorption is neglected, then $F_{-H} = F_{H}*$

PbTiO3
3.9051 3.9051 4.156 90.0 90.0 90.0
400.
1.0 1.0 1.0
3 1 1 1
Atoms in basis # 1
Pb
82 1.000
1
0.00 0.00 0.1159
Atoms in basis # 2
Ti
22 1.000
1
0.500 0.500 0.5769
Atoms in basis # 3
O
08 1.000
3
0.50 0.00 0.5
0.50 0.5 0.00
0.0 0.50 0.50

Experimental Reflectivity along the (00L) Crystal Truncation Rod (CTR) for ferro-electric PZT thin-film capacitor heterostructure $PZT = Pb(Zr_{0.3}Ti_{0.7})O_3$

In Figure 149.7, the experimentally measured X-ray reflectivity of the PZT capacitor structure is taken at $\lambda = 0.914\mathring{A}$. The inset shows the layers within the heteroepitaxial structure, a $SrTiO_3$(001) substrate, a 136-nm-thick $SruO_3$bottom electrode, a 20-nm-thick PZT film and a 30-nm-thick Ag top electrode. The angle positions for the three distinct (001) Bragg peaks mark off the the three distinct c lattice constants of $SrTiO_3$, $SrRuO_3$,and PZT.

$s_{001} = 3.905, 3.970, 4.130\mathring{A}$

The two patterns of oscillation are due to the thickness of the $SrRuO_3$and PZT films, where the smaller period oscillation belongs to the thicker film. The Ag top electrode has a polycrystalline structure and does not measurably contribute to the scattered intensity in this high-resolution scan.

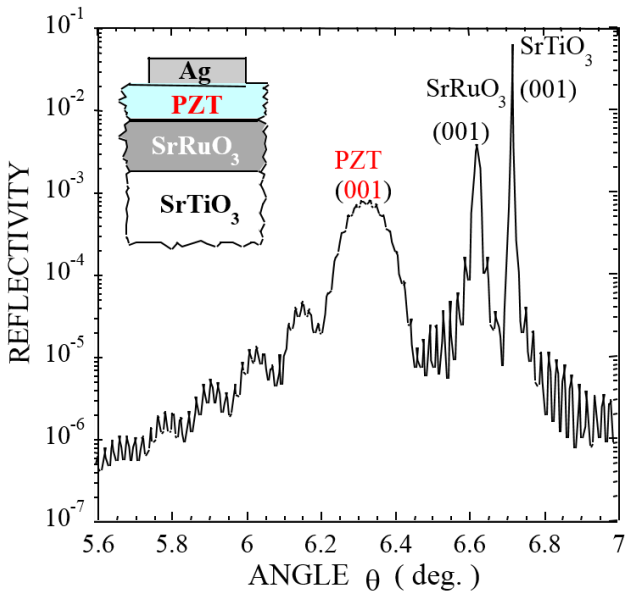


Figure 149.7: Reflectivity of PZT thin film

As predicted by the numerator of the interference function, the widths of the peaks and their Laue fringes are inversely proportional to the film thickness.

$$t = 2\pi/\Delta q = \lambda/\underbrace{(2\Delta_\theta)}_{\hookrightarrow radians}$$

What about diffraction from 2D crystals?
In one dimension, there is a set of periodic scatterers and in reciprocal space, there exists a set of parallel periodic planes with a $\frac{1}{a}$ spacing. In two dimensions, there are two constants for the laue condition, yielding a net of parallel rods in reciprocal space as shown in Figure 149.8.

Crystal D	Real-Space Basis	Laue Condition	Recip. Space Periodicity
1D	a	$ha* = const.$	$\|$planes, $\perp a$
2D	a_1, a_2	$h\bar{a}_1^* = const.\ k\bar{a}_2^* = const.$	$\|$rods, $\perp a_1 a_2$ plane
3D	a_1, a_2, a_3	$h\bar{a}_1^* = const\ k\bar{a}_2^* = const.\ l\bar{a}_3^* = const.$	points

Table 149.1: Real Space ⇔Reciprocal Space

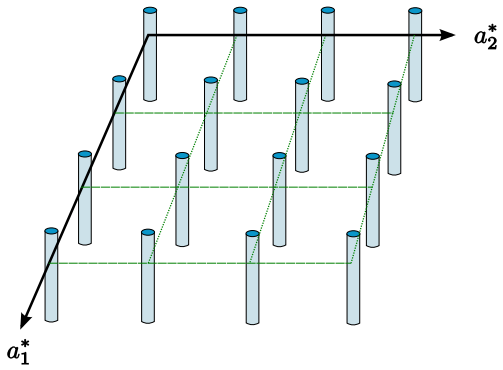


Figure 149.8: 2D Reciprocal Space Periodicity

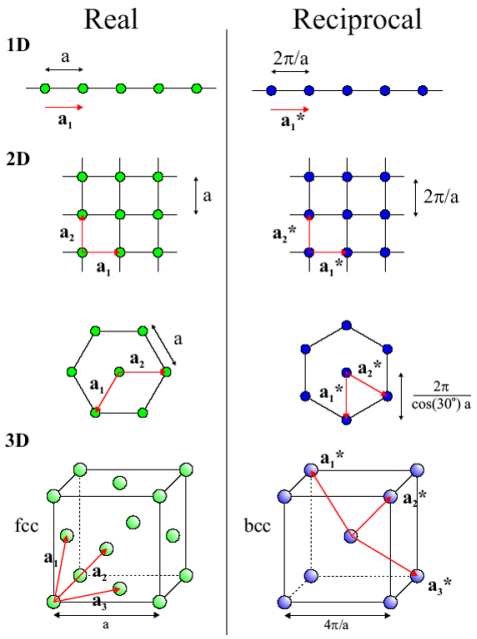


Figure 149.9: Real ⇔Reciprocal Space (From :[ANM])

In two dimensions, there is a rod going through the origin, and the incident beam is perpendicular to the plane of the two dimensional crystal. The radius is $\frac{1}{\lambda}$.

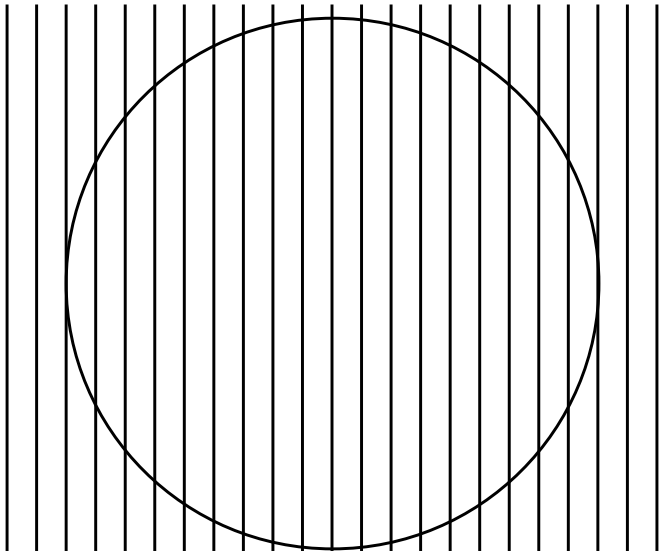


Figure 149.10: Ewald reciprocal-space diagram for diffraction from a 2D square lattice with the incident beam k_0 perpendicular to the ab-plane

Figure 149.10 illustrates an example of low energy electron diffraction. It is the k=0 plane cut with a ~ 13λ.The 2D periodicity in real space leads to rods in reciprocal space. The 2D lattice is at C (the center of the Ewald sphere), and k_0 extends from C to (00).

Diffracted wave vectors k_{hk} extend from C to the intersection point of the (hk) rod and the Ewald sphere. Note that each rod can produce a forward diffracted beam and a back reflected beam, due to two intersections with the Ewald sphere. This diagram also portrays diffraction from a 1D periodic grating, in which case the (h0) rods become (h) planes seen edge-on in k-space.

Low-Energy Electron Diffraction

For a low energy diffraction instrument, the accelerating voltage is controlled.

Recall the de Broglie relation for X-rays: $\lambda = \frac{hc}{E}$, where $hc = 12.4keV \cdot A$

More generally, $\lambda = h/p$, where momentum $p = mv$ and $p = Ec$ for a photon since $m = E/c^2$
For nonrelativistic massive particles, $\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2m(\frac{1}{2}mv^2)}}$
For electros: $\lambda(\AA) = \frac{12.3}{\sqrt{V(volts)}} \rightarrow \lambda = 1\AA$ for 150V (electron beam with a wavelength on one angstrom requires 150 V)
$\lambda = 1\AA$ for neturons
$KE_n = \frac{150eV}{m_n/m_e} = \frac{150eV}{1839} = 82meV$ Greater kinetic energy required for neutron of greater mass relative to electrons.
Thermal neutrons kT=26meV at RT. Neutron diffraction has a different sensitivity to elements than x-rays.
Filament in vacuum, negative potential relative to sample at ground, concentric spheres are grids that, fluorescent screen gives spot where electron diffracted beam hitts

From 1D Laue Condition:

h^{th} diffraction peak at $Q = ha*$

$\rightarrow Qa = h$

$\rightarrow h\lambda = 2asin\theta_h$

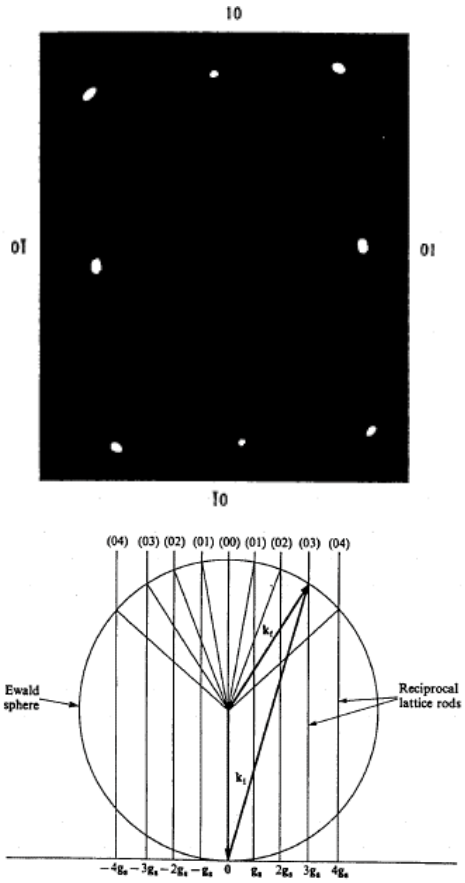


Figure 149.11: LEED pattern from Cu(110) surface

Wavelength gives radius, spacing is $\frac{1}{a}$ for copper.
Why is electron diffraction surface sensitive, in comparison to x-ray diffraction?
Electron beam interacts very strongly with matter, showing scattering from top one or two layers whereas x-rays interact very weakly.
Universal mean free path for electrons based on inelastic scattering (plot vs KE)
To achieve surface sensitivity with x-rays we must use total external reflaection which produces and evanescent wave with nm penetration depth.

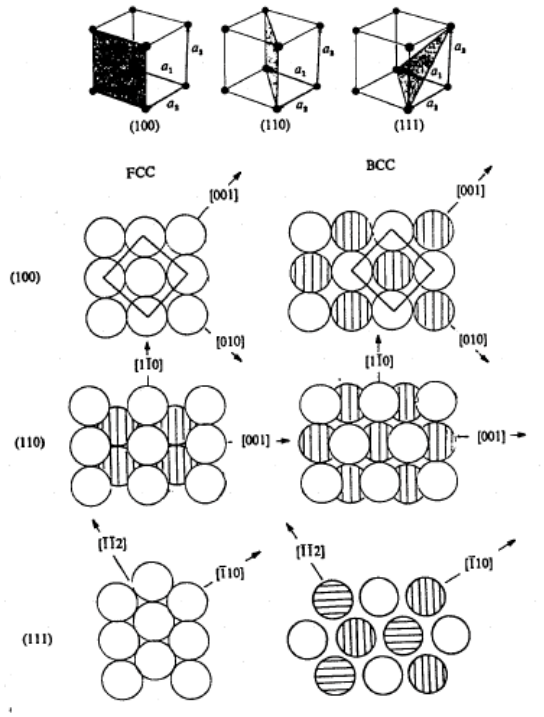


Figure 149.12: Low-index ideal surfaces of hard-sphere cubic-F & cubic-I crystals

4-Bromostyrene SAM on Si (111) surface.
R(q_z) calculated for 1/2 ML coverage.
A z-axis projection can be performed for the structure.

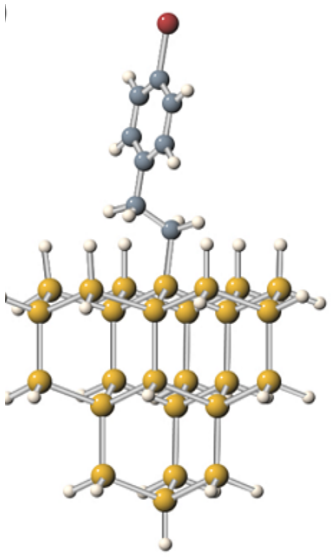


Figure 149.13

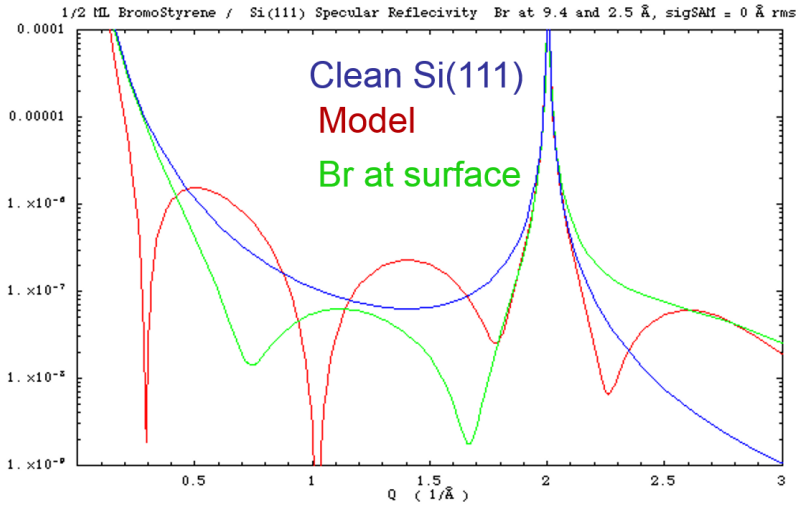


Figure 149.14

Substrate Reflectivity:
$R_s = |r_s(q)|^2$

like 1D interference function. The period of the subsidiary peaks has gone to zero since N is so large.

$$r_s(q) = \frac{4\pi i r_e F_s(q)}{AQ(1 - e^{-iqc})}$$

$A =$ Area of 1x1 u.c. =12.8$\AA^2$

$$F_s(q) = 2f_{Si}cos(qc/8),$$

$$c = d_{111} = 3.135\AA$$

$$R_{Tot}(q) = |r_s(q) + r_M(q)^2|^2$$

$$r_M(q) = \frac{4\pi i r_e \Theta F_M(q)}{Aq}$$

Θ =Molecules per 1x1 u.c = 1/2

$$F_M(q) = f_c(q) \sum_{j=1}^{8} e^{iqz_i^C} + f_{Br}(q)e^{iqz^{Br}}$$

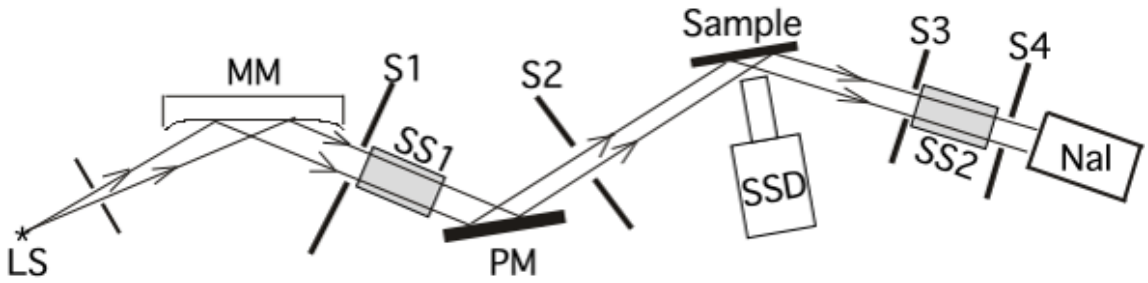


Figure 149.15: Overhead schematic view of 18kW rotating anode setup in Nu X-ray Lab used for XRR, XRF and XSW measurements of SAM-Si (111) samples

There is a line source coming out of a page.

- LS: Vertical Line Source, at focus of parabola.
- Cu Anode (8.04 keV) used for XRR;
- Mo anode (17.44 keV) used for XRF and XSW
- MM: Parabolic multilayer-mirror
- S1-4: slits
- SS1&2: solar slit for 1vertical collimation
- PM: Si (111) postmonochromator to increase angle resolution
- Sample on stepping motor driven $\theta - 2\theta$2-circle, X-translation, and x- tilt
- Nal: pulse counting X-ray detector for reflected intensity on 2θarm
- SSD: Solid-state XRF detector with pulse-height analysis and multichannel analyzer

Total External Reflection of X-rays from a Mirror Surface Start with the zeroth order structure factor

$F_{000} = F_0$ =number of electrons in the unit cell

F_0 is related to the zeroth order Bragg diffraction peak

For x-rays, the index of refraction: $n <\approx 1$, beacuse x-rays are above the natural frequency of the electrons in the sample.

