# Lecture Notes on Thermodynamics and Statistical Mechanics (A Work in Progress)

Daniel Arovas
Department of Physics
University of California, San Diego

October 5, 2026

# Contents

**Contents** i

**List of Tables** xvi

**List of Figures** xvii

0.1 Preface . . . 1

0.2 General references . . . 1

**1 Fundamentals of Probability** 3

1.1 References . . . 3

1.2 Statistical Properties of Random Walks . . . 4

1.2.1 One-dimensional random walk . . . 4

1.2.2 Thermodynamic limit . . . 6

1.2.3 Entropy and energy . . . 7

1.3 Basic Concepts in Probability Theory . . . 8

1.3.1 Fundamental definitions . . . 8

1.3.2 Bayesian statistics . . . 9

1.3.3 Random variables and their averages . . . 10

1.4 Entropy and Probability . . . 12

1.4.1 Entropy and information theory . . . 12

1.4.2 Rényi entropy . . . 13

1.4.3 Kullback-Leibler divergence . . . 14

1.4.4 Probability distributions from maximum entropy . . . 14
1.4.5 Continuous probability distributions . . . 18
1.5 General Aspects of Probability Distributions . . . 19
1.5.1 Discrete and continuous distributions . . . 19
1.5.2 Central limit theorem . . . 21
1.5.3 Moments and cumulants . . . 23
1.5.4 Multidimensional Gaussian integral . . . 24
1.6 Appendix: Bayesian Statistical Inference . . . 24
1.6.1 Frequentists and Bayesians . . . 24
1.6.2 Updating Bayesian priors . . . 25
1.6.3 Hyperparameters and conjugate priors . . . 27
1.6.4 The problem with priors . . . 28

**2 Thermodynamics** **31**
2.1 References . . . 31
2.2 What is Thermodynamics? . . . 32
2.2.1 Thermodynamic systems and state variables . . . 32
2.2.2 Mathematical interlude : exact and inexact differentials . . . 34
2.2.3 Equations of state . . . 36
2.2.4 Heat . . . 37
2.2.5 Work . . . 37
2.2.6 Pressure and temperature . . . 38
2.2.7 Standard temperature and pressure . . . 39
2.3 The Zeroth Law of Thermodynamics . . . 40
2.3.1 Gibbs phase rule . . . 41
2.4 The First Law of Thermodynamics . . . 42
2.4.1 Conservation of energy . . . 42
2.4.2 Single component systems . . . 42
2.4.3 Heat capacity . . . 44

2.4.4 Ideal gases . . . 46
2.4.5 Adiabatic transformations of ideal gases . . . 48
2.4.6 Adiabatic free expansion . . . 49
2.5 Heat Engines and the Second Law of Thermodynamics . . . 50
2.5.1 There's no free lunch so quit asking . . . 50
2.5.2 Engines and refrigerators . . . 51
2.5.3 Nothing beats a Carnot engine . . . 53
2.5.4 The Carnot cycle . . . 54
2.5.5 Carnot engine at maximum power output . . . 56
2.6 The Entropy . . . 58
2.6.1 Entropy and heat . . . 58
2.6.2 The third law of thermodynamics . . . 59
2.6.3 Entropy changes in cyclic processes . . . 60
2.6.4 Gibbs-Duhem relation . . . 61
2.6.5 Entropy for an ideal gas . . . 62
2.6.6 Example system . . . 63
2.6.7 Measuring the entropy of a substance . . . 65
2.7 Thermodynamic Potentials . . . 65
2.7.1 Energy $E$ . . . 66
2.7.2 Helmholtz free energy $F$ . . . 66
2.7.3 Enthalpy $\mathsf{H}$ . . . 67
2.7.4 Gibbs free energy $G$ . . . 68
2.7.5 Grand potential $\Omega$ . . . 69
2.8 Maxwell Relations . . . 69
2.8.1 Relations deriving from $E(S, V, N)$ . . . 70
2.8.2 Relations deriving from $F(T, V, N)$ . . . 70
2.8.3 Relations deriving from $\mathsf{H}(S, p, N)$ . . . 71
2.8.4 Relations deriving from $G(T, p, N)$ . . . 71

2.8.5 Relations deriving from $\Omega(T, V, \mu)$ . . . 72
2.8.6 Generalized thermodynamic potentials . . . 73
2.9 Equilibrium and Stability . . . 74
2.9.1 Equilibrium . . . 74
2.9.2 Stability . . . 74
2.10 Applications of Thermodynamics . . . 77
2.10.1 Adiabatic free expansion revisited . . . 77
2.10.2 Energy and volume . . . 78
2.10.3 van der Waals equation of state . . . 79
2.10.4 General equation of state . . . 81
2.10.5 Thermodynamic response functions . . . 81
2.10.6 Joule effect: free expansion of a gas . . . 84
2.10.7 Throttling: the Joule-Thomson effect . . . 85
2.11 Phase Transitions and Phase Equilibria . . . 88
2.11.1 $p$-$v$-$T$ surfaces . . . 89
2.11.2 The Clausius-Clapeyron relation . . . 89
2.11.3 Liquid-solid line in $H_2O$ . . . 92
2.11.4 Slow melting of ice : a quasistatic but irreversible process . . . 94
2.11.5 Generalized Gibbs phase rule . . . 95
2.12 Entropy of Mixing and the Gibbs Paradox . . . 97
2.12.1 Computing the entropy of mixing . . . 97
2.12.2 Entropy and combinatorics . . . 99
2.12.3 Weak solutions and osmotic pressure . . . 101
2.12.4 Effect of impurities on boiling and freezing points . . . 103
2.12.5 Binary solutions . . . 105
2.13 Some Concepts in Thermochemistry . . . 113
2.13.1 Chemical reactions and the law of mass action . . . 113
2.13.2 Enthalpy of formation . . . 115

2.13.3 Bond enthalpies . . . 118
2.14 Appendix I: Integrating Factors . . . 120
2.15 Appendix II: Other Thermodynamic Cycles . . . 121
2.15.1 The Stirling cycle . . . 121
2.15.2 The Otto and Diesel cycles . . . 123
2.15.3 The Joule-Brayton cycle . . . 125
2.16 Appendix III: Legendre Transformations . . . 127
2.16.1 Partial Legendre transformation . . . 129
2.17 Appendix IV: Useful Mathematical Relations . . . 130

**3 Ergodicity and the Approach to Equilibrium 135**
3.1 References . . . 135
3.2 The Master Equation and the Approach to Equilibrium . . . 136
3.2.1 Equilibrium . . . 136
3.2.2 The master equation . . . 136
3.2.3 Solution to the master equation . . . 137
3.2.4 Boltzmann's $\mathcal{H}$-theorem . . . 140
3.2.5 Markov chains . . . 141
3.3 Phase Flows in Classical Mechanics . . . 143
3.3.1 Hamiltonian evolution . . . 143
3.3.2 Dynamical systems and the evolution of phase space volumes . . . 144
3.3.3 Liouville's equation and the microcanonical distribution . . . 147
3.4 Irreversibility and Poincaré Recurrence . . . 149
3.4.1 Poincaré recurrence theorem . . . 149
3.4.2 Kac ring model . . . 151
3.5 Remarks on Ergodic Theory . . . 154
3.5.1 Definition of ergodicity . . . 154
3.5.2 The microcanonical ensemble . . . 156
3.5.3 Ergodicity and mixing . . . 156

3.6 Thermalization of Quantum Systems . . . 161
3.6.1 Quantum dephasing . . . 161
3.6.2 Eigenstate thermalization hypothesis . . . 162
3.6.3 Entanglement and entanglement entropy . . . 164
3.6.4 Singular value decomposition and entanglement . . . 165
3.6.5 Digression on random matrices . . . 166
3.6.6 More precise formulation of ETH . . . 169
3.6.7 When is the ETH true? . . . 170
3.7 Appendix I: Matrices and Jordan Canonical Form . . . 171
3.7.1 Normal and non-normal matrices . . . 171
3.7.2 Contrast with singular value decomposition . . . 173
3.7.3 Examples . . . 173
3.7.4 Exponentiating Jordan blocks . . . 174
3.8 Appendix II: Poisson Processes and Radioactive Decay . . . 175
3.8.1 Poisson process . . . 175
3.8.2 Radioactive decay . . . 175
3.9 Appendix III: The Markovian Mpemba Effect . . . 177
3.9.1 Introduction . . . 177
3.9.2 Distance between distributions . . . 178
3.9.3 Conditions for the Markovian Mpemba effect . . . 179
3.9.4 Numerical simulation . . . 179
3.10 Appendix IV: Transition to Ergodicity in a Simple Model . . . 180

**4 Statistical Ensembles** **189**
4.1 References . . . 189
4.2 Microcanonical Ensemble ($\mu$CE) . . . 190
4.2.1 The microcanonical distribution function . . . 190
4.2.2 Density of states . . . 191
4.2.3 Arbitrariness in the definition of $S(E)$ . . . 194

4.2.4 Ultra-relativistic ideal gas . . . 194
4.2.5 Discrete systems . . . 194
4.2.6 Two systems in thermal contact . . . 195
4.3 The Quantum Mechanical Trace . . . 197
4.3.1 The density matrix . . . 197
4.3.2 Averaging the DOS . . . 200
4.3.3 Coherent states . . . 201
4.4 Ordinary Canonical Ensemble (OCE) . . . 203
4.4.1 Canonical distribution and partition function . . . 203
4.4.2 The difference between $P(E_n)$ and $P_n$ . . . 204
4.4.3 Additional remarks . . . 204
4.4.4 Averages within the OCE . . . 205
4.4.5 Entropy and free energy . . . 205
4.4.6 Fluctuations in the OCE . . . 206
4.4.7 Energy distribution within the OCE . . . 206
4.4.8 Thermodynamics revisited . . . 207
4.4.9 Generalized susceptibilities . . . 208
4.5 Grand Canonical Ensemble (GCE) . . . 209
4.5.1 Grand canonical distribution and partition function . . . 209
4.5.2 Entropy and Gibbs-Duhem relation . . . 210
4.5.3 Fluctuations in the GCE . . . 211
4.5.4 Generalized susceptibilities in the GCE . . . 212
4.5.5 Gibbs ensemble . . . 213
4.6 Statistical Ensembles from Maximum Entropy . . . 213
4.6.1 Microcanonical ensemble . . . 214
4.6.2 Ordinary canonical ensemble . . . 214
4.6.3 Grand canonical ensemble . . . 215
4.6.4 Generalized Gibbs ensembles . . . 215

4.7 Ideal Gas Statistical Mechanics . . . 217
4.7.1 Maxwell velocity distribution . . . 217
4.7.2 Equipartition . . . 219
4.8 Selected Examples . . . 220
4.8.1 Spins in an external magnetic field . . . 220
4.8.2 Negative temperature (!) . . . 222
4.8.3 Adsorption . . . 223
4.8.4 Dipole gas . . . 224
4.8.5 Elasticity of wool . . . 226
4.9 Statistical Mechanics of Molecular Gases . . . 228
4.9.1 Separation of translational and internal degrees of freedom . . . 228
4.9.2 The internal coordinate partition function . . . 229
4.9.3 Quantum theory of rotations . . . 230
4.9.4 Vibrations . . . 232
4.9.5 Two-level systems : Schottky anomaly . . . 233
4.9.6 Electronic and nuclear excitations . . . 235
4.10 Appendix I: Constrained Extremization of Functions . . . 237
4.10.1 Example : volume of a cylinder . . . 238
4.11 Appendix II: Additional Examples of Statistical Ensembles . . . 239
4.11.1 Noninteracting spin dimers . . . 239
4.11.2 Three-state system . . . 240
4.11.3 Spins and vacancies on a surface . . . 241
4.11.4 Fluctuating interface . . . 243
4.11.5 Dissociation of atomic hydrogen . . . 245

**5 Noninteracting Quantum Systems** **247**
5.1 References . . . 247
5.2 Statistical Mechanics of Noninteracting Quantum Systems . . . 248
5.2.1 Bose and Fermi systems in the grand canonical ensemble . . . 248

5.2.2 Quantum statistics and the Maxwell-Boltzmann limit . . . 250
5.2.3 Entropy and counting states . . . 251
5.2.4 Single particle density of states . . . 253
5.3 Quantum Ideal Gases : Low Density Expansions . . . 254
5.3.1 Expansion in powers of the fugacity . . . 254
5.3.2 Virial expansion of the equation of state . . . 255
5.3.3 Ballistic dispersion . . . 256
5.4 Photon Statistics . . . 257
5.4.1 Thermodynamics of the photon gas . . . 257
5.4.2 Classical arguments for the photon gas . . . 259
5.4.3 Surface temperature of the earth . . . 260
5.4.4 Distribution of blackbody radiation . . . 261
5.4.5 What if the sun emitted ferromagnetic spin waves? . . . 262
5.5 Lattice Vibrations : Einstein and Debye Models . . . 263
5.5.1 One-dimensional chain . . . 263
5.5.2 General theory of lattice vibrations . . . 265
5.5.3 Einstein and Debye models . . . 269
5.5.4 Melting and the Lindemann criterion . . . 270
5.5.5 Goldstone bosons . . . 274
5.6 The Ideal Bose Gas . . . 275
5.6.1 General formulation for noninteracting systems . . . 275
5.6.2 Ballistic dispersion . . . 277
5.6.3 Energy and entropy, heat capacity, and compressibility . . . 278
5.6.4 Isotherms for the ideal Bose gas . . . 282
5.6.5 Mean field theory of the interacting Bose gas . . . 283
5.6.6 Properties of liquid $^4$He . . . 285
5.6.7 Bose condensation in optical traps . . . 291
5.7 The Ideal Fermi Gas . . . 293

5.7.1 Grand potential and particle number . . . 293
5.7.2 The Fermi distribution . . . 293
5.7.3 $T = 0$ and the Fermi surface . . . 293
5.7.4 The Sommerfeld expansion . . . 296
5.7.5 Magnetic susceptibility . . . 301
5.7.6 Moment formation in interacting itinerant electron systems . . . 304
5.7.7 White dwarf stars . . . 312
5.8 Appendix I: Second Quantization . . . 316
5.8.1 Basis states and creation/annihilation operators . . . 316
5.8.2 Second quantized operators . . . 318
5.9 Appendix II: Ideal Bose Gas Condensation . . . 319
5.10 Appendix III: Example Bose Condensation Problem . . . 321
5.11 Appendix IV: $C_{p,N}$ for the Ideal Bose Gas . . . 323

**6 Classical Interacting Systems** **325**
6.1 References . . . 325
6.2 The Ising Model . . . 326
6.2.1 Spontaneous symmetry breaking . . . 327
6.2.2 Solution in $d = 1$ dimension . . . 328
6.2.3 Domain walls . . . 331
6.2.4 Ising model in two dimensions : Peierls' argument . . . 333
6.2.5 High temperature expansion . . . 337
6.2.6 Kramers-Wannier duality . . . 340
6.3 Nonideal Classical Gases . . . 341
6.3.1 The configuration integral . . . 341
6.3.2 One-dimensional Tonks gas . . . 342
6.3.3 Mayer cluster expansion . . . 343
6.3.4 Examples . . . 350
6.4 Liquid State Physics . . . 357

6.4.1 The many-particle distribution function . . . 357
6.4.2 Averages over the distribution . . . 359
6.4.3 Virial equation of state . . . 362
6.4.4 Correlations and scattering . . . 364
6.4.5 Correlation and response . . . 367
6.4.6 BBGKY hierarchy . . . 369
6.4.7 Ornstein-Zernike theory . . . 370
6.4.8 Percus-Yevick and hypernetted chains approximations . . . 373
6.5 Coulomb Systems : Plasmas and the Electron Gas . . . 375
6.5.1 Electrostatic potential . . . 375
6.5.2 Debye-Hückel theory . . . 376
6.5.3 The electron gas : Thomas-Fermi screening . . . 379
6.6 Polymers . . . 382
6.6.1 Basic concepts . . . 382
6.6.2 Polymers as random walks . . . 383
6.6.3 The Rouse model . . . 386
6.6.4 Flory theory of self-avoiding walks . . . 390
6.6.5 Polymers and solvents . . . 391
6.7 Lee-Yang Theory . . . 393
6.7.1 Analytic properties of the partition function . . . 393
6.7.2 Electrostatic analogy . . . 395
6.7.3 Example . . . 396
6.8 Appendix: Potts Model in One Dimension . . . 397
6.8.1 Hamiltonian . . . 397
6.8.2 Transfer matrix . . . 398

**7 Mean Field Theory of Phase Transitions 401**
7.1 References . . . 401
7.2 The van der Waals system . . . 402

7.2.1 Equation of state . . . 402
7.2.2 Analytic form of the coexistence curve near the critical point . . . 406
7.2.3 History of the van der Waals equation . . . 410
7.3 Fluids, Magnets, and the Ising Model . . . 411
7.3.1 Lattice gas description of a fluid . . . 411
7.3.2 Phase diagrams and critical exponents . . . 412
7.3.3 Gibbs-Duhem relation for magnetic systems . . . 415
7.3.4 Order-disorder transitions . . . 415
7.4 Mean Field Theory . . . 417
7.4.1 The mean field *Ansatz* . . . 417
7.4.2 Zero external field . . . 418
7.4.3 Finite external field . . . 420
7.4.4 Magnetization dynamics . . . 423
7.4.5 Beyond nearest neighbors . . . 426
7.4.6 Ising model with long-ranged forces . . . 427
7.5 Landau Theory of Phase Transitions . . . 428
7.5.1 Quartic free energy with Ising symmetry . . . 429
7.5.2 Cubic terms in Landau theory : first order transitions . . . 431
7.5.3 Order parameter dynamics . . . 432
7.5.4 Sixth order Landau theory : tricritical point . . . 434
7.5.5 Hysteresis for the sextic potential . . . 436
7.6 Variational Density Matrix Method . . . 438
7.6.1 The variational principle . . . 438
7.6.2 Variational density matrix for the Ising model . . . 439
7.6.3 $q$-state Potts model . . . 442
7.6.4 $XY$ Model . . . 446
7.7 Mean Field Theory of Fluctuations . . . 449
7.7.1 Correlation and response in mean field theory . . . 449

7.7.2 Calculation of the response functions . . . 450
7.7.3 Beyond the two-state Ising model . . . 453
7.7.4 Magnetic ordering . . . 456
7.8 Global Symmetries . . . 459
7.8.1 Symmetries and symmetry groups . . . 459
7.8.2 Domain walls and lower critical dimension . . . 460
7.8.3 Continuous symmetries . . . 463
7.8.4 Random systems : Imry-Ma argument . . . 464
7.9 Ginzburg-Landau Theory . . . 465
7.9.1 Ginzburg-Landau free energy . . . 465
7.9.2 Domain wall profile . . . 466
7.9.3 Ginzburg criterion . . . 467
7.10 Appendix I: Bifurcations . . . 469
7.10.1 $N = 1$ dynamical systems . . . 469
7.10.2 Saddle-node bifurcation . . . 469
7.10.3 Transcritical bifurcation . . . 470
7.10.4 Pitchfork bifurcation . . . 471
7.10.5 Imperfect bifurcation . . . 472
7.11 Appendix II: Equivalence of the Mean Field Descriptions . . . 475
7.11.1 Variational density matrix . . . 476
7.11.2 Mean field approximation . . . 477
7.12 Appendix III: Additional Examples of Mean Field Theory . . . 478
7.12.1 Blume-Capel model . . . 478
7.12.2 Ising antiferromagnet in an external field . . . 480
7.12.3 Canted quantum antiferromagnet . . . 483
7.12.4 Coupled order parameters . . . 485
7.13 Appendix IV: Derivation of Ginzburg-Landau Free Energy . . . 490

**8 The Boltzmann Equation** **495**

8.1 References . . . 495
8.2 Equilibrium, Nonequilibrium and Local Equilibrium . . . 496
8.3 Boltzmann Transport Theory . . . 498
8.3.1 Derivation of the Boltzmann equation . . . 498
8.3.2 Collisionless Boltzmann equation . . . 500
8.3.3 Scattering processes . . . 501
8.3.4 Collisional invariants . . . 503
8.3.5 Detailed balance . . . 504
8.3.6 Kinematics and cross section . . . 506
8.3.7 $\mathcal{H}$-theorem . . . 506
8.4 Weakly Inhomogeneous Gas . . . 509
8.5 Relaxation Time Approximation . . . 511
8.5.1 Approximation of collision integral . . . 511
8.5.2 Computation of the scattering time . . . 512
8.5.3 Thermal conductivity . . . 513
8.5.4 Viscosity . . . 514
8.5.5 Oscillating external force . . . 517
8.5.6 Quick and dirty calculation of transport coefficients . . . 517
8.5.7 Thermal diffusivity, kinematic viscosity, and Prandtl number . . . 518
8.6 Diffusion and the Lorentz model . . . 519
8.6.1 Failure of the relaxation time approximation . . . 519
8.6.2 Modified Boltzmann equation and its solution . . . 520
8.7 Linearized Boltzmann Equation . . . 523
8.7.1 Linearizing the collision integral . . . 523
8.7.2 Linear algebraic properties of $\hat{L}$ . . . 524
8.7.3 Steady state solution to the linearized Boltzmann equation . . . 525
8.7.4 Variational approach . . . 526
8.8 The Equations of Hydrodynamics . . . 529

8.9 Nonequilibrium Quantum Transport . . . 530
8.9.1 Boltzmann equation for quantum systems . . . 530
8.9.2 The Heat Equation . . . 534
8.9.3 Calculation of Transport Coefficients . . . 534
8.9.4 Onsager Relations . . . 535
8.10 Appendix I: The BBGKY Hierarchy and the Boltzmann Equation . . . 537
8.10.1 BBGKY hierarchy . . . 537
8.10.2 Boltzmann equation . . . 539
8.11 Appendix II: Boltzmann Equation and Collisional Invariants . . . 543

**9 Stochastic Processes** **547**
9.1 References . . . 547
9.2 Introduction to Stochastic Processes . . . 548
9.2.1 Diffusion and Brownian motion . . . 548
9.2.2 Langevin equation . . . 549
9.3 Distributions and Functionals . . . 552
9.3.1 Basic definitions . . . 552
9.3.2 Correlations for the Langevin equation . . . 555
9.4 The Fokker-Planck Equation . . . 557
9.4.1 Basic derivation . . . 557
9.4.2 Brownian motion redux . . . 558
9.4.3 Ornstein-Uhlenbeck process . . . 559
9.4.4 Equilibrium distribution . . . 561
9.5 Formal Theory of Stochastic Processes . . . 561
9.5.1 Markov processes . . . 562
9.5.2 Martingales . . . 564
9.5.3 Differential Chapman-Kolmogorov equations . . . 565
9.5.4 Stationary Markov processes and ergodic properties . . . 568
9.5.5 Approach to stationary solution . . . 569

9.6 Appendix I: Nonlinear diffusion . . . 571
9.6.1 PDEs with infinite propagation speed . . . 571
9.6.2 The porous medium and $p$-Laplacian equations . . . 572
9.6.3 Illustrative solutions . . . 573
9.7 Appendix II: Langevin equation for a particle in a harmonic well . . . 576
9.8 Appendix III: General Linear Autonomous Inhomogeneous ODEs . . . 577
9.8.1 Solution by Fourier transform . . . 577
9.8.2 Solution by Laplace transform . . . 579
9.8.3 Higher order ODEs . . . 581
9.8.4 Kramers-Krönig relations . . . 585
9.9 Appendix IV: Method of Characteristics . . . 586
9.9.1 Quasilinear partial differential equations . . . 586
9.9.2 Example . . . 587

# List of Tables


2.1 Specific heat of some common substances . . . 44
2.2 Performances of real heat engines . . . 58
2.3 Van der Waals parameters for some common gases . . . 80
2.4 Latent heats of fusion and vaporization at $p = 1$ atm . . . 104
2.5 Enthalpies of formation of some common substances . . . 116
2.6 Average bond enthalpies for some common bonds . . . 118

3.1 States and their degeneracies. . . . 142
3.2 Comparison of time and microcanonical averages (I) . . . 186
3.3 Comparison of time and microcanonical averages (II) . . . 188

4.1 Rotational and vibrational temperatures of common molecules . . . 230
4.2 Nuclear angular momentum states for homonuclear diatomic molecules . . . 236

5.1 Debye temperatures and melting points for some common elements . . . 272

6.1 Exact, Percus-Yevick, and hypernetted chains results for hard spheres . . . 375

7.1 van der Waals parameters for some common gases . . . 404
7.2 Critical exponents . . . 422

8.1 Viscosities, thermal conductivities, and Prandtl numbers for some common gases . . . 519

# List of Figures


1 My father and my dog. . . . . 1

1.1 The falling ball novelty . . . . . 5

1.2 Approaching the Gaussian distribution as $N \to \infty$ . . . . . 7

2.1 Microscale to macroscale in physics *vs.* social sciences . . . . . 33

2.2 Gas pressure as a space-time averaged quantity . . . . . 34

2.3 Two distinct paths with identical endpoints . . . . . 35

2.4 The constant volume gas thermometer . . . . . 39

2.5 The phase diagram of $H_2O$ . . . . . 40

2.6 Convergence of measurements in the gas thermometer . . . . . 41

2.7 The first law of thermodynamics . . . . . 43

2.8 $C_V$ for one mole of $H_2$ gas . . . . . 46

2.9 Molar heat capacities $c_V$ for three solids . . . . . 47

2.10 Adiabatic free expansion . . . . . 50

2.11 A perfect engine . . . . . 51

2.12 Heat engine and refrigerator . . . . . 52

2.13 A wonder engine driving a Carnot refrigerator . . . . . 54

2.14 The Carnot cycle . . . . . 55

2.15 Decomposition of a thermodynamic loop into Carnot cycles . . . . . 59

2.16 Check for instability . . . . . 75

2.17 Adiabatic free expansion via a thermal path . . . . . 78

2.18 Throttling . . . 85
2.19 Temperature inversion for the van der Waals gas . . . 86
2.20 Phase diagrams of single component systems . . . 87
2.21 $p(v,T)$ surface for the ideal gas equation of state . . . 88
2.22 A $p$-$v$-$T$ surface for a substance which contracts upon freezing . . . 90
2.23 $p$-$v$-$T$ surfaces for a substance which expands upon freezing . . . 91
2.24 Phase diagram for $CO_2$ in the $(p,T)$ plane . . . 91
2.25 Projection of the $p$-$v$-$T$ surface onto the $(v,p)$ plane . . . 93
2.26 Surface melting data for ice . . . 93
2.27 Multicomponent system . . . 98
2.28 Mixing among three different species of particles . . . 100
2.29 Osmotic pressure . . . 102
2.30 Gibbs free energy for a binary solution . . . 105
2.31 Binary systems: chemical potential shifts and $(T,\mu)$ phase diagrams . . . 107
2.32 Phase diagram for the binary system in the $(x,T)$ plane . . . 108
2.33 Gibbs free energy for an ideal binary solution . . . 109
2.34 Phase diagrams for azeotropes . . . 110
2.35 Binary fluid mixture in equilibrium with a vapor . . . 111
2.36 Eutectic phase diagram . . . 112
2.37 Reaction enthalpy . . . 117
2.38 Reaction enthalpy for hydrogenation of ethene . . . 119
2.39 The Stirling cycle . . . 122
2.40 The Otto cycle . . . 123
2.41 The Diesel cycle . . . 125
2.42 The Joule-Brayton cycle . . . 126
2.43 The Legendre transformation . . . 128

3.1 Time evolution of two immiscible fluids . . . 146
3.2 Projected evolution of a set $\mathcal{R}_0$ under the mapping $g_\tau$ . . . 150