$n = 1 - \delta$ where $\delta = \frac{1}{2}\Gamma F_0$ $\Gamma = \frac{r_e \lambda^2}{\pi V_{u.c.}}$ $r_e = 2.818 \times 10^{-5} \mathring{A}$

N_e= electron density

$\theta_2 = 0 \rightarrow TER \rightarrow \theta_1 = \theta_c$ critical angle

snell's law

$n_1 cos\theta_1 = n_2 cos\theta_2$

therefore $cos\theta_c = n_2/n_1$=n_2 for $n_1 = 1$

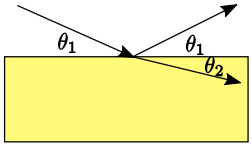


Figure 149.16

Dynamical Theory for Total External Rflection of X-rays form a Mirror Surface

$$n = 1 - \delta - i\beta$$

$$\delta = \frac{N_e r_e \lambda^2}{2\pi} \text{ , } \beta = \frac{\lambda \mu}{4\pi}$$

Fresnel Theory for reflectivity from ideal mirror

$$\frac{E_R}{E_0} = \frac{x - \sqrt{x^2 - 1 - iy}}{x + \sqrt{x^2 - 1 - iy}}$$

where $x = Q/Q_c$ and $y = \beta/\delta$

$$R = \left|\frac{E_R}{E_0}\right|^2$$

Penetration Depth: For $\theta \gg \theta_c \rightarrow \Lambda = sin\theta/\mu$
For $\theta < \theta_C \rightarrow \Lambda$

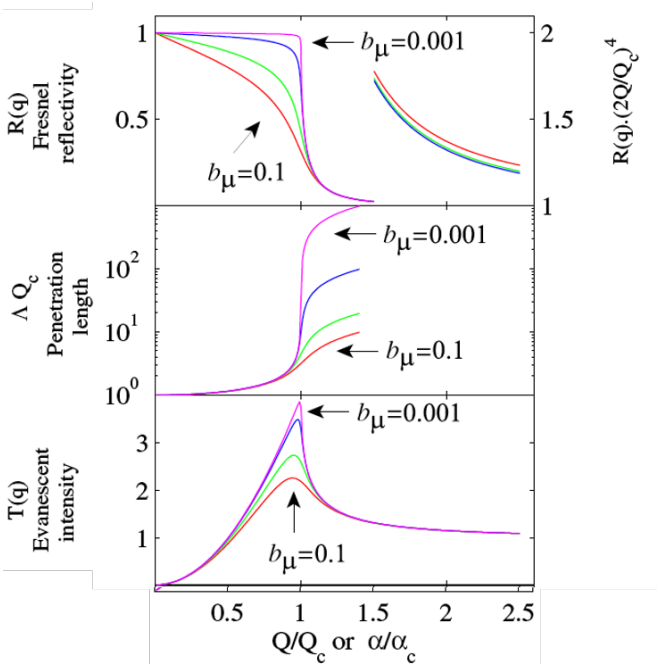


Figure 149.17

This is a Dynamical Theory
X-ray Reflectivity (XRR) Analysis
XRR is very sensitive to surfaces

Example: SAM/Si (111)

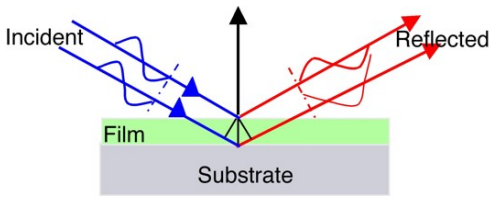


Figure 149.18: SAM/Si (111)

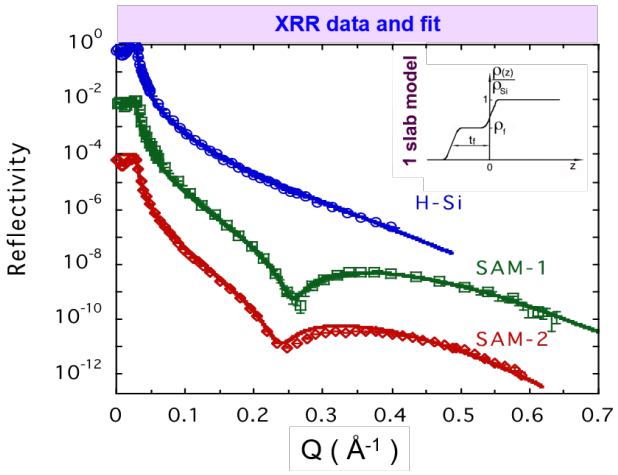


Figure 149.19: XRR data and fit

$$q = 4\pi sin\theta/\lambda$$

Fresnel Theory:

$R \sim q$ for $q < q_c = 0.031 \mathring{A}^{-1}$ Si mirror TER

$R_F = (2q/q_c)^{-4}$ for $q \gg q_c$. Fourier transform of a step function

At 1^{st} dip, the scattered plane-waves from the top and bottom interfaces have a $\lambda/2$ path-length difference (or π phase difference).

Modulation period →film thickness Range: 1 to 100 nm

Modulation Amplitude →relative electron density of film

Modultion damping →roughness of interface(s) Range: < 2nm

X-ray Reflectiviity of UDAME monolayer on Si (111) Substrate First step in DNA covalent attachment
Sample: Reagan Kinser (Prof. Mark Hersam)
UDAME: $CH_2 = CH - C_8H_{16} - COO - CH_3$
$H - Si(111)$
UDAME-Si (111)

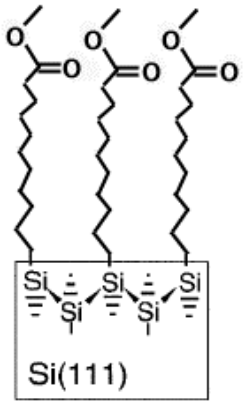


Figure 149.20: Sample

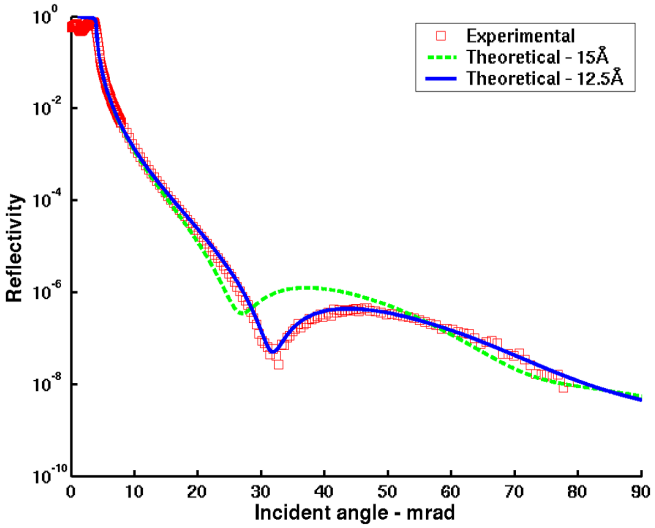


Figure 149.21: Tilt Angle = 33, 8.04keV @NWU X-ray Lab March 2003

	Exper.	Theor.
$N_e(e^-/A^3)$	0.28	0.30(liq)
$t(\mathring{A})$	12.5	15.0
$\sigma_s(\mathring{A})$	1.8	
$\sigma_l(\mathring{A})$	3.8	

$0.28e^-/\mathring{A}^3 = 0.41$ML

Table 149.2

X-ray Reflectivity Analysis Fundamentals Kinematical approach:

$q = 4\pi sin\theta/\lambda$

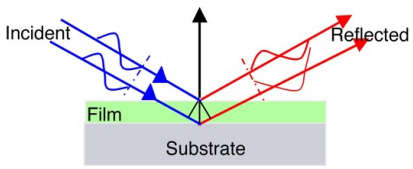


Figure 149.22

$$R(q) = R_F(q)|\Phi(q)|^2$$

$$\Phi(q) = \frac{1}{\rho_\infty}\int \frac{d\rho}{dz}e^{iqz}dz$$

FT of gradient of e^- density profile

$$\frac{R(q)}{R_F(q)} = [1 - 4b(1-b)sin^2(qt/2)]e^{-q^2\sigma^2}$$

For 1 slab model

$\rho(z) = e^-$density profile, $b = \rho_F/\rho_{Si}$, $\sigma = \sigma_s = \sigma_l = (rms)roughness$, t =film thickness

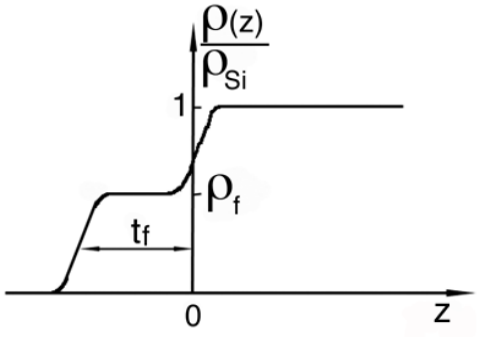


Figure 149.23

b	t(Å)	σ(Å)
0.5,0	12.2	3.2
0.5,3	13.3	3.6

Table 149.3

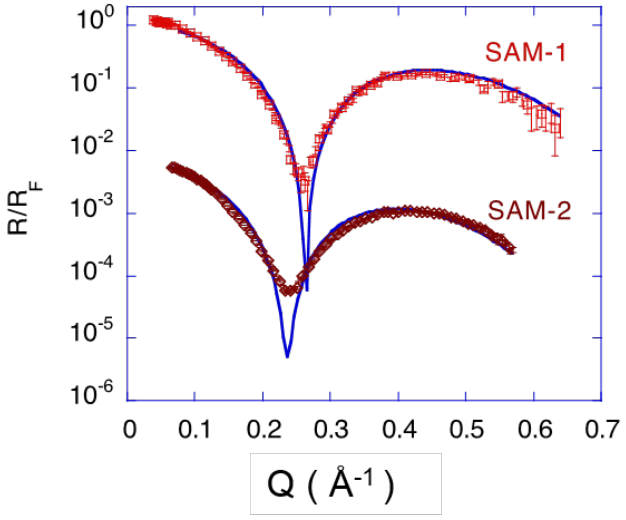


Figure 149.24

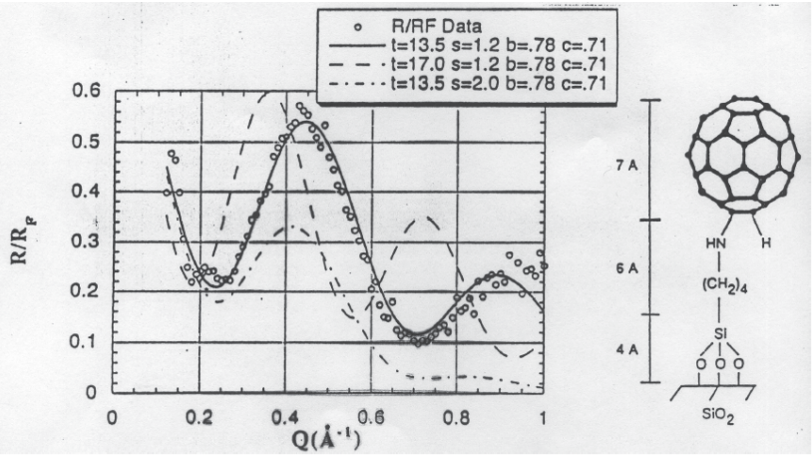


Figure 149.25: XRR Analysis of C_{60}attached to Silica surface; Grown by the Mirkin Group at NU

150 Order-Disorder

150.1 Order-disorder geometry

Consider β Brass CuZn (Cubic Structure)

- **Ordered** from →Cu at 000, Zn at $\frac{1}{2}\frac{1}{2}\frac{1}{2}$, or CsCl structure
  - Diffraction pattern would indicate Cubic P structure
- **Disordered** →some Cu occupy Zn sites, some Zn occupy Cu sites. In a randomly disordered structure, each site has an equal likelihood of Zn or Cu occupation Cubic I.
  - Diffraction pattern would indicate Cubic I structure

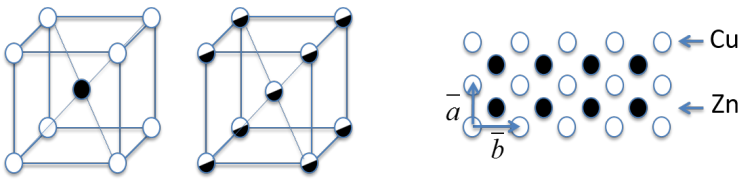


Figure 150.1: Ordered β brass CuZn; Superlattice reflections (100),(300),(111), etc. →allowed by Cubic P and disallowed by Cubic I

The Zn is out of phase, and the zinc and copper do not scatter equally. The absent reflections for the randomly disordered phase are called superlattice reflections, and are sensitive to the order of the crystal.
Fundamental reflections are allowed by both P and I :$h+k+l=2n$
The intensity in the superlattice reflection indicates the degree of order present in the crystal. A high temperature form is generally disordered, while one that is cooling down has ordered domains meeting at domain walls. Zinc would like to have copper nearest neighbors and copper would like to have zinc nearest neighbors, therefore favoring an ordered structure at cooling.

150.2 Long-range order (W.L. Bragg and E.J. Williams)

Consider a binary compound (CuZn, Au_3Cu, etc.) with 2 types of atoms - A and B - and two correct sites in the unit cell - α and β.
There are N_A atoms of type A and sites α and N_B atoms of type B and sites β. Therefore the total number of atoms in the crystal is given by the following:
$$N = N_A + N_B$$

The following are fractions that change with temperature:

- r_α= fraction of rightly occupied α sites
- w_α =fraction of wrongly occupied α sites
- r_β= fraction of rightly occupied β sites
- w_β= fraction of wrongly occupied β sites

The fraction of wrongly occupied sites and rightly occupied sites sum to one:
$r_\alpha + w_\alpha = r_\beta + w_\beta = 1.0$
In the complete order case, $r_\alpha = r_\beta = 1.0$ $w_\alpha = w_\beta = 0$.
For complete or random disorder $\rightarrow r_\alpha = N_A/N = m_A$
In general $N_a w_\alpha = N_B w_\beta \rightarrow m_A w_\alpha = m_B w_\beta$

150.2.1 Long range order parameter

The unitless quantity S is the long range order parameter, ranging from 0 to 1. Complete order has a value of S=1 and random disorder has a value of S=0.
$S = r_\alpha - w_\beta = r_\beta - w_\alpha$
ConsiderAu_3Cu in P cubic cell

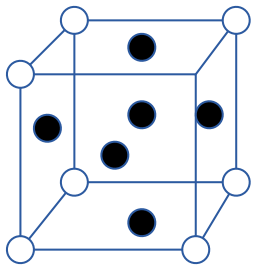


Figure 150.2: P cubic cell

$B \rightarrow$Cu at (000) β site

$A \rightarrow$Au at $\{\frac{1}{2}\frac{1}{2}0, \frac{1}{2}0\frac{1}{2}, 0\frac{1}{2}\frac{1}{2}\}\alpha$sites

$m_A = \frac{3}{4} = \frac{N_A}{N}$ $m_B = \frac{1}{4}$ $2w_\alpha = w_\beta$

Effective Atomic Scattering Factor The ensemble average of the atomic scattering factor from the α and β sites is calculated by the following equations:

$f_\alpha = r_\alpha f_A + w_\alpha f_B$ α **site scattering factor**

$f_\beta = r_\beta f_B + w_\beta f_A$ β **site scattering factor**

For hkl unmixed

$F_{hkl} = f_\beta e^0 + f_\alpha(e^{\pi i(h+k)} + e^{\pi i(h+l)} + e^{\pi i(k+l)})$ - The beta site is at 000 and the alpha sites are face centered.

$= f_\beta + 3f_\alpha$

$= r_\beta f_B + w_\beta f_A + 3r_\alpha f_A + 3w_\alpha f_B = (3r_\alpha + w_\beta)f_A + (r_\beta + 3w_\alpha)f_B$

Simplifying with $3r_\alpha + w_\beta = 3(1 - w_\alpha) + w_\beta = 3 - 3w_\alpha + w_\beta = 3$

yields$F_{hkl} = 3f_A + f_B$ **for hkl unmixed (fundamental reflection) independent of S**

For hkl mixed (superlattice reflections)

$$F_{hkl} = f_\beta e^0 + f_\alpha(e^{\pi i(h+k)} + e^{\pi i(h+l)} + e^{\pi i(k+l)}) = f_\beta + f_\alpha(-1)$$
$$= (r_\beta f_B + w_\beta f_A) - (r_\alpha f_A + w_\alpha f_B)$$
$$= (r_\beta - w_\alpha) f_B - (r_\alpha - w_\beta) f_A$$
$$= (f_B - f_A)S \text{ recall } S = r_\alpha - w_\beta = r_\beta - w_\alpha$$ This is completely sensitive to the long range order parameter.

Therefore, for Au_3Cu

$F_{hkl} = 3f_{Au} + f_{Cu}$ for hkl unmixed and $F_{hkl} = (f_{Au} - f_{Cu})(-S)$ for hkl mixed

Note $f_{Au} > f_{Cu}$, $z_{Au} > z_{Cu}$, $79 > 29$

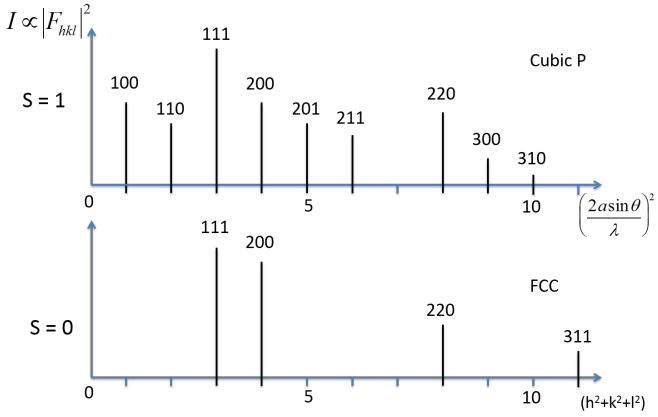


Figure 150.3: Au_3Cu Powder XRD Patterns

Powder diffraction pattern, as S goes from 1 to zero, there is a

Order-disorder phase transitions Measured by comparing a superlattice XRD peak to a fundamental Temperature is normalized by the critical temperature. Measure S by comparing a fundamental diffraction peak intensity to a superlattice diffraction peak intensity.

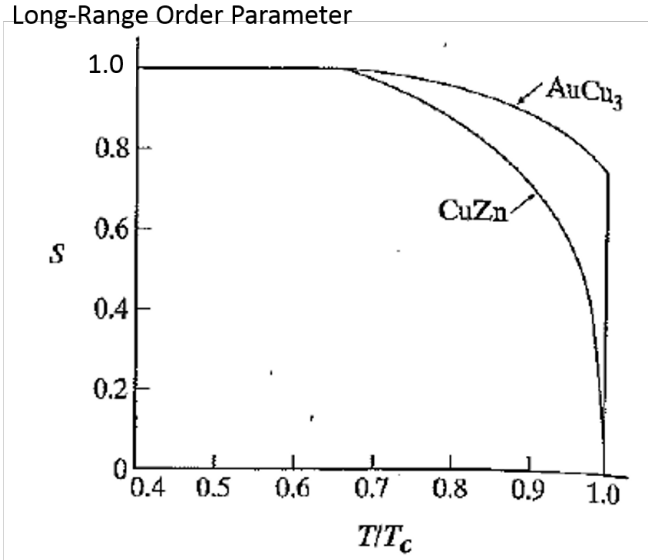


Figure 150.4: Order-disorder phase transitions

1^{st}order phase transition: $AuCu_3$and Au_3Cu both have a discontinuity at T_C , $T_C = 633K$ for $AuCu_3$
2^{nd}order (continuous) phase transition $CuZn$ $T_C = 730$K

151 Special Topics in X-ray Physics: Extended X-ray Absorption Fine Structure (EXAFS)

EXAFS: Why Look at photoelectron scattering?
Unlike XRD, EXAFS does not require a crystalline or semi-crystalline material to determine a material's structure. The probability of dissociating a core photoelectron is related to ionization energy, which is different for every element. This allows for an element-specific structural probe.

151.1 X-ray Absorption of Fine Structure (XAFS): an overview

The incident x ray beam cases a dissociation of core photoelectron. The electron "feels" the surrounding atoms, which yields an absorption spectrum.

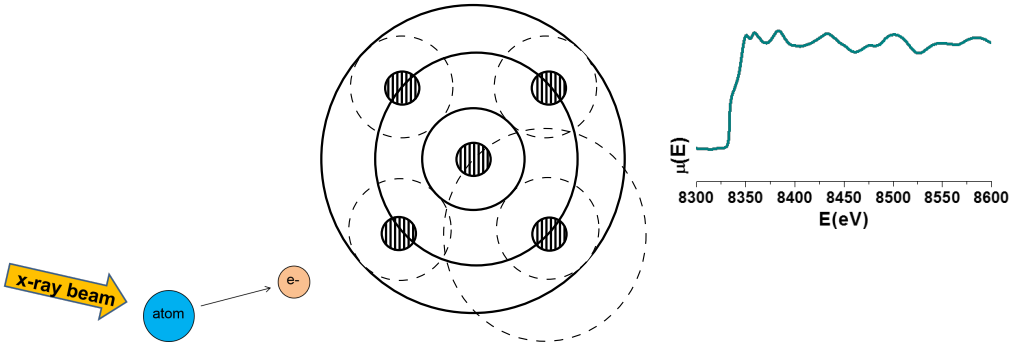


Figure 151.1: XAFS overview

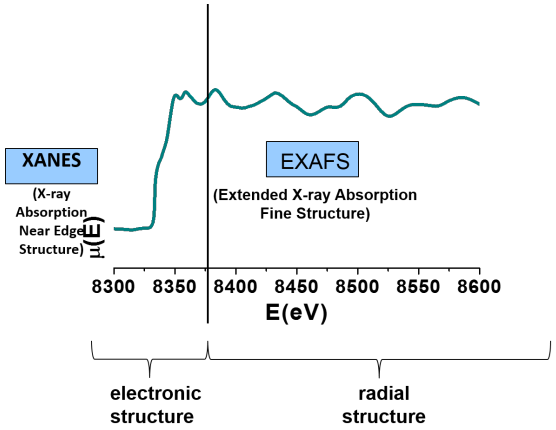


Figure 151.2

X-ray Absorption Spectroscopy proves the perfect tool for studying in-depth nanoparticle structure.
From the lower energy **XANES** region of the spectrum, we can determine the oxidation state, charge transfer, and chemical fingerprint of our element of interest.
From the extended **EXAFS** region, we obtain information about atomic coordination, interatomic spacing, as well as structural and thermal disorder. This ultimately enables us to answer the question: How many of what type of atom are at what distance from my element of interest?

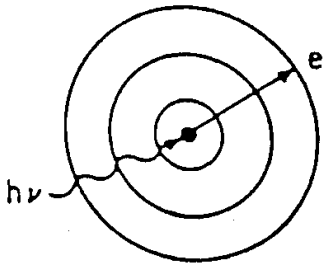


Figure 151.3: Outgoing photoelectron

Outgoing photoelectron: electrons as spherical waves Relationships from classical physics:

$$k = \frac{2\pi}{\lambda} = p/\hbar$$

$$KE = \hbar\omega - E = \frac{p^2}{2m} = \hbar^2 \frac{k^2}{2m}$$

$$k = \frac{\sqrt{2m(\hbar\omega - E)}}{\hbar}, \text{ or } k(\mathring{A}) = 0.521\sqrt{KE(eV)}$$

Due to electron delocalization, we can describe the dissociated core photoelectron as the spherical wave to describe its trajectory

$$\Psi_{out}(\overrightarrow{r}) = \frac{\Psi_0 e^{i\overrightarrow{k}\cdot\overrightarrow{r}}}{r} = \frac{\Psi_0 e^{ikr}}{r}$$

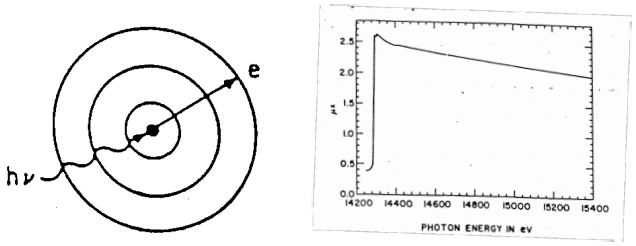


Figure 151.4: Monatomic gas: Kr

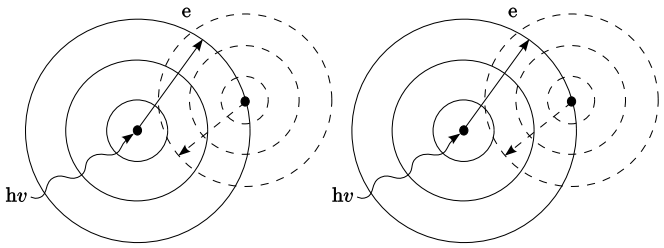


Figure 151.5: Diatomic gas: Br_2

XAFS spectra of mono and diatomic gases How the photoelectron interacts with surrounding atoms

Here, Ψ_{inc} is the wave incident on the second atom, f is the scattered fraction, and $e^{i(k|\overrightarrow{r}-\overrightarrow{R}|+\phi)}$is the new spherical wave centered on the second atom

$$\Psi_{scattered}(\overrightarrow{r}) = \frac{\Psi_{inc.} f e^{i(k|\overrightarrow{r}-\overrightarrow{R}|+\phi)}}{|\overrightarrow{r}-\overrightarrow{R}|}$$

$$\Psi_{inc.} = \Psi_{out.}(\overrightarrow{r} = \overrightarrow{R})$$

$$\Psi_{scattered}(\overrightarrow{r}) = \frac{\Psi_0 e^{ikR}}{R} \frac{f e^{i(k|\overrightarrow{r}-\overrightarrow{R}|+\psi}}{|\overrightarrow{r}-\overrightarrow{R}|}$$

simplification when r=0:

$$\Psi_{scattered}(\overrightarrow{r} = 0) = \frac{\Psi_0}{R^2} f e^{i(2kR+\phi)}$$

How surrounding atoms affect the modulations Nearby atoms can affect the probability of creating a core hole

$$\Psi_{total} = \Psi_0 + \Psi_{scat.} = \Psi_0(1 + \frac{f}{R^2}e^{i(2kR+\phi)})$$

$$I = \Psi\Psi* \propto \underbrace{1 + \frac{2f}{R^2}sin(2kR + \phi)}_{\textbf{interference between incoming and outgoing waves}} + ...$$

Recall from earlier in the course: $I = |E|^2$

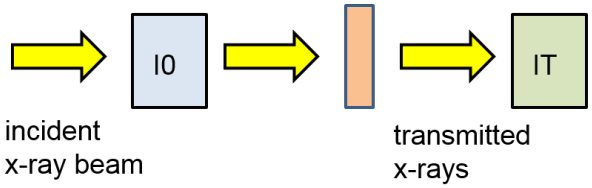


Figure 151.6: Transmission

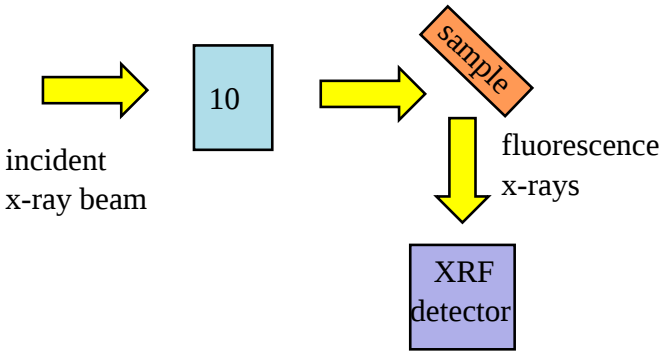


Figure 151.7: Fluorescence

How XAFS date is measured: transmission and fluorescence detection Transmission $\mu \propto -ln(\frac{I_T}{I_0})$ and Fluorescence $\mu \propto (\frac{I_F}{I_0})$ ↷

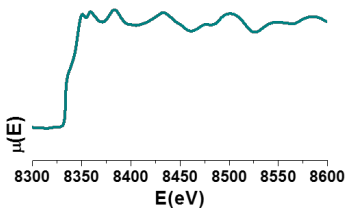
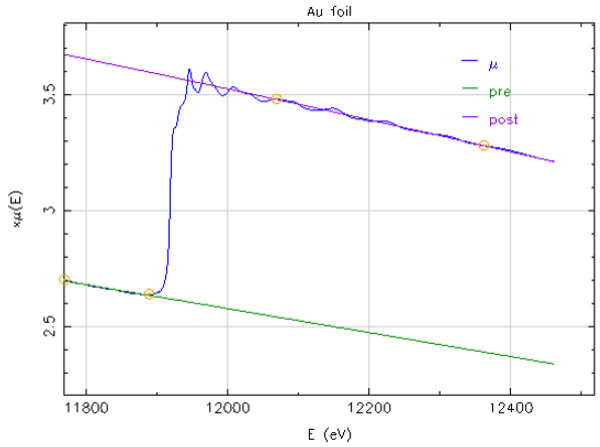


Figure 151.8: Absorption Spectrum?



$\mu \rightarrow \gamma_0$, pre$\rightarrow \gamma$, post $\rightarrow \gamma*$

Figure 151.9: Subtracting the background: fitting pre-edge and post-edge lines

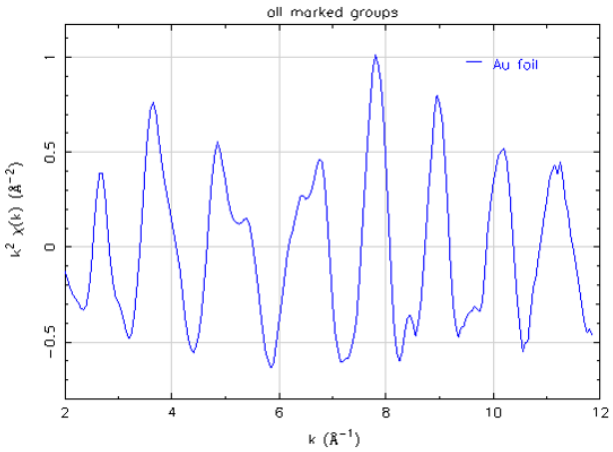


Figure 151.10: Background subtracted data in k-space

Background subtraction: extracting the quasi-periodic modulations

$$\chi(E) = \frac{\gamma - \gamma_0}{(\gamma_0 - \gamma*)}$$

E-Space to K-Space conversion $E(eV) = h\nu - E_0 = \hbar^2 \times \frac{k^2}{2m} \rightarrow k(\mathring{A}^{-1})0.5123(h\nu - E_0)^{1/2}$

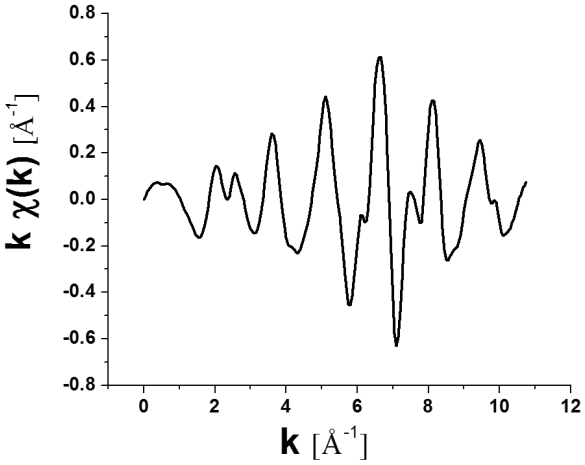


Figure 151.11: K-Space conversion

The EXAFS equation: Describing the k-space data mathematically

$$\chi(k) = \frac{NS_0^2 f(k)}{kR^2} e^{-2k^2\sigma^2} sin[2kR + \delta(k)]$$

N - Degeneracy; S_0^2- Amplitude Reaction Factor

$f(k)$- Scattering Factor; k - Wave number

R- Interatomic distance;σ^2- Mean-squared disorder

$\delta(k)$- Phase shift

What affects the k-space data?
Reduction in CN →reduction in amplitude of k-space oscillations

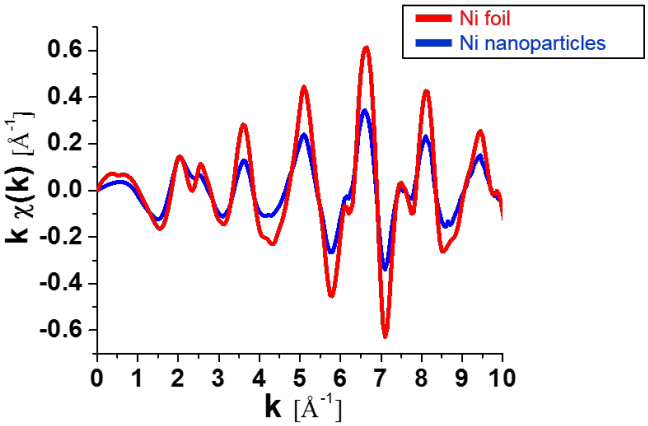


Figure 151.12: Reduction in CN

Change in interatomic spacing →change in period of k-space oscillations

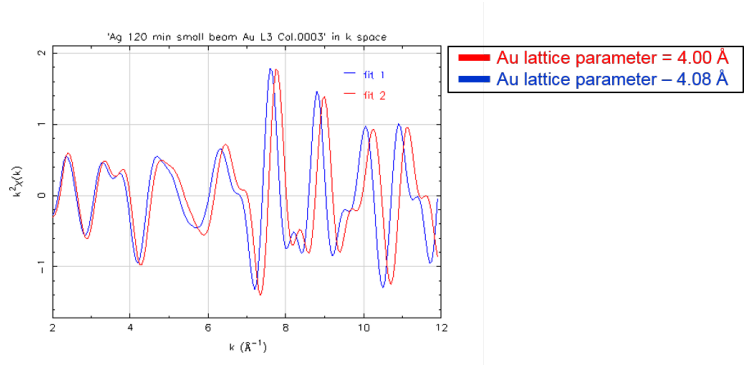


Figure 151.13: Change in interatomic spacing

Fourier transform $k \rightarrow r$ (Fourier transform):

$$\tilde{A}(r) = \left(\frac{1}{\sqrt{\pi}}\right) \int A(k) e^{2ikr} dk$$

$r \rightarrow k$ (inverse Fourier transform):

$$A(k) = \left(\frac{1}{\sqrt{\pi}}\right) \tilde{A}(r') e^{-2ikr'} dr'$$

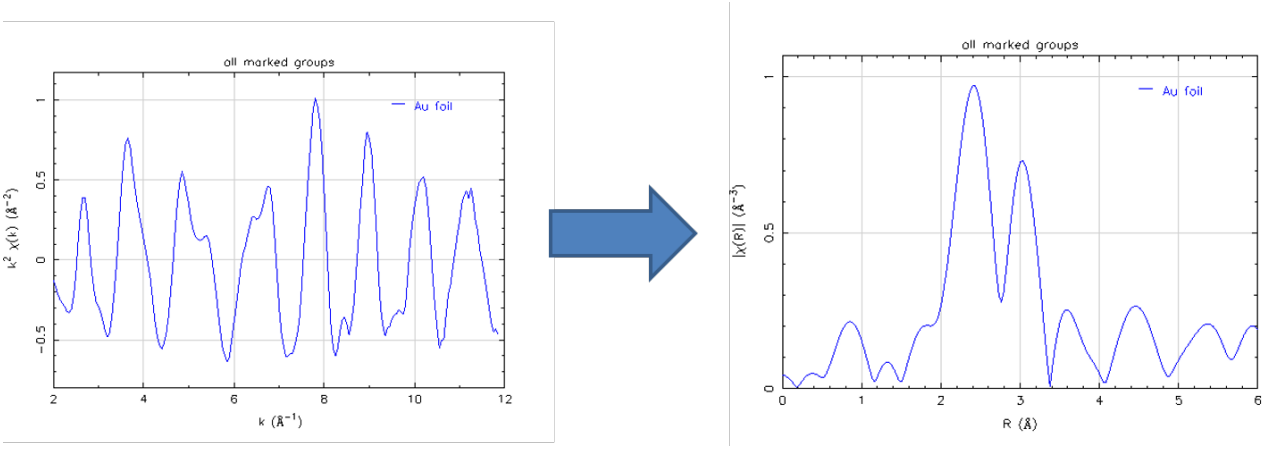


Figure 151.14: Transforming EXAFS data in k-space to EXAFS data in R space

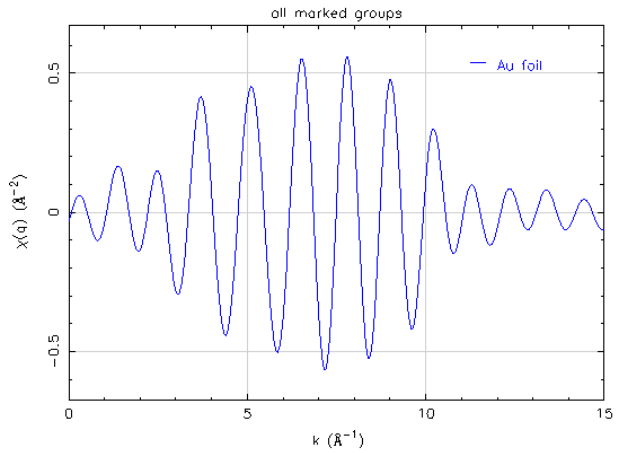


Figure 151.15: Inverse Fourier transform

How do multiple coordination shells affect the spectrum? In the figure below, the Au spectrum with only the first shell is shown in red, and the Au spectrum with the first four shells is shown in blue.