3.3 Poincaré recurrence . . . 151
3.4 The Kac ring model . . . 152
3.5 Simulation of the Kac ring model . . . 153
3.6 More simulations of the Kac ring model . . . 154
3.7 An ergodic flow which is not mixing . . . 157
3.8 The baker's transformation . . . 158
3.9 Multiply iterated baker's transformation . . . 159
3.10 Arnold's cat map . . . 160
3.11 Hierarchy of dynamical systems . . . 161
3.12 Three state Markovian Mpemba effect . . . 179
3.13 Evolution of the distribution in the MME . . . 180
3.14 Poincaré sections for the ball and piston problem . . . 185
3.15 Long time averages of $X_{\rm av}(t) \equiv t^{-1} \int_0^t dt'\, X(t')$ . . . 186

4.1 Complex integration contours $\mathcal{C}$ for inverse Laplace transform . . . 192
4.2 Two systems in thermal contact . . . 196
4.3 A system S in contact with a world W . . . 198
4.4 Averaging the quantum mechanical discrete density of states . . . 201
4.5 Microscopic interpretation of the First Law of Thermodynamics . . . 208
4.6 Maxwell distribution of speeds $\varphi(v/v_0)$ . . . 219
4.7 Entropy *versus* energy for a spin system . . . 223
4.8 A crude model for the elasticity of wool . . . 226
4.9 Length *versus* temperature and force . . . 227
4.10 Heat capacity per molecule as a function of temperature . . . 234
4.11 Noninteracting spin dimers on a lattice . . . 239

5.1 Occupation number *versus* $\varepsilon - \mu$ for Bose and Fermi distributions . . . 250
5.2 Partitions of bosonic occupation states . . . 252
5.3 Spectral density of blackbody radiation . . . 262

5.4 A linear chain of masses and springs . . . 264
5.5 Crystal structure, Bravais lattice, and basis . . . 266
5.6 Phonon spectra . . . 267
5.7 The polylogarithm function $\mathrm{Li}_s(z)$ . . . 277
5.8 Molar heat capacity of the ideal Bose gas . . . 280
5.9 Phase diagrams for the ideal Bose gas . . . 283
5.10 Fugacity *versus* dimensionless density for the mean field interacting Bose gas . . . 285
5.11 Phase diagram of $^4$He . . . 286
5.12 Specific heat of liquid $^4$He in the vicinity of the $\lambda$-transition . . . 287
5.13 The fountain effect . . . 288
5.14 Strange properties of superfluid $^4$He . . . 290
5.15 Fermi surfaces for two and three-dimensional structures . . . 295
5.16 Deformation of the complex integration contour in eqn. 5.235 . . . 298
5.17 Fermi distributions in the presence of a magnetic field . . . 302
5.18 A graduate student experiences the Stoner enhancement . . . 308
5.19 Mean field phase diagram of the Hubbard model . . . 311
5.20 Mass-radius relationship for white dwarf stars . . . 314
5.21 Solutions to the Lane-Emden equation . . . 316

6.1 Domain walls in the $d = 2$ Ising model . . . 332
6.2 Clusters and boundaries for the square lattice Ising model . . . 334
6.3 A two-dimensional square lattice mapped onto a one-dimensional chain . . . 336
6.4 High temperature expansion diagrams . . . 338
6.5 High temperature expansion for the correlation function . . . 339
6.6 The Lennard-Jones potential . . . 344
6.7 Diagrams and the Mayer cluster expansion . . . 345
6.8 Vertex labels in the configuration integral . . . 346
6.9 Symmetry factors for cluster diagrams . . . 347
6.10 Connected *versus* irreducible clusters . . . 350

6.11 The overlap of hard sphere Mayer functions . . . 351
6.12 Phase diagram of the $d = 3$ hard sphere gas . . . 352
6.13 Mayer function for an attractive spherical well with a repulsive core . . . 354
6.14 Density of hard spheres in the presence of a hard wall . . . 356
6.15 Hard sphere pair distribution functions: simulation and experiment . . . 361
6.16 Monte Carlo pair distribution functions for liquid water . . . 362
6.17 Elastic and inelastic scattering . . . 365
6.18 Static structure factor of the Lennard-Jones fluid . . . 367
6.19 The Thomas-Fermi atom . . . 382
6.20 Some examples of linear chain polymers . . . 383
6.21 *trans* and *gauche* orientations in carbon chains . . . 384
6.22 The polymer chain as a random coil . . . 386
6.23 Radius of gyration $R_{\mathrm{g}}$ of polystyrene in a toluene and benzene solvent . . . 391
6.24 Singularities of the partition function . . . 394
6.25 Fugacity $z$ and $pv_0/k_{\mathrm{B}}T$ *versus* dimensionless specific volume $v/v_0$ . . . 397

7.1 Pressure *versus* volume for the van der Waals gas . . . 403
7.2 Molar free energy $f(T, v)$ of the van der Waals system . . . 405
7.3 Pressure-volume isotherms for the van der Waals system, with Maxwell construction . . . 406
7.4 Maxwell construction in the $(v, p)$ plane . . . 407
7.5 Universality of the liquid-gas transition . . . 409
7.6 The lattice gas model . . . 412
7.7 Comparison of liquid-gas and Ising ferromagnet phase diagrams . . . 413
7.8 Order-disorder transition on the square lattice . . . 416
7.9 Ising mean field theory at $h = 0$ . . . 419
7.10 Ising mean field theory at $h = 0.1$ . . . 421
7.11 Dissipative magnetization dynamics $\dot{m} = -f'(m)$ . . . 424
7.12 Magnetization hysteresis . . . 426
7.13 Phase diagram for the quartic Landau free energy . . . 430

7.14 Behavior of the quartic free energy $f(m) = \frac{1}{2}am^2 - \frac{1}{3}ym^3 + \frac{1}{4}bm^4$ . . . 432
7.15 Fixed points for $\varphi(u) = \frac{1}{2}ru^2 - \frac{1}{3}u^3 + \frac{1}{4}u^4$ and flow $\dot{u} = -\varphi'(u)$ . . . 433
7.16 Behavior of the sextic free energy $f(m) = \frac{1}{2}am^2 + \frac{1}{4}bm^4 + \frac{1}{6}cm^6$ . . . 435
7.17 Sextic free energy $\varphi(u) = \frac{1}{2}ru^2 - \frac{1}{4}u^4 + \frac{1}{6}u^6$ for different values of $r$ . . . 436
7.18 Fixed point branches $\varphi'(u^*) = 0$ under the flow $\dot{u} = -\varphi'(u)$ for the sextic potential . . . 437
7.19 Variational field free energy *versus* magnetization . . . 441
7.20 Variational mean field theory of the $q = 7$ state Potts model . . . 445
7.21 Domain walls in the one, two and three dimensional Ising model . . . 462
7.22 A domain wall in an $XY$ ferromagnet . . . 463
7.23 Imry-Ma domains and free energy *versus* domain size . . . 465
7.24 Evolution of $G(x, \alpha)$ as a function of the control parameter $\alpha$ . . . 469
7.25 Flow diagrams for $\dot{u} = r + u^2$ and $\dot{u} = ru - u^2$ . . . 470
7.26 Extended phase space flow for the saddle-node and transcritical bifurcations . . . 471
7.27 Supercritical and subcritical pitchfork bifurcations . . . 472
7.28 Extended phase space flow for supercritical and subcritical pitchfork bifurcations . . . 472
7.29 Scaled free energy and phase diagram . . . 473
7.30 Imperfect pitchfork bifurcation . . . 474
7.31 Mean field phase diagram for the Blume-Capel model . . . 480
7.32 Mean field solution for an Ising antiferromagnet in an external field . . . 482
7.33 Mean field phase diagram for an Ising antiferromagnet in an external field . . . 483
7.34 Mean field phase diagram for the model of eqn. 7.356 . . . 485
7.35 Phase diagram for the model of eqn. 7.368 . . . 489

8.1 Level sets for a sample $f(\bar{x}, \bar{p}, \bar{t}) = C\, e^{-\frac{1}{2}(\bar{x}-\bar{p}\bar{t})^2}\, e^{-\frac{1}{2}\bar{p}^2}$ . . . 501
8.2 One and two particle scattering processes . . . 503
8.3 Graphic representation of the equation $n\,\sigma\,\bar{v}_{\rm rel}\,\tau = 1$ . . . 512
8.4 *Gedankenexperiment* to measure shear viscosity $\eta$ in a fluid . . . 515
8.5 Experimental data on thermal conductivity and shear viscosity . . . 516
8.6 Scattering in the center of mass frame . . . 527

8.7 Thermocouple and Peltier effect . . . . . 533

9.1 Discretization of a continuous function $\eta(t)$ . . . . . 553
9.2 Interpretation of the Chapman-Kolmogorov equation . . . . . 558
9.3 Wiener and Cauchy processes . . . . . 564
9.4 Repeating unit of a transmission line . . . . . 573
9.5 Diffusion and porous medium equations . . . . . 575

## 0.1 Preface

This is a proto-preface. A more complete preface will be written after these notes are completed.

These lecture notes are intended to supplement a course in statistical physics at the upper division undergraduate or beginning graduate level.

I was fortunate to learn this subject from one of the great statistical physicists of our time, John Cardy.

I am grateful to my wife Joyce and to my children Ezra and Lily for putting up with all the outrageous lies I've told them about getting off the computer 'in just a few minutes' while working on these notes.

These notes are dedicated to the only two creatures I know who are never angry with me: my father and my dog.



Figure 1: My father (Louis) and my dog (Henry).

## 0.2 General references

- L. Peliti, *Statistical Mechanics in a Nutshell* (Princeton University Press, 2011)
  The best all-around book on the subject I've come across thus far. Appropriate for the graduate or advanced undergraduate level.
- J. P. Sethna, *Entropy, Order Parameters, and Complexity* (Oxford, 2006)
  An excellent introductory text with a very modern set of topics and exercises. Available online at http://www.physics.cornell.edu/sethna/StatMech
- M. Kardar, *Statistical Physics of Particles* (Cambridge, 2007)
  A superb modern text, with many insightful presentations of key concepts.
- M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* ($3^{\rm rd}$ edition, World Scientific, 2006)
  An excellent graduate level text. Less insightful than Kardar but still a good modern treatment of the subject. Good discussion of mean field theory.
- E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics* (part I, $3^{\rm rd}$ edition, Pergamon, 1980)
  This is volume 5 in the famous Landau and Lifshitz *Course of Theoretical Physics*. Though dated, it still contains a wealth of information and physical insight.
- F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, 1987)
  This has been perhaps the most popular undergraduate text since it first appeared in 1967, and with good reason.

# Chapter 1

# Fundamentals of Probability

## 1.1 References

- C. Gardiner, *Stochastic Methods* ($4^{\text{th}}$ edition, Springer-Verlag, 2010)
  Very clear and complete text on stochastic methods with many applications.
- J. M. Bernardo and A. F. M. Smith, *Bayesian Theory* (Wiley, 2000)
  A thorough textbook on Bayesian methods.
- D. Williams, *Weighing the Odds: A Course in Probability and Statistics* (Cambridge, 2001)
  A good overall statistics textbook, according to a mathematician colleague.
- E. T. Jaynes, *Probability Theory* (Cambridge, 2007)
  An extensive, descriptive, and highly opinionated presentation, with a strongly Bayesian approach.
- A. N. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea, 1956)
  The *Urtext* of mathematical probability theory.

## 1.2 Statistical Properties of Random Walks

### 1.2.1 One-dimensional random walk

Consider the mechanical system depicted in fig. 1.1, a version of which is often sold in novelty shops. A ball is released from the top, which cascades consecutively through $N$ levels. The details of each ball's motion are governed by Newton's laws of motion. However, to predict where any given ball will end up in the bottom row is difficult, because the ball's trajectory depends sensitively on its initial conditions, and may even be influenced by random vibrations of the entire apparatus. We therefore abandon all hope of integrating the equations of motion and treat the system statistically. That is, we assume, at each level, that the ball moves to the right with probability $p$ and to the left with probability $q = 1 - p$. If there is no bias in the system, then $p = q = \frac{1}{2}$. The position $X_N$ after $N$ steps may be written

$$X_N = \sum_{j=1}^{N} \sigma_j \quad , \tag{1.1}$$

where $\sigma_j = +1$ if the ball moves to the right at level $j$, and $\sigma_j = -1$ if the ball moves to the left at level $j$. At each level, the probability for these two outcomes is given by

$$P_\sigma = p\,\delta_{\sigma,+1} + q\,\delta_{\sigma,-1} = \begin{cases} p & \text{if } \sigma = +1 \\ q & \text{if } \sigma = -1 \end{cases} \quad . \tag{1.2}$$

This is a normalized discrete probability distribution of the type discussed in section 1.5 below. The multivariate distribution for all the steps is then

$$P(\sigma_1\,,\,\ldots\,,\sigma_N) = \prod_{j=1}^{N} P(\sigma_j) \quad . \tag{1.3}$$

Our system is equivalent to a one-dimensional *random walk*. Imagine an inebriated pedestrian on a sidewalk taking steps to the right and left at random. After $N$ steps, the pedestrian's location is $X$.

Now let's compute the average of $X$:

$$\langle X \rangle = \Big\langle \sum_{j=1}^{N} \sigma_j \Big\rangle = N\langle\sigma\rangle = N \sum_{\sigma=\pm1} \sigma\, P(\sigma) = N(p-q) = N(2p-1) \quad . \tag{1.4}$$

This could be identified as an *equation of state* for our system, as it relates a measurable quantity $X$ to the number of steps $N$ and the local bias $p$. Next, let's compute the average of $X^2$:

$$\langle X^2 \rangle = \sum_{j=1}^{N} \sum_{j'=1}^{N} \langle \sigma_j \sigma_{j'} \rangle = N^2(p-q)^2 + 4Npq \quad . \tag{1.5}$$

Here we have used

$$\langle \sigma_j \sigma_{j'} \rangle = \delta_{jj'} + \big(1 - \delta_{jj'}\big)(p-q)^2 = \begin{cases} 1 & \text{if } j = j' \\ (p-q)^2 & \text{if } j \neq j' \end{cases} \quad . \tag{1.6}$$

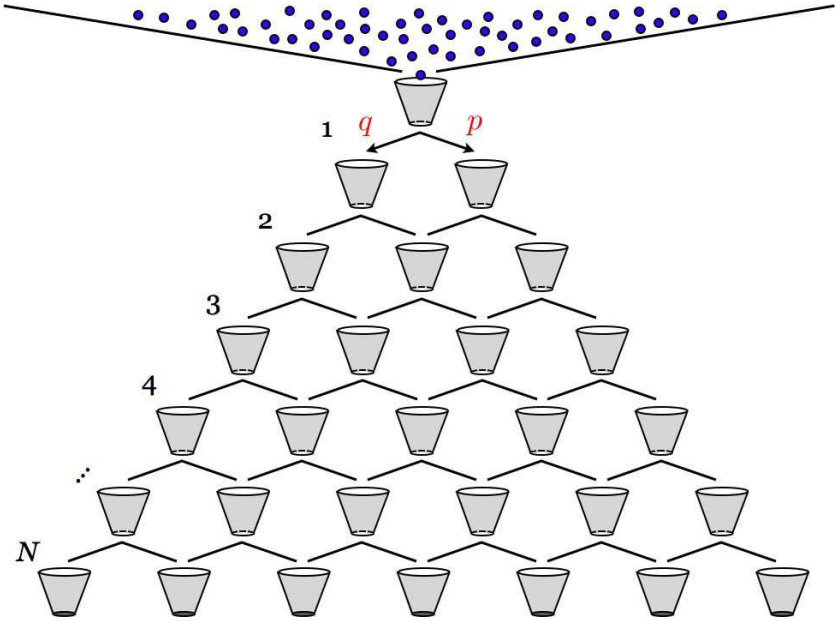



Figure 1.1: The falling ball system, which mimics a one-dimensional random walk.

Note that $\langle X^2\rangle \geq \langle X\rangle^2$, which must be so because

$$\mathrm{Var}(X) = \langle (\Delta X)^2\rangle \equiv \big\langle \big(X - \langle X\rangle\big)^2\big\rangle = \langle X^2\rangle - \langle X\rangle^2 \quad . \tag{1.7}$$

This is called the *variance* of $X$. We have $\mathrm{Var}(X) = 4Npq$. The *root mean square* deviation, $\Delta X_{\rm rms}$, is the square root of the variance: $\Delta X_{\rm rms} = \sqrt{\mathrm{Var}(X)}$. Note that the mean value of $X$ is linearly proportional to $N$ (for all cases other than $p = q = \frac{1}{2}$), but the RMS fluctuations $\Delta X_{\rm rms}$ are proportional to $N^{1/2}$. In the limit $N \to \infty$ then, the ratio $\Delta X_{\rm rms}/\langle X\rangle$ vanishes as $N^{-1/2}$. This is a consequence of the central limit theorem (see §1.5.2 below), and we shall meet up with it again on several occasions.

We can do even better. We can find the complete probability distribution for $X$. It is given by

$$P_{N,X} = \binom{N}{N_{\rm R}} p^{N_{\rm R}}\, q^{N_{\rm L}} \quad , \tag{1.8}$$

where $N_{\rm R/L}$ are the numbers of steps taken to the right/left, with $N = N_{\rm R} + N_{\rm L}$, and $X = N_{\rm R} - N_{\rm L}$. There are many independent ways to take $N_{\rm R}$ steps to the right. For example, our first $N_{\rm R}$ steps could all be to the right, and the remaining $N_{\rm L} = N - N_{\rm R}$ steps would then all be to the left. Or our final $N_{\rm R}$ steps could all be to the right. For each of these independent possibilities, the probability is $p^{N_{\rm R}}\, q^{N_{\rm L}}$. How many possibilities are there? Elementary combinatorics tells us this number is

$$\binom{N}{N_{\rm R}} = \frac{N!}{N_{\rm R}!\, N_{\rm L}!} \quad . \tag{1.9}$$

Note that $N \pm X = 2N_{\rm R/L}$, so we can replace $N_{\rm R/L} = \frac{1}{2}(N \pm X)$. Thus,

$$P_{N,X} = \frac{N!}{\left(\frac{N+X}{2}\right)!\,\left(\frac{N-X}{2}\right)!}\, p^{(N+X)/2}\, q^{(N-X)/2} \quad . \tag{1.10}$$

### 1.2.2 Thermodynamic limit

Consider the limit $N \to \infty$ but with $x \equiv X/N$ finite. This is analogous to what is called the *thermodynamic limit* in statistical mechanics. Since $N$ is large, $x$ may be considered a continuous variable. We evaluate $\log P_{N,X}$ using Stirling's asymptotic expansion

$$\log N! \simeq N \log N - N + \mathcal{O}(\log N) \quad . \tag{1.11}$$

We then have

$$\begin{aligned} \log P_{N,X} &\simeq N \log N - N - \tfrac{1}{2}N(1+x)\log\Big[\tfrac{1}{2}N(1+x)\Big] + \tfrac{1}{2}N(1+x) \\ &\quad - \tfrac{1}{2}N(1-x)\log\Big[\tfrac{1}{2}N(1-x)\Big] + \tfrac{1}{2}N(1-x) + \tfrac{1}{2}N(1+x)\,\log p + \tfrac{1}{2}N(1-x)\,\log q \\ &= -N\Big[\big(\tfrac{1+x}{2}\big)\log\big(\tfrac{1+x}{2}\big) + \big(\tfrac{1-x}{2}\big)\log\big(\tfrac{1-x}{2}\big)\Big] + N\Big[\big(\tfrac{1+x}{2}\big)\log p + \big(\tfrac{1-x}{2}\big)\log q\Big] \quad . \end{aligned} \tag{1.12}$$

Notice that the terms proportional to $N \log N$ have all cancelled, leaving us with a quantity which is linear in $N$. We may therefore write $\log P_{N,X} = -N f(x) + \mathcal{O}(\log N)$, where

$$f(x) = \Big[\big(\tfrac{1+x}{2}\big)\log\big(\tfrac{1+x}{2}\big) + \big(\tfrac{1-x}{2}\big)\log\big(\tfrac{1-x}{2}\big)\Big] - \Big[\big(\tfrac{1+x}{2}\big)\log p + \big(\tfrac{1-x}{2}\big)\log q\Big] \quad . \tag{1.13}$$

We have just shown that in the large $N$ limit we may write

$$P_{N,X} \approx \mathcal{C}\, e^{-N f(X/N)} \quad , \tag{1.14}$$

where $\mathcal{C}$ is a normalization factor[1]. Since $N$ is by assumption large, the function $P_{N,X}$ is dominated by the minimum (or minima) of $f(x)$, where the probability is maximized. To find the minimum of $f(x)$, we set $f'(x) = 0$, where

$$f'(x) = \tfrac{1}{2}\log\left(\frac{q}{p}\cdot\frac{1+x}{1-x}\right) \quad . \tag{1.15}$$

Setting $f'(x) = 0$, we obtain

$$\frac{1+x}{1-x} = \frac{p}{q} \quad \Rightarrow \quad \bar{x} = p - q \quad . \tag{1.16}$$

We also have

$$f''(x) = \frac{1}{1-x^2} \quad , \tag{1.17}$$

so invoking Taylor's theorem,

$$f(x) = f(\bar{x}) + \tfrac{1}{2}f''(\bar{x})\,(x-\bar{x})^2 + \ldots \quad . \tag{1.18}$$

Putting it all together, we have

$$P_{N,X} \approx \mathcal{C}\,\exp\left[-\frac{N(x-\bar{x})^2}{8pq}\right] = \mathcal{C}\,\exp\left[-\frac{(X-\bar{X})^2}{8Npq}\right] \quad , \tag{1.19}$$

[1]The origin of $\mathcal{C}$ lies in the $\mathcal{O}(\log N)$ and $\mathcal{O}(N^0)$ terms in the asymptotic expansion of $\log N!$ – these terms will in general also depend on $x = X/N$. We have ignored these terms here. Accounting for them carefully reproduces the correct value of $\mathcal{C}$ in eqn. 1.20.

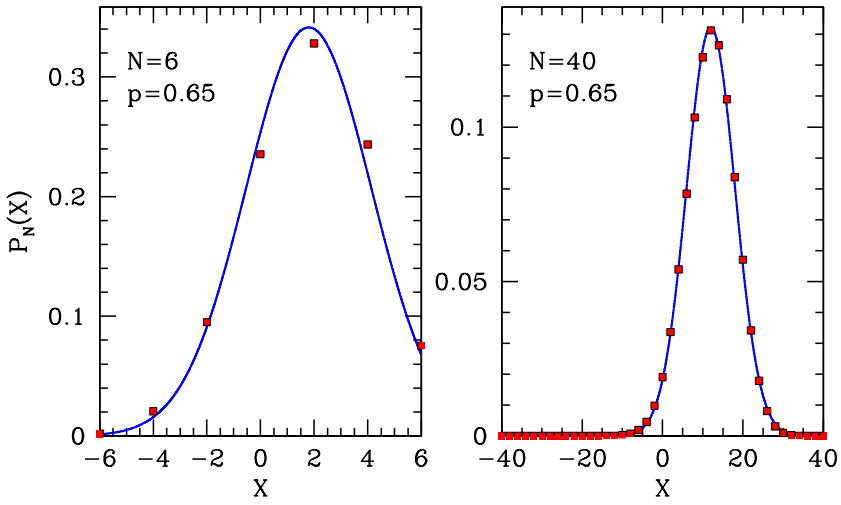



Figure 1.2: Comparison of exact distribution of eqn. 1.10 (red squares) with the Gaussian distribution of eqn. 1.19 (blue line).

where $\bar{X} = \langle X \rangle = N(p-q) = N\bar{x}$. The constant $\mathcal{C}$ is determined by the normalization condition,

$$\sum_{X=-\infty}^{\infty} P_{N,X} \approx \tfrac{1}{2}\mathcal{C} \int_{-\infty}^{\infty} dX \ \exp\left[ -\frac{(X-\bar{X})^2}{8Npq} \right] = \sqrt{2\pi Npq}\,\mathcal{C} \quad , \tag{1.20}$$

and thus $\mathcal{C} = 1/\sqrt{2\pi Npq}$. Why don't we go beyond second order in the Taylor expansion of $f(x)$? We will find out in §1.5.2 below.

### 1.2.3 Entropy and energy

The function $f(x)$ can be written as a sum of two contributions, $f(x) = e(x) - s(x)$, where

$$\begin{aligned} s(x) &= -\left(\tfrac{1+x}{2}\right)\log\left(\tfrac{1+x}{2}\right) - \left(\tfrac{1-x}{2}\right)\log\left(\tfrac{1-x}{2}\right) \\ e(x) &= -\tfrac{1}{2}(1+x)\log p - \tfrac{1}{2}(1-x)\log q \quad . \end{aligned} \tag{1.21}$$

The function $S(N,x) \equiv Ns(x)$ is analogous to the *statistical entropy* of our system, and $E(N,x) \equiv Ne(x)$ to the *energy* of the system[2]. The statistical entropy is the logarithm of the number of ways, at fixed $N$, that the system can be configured so as to yield the same value of $X$. The energy biases the probability $P_{N,X} = \exp(S-E)$ so that low energy configurations are more probable than high energy configurations. For our system, we see that when $p < \frac{1}{2}$ the energy is minimized by taking $x$ as small as possible, *i.e.* $x = -1$. Conversely, when $p > \frac{1}{2}$ the energy is minimized by taking $x$ as large as possible, *i.e.* $x = +1$. The average value of $x$, as we have computed explicitly, is $\bar{x} = p - q = 2p - 1$, which falls somewhere in between these two extremes.

[2]The functions $s(x)$ and $e(x)$ are the *specific entropy* and *specific energy*, respectively.

In actual thermodynamic systems, entropy and energy are not dimensionless. What we have called $S$ here is really $S/k_{\rm B}$, which is the entropy in units of Boltzmann's constant. And what we have called $E$ here is really $E/k_{\rm B}T$, which is energy in units of Boltzmann's constant times temperature.

## 1.3 Basic Concepts in Probability Theory

Here we recite the basics of probability theory.

### 1.3.1 Fundamental definitions

The natural mathematical setting is set theory. *Sets* are generalized collections of *objects*. The basics: $\omega \in A$ is a binary relation which says that the object $\omega$ is an *element* of the set $A$. Another binary relation is *set inclusion*. If all members of $A$ are in $B$, we write $A \subseteq B$. The *union* of sets $A$ and $B$ is denoted $A \cup B$ and the *intersection* of $A$ and $B$ is denoted $A \cap B$. The *Cartesian product* of $A$ and $B$, denoted $A \times B$, is the set of all ordered pairs $(a, b)$ where $a \in A$ and $b \in B$.

Some details: If $\omega$ is not in $A$, we write $\omega \notin A$. Sets may also be objects, so we may speak of sets of sets, but typically the sets which will concern us are simple discrete collections of numbers, such as the possible rolls of a die {1,2,3,4,5,6}, or the real numbers $\mathbb{R}$, or Cartesian products such as $\mathbb{R}^N$. If $A \subseteq B$ but $A \neq B$, we say that $A$ is a *proper subset* of $B$ and write $A \subset B$. Another binary operation is the *set difference* $A \backslash B$, which contains all $\omega$ such that $\omega \in A$ and $\omega \notin B$.

In probability theory, each object $\omega$ is identified as an *event*. We denote by $\Omega$ the set of all events, and $\emptyset$ denotes the set of no events. There are three basic axioms of probability:

i) To each set $A$ is associated a non-negative real number $P(A)$, which is called the probability of $A$.

ii) $P(\Omega) = 1$.

iii) If $\{A_i\}$ is a countable collection of disjoint sets, *i.e.* if $A_i \cap A_j = \emptyset$ for all $i \neq j$, then

$$P\Big(\bigcup_i A_i\Big) = \sum_i P(A_i) \quad . \tag{1.22}$$

From these axioms follow a number of conclusions. Among them, let $\neg A = \Omega \backslash A$ be the *complement* of $A$, *i.e.* the set of all events *not* in $A$. Then since $A \cup \neg A = \Omega$, we have $P(\neg A) = 1 - P(A)$. Taking $A = \Omega$, we conclude $P(\emptyset) = 0$.

The operational definition of $P(A)$ is that if events $\omega$ are chosen from $\Omega$ *at random*[3], then the relative frequency for $\omega \in A$ approaches $P(A)$ as the number of trials tends to infinity. But what do we mean by 'at random'? One meaning we can impart to the notion of randomness is that a process is random if its outcomes can be accurately modeled using the axioms of probability. This entails the identification

[3]We assume independent and identically distributed trials, and that the law of large numbers holds.

of a *probability space* $\Omega$ as well as a *probability measure* $P$. For example, in the microcanonical ensemble of classical statistical physics, the space $\Omega$ is the collection of phase space points $\boldsymbol{\varphi} = \{q_1, \ldots, q_n, p_1, \ldots, p_n\}$ and the probability measure is $d\mu = \Sigma^{-1}(E) \prod_{i=1}^n dq_i\, dp_i\, \delta\big(E - H(q,p)\big)$, so that for $A \subseteq \Omega$ the probability of $A$ is $P(A) = \int d\mu\, \chi_{_A}(\boldsymbol{\varphi})$, where $\chi_{_A}(\boldsymbol{\varphi}) = 1$ if $\boldsymbol{\varphi} \in A$ and $\chi_{_A}(\boldsymbol{\varphi}) = 0$ if $\boldsymbol{\varphi} \notin A$ is the *characteristic function* of $A$. The quantity $\Sigma(E)$ is determined by normalization: $\int d\mu = 1$.

### 1.3.2 Bayesian statistics

We now introduce two additional probabilities. The *joint probability* for sets $A$ and $B$ together is written $P(A \cap B)$. That is, $P(A \cap B) = \mathsf{Prob}[\omega \in A \text{ and } \omega \in B]$. For example, $A$ might denote the set of all politicians, $B$ the set of all American citizens, and $C$ the set of all living humans who are allergic to cats. Then $A \cap B$ would be the set of all politicians who are also American citizens, *etc.*

The *conditional probability* of $B$ *given* $A$ is written $P(B \,|\, A)$. The joint probability $P(A \cap B) = P(B \cap A)$ may be expressed in two ways:

$$P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A) \quad . \tag{1.23}$$

Thus,

$$P(A \,|\, B) = \frac{P(B \,|\, A)\, P(A)}{P(B)} \quad , \tag{1.24}$$

a result known as *Bayes' theorem*. Now suppose the 'event space' is partitioned as $\{A_i\}$. Then

$$P(B) = \sum_i P(B|A_i)\, P(A_i) \quad . \tag{1.25}$$

We then have

$$P(A_i \,|\, B) = \frac{P(B \,|\, A_i)\, P(A_i)}{\sum_j P(B \,|\, A_j)\, P(A_j)} \quad , \tag{1.26}$$

a result sometimes known as the *extended form of Bayes' theorem*. When the event space is a 'binary partition' $\{A, \neg A\}$, we have

$$P(A \,|\, B) = \frac{P(B \,|\, A)\, P(A)}{P(B \,|\, A)\, P(A) + P(B \,|\, \neg A)\, P(\neg A)} \quad . \tag{1.27}$$

Note that $P(A \,|\, B) + P(\neg A \,|\, B) = 1$ (which follows from $\neg\neg A = A$).

As an example, consider the following problem in epidemiology. Suppose there is a rare but highly contagious disease $A$ which occurs in $0.01\%$ of the general population. Suppose further that there is a simple test for the disease which is accurate $99.99\%$ of the time. That is, out of every 10,000 tests, the correct answer is returned 9,999 times, and the incorrect answer is returned only once. Now let us administer the test to a large group of people from the general population. Those who test positive are quarantined. Question: what is the probability that someone chosen at random from the quarantine group actually has the disease? We use Bayes' theorem with the binary partition $\{A, \neg A\}$. Let $B$ denote the event that an individual tests positive. Anyone from the quarantine group has tested positive. Given

this datum, we want to know the probability that that person has the disease. That is, we want $P(A \mid B)$. Applying eqn. 1.27 with

$$P(A) = 0.0001 \quad , \quad P(\neg A) = 0.9999 \quad , \quad P(B \mid A) = 0.9999 \quad , \quad P(B \mid \neg A) = 0.0001 \quad ,$$

we find $P(A \mid B) = \frac{1}{2}$. That is, there is only a $50\%$ chance that someone who tested positive actually has the disease, despite the test being $99.99\%$ accurate! The reason is that, given the rarity of the disease in the general population, the number of false positives is statistically equal to the number of true positives.

In the above example, we had $P(B \mid A) + P(B \mid \neg A) = 1$, but this is not generally the case. What is true instead is $P(B \mid A) + P(\neg B \mid A) = 1$. Epidemiologists define the *sensitivity* of a binary classification test as the fraction of actual positives which are correctly identified, and the *specificity* as the fraction of actual negatives that are correctly identified. Thus, $\mathsf{se} = P(B \mid A)$ is the sensitivity and $\mathsf{sp} = P(\neg B \mid \neg A)$ is the specificity. We then have $P(B \mid \neg A) = 1 - P(\neg B \mid \neg A)$. Therefore,

$$P(B \mid A) + P(B \mid \neg A) = 1 + P(B \mid A) - P(\neg B \mid \neg A) = 1 + \mathsf{se} - \mathsf{sp} \quad . \tag{1.28}$$

In our previous example, $\mathsf{se} = \mathsf{sp} = 0.9999$, in which case the RHS above gives 1. In general, if $P(A) \equiv f$ is the fraction of the population which is afflicted, then

$$P(\mathsf{infected} \mid \mathsf{positive}) = \frac{f \cdot \mathsf{se}}{f \cdot \mathsf{se} + (1-f) \cdot (1 - \mathsf{sp})} \quad . \tag{1.29}$$

For continuous distributions, we speak of a probability *density*. We then have

$$P(y) = \int dx \, P(y \mid x) \, P(x) \tag{1.30}$$

and

$$P(x \mid y) = \frac{P(y \mid x) \, P(x)}{\int dx' \, P(y \mid x') \, P(x')} \quad . \tag{1.31}$$

The range of integration may depend on the specific application.

The quantities $P(A_i)$ are called the *prior distribution*. Clearly in order to compute $P(B)$ or $P(A_i \mid B)$ we must know the priors, and this is usually the weakest link in the Bayesian chain of reasoning. If our prior distribution is not accurate, Bayes' theorem will generate incorrect results. One approach to approximating prior probabilities $P(A_i)$ is to derive them from a *maximum entropy construction*.

### 1.3.3 Random variables and their averages

Consider an abstract probability space $\mathcal{X}$ whose elements (*i.e.* events) are labeled by $x$. A *random variable* $X$ can take values $x \in \mathcal{X}$.[4] The average of a function $f(X)$ of the random variable $X$ is denoted as $\mathbb{E}f(X)$ or $\langle f(X) \rangle$, and is defined for discrete sets as

$$\mathbb{E}f(X) = \langle f(X) \rangle = \sum_{x \in \mathcal{X}} f(x) \, P(x) \quad , \tag{1.32}$$

[4] Formally, a random variable $X$ is a measurable function $X : \Omega \to \mathcal{X}$ from the underlying sample space $\Omega$ to a value space $\mathcal{X}$

where $P(x) = \mathsf{Prob}(X = x)$. For continuous sets, we have

$$\mathbb{E}f(X) = \langle f(X)\rangle = \int\limits_{\mathcal{X}} dx\; f(x)\, P(x) \quad , \tag{1.33}$$

where now $P(x)$ is the *probability density* at $x$. We then have that

$$P(x)\, dx = \mathsf{Prob}\big(X \in [x, x + dx]\big) \quad . \tag{1.34}$$

Typically for continuous sets we have $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{R}_{\geq 0}$. While it is formally useful to use symbols such as $X$ and $Y$ for random variables, at times we may slip and loosely speak of $x$ and $y$ as random variables.

When there are two random variables $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, we have $\Omega = \mathcal{X} \times \mathcal{Y}$ is the product space, and

$$\mathbb{E}f(X,Y) = \langle f(X,Y)\rangle = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} f(x,y)\, P(x,y) \quad , \tag{1.35}$$

with the obvious generalization to continuous sets. This generalizes to higher rank product probability spaces, *i.e.* $\Omega = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_N$ with $x_i \in \mathcal{X}_i$ for $i \in \{1, \ldots, N\}$. The *covariance* of $X_i$ and $X_j$ is defined as

$$C_{ij} \equiv \big\langle \big(X_i - \langle X_i\rangle\big)\big(X_j - \langle X_j\rangle\big)\big\rangle = \langle X_i X_j\rangle - \langle X_i\rangle\langle X_j\rangle \quad . \tag{1.36}$$

If $f(x)$ is a convex function then one has

$$\mathbb{E}f(X) \geq f(\mathbb{E}X) \quad . \tag{1.37}$$

If a function $f(x)$ is twice differentiable on the interval $[a, b]$, it is convex if and only if $f''(x) \geq 0$ everywhere[5]. In such cases, then for $x_{1,2} \in [a, b]$ we must have

$$f\big(\lambda x_1 + (1-\lambda)x_2\big) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \quad , \tag{1.38}$$

where $\lambda \in [0, 1]$. This is easily generalized to

$$f\Big(\sum_n p_n\, x_n\Big) \leq \sum_n p_n\, f(x_n) \quad , \tag{1.39}$$

where $p_n = P(x_n)$, a result known as *Jensen's theorem*. For continuous random variables,

$$f\left(\int\limits_{\mathcal{X}} dx\, P(x)\, x\right) \leq \int\limits_{\mathcal{X}} dx\, P(x)\, f(x) \quad , \tag{1.40}$$

[5] A function $g(x)$ is *concave* if $-g(x)$ is convex.

## 1.4 Entropy and Probability

### 1.4.1 Entropy and information theory

It was shown in the classic 1948 work of Claude Shannon that entropy is in fact a measure of *information*[6]. Suppose we observe that a particular event occurs with probability $p$. We associate with this observation an amount of information $I(p)$. The information $I(p)$ should satisfy certain desiderata:

1 Information is non-negative, *i.e.* $I(p) \geq 0$.

2 If two events occur independently so their joint probability is $p_1 p_2$, then their information is additive, *i.e.* $I(p_1 p_2) = I(p_1) + I(p_2)$.

3 $I(p)$ is a continuous function of $p$.

4 There is no information content to an event which is always observed, *i.e.* $I(1) = 0$.

From these four properties, it is easy to show that the only possible function $I(p)$ is

$$I(p) = -A \log p \quad , \tag{1.41}$$

where $A$ is an arbitrary constant that can be absorbed into the base of the logarithm, since $\log_b x = \log x / \log b$. In statistical physics we will take $A = 1$ and use $e$ as the base, so $I(p) = -\log p$. These are so-called *natural units* of information. Another common choice, typical in computer science, is to take the base of the logarithm to be 2, so $I(p) = -\log_2 p$. In this latter case, the units of information are known as *bits*. Note that, in either case, $I(0) = \infty$, which means that the observation of an extremely rare event carries a great deal of information[7].

Now suppose we have a set of events labeled by an integer $n$ which occur with probabilities $\{p_n\}$. What is the expected amount of information in $N$ observations? Since event $n$ occurs an average of $Np_n$ times, and the information content in $p_n$ is $-\log p_n$, we have that the average information per observation is

$$S = \frac{\langle I_N \rangle}{N} = -\sum_n p_n \log p_n \quad , \tag{1.42}$$

which is known as the entropy of the distribution. Thus, maximizing $S$ is equivalent to maximizing the *information* content per observation. $S$ is known as the *Shannon entropy* of the distribution $\{p_n\}$. The base of the logarithm determines the units: bits for base 2, nats for base $e$, and hartleys or decimal digits for base 10.

Consider, for example, the information content of course grades. As we shall see, if the only constraint on the probability distribution is that of overall normalization, then $S$ is maximized when all the probabilities $p_n$ are equal. The Shannon entropy is then $S = \log_2 \Gamma$, where $\Gamma$ is the total size of our discrete

[6] See 'An Introduction to Information Theory and Entropy' by T. Carter, Santa Fe Complex Systems Summer School, June 2011: http://astarte.csustan.edu/$\sim$tom/SFI-CSSS/info-theory/info-lec.pdf.

[7] My colleague John McGreevy refers to $I(p)$ as the *surprise* of observing an event occurring with probability $p$. I like this very much.

space of states, since $p_n = 1/\Gamma$. Thus, for pass/fail grading, the maximum average information per grade is $-\log_2(\frac{1}{2}) = \log_2 2 = 1$ bit. If only A, B, C, D, and F grades are assigned, then the maximum average information per grade is $\log_2 5 = 2.32$ bits. If we expand the grade options to include {A+, A, A-, B+, B, B-, C+, C, C-, D, F}, then the maximum average information per grade is $\log_2 11 = 3.46$ bits.

Equivalently, consider, following the discussion in vol. 1 of Kardar, a random sequence $\{n_1, n_2, \ldots, n_N\}$ where each element $n_j$ takes one of $K$ possible values. There are then $K^N$ such possible sequences, and to specify one of them requires $\log_2(K^N) = N \log_2 K$ bits of information. However, if the value $n$ occurs with probability $p_n$, then on average it will occur $N_n = Np_n$ times in a sequence of length $N$, and the total number of such sequences will be

$$g(N) = \frac{N!}{\prod_{n=1}^{K} N_n!} \quad . \tag{1.43}$$

In general, this is far less than the total possible number $K^N$, and the number of bits necessary to specify one from among these $g(N)$ possibilities is

$$\log_2 g(N) = \log_2(N!) - \sum_{n=1}^{K} \log_2(N_n!) \approx -N \sum_{n=1}^{K} p_n \log_2 p_n \quad , \tag{1.44}$$

up to terms of order $\log N$. Here we have invoked Stirling's approximation. If the distribution is uniform, then we have $p_n = K^{-1}$ for all $n \in \{1, \ldots, K\}$, and $\log_2 g(N) = N \log_2 K$.

A quantum mechanical *density matrix* $\hat{\varrho} = \sum_{j,k} \varrho_{jk} \, |j\rangle\langle k|$ is a Hermitian operator normalized such that $\mathsf{Tr}\ \hat{\varrho} = 1$ and subject to the condition that each of its eigenvalues $p_n$ lies on the interval $[0, 1]$. Thus we have $\mathsf{Tr}\ \hat{\varrho} = \sum_n p_n = 1$. The $q^{\rm th}$ power of $\hat{\varrho}$ is then

$$\hat{\varrho}^q = \sum_{j,k} (\varrho^q)_{jk} \, |\, j \,\rangle\langle\, k \,| \quad , \tag{1.45}$$

where $(\varrho^q)_{jk}$ is the $j, k$ matrix element of the matrix $\varrho^q$. We can write the entropy of the system as

$$S = -\,\mathsf{Tr}\,\big(\varrho \log \varrho\big) = -\sum_n p_n \log p_n \quad . \tag{1.46}$$

We define $\Omega \equiv \sum_n 1$ to be the total number of events. It may be the case that $\Omega = \infty$.

### 1.4.2 Rényi entropy

The Rényi entropy $S_q\big(\{p_n\}\big)$ is a one-parameter family of entropy functions parameterized by a real number $q \in [0, \infty)$, with

$$S_q = \frac{\log \mathsf{Tr}\,(\varrho^q)}{1-q} = \frac{1}{1-q} \log\Big( \sum_n (p_n)^q \Big) \quad . \tag{1.47}$$

With $q = 1 + \varepsilon$, we have $p_n^{1+\varepsilon} = p_n \cdot \exp(\varepsilon \log p_n) = p_n + \varepsilon\, p_n \log p_n + \mathcal{O}(\varepsilon^2)$ and thus in the limit $q \to 1$ we recover the Shannon entropy (up to a multiplicative constant depending on the base of the logarithm),

$$\lim_{q \to 1} S_q = -\sum_n p_n \log p_n \quad , \tag{1.48}$$

Note that the Rényi entropy is additive for all $q$. Thus, if $\varrho^{\mathsf{A}}$ and $\varrho^{\mathsf{B}}$ are density matrices for two independent systems, then $\varrho^{\mathsf{AB}} = \varrho^{\mathsf{A}} \otimes \varrho^{\mathsf{B}}$ and

$$S_q^{\mathsf{AB}} = S_q^{\mathsf{A}} + S_q^{\mathsf{B}} \quad . \tag{1.49}$$

Again, the Rényi entropy $S_q$ is defined for all $q \in \mathbb{R}_{\geq 0}$. Some special cases:

- In the limit $q \to 0^+$ we have $S_{0^+} = \log \Omega'$, where $\Omega'$ is the number of events for which the probability of occurrence is nonzero, *i.e.* $\Omega' = \sum_n \Theta(p_n)$ where $\Theta(x)$ is the step function.
- For $q \in \mathbb{Z}_+$ a positive integer, we can think of the Rényi entropy in terms of *replicas*. Imagine we take $q$ identical systems, each governed by the same probability distribution. Then the probability of observing the same event among each of the $q$ replicas is $\sum_n (p_n)^q = e^{(1-q)S_q}$.
- As $q \to \infty$, the sum $\sum_n (p_n)^q$ is dominated by the maximum value of $p_n$. Let $p^* = \max_n\{p_n\}$. Then $S_\infty = -\log p^*$. One can check that the result is independent of the degeneracy $g^*$ of the maximum probability[8].

The following inequality holds in general for $q > 1$:

$$S_\infty \leq S_q \leq \left(\frac{q}{q-1}\right) S_\infty \quad . \tag{1.50}$$

### 1.4.3 Kullback-Leibler divergence

Suppose we approximate a probability distribution $P(x)$ by another function $Q(x)$. The expected difference in the logarithms of the probabilities is then given by

$$\begin{aligned} D_{\mathsf{KL}}(P \parallel Q) &= -\sum_{x\in\mathcal{X}} P(x) \Big\{ \log Q(x) - \log P(x) \Big\} \\ &= \sum_{x\in\mathcal{X}} P(x) \log\left(\frac{P(x)}{Q(x)}\right) = \mathbb{E}_P \log\left(\frac{P(X)}{Q(X)}\right) \quad , \end{aligned} \tag{1.51}$$

called the *Kullback-Leibler divergence* or *relative entropy*. If $P(x) > 0$ and $Q(x) = 0$ for any event $x$ then $D_{\mathsf{KL}}(P \parallel Q) = \infty$.

### 1.4.4 Probability distributions from maximum entropy

We have shown how one can proceed from a probability distribution and compute various averages. We now seek to go in the other direction, and determine the full probability distribution based on a knowledge of certain averages.

[8] Provided that $g^*$ is finite, or more generally, that $\log(g^*)/q \to 0$ as $q \to \infty$.

At first, this seems impossible. Suppose we want to reproduce the full probability distribution for an $N$-step random walk from knowledge of the average $\langle X\rangle = (2p-1)N$, where $p$ is the probability of moving to the right at each step (see §1.2 above). The problem seems ridiculously underdetermined, since there are $2^N$ possible configurations for an $N$-step random walk: $\sigma_j = \pm 1$ for $j = 1, \ldots, N$. Overall normalization requires

$$\sum_{\{\sigma_j\}} P(\sigma_1, \ldots, \sigma_N) = 1 \quad , \tag{1.52}$$

but this just imposes one constraint on the $2^N$ probabilities $P(\sigma_1, \ldots, \sigma_N)$, leaving $2^N - 1$ overall parameters. What principle allows us to reconstruct the full probability distribution

$$P(\sigma_1, \ldots, \sigma_N) = \prod_{j=1}^{N} \left( p\,\delta_{\sigma_j,1} + q\,\delta_{\sigma_j,-1} \right) = \prod_{j=1}^{N} p^{(1+\sigma_j)/2}\, q^{(1-\sigma_j)/2} \quad , \tag{1.53}$$

corresponding to $N$ independent steps?

**The principle of maximum entropy**

The entropy of a discrete probability distribution $\{p_n\}$ is defined as

$$S = -\sum_n p_n \log p_n \quad , \tag{1.54}$$

where here we take $e$ as the base of the logarithm. The entropy may therefore be regarded as a function of the probability distribution: $S = S\big(\{p_n\}\big)$. One special property of the entropy is the following. Suppose we have two independent normalized distributions $\{p_a^{\mathsf{A}}\}$ and $\{p_b^{\mathsf{B}}\}$. The joint probability for events $a$ and $b$ is then $P_{a,b} = p_a^{\mathsf{A}}\, p_b^{\mathsf{B}}$. The entropy of the joint distribution is then

$$\begin{aligned} S &= -\sum_a \sum_b P_{a,b} \log P_{a,b} = -\sum_a \sum_b p_a^{\mathsf{A}}\, p_b^{\mathsf{B}} \log\left(p_a^{\mathsf{A}}\, p_b^{\mathsf{B}}\right) = -\sum_a \sum_b p_a^{\mathsf{A}}\, p_b^{\mathsf{B}} \left(\log p_a^{\mathsf{A}} + \log p_b^{\mathsf{B}}\right) \\ &= -\sum_a p_a^{\mathsf{A}} \log p_a^{\mathsf{A}} \cdot \sum_b p_b^{\mathsf{B}} - \sum_b p_b^{\mathsf{B}} \log p_b^{\mathsf{B}} \cdot \sum_a p_a^{\mathsf{A}} = -\sum_a p_a^{\mathsf{A}} \log p_a^{\mathsf{A}} - \sum_b p_b^{\mathsf{B}} \log p_b^{\mathsf{B}} = S^{\mathsf{A}} + S^{\mathsf{B}} \quad . \end{aligned} \tag{1.55}$$

Thus, the entropy of a joint distribution formed from two independent distributions is additive.

Suppose all we knew about $\{p_n\}$ was that it was normalized. Then $\sum_n p_n = 1$. This is a constraint on the values $\{p_n\}$. Let us now extremize the entropy $S$ with respect to the distribution $\{p_n\}$, but subject to the normalization constraint. We do this using Lagrange's method of undetermined multipliers. We define

$$S^*\big(\{p_n\}, \lambda\big) = -\sum_n p_n \log p_n - \lambda\Big(\sum_n p_n - 1\Big) \tag{1.56}$$

and we freely extremize $S^*$ over all its arguments. Thus, for all $n$ we have

$$\begin{aligned} 0 &= \frac{\partial S^*}{\partial p_n} = -\big(\log p_n + 1 + \lambda\big) \\ 0 &= \frac{\partial S^*}{\partial \lambda} = \sum_n p_n - 1 \quad . \end{aligned} \tag{1.57}$$

From the first of these equations, we obtain $p_n = e^{-(1+\lambda)}$, and from the second we obtain

$$\sum_n p_n = e^{-(1+\lambda)} \cdot \sum_n 1 = \Gamma\, e^{-(1+\lambda)} \quad , \tag{1.58}$$

where $\Gamma \equiv \sum_n 1$ is the total number of possible events. Thus, $p_n = \Gamma^{-1}$, which says that all events are equally probable.

Now suppose we know one other piece of information, which is the average value $X = \sum_n X_n\, p_n$ of some quantity. We now extremize $S$ subject to two constraints, and so we define

$$S^*\big(\{p_n\}, \lambda_0, \lambda_1\big) = -\sum_n p_n \log p_n - \lambda_0\Big(\sum_n p_n - 1\Big) - \lambda_1\Big(\sum_n X_n\, p_n - X\Big) \quad . \tag{1.59}$$

We then have

$$\frac{\partial S^*}{\partial p_n} = -\big(\log p_n + 1 + \lambda_0 + \lambda_1\, X_n\big) = 0 \quad , \tag{1.60}$$

which yields the two-parameter distribution

$$p_n = e^{-(1+\lambda_0)}\, e^{-\lambda_1 X_n} \quad . \tag{1.61}$$

To fully determine the distribution $\{p_n\}$, we invoke the two equations $\sum_n p_n = 1$ and $\sum_n X_n\, p_n = X$, which come from extremizing $S^*$ with respect to $\lambda_0$ and $\lambda_1$, respectively:

$$\begin{aligned} 1 &= e^{-(1+\lambda_0)} \sum_n e^{-\lambda_1 X_n} \\ X &= e^{-(1+\lambda_0)} \sum_n X_n\, e^{-\lambda_1 X_n} \quad . \end{aligned} \tag{1.62}$$

**General formulation**

The generalization to $K+1$ constraints is immediately apparent. We have

$$X^a = \sum_n X_n^a\, p_n \quad , \tag{1.63}$$

and therefore we define

$$S^*\big(\{p_n\}, \{\lambda_a\}\big) = -\sum_n p_n \log p_n - \sum_{a=0}^{K} \lambda_a \Big(\sum_n X_n^a\, p_n - X^a\Big) \quad , \tag{1.64}$$

with $X_n^{(a=0)} \equiv X^{(a=0)} = 1$. Then the optimal distribution which extremizes $S$ subject to the $K+1$ constraints is

$$p_n = \exp\left\{-1 - \sum_{a=0}^{K} \lambda_a\, X_n^a\right\} = \frac{1}{Z} \exp\left\{-\sum_{a=1}^{K} \lambda_a\, X_n^a\right\} \quad , \tag{1.65}$$

where $Z = e^{1+\lambda_0}$ is determined by normalization: $\sum_n p_n = 1$. This is a $(K+1)$-parameter distribution, with $\{\lambda_0, \lambda_1, \ldots, \lambda_K\}$ determined by the $K+1$ constraints in eqn. 1.63.

**Example**

As an example, consider the random walk problem. We have two pieces of information:

$$\begin{aligned}\sum_{\sigma_1}\cdots\sum_{\sigma_N} P(\sigma_1,\ldots,\sigma_N) &= 1 \\ \sum_{\sigma_1}\cdots\sum_{\sigma_N} P(\sigma_1,\ldots,\sigma_N)\sum_{j=1}^N \sigma_j &= X \quad .\end{aligned} \tag{1.66}$$

Here the discrete label $n$ from §1.4.4 ranges over $2^N$ possible values, and may be written as an $N$ digit binary number $r_N\cdots r_1$, where $r_j = \frac{1}{2}(1+\sigma_j)$ is 0 or 1. Extremizing $S$ subject to these constraints, we obtain

$$P(\sigma_1,\ldots,\sigma_N) = \mathcal{C}\,\exp\left\{-\lambda\sum_j \sigma_j\right\} = \mathcal{C}\prod_{j=1}^N e^{-\lambda\,\sigma_j} \quad , \tag{1.67}$$

where $\mathcal{C} \equiv e^{-(1+\lambda_0)}$ and $\lambda \equiv \lambda_1$. Normalization then requires

$$\mathsf{Tr}\;P \equiv \sum_{\{\sigma_j\}} P(\sigma_1,\ldots,\sigma_N) = \mathcal{C}\left(e^{\lambda}+e^{-\lambda}\right)^N \quad , \tag{1.68}$$

hence $\mathcal{C} = (2\cosh\lambda)^{-N}$. We then have

$$P(\sigma_1,\ldots,\sigma_N) = \prod_{j=1}^N \frac{e^{-\lambda\sigma_j}}{e^{\lambda}+e^{-\lambda}} = \prod_{j=1}^N \left(p\,\delta_{\sigma_j,1} + q\,\delta_{\sigma_j,-1}\right) \quad , \tag{1.69}$$

where

$$p = \frac{e^{-\lambda}}{e^{\lambda}+e^{-\lambda}} \quad , \quad q = 1-p = \frac{e^{\lambda}}{e^{\lambda}+e^{-\lambda}} \quad . \tag{1.70}$$

We then have $X = (2p-1)N$, which determines $p = \frac{1}{2}\big(1+\frac{X}{N}\big)$, and we have recovered the Bernoulli distribution.