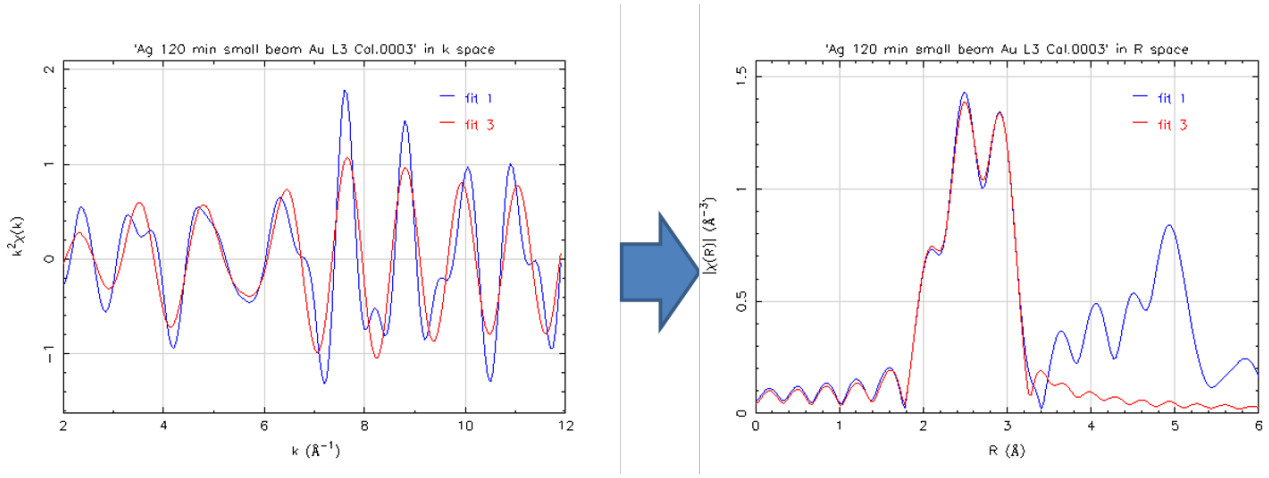


Figure 151.16: Single vs. multiple coordination shells

Recap: Extended X-ray Absorption Fine Structure (EXAFS)

The EXAFS method allows one to determine how many of what type of atom is at what distance from the absorbing atom.

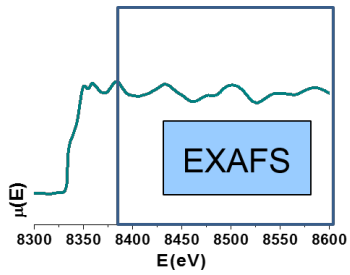


Figure 151.17: EXAFS region

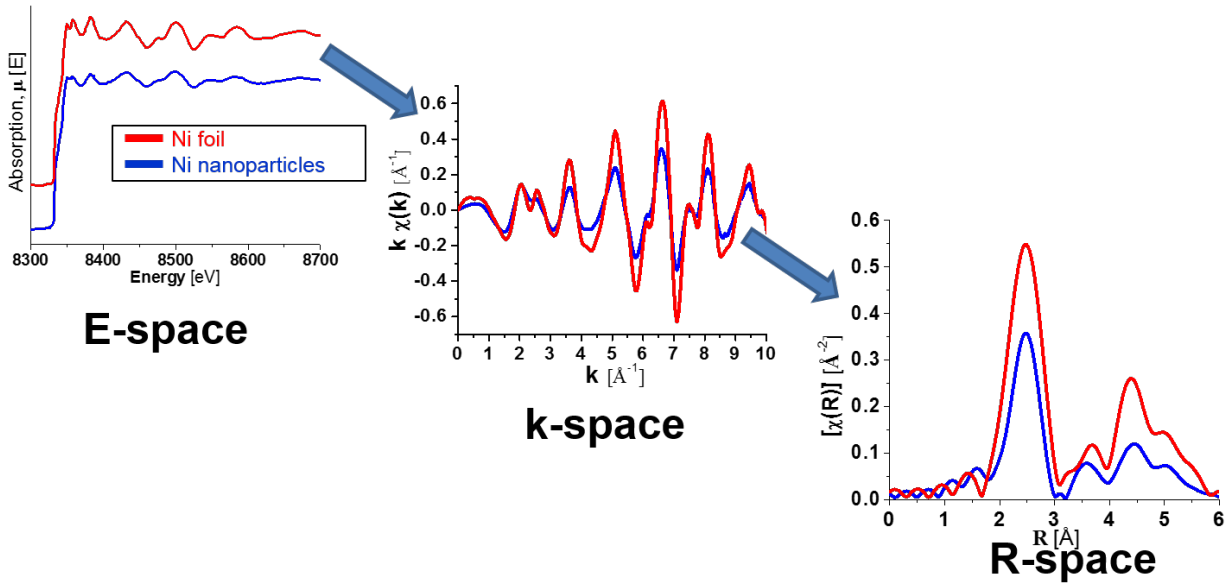


Figure 151.18

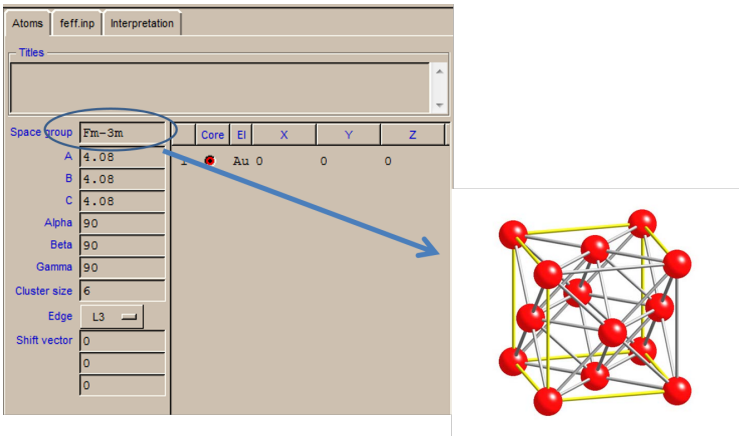


Figure 151.19: $Au(FCC)$, space group $Fm-3m$

The modeling process (crystallography) Pathways generated from crystal structure

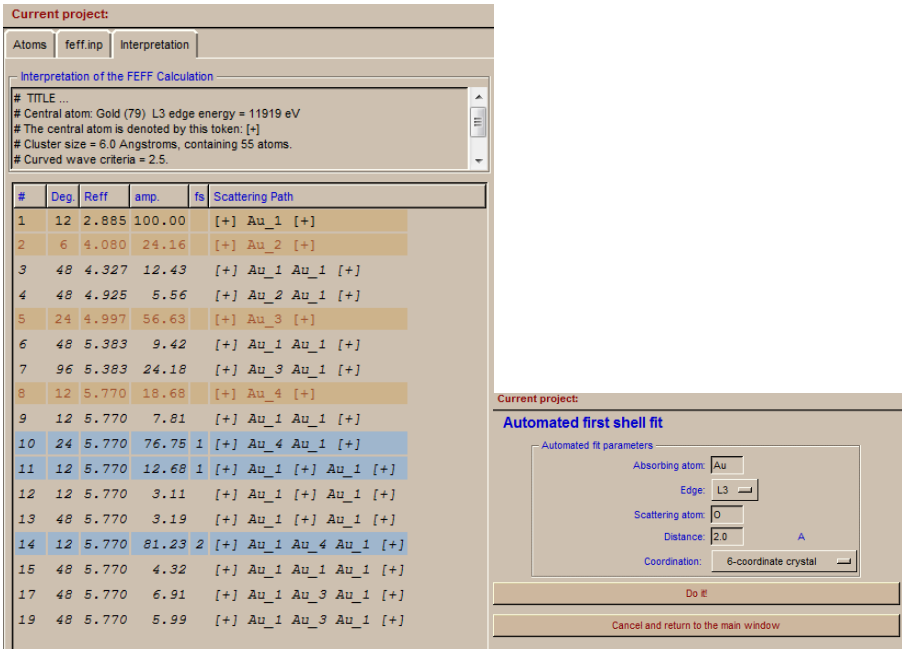


Figure 151.20: Pathways

You can also generate own pathways based on pre-existing knowledge. Each pathway has its own set of variables which must be considered.

The EXAFS equation: how do we treat this in creating a model? Recall:

$$\chi(k) = \frac{NS_0^2 f(k)}{kR^2} e^{-2k^2\sigma^2} sin[2kR + \delta(k)]$$

- Parameters we can determine through theoretical modeling
- Parameters calculated using fitting program, a function of the material
- Parameters which are related to the measurement

Variables defined for each pathway

(1) N is related to the average coordination number

(2) R is defined in fitting as ΔR , or the deviation from the interatomic distance defined in the pathway creation

(3) σ^2 is the mean squared disorder, a convolution of the Debye-Waller factor (thermal disorder) and the radial disorder

Evaluating best-fit model (Statistics)

$$\chi^2 = \frac{N_{idp}}{\epsilon N_{data}} \sum_{min}^{max} [Re(\chi_d(r_i) - \chi_t(r_i))^2 + lm(\chi_d(r_i) - \chi_t(r_i))^2]$$

$$\chi_v^2 = \frac{\chi^2}{v}; v = N_{idp} - N_{var}\ ; \epsilon = \text{measurement uncertainty}$$

$$R = \frac{\sum_{i=min}^{maz} [Re(\chi_d(r_i) - \chi_t(r_i))^2 + lm(\chi_d(r_i) - \chi_t(r_i))^2]}{\sum_{i=min}^{max} [Re(\chi_d(r_i))^2 + lm(\chi_d(r_i))^2]}$$

A "good" model minimises the χ^2 factor and R factor with physically reasonable parameters. There are many "well-fitting" models but only one "correct" model.

Some Examples: Materials Science problems solved by EXAFS EXAFS is particularly useful in systems which involve amorphous phases or dopant atoms, such as amorphous thin gilms, heavy atoms in glass or polymeric materials, nanoparticles, phase-change materials, as well as systems where element-specifc structural details are not available by other methods

(1) EXAFS discovers amorphous phases

Moreau, Ha et. al, 2013. Chem. Mater. 25 (12), 2394.

Moreau, Ha et. al. Nano Letters, 2012, 12, 4530

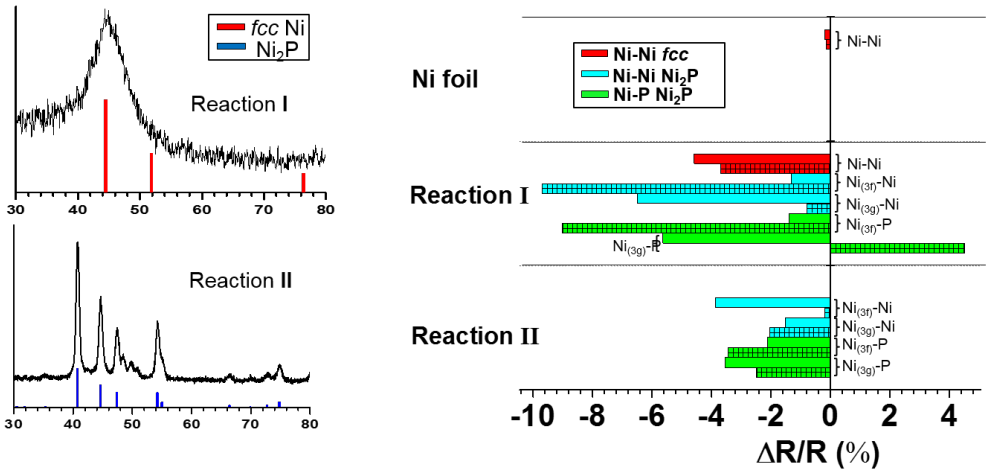


Figure 151.21: Amorphous phases

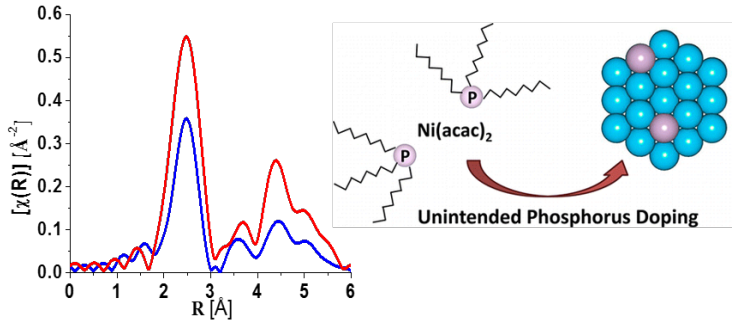


Figure 151.22

(2) EXAFS as a means to study nanoscale transformation

Ha, Moreau et. al. J. Mater. Chem, 2011, 21, 11498.

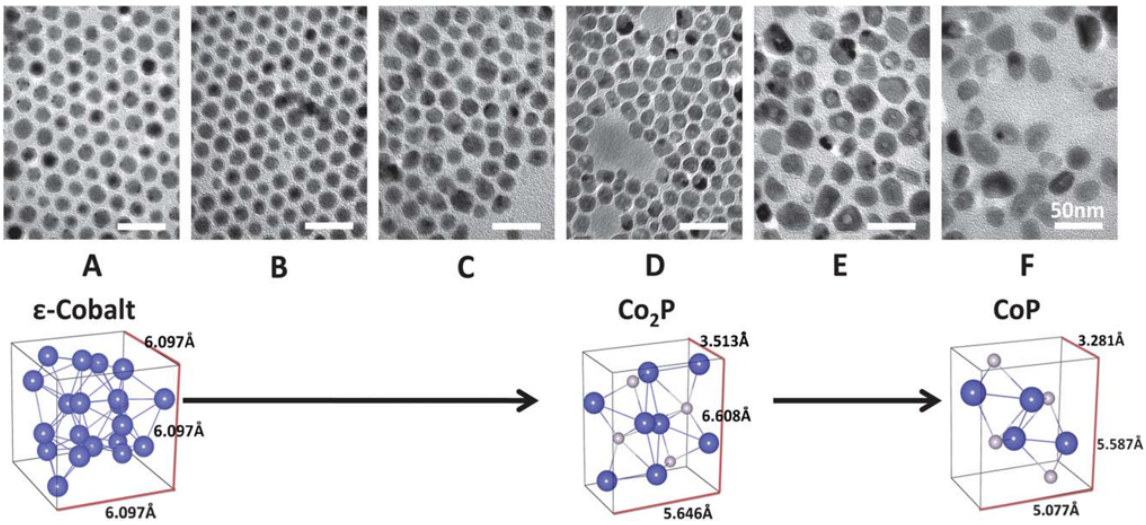


Figure 151.23: Nanoscale transformation

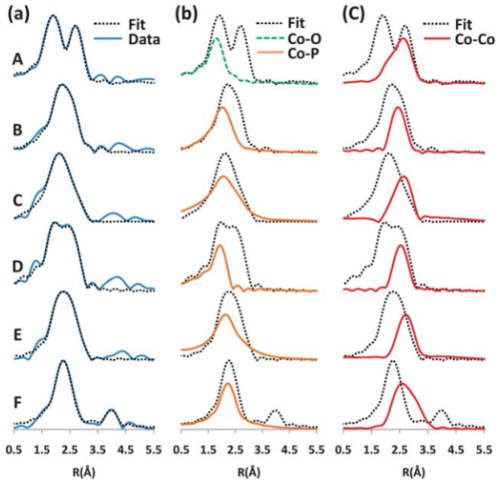


Figure 151.24

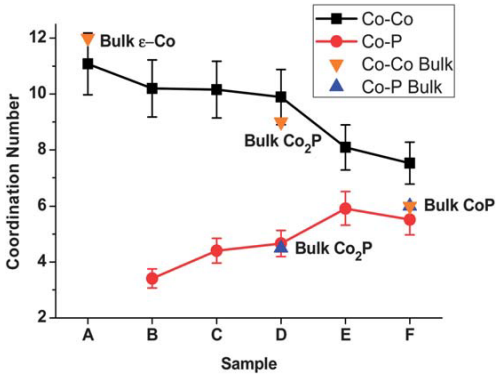


Figure 151.25

(3) EXAFS elucidates nanoparticle growth mechanisms

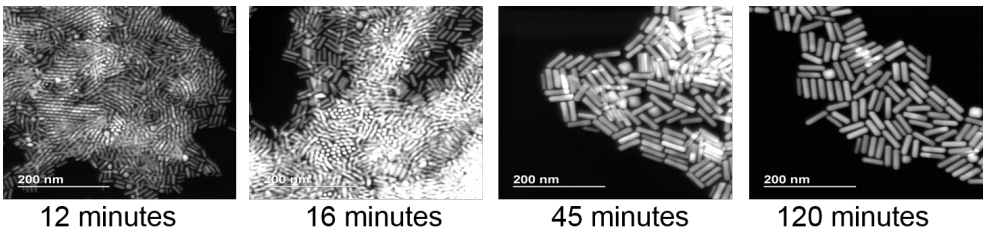


Figure 151.26

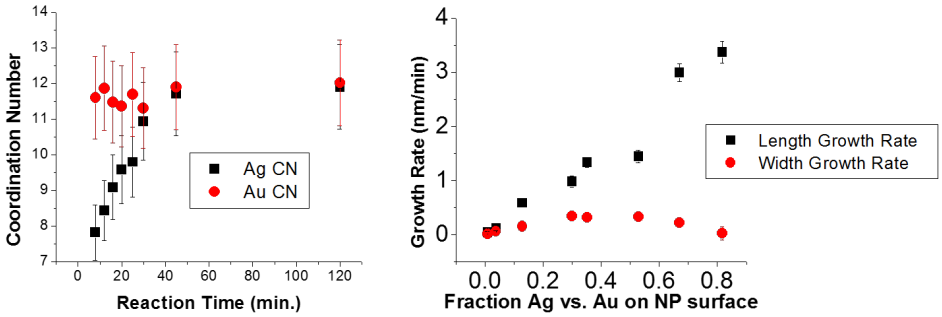


Figure 151.27

From EXAFS results (top) we can determine the coordination number of an average Au atom and Ag atom within the nanorods as a function of reaction time (still time course). We expect that the coordination number will be 12 when the atoms are in the bulk and less than 12 when atoms are on the surface. From the plot we see that although the CN for Au is approximately 12 in all cases, the CN for Ag increases over time until it reaches 12 by ~30 minutes. This reveals that Ag is initially on the nanoparticle surface and over time as deposition of Ag slows, very little to no Ag is on the nanoparticle surface.

Combining EXAFS, XRF and TEM results (bottom) we can relate the growth rate of the length and width of the nanorods (as determined from TEM) to the fraction of Ag vs. Au on the nanoparticle surface (determined from XRF and EXAFS data). We see that the length growth rate of the nanorods is directly correlated with surface Ag. Early on in the reaction (higher fraction Ag vs. Au on the surface), Ag on preferential facets causes the length growth rate to exceed the width and anisotropy is induced. As the reaction progresses and less Ag is on the surface, the length and width growth rates equilibrate and the aspect ratio of the nanorods is maintained.

(4) Probing nanoscale surface structure

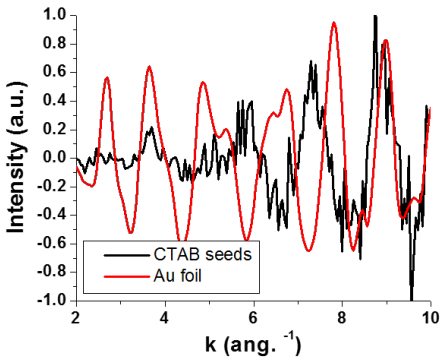


Figure 151.28: Au EXAFS

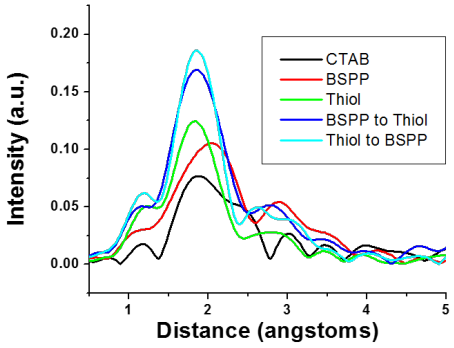


Figure 151.29

2nm particle structure cannot be determined with bulk pathways alone.

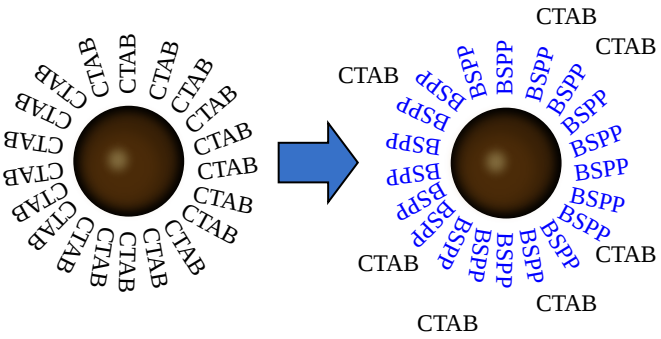


Figure 151.30

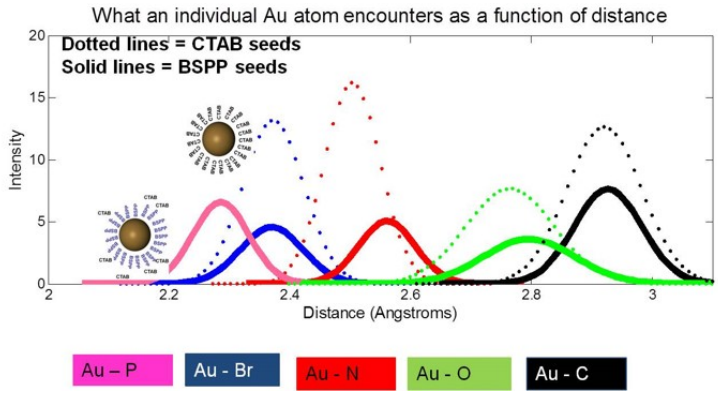


Figure 151.31: What an individual Au atom encounters as a function of distance

152 361 Problems

Homework Policy: You may consult with other class members when preparing your homework answers; indicate clearly on your paper the names of the students with whom you worked. The solutions you submit should, nevertheless, be your own and in your own hand.
Reading Assignments: See the Syllabus, which is published on the NU Blackboard for this course.

(1) By operating on a "7" in the plane of the paper, show that an inversion center is equivalent to a 2-fold rotation axis (perpendicular to the plane of the paper) in two dimensions.

Show that the inversion center and 2-fold rotation axis are not equivalent in three dimensions by doing the same type of operation on a 3D asymmetrical unit " ."

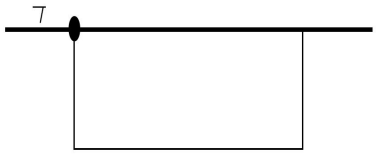
(2) Below are sketched a 2-fold axis parallel to a mirror and two perpendicular mirrors. Operate on the "7" with the indicated symmetry elements and show that a third symmetry element is present in each case and that the two cases are equivalent. Which point group does this pattern of 7's belong to ?

(3) For the following 2D primitive rectangular unit cell:

a Fill in the pattern of 7's based on the indicated unit cell and symmetry elements.

b Draw in the additional symmetry elements that are present.

c Indicate which 2D space group this pattern belongs to.



(4) From Azaroff 1-7

Examine the crystal on page 22 of Week 1 notes (Azaroff Fig. 1-29) and list all the symmetry elements that you find present. Determine the correct crystal class and system.

(5) From Azaroff 1-8

Consider the tetragonal crystal shown on page 21 of Week 1 notes (Azaroff Fig. 1-24). Using the axis labeled a1, a2, and c in that figure, what are the indices of the four faces visible? What is the form?

(6) C&S Prob. 2-6

(7) Hammond Prob. 2-3

(8) C&S Prob. 2-10

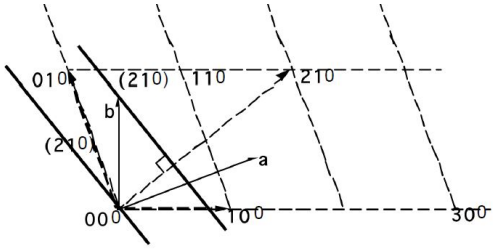
(9) C&S Problem 2-13 (Use tracing paper on top of a Wulff net for this problem, in addition to the following two).

(10) C&S Problem 2-14

(11) C&S Problem 2-15

(12) Prepare a drawing of the *hk*0 layer in the reciprocal lattice of a monoclinic crystal for which a = 2.5 Å, b = 3.0 Å, c = 5.12 Å, and γ=120°. Note that γ^* = 180° - γ = 60°.

(13) Using the figure from lecture notes Week 2 slide 11, draw two lines that represent the edge-on view of the (010), (110) and (210) hkl planes. Draw the corresponding reciprocal lattice vectors, noting that they are perpendicular to the hkl planes. This is already done below for the (210) plane. Using a ruler, measure the interplanar spacing, d_{hkl}, and the length of the reciprocal lattice vector r^*_{hkl} . Verify that the product $d_{hkl} \bullet r^*_{hkl}$ is the same for each set of planes, thus demonstrating that reciprocal lattice vector lengths are inversely proportional to d-spacings in direct space.



(14) Using the general definition of the reciprocal lattice vector (see Week 2 Lecture Notes), derive the relation for d_{hkl} for the tetragonal, orthorhombic, and monoclinic systems in terms of the unit cell parameters a, b, c, α, β, γ and the Miller indices h, k, l. Check answer with C&S A3-1.

(15) (a) Determine the lattice constant "a" in terms of R for a BCC (body centered cubic) structure with spheres of radius R at each lattice point packed as closely as possible. (b) repeat part (a) for an FCC lattice.

(16) GaAs is a zinc blende structure, which has Ga atoms at FCC positions and As atoms at FCC + (1/4, 1/4, 1/4) positions. Write out the 4 coordinates for Ga and the 4 coordinates for As for a unit cell. Make a sketch of this unit cell with the 8 atoms or use Crystal Maker.

(17) Hammond Prob. 5.6. Plus use MatLab to perform the matrix transformations between cubic-I (HKL) and rhombohedral (hkl) for HKL = {100, 111, 110} and hkl = {100, 111, 110}.

(18) C&S Problem 2-7

(19) C&S Problem 2-8

(20) Hammond Exercise 5.2

(21) One formula weight of NiO equals 74.69 amu, and the density of a NiO crystal is 6.67 g/cm^3. (a) Calculate the length of the unit cell edge, a. (b) Using λ = 1.54 Å prepare a list of the first 6 allowed reflections (hkl) and the corresponding Bragg angles, θ_{hkl}. Refer to CS Appendix 9 which lists the allowed reflections for various Bravais direct-space lattices.

(22) Prepare a 3-d drawing of the reciprocal lattice for a body-centered-cubic direct space lattice. Show and label all reciprocal lattice points for $0 \leq$ h,k,l ≤ 2. Show and label the reciprocal space axes, **a***, **b***, and **c***. Identify the Bravais lattice type for this reciprocal space lattice.

(23) a) What are the (hkl) Rhombohedral unit cell Miller indices for the three nonparallel faces of the Rhombohedral unit cell?

b) Read Hammond Section 5.9, make sure you understand how the transformation matrices were generated for trigonal crystals with rhombohedral lattices. Using the transformation matrix derived in Sec. 5.9, what are the (HKL) Hexagonal unit cell Miller indices for the three nonparallel faces of the Rhombohedral unit cell?

c) Now list these as (hkil) Miller- Bravais indices.

d) What is the [uvw] direction in the Rhombohedral unit cell for the hexagonal c-axis direction [001]?

EXTRA CREDIT: Hematite α-Fe2O3 has a rhombohedral crystal structure with a=5.43Å and $\alpha = 55.3$. What is the angle between the hexagonal c-axis and the rhombohedral c-axis? What is the angle between the hexagonal c-face and the rhombohedral c-face?

(24) Find Azaroff Chapters 6 and 8 on BlackBoard, Read Azaroff Ch 6. pp 86-113, an Azaroff Ch. 8 pp. 155-165

(25) [Azaroff 6-1]

(26) Refer to Azaroff Fig. 6-2 on page 88. This experiment is with continuous radiation.

a The x ray beam from the x-ray tube is traveling along y. Is this x ray beam:

i. preferentially polarized along x ?
ii. preferentially polarized along y ?

iii. preferentially polarized along z ?

iv. not preferentially polarized ? Explain your answer.

b Is radiation scattered by the block stronger in detector 1 (along x) or in detector 2 (along z)? Again explain your answer.

(27) [Azaroff 6-2]

(28) [Azaroff 6-3]

(29) [Azaroff 8-1] **Hint** : According to the Bohr model, the radius rn of the nth orbit is proportional to n^2/Z, where Z is the atomic number. Plot the scattering-factor curves f_K,f_L and $f$as functions of kr_K.

$$f = 2\frac{sinkr_K}{kr_K} + 8\frac{sinkr_L}{kr_L}$$

Note that there is a typo in the text book. "a missing factor of 8".

(30) [Azaroff 8-2] (see Cullity & Stock Appendix 10)

(31) [Azaroff 8-9]

(32) The one dimensional interference function is:

$$\frac{sin^2\pi Q \cdot aM}{sin^2\pi Q \cdot a} = \frac{sin^2\beta M}{sin^2\beta}$$

$\mathbf{Q} = (\mathbf{S} - \mathbf{S}_o)/\lambda$ is the scattering vector. $\mathbf{S}$ and $\mathbf{S}_o$are unit vectors in the scattered and incident directions where $\beta = \pi Q \cdot a$ is one-half the phase angle difference (in radians) for rays scattered by planes separated by a. The crystal contains M planes separated by **a**.

a Use a computer with a suitable program or spreadsheet and plotting program (MSE Ugrads must use MatLab) to plot this function over the domain $0 \leq \beta \leq \pi$. Note that this is a periodic function repeated every π radians. Make separate plots of the function for the cases of M = 2, 4, and 16. You will have to appropriately adjust the evaluation interval $\Delta\beta$ to accommodate the sharpness of the function that varies with M.

b For each M above, tabulate the height of the primary maxima, the height of the first subsidiary maxima, and the ratio of the height of the first subsidiary maxima to the height of the primary maxima. This demonstrates how the subsidiary maxima become less significant as M increases.

c By numerical integration determine the area under the β = 0 primary diffraction peak between the first intensity minima (nulls) for each M above. How does this vary with M?

d The integral breadth of a peak is the quotient of the area over the height. This represents the full width of a rectangle that has the same height and area as the peak in question. Calculate the integral breadth of the primary maximum for the three cases, and consider its relation to M.

(33) Consider powder diffraction from CsCl (a = 4.11 Å) with Cu Kα radiation (λ = 1.542 Å). This is a simple cubic structure with Cs^+ at 0,0,0 and Cl^-at 1/2,1/2,1/2. (Note: CsCl is not BCC.)

a What are the *hkl* indices for the first 7 observed reflections from CsCl?

b For the first four of these calculate and tabulate sinϑ/λ ,f_{Cs^+}, f_{Cl^-}, F_{hkl}, 2ϑ, and LP.

c Look up multiplicities m_{hkl} (Cullity & Stock Appendix 11). Combine these with your results from part B to calculate the relative intensities of the four peaks in question. Assign the value of 100 to the most intense peak.

(34) The relation of multiplicity factors m_{hkl} to powder diffraction intensities can be appreciated with the reciprocal lattice.

a Sketch the hk0 plane of the reciprocal lattice for an HCP crystal, using solid dots (•) to indicate nonzero $|F_{hkl}|^2$. Cover the range from -2a* to +2a* and -2b* to +2b*.

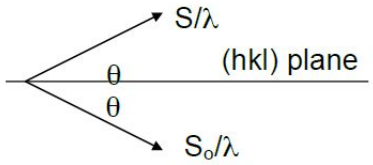


Figure 152.1: $\mathbf{S}/\lambda$ and $\mathbf{S_o}/\breve{}$

b Show that there are 6 reciprocal lattice vectors of the same length $r_{hkl}* = a* = b*$. Give the *hkl* indices for these six vectors. In a powder sample, these will all diffract at the same value of (scalar) Q (or in the same 2θ direction).

c The six planes responsible for this diffraction are of the form {100}, though this is seen more clearly in the 4 axis Bravais-Miller system hkil , where i = -(h+k). Write the indices hkil for these six planes.

(35) Write the structure factor for a crystal with two atoms of the same type per unit cell atx, y, z and $x + 1/2, y + 1/2, z + 1/2$. This conforms to a body centered (I) unit cell with an arbitrarily displaced origin at x,y,z.

Now evaluate the product $FF*$ for all possible values of h, k, and l. You should see that the result is independent of x,y,z. Therefore, the intensity (FF*) depends only on the separation between atoms; absolute positions are not relevant.

Make a perspective sketch of the reciprocal lattice for an orthorhombic I crystal, a = 3Å, b = 4 Å and c = 5 Å. Include 8 reciprocal lattice unit cells (from origin to 2a*, 2b*, 2c*). Indicate nonzero values of FF* by dots (•).

You should notice two things: the "missing" reciprocal lattice points result from the choice of the nonprimitive unit cell; the pattern of nonzero points forms a face centered (F) array in reciprocal space. The same is true for cubic I and tetragonal I Bravais lattices, the reciprocal lattices of which are face centered.

(36) (Cullity & Stock Prob. 4-4) Derive simplified expressions for F and F^2 for diamond, including the rules governing observed reflections. This crystal is FCC and contains 8 atoms per unit cell located at: FCC and FCC + 1/4 1/4 1/4 . Note, silicon and germanium are diamond structure.

(37) Construct a diagram that illustrates the *hkl* Bragg condition, showing $\mathbf{S}/\lambda$, $\mathbf{S}_o/\lambda$ and an (hkl) reflecting plane such that the two unit vectors originate from the same point as shown below. Continue the diagram and show that the vector $\mathbf{Q} = (\mathbf{S}_o/\lambda - \mathbf{S}/\lambda)$ must be perpendicular to the hkl plane. From the construction derive Braggs' Law recalling that **Q** must be the same as a reciprocal lattice vector.

(38) Figure P2 is a c-axis-rotation photograph of an orthorhombic crystal which has **a*** = 0.25 $\AA^{-1}$, **b*** = 0.15 $\AA^{-1}$ and **c*** = 0.36 $\AA^{-1}$. The radius of the cylindrical camera used for taking this photograph is 50 mm and the x-ray wavelength is 1.542 $\AA$. Construct the appropriate reciprocal lattice nets (l=0 and l=1) to index all the reflections shown. What is the lattice type of this crystal?