Of course there are no miracles[9], and there are an infinite family of distributions for which $X = (2p-1)N$ that are not Bernoulli. For example, we could have imposed another constraint, such as $E = \sum_{j=1}^{N-1}\sigma_j\,\sigma_{j+1}$. This would result in the distribution

$$P(\sigma_1,\ldots,\sigma_N) = \frac{1}{Z}\exp\left\{-\lambda_1\sum_{j=1}^N \sigma_j - \lambda_2\sum_{j=1}^{N-1}\sigma_j\,\sigma_{j+1}\right\} \quad , \tag{1.71}$$

with $Z(\lambda_1,\lambda_2)$ determined by normalization: $\sum_{\boldsymbol{\sigma}} P(\boldsymbol{\sigma}) = 1$. This is the one-dimensional Ising chain of classical equilibrium statistical physics. Defining the transfer matrix $R_{ss'} = e^{-\lambda_1(s+s')/2}\,e^{-\lambda_2 ss'}$ with $s, s' = \pm 1$,

$$\begin{aligned}R &= \begin{pmatrix} e^{-\lambda_1-\lambda_2} & e^{\lambda_2} \\ e^{\lambda_2} & e^{\lambda_1-\lambda_2}\end{pmatrix} \\ &= e^{-\lambda_2}\cosh(\lambda_1)\,\mathbb{I} - e^{-\lambda_2}\sinh(\lambda_1)\,\tau^z + e^{\lambda_2}\,\tau^x \quad ,\end{aligned} \tag{1.72}$$

[9] See §10 of *An Enquiry Concerning Human Understanding* by David Hume (1748).

where $\tau^x$ and $\tau^z$ are Pauli matrices, we have that

$$Z_{\rm ring} = {\rm Tr}\left(R^N\right) \quad , \quad Z_{\rm chain} = {\rm Tr}\left(R^{N-1}S\right) \quad , \tag{1.73}$$

where $S_{ss'} = \exp\left(-\frac{1}{2}\lambda_1(s+s')\right)$, *i.e.*

$$S = \begin{pmatrix} e^{-\lambda_1} & 1 \\ 1 & e^{\lambda_1} \end{pmatrix} = \cosh(\lambda_1)\,\mathbb{I} - \sinh(\lambda_1)\,\tau^z + \tau^x \quad . \tag{1.74}$$

The appropriate case here is that of the chain, but in the thermodynamic limit $N \to \infty$ both chain and ring yield identical results, so we will examine here the results for the ring, which are somewhat easier to obtain. Clearly $Z_{\rm ring} = \zeta_+^N + \zeta_-^N$, where $\zeta_\pm$ are the eigenvalues of $R$:

$$\zeta_\pm = e^{-\lambda_2}\cosh\lambda_1 \pm \sqrt{e^{-2\lambda_2}\sinh^2\lambda_1 + e^{2\lambda_2}} \quad . \tag{1.75}$$

In the thermodynamic limit, the $\zeta_+$ eigenvalue dominates, and $Z_{\rm ring} \simeq \zeta_+^N$. We now have

$$X = \Big\langle \sum_{j=1}^N \sigma_j \Big\rangle = -\frac{\partial \log Z}{\partial \lambda_1} = -\frac{N\sinh\lambda_1}{\sqrt{\sinh^2\lambda_1 + e^{4\lambda_2}}} \quad . \tag{1.76}$$

We also have $E = -\partial \log Z/\partial\lambda_2$. These two equations determine the Lagrange multipliers $\lambda_1(X,E,N)$ and $\lambda_2(X,E,N)$. In the thermodynamic limit, we have $\lambda_i = \lambda_i(X/N, E/N)$. Thus, if we fix $X/N = 2p-1$ alone, there is a continuous one-parameter family of distributions, parametrized by $\varepsilon = E/N$, which satisfy the constraint on $X$.

So what is it about the maximum entropy approach that is so compelling? Maximum entropy gives us a calculable distribution which is consistent with "maximum ignorance" given our known constraints[10]. In that sense, it is as unbiased as possible, from an information theoretic point of view. As a starting point, a maximum entropy distribution may be improved upon, using Bayesian methods for example (see §1.6.2 below).

### 1.4.5 Continuous probability distributions

Suppose we have a continuous probability density $P(\boldsymbol{\varphi})$ defined over some set $\Omega$. We have observables

$$X^a = \int\limits_\Omega d\mu\, X^a(\boldsymbol{\varphi})\, P(\boldsymbol{\varphi}) \quad , \tag{1.77}$$

where $d\mu$ is the appropriate integration measure. We assume $d\mu = \prod_{j=1}^D d\varphi_j$, where $D$ is the dimension of $\Omega$. Then we extremize the functional

$$S^*\big[P(\boldsymbol{\varphi}), \{\lambda_a\}\big] = -\int\limits_\Omega d\mu\, P(\boldsymbol{\varphi})\log P(\boldsymbol{\varphi}) - \sum_{a=0}^K \lambda_a \left( \int\limits_\Omega d\mu\, P(\boldsymbol{\varphi})\, X^a(\boldsymbol{\varphi}) - X^a \right) \tag{1.78}$$

[10] This is conventional Jaynesian language, but more accurately the result depends critically on the chosen state space, reference measure, and constraints. There is no observer-independent definition of ignorance.

with respect to $P(\boldsymbol{\varphi})$ and with respect to $\{\lambda_a\}$. Again, $X^0(\boldsymbol{\varphi}) \equiv X^0 \equiv 1$. This yields the following result:

$$\log P(\boldsymbol{\varphi}) = -1 - \sum_{a=0}^{K} \lambda_a \, X^a(\boldsymbol{\varphi}) \quad . \tag{1.79}$$

The $(K+1)$ Lagrange multipliers $\{\lambda_a\}$ are then determined from the $(K+1)$ constraint equations in eqn. 1.77.

As an example, consider a distribution $P(x)$ over the real numbers $\mathbb{R}$. We constrain

$$\int\limits_{-\infty}^{\infty} dx \, P(x) = 1 \qquad , \qquad \int\limits_{-\infty}^{\infty} dx \, x \, P(x) = \mu \qquad , \qquad \int\limits_{-\infty}^{\infty} dx \, x^2 \, P(x) = \mu^2 + \sigma^2 \quad . \tag{1.80}$$

Extremizing the entropy, we then obtain

$$P(x) = \mathcal{C} \, e^{-\lambda_1 x - \lambda_2 x^2} \quad , \tag{1.81}$$

where $\mathcal{C} = e^{-(1+\lambda_0)}$. We already know the answer:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \, e^{-(x-\mu)^2/2\sigma^2} \quad . \tag{1.82}$$

In other words, $\lambda_1 = -\mu/\sigma^2$ and $\lambda_2 = 1/2\sigma^2$, with $\mathcal{C} = (2\pi\sigma^2)^{-1/2} \, \exp(-\mu^2/2\sigma^2)$.

## 1.5 General Aspects of Probability Distributions

### 1.5.1 Discrete and continuous distributions

Consider a system whose possible configurations $|\, n \,\rangle$ can be labeled by a discrete variable $n \in \mathcal{C}$, where $\mathcal{C}$ is the set of possible configurations. The total number of possible configurations, which is to say the *order* of the set $\mathcal{C}$, may be finite or infinite. Next, consider an ensemble of such systems, and let $P_n$ denote the probability that a given random element from that ensemble is in the state (configuration) $|\, n \,\rangle$. The collection $\{P_n\}$ forms a *discrete probability distribution*. We assume that the distribution is *normalized*, meaning

$$\sum_{n \in \mathcal{C}} P_n = 1 \quad . \tag{1.83}$$

Now let $A_n$ be a quantity which takes values depending on $n$. The average of $A$ is given by

$$\langle A \rangle = \sum_{n \in \mathcal{C}} P_n \, A_n \quad . \tag{1.84}$$

Typically, $\mathcal{C}$ is the set of integers ($\mathbb{Z}$) or some subset thereof, but it could be any countable set. As an example, consider the throw of a single six-sided die. Then $P_n = \frac{1}{6}$ for each $n \in \{1, \ldots, 6\}$. Let $A_n = 0$ if

$n$ is even and 1 if $n$ is odd. We then find $\langle A \rangle = \frac{1}{2}$, *i.e.* on average half the throws of the die will result in an even number.

It may be that the system's configurations are described by several discrete variables $\{n_1, n_2, n_3, \ldots\}$. We can combine these into a vector $\boldsymbol{n}$ and then we write $P_{\boldsymbol{n}}$ for the discrete distribution, with $\sum_{\boldsymbol{n}} P_{\boldsymbol{n}} = 1$.

Another possibility is that the system's configurations are parameterized by a collection of continuous variables, $\boldsymbol{\varphi} = \{\varphi_1, \ldots, \varphi_n\}$. We write $\boldsymbol{\varphi} \in \Omega$, where $\Omega$ is the phase space (or configuration space) of the system. Let $d\mu$ be a *measure* on this space. In general, we can write

$$d\mu = W(\varphi_1, \ldots, \varphi_n)\, d\varphi_1\, d\varphi_2 \cdots d\varphi_n \quad . \tag{1.85}$$

The phase space measure used in classical statistical mechanics gives equal weight $W$ to equal phase space volumes:

$$d\mu = \mathcal{C} \prod_{\sigma=1}^{r} dq_\sigma\, dp_\sigma \quad , \tag{1.86}$$

where $\mathcal{C}$ is a constant we shall discuss below[11].

Any continuous probability distribution $P(\boldsymbol{\varphi})$ is normalized according to

$$\int\limits_{\Omega} d\mu\, P(\boldsymbol{\varphi}) = 1 \quad . \tag{1.87}$$

The average of a function $A(\boldsymbol{\varphi})$ on configuration space is then

$$\langle A \rangle = \int\limits_{\Omega} d\mu\, P(\boldsymbol{\varphi})\, A(\boldsymbol{\varphi}) \quad . \tag{1.88}$$

For example, consider the Gaussian distribution

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2} \quad . \tag{1.89}$$

From the result[12]

$$\int\limits_{-\infty}^{\infty} dx\, e^{-\alpha x^2}\, e^{-\beta x} = \sqrt{\frac{\pi}{\alpha}}\, e^{\beta^2/4\alpha} \quad , \tag{1.90}$$

we see that $P(x)$ is normalized. One can then compute

$$\langle x \rangle = \mu \qquad , \qquad \langle x^2 \rangle - \langle x \rangle^2 = \sigma^2 \quad . \tag{1.91}$$

We call $\mu$ the *mean* and $\sigma$ the *standard deviation* of the distribution, eqn. 1.89.

[11] Such a measure is invariant with respect to canonical transformations, which are the broad class of transformations among coordinates and momenta which leave Hamilton's equations of motion invariant, and which preserve phase space volumes under Hamiltonian evolution. For this reason $d\mu$ is called an *invariant phase space measure*.

[12] Memorize this!

The quantity $P(\boldsymbol{\varphi})$ is called the *distribution* or *probability density*. One has

$$P(\boldsymbol{\varphi})\,d\mu = \text{probability that configuration lies within volume } d\mu \text{ centered at } \boldsymbol{\varphi}$$

For example, consider the probability density $P = 1$ normalized on the interval $x \in \big[0, 1\big]$. The probability that some $x$ chosen at random will be *exactly* $\frac{1}{2}$, say, is *zero* – one would have to specify each of the infinitely many digits of $x$. However, we can say that $x \in \big[0.45, 0.55\big]$ with probability $\frac{1}{10}$.

If $x$ is distributed according to $P_1(x)$, then if $x_1$ and $x_2$ are independent copies drawn from $P_1(x)$, the probability distribution on the product space $(x_1\,,\,x_2)$ is simply the product of the distributions: $P_2(x_1, x_2) = P_1(x_1)\,P_1(x_2)$. Suppose we have a function $\phi(x_1, \ldots, x_N)$. How is it distributed? Let $P(\phi)$ be the distribution for $\phi$. We then have

$$\begin{aligned} P(\phi) &= \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P_N(x_1, \ldots, x_N)\,\delta\Big(\phi(x_1, \ldots, x_N) - \phi\Big) \\ &= \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P_1(x_1) \cdots P_1(x_N)\,\delta\Big(\phi(x_1, \ldots, x_N) - \phi\Big) \quad , \end{aligned} \tag{1.92}$$

where the second line is appropriate if the $\{x_j\}$ are themselves distributed independently. Note that the distribution $P(\phi)$ is normalized, *i.e.* $\int\limits_{-\infty}^{\infty} d\phi\, P(\phi) = 1$.

### 1.5.2 Central limit theorem

In particular, consider the distribution function of the sum $X = \sum_{i=1}^{N} x_i$. We will be particularly interested in the case where $N$ is large. For general $N$, though, we have

$$P_N(X) = \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P_1(x_1) \cdots P_1(x_N)\,\delta\big(x_1 + x_2 + \ldots + x_N - X\big) \quad . \tag{1.93}$$

It is convenient to compute the Fourier transform of $P(X)$:

$$\begin{aligned} \hat{P}_N(k) &= \int\limits_{-\infty}^{\infty} dX\, P_N(X)\, e^{-ikX} \\ &= \int\limits_{-\infty}^{\infty} dX \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P_1(x_1) \cdots P_1(x_N)\,\delta\big(x_1 + \ldots + x_N - X\big)\, e^{-ikX} = \big[\hat{P}_1(k)\big]^N \quad , \end{aligned} \tag{1.94}$$

where

$$\hat{P}_1(k) = \int\limits_{-\infty}^{\infty} dx\, P_1(x)\, e^{-ikx} \tag{1.95}$$

is the Fourier transform[13] of the single variable distribution $P_1(x)$. The distribution $P_N(X)$ is a *convolution* of the individual $P_1(x_i)$ distributions. We have therefore proven that *the Fourier transform of a convolution is the product of the Fourier transforms*.

OK, now we can write for $\hat{P}_1(k)$

$$\begin{aligned}\hat{P}_1(k) &= \int_{-\infty}^{\infty} dx\, P_1(x)\left(1 - ikx - \tfrac{1}{2}\,k^2x^2 + \tfrac{1}{6}ik^3\,x^3 + \ldots\right) \\ &= 1 - ik\langle x\rangle - \tfrac{1}{2}\,k^2\langle x^2\rangle + \tfrac{1}{6}\,ik^3\langle x^3\rangle + \ldots \quad .\end{aligned} \tag{1.96}$$

Thus,

$$\log\hat{P}_1(k) = -i\mu k - \tfrac{1}{2}\sigma^2k^2 + \tfrac{1}{6}\,i\gamma^3\,k^3 + \ldots \quad , \tag{1.97}$$

where

$$\mu = \langle x\rangle \qquad , \qquad \sigma^2 = \langle x^2\rangle - \langle x\rangle^2 \qquad , \qquad \gamma^3 = \langle x^3\rangle - 3\,\langle x^2\rangle\,\langle x\rangle + 2\,\langle x\rangle^3 \tag{1.98}$$

We can now write

$$\left[\hat{P}_1(k)\right]^N = e^{-iN\mu k}\,e^{-N\sigma^2k^2/2}\,e^{iN\gamma^3k^3/6}\ldots \quad . \tag{1.99}$$

Now for the inverse transform. In computing $P_N(X)$, we will expand the term $e^{iN\gamma^3k^3/6}$ and all subsequent terms in the above product as a power series in $k$. We then have

$$\begin{aligned}P_N(X) &= \int_{-\infty}^{\infty}\frac{dk}{2\pi}\,e^{ik(X-N\mu)}\,e^{-N\sigma^2k^2/2}\left\{1 + \tfrac{1}{6}\,i\,N\gamma^3k^3 + \ldots\right\} \\ &= \left(1 - \frac{\gamma^3}{6}N\,\frac{\partial^3}{\partial X^3} + \ldots\right)\frac{1}{\sqrt{2\pi N\sigma^2}}\,e^{-(X-N\mu)^2/2N\sigma^2} \\ &= \left(1 - \frac{\gamma^3}{6}\,N^{-1/2}\,\frac{\partial^3}{\partial\xi^3} + \ldots\right)\frac{1}{\sqrt{2\pi N\sigma^2}}\,e^{-\xi^2/2\sigma^2} \quad .\end{aligned} \tag{1.100}$$

[13] Jean Baptiste Joseph Fourier (1768-1830) had an illustrious career. The son of a tailor, and orphaned at age eight, Fourier's ignoble status rendered him ineligible to receive a commission in the scientific corps of the French army. A Benedictine minister at the École Royale Militaire of Auxerre remarked, "Fourier, not being noble, could not enter the artillery, although he were a second Newton." Fourier prepared for the priesthood but his affinity for mathematics proved overwhelming, and so he left the abbey and soon thereafter accepted a military lectureship position. Despite his initial support for revolution in France, in 1794 Fourier ran afoul of a rival sect while on a trip to Orléans and was arrested and very nearly guillotined. Fortunately the Reign of Terror ended soon after the death of Robespierre, and Fourier was released. He went on Napoleon Bonaparte's 1798 expedition to Egypt, where he was appointed governor of Lower Egypt. His organizational skills impressed Napoleon, and upon return to France he was appointed to a position of prefect in Grenoble. It was in Grenoble that Fourier performed his landmark studies of heat, and his famous work on partial differential equations and Fourier series. It seems that Fourier's fascination with heat began in Egypt, where he developed an appreciation of desert climate. His fascination developed into an obsession, and he became convinced that heat could promote a healthy body. He would cover himself in blankets, like a mummy, in his heated apartment, even during the middle of summer. On May 4, 1830, Fourier, so arrayed, tripped and fell down a flight of stairs. This aggravated a developing heart condition, which he refused to treat with anything other than more heat. Two weeks later, he died. Fourier's is one of the 72 names of scientists, engineers and other luminaries which are engraved on the Eiffel Tower.

In going from the second line to the third, we have written $X = N\mu + \sqrt{N}\,\xi$, in which case $\partial_X = N^{-1/2}\,\partial_\xi$, and the non-Gaussian terms give a subleading contribution which vanishes in the $N \to \infty$ limit. We have just proven the *central limit theorem*: for large but finite $N$, the distribution of a sum of $N$ independent random variables $x_i$ is approximately a Gaussian with mean $N\mu$ and standard deviation $\sqrt{N}\,\sigma$. Our only assumptions are that the mean $\mu$ and standard deviation $\sigma$ exist for the distribution $P_1(x)$. Note that $P_1(x)$ itself need not be a Gaussian – it could be a very peculiar distribution indeed, but so long as its first and second moments exist, where the $k^{\rm th}$ moment is simply $\langle x^k\rangle$, the distribution of the sum $X = \sum_{i=1}^N x_i$ is a Gaussian.

Note that we have assumed the existence of a finite third moment $\gamma$ in the above. Without this assumption, we may still conclude

$$\hat{P}(k) = 1 - i\mu k - \tfrac{1}{2}\langle x^2\rangle\, k^2 + o(k^2) \quad . \tag{1.101}$$

### 1.5.3 Moments and cumulants

Consider a general multivariate distribution $P(x_1, \ldots, x_N)$ and define the multivariate Fourier transform

$$\hat{P}(k_1, \ldots, k_N) = \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P(x_1, \ldots, x_N)\, \exp\left(-i \sum_{j=1}^N k_j x_j\right) \quad . \tag{1.102}$$

The inverse relation is

$$P(x_1, \ldots, x_N) = \int\limits_{-\infty}^{\infty} \frac{dk_1}{2\pi} \cdots \int\limits_{-\infty}^{\infty} \frac{dk_N}{2\pi}\, \hat{P}(k_1, \ldots, k_N)\, \exp\left(+i \sum_{j=1}^N k_j x_j\right) \quad . \tag{1.103}$$

Acting on $\hat{P}(\boldsymbol{k})$, the differential operator $i\,\frac{\partial}{\partial k_i}$ brings down from the exponential a factor of $x_i$ inside the integral. Thus,

$$\left[\left(i\,\frac{\partial}{\partial k_1}\right)^{m_1} \cdots \left(i\,\frac{\partial}{\partial k_N}\right)^{m_N} \hat{P}(\boldsymbol{k})\right]_{\boldsymbol{k}=0} = \left\langle x_1^{m_1} \cdots x_N^{m_N}\right\rangle \quad . \tag{1.104}$$

Similarly, we can reconstruct the distribution from its moments, *viz.*

$$\hat{P}(\boldsymbol{k}) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_N=0}^{\infty} \frac{(-ik_1)^{m_1}}{m_1!} \cdots \frac{(-ik_N)^{m_N}}{m_N!} \left\langle x_1^{m_1} \cdots x_N^{m_N}\right\rangle \quad . \tag{1.105}$$

The *cumulants* $\langle\!\langle x_1^{m_1} \cdots x_N^{m_N}\rangle\!\rangle$ are defined by the Taylor expansion of $\log \hat{P}(\boldsymbol{k})$:

$$\log \hat{P}(\boldsymbol{k}) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_N=0}^{\infty} \frac{(-ik_1)^{m_1}}{m_1!} \cdots \frac{(-ik_N)^{m_N}}{m_N!} \left\langle\!\left\langle x_1^{m_1} \cdots x_N^{m_N}\right\rangle\!\right\rangle \quad . \tag{1.106}$$

There is no general form for the cumulants. It is straightforward to derive the following low order results:

$$\begin{aligned}
\langle\!\langle x_i\rangle\!\rangle &= \langle x_i\rangle \\
\langle\!\langle x_i x_j\rangle\!\rangle &= \langle x_i x_j\rangle - \langle x_i\rangle\langle x_j\rangle \\
\langle\!\langle x_i x_j x_k\rangle\!\rangle &= \langle x_i x_j x_k\rangle - \langle x_i x_j\rangle\langle x_k\rangle - \langle x_j x_k\rangle\langle x_i\rangle - \langle x_k x_i\rangle\langle x_j\rangle + 2\langle x_i\rangle\langle x_j\rangle\langle x_k\rangle \quad .
\end{aligned} \tag{1.107}$$

### 1.5.4 Multidimensional Gaussian integral

Consider the multivariable Gaussian distribution,

$$P(\boldsymbol{x}) \equiv \left(\frac{\det A}{(2\pi)^n}\right)^{1/2} \exp\Big(-\tfrac{1}{2}\, x_i\, A_{ij}\, x_j\Big) \quad , \tag{1.108}$$

where $A$ is a real, symmetric, positive definite matrix of rank $n$. A mathematical result which is extremely important throughout physics is the following:

$$Z(\boldsymbol{b}) = \left(\frac{\det A}{(2\pi)^n}\right)^{1/2} \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_n\ \exp\Big(-\tfrac{1}{2}\, x_i\, A_{ij}\, x_j + b_i\, x_i\Big) = \exp\left(\tfrac{1}{2}\, b_i\, A^{-1}_{ij}\, b_j\right) \quad . \tag{1.109}$$

Here, the vector $\boldsymbol{b} = (b_1\,,\,\ldots\,,\,b_n)$ is identified as a *source*. Since $Z(0) = 1$, we have that the distribution $P(\boldsymbol{x})$ is normalized. Now consider averages of the form

$$\begin{aligned} \langle\, x_{j_1} \cdots\, x_{j_{2k}}\,\rangle &= \int d^n\!x\; P(\boldsymbol{x})\; x_{j_1} \cdots\, x_{j_{2k}} = \left.\frac{\partial^{2k} Z(\boldsymbol{b})}{\partial b_{j_1} \cdots\, \partial b_{j_{2k}}}\right|_{\boldsymbol{b}=0} \\ &= \sum_{\text{contractions}} A^{-1}_{j_{\sigma(1)} j_{\sigma(2)}} \cdots A^{-1}_{j_{\sigma(2k-1)} j_{\sigma(2k)}} \quad . \end{aligned} \tag{1.110}$$

The sum in the last term is over all *contractions* of the indices $\{j_1\,,\,\ldots\,,\,j_{2k}\}$. A contraction is an arrangement of the $2k$ indices into $k$ pairs. There are $C_{2k} = (2k)!/2^k k!$ possible such contractions. To obtain this result for $C_{2k}$, we start with the first index and then find a mate among the remaining $2k-1$ indices. Then we choose the next unpaired index and find a mate among the remaining $2k-3$ indices. Proceeding in this manner, we have

$$C_{2k} = (2k-1)\cdot(2k-3)\cdots 3\cdot 1 = \frac{(2k)!}{2^k k!} \quad . \tag{1.111}$$

Equivalently, we can take all possible permutations of the $2k$ indices, and then divide by $2^k k!$ since permutation within a given pair results in the same contraction and permutation among the $k$ pairs results in the same contraction. For example, for $k = 2$, we have $C_4 = 3$, and

$$\langle\, x_{j_1} x_{j_2} x_{j_3} x_{j_4}\,\rangle = A^{-1}_{j_1 j_2} A^{-1}_{j_3 j_4} + A^{-1}_{j_1 j_3} A^{-1}_{j_2 j_4} + A^{-1}_{j_1 j_4} A^{-1}_{j_2 j_3} \quad . \tag{1.112}$$

If we define $b_i = ik_i$, we have

$$\hat{P}(\boldsymbol{k}) = \exp\Big(-\tfrac{1}{2}\, k_i\, A^{-1}_{ij}\, k_j\Big) \quad , \tag{1.113}$$

from which we read off the cumulants $\langle\!\langle x_i x_j \rangle\!\rangle = A^{-1}_{ij}$, with all higher order cumulants vanishing.

## 1.6 Appendix: Bayesian Statistical Inference

### 1.6.1 Frequentists and Bayesians

The field of statistical inference is roughly divided into two schools of practice: frequentism and Bayesianism. You can find several articles on the web discussing the differences in these two approaches. In both

cases we would like to model observable data $\boldsymbol{x}$ by a distribution. The distribution in general depends on one or more parameters $\boldsymbol{\theta}$. The basic worldviews of the two approaches are as follows:

> **Frequentism:** Data $\boldsymbol{x}$ are a random sample drawn from an infinite pool at some *frequency*. The underlying parameters $\boldsymbol{\theta}$, which are to be estimated, remain fixed during this process. There is no information prior to the model specification. The experimental conditions under which the data are collected are presumed to be controlled and repeatable. Results are generally expressed in terms of *confidence intervals* and *confidence levels*, obtained via *statistical hypothesis testing*. Probabilities have meaning only for data yet to be collected. Calculations generally are computationally straightforward.

> **Bayesianism:** The only data $\boldsymbol{x}$ which matter are those which have been observed. The parameters $\boldsymbol{\theta}$ are unknown and due to uncertainty are described probabilistically using a *prior distribution*, which is generally based on some available information but which also may be at least partially subjective. The priors are then to be *updated* based on observed data $\boldsymbol{x}$. Results are expressed in terms of *posterior distributions* and *credible intervals*. Calculations can be computationally intensive.

In essence, frequentists say *the data are random and the parameters are fixed*. While Bayesians say *the data are fixed and the parameters are random*[14]. Overall, frequentism has dominated over the past several hundred years, but Bayesianism has been coming on strong of late, and many physicists seem naturally drawn to the Bayesian perspective.