(39) Use the Scherrer equation $(\delta(2\theta) = 2\varepsilon_{1/2} = 0.94\lambda/Dcos\theta)$ to evaluate the FWHM peak width $\delta(2\theta)$ (in degrees) for crystallite sizes D = 25, 75, 125, 450, and 900 Å. Assume that λ = 1.542 Å and that the peak in this $\theta - 2\theta$ scan whose width we are considering occurs at $2\theta = 80$.

The ability to measure small peak widths is hampered by instrumental broadening. Even if the incident and diffracted beams were perfectly collimated, the peaks would still be broadened by the$K_{\alpha 1} - K_{\alpha 2}$ splitting. Differentiate Braggs law to obtain an expression for $d\theta$ as a function of θ, λ, and dλ. Using this expression, calculate dθ for the $K_{\alpha 1} - K_{\alpha 2}$ splitting of Cu radiation at $2\theta = 90$.

$\lambda(K_{\alpha 1}) = 1.5405\AA$.

$\lambda(K_{\alpha 2}) = 1.5433\AA$

From this you should appreciate that it is difficult to measure D values greater than about 500 $\AA$. when using the Cu K_α doublet as a radiation source. Note that the wavelength resolution $\lambda/d\lambda$ can be improved by an order of magnitude by using a single crystal monochromator to select out the Cu $K_{\alpha 1}$, whose width is approximately 2 eV in comparison to the 20 eV energy separation between the $K_{\alpha 1}$ and $K_{\alpha 2}$.

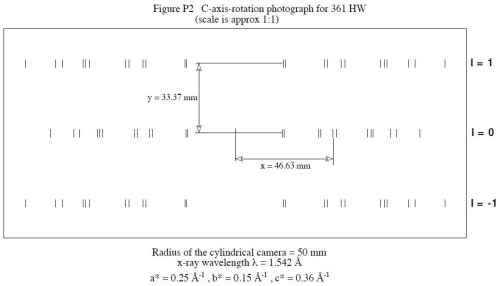
(40) Identical powder diffraction scans are made of a diamond powder sample and a germanium powder sample with Cu K_α radiation. Each has "diamond cubic" crystal structure with 8 atoms per unit cell. Other pertinent information for this problem is as follows:

Sample	Z	a[Å]	μ/ρ $[cm^2/g]$	ρ [g/cm^3]
Carbon	6	3.5671	4.22	3.51
Germanium	32	5.6577	67.9	5.91

The scans were done in symmetric reflection. One might think that the ratio of observed intensities would be given by the ratio of the quantities

$$K_x = \frac{|F_{hkl}|^2 m_{hkl} LP}{V_{C^2}}$$

where x is C or Ge. However, the observed intensity of diamond, relative to that of germanium, was 25 times larger than that calculated from the ratio K_C/K_{Ge}. Explain where this factor of 25 comes from.



Part XI
390: Materials Design

Gregory B. Olson
Department of Materials Science and Engineering
Northwestern University

153 Catalog Description

Analysis and control of microstructures. Quantitative process/structure/property/ performance relations, with case studies. Computer lab for modeling multicomponent thermodynamics and transformation kinetics. Prerequisites: 315, 316 1,2, or consent of instructor.

154 Course Outcomes

155 390: Materials Design

At the conclusion of the course students will be able to:

(1) Apply a systems approach to complex problem identification and formulation.

(2) Apply property/performance relations to materials selection, and to the specification of property objectives to meet user performance needs in evolving environments.

(3) Apply mechanistic models in design integration of process/structure and structure/property relations in dynamic multilevel-structured materials systems

(4) Synthesize results of interdisciplinary design research.

(5) Apply computational skills in materials selection (CMS Selector), multicomponent thermodynamics (THERMOCALC) and diffusion (DICTRA), and integrate them in the theoretical conceptual design of materials.

(6) Perform in cross-functional teams.

(7) Communicate effectively in oral, written and graphical form.

Part XII
391: Process Design

156 391 Problems

(1) Given that $V(y_1) = E\left[(y_1 - \mu_1)^2\right]$ and $V(y_2) = E\left[(y_2 - \mu_2)^2\right]$ and y_1 and y_2 are independent, prove that $V(y_1 - y_2) = V(y_1) + V(y_2)$

(2) A computer program has produced the following output for the hypothesis-testing problem:

Difference in sample means: 2.35

Degrees of freedom: 18

Standard error of the difference in the sample means: ?

Test statistic: t_o = 2.01

p-value: 0.0298

a What is the missing value for the standard error?

b Is this a two-sided or one-sided test?

c If α = 0.05, what are your conclusions?

d Find a 90% two-sided confidence interval on the difference in the means.

(3) One of my students has developed two methods to make Ni films. Method A gives an average hardness of 15.0 GPa based on 10 random samples with sample standard deviation of 1.0 GPa. Method B gives an average hardness of 14.0 GPa based on 10 random samples with sample standard deviation of 1.0 GPa. Show that method A gives harder Ni films. Assume significance level α= 0.05.

(4) If the student in the above exercise were to make another set of 10 Ni films using method B, what is the probability that its average hardness is the same as that made by method A? Assume significance level α= 0.05. Hint: Think type II error.

(5) Suppose we are testing:

$$H_o : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

where σ_1^2 and σ_2^2 are known. Our sampling resources are constrained such that $n_1 + n_2 = N$, where N is fixed. How should we allocate the N observations between the two populations to obtain the most powerful test? For a given $\bar{y_1} - \bar{y}_2$, the most powerful test is attained by choosing n_1 and n_2 so that z_o is maximized, where

$$z_o = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Show that this happens when $n_1/n_2 = \sigma_1/\sigma_2$

(6) An experimenter has conducted a single-factor experiment with six levels of the factor, and each factor level has been replicated three times. The computer value of the$F$-statistic is 5.81. Determine the$p$-value.

Use the tool from: http://www.danielsoper.com/statcalc3/calc.aspx?id=7

Hint: determine the degrees of freedom for the factor (numerator) and degrees of freedom for the error (denominator).

(7) Five randomly selected batteries are tested for their lives. The result is shown below:

108,100,96,98,100 weeks

a Determine the sample mean and standard deviation.

b What is the probability that a given battery will fail at or before 85 weeks?

(8) An article in*Solid State Technology*, "Orthogonal Design of Process Optimization and Its Application to Plasma Etching" by G.Z. Yin and D.W. Jillie (May, 1987) describes an experiment to determine the effect ofC_2F_6 flow rate on the uniformity of the etch on a silicon wafer used in integrated circuit manufacturing. Data for two flow rates are as follows:

C_2F_6	Uniformity Observation					
(SCCM)	1	2	3	4	5	6
125	2.7	4.6	2.6	3.0	3.2	3.8
200	4.6	3.4	2.9	3.5	4.1	5.1

a Does theC_2F_6 flow rate affect average etch uniformity? Useα= 0.05. Hint: compare the two means

b Does theC_2F_6 flow rate affect the wafer-to-wafer variability in etch uniformity. Useα = 0.05. Hint: compare the ratio of the two variances. Think F distribution.

(9) A pharmaceutical manufacturer wants to investigate the bioactivity of a new drug. A completely randomized single-factor experiment was conducted with three dosage levels, and the following results were obtained.

Dosage	Observations			
20g	24	28	37	30
30g	37	44	31	35
40g	42	47	52	38

Is there evidence to indicate that dosage level affects bioactivity? Use$\alpha = 0.05$.

(10) A regional opera company has tried three approaches to solicit donations from 24 potential sponsors. The 24 potential sponsors were randomly divided into three groups of eight, and one approach was used for each group. The dollar amounts of the resulting contributions are shown in the following table.

Approach	Contributions ($)							
1	1000	1500	1200	1800	1600	1100	1000	1250
2	1500	1800	2000	1200	2000	1700	1800	1900
3	900	1000	1200	1500	1200	1550	1000	1100

a Do the data indicate that there is a difference in results obtained from the three different approaches? Use $\alpha = 0.05$.

b Analyze the residuals from this experiment and comment on the model adequacy.

(11) In the manufacturing of solar panels, there are four process parameters that may affect the efficiency of such panels. You plan to run a 2^4 factorial design experiment to optimize the manufacturing process. The standard deviation in determining solar panel efficiency is estimated to be about 1.0% (t distribution should therefore be used). In order to distinguish the outcome due to different process parameters, you want the 95-percent confidence interval of any effect to be ± 0.5% or less. How many replicates of the design do you need to run?

(12) The following output was obtained from a computer program that performed a two-factor ANOVA on a factorial experiment.

Two-way ANOVA: y versus A,B

Source	DF	SS	MS	F	P
A	2	?	0.38	?	0.0004
B	?	0.30	?	?	0.0085
Interaction	4	0.069	?	?	?
Error	9	0.16	?		
Total	17	1.30			

a Fill in the blanks in the ANOVA table

b How many levels were used for factor B?

c How many replicates of the experiment were performed?

d What conclusions would you draw about this experiment?

(13) Consider the wine tasting problem discussed in class

Judges	Wine X	Wine Y	Wine Z
A	16	17	15
B	15	16	14
C	15	16	15
D	16	18	16
E	16	16	16

Source of Variation	Sum of Squares	DoF	Mean Square	F_o
Total	12.4	14		
Wines	5.2	2	2.6	9.75
Judges	5.07	4	1.27	4.75
Residual	2.13	8	0.27	

Assume significant level$\alpha = 0.05$.

a Is the quality of wines Y and Z different?

b Are judges B and C about equally tough in rating wines?
Hint: think*LSD*.

Score difference = 1.4 >*LSD*. Therefore there is statistically no difference in the rating given by these two judges.

(14) In the development of aluminum alloys, one normally subjects them first to a solution treatment. They are then quenched and aged at 300ºC for different times. The following table shows the strength of two aluminum alloys (MPa) as a function of the aging times:

Aluminum Alloy type	Strength (MPa) after aging at 300ºC		
	1 hr	2hr	3hr
AA_1	280	300	290
	290	310	285
	285	295	290
AA_2	230	260	220
	235	240	225
	240	235	230

a Does the alloy type or aging time at 300ºC affect the strength?

b Do these two factors interact? Useα = 0.05.

c Check the adequacy of the model by performing a normal plot of the residuals.

(15) It turns out that in making aluminum alloy prototypes for the above problem, there is only enough material in one batch to perform one replicate for the three aging times. Therefore, one should reformat the above table in three blocks as shown below:

Aluminum Alloy type	Blocks	Strength (MPa) after aging at 300ºC		
		1 hr	2hr	3hr
AA_1	1	280	300	290
	2	290	310	285
	3	285	295	290
AA_2	1	230	260	220
	2	235	240	225
	3	240	235	230

Re-analyze the above data in blocks. Is blocking useful in this experiment?

(16) An engineer is interested in the effects of cutting speed (A), tool geometry (B), and cutting angle on the life (in hours) of a machine tool. Two levels of each factor are chosen, and three replicates of a2^3 factorial design are run. The results are as follows:

A	B	C	Treatment Combination	Replicate		
				I	II	III
-	-	-	(1)	22	31	25
+	-	-	*a*	32	43	29
-	+	-	*b*	35	34	50
+	+	-	*ab*	55	47	46
-	-	+	*c*	44	45	38
+	-	+	*ac*	40	37	36
-	+	+	*bc*	60	50	54
+	+	+	*abc*	39	41	47

a Estimate the factor effects. Which effects appear to be large?

b Use ANOVA to confirm the above conclusions.

c Write down the regression equation for predicting tool life (in hours) based on results of this experiment.

d Show that the residuals follow the normal distribution and that there is nothing unusual about the relationship between residuals and fitted values.

(17) (a) Based on the regression equation derived in part (c) of the previous question, it is obvious that we need to keep B high for maximum tool life. Determine the coded factor levels of A and C to achieve maximum tool life by making a contour plot (2D or 3D) when B is high.

(b) Find the standard error of the factor effects and approximate 95 percent confidence limits for the factor effects in the previous question. Do the results of this analysis agree with the conclusions from the analysis of variance?

(18) Consider the data shown in the previous question, but with only data from replicate I:

A	B	C	Treatment Combination	Replicate		
				I	II	III
-	-	-	(1)	22	31	25
+	-	-	*a*	32	43	29
-	+	-	*b*	35	34	50
+	+	-	*ab*	55	47	46
-	-	+	*c*	44	45	38
+	-	+	*ac*	40	37	36
-	+	+	*bc*	60	50	54
+	+	+	*abc*	39	41	47

Let us assume that the experiment ab could not be done, and we therefore cannot proceed with ANOVA. One solution is to estimate the missing value with a number that makes the highest order effect or contrast equal to zero (ABC in this case). On that basis, estimateab.

(19) Consider a2^3 design withABC confounded in the first two replicates andBC confounded in the third. Show that the degree of freedom for error is equal to 11. Hint: DoF(total) = DoF(factors) + DoF(all blocks) + DoF(error)

Replicate I		Replicate II		Replicate III	
(ABC Confounded)		(ABC confounded)		(BC Confounded)	
Block 1	Block 2	Block 1	Block 2	Block 1	Block 2
(1)	*a*	(1)	*a*	(1)	*b*
ab	*b*	*ab*	*b*	*bc*	*c*
ac	*c*	*ac*	*c*	*abc*	*ab*
bc	*abc*	*bc*	*abc*	*a*	*ac*

(20) It is known that heat treatment may affect the yield strength of steels. For a given category of steels, one heat treatment procedure gives an average yield strength ofY_1 with standard deviation ofS_1, while another heat treatment procedure givesY_2 with standard deviation ofS_2. To determine if heat treatment will indeed affect the yield strength of this category of steel, I takeN_1 steel samples after they are subjected to the first heat treatment procedure andN_2 samples after the second heat treatment procedure. Because of resource limitation, I am limited to the analysis of a total of$N$ measurements, i.e.,$N = N1 + N2$ is a constant. As guidance, I will chooseN_1 (and henceN_2) so that the standard error of the difference between the two means is as small as possible. This will enhance my ability to distinguish between two means. On that basis, show that my best choice forN_1 is given by:

$$N_1 = N\frac{S_1}{S_1 + S_2}$$

(21) You are working in a research lab, and you have 11 factors that might affect the performance of devices you fabricate. A full2^{11} design would take2^{11} runs, and you are only able to perform 32 runs/week. You just want to estimate the 11 main effects and get some insight regarding the two-factor interactions. That is, you don't want to have any aliasing with main factors, but aliasing of two-factor interactions are acceptable. You don't think three-factor and higher interactions will be important.

a What resolution design would be appropriate?

b What2^{k-p} fractional factorial design should you use?

c How many weeks will this design take you? How much will you save with this approach instead of the full2^{11}design?

(22) R.D. Snee ("Experimenting with a Large Number of Variables," in*Experiments in Industry: Design, Analysis and Interpretation of Results*, by R.D. Snee, L.B. Hare, and J.B. Trout, Editors, ASQC, 1985) describes an experiment in which a2^{5-1} design with$I$=$ABCDE$ was used to investigate the effects of five factors on the color of a chemical product. The factors are$A$ = solvent/reactant,B = catalyst/reactant,C = temperature,D = reactant purity, and$E$ = reactant pH. The results obtained were as follows:

$e = -0.63$	$d = 6.79$
$a = 2.51$	$ade = 5.47$
$b = -2.68$	$bde = 3.45$
$abe = 1.66$	$abd = 5.68$
$c = 2.06$	$cde = 5.22$
$ace = 1.22$	$acd = 4.38$
$bce = -2.09$	$bcd = 4.30$
$abc = 1.93$	$abcde = 4.05$

a Prepare a normal probability plot of the effects. Which effects seem active?

b Construct a normal probability plot of the residuals and plot the residuals versus the fitted values. Comment on the plots.

c If any factors are negligible, collapse the2^{5-1} design into a full factorial in the active factors. Comment on the resulting design, and interpret the results using ANOVA.

(23) An article in*Industrial and Engineering Chemistry* ("More on Planning Experiments to Increase Research Efficiency," 1970, pp. 60-65) uses a2^{5-2} design to investigate the effect of$A$ = condensation,B = amount of material 1,C =solvent volume, D = condensation time, and E = amount of material 2 on yield. The results are obtained are as follows:

$e = 23.2$	$ad = 16.9$	$cd = 23.8$	$bde = 16.8$
$ab = 15.5$	$bc = 16.2$	$ace = 23.4$	$abcde = 18.1$

a Verify that the design generators used were$I = ACE$and$I = BDE$.

b Estimate the main factor effects.

c If all three-factor interactions are negligible, prove thatAB andAD interactions are available to use as error. Hint: prove that they are aliased with other three-factor interactions.

(24) Consider the following fractional factorial experiment in which one investigates the effects of five factors$A$, B, C, D, and$E$:

Run	Treatment Combination
1	d
2	ae
3	b
4	abde
5	cde
6	ac
7	bce
8	abcd

a How many factors are in the basic design? That is, if the design is2^{k-p}, the number of factors in the basic design = k – p.

b Assume that the factors in the experiment are represented by the initial letters of the alphabet (i.e.A, B, etc.), prove that the design generators are$D = AB$ and$E = -AC$.

c What is the complete defining relation?

d What is the resolution of this design?

(25) Example 8.4 in the text (shrinkage in an injection molding process) shows that the variability of the data sometimes depends on the setting of a given parameter (high or low). Let us examine this variability question using the following data from a2^{4-1} fractional factorial design,y being the response:

Run	Treatment	A	B	C	$D = ABC$	y
1	(1)	-1	-1	-1	-1	8
2	*ad*	1	-1	-1	1	10
3	*bd*	-1	1	-1	1	12
4	*ab*	1	1	-1	-1	7
5	*cd*	-1	-1	1	1	13
6	*ac*	1	-1	1	-1	6
7	*bc*	-1	1	1	-1	5
8	*abcd*	1	1	1	1	11

a What is the resolution of this design?

b Calculate the effect of B.

c Determine$S^2(B = +)$ and$S^2(B = -)$

(26) A2^{6-2} factorial experiment with three replicates has been run in a pharmaceutical drug manufacturing process. The experimenter has used the following factors:

Factor	Natural Levels	Cooled levels (x's)
A	50,100	-1,1
B	20,60	-1,1
C	10,30	-1,1
D	12,18	-1,1
E	15,30	-1,1
F	60,100	-1,1

If two main effects and one two-factor interaction are included in the final model, how many degrees of freedom for error will be available?