### 1.6.2 Updating Bayesian priors

Given *data* $D$ and a *hypothesis* $H$, Bayes' theorem tells us

$$P(H \,|\, D) = \frac{P(D \,|\, H)\, P(H)}{P(D)} \quad . \tag{1.114}$$

Typically the data are in the form of a set of values $\boldsymbol{x} = \{x_1, \ldots, x_N\}$, and the hypothesis in the form of a set of parameters $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_K\}$. It is notationally helpful to express distributions of $\boldsymbol{x}$ and distributions of $\boldsymbol{x}$ conditioned on $\boldsymbol{\theta}$ using the symbol $f$, and distributions of $\boldsymbol{\theta}$ and distributions of $\boldsymbol{\theta}$ conditioned on $\boldsymbol{x}$ using the symbol $\pi$, rather than using the symbol $P$ everywhere. We then have

$$\pi(\boldsymbol{\theta} \,|\, \boldsymbol{x}) = \frac{f(\boldsymbol{x} \,|\, \boldsymbol{\theta})\, \pi(\boldsymbol{\theta})}{\int\limits_{\Theta} d\boldsymbol{\theta}'\, f(\boldsymbol{x} \,|\, \boldsymbol{\theta}')\, \pi(\boldsymbol{\theta}')} \quad , \tag{1.115}$$

where $\Theta \ni \boldsymbol{\theta}$ is the space of parameters. Note that $\int_\Theta d\boldsymbol{\theta}\, \pi(\boldsymbol{\theta} \,|\, \boldsymbol{x}) = 1$. The denominator of the RHS is simply $f(\boldsymbol{x})$, which is independent of $\boldsymbol{\theta}$, hence $\pi(\boldsymbol{\theta} \,|\, \boldsymbol{x}) \propto f(\boldsymbol{x} \,|\, \boldsymbol{\theta})\, \pi(\boldsymbol{\theta})$. We call $\pi(\boldsymbol{\theta})$ the *prior* for

[14] "A frequentist is a person whose long-run ambition is to be wrong 5% of the time. A Bayesian is one who, vaguely expecting a horse, and catching glimpse of a donkey, strongly believes he has seen a mule." – Charles Annis.

$\boldsymbol{\theta}$, $f(\boldsymbol{x}\,|\,\boldsymbol{\theta})$ the *likelihood* of $\boldsymbol{x}$ given $\boldsymbol{\theta}$, and $\pi(\boldsymbol{\theta}\,|\,\boldsymbol{x})$ the *posterior* for $\boldsymbol{\theta}$ given $\boldsymbol{x}$. The idea here is that while our initial guess at the $\boldsymbol{\theta}$ distribution is given by the prior $\pi(\boldsymbol{\theta})$, after taking data, we should *update* this distribution to the posterior $\pi(\boldsymbol{\theta}\,|\,\boldsymbol{x})$. The likelihood $f(\boldsymbol{x}\,|\,\boldsymbol{\theta})$ is entailed by our model for the phenomenon which produces the data. We can use the posterior to find the distribution of new data points $\boldsymbol{y}$, called the *posterior predictive distribution*,

$$f(\boldsymbol{y}\,|\,\boldsymbol{x}) = \int\limits_{\Theta} d\boldsymbol{\theta}\, f(\boldsymbol{y}\,|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta}\,|\,\boldsymbol{x}) \quad . \tag{1.116}$$

This is the update of the *prior predictive distribution*,

$$f(\boldsymbol{x}) = \int\limits_{\Theta} d\boldsymbol{\theta}\, f(\boldsymbol{x}\,|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta}) \quad . \tag{1.117}$$

**Example: coin flipping**

Consider a model of coin flipping based on a standard Bernoulli distribution, where $\theta \in [0,1]$ is the probability for heads ($x=1$) and $1-\theta$ the probability for tails ($x=0$). That is,

$$\begin{aligned} f(x_1,\ldots,x_N\,|\,\theta) &= \prod_{j=1}^{N}\Big[(1-\theta)\,\delta_{x_j,0} + \theta\,\delta_{x_j,1}\Big] \\ &= \theta^X(1-\theta)^{N-X} \quad , \end{aligned} \tag{1.118}$$

where $X = \sum_{j=1}^N x_j$ is the observed total number of heads, and $N-X$ the corresponding number of tails. We now need a prior $\pi(\theta)$. We choose the Beta distribution,

$$\pi(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\mathsf{B}(\alpha,\beta)} \quad , \tag{1.119}$$

where $\mathsf{B}(\alpha,\beta) = \Gamma(\alpha)\,\Gamma(\beta)/\Gamma(\alpha+\beta)$ is the Beta function. One can check that $\pi(\theta)$ is normalized on the unit interval: $\int_0^1 d\theta\,\pi(\theta) = 1$ for all positive $\alpha,\beta$. Even if we limit ourselves to this form of the prior, different Bayesians might bring different assumptions about the values of $\alpha$ and $\beta$. Note that if we choose $\alpha=\beta=1$, the prior distribution for $\theta$ is flat, with $\pi(\theta)=1$.

We now compute the posterior distribution for $\theta$:

$$\pi(\theta\,|\,x_1,\ldots,x_N) = \frac{f(x_1,\ldots,x_N\,|\,\theta)\,\pi(\theta)}{\int_0^1 d\theta'\, f(x_1,\ldots,x_N\,|\,\theta')\,\pi(\theta')} = \frac{\theta^{X+\alpha-1}(1-\theta)^{N-X+\beta-1}}{\mathsf{B}(X+\alpha,N-X+\beta)} \quad . \tag{1.120}$$

Thus, we retain the form of the Beta distribution, but with updated parameters,

$$\begin{aligned} \alpha' &= X+\alpha \\ \beta' &= N-X+\beta \quad . \end{aligned} \tag{1.121}$$

The fact that the functional form of the prior is retained by the posterior is generally *not* the case in Bayesian updating. We can also compute the prior predictive,

$$\begin{aligned} f(x_1,\ldots,x_N) &= \int_0^1 d\theta\; f(x_1,\ldots,x_N\,|\,\theta)\,\pi(\theta) \\ &= \frac{1}{\mathsf{B}(\alpha,\beta)}\int_0^1 d\theta\;\theta^{X+\alpha-1}(1-\theta)^{N-X+\beta-1} = \frac{\mathsf{B}(X+\alpha,N-X+\beta)}{\mathsf{B}(\alpha,\beta)} \quad . \end{aligned} \tag{1.122}$$

The posterior predictive is then

$$\begin{aligned} f(y_1,\ldots,y_M\,|\,x_1,\ldots,x_N) &= \int_0^1 d\theta\; f(y_1,\ldots,y_M\,|\,\theta)\,\pi(\theta\,|\,x_1,\ldots,x_N) \\ &= \frac{1}{\mathsf{B}(X+\alpha,N-X+\beta)}\int_0^1 d\theta\;\theta^{X+Y+\alpha-1}(1-\theta)^{N-X+M-Y+\beta-1} \\ &= \frac{\mathsf{B}(X+Y+\alpha,N-X+M-Y+\beta)}{\mathsf{B}(X+\alpha,N-X+\beta)} \quad . \end{aligned} \tag{1.123}$$

### 1.6.3 Hyperparameters and conjugate priors

In the above example, $\theta$ is a *parameter* of the Bernoulli distribution, *i.e.* the likelihood, while quantities $\alpha$ and $\beta$ are *hyperparameters* which enter the prior $\pi(\theta)$. Accordingly, we could have written $\pi(\theta\,|\,\alpha,\beta)$ for the prior. We then have for the posterior

$$\pi(\boldsymbol{\theta}\,|\,\boldsymbol{x},\boldsymbol{\alpha}) = \frac{f(\boldsymbol{x}\,|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta}\,|\,\boldsymbol{\alpha})}{\int\limits_{\Theta} d\boldsymbol{\theta}'\,f(\boldsymbol{x}\,|\,\boldsymbol{\theta}')\,\pi(\boldsymbol{\theta}'\,|\,\boldsymbol{\alpha})} \quad , \tag{1.124}$$

replacing eqn. 1.115, *etc.*, where $\boldsymbol{\alpha} \in A$ is the vector of hyperparameters. The hyperparameters can also be distributed, according to a *hyperprior* $\rho(\boldsymbol{\alpha})$, and the hyperpriors can further be parameterized by *hyperhyperparameters*, which can have their own distributions, *ad nauseam*.

What use is all this? We've already seen a compelling example: when the posterior is of the same form as the prior, the Bayesian update can be viewed as an automorphism of the hyperparameter space $A$, *i.e.* one set of hyperparameters $\boldsymbol{\alpha}$ is mapped to a new set of hyperparameters $\widetilde{\boldsymbol{\alpha}}$.

**Definition:** A parametric family of distributions $\mathcal{P} = \big\{\pi(\boldsymbol{\theta}\,|\,\boldsymbol{\alpha}) \mid \boldsymbol{\theta}\in\Theta,\,\boldsymbol{\alpha}\in A\big\}$ is called a *conjugate family* for a family of distributions $\big\{f(\boldsymbol{x}\,|\,\boldsymbol{\theta}) \mid \boldsymbol{x}\in\mathcal{X},\,\boldsymbol{\theta}\in\Theta\big\}$ if, for all $\boldsymbol{x}\in\mathcal{X}$ and $\boldsymbol{\alpha}\in A$,

$$\pi(\boldsymbol{\theta}\,|\,\boldsymbol{x},\boldsymbol{\alpha}) \equiv \frac{f(\boldsymbol{x}\,|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta}\,|\,\boldsymbol{\alpha})}{\int\limits_{\Theta} d\boldsymbol{\theta}'\,f(\boldsymbol{x}\,|\,\boldsymbol{\theta}')\,\pi(\boldsymbol{\theta}'\,|\,\boldsymbol{\alpha})} \in \mathcal{P} \quad . \tag{1.125}$$

That is, $\pi(\boldsymbol{\theta}\,|\,\boldsymbol{x},\boldsymbol{\alpha}) = \pi(\boldsymbol{\theta}\,|\,\widetilde{\boldsymbol{\alpha}})$ for some $\widetilde{\boldsymbol{\alpha}}\in A$, with $\widetilde{\boldsymbol{\alpha}} = \widetilde{\boldsymbol{\alpha}}(\boldsymbol{\alpha},\boldsymbol{x})$.

As an example, consider the conjugate Bayesian analysis of the Gaussian distribution. We assume a likelihood

$$f(\boldsymbol{x}\,|\,u,s) = (2\pi s^2)^{-N/2}\exp\left\{-\frac{1}{2s^2}\sum_{j=1}^{N}(x_j-u)^2\right\} \tag{1.126}$$

with $s$ regarded as *fixed*. There is now a single parameter: $\boldsymbol{\theta} = \{u\}$. Now consider the prior distribution

$$\pi(u\,|\,\mu_0,\sigma_0) = (2\pi\sigma_0^2)^{-1/2}\,\exp\left\{-\frac{(u-\mu_0)^2}{2\sigma_0^2}\right\} \quad . \tag{1.127}$$

Note that the prior distribution is independent of the parameter $s$ and only depends on $u$ and the hyperparameters $\boldsymbol{\alpha} = (\mu_0,\sigma_0)$. We now compute the posterior:

$$\begin{aligned}\pi(u\,|\,\boldsymbol{x},s,\mu_0,\sigma_0) &\propto f(\boldsymbol{x}\,|\,u,s)\,\pi(u\,|\,\mu_0,\sigma_0)\\ &= \exp\left\{-\left(\frac{1}{2\sigma_0^2}+\frac{N}{2s^2}\right)u^2+\left(\frac{\mu_0}{\sigma_0^2}+\frac{N\langle x\rangle}{s^2}\right)u-\left(\frac{\mu_0^2}{2\sigma_0^2}+\frac{N\langle x^2\rangle}{2s^2}\right)\right\} \quad ,\end{aligned} \tag{1.128}$$

with $\langle x\rangle = N^{-1}\sum_{j=1}^N x_j$ and $\langle x^2\rangle = N^{-1}\sum_{j=1}^N x_j^2$. This is also a Gaussian distribution for $u$, and after supplying the appropriate normalization one finds

$$\pi(u\,|\,\boldsymbol{x},s,\mu_0,\sigma_0) = (2\pi\sigma_1^2)^{-1/2}\exp\left\{-\frac{(u-\mu_1)^2}{2\sigma_1^2}\right\} \quad , \tag{1.129}$$

which is normalized over $u$, with

$$\mu_1 = \mu_0 + \frac{N\big(\langle x\rangle-\mu_0\big)\sigma_0^2}{s^2+N\sigma_0^2} \qquad , \qquad \sigma_1^2 = \frac{s^2\sigma_0^2}{s^2+N\sigma_0^2} \quad . \tag{1.130}$$

Thus, the posterior is among the same family as the prior, and we have derived the update rule for the hyperparameters $(\mu_0,\sigma_0)\to(\mu_1,\sigma_1)$. Note that $\sigma_1<\sigma_0$, so the updated Gaussian prior is sharper than the original. The updated mean $\mu_1$ shifts in the direction of $\langle x\rangle$ obtained from the data set.

### 1.6.4 The problem with priors

We might think that, for the coin flipping problem, the flat prior $\pi(\theta)=1$ is an appropriate initial one, since it does not privilege any value of $\theta$. This prior therefore seems 'objective' or 'unbiased', also called 'uninformative'. But suppose we make a change of variables, mapping the interval $\theta\in[0,1]$ to the entire real line according to $\zeta = \log\big[\theta/(1-\theta)\big]$. In terms of the new parameter $\zeta$, we write the prior as $\tilde\pi(\zeta)$. Clearly $\pi(\theta)\,d\theta = \tilde\pi(\zeta)\,d\zeta$, so $\tilde\pi(\zeta) = \pi(\theta)\,d\theta/d\zeta$. For our example, find $\tilde\pi(\zeta) = \frac{1}{4}\mathrm{sech}^2(\zeta/2)$, which is not flat. Thus what was uninformative in terms of $\theta$ has become very informative in terms of the new parameter $\zeta$. Is there any truly unbiased way of selecting a Bayesian prior?

One approach, advocated by E. T. Jaynes, is to choose the prior distribution $\pi(\boldsymbol{\theta})$ according to the principle of maximum entropy. For continuous parameter spaces, we must first define a parameter space metric so as to be able to 'count' the number of different parameter states. The entropy of a distribution $\pi(\boldsymbol{\theta})$ is then dependent on this metric: $S = -\int d\mu(\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,\log\pi(\boldsymbol{\theta})$.

Another approach, due to Jeffreys, is to derive a parameterization-independent prior from the likelihood $f(\boldsymbol{x}\,|\,\boldsymbol{\theta})$ using the so-called *Fisher information matrix*,

$$I_{ij}(\boldsymbol{\theta}) = -\mathbb{E}_\theta\left(\frac{\partial^2 \log f(\boldsymbol{x}\,|\,\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\right) = -\int d\boldsymbol{x}\; f(\boldsymbol{x}\,|\,\boldsymbol{\theta})\;\frac{\partial^2 \log f(\boldsymbol{x}\,|\,\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j} \quad . \tag{1.131}$$

The *Jeffreys prior* $\pi_{\rm J}(\boldsymbol{\theta})$ is defined as

$$\pi_{\rm J}(\boldsymbol{\theta}) \propto \sqrt{\det I(\boldsymbol{\theta})} \quad . \tag{1.132}$$

One can check that the Jeffreys prior is invariant under reparameterization. As an example, consider the Bernoulli process, for which $\log f(\boldsymbol{x}\,|\,\theta) = X\log\theta + (N-X)\log(1-\theta)$, where $X = \sum_{j=1}^N x_j$. Then

$$-\frac{d^2 \log p(\boldsymbol{x}\,|\,\theta)}{d\theta^2} = \frac{X}{\theta^2} + \frac{N-X}{(1-\theta)^2} \quad , \tag{1.133}$$

and since $\mathbb{E}_\theta\, X = N\theta$, we have

$$I(\theta) = \frac{N}{\theta(1-\theta)} \qquad \Rightarrow \qquad \pi_{\rm J}(\theta) = \frac{1}{\pi}\,\frac{1}{\sqrt{\theta(1-\theta)}} \quad , \tag{1.134}$$

which felicitously corresponds to a Beta distribution with $\alpha = \beta = \frac{1}{2}$. In this example the Jeffreys prior turned out to be a conjugate prior, but in general this is not the case.

We can try to implement the Jeffreys procedure for a two-parameter family where each $x_j$ is normally distributed with mean $\mu$ and standard deviation $\sigma$. Let the parameters be $(\theta_1, \theta_2) = (\mu, \sigma)$. Then

$$-\log f(\boldsymbol{x}\,|\,\boldsymbol{\theta}) = N\log\sqrt{2\pi} + N\log\sigma + \frac{1}{2\sigma^2}\sum_{j=1}^N (x_j-\mu)^2 \quad , \tag{1.135}$$

and the Fisher information matrix is

$$I(\boldsymbol{\theta}) = -\frac{\partial^2 \log f(\boldsymbol{x}\,|\,\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j} = \begin{pmatrix} N\sigma^{-2} & \sigma^{-3}\sum_j (x_j-\mu) \\ \sigma^{-3}\sum_j (x_j-\mu) & -N\sigma^{-2} + 3\sigma^{-4}\sum_j (x_j-\mu)^2 \end{pmatrix} \quad . \tag{1.136}$$

Taking the expectation value, we have $\mathbb{E}\,(x_j-\mu) = 0$ and $\mathbb{E}\,(x_j-\mu)^2 = \sigma^2$, hence

$$\mathbb{E}\,I(\boldsymbol{\theta}) = \begin{pmatrix} N\sigma^{-2} & 0 \\ 0 & 2N\sigma^{-2} \end{pmatrix} \tag{1.137}$$

and the Jeffreys prior is $\pi_{\rm J}(\mu,\sigma) \propto \sigma^{-2}$. This is problematic because if we choose a flat metric on the $(\mu,\sigma)$ upper half-plane, the Jeffreys prior is not normalizable. Note also that the Jeffreys prior no longer resembles a Gaussian, and hence is not a conjugate prior.

# Chapter 2

# Thermodynamics

## 2.1 References

- E. Fermi, *Thermodynamics* (Dover, 1956)
  This outstanding and inexpensive little book is a model of clarity.
- A. H. Carter, *Classical and Statistical Thermodynamics*
  (Benjamin Cummings, 2000)
  A very relaxed treatment appropriate for undergraduate physics majors.
- H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*
  ($2^{\text{nd}}$ edition, Wiley, 1985)
  A comprehensive text appropriate for an extended course on thermodynamics.
- D. V. Schroeder, *An Introduction to Thermal Physics* (Addison-Wesley, 2000)
  An excellent thermodynamics text appropriate for upper division undergraduates. Contains many illustrative practical applications.
- D. Kondepudi and I. Prigogine, *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (Wiley, 1998)
  Lively modern text with excellent choice of topics and good historical content. More focus on chemical and materials applications than in Callen.
- L. E. Reichl, *A Modern Course in Statistical Physics* (2nd edition, Wiley, 1998)
  A graduate level text with an excellent and crisp section on thermodynamics.

## 2.2 What is Thermodynamics?

Thermodynamics is the study of relations among the *state variables* describing a thermodynamic system, and of transformations of heat into work and *vice versa*.

### 2.2.1 Thermodynamic systems and state variables

Thermodynamic systems contain large numbers of constituent particles, and are described by a set of *state variables* which describe the system's properties in an average sense. State variables are classified as being either *extensive* or *intensive*.

Extensive variables, such as volume $V$, particle number $N$, total internal energy $E$, magnetization $M$, *etc.*, scale linearly with the system size, *i.e.* as the first power of the system volume. If we take two identical thermodynamic systems, place them next to each other, and remove any barriers between them, then all the extensive variables will double in size.

Intensive variables, such as the pressure $p$, the temperature $T$, the chemical potential $\mu$, the electric field $\boldsymbol{E}$, *etc.*, are independent of system size, scaling as the zeroth power of the volume. They are the same throughout the system, if that system is in an appropriate state of *equilibrium*. The ratio of any two extensive variables is an intensive variable. For example, we write $n = N/V$ for the number density, which scales as $V^0$. Intensive variables may also be *inhomogeneous*. For example, $n(\boldsymbol{r})$ is the number density at position $\boldsymbol{r}$, and is defined as the limit of $\Delta N/\Delta V$ of the number of particles $\Delta N$ inside a volume $\Delta V$ which contains the point $\boldsymbol{r}$, in the limit $V \gg \Delta V \gg V/N$.

Heat and work are *not* state variables. That is to say, we cannot meaningfully speak of the heat or work of a thermodynamic system the way we can speak of its energy or pressure. Rather, what is meaningful is the heat $Q$ added to a system, say that supplied to a kettle of water on a hot stove, or the work $W$ done by a system, say by an expanding gas which pushes on a piston head. Both $Q$ and $W$ have dimensions of energy, and conservation of energy then entails that the internal energy change $\Delta E$ of a thermodynamic system is given by $\Delta E = Q - W$, which is the First Law of Thermodynamics. As a crude analogy, think of your bank account. You can deposit or withdraw funds in the form of checks or cash, which might be likened to thermodynamic heat and work, each having the units of US dollars. But the bank just keeps track of the total dollar value of your account, which is analogous to the total energy. You cannot say that your account consists of so much in checks, so much in bills, or, at a more refined level, so much in singles, fives, tens, twenties, *etc.*

Classically, the full motion of a system of $N$ point particles requires $6N$ variables to fully describe it ($3N$ positions and $3N$ velocities or momenta, in three space dimensions)[1]. Since the constituents are very small, $N$ is typically very large. A typical solid or liquid, for example, has a mass density on the order of $\varrho \sim 1\,\mathrm{g/cm^3}$; for gases, $\varrho \sim 10^{-3}\,\mathrm{g/cm^3}$. The constituent atoms have masses of $10^0$ to $10^2$ grams per mole, where one mole of $X$ contains $N_{\rm A}$ of $X$, and $N_{\rm A} = 6.022140765 \times 10^{23}$ is Avogadro's number[2]. Thus, for solids and liquids we roughly expect number densities $n$ of $10^{-2} - 10^0\,\mathrm{mol/cm^3}$ for

[1] For a system of $N$ molecules which can freely rotate, we must then specify $3N$ additional orientational variables – the Euler angles – and their $3N$ conjugate momenta. The dimension of phase space is then $12N$.

[2] Hence, 1 guacamole = $6.0221415 \times 10^{23}$ guacas.

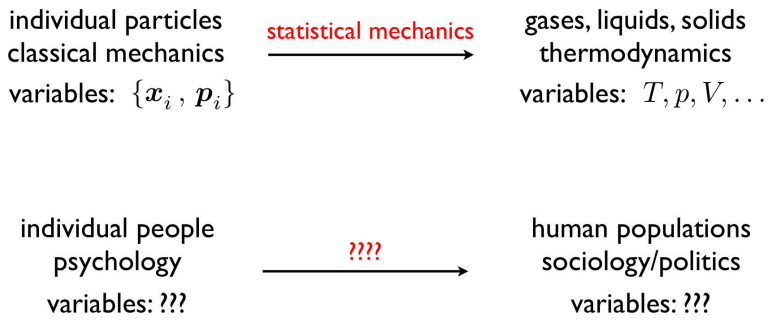



Figure 2.1: From microscale to macroscale : physical *versus* social sciences.

solids and liquids, and $10^{-5} - 10^{-3}\,\mathrm{mol/cm^3}$ for gases. Clearly we are dealing with fantastically large numbers of constituent particles in a typical thermodynamic system. The underlying theoretical basis for thermodynamics, where we use a small number of state variables to describe a system, is provided by the microscopic theory of statistical mechanics, which we shall study in the weeks ahead.

Intensive quantities such as $p$, $T$, and $n$ ultimately involve averages over both space and time. Consider for example the case of a gas enclosed in a container. We can measure the pressure (relative to atmospheric pressure) by attaching a spring to a moveable wall, as shown in fig. 2.2. From the displacement of the spring and the value of its spring constant $k$ we determine the force $F$. This force is due to the difference in pressures, so $p = p_0 + F/A$. Microscopically, the gas consists of constituent atoms or molecules, which are constantly undergoing collisions with each other and with the walls of the container. When a particle bounces off a wall, it imparts an impulse $2\hat{\boldsymbol{n}}(\hat{\boldsymbol{n}} \cdot \boldsymbol{p})$, where $\boldsymbol{p}$ is the particle's momentum and $\hat{\boldsymbol{n}}$ is the unit vector normal to the wall. (Only particles with $\boldsymbol{p} \cdot \hat{\boldsymbol{n}} > 0$ will hit the wall.) Multiply this by the number of particles colliding with the wall per unit time, and one finds the net force on the wall; dividing by the area gives the pressure $p$. Within the gas, each particle travels for a distance $\ell$, called the *mean free path*, before it undergoes a collision. We can write $\ell = \bar{v}\tau$, where $\bar{v}$ is the average particle speed and $\tau$ is the *mean free time*. When we study the kinetic theory of gases, we will derive formulas for $\ell$ and $\bar{v}$ (and hence $\tau$). For now it is helpful to quote some numbers to get an idea of the relevant distance and time scales. For $O_2$ gas at standard temperature and pressure ($T = 0^\circ$ C, $p = 1$ atm), the mean free path is $\ell \approx 1.1 \times 10^{-5}$ cm, the average speed is $\bar{v} \approx 425\,\mathrm{m/s}$, and the mean free time is $\tau \approx 2.6 \times 10^{-10}$ s. Thus, particles in the gas undergo collisions at a rate $\tau^{-1} \approx 4.0 \times 10^9\,\mathrm{s^{-1}}$. A measuring device, such as our spring, or a thermometer, effectively performs time and space averages. If there are $N_\mathrm{c}$ collisions with a particular patch of wall during some time interval on which our measurement device responds, then the root mean square relative fluctuations in the local pressure will be on the order of $N_\mathrm{c}^{-1/2}$ times the average. Since $N_\mathrm{c}$ is a very large number, the fluctuations are negligible.

If the system is in *steady state*, the state variables do not change with time. If furthermore there are no macroscopic currents of energy or particle number flowing through the system, the system is said to be in *equilibrium*. A continuous succession of equilibrium states is known as a *thermodynamic path*, which can be represented as a smooth curve in a multidimensional space whose axes are labeled by state variables. A thermodynamic *process* is any change or succession of changes which results in a change of the state variables. In a *cyclic* process, the initial and final states are the same. In a *quasistatic* process, the system passes through a continuous succession of equilibria. A *reversible* process is one where the external conditions and the thermodynamic path of the system can be reversed; it is both quasi-static

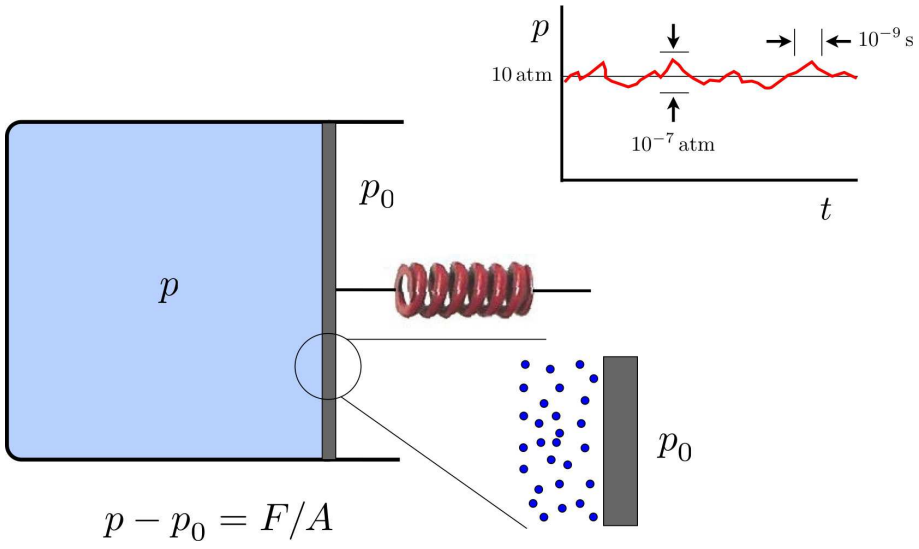



Figure 2.2: The pressure $p$ of a gas is due to an average over space and time of the impulses due to the constituent particles.