(27) Heat treating is often used to carbonize metal parts, such as gears. The thickness of the carbonized layer is a critical output variable from this process, and it is usually measured by performing a carbon analysis on the gear pitch (top of the gear tooth). Six factors were studied on a 2^{6-2}_{IV} design: A = furnace temperature, B = cycle time, C = carbon concentration, D = duration of the carbonizing cycle, E = carbon concentration of the diffuse cycle, and F = duration of the diffuse cycle. The experiment is shown below:

Standard Order	Run Order	A	B	C	D	E	F	Pitch
1	5	-	-	-	-	-	-	74
2	7	+	-	-	-	+	-	190
3	8	-	+	-	-	+	+	133
4	2	+	+	-	-	-	+	127
5	10	-	-	+	-	+	+	115
6	12	+	-	+	-	-	+	101
7	16	-	+	+	-	-	-	54
8	1	+	+	+	-	+	-	144
9	6	-	-	-	+	-	+	121
10	9	+	-	-	+	+	+	177
11	14	-	+	-	+	+	-	135
12	13	+	+	-	+	-	-	170
13	11	-	-	+	+	+	-	126
14	3	+	-	+	+	-	-	175
15	15	-	+	+	+	-	+	126
16	4	+	+	+	+	+	+	193

a Estimate the factor effects, and plot them on a normal probability plot. Select a tentative model.

b Make an interaction graph between factors C and D, i.e., plot the pitch versus C for high and low values of D

(28) A company asks you to design a new alloy. You know there are 5 major elements that you will use. You know from prior experience that most two-factor interactions will be important due to formation of precipitates and that three-factor interactions are most likely of secondary significance. In addition, you are fairly confident that higher order interactions (4th and 5th order) are unimportant. Your lab only has enough material to run 16 experiments. What design resolution should you run to estimate all main and two-factor effects? What is the defining relation for such a design? How do you know the two factor effects aren't aliased with each other?

(29) A plant distills liquid air to produce oxygen, nitrogen, and argon. The percentage of impurity in the oxygen is thought to be linearly related to the amount of impurities in the air as measured by the "pollution count" in part per million (ppm). A sample of plant operating data is shown below.

Purity (%)	93.3	92.0	92.4	91.7	94.0	94.6	93.6	93.1	93.2	92.9	92.2	91.3	90.1	91.6	91.9
Pollution (ppm)	1.10	1.45	1.36	1.59	1.08	0.75	1.20	0.99	0.83	1.22	1.47	1.81	2.03	1.75	1.68

a Fit a linear regression model to the data (pollution is the independent or regressor variable).

b Test for significance of regression.

c Find a 95% confidence interval on β_1 (slope of the regression line)

(30) Repeat the above problem with design #2, assuming a quadratic function, i.e.,

$$y = \beta_o + \beta_1 x + \beta_2 x^2 + \epsilon$$

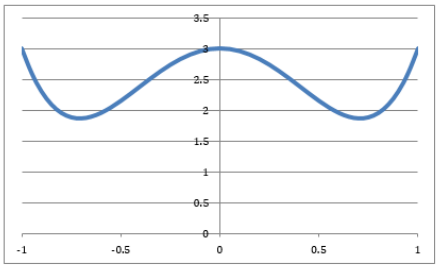
Derive the expression for the scaled prediction variation and plot it over the concentration range.

Solution

$$X = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, X^T X = \begin{bmatrix} 12 & 0 & 8 \\ 0 & 8 & 0 \\ 8 & 0 & 8 \end{bmatrix}, (X^T X)^{-1} = \begin{bmatrix} 0.25 & 0 & -0.25 \\ 0 & 0.12 & 0 \\ -0.25 & 0 & 0.37 \end{bmatrix}$$

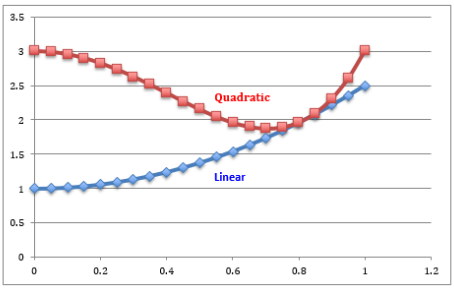
Therefore, the scaled prediction variance for design #2

$$= \begin{bmatrix} 1 x x^2 \end{bmatrix} \begin{bmatrix} 0.25 & 0 & -0.25 \\ 0 & 0.125 & 0 \\ -0.25 & 0 & 0.375 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} = 3.0 - 4.5x^2 + 4.5x^4$$



Supplemental discussion:

Here is the comparison of SPV for the linear and quadratic model:



Notice that the quadratic fit gives larger SPV.

The variance of prediction $V(\hat{y})$ is equal to SPV multiplied by $\frac{\sigma^2}{N}$. The quadratic fit should give a smaller σ^2. Depending on the relative magnitudes of σ^2 for the two models and the region of interest, one may be better off with one design versus the other.

(31) The data shown in the file "HW4 worksheet.xlsx" were collected in an experiment to optimize crystal growth as a function of three variables x_1, x_2 and x_3. Large values of y (yield in grams) are desirable.

a Fit a second-order model, i.e., include $A, B, C, AB, AC, BC, A^2, B^2$ and C^2. Show that the most important factors are B^2 and C^2.

b Re-do the analysis by including only B^2, C^2, and B and C, for hierarchical considerations. Under what set of conditions is maximum growth achieved?

(32) **Bonus Question**

You may notice a strange output from Minitab: the total sum of squares for the model is different from the sum of squares contributions from the 9 factors mentioned above. Can you explain the discrepancy?

(33) A manufacturer of cutting tools has developed two empirical equations for tool life in hours (y_1) and for tool cost in dollars (y_2). Both models are linear functions of steel hardness (x_1) and manufacturing time (x_2). The two equations are$y_1 = 10 + 5x_1 + 2x_2$ and$y_2 = 23 + 3x_1 + 4x_2$. Both equations are valid over the range -1.5$\leq x_i \leq$1.5. Unit tool cost must be below \$27.50, and life must exceed 12 hours for the product to be competitive. Is there a feasible set of operating conditions for this process? Where would you recommend that the process be run?

(34) Consider the following experimental design:

x_1	x_2	x_3
1	1	1
1	1	-1
1	-1	1
1	-1	-1
-1	1	1
-1	1	-1
-1	-1	1
-1	-1	-1
0	0	0
0	0	0
0	0	0
0	0	0

If the experiment*abc* cannot be performed, what will be the resulting design matrix$X$? Solve for$X^T X$.

Part XIII
396: Senior Projects

Kathleen A. Stair
Department of Materials Science and Engineering
Northwestern University

157 Catalog Description

To be taken in two consecutive quarters. Independent basic or applied research project, conceived and performed under the direction of a department faculty member. Prerequisite: senior standing in materials science program.

158 Course Outcomes

159 396: Senior Projects

At the conclusion of the course students will be able to:

(1) Understand the basis for making ethical decisions in the practice of science and engineering.

(2) Write a review of the literature relevant to their research topic.

(3) Give effective oral and written reports of ongoing and completed work.

(4) Write a research report consistent with standard practices.

160 Writing a Literature Review

160.1 General Advice

A literature review is a critical evaluation of work published on a specific topic. Also known as a review paper, the purpose of this type of document is to provide an overview of the state-of-the-art in a given field. Reading a literature review on a topic related to your research project is a good way to learn about the subject quickly. It is an attempt to evaluate what work has been done, what results have been obtained, and what issues remain in a field.

Your literature review should summarize and analyze prior work, organized by topic, not author(s). You should indicate what was done, how it was done, what the results were and what the results mean. (What did the authors conclude?) Ultimately, it will serve as the basis for the Introduction section of your final thesis. The scope of the review should not be too broad; it should focus on the specific topics related to your work. On the other hand, the review might encompass several related topics, i.e. synthesis, characterization techniques, results of tests on materials similar to yours, details about specific materials properties that motivate this work, etc. Consider the related subjects you must understand to successfully understand and carry-out your work. Your discussion of prior work should logically motivate your project. Most importantly, the body of your review should not consist of sequential summaries of individual papers, by author – this makes for an unsynthesized "laundry list" that is difficult for most readers to follow. Your review should summarize what is known from the literature, organized topically, and conclude with what questions remain. You could also summarize the approach your project will take to address those issues.

As you read your sources, take note of the main concepts and variables related to your topic. (Really – take notes! Read with questions in mind: Why was this work done? What questions were the researchers trying to answer? Why? How? What was measured? Computed? What were the results? What conclusions could be made?) Different researchers might have used different techniques (experimental or computational) to address the same topic. Their results may or may not agree. Your job is to group related information from different sources and discuss it in a way that gives your reader an overview of the specific issues related to your research. It will be useful to start with an outline of main topics and subtopics, before you write complete sentences or paragraphs.

This makes it easier to move information around, so you can more easily discuss work based on one topic at a time, not one paper at a time.
Structure your writing so that you start with general information and move to details. Your goal is to inform your readers. Your writing should follow a logical progression. Providing an overview before discussing details makes the document easier to read.
References in most peer-reviewed science or engineering journals are cited using a superscript Arabic numeral in the text. These are numbered in the sequence in which they appear, and are listed in the same order at the end of the paper. This is the method you should use. The citation style varies somewhat from journal to journal. One approach is to decide on the journal that would be most appropriate for submission of your work and use the guidelines provided to authors. One caveat – the Graduate School requires that graduate students include titles of journal articles in their citations, and this makes the list of references more useful to readers. You should do the same.
Finally, after you finish writing the first draft, set the document aside. Then return to it a day or two later and re-read it. Does it still make sense? Is it free from obvious grammatical and spelling errors? Is the wording concise and unambiguous? Would an educated reader unfamiliar with the details of your research still be able to appreciate the motivation behind it? Proofread.....revise.....proofread....submit only a carefully proofread report.

160.2 Checklist

Note: These are guidelines for content; please do not use "Introduction", "Body" as subtitles in your review.

160.2.1 Title

Review of — what? Create a title that will define the scope of the review.

160.2.2 Body

Includes a discussion of prior work in the context of the general problem outlined. Topics are introduced with an overview, and then covered in sufficient depth. Results from different authors on the same topic are grouped and discussed/ analyzed in depth. (If your paragraphs begin with author names, consider how to re-organize the information.) Subtitles may be used to organize the document in a reader-friendly manner. Transitions from one section to the next should provide continuity and logical progression.

160.2.3 Conclusions

Information presented in the body is synthesized and summarized. Author provides insight into the problem. Thesis or topic for research clearly follows and is supported in the text.

160.2.4 References

Use appropriately, with superscript Arabic numerals in the text citing references listed at the end in numerical order. Use a reference manager like Endnote or Zotero.

160.2.5 Writing Mechanics

Writing is clear, precise and concise. Jargon is avoided. Important terms or concepts are defined, described, and used in context. Author uses parallel construction, complete sentences, proper verb tense, spelling & modifiers. Nouns and verbs agree in number. Nouns and verbs are in close proximity. Writing is neither redundant nor ambiguous. Word choice indicates clear thinking and precise meaning. White space is used well; pages are numbered.

160.3 Additional Information

Some excellent advice is available from Dena Taylor, Director, Health Sciences Writing Centre, and Margaret Procter, Coordinator, Writing Support, University of Toronto [34]. They distinguish between an annotated bibliography and a literature review (you are writing the latter). They provide questions which you can use to guide your reading of the literature as well as writing about it. The writing center at UW Madison also has some useful information describing the sequence of tasks related to writing a research paper [1].

160.4 Assessment Rubric

160.4.1 Title

Review should have a title that introduces the topic.

160.4.2 Introduction

Excellent: The thesis and scope of the problem are clearly introduced and groundwork is provided for the research effort.

Good: The research topic is introduced but the general context is missing, or the scope of the review is not well-enough defined.

Poor: Scope of the review is unclear.

160.4.3 Body

Excellent: Includes a discussion of prior work (including results) in the context of the general problem outlined, organized by topic (not by author or individual papers). Transitions from one section to the next provide continuity and logical progression. Topics are covered in sufficient depth.

Good: Some, but not all, sections follow in a logical order. Discussion needs more evaluation or analysis. Topics not covered in enough depth.

Poor: Subtopics appear disjointed. No evaluation of topics. Review presents sources serially (i.e. each paragraph cites a different author) without synthesis.

160.4.4 Conclusions

Excellent: Information presented in the body is synthesized and summarized. Author provides insight into the problem. Thesis or topic for research clearly follows and is supported in the text.

Good: Summary comments indicate some analysis, but overall thesis or research topic is not clearly defined based on preceding discussion.

Poor: No synthesis of information. No conclusions drawn.

160.4.5 Writing Mechanics

Excellent: Writing is clear, precise and concise. Author uses parallel construction, proper verb tense, spelling & modifiers. Writing is neither redundant nor ambiguous. Word choice indicates clear thinking and precise meaning. Grammar is correct. Spelling is correct. Pages are numbered.

Good: Generally clear, but sometimes repetitive, poor word choice or unnecessary words. Text is difficult to read. Subjects are too removed from verbs.

Poor: Construction is not parallel. Noun-verb agreement in number or meaning is poor. Verb tense is not consistent. Meaning is ambiguous, sentences are convoluted.

160.4.6 References

- Numbers in text refer to references at end, in numerical order.
- Placement of references is appropriate.
- An appropriate number of references are cited. Work is properly cited.
- Direct quotes are avoided
- Phrases are not "lifted" from reference texts (author uses own words)
- References are cited properly using a recognizable journal format.
- References are complete: titles, all authors, publisher, dates, page numbers, web titles/authors/ access dates, etc. are included.

Part XIV
References, Nomenclature and Index

It's not really possible to have one set of nomenclature for the entire core course sequence, but we can at least try to be as consistent as possible. Below we list all the symbol definitions that are defined in the individual course documents.

References

[1] Academic and Professional Writing: Writing a Research Paper. 975

[2] "Bohr-atom-PAR" by JabberWok at the English language Wikipedia. Licensed under CC BY-SA 3.0 via Commons. 686

[3] Differential Scanning Calorimetry. Explains the differential scanning calorimetry technique used in polymer science. 573

[4] The Glass Transition. Explains the glass transition which is unique to polymers. 573

[5] Teflon. 564

[6] Thermal Transitions of Homopolymers: Glass Transition & Melting Point. Technical report. 572

[7] Cross product, April 2014. Page Version ID: 605766652. 254

[8] Grain boundary strengthening, May 2014. Page Version ID: 607460666. 292

[9] Miller index, May 2014. Page Version ID: 607139878. 257, 281

[10] Torsion spring, October 2014. Page Version ID: 618346044. 490

[11] Seawater, March 2015. Page Version ID: 650113695. 640

[12] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, 0009-revised edition edition, June 1965. 181

[13] Frank S. Bates and Glenn H. Fredrickson. Block Copolymers—Designer Soft Materials. *Physics Today*, 52(2):32–38, January 2008. 646

[14] James Edward Bauerle and J. S. Koehler. Quenched-in lattice defects in Gold. *Physical Review*, 107(6):1493, 1957. 197

[15] A. M. Brown and M. F. Ashby. Correlations for diffusion constants. *Acta Metallurgica*, 28(8):1085–1101, August 1980. 192, 193

[16] Robert T DeHoff. *Thermodynamics in materials science*. CRC/Taylor & Francis, Boca Raton, 2006. 38

[17] N. A. Fleck, V. S. Deshpande, and M. F. Ashby. Micro-architectured materials: past, present and future. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2121):2495–2516, June 2010. 427, 428

[18] Walter M. Hirthe and John O. Brittain. High-Temperature Steady-State Creep in Rutile. *Journal of the American Ceramic Society*, 46(9):411–417, September 1963. 501

[19] H. B. Huntington, G. A. Shirn, and E. S. Wajda. Calculation of the Entropies of Lattice Defects. *Physical Review*, 99(4):1085–1091, August 1955. 188

[20] Carl E. Johnson and Ellen J. Hathaway. Solid-Liquid Phase Equilibria for the Ternary Systems Li(F,Cl,I) and Na(F,Cl,I). *Journal of The Electrochemical Society*, 118(4):631–634, April 1971. 227

[21] E.W. Johnson and M.L. Hill. The Diffusivity of Hydrogen in Alpha Iron. *Trans. AIME*, 218(6):1104–1112. 189

[22] K.L. Johnson. *Contact Mechanics*. Cambridge University Press, Cambridge, 1985. 445, 451

[23] E.J. Kramer. Microscopic and Molecular Fundamentals of Crazing. In *Adv. Polym. Sci.*, volume 52/53, page 1. 1983. 464

[24] G. Masing and Bruce A. Rogers. *Ternary systems; introduction to the theory of three component systems*. Ternäre Systeme.English. Reinhold publishing corporation, New York, 1944. 165

[25] Bob McGinty. Coordinate Transforms. 432

[26] Rachel L. McSwain and Kenneth R. Shull. Cavity Nucleation and Delamination During Adhesive Transfer of a Thin Viscoelastic Film. *J. Appl. Phys.*, 99:053533, 2006. 504

[27] Metallos. Ellingham diagram giving the free energy of formation of metal oxides and the corresponding oxygen partial pressure at equilibrium. Labels in Greek., October 2007. 120

[28] Hamdy F.M. Mohamed, E.E. Abdel-Hady, and Salwa S. Mohamed. Temperature dependence of the free volume in polytetrafluoroethylene studied by positron annihilation spectroscopy. *Radiation Physics and Chemistry*, 76(2):160–164, 2007. 564

[29] A. D. Pelton and W. T. Thompson. Phase diagrams. *Progress in Solid State Chemistry*, 10, Part 3:119–155, 1975. 117, 141

[30] K.R. Shull. Contact Mechanics and the Adhesion of Soft Solids. *Mat. Sci. and Eng. R.*, 36:1–45, 2002. 443

[31] R. O. Simmons and R. W. Balluffi. Measurements of Equilibrium Vacancy Concentrations in Aluminum. *Physical Review*, 117(1):52–61, January 1960. 195

[32] Ad Smigelskas and Eo Kirkendall. Zinc Diffusion in Alpha-Brass. *Transactions of the American Institute of Mining and Metallurgical Engineers*, 171:130–142, 1947. WOS:A1947XR33600008. 238

[33] Andrea R. Tao, Susan Habas, and Peidong Yang. Shape control of colloidal metal nanocrystals. *Small*, 4(3):310–325, April 2008. 374

[34] Dena Taylor. The Literature Review: A Few Tips On Conducting It. 975

[35] J. Weertman. Dislocation Climb Theory of Steady-State Creep. *Asm Transactions Quarterly*, 61(4):681–&, 1968. WOS:A1968C550500002. 501

[36] C. A. Wert. Diffusion Coefficient of C in α-Iron. *Physical Review*, 79(4):601, 1950. 189

[37] C. A. Wert. Measurements on the Diffusion of Interstitial Atoms in B.C.C. Lattices. *Journal of Applied Physics*, 21(11):1196, 1950. 189

[38] Alan T. Zehnder. Linear Elastic Stress Analysis of 2d Cracks. In *Fracture Mechanics*, volume 62, pages 7–32. Springer Netherlands, Dordrecht, 2012. 456, 457

Nomenclature

α	Coefficient of thermal expansion
β	Coefficient of compressibility
ΔG_v	Gibbs free energy change per unit volume
ΔH_m	Molar enthalpy of melting
ΔP	Pressure difference
ΔS_m	Molar entropy of melting
ΔT	Undercooling
δ	Compressive displacement.
δ_t	Tensile displacement.
ℓs	Liquid solution
ℓ	Liquid
ℓ_i^D	Diffusion length for component i (m)
ϵ_0	Permittivity of Free Space ($8.854 \times 10^{12} \mathrm{C}^2/\mathrm{Jm}$)
γ	Interfacial free energy
$\gamma_{\alpha\beta}$	Interfacial free energy between α and β phases ($\mathrm{J/m}^2$)
$\gamma_{\alpha\beta}$	Interfacial free energy between α and β phases
$\hat{s}$	Unit vector directed along a dislocation core (dimensionless)
$\mathcal{G}$	Energy release rate
μ	Chemical potential
μ	Chemical potential
μ_A, μ_B	chemical potentials of A and B

μ_i	Chemical potential of component i (J/mole)
μ_i	Chemical potential of component i
ν	Poisson's ratio
Ω	Omega potential
ρ	Density (kg/m^3)
ρ_c	Crack tip radius of curvature
σ	Stefan's Constant ($5.67 \times 10^{-8} \frac{\mathrm{W}}{\mathrm{m^2 k^4}}$)
σ	Stress
σ	Tensile stress (Pa)
σ_f	Fracture strength
τ	Shear stress (Pa)
τ_{crss}	Critical resolved shear stress (Pa)
τ^0_{crss}	critical resloved shear stress in the absence of dislocations
τ_{rss}	Resolved shear stress (Pa)
$\tilde{D}$	Interdiffusion coefficient (m^2/s)
$\underline{F}$	Degrees of freedom
$\vec{b}$	Burgers vector (m)
$\vec{n}_d$	Vector cross product of $\vec{s}$ and $\vec{b}$ (m)
A	Contact area
A	Helmoltz free energy
A^*	Surface area of the critical nucleus
A_0	Undeformed cross section
a_i	Activity coeffiienct for component i (dimensionless)
b	Magnitude of $\vec{b}$ (m)
b_c	Minor axis of elliptical crack
C	Compliance
c	Speed of Light (2.998×10^8m/s)
C_P	Heat capacity at constant pressure
C_V	Heat capacity at constant volume
C_0	Flat punch compliance for a rigid punch with a circular cross section in contact with an elastic half space.
C_0	Overall concentration of atoms
C_i	Concentration of component i
d	Diatance from crack tip.
D_i	Intrinsic diffusion coefficient for component i (m^2/s)
D^*_i	Tracer diffusion coefficient for component i (m^2/s)
E	Energy (J)
E	Young's modulus
e	Electron Charge (1.6022×10^{-19} C)
E_r	Reduced modulus
E_s	Dislocation Energy (J)
e_{xy}	Shear strain in the x-y plane
F	Helmholtz free energy
F	Helmholtz free energy
F	Helmholtz free energy
f	Reactive functionality - the number of times a given molecule can react in a step growth polymerization.
F_s	Force per unit length acting on a dislocation (N)
F_s	Total force per unit length acting on a dislocation (N/m)
F^τ_s	Stress-induced force per unit length acting on a dislocation (N/m)
F^r_s	Curvature-induced force per unit length acting on a dislocation (N/m)
G	Gibbs free energy
G	Gibbs free energy
G	Gibbs free energy
G	Shear modulus (Pa)
G	Shear modulus
g	Gas
G^L_m	Molar Gibbs free energy of the liquid phase

G_m^S	Molar Gibbs free energy of the solid phase
G_N^0	Plateau modulus (also G_e)
H	Enthalpy
H	Enthalpy
h	Planck's Constant (6.626×10^{-34}Js)
h	Thickness of the compliant layer
h_p	Plastic zone size
H_i	Henry' law coefficient for component i (dimensionless)
I	Moment of intertia (Kg-m^2)
J_i	Diffusive flux of component i with respect to a cooridnate system fixed to the lattice planes (atoms/m^2/s)
$J_i^{'}$	Diffusive flux of component i with respect to a cooridnate system fixed to the external dimensions of the sample (atoms/m^2/s)
K_b	Bulk modulus
k_B	Boltzmann's constant (1.38x10^{-23} J/K)
k_B	Boltzmann's constant
K_t	Torsional stiffness
L_m	Molar heat of sublimation (Joules)
m_e	electron mass
m_e	Mass of an Electron (9.109×10^{-31}kg)
M_i	Mobility of component i (units of velocity/force)
M_n	Number average molecular weight
M_n^{max}	Maximum value of M_n that can be obtained from a stoichiometrically unbalanced, linear step growth polymerization
M_w	Weight average molecular weight
N	Degree of polymerization.
n	Number of moles
n	Number of moles
n	Total number of molecules
n_k	Moles of component k
n_i	Number of moles of component i
n_i	Total number of atoms of component i.
N_n	Number average degree of polymerization
n_N	Number avg. N
P	Compressive force.
P	Force (with components P_1, P_2, P_3)
P	Pressure
P	Pressure
p	Enumerate Reaction
p	Hydrostatic pressure
P^L	Liquid phase pressure
P^{S*}	Pressure inside the critical nucleus
P_t	Tensile force
P_t	Tensile force.
Pr	Probability density
Q	Heat
R	Gas constant (8.314 J/mole· K)
R	Ideal gas constant, 8.314 $J\,K^{-1}\,mol^{-1}$
R	Radius (or radius of curvature)
r	Particle radius (m)
R^*	Radius of the critical nucleus
S	Entropy
S	Entropy
s	Solid
S_m^L	Molar entropy of the liquid phase
S_m^L	Molar entropy of the liquid phase
S_m^S	Molar entropy of the critical nucleus

S_m^S	Molar entropy of the critical nucleus
T	Absolute Temperature
T	Absolute temperature (K)
T	Temperature
T_m	Equilibrium melting temperature
T_s	Dislocation line tension (N or J/m)
U	Internal energy
U	Internal energy
V	Volume (m^3)
V	Volume
V	Volume
V^*	Volume of the critical nucleus
v_ℓ	Velocity of the lattice planes with respect to a laboratory coordinate system (m/s)
V_m^S	Molar volume of the solid phase
V_m^S	Molar volume of the solid phase
W	Work done on system
W	Work
w	Total weight of polymer
w_N	weight of all molecules with degreee of polymerization = N
W_R^*	Pressure difference
X_A, X_B	Mole fractions of A and B
F	Helmholtz free energy
G	Gibbs free energy
T	Absolute temperature

Index

7 Crystal Systems., 834

absorption length, 873
Activation volume, 497
Activity coefficient, 133
adsorbed intensity, 872
Alkemade Theorem, 154
Anomalous dispersion, 889
Aromatic compounds, 540
Atactic, 560
atomic scattering factor, 884
Auger effect, 870
Auger electron, 870
Avogadro's number, 597

Benzene, 540
Berkovich tip, 452
Bernal chart, 920
Birefringence, 617, 621
Block copolymers, 569
Bohr radius, 686
Boltmann's Constant, 536
Boltzmann superposition principle, 599
Bragg ring, 909
Bravais-Miller, 827
Bremsstrahlung radiation, 865
Bulk modulus, 440
Burgers circuit, 251
Burgers vector, 249

Chain Folding, 613
Chain-Growth Polymerization, 554
Characteristic ratio, 585
characteristic spectrum, 868
Charpy impact test, 467
Chemical potential, 622
Chi parameter, 627
Classical thomson scattering, 877
Classification scheme, 538
Climb (of a dislocation), 257
Closed system, 39
Coble Creep, 503
Coefficient of Compressibility, 46
coherent, 859
Compton scattering, 886
Considére construction, 480
coordination number, 856
Copolymers, 569
Correlation length, 637
Covalent Bonding, 539
Crack opening displacement, 465
Crazing, 463
Creep compliance function, 494
Creep rupture, 499
Critical Point, 630
Critical resolved shear stress
limiting value in the absence of dislocations, 247
critical resolved shear stress, 477
Cross slip, 257
Crosslink points, 594
Crosslinking, 568
Crystal, 818
Crystal axes, 826
Crystal Classes, 823
crystal lattice, 830
crystal period, 615
Crystal Systems, 823
Crystal Thickness, 613
crystal thickness, 615
Crystalline Unit Cells, 611

d-spacing, 840
Dashpot, 491
Deformation twinning, 483
Degree of polymerization, 544
degrees of freedom, 115, 116
Diblock copolymer, 569
Diblock Copolymer Morphologies, 644
Dienes, 559
Differential scanning calorimetry, 573
Diffusion equation, 233
director, 620
Dislocation
Edge, 248
Dislocation density, 256
Dislocations
Screw, 250
Sense vector, 249
$\vec{b} \times \vec{s}$ cross product, 253
Double cantilever beam geometry, 459
Dugdale Model, 465

Edge dislocation, 248
Eigenvalues, 436
Elastomers, 538
electron mass, 686
Ellingham diagram, 119
Energy release rate, 458
Engineering strains, 590
Engineering stress, 429, 590
Entanglements, 601
Epoxies, 541, 543
Equilibrium Contact Angle, 641
Equilibrium Swelling, 640
Ewald sphere., 898
Extension ratios, 438
Extensive properties, 41
Extent of reaction, 541
Eyring model, 497

Fiber torsion, 471

Fick's first law, 233
First law of thermodynamics, 42, 114, 591
Flory-Huggins Free Energy of Mixing, 627
fluorescence yield, 870
fluorescent lines, 868
form, 828
four axial systems, 831
Fracture modes, 455
Frank-Read source, 263
Free Volume, 570

gas constant, 50
Gauche configuration, 585
Gel fraction, 552
Gel point, 552
gelation, 552
Gibbs free energy, 267
Gibbs phase rule, 115
Gibbs-Duhem equation, 113
Glass transition temperature, 570
Glassy Polymers, 538
Glide, 256, 833
Glide plane, 256
Great Circles, 837

Hall-Petch relationship, 292
Heat Capacity, 47
Heat of sublimation, 282
Helix Formation, 609
Helmholtz free energy, 267
Henry's Law, 133
Henry's law, 279
Henry's law coefficient, 279
Heteroepitaxy, 297
Heterogeneous system, 38
High impact polystyrene (HIPS), 465
Homogeneous system, 38
hydrolysis, 543
Hydrostatic pressure, 436

improper rotation axis, 819
incident intensity, 872
Intensive properties, 41
Interdiffusion coefficient, 240
Interfacial energy, 271
Interfacial Free Energy, 642
Interference, 859
Intrinsic diffusion coefficient, 233
Inversion center, 818
ionic conductivity, 221
Irreversible processes, 43
Irwin Model, 455
Isolated system, 39
Isotactic, 560
Izod impact test, 467

JKR equation, 452
Johnson-Mehl-Avrami-Kolmogorov (JMAK) equation, 292

K\alpha doublet, 868
Kevlar, 566

lamellar crystallites, 614
Laplace pressure equation, 273
Larmor Formula, 687
Lattice points, 830
Lewis Diagrams, 539
linear absorption coefficient, 872
Liquid Crystals, 620
Living Polymerizations, 560
Loss modulus, 489
loss tangent, 605

Macrostate, 65
Maltese cross, 617
mass absorption coefficient, 873
MATLAB
fsolve, 294
Function plot, 234
loading and plotting data from .mat file, 307
polar plot, 318
setting plotting defaults, 308
Solve nonlinear equation, 681
Symbolic math, 681
Vacancy diffusion simulation, 237
Wulff Construction, 284
Maxwell model, 601
Maxwell model - Generalized, 602
Microstate, 65
Miller indices, 281, 826
Mirror plane, 818
Modulus
reduced, 444
Mohr circle construction, 434
Mohr's circle, 440
Monomeric repeat unit, 609
Multicomponent system, 38
multiplicity, 910

Natural Rubber, 559
nematic liquid crystal phase, 620
Neohookean model, 595
Neoprene, 563
Network, 551
Network strands, 594
Newman projection, 585
Non-Linear Step Growth Polymerization, 551
non-primitive unit cell, 832
Non-reacting system, 39
Number average degree of polymerization ($_n$), 545
Number average molecular weight (M_n), 545
Nylon 6, 565

Occupied volume, 570

Octahedral voids, 856
Open system, 39
Osmotic Pressure, 634
Overlap concentration, 636

Partial wetting, 294
photoelectric effect, 876
\pi polarization case, 879
Planck's constant, 686
plane group, 832
plane groups, 832
Point groups, 474
relationships among the matrix components for linear properties, 476
Point of projection, 838
Poisson's ratio, 441, 590
Poly(dimethyl siloxane), 568
Poly(ethylene oxide), 565
Poly(methyl acrylate), 562
Poly(methyl methacrylate), 562
Poly(phenylene oxide), 567
Poly(tetrafluoroethylene) (PTFE), 564
Poly(vinyl acetate), 564
Poly(vinyl chloride) PVC, 564
Poly(vinyl pyridine), 565
Polyamides, 540–542
Polybutene-1, 562
Polycaprolactam, 565
Polycaprolactone, 566
Polycarbonate, 566
Polychloroprene, 563
polydispersity index, 545
Polyesters, 540–542
Polyethylene, 561
Polyethylene Terephthalate (PET), 567
Polyisobutylene, 563
Polypropylene, 562
Polystyrene, 563
powder method, 908
prepolymer, 552
primitive unit cell, 832
Principal extension ratios, 590
Principal stresses, 589
process variables, 39
Projection plane, 838
Proper rotation axes, 819

Random Walk in One Dimension, 575
Random Walk in Two Dimensions, 578
Reacting system, 39
Reactive Functionality, 551
reciprocal lattice, 842, 843
Reduced modulus, 444
regular solution model, 129
Relaxation modulus, 488
Repeat units, 537
Resolved shear stress, 477
Reversible processes, 43
rotating crystal method, 917
Rotational axis, 818
Rubber Elasticity, 536, 593

Saddle-point, 187
Screw Dislocations, 250
Second virial coefficient, 636
Self avoiding Random Walks, 586
Sense vector, 249
Shear modulus, 439, 591
Shear strain, 591
Shear stress, 591
Shear viscosity, 598
\sigma polarization, 879
Silicones, 567
Simple shear, 439
Slip Plane, 248
small circles, 839
Sol fraction, 552
Solid solution strengthening, 266
Sound velocity, 442
Space Elevator, 510
spherulite, 615
Spinodal curve, 629
Stability limit, 629
standard linear solid, 493
State Variables, 39
Statistical copolymers, 569
Statistical segment length, 578
Steady state diffusion, 172
Step-Growth Polymerizations, 541
Stereographic projections, 837
stereographic projections, 820
Storage modulus, 489
Strain, 437
Strain ellipsoid, 590
Strains
Sample Displacements for Small Strains, 437
Stress intensity factor, 456
Stress invariants, 436
Structural repeat unit, 609
Symmetry, 818
Synchrotron radiation, 866
Syndiotactic, 560

Tacticity, 560
Temperature
as a state function, 39
Tempered glass, 509
Ternary Lever Rule, 146
Ternary Phase Diagrams
Isothermal Sections, 157
tetrahedral void, 856
Thermal expansion coefficient, 571
Thermite reaction, 56
Thermodynamic laws

first law, 42
Theta temperature, 632
Thin Film Growth, 300
Three phase contact lines, 292
tie line
binary system, 139
Tie molecules, 614
Tilt boundary, 286
Trans configuration, 585
Translation center, 818
transmitted intensity, 872
transverse scan, 907
Tresca yield criterion, 478
True strain, 438
True strains, 590
True stress, 590
Twin boundaries, 287
Twin boundary, 286
Twist boundary, 286

Ultem Polyetherimide, 567
Unary system, 38
Undercooling, 618
Unit cell, 611
unit cell, 832

Vinyl polymers, 560
Virial expansion, 636
Viscoelasticity, 597
Vogel equation, 603
Von Mises stress, 479

Weibull distribution, 504
Weibull modulus, 504
Weight average molecular weight (M_w), 545
Wulff construction, 284

X-ray, 858
X-ray filament tube, 864

Young's modulus, 441

Zero shear viscosity, 601
zone, 829
zone axis, 829