*and* non-dissipative (*i.e.* no friction). The slow expansion of a gas against a piston head, whose counterforce is always infinitesimally less than the force $pA$ exerted by the gas, is reversible. To reverse this process, we simply add infinitesimally more force to $pA$ and the gas compresses. An example of a quasistatic process which is not reversible: slowly dragging a block across the floor, or the slow leak of air from a tire. Irreversible processes, as a rule, are dissipative. Oftentimes we will concern ourselves with infinitesimal amounts of heat and work, and how these are related to infinitesimal changes in state variables. Expressed in terms of differentials, there are several special processes we will discuss:

$$\begin{aligned} \textit{reversible: } đQ &= T\,dS & \qquad \textit{isothermal: } dT &= 0 \\ \textit{spontaneous: } đQ &< T\,dS & \qquad \textit{isochoric: } dV &= 0 \\ \textit{adiabatic: } đQ &= 0 & \qquad \textit{isobaric: } dp &= 0 \quad . \end{aligned}$$

If $\phi$ is a state variable or a function of state variables, its infinitesimal change is expressed as $d\phi$, which is an *exact differential*. If $\phi$ is not a state variable, then its differential is in general *inexact*, and we express its infinitesimal change as $đ\phi$. We shall discuss the entropy $S$ and its relation to irreversibility in §2.6.

### 2.2.2 Mathematical interlude : exact and inexact differentials

The differential

$$dF = \sum_{i=1}^{k} A_i\,dx_i \tag{2.1}$$

is called *exact* if there is a smooth function[3] $F(x_1, \ldots, x_k)$ whose differential gives the right hand side of eqn. 2.1. On a simply connected open domain, a differential with continuously differentiable coefficients

[3]Smoothness means that all partial derivatives exist and are continuous. In fact, this condition is needlessly stringent for our purposes. For exactness we only require that all the second partial derivatives exist and are

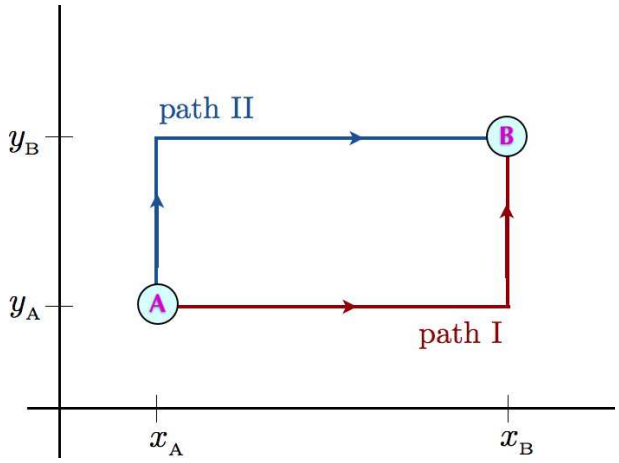



Figure 2.3: Two distinct paths with identical endpoints.

$A_i$ is exact if and only if its cross derivatives agree, *viz.*

$$A_i = \frac{\partial F}{\partial x_i} \qquad \Longleftrightarrow \qquad \frac{\partial A_i}{\partial x_j} = \frac{\partial A_j}{\partial x_i} = \frac{\partial^2 F}{\partial x_i \partial x_j} \quad \forall\, i,j \tag{2.2}$$

for some single-valued function $F$. For exact differentials, the integral between fixed endpoints is path-independent:

$$\int_{\rm A}^{\rm B} dF = F(x_1^{\rm B},\ldots,x_k^{\rm B}) - F(x_1^{\rm A},\ldots,x_k^{\rm A}) \quad , \tag{2.3}$$

from which it follows that the integral of $dF$ around any closed path must vanish: $\oint dF = 0$.

When the cross derivatives are not identical, *i.e.* when $\partial A_i/\partial x_j \neq \partial A_j/\partial x_i$, the differential is *inexact*. In this case, the integral of $dF$ is path dependent, and does not depend solely on the endpoints.

As an example, consider the differential $dF = K_1\, y\, dx + K_2\, x\, dy$. Let's evaluate the integral of $dF$ along each of the two paths in fig. 2.3:

$$F^{(\rm I)} = \int\limits_{\mathcal{C}_{\rm I}} dF = K_1 \int\limits_{x_{\rm A}}^{x_{\rm B}} dx\; y_{\rm A} + K_2 \int\limits_{y_{\rm A}}^{y_{\rm B}} dy\; x_{\rm B} = K_1\, y_{\rm A}\,(x_{\rm B} - x_{\rm A}) + K_2\, x_{\rm B}\,(y_{\rm B} - y_{\rm A}) \tag{2.4}$$

$$F^{(\rm II)} = \int\limits_{\mathcal{C}_{\rm II}} dF = K_1 \int\limits_{x_{\rm A}}^{x_{\rm B}} dx\; y_{\rm B} + K_2 \int\limits_{y_{\rm A}}^{y_{\rm B}} dy\; x_{\rm A} = K_1\, y_{\rm B}\,(x_{\rm B} - x_{\rm A}) + K_2\, x_{\rm A}\,(y_{\rm B} - y_{\rm A}) \quad , \tag{2.5}$$

continuous, which entails eqn. 2.2. The domain matters for global exactness. For example, the differential $(x\,dy - y\,dx)/(x^2+y^2)$ has equal cross derivatives on the punctured plane, but its integral around a counterclockwise circle enclosing the origin is $2\pi$, so it is not exact there. Locally it is the differential $d\theta$ of an angular coordinate $\theta$, but no single-valued angular coordinate exists on the entire punctured plane.

where $\mathcal{C}_{\rm I,II}$ are, respectively, paths I and II. Note that in general $F^{\rm (I)} \neq F^{\rm (II)}$. Thus, if we start at point A, the value of $F$ at point B will depend on the path taken, since the change $\Delta F = \int_{\rm A}^{\rm B} dF$ is path-dependent. The difference in $\Delta F$ associated with each of the two paths is

$$F^{\rm (I)} - F^{\rm (II)} = \oint_{\mathcal{C}} dF = (K_2 - K_1)\,(x_{\rm B} - x_{\rm A})\,(y_{\rm B} - y_{\rm A}) \quad , \tag{2.6}$$

where $\mathcal{C} = \mathcal{C}_{\rm II}^{-1} \circ \mathcal{C}_{\rm I}$ is the closed path formed by appending the reverse of path II (*i.e.* $C_{\rm II}^{-1}$) to path I. Thus, we see that if $K_1 = K_2$, the work is the same for the two paths. In fact, if $K_1 = K_2$, the work would be path-independent, and would depend only on the endpoints. This is true for *any* path, and not just piecewise linear paths of the type depicted in fig. 2.3. Thus, if $K_1 = K_2$, we are justified in using the notation $dF$ for the differential. Explicitly, we then have $F = K_1 xy$. However, if $K_1 \neq K_2$, the differential is inexact, and we will henceforth write $đF$ in such cases.

### 2.2.3 Equations of state

How many state variables are necessary to fully specify the equilibrium state of a thermodynamic system? For a single component system, such as water which is composed of one molecular species, the answer is three. These can be taken to be $T$, $p$, and $V$. One always must specify at least one extensive variable, else we cannot determine the overall size of the system. We shall discuss this question in detail in §2.3.1 and §2.11.5 below. For a multicomponent system with $\mathsf{c}$ different species, we must specify $\mathsf{c}+2$ state variables, which may be $\{T, p, N_1, \ldots, N_{\mathsf{c}}\}$, where $N_a$ is the number of particles of species $a$. Another possibility is the set $\{T, p, V, x_1, \ldots, x_{\mathsf{c}-1}\}$, where the *concentration* of species $a$ is $x_a = N_a/N$. Here, $N = \sum_{a=1}^{\mathsf{c}} N_a$ is the total number of particles. Note that $\sum_{a=1}^{\mathsf{c}} x_a = 1$.

It then follows that if we specify more than $\mathsf{c}+2$ state variables, there must exist a relation among them. Such relations are known as *equations of state*. The most famous example is the ideal gas law,

$$pV = Nk_{\rm B}T \quad , \tag{2.7}$$

relating the four state variables $T$, $p$, $V$, and $N$. Here $k_{\rm B} = 1.380649 \times 10^{-16}\,\mathrm{erg/K}$ is Boltzmann's constant. Another example is the van der Waals equation,

$$\left(p + \frac{\bar{a}N^2}{V^2}\right)(V - \bar{b}N) = Nk_{\rm B}T \quad , \tag{2.8}$$

where $\bar{a}$ and $\bar{b}$ are constants which depend on the molecule which forms the gas. For a third example, consider a paramagnet, where

$$\frac{M}{V} = \frac{CH}{T} \quad , \tag{2.9}$$

where $M$ is the magnetization, $H$ the magnetic field, and $C$ the Curie constant.

Any quantity which, in equilibrium, depends only on the state variables is called a *state function*. For example, the total internal energy $E$ of a thermodynamic system is a state function, and we may write $E = E(T, p, V)$. State functions can also serve as state variables, although the most natural state variables are those which can be directly measured.

### 2.2.4 Heat

Once thought to be a type of fluid, heat is now understood in terms of the kinetic theory of gases, liquids, and solids as a form of energy stored in the disordered motion of constituent particles. The units of heat are therefore units of energy, and it is appropriate to speak of *heat energy*, which we shall simply abbreviate as *heat*:[4]

$$1\,\mathrm{J} = 10^7\,\mathrm{erg} = 6.242\times10^{18}\,\mathrm{eV} = 2.390\times10^{-4}\,\mathrm{kcal} = 9.478\times10^{-4}\,\mathrm{BTU} \quad . \tag{2.10}$$

We will use the symbol $Q$ to denote the amount of heat energy absorbed by a system during some given thermodynamic process, and $đQ$ to denote a differential amount of heat energy. The symbol $đ$ indicates an inexact differential (see §2.2.2). This means that heat is not a state function: there is no 'heat function' $Q(T,p,V)$.

### 2.2.5 Work

In general we will write the differential element of work $đW$ done *by* the system as

$$đW = \sum_i F_i\,dX_i \quad , \tag{2.11}$$

where $F_i$ is a *generalized force* and $dX_i$ a *generalized displacement*[5]. The generalized forces and displacements are themselves state variables, and by convention we will take the generalized forces to be *intensive* and the generalized displacements to be *extensive*. As an example, in a simple one-component system, we have $đW = p\,dV$. More generally, we write

$$đW = \overbrace{\big(p\,dV - \boldsymbol{H}\cdot d\boldsymbol{M} - \boldsymbol{E}\cdot d\boldsymbol{P} - \sigma\,dA + \ldots\big)}^{-\sum_j y_j\,dX_j} - \overbrace{\big(\mu_1\,dN_1 + \mu_2\,dN_2 + \ldots\big)}^{\sum_a \mu_a\,dN_a} \tag{2.12}$$

Here we distinguish between two types of work. The first involves changes in quantities such as volume, magnetization, electric polarization, area, *etc.* The conjugate forces $y_i$ applied *to* the system are then $-p$, the magnetic field $\boldsymbol{H}$, the electric field $\boldsymbol{E}$, the surface tension $\sigma$, respectively. The second type of work involves changes in the number of constituents of a given species. For example, energy is required in order to dissociate two hydrogen atoms in an $\mathrm{H}_2$ molecule. The effect of such a process is $dN_{\mathrm{H}_2} = -1$ and $dN_{\mathrm{H}} = +2$. Thus, $đW = -\boldsymbol{y}\cdot d\boldsymbol{X} - \boldsymbol{\mu}\cdot d\boldsymbol{N}$, with $\boldsymbol{X} = \{X_1,\ldots,X_{\mathsf{d}}\}$ and $N = \{N_1,\ldots,N_{\mathsf{c}}\}$. In this general setting there are $\mathsf{d}$ generalized displacements and $\mathsf{c}$ component species. When $\mathsf{d} = 1$, the volume is the only displacement.

As with heat, $đW$ is an inexact differential, and work $W$ is not a state variable, since it is path-dependent. There is no 'work function' $W(T,p,V)$.

[4] One calorie (cal) is the amount of heat needed to raise $1\,\mathrm{g}$ of $\mathrm{H_2O}$ from $T_0 = 14.5^\circ\,\mathrm{C}$ to $T_1 = 15.5^\circ\,\mathrm{C}$ at a pressure of $p_0 = 1\,\mathrm{atm}$. One British Thermal Unit (BTU) is the amount of heat needed to raise $1\,\mathrm{lb}$. of $\mathrm{H_2O}$ from $T_0 = 63^\circ\,\mathrm{F}$ to $T_1 = 64^\circ\,\mathrm{F}$ at a pressure of $p_0 = 1\,\mathrm{atm}$.

[5] We use the symbol $đ$ in the differential $đW$ to indicate that this is not an exact differential. More on this in section 2.2.2.

### 2.2.6 Pressure and temperature

The units of pressure ($p$) are force per unit area. The SI unit is the Pascal (Pa): $1\,\mathrm{Pa} = 1\,\mathrm{N/m^2} = 1\,\mathrm{kg/m\,s^2}$. Other units of pressure we will encounter:

$$1\,\mathrm{bar} \equiv 10^5\,\mathrm{Pa} \quad , \quad 1\,\mathrm{atm} \equiv 1.01325 \times 10^5\,\mathrm{Pa} \quad , \quad 1\,\mathrm{torr} \equiv 133.3\,\mathrm{Pa} \quad .$$

Temperature ($T$) has a very precise definition from the point of view of statistical mechanics, as we shall see. Many physical properties depend on the temperature – such properties are called *thermometric properties*. For example, the resistivity of a metal $\rho(T,p)$ or the number density of a gas $n(T,p)$ are both thermometric properties, and can be used to define a temperature scale. Consider the device known as the 'constant volume gas thermometer' depicted in fig. 2.4, in which the volume or pressure of a gas may be used to measure temperature. The gas is assumed to be in equilibrium at some pressure $p$, volume $V$, and temperature $T$. An incompressible fluid of density $\varrho$ is used to measure the pressure difference $\Delta p = p - p_0$, where $p_0$ is the ambient pressure at the top of the reservoir:

$$p - p_0 = \varrho g(h_2 - h_1) \quad , \tag{2.13}$$

where $g$ is the acceleration due to gravity. The height $h_1$ of the left column of fluid in the U-tube provides a measure of the change in the volume of the gas:

$$V(h_1) = V(0) - Ah_1 \quad , \tag{2.14}$$

where $A$ is the (assumed constant) cross-sectional area of the left arm of the U-tube. The device can operate in two modes:

- Constant pressure mode : The height of the reservoir is adjusted so that the height difference $h_2 - h_1$ is held constant. This fixes the pressure $p$ of the gas. The gas volume still varies with temperature $T$, and we can define
$$\frac{T}{T_{\rm ref}} = \frac{V}{V_{\rm ref}} \quad , \tag{2.15}$$
where $T_{\rm ref}$ and $V_{\rm ref}$ are the reference temperature and volume, respectively.
- Constant volume mode : The height of the reservoir is adjusted so that $h_1 = 0$, hence the volume of the gas is held fixed, and the pressure varies with temperature. We then define
$$\frac{T}{T_{\rm ref}} = \frac{p}{p_{\rm ref}} \quad , \tag{2.16}$$
where $T_{\rm ref}$ and $p_{\rm ref}$ are the reference temperature and pressure, respectively.

What should we use for a reference? One might think that a pot of boiling water will do, but anyone who has gone camping in the mountains knows that water boils at lower temperatures at high altitude (lower pressure). This phenomenon is reflected in the *phase diagram* for $H_2O$, depicted in fig. 2.5. There are two special points in the phase diagram, however. One is the *triple point*, where the solid, liquid, and vapor (gas) phases all coexist. The second is the *critical point*, which is the terminus of the curve separating

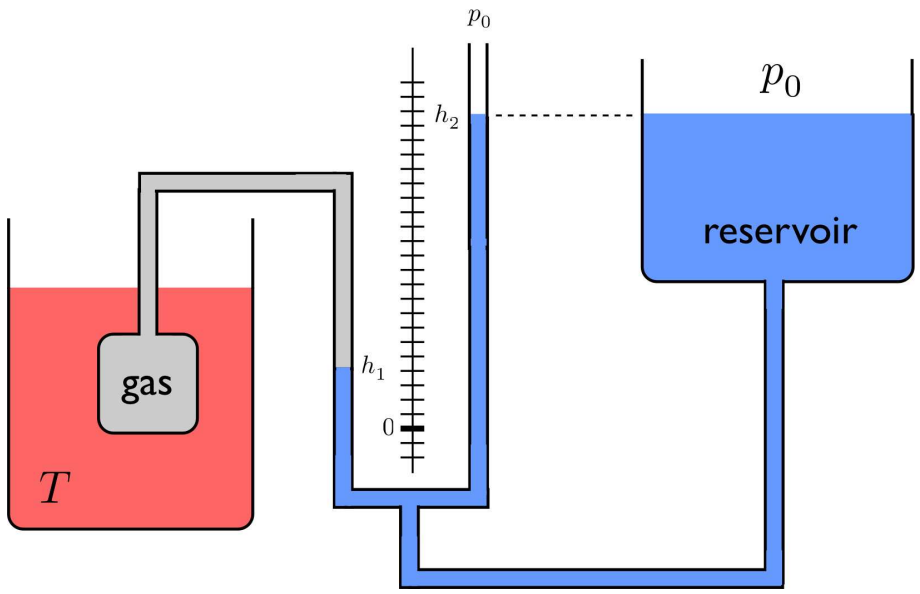



Figure 2.4: The constant volume gas thermometer. The gas is placed in thermal contact with an object of temperature $T$. An incompressible fluid of density $\varrho$ is used to measure the pressure difference $\Delta p = p_{\rm gas} - p_0$.

liquid from gas. At the critical point, the *latent heat of transition* between liquid and gas phases vanishes (more on this later on). The triple point temperature $T_{\rm t}$ is thus *unique* and until 2019 was defined to be $T_{\rm t} = 273.16\,{\rm K}$. The pressure at the triple point is $611.7\,{\rm Pa} = 6.037 \times 10^{-3}\,{\rm atm}$. Since 2019, the kelvin has been defined by fixing Boltzmann's constant.

A question remains: are the two modes of the thermometer compatible? *E.g.* if we boil water at $p = p_0 = 1\,{\rm atm}$, do they yield the same value for $T$? And what if we use a different gas in our measurements? In fact, all these measurements will in general be *incompatible*, yielding different results for the temperature $T$. However, in the limit that we use a very low density gas, all the results converge. This is because all low density gases behave as *ideal gases*, and obey the ideal gas equation of state $pV = Nk_{\rm B}T$.

### 2.2.7 Standard temperature and pressure

It is customary in the physical sciences to define certain standard conditions with respect to which any arbitrary conditions may be compared. In thermodynamics, there is a notion of *standard temperature and pressure*, abbreviated STP. Unfortunately, there are two *different* definitions of STP currently in use, one from the International Union of Pure and Applied Chemistry (IUPAC), and the other from the U.S. National Institute of Standards and Technology (NIST). The two standards are:

$$\begin{aligned} {\rm IUPAC}\ &:\ T_0 = 0^\circ\,{\rm C} = 273.15\,{\rm K} \quad , \quad p_0 = 10^5\,{\rm Pa} \\ {\rm NIST}\ &:\ T_0 = 20^\circ\,{\rm C} = 293.15\,{\rm K} \quad , \quad p_0 = 1\,{\rm atm} = 1.01325 \times 10^5\,{\rm Pa} \end{aligned}$$

To make matters worse, in the past it was customary to define STP as $T_0 = 0^\circ\,{\rm C}$ and $p_0 = 1\,{\rm atm}$. We will use the NIST definition in this course. Unless I slip and use the IUPAC definition. Figuring out what I

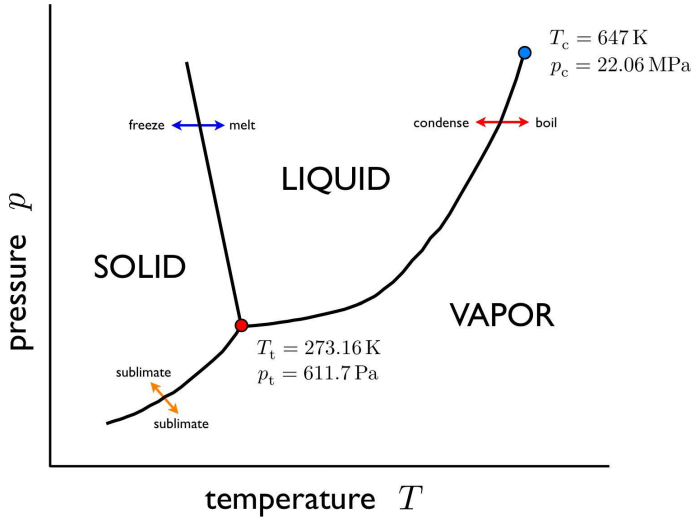



Figure 2.5: A sketch of the phase diagram of $H_2O$ (water). Two special points are identified: the triple point $(T_t, p_t)$ at which there is three phase coexistence, and the critical point $(T_c, p_c)$, where the latent heat of transformation from liquid to gas vanishes. Not shown are transitions between several different solid phases.

mean by STP will keep you on your toes.

The volume of one mole of ideal gas at STP is then

$$V = \frac{N_A k_B T_0}{p_0} = \begin{cases} 22.711\ \ell & \text{(IUPAC)} \\ 24.219\ \ell & \text{(NIST)} \end{cases}, \tag{2.17}$$

where $1\,\ell = 10^3\,\mathrm{cm}^3 = 10^{-3}\,\mathrm{m}^3$ is one liter. Under the old definition of STP as $T_0 = 0^\circ\,\mathrm{C}$ and $p_0 = 1\,\mathrm{atm}$, the volume of one mole of gas at STP is $22.414\ \ell$, which is a figure I remember from my $10^{\mathrm{th}}$ grade chemistry class with Mr. Lawrence.

## 2.3 The Zeroth Law of Thermodynamics

Equilibrium is established by the exchange of energy, volume, or particle number between different systems or subsystems:

$$\begin{aligned} &\text{energy exchange} &&\Longrightarrow &&T = \text{ constant} &&\Longrightarrow &&\text{thermal equilibrium} \\ &\text{volume exchange} &&\Longrightarrow &&\frac{p}{T} = \text{ constant} &&\Longrightarrow &&\text{mechanical equilibrium} \\ &\text{particle exchange} &&\Longrightarrow &&\frac{\mu}{T} = \text{ constant} &&\Longrightarrow &&\text{chemical equilibrium} \end{aligned}$$

Equilibrium is transitive, so

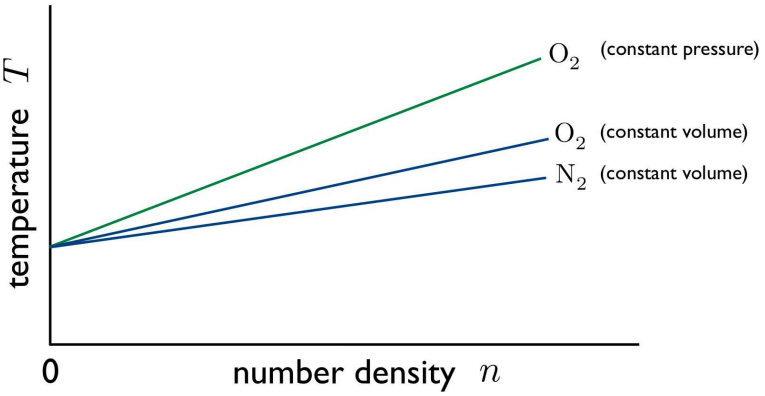



Figure 2.6: As the gas density tends to zero, the readings of the constant volume gas thermometer converge.

**If A is in equilibrium with B, and B with C, then A is in equilibrium with C.**

This is known as the Zeroth Law of Thermodynamics[6].

### 2.3.1 Gibbs phase rule

We see in fig. 2.5 that there are special lower-dimensional subsets of the phase diagram (*i.e.* curves and points) where different phases may *coexist* – a nongeneric state of affairs. For a system with $\mathsf{c}$ distinct components, $\mathsf{c}+2$ state variables are required for a complete specification of any equilibrium phase, at least one of which must be extensive. If we choose *only* one extensive variable, say the volume $V$, then we may construct a *phase diagram* where the axes correspond to the $\mathsf{c}+1$ quantities $\{T, p, \mu_1, \ldots, \mu_{\mathsf{c}-1}\}$. The Gibbs-Duhem relation (see eqn. 2.111) then guarantees that *any* intensive quantity $\phi$ may be expressed as a state function $\phi = \phi(T, p, \mu_1, \ldots, \mu_{\mathsf{c}-1})$. In particular, the $\mathsf{c}^{\rm th}$ chemical potential $\mu_{\mathsf{c}}$ may be expressed as such a state function. The phase diagram doesn't depend on the extensive variable – $H_2O$ at atmospheric pressure freezes at $0^\circ$ C regardless of the amount of substance present.

As we shall derive in §2.11.5 below, a number $\varphi$ distinct thermodynamic phases may coexist in a state of *phase equilibrium* throughout a $d(\mathsf{c}, \varphi) = \mathsf{c} + 2 - \varphi$ dimensional subset of the full $(\mathsf{c}+1)$-dimensional space $\Sigma = \{T, p, \mu_1, \ldots, \mu_{\mathsf{c}-1}\}$[7]. Since $d \geq 0$, we must have $\varphi \leq \mathsf{c} + 2$. Thus, for a single component ($\mathsf{c} = 1$) system like $H_2O$, we can have up to three-phase coexistence ($\varphi \leq 3$), as exemplified by the existence of a triple point in the phase diagram. There are also curves along which there is two-phase coexistence ($\varphi = 2$) between solid and liquid, liquid and gas, and solid and gas. Away from these special phase coexistence regions, full equilibrium entails that $p$, $T$, and $\mu$ are constant throughout the system. Furthermore, *all* intensive quantities, such as the specific volume $v = n^{-1}$ or energy density $\varepsilon$, are also

[6] As we shall see further below, thermomechanical equilibrium in fact leads to constant $p/T$, and thermochemical equilibrium to constant $\mu/T$. If there is thermal equilibrium, then $T$ is already constant, and so thermomechanical and thermochemical equilibria then guarantee the constancy of $p$ and $\mu$.

[7] Thus the codimension of the phase coexistence space is $\bar{d}(\mathsf{c}, \varphi) = \mathsf{c} + 1 - d(\mathsf{c}, \varphi) = \varphi - 1$. Thus for $\varphi = 1$ we have $\bar{d} = 0$.

constant. However, in the phase coexistence regions, this latter condition is no longer the case. Along the liquid-gas boundary, for example, the liquid and gas regions – in the presence of a gravitational field, the denser liquid component will separate out at the bottom – will have *different* specific volumes and energy densities.

## 2.4 The First Law of Thermodynamics

### 2.4.1 Conservation of energy

The first law is a statement of energy conservation, and is depicted in fig. 2.7. It says, quite simply, that during a thermodynamic process, the change in a system's internal energy $E$ is given by the heat energy $Q$ added *to* the system, minus the work $W$ done *by* the system:

$$\Delta E = Q - W \quad . \tag{2.18}$$

The differential form of this, the First Law of Thermodynamics, is

$$dE = đQ - đW \quad . \tag{2.19}$$

We use the symbol $đ$ in the differentials $đQ$ and $đW$ to remind us that these are inexact differentials. The energy $E$, however, is a state function, hence $dE$ is an exact differential.

Consider a volume $V$ of fluid held in a flask, initially at temperature $T_0$, and held at atmospheric pressure. The internal energy is then $E_0 = E(T_0, p, V)$. Now let us contemplate changing the temperature in two different ways. The first method (A) is to place the flask on a hot plate until the temperature of the fluid rises to a value $T_1$. The second method (B) is to stir the fluid vigorously. In the first case, we add heat $Q_{\rm A} > 0$ but no work is done[8], so $W_{\rm A} = 0$. In the second case, if we thermally insulate the flask and use a stirrer of very low thermal conductivity, then no heat is added, *i.e.* $Q_{\rm B} = 0$. However, the stirrer does work $-W_{\rm B} > 0$ on the fluid (remember $W$ is the work done by the system). If we end up at the same temperature $T_1$, then the final energy is $E_1 = E(T_1, p, V)$ in both cases. We then have

$$\Delta E = E_1 - E_0 = Q_{\rm A} = -W_{\rm B} \quad . \tag{2.20}$$

It also follows that for any cyclic transformation, where the state variables are the same at the beginning and the end, we have

$$\Delta E_{\rm cyclic} = Q - W = 0 \quad \Longrightarrow \quad Q = W \quad \text{(cyclic)} \quad . \tag{2.21}$$

### 2.4.2 Single component systems

A single component system is specified by three state variables. In many applications, the total number of particles $N$ is conserved, so it is useful to take $N$ as one of the state variables. The remaining two can

[8] In fact, the assumptions of constant pressure and zero work are inconsistent here. A fluid heated at atmospheric pressure generally will expand, so work $W_{\rm A} = p\,\Delta V$ is done. Typically this is a small amount if the fluid remains in its liquid state. More correctly, we can replace our flask with a rigid, closed vessel.

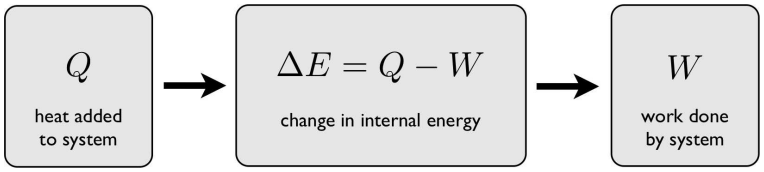



Figure 2.7: The first law of thermodynamics is a statement of energy conservation.

be $(T, V)$ or $(T, p)$ or $(p, V)$. The differential form of the first law says

$$\begin{aligned} dE &= đQ - đW \\ &= đQ - p\,dV + \mu\,dN \quad . \end{aligned} \tag{2.22}$$

The quantity $\mu$ is called the *chemical potential*. We ask: how much heat is required in order to make an infinitesimal change in temperature, pressure, volume, or particle number? We start by rewriting eqn. 2.22 as

$$đQ = dE + p\,dV - \mu\,dN \quad . \tag{2.23}$$

We now must roll up our sleeves and do some work with partial derivatives.

• $(T, V, N)$ systems : If the state variables are $(T, V, N)$, we write

$$dE = \left(\frac{\partial E}{\partial T}\right)_{V,N} dT + \left(\frac{\partial E}{\partial V}\right)_{T,N} dV + \left(\frac{\partial E}{\partial N}\right)_{T,V} dN \quad . \tag{2.24}$$

Then

$$đQ = \left(\frac{\partial E}{\partial T}\right)_{V,N} dT + \left[\left(\frac{\partial E}{\partial V}\right)_{T,N} + p\right] dV + \left[\left(\frac{\partial E}{\partial N}\right)_{T,V} - \mu\right] dN \quad . \tag{2.25}$$

• $(T, p, N)$ systems : If the state variables are $(T, p, N)$, we write

$$dE = \left(\frac{\partial E}{\partial T}\right)_{p,N} dT + \left(\frac{\partial E}{\partial p}\right)_{T,N} dp + \left(\frac{\partial E}{\partial N}\right)_{T,p} dN \quad . \tag{2.26}$$

We also write

$$dV = \left(\frac{\partial V}{\partial T}\right)_{p,N} dT + \left(\frac{\partial V}{\partial p}\right)_{T,N} dp + \left(\frac{\partial V}{\partial N}\right)_{T,p} dN \quad . \tag{2.27}$$

Then

$$\begin{aligned} đQ = &\left[\left(\frac{\partial E}{\partial T}\right)_{p,N} + p\left(\frac{\partial V}{\partial T}\right)_{p,N}\right] dT + \left[\left(\frac{\partial E}{\partial p}\right)_{T,N} + p\left(\frac{\partial V}{\partial p}\right)_{T,N}\right] dp \\ &+ \left[\left(\frac{\partial E}{\partial N}\right)_{T,p} + p\left(\frac{\partial V}{\partial N}\right)_{T,p} - \mu\right] dN \quad . \end{aligned} \tag{2.28}$$

• $(p, V, N)$ systems : If the state variables are $(p, V, N)$, we write

$$dE = \left(\frac{\partial E}{\partial p}\right)_{V,N} dp + \left(\frac{\partial E}{\partial V}\right)_{p,N} dV + \left(\frac{\partial E}{\partial N}\right)_{p,V} dN \quad . \tag{2.29}$$

| SUBSTANCE | $c_p$ (J/mol K) | $\tilde{c}_p$ (J/g K) | SUBSTANCE | $c_p$ (J/mol K) | $\tilde{c}_p$ (J/g K) |
|---|---|---|---|---|---|
| Air | 29.07 | 1.01 | $H_2O$ (25° C) | 75.34 | 4.181 |
| Aluminum | 24.2 | 0.897 | $H_2O$ ($100^\circ_+$ C) | 37.47 | 2.08 |
| Copper | 24.47 | 0.385 | Iron | 25.1 | 0.450 |
| $CO_2$ | 36.94 | 0.839 | Lead | 26.4 | 0.127 |
| Diamond | 6.115 | 0.509 | Lithium | 24.8 | 3.58 |
| Ethanol | 112 | 2.44 | Neon | 20.786 | 1.03 |
| Gold | 25.42 | 0.129 | Oxygen | 29.38 | 0.918 |
| Helium | 20.786 | 5.193 | Paraffin (wax) | 900 | 2.5 |
| Hydrogen | 28.82 | 14.30 | Uranium | 27.7 | 0.116 |
| $H_2O$ (−10° C) | 38.09 | 2.05 | Zinc | 25.3 | 0.387 |

Table 2.1: Specific heat (at 25° C, unless otherwise noted) of some common substances. (*Source: Wikipedia.*)

Then

$$đQ = \left(\frac{\partial E}{\partial p}\right)_{V,N} dp + \left[\left(\frac{\partial E}{\partial V}\right)_{p,N} + p\right] dV + \left[\left(\frac{\partial E}{\partial N}\right)_{p,V} - \mu\right] dN \quad . \tag{2.30}$$

### 2.4.3 Heat capacity

The *heat capacity* of a body, $C$, is by definition the ratio $đQ/dT$ of the amount of heat absorbed by the body to the associated infinitesimal change in temperature $dT$. The heat capacity will in general be different if the body is heated at constant volume or at constant pressure. Setting $dV = 0$ gives, from eqn. 2.25,

$$C_{V,N} = \left(\frac{đQ}{dT}\right)_{V,N} = \left(\frac{\partial E}{\partial T}\right)_{V,N} \quad . \tag{2.31}$$

Similarly, if we set $dp = 0$, then eqn. 2.28 yields

$$C_{p,N} = \left(\frac{đQ}{dT}\right)_{p,N} = \left(\frac{\partial E}{\partial T}\right)_{p,N} + p\left(\frac{\partial V}{\partial T}\right)_{p,N} \quad . \tag{2.32}$$

Unless otherwise stated, we shall assume that $N$ is fixed, and will write $C_V$ for $C_{V,N}$ and $C_p$ for $C_{p,N}$.

The units of heat capacity are energy divided by temperature, *e.g.* J/K. The heat capacity is an extensive quantity, scaling with the size of the system. If we divide by the number of moles $N/N_{\rm A}$, we obtain the *molar heat capacity*, sometimes called the *molar specific heat*: $c = C/\nu$, where $\nu = N/N_{\rm A}$ is the number of moles of substance. Specific heat is also sometimes quoted in units of heat capacity per gram of

substance. We shall define

$$\tilde{c} = \frac{C}{mN} = \frac{c}{M} = \frac{\text{heat capacity per mole}}{\text{mass per mole}} \quad . \tag{2.33}$$

Here $m$ is the mass per particle and $M$ is the mass per mole: $M = N_{\rm A}\, m$.

Suppose we raise the temperature of a body from $T = T_{\rm A}$ to $T = T_{\rm B}$. How much heat is required? We have

$$Q = \int_{T_{\rm A}}^{T_{\rm B}} dT\, C(T) \quad , \tag{2.34}$$

where $C = C_V$ or $C = C_p$ depending on whether volume or pressure is held constant. For ideal gases with no internal degrees of freedom, as we shall discuss below, $C(T)$ is constant, and thus

$$Q = C(T_{\rm B} - T_{\rm A}) \quad \Longrightarrow \quad T_{\rm B} = T_{\rm A} + \frac{Q}{C} \quad . \tag{2.35}$$

In metals at very low temperatures one finds $C = \gamma T$, where $\gamma$ is a constant[9]. We then have

$$Q = \int_{T_{\rm A}}^{T_{\rm B}} dT\, C(T) = \tfrac{1}{2}\gamma\big(T_{\rm B}^2 - T_{\rm A}^2\big) \quad \Longrightarrow \quad T_{\rm B} = \sqrt{T_{\rm A}^2 + 2\gamma^{-1}Q} \quad . \tag{2.36}$$

**Calorimetry**

We measure specific heat using a device called a *calorimeter*, of which there are many types. The basic physics is that a substance initially in equilibrium at temperature $T_1$ put in thermal contact with a reference fluid (typically $H_2O$) in equilibrium at temperature $T_2 > T_1$ will come to equilibrium at some final temperature $T_{\rm f} \in (T_1, T_2)$. If the substance and the reference fluid are thermally isolated from the rest of the universe, then the heat energy gained by the solid is the heat energy lost by the reference fluid. Assuming the specific heats $\tilde{c}_{\rm s,r}(T)$ are constant throughout the regime $T \in [T_1, T_2]$, we then have

$$Q = M_{\rm s}\,\tilde{c}_{\rm s}\cdot(T_{\rm f} - T_1) = M_{\rm r}\,\tilde{c}_{\rm r}\cdot(T_2 - T_{\rm f}) \quad \Rightarrow \quad \tilde{c}_{\rm s} = \frac{M_{\rm r}}{M_{\rm s}}\cdot\frac{T_2 - T_{\rm f}}{T_{\rm f} - T_1} \quad . \tag{2.37}$$

Thus, a measurement of the final temperature $T_{\rm f}$ can be used to determine the unknown $\tilde{c}_{\rm s}$. For $H_2O$, it requires $4.1855\,\mathrm{J}$ per gram of energy to raise the temperature from $T = 14.5^\circ\mathrm{C}$ to $T' = 15.5^\circ\mathrm{C}$, which is to say $\tilde{c}_{\rm H_2O}(T = 15^\circ\mathrm{C}) = 4.1855\,\mathrm{J/g\,^\circ C}$.[10] In practice, the fluid is contained in a vessel which has its own heat capacity and which must be accounted for in the determination of $\tilde{c}_{\rm s}$.

[9] In most metals, the difference between $C_V$ and $C_p$ is negligible.

[10] The specific heat $\tilde{c}_{\rm H_2O}(T)$ is weakly temperature dependent throughout much of the liquid phase. It turns out that $\tilde{c}_{\rm H_2O}(T = 20^\circ\mathrm{C}) = 4.1819\,\mathrm{J/g\,^\circ C}$. The conversion factor between Joules and *thermochemical calories* is $1\,\mathrm{J} = 0.2390\,\mathrm{cal}$. The so-called "15° calorie" is equivalent to 4.1855 Joules.

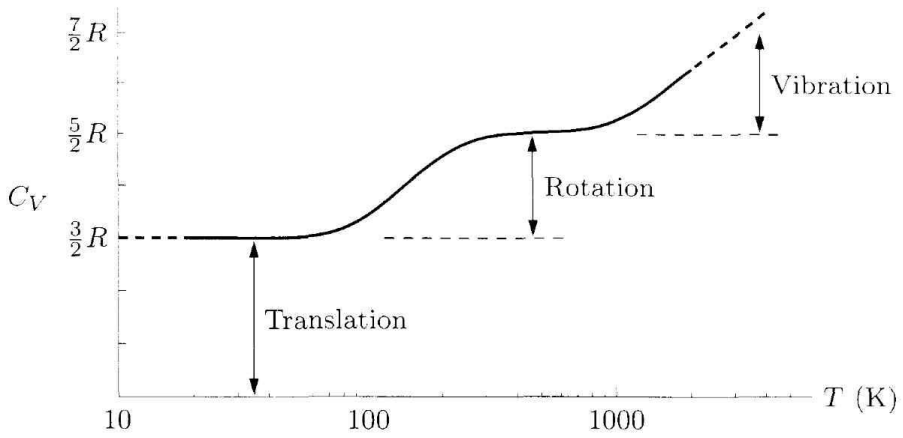



Figure 2.8: Heat capacity $C_V$ for one mole of hydrogen ($H_2$) gas. At the lowest temperatures, only translational degrees of freedom are relevant, and $f = 3$. At around 200 K, two rotational modes are excitable and $f = 5$. Above 1000 K, the vibrational excitations begin to contribute. Note the logarithmic temperature scale. (Data from H. W. Wooley *et al.*, *Jour. Natl. Bureau of Standards*, **41**, 379 (1948).)

### 2.4.4 Ideal gases

The ideal gas equation of state is $pV = Nk_{\rm B}T$. In order to invoke the formulae in eqns. 2.25, 2.28, and 2.30, we need to know the state function $E(T, V, N)$. A landmark experiment by Joule in the mid-19th century established that the energy of a low density gas is independent of its volume[11]. Essentially, a gas at temperature $T$ was allowed to freely expand from one volume $V$ to a larger volume $V' > V$, with no added heat $Q$ and no work $W$ done. Therefore the energy cannot change. What Joule found was that the temperature also did not change. This means that $E(T, V, N) = E(T, N)$ cannot be a function of the volume.

Since $E$ is extensive, we conclude that

$$E(T, V, N) = \nu\,\varepsilon(T) \quad , \tag{2.38}$$

where $\nu = N/N_{\rm A}$ is the number of moles of substance. Note that $\nu$ is an extensive variable. From eqns. 2.31 and 2.32, we conclude

$$C_V(T) = \nu\,\varepsilon'(T) \qquad , \qquad C_p(T) = C_V(T) + \nu R \quad , \tag{2.39}$$

where we invoke the ideal gas law to obtain the second of these. Empirically it is found that $C_V(T)$ is temperature independent over a wide range of $T$, far enough from boiling point. We can then write $C_V = \nu\, c_V$, where $\nu \equiv N/N_{\rm A}$ is the number of moles, and where $c_V$ is the molar heat capacity. We then have $c_p = c_V + R$, where $R = N_{\rm A}k_{\rm B} = 8.314462618\,{\rm J/mol\,K}$ is the *gas constant*. We denote by $\gamma = c_p/c_V$ the ratio of specific heat at constant pressure and at constant volume.

[11] See the description in E. Fermi, *Thermodynamics*, pp. 22-23.

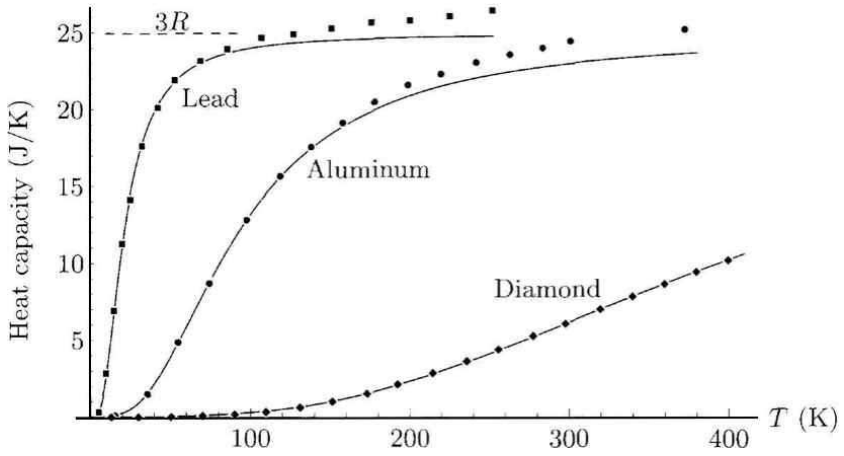



Figure 2.9: Molar heat capacities $c_V$ for three solids. The solid curves correspond to the predictions of the Debye model, which we shall discuss later.

From the kinetic theory of gases, one can show that

$$\begin{aligned} \text{monatomic gases:} \quad & c_V = \tfrac{3}{2}R \quad , \quad c_p = \tfrac{5}{2}R \quad , \quad \gamma = \tfrac{5}{3} \\ \text{diatomic gases:} \quad & c_V = \tfrac{5}{2}R \quad , \quad c_p = \tfrac{7}{2}R \quad , \quad \gamma = \tfrac{7}{5} \\ \text{polyatomic gases:} \quad & c_V = 3R \quad , \quad c_p = 4R \quad , \quad \gamma = \tfrac{4}{3} \quad . \end{aligned}$$

In the polyatomic case we assume the molecule is not linear; linear molecules, even if polyatomic, have only two rotational modes for the same reason that diatomic molecules do, *i.e.* the rotational inertial about their axis is so small that the associated quantized rotational energy levels are typically well above thermal energy scales. For linear polyatomic molecules, $f = 5$.

**Digression : kinetic theory of gases**

We will conclude in general from noninteracting classical statistical mechanics that the specific heat of a substance is $c_v = \frac{1}{2}fR$, where $f$ is the number of phase space coordinates, per particle, for which there is a quadratic kinetic or potential energy function. For example, a point particle has three translational degrees of freedom, and the kinetic energy is a quadratic function of their conjugate momenta: $H_0 = (p_x^2 + p_y^2 + p_z^2)/2m$. Thus, $f = 3$. Diatomic molecules have two additional rotational degrees of freedom – we don't count rotations about the symmetry axis – and their conjugate momenta also appear quadratically in the kinetic energy, leading to $f = 5$. For polyatomic molecules, all three Euler angles and their conjugate momenta are in play, and $f = 6$.

The reason that $f = 5$ for diatomic molecules rather than $f = 6$ is due to quantum mechanics. While translational eigenstates form a continuum, or are quantized in a box with $\Delta k_\alpha = 2\pi/L_\alpha$ being very small, since the dimensions $L_\alpha$ are macroscopic, angular momentum, and hence rotational kinetic energy, is quantized. For rotations about a principal axis with very low moment of inertia $I$, the corresponding energy scale $\hbar^2/2I$ is very large, and a high temperature is required in order to thermally populate these states. Thus, degrees of freedom with a quantization energy on the order of or greater than $\varepsilon_0$ are 'frozen out' for temperatures $T \lesssim \varepsilon_0/k_{\rm B}$.

In solids, each atom is effectively connected to its neighbors by springs; such a potential arises from quantum mechanical and electrostatic consideration of the interacting atoms. Thus, each degree of freedom contributes to the potential energy, and its conjugate momentum contributes to the kinetic energy. This results in $f = 6$. Assuming only lattice vibrations, then, the high temperature limit for $c_V(T)$ for any solid is predicted to be $3R = 24.944\,\mathrm{J/mol\,K}$. This is called the *Dulong-Petit law*. The high temperature limit is reached above the so-called *Debye temperature*, which is roughly proportional to the melting temperature of the solid.

In table 2.1, we list $c_p$ and $\tilde{c}_p$ for some common substances at $T = 25^\circ\,\mathrm{C}$ (unless otherwise noted). Note that $c_p$ for the monatomic gases He and Ne is to high accuracy given by the value from kinetic theory, $c_p = \frac{5}{2}R = 20.7864\,\mathrm{J/mol\,K}$. For the diatomic gases oxygen ($O_2$) and air (mostly $N_2$ and $O_2$), kinetic theory predicts $c_p = \frac{7}{2}R = 29.10$, which is close to the measured values. Kinetic theory predicts $c_p = 4R = 33.258$ for polyatomic gases; the measured values for $CO_2$ and $H_2O$ are both about 10% higher.

### 2.4.5 Adiabatic transformations of ideal gases

Assuming $dN = 0$ and $E = \nu\,\varepsilon(T)$, eqn. 2.25 tells us that

$$đQ = C_V\,dT + p\,dV \quad . \tag{2.40}$$

Invoking the ideal gas law to write $p = \nu RT/V$, and remembering $C_V = \nu\,c_V$, we have, setting $đQ = 0$,

$$\frac{dT}{T} + \frac{R}{c_V}\frac{dV}{V} = 0 \quad . \tag{2.41}$$

We can immediately integrate to obtain

$$đQ = 0 \quad \Longrightarrow \quad \begin{cases} TV^{\gamma-1} = \text{constant} \\ pV^{\gamma} = \text{constant} \\ T^{\gamma}p^{1-\gamma} = \text{constant} \end{cases} \tag{2.42}$$

where the second two equations are obtained from the first by invoking the ideal gas law. These are all *adiabatic equations of state*. Note the difference between the adiabatic equation of state $d(pV^\gamma) = 0$ and the isothermal equation of state $d(pV) = 0$. Equivalently, we can write these three conditions as

$$V^2\,T^f = V_0^2\,T_0^f \qquad , \qquad p^f\,V^{f+2} = p_0^f\,V_0^{f+2} \qquad , \qquad T^{f+2}\,p^{-2} = T_0^{f+2}\,p_0^{-2} \quad . \tag{2.43}$$

It turns out that air is a rather poor conductor of heat. This suggests the following model for an *adiabatic atmosphere*. The hydrostatic pressure decrease associated with an increase $dz$ in height is $dp = -\varrho g\,dz$, where $\varrho$ is the density and $g$ the acceleration due to gravity. Assuming the gas is ideal, the density can be written as $\varrho = Mp/RT$, where $M$ is the molar mass. Thus,

$$\frac{dp}{p} = -\frac{Mg}{RT}\,dz \quad . \tag{2.44}$$

If the atmosphere is isothermal, then $T$ is constant and $p(z) = p(0)\exp(-z/\lambda)$, where $\lambda = RT/Mg$. This also entails an exponential profile of the density: $\varrho(z) = \varrho(0)\,\exp(-z/\lambda)$. With $M = 28.88\,\mathrm{g}$ for air and $T = 293\,\mathrm{K}$, we have $\lambda = 8.6\,\mathrm{km}$.

If, on the other hand, the atmosphere is adiabatic, then from $d(T^\gamma p^{1-\gamma}) = 0$, we have

$$dT = \frac{\gamma-1}{\gamma}\,\frac{Tdp}{p} = -\frac{\gamma-1}{\gamma}\,\frac{Mg}{R}\,dz \quad , \tag{2.45}$$

with the solution

$$T(z) = T_0 - \frac{\gamma-1}{\gamma}\,\frac{Mg}{R}\,z = \left(1 - \frac{\gamma-1}{\gamma}\,\frac{z}{\lambda}\right)T_0 \quad , \tag{2.46}$$

where $T_0 = T(0) = 293\,\mathrm{K}$ is the temperature at the earth's surface, and $\lambda = RT_0/Mg$ as above. Air is mostly diatomic, and taking $\gamma = \frac{7}{5}$ we obtain $dT/dz = -(1-\gamma^{-1})\,T_0/\lambda = -9.7\,\mathrm{K/km}$. Note that in this model the atmosphere ends at a height $z_{\max} = \gamma\lambda/(\gamma-1) = 30\,\mathrm{km}$.

Again invoking the adiabatic equation of state, we can find $p(z)$:

$$\frac{p(z)}{p_0} = \left(\frac{T}{T_0}\right)^{\frac{\gamma}{\gamma-1}} = \left(1 - \frac{\gamma-1}{\gamma}\,\frac{z}{\lambda}\right)^{\frac{\gamma}{\gamma-1}} \tag{2.47}$$

Since $\varrho \propto p/T$ from the ideal gas law, we have

$$\frac{\varrho(z)}{\varrho_0} = \left(1 - \frac{\gamma-1}{\gamma}\,\frac{z}{\lambda}\right)^{\frac{1}{\gamma-1}} \quad . \tag{2.48}$$

Finally, recall the mathematical result $e^x = \lim_{k\to\infty}\left(1+\frac{x}{k}\right)^k$. Thus, in the limit $\gamma \to 1$, where $k = \gamma/(\gamma-1) \to \infty$, we recover the isothermal result $p(z) = p_0\,\exp(-z/\lambda)$. This makes sense because $d(T^\gamma p^{1-\gamma}) = dT = 0$ for $\gamma = 1$.

### 2.4.6 Adiabatic free expansion

Consider the situation depicted in fig. 2.10. A quantity ($\nu$ moles) of gas in equilibrium at temperature $T$ and volume $V_1$ is allowed to expand freely into an evacuated chamber of volume $V_2$ by the removal of a barrier. Clearly no work is done on or by the gas during this process, hence $W = 0$. If the walls are everywhere insulating, so that no heat can pass through them, then $Q = 0$ as well. The First Law then gives $\Delta E = Q - W = 0$, and there is no change in energy.

If the gas is ideal, then since $E(T,V,N) = \nu c_V T$ with $\nu = N/N_{\rm A}$, $\Delta E = 0$ gives $\Delta T = 0$, and there is no change in temperature. (If the walls are insulating against the passage of heat, they must also prevent the passage of particles, so $\Delta N = 0$.) There is of course a change in volume: $\Delta V = V_2$, hence there is a change in pressure. The initial pressure is $p = Nk_{\rm B}T/V_1$ and the final pressure is $p' = Nk_{\rm B}T/(V_1+V_2)$.

If the gas is nonideal, then the temperature will in general change. Suppose $E(T,V,N) = \alpha\,V^x\,N^{1-x}\,T^y$, where $\alpha$, $x$, and $y$ are constants. This form is properly extensive: if $V$ and $N$ double, then $E$ doubles. If the volume changes from $V$ to $V'$ under an adiabatic free expansion, then we must have, from $\Delta E = 0$,

$$\left(\frac{V}{V'}\right)^x = \left(\frac{T'}{T}\right)^y \qquad \Longrightarrow \qquad T' = T\cdot\left(\frac{V}{V'}\right)^{x/y} \quad . \tag{2.49}$$

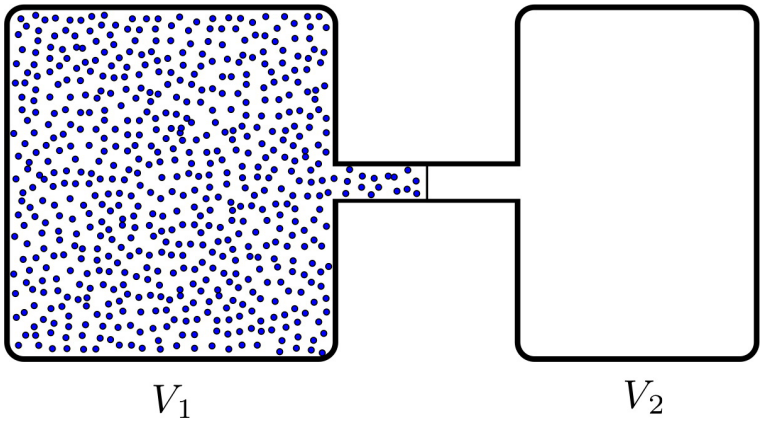



Figure 2.10: In the adiabatic free expansion of a gas, there is volume expansion with no work or heat exchange with the environment: $\Delta E = Q = W = 0$.

If $x/y > 0$, the temperature decreases upon the expansion. If $x/y < 0$, the temperature increases. Without an equation of state, we can't say precisely what happens to the pressure. Does the pressure decrease? Not necessarily! While thermodynamic stability entails a positive isothermal compressibility $\kappa_T = -V^{-1}(\partial V/\partial p)_{T,N} > 0$, thermodynamics does not impose a universal sign on $(\partial p/\partial V)_{E,N}$.

Adiabatic free expansion of a gas is a *spontaneous process*, arising due to the natural internal dynamics of the system. It is also *irreversible*. If we wish to take the gas back to its original state, we must do work on it to compress it. If the gas is ideal, then the initial and final temperatures are identical, so we can place the system in thermal contact with a reservoir at temperature $T$ and follow a thermodynamic path along an isotherm. The work done *on the gas* during compression is then

$$\mathcal{W} = -Nk_{\rm B}T \int_{V_{\rm f}}^{V_{\rm i}} \frac{dV}{V} = Nk_{\rm B}T \, \log\left(\frac{V_{\rm f}}{V_{\rm i}}\right) = Nk_{\rm B}T \, \log\left(1 + \frac{V_2}{V_1}\right) \quad . \tag{2.50}$$

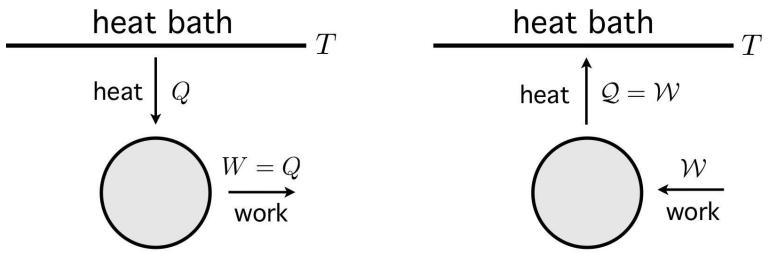
The work done *by the gas* is $W = \int p\, dV = -\mathcal{W}$. During the compression, heat energy $Q = W < 0$ is transferred to the gas from the reservoir. Thus, $\mathcal{Q} = \mathcal{W} > 0$ is given off by the gas to its environment.

## 2.5 Heat Engines and the Second Law of Thermodynamics

### 2.5.1 There's no free lunch so quit asking

A *heat engine* is a device which takes a thermodynamic system through a repeated cycle which can be represented as a succession of equilibrium states: $\mathrm{A} \to \mathrm{B} \to \mathrm{C} \cdots \to \mathrm{A}$. The net result of such a cyclic process is to convert heat into mechanical work, or *vice versa*.

For a system in equilibrium at temperature $T$, there is a thermodynamically large amount of internal energy stored in the random internal motion of its constituent particles. Later, when we study statistical mechanics, we will see how each 'quadratic' degree of freedom in the Hamiltonian contributes $\frac{1}{2}k_{\rm B}T$




(a) a perfect engine (b) a perfect waste of time

Figure 2.11: A *perfect engine* would extract heat $Q$ from a thermal reservoir at some temperature $T$ and convert it into useful mechanical work $W$. This process is alas impossible, according to the Second Law of thermodynamics. The inverse process, where work $\mathcal{W}$ is converted into heat $\mathcal{Q}$, is always possible.

to the total internal energy. An immense body in equilibrium at temperature $T$ has an enormous heat capacity $C$, so extracting a finite quantity of heat $Q$ from it results in a temperature change $\Delta T = -Q/C$ which is utterly negligible. Such a body is called a *heat bath*, or *thermal reservoir*. A *perfect engine* would, in each cycle, extract an amount of heat $Q$ from the bath and convert it into work. Since $\Delta E = 0$ for a cyclic process, the First Law then gives $W = Q$. This situation is depicted schematically in fig. 2.11. One could imagine running this process virtually indefinitely, slowly sucking energy out of an immense heat bath, converting the random thermal motion of its constituent molecules into useful mechanical work. Sadly, this is not possible:

> *A transformation whose only final result is to extract heat from a source at fixed temperature and transform that heat into work is impossible.*

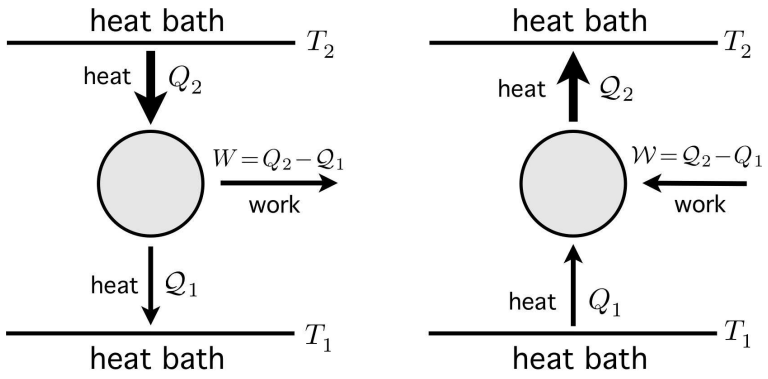
This is known as the *Postulate of Lord Kelvin*. It is equivalent to the *postulate of Clausius*,

> *A transformation whose only result is to transfer heat from a body at a given temperature to a body at higher temperature is impossible.*

These postulates which have been repeatedly validated by empirical observations, constitute the Second Law of Thermodynamics.

### 2.5.2 Engines and refrigerators

While it is not possible to convert heat into work with 100% efficiency, it is possible to transfer heat from one thermal reservoir to another one, at lower temperature, and to convert some of that heat into work. This is what an engine does. The energy accounting for one cycle of the engine is depicted in the left hand panel of fig. 2.12. An amount of heat $Q_2 > 0$ is extracted from the reservoir at temperature $T_2$. Since the reservoir is assumed to be enormous, its temperature change $\Delta T_2 = -Q_2/C_2$ is negligible, and its temperature remains constant – this is what it means for an object to be a reservoir. A lesser amount of heat, $\mathcal{Q}_1$, with $0 < \mathcal{Q}_1 < Q_2$, is deposited in a second reservoir at a lower temperature $T_1$. Its temperature change $\Delta T_1 = +\mathcal{Q}_1/C_1$ is also negligible. The difference $W = Q_2 - \mathcal{Q}_1$ is extracted as




(a) engine (b) refrigerator

Figure 2.12: An engine (left) extracts heat $Q_2$ from a reservoir at temperature $T_2$ and deposits a smaller amount of heat $\mathcal{Q}_1$ into a reservoir at a lower temperature $T_1$, during each cycle. The difference $W = Q_2 - \mathcal{Q}_1$ is transformed into mechanical work. A refrigerator (right) performs the inverse process, drawing heat $Q_1$ from a low temperature reservoir and depositing heat $\mathcal{Q}_2 = Q_1 + \mathcal{W}$ into a high temperature reservoir, where $\mathcal{W}$ is the mechanical (or electrical) work done per cycle.

useful work. We define the *efficiency*, $\eta$, of the engine as the ratio of the work done to the heat extracted from the upper reservoir, per cycle:

$$\eta = \frac{W}{Q_2} = 1 - \frac{\mathcal{Q}_1}{Q_2} \quad . \tag{2.51}$$

This is a natural definition of efficiency, since it will cost us fuel to maintain the temperature of the upper reservoir over many cycles of the engine. Thus, the efficiency is proportional to the ratio of the work done to the cost of the fuel.

A refrigerator works according to the same principles, but the process runs in reverse. An amount of heat $Q_1$ is extracted from the lower reservoir – the inside of our refrigerator – and is pumped into the upper reservoir. As Clausius' form of the Second Law asserts, it is impossible for this to be the only result of our cycle. Some amount of work $\mathcal{W}$ must be performed on the refrigerator in order for it to extract the heat $Q_1$. Since $\Delta E = 0$ for the cycle, a heat $\mathcal{Q}_2 = \mathcal{W} + Q_1$ must be deposited into the upper reservoir during each cycle. The analog of efficiency here is called the *coefficient of refrigeration*, $\kappa$, defined as

$$\kappa = \frac{Q_1}{\mathcal{W}} = \frac{Q_1}{\mathcal{Q}_2 - Q_1} \quad . \tag{2.52}$$

Thus, $\kappa$ is proportional to the ratio of the heat extracted to the cost of electricity, per cycle.

Please note the deliberate notation here. I am using symbols $Q$ and $W$ to denote the heat supplied *to* the engine (or refrigerator) and the work done *by* the engine, respectively, and $\mathcal{Q}$ and $\mathcal{W}$ to denote the heat taken *from* the engine and the work done *on* the engine.

A perfect engine has $\mathcal{Q}_1 = 0$ and $\eta = 1$; a perfect refrigerator has $Q_1 = \mathcal{Q}_2$ and $\kappa = \infty$. Both violate the Second Law. Sadi Carnot[12] (1796 – 1832) realized that a *reversible* cyclic engine operating between

[12]Carnot died during the cholera epidemic of 1832. His is one of the 72 names engraved on the Eiffel Tower.

two thermal reservoirs must produce the maximum amount of work $W$, and that the amount of work produced is *independent of the material properties of the engine*. We call any such engine a *Carnot engine*.

The efficiency of a Carnot engine may be used to define a temperature scale. We know from Carnot's observations that the efficiency $\eta_{\rm C}$ can only be a function of the temperatures $T_1$ and $T_2$: $\eta_{\rm C} = \eta_{\rm C}(T_1, T_2)$. We can then define

$$\frac{T_1}{T_2} \equiv 1 - \eta_{\rm C}(T_1, T_2) \quad . \tag{2.53}$$

Below, in §2.5.4, we will see how, using an ideal gas as the 'working substance' of the Carnot engine, this temperature scale coincides precisely with the ideal gas temperature scale from §2.2.6.

### 2.5.3 Nothing beats a Carnot engine

The Carnot engine is the most efficient engine possible operating between two thermal reservoirs. To see this, let's suppose that an amazing wonder engine has an efficiency even greater than that of the Carnot engine. A key feature of the Carnot engine is its reversibility – we can just go around its cycle in the opposite direction, creating a Carnot refrigerator. Let's use our notional wonder engine to drive a Carnot refrigerator, as depicted in fig. 2.13.

We assume that

$$\frac{W}{Q_2} = \eta_{\rm wonder} > \eta_{\rm Carnot} = \frac{\mathcal{W}'}{\mathcal{Q}_2'} \quad . \tag{2.54}$$

But from the figure, we have $W = \mathcal{W}'$, and therefore the heat energy $\mathcal{Q}_2' - Q_2$ transferred *to* the upper reservoir is positive. From

$$W = Q_2 - \mathcal{Q}_1 = \mathcal{Q}_2' - Q_1' = \mathcal{W}' \quad , \tag{2.55}$$

we see that this is equal to the heat energy extracted *from* the lower reservoir, since no external work is done on the system:

$$\mathcal{Q}_2' - Q_2 = Q_1' - \mathcal{Q}_1 > 0 \quad . \tag{2.56}$$

Therefore, the existence of the wonder engine entails a violation of the Second Law. Since the Second Law is correct – *Lord* Kelvin articulated it, and who are we to argue with a *Lord*? – the wonder engine cannot exist.

We further conclude that *all reversible engines running between two thermal reservoirs have the same efficiency, which is the efficiency of a Carnot engine.* In general, we have

$$\eta = \frac{W}{Q_2} = 1 - \frac{\mathcal{Q}_1}{Q_2} \le 1 - \frac{T_1}{T_2} = \eta_{\rm C} \quad , \tag{2.57}$$

and therefore

$$\frac{Q_2}{T_2} - \frac{\mathcal{Q}_1}{T_1} \le 0 \quad , \tag{2.58}$$

with the equalities holding only for reversible engines.

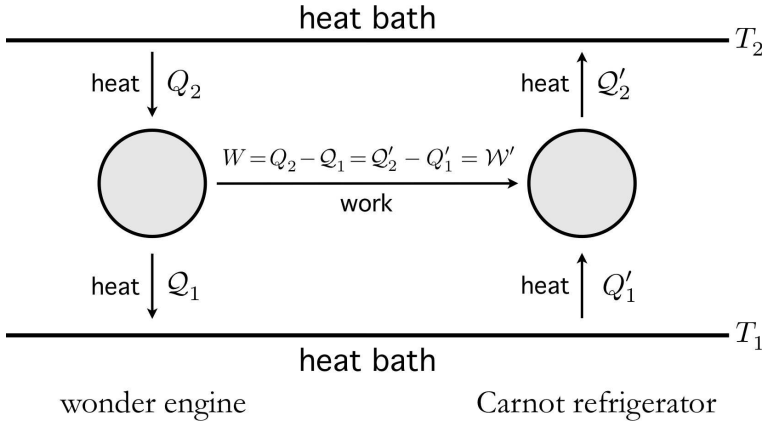



Figure 2.13: A wonder engine driving a Carnot refrigerator.

### 2.5.4 The Carnot cycle

Let us now consider a specific cycle, known as the Carnot cycle, depicted in fig. 2.14. The cycle consists of two adiabats and two isotherms. The work done per cycle is simply the area inside the curve on our $p - V$ diagram:

$$W = \oint p\, dV \quad . \tag{2.59}$$

The gas inside our Carnot engine is called the 'working substance'. Whatever it may be, the system obeys the First Law,

$$dE = đQ - đW = đQ - p\, dV \quad . \tag{2.60}$$

We will now assume that the working material is an ideal gas, and we compute $W$ as well as $\mathcal{Q}_1$ and $Q_2$ to find the efficiency of this cycle. In order to do this, we will rely upon the ideal gas equations,

$$E = \frac{\nu RT}{\gamma - 1} \qquad , \qquad pV = \nu RT \quad , \tag{2.61}$$

where $\gamma = c_p/c_v = 1 + \frac{2}{f}$, where $f$ is the effective number of molecular degrees of freedom contributing to the internal energy. Recall $f = 3$ for monatomic gases, $f = 5$ for diatomic gases, and $f = 6$ for polyatomic gases. The finite difference form of the first law is

$$\Delta E = E_{\rm f} - E_{\rm i} = Q_{\rm if} - W_{\rm if} \quad , \tag{2.62}$$

where i denotes the initial state and f the final state.

AB: This stage is an isothermal expansion at temperature $T_2$. It is the 'power stroke' of the engine. We have

$$W_{\rm AB} = \int_{V_{\rm A}}^{V_{\rm B}} dV\, \frac{\nu RT_2}{V} = \nu RT_2\, \log\left(\frac{V_{\rm B}}{V_{\rm A}}\right) \tag{2.63}$$

$$E_{\rm A} = E_{\rm B} = \frac{\nu RT_2}{\gamma - 1} \quad , \tag{2.64}$$

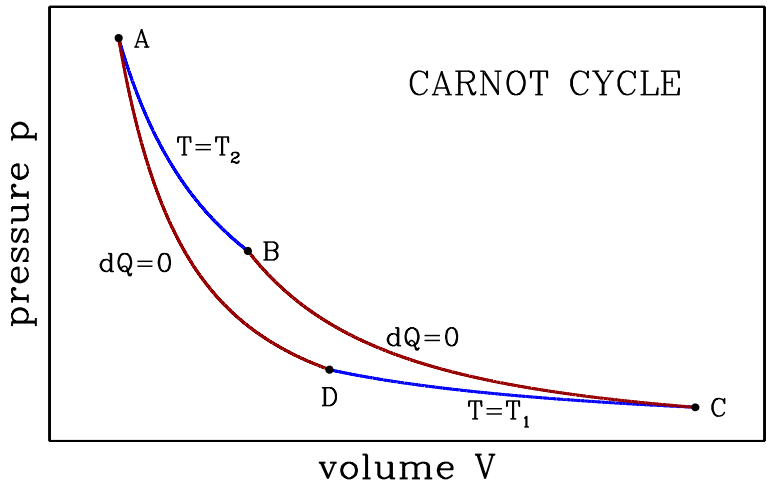



Figure 2.14: The Carnot cycle consists of two adiabats (dark red) and two isotherms (blue).

hence

$$Q_{\rm AB} = \Delta E_{\rm AB} + W_{\rm AB} = \nu R T_2 \, \log\left(\frac{V_{\rm B}}{V_{\rm A}}\right) \quad . \tag{2.65}$$

BC: This stage is an adiabatic expansion. We have

$$Q_{\rm BC} = 0 \tag{2.66}$$

$$\Delta E_{\rm BC} = E_{\rm C} - E_{\rm B} = \frac{\nu R}{\gamma - 1}\,(T_1 - T_2) \quad . \tag{2.67}$$

The energy change is negative, and the heat exchange is zero, so the engine still does some work during this stage:

$$W_{\rm BC} = Q_{\rm BC} - \Delta E_{\rm BC} = \frac{\nu R}{\gamma - 1}\,(T_2 - T_1) \quad . \tag{2.68}$$

CD: This stage is an isothermal compression, and we may apply the analysis of the isothermal expansion, *mutatis mutandis*:

$$W_{\rm CD} = \int_{V_{\rm C}}^{V_{\rm D}} dV \, \frac{\nu R T_1}{V} = \nu R T_1 \, \log\left(\frac{V_{\rm D}}{V_{\rm C}}\right) \tag{2.69}$$

$$E_{\rm C} = E_{\rm D} = \frac{\nu R T_1}{\gamma - 1} \quad , \tag{2.70}$$

hence

$$Q_{\rm CD} = \Delta E_{\rm CD} + W_{\rm CD} = \nu R T_1 \, \log\left(\frac{V_{\rm D}}{V_{\rm C}}\right) \quad . \tag{2.71}$$

DA: This last stage is an adiabatic compression, and we may draw on the results from the adiabatic

expansion in BC:

$$Q_{\rm DA} = 0 \tag{2.72}$$

$$\Delta E_{\rm DA} = E_{\rm A} - E_{\rm D} = \frac{\nu R}{\gamma - 1}\,(T_2 - T_1) \quad . \tag{2.73}$$

The energy change is positive, and the heat exchange is zero, so work is done on the engine:

$$W_{\rm DA} = Q_{\rm DA} - \Delta E_{\rm DA} = \frac{\nu R}{\gamma - 1}\,(T_1 - T_2) \quad . \tag{2.74}$$

We now add up all the work values from the individual stages to get for the cycle

$$\begin{aligned} W &= W_{\rm AB} + W_{\rm BC} + W_{\rm CD} + W_{\rm DA} \\ &= \nu R T_2\,\log\left(\frac{V_{\rm B}}{V_{\rm A}}\right) + \nu R T_1\,\log\left(\frac{V_{\rm D}}{V_{\rm C}}\right) \quad . \end{aligned} \tag{2.75}$$

Since we are analyzing a cyclic process, we must have $\Delta E = 0$, hence $Q = W$, which can of course be verified explicitly, by computing $Q = Q_{\rm AB} + Q_{\rm BC} + Q_{\rm CD} + Q_{\rm DA}$. To finish up, recall the adiabatic ideal gas equation of state, $d(TV^{\gamma-1}) = 0$. This tells us that

$$T_2\,V_{\rm B}^{\gamma-1} = T_1\,V_{\rm C}^{\gamma-1} \tag{2.76}$$

$$T_2\,V_{\rm A}^{\gamma-1} = T_1\,V_{\rm D}^{\gamma-1} \quad . \tag{2.77}$$

Dividing these two equations, we find

$$\frac{V_{\rm B}}{V_{\rm A}} = \frac{V_{\rm C}}{V_{\rm D}} \quad , \tag{2.78}$$

and therefore

$$W = \nu R(T_2 - T_1)\,\log\left(\frac{V_{\rm B}}{V_{\rm A}}\right) \tag{2.79}$$

$$Q_{\rm AB} = \nu R T_2\,\log\left(\frac{V_{\rm B}}{V_{\rm A}}\right) \quad . \tag{2.80}$$

Finally, the efficiency is given by the ratio of these two quantities:

$$\eta = \frac{W}{Q_{\rm AB}} = 1 - \frac{T_1}{T_2} \quad . \tag{2.81}$$

Examples of other useful thermodynamic cycles are discussed in the appendix, §2.15.

### 2.5.5 Carnot engine at maximum power output

While the Carnot engine described above in §2.5.4 has maximum efficiency, it is practically useless, because the isothermal processes must take place infinitely slowly in order for the working material to

remain in thermal equilibrium with each reservoir. Thus, while the work done per cycle is finite, the cycle period is infinite, and the engine *power* is zero.

A modification of the ideal Carnot cycle is necessary to create a practical engine. The idea[13] is as follows. During the isothermal expansion stage, the working material is maintained at a temperature $T_{2\rm w} < T_2$. The temperature difference between the working material and the hot reservoir drives a thermal current,

$$\frac{đQ_2}{dt} = \kappa_2 \, (T_2 - T_{2\rm w}) \quad . \tag{2.82}$$

Here, $\kappa_2$ is a *transport coefficient* which describes the *thermal conductivity* of the chamber walls, multiplied by a geometric parameter (which is the ratio of the total wall area to its thickness). Similarly, during the isothermal compression, the working material is maintained at a temperature $T_{1\rm w} > T_1$, which drives a thermal current to the cold reservoir,

$$\frac{đ\mathcal{Q}_1}{dt} = \kappa_1 \, (T_{1\rm w} - T_1) \quad . \tag{2.83}$$

Now let us assume that the upper isothermal stage requires a duration $\Delta t_2$ and the lower isotherm a duration $\Delta t_1$. Then

$$Q_2 = \kappa_2 \, \Delta t_2 \, (T_2 - T_{2\rm w}) \tag{2.84}$$

$$\mathcal{Q}_1 = \kappa_1 \, \Delta t_1 \, (T_{1\rm w} - T_1) \quad . \tag{2.85}$$

Assuming that the working fluid undergoes an internally reversible cycle, even though heat transfer between the working fluid and the reservoirs is irreversible, we have

$$\frac{\mathcal{Q}_1}{T_{1\rm w}} = \frac{Q_2}{T_{2\rm w}} \quad , \tag{2.86}$$

which says

$$\frac{\Delta t_1}{\Delta t_2} = \frac{\kappa_2 \, T_{1\rm w} \, (T_2 - T_{2\rm w})}{\kappa_1 \, T_{2\rm w} \, (T_{1\rm w} - T_1)} \quad . \tag{2.87}$$

The power is

$$P = \frac{Q_2 - \mathcal{Q}_1}{(1+\alpha) \, (\Delta t_1 + \Delta t_2)} \quad , \tag{2.88}$$

where we assume that the adiabatic stages require a combined time of $\alpha \, (\Delta t_1 + \Delta t_2)$. Thus, we find

$$P = \frac{\kappa_1 \, \kappa_2}{1+\alpha} \cdot \frac{(T_{2\rm w} - T_{1\rm w}) \, (T_{1\rm w} - T_1) \, (T_2 - T_{2\rm w})}{\kappa_1 \, T_{2\rm w} \, (T_{1\rm w} - T_1) + \kappa_2 \, T_{1\rm w} \, (T_2 - T_{2\rm w})} \quad . \tag{2.89}$$

We optimize the engine by maximizing $P$ with respect to the temperatures $T_{1\rm w}$ and $T_{2\rm w}$. This yields

$$T_{2\rm w} = T_2 - \frac{T_2 - \sqrt{T_1 T_2}}{1 + \sqrt{\kappa_2/\kappa_1}} \tag{2.90}$$

$$T_{1\rm w} = T_1 + \frac{\sqrt{T_1 T_2} - T_1}{1 + \sqrt{\kappa_1/\kappa_2}} \quad . \tag{2.91}$$

[13] See F. L. Curzon and B. Ahlborn, *Am. J. Phys.* **43**, 22 (1975). I am grateful to Professor Asle Sudbø for correcting a typo in one expression and providing a simplified form of another.

| Power source | $T_1$ (°C) | $T_2$ (°C) | $\eta_{\rm Carnot}$ | $\eta$ (theor.) | $\eta$ (obs.) |
|---|---|---|---|---|---|
| West Thurrock (UK) Coal Fired Steam Plant | $\sim 25$ | 565 | 0.641 | 0.40 | 0.36 |
| CANDU (Canada) PHW Nuclear Reactor | $\sim 25$ | 300 | 0.480 | 0.28 | 0.30 |
| Larderello (Italy) Geothermal Steam Plant | $\sim 80$ | 250 | 0.323 | 0.175 | 0.16 |

Table 2.2: Observed performances of real heat engines, taken from table 1 from Curzon and Ahlborn (1975).

The efficiency at maximum power is then

$$\eta = \frac{Q_2 - \mathcal{Q}_1}{Q_2} = 1 - \frac{T_{1\rm w}}{T_{2\rm w}} = 1 - \sqrt{\frac{T_1}{T_2}} \quad . \tag{2.92}$$

One also finds at maximum power

$$\frac{\Delta t_2}{\Delta t_1} = \sqrt{\frac{\kappa_1}{\kappa_2}} \quad . \tag{2.93}$$

Finally, the maximized power is

$$P_{\rm max} = \frac{\kappa_1 \kappa_2}{1+\alpha} \left( \frac{\sqrt{T_2} - \sqrt{T_1}}{\sqrt{\kappa_1} + \sqrt{\kappa_2}} \right)^2 \quad . \tag{2.94}$$

Table 2.2, taken from the article of Curzon and Ahlborn (1975), shows how the efficiency of this practical Carnot cycle, given by eqn. 2.92, rather accurately predicts the efficiencies of functioning power plants.

## 2.6 The Entropy

### 2.6.1 Entropy and heat

The Second Law guarantees us that an engine operating between two heat baths at temperatures $T_1$ and $T_2$ must satisfy

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} \le 0 \quad , \tag{2.95}$$

with the equality holding for reversible processes. This is a restatement of eqn. 2.58, after writing $Q_1 = -\mathcal{Q}_1$ for the heat transferred *to* the engine from reservoir #1. Consider now an arbitrary curve in the $p - V$ plane. We can describe such a curve, to arbitrary accuracy, as a combination of Carnot cycles,

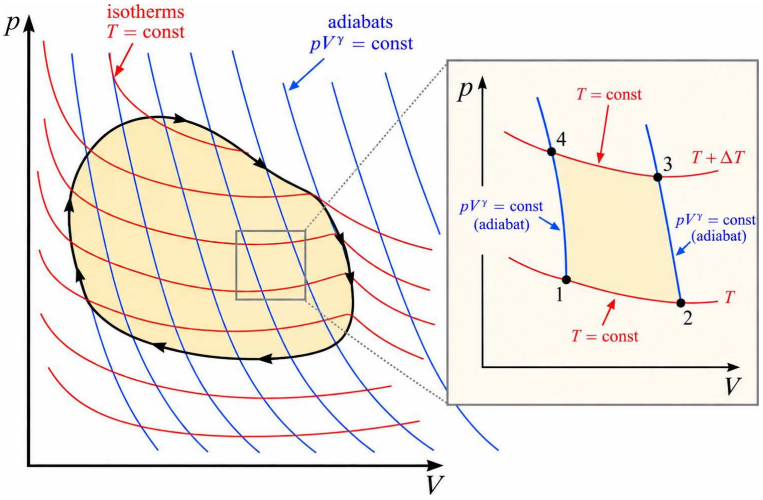



Figure 2.15: An arbitrarily shaped cycle in the $p-V$ plane can be decomposed into a number of smaller Carnot cycles. Red curves indicate isotherms and blue curves adiabats, with $\gamma = \frac{5}{3}$.

as shown in fig. 2.15. Each little Carnot cycle consists of two adiabats and two isotherms. We then conclude

$$\sum_i \frac{Q_i}{T_i} \longrightarrow \oint_{\mathcal{C}} \frac{đQ}{T} \leq 0 \quad , \tag{2.96}$$

with equality holding if all the cycles are reversible. Rudolf Clausius, in 1865, realized that one could then define a new state function, which he called the *entropy*, $S$, that depended only on the initial and final states of a reversible process:

$$dS = \frac{đQ}{T} \quad \Longrightarrow \quad S_{\rm B} - S_{\rm A} = \int_{\rm A}^{\rm B} \frac{đQ}{T} \quad . \tag{2.97}$$

Since $Q$ is extensive, so is $S$; the units of entropy are $[S] = \mathrm{J/K}$.

### 2.6.2 The third law of thermodynamics

Eqn. 2.97 determines the entropy up to a constant. By choosing a standard state $\Upsilon$, we can define $S_\Upsilon = 0$, and then by taking $\mathrm{A} = \Upsilon$ in the above equation, we can define the absolute entropy $S$ for any state. However, it turns out that this seemingly arbitrary constant $S_\Upsilon$ in the entropy does have consequences, for example in the theory of gaseous equilibrium. The proper definition of entropy, from the point of view of statistical mechanics, will lead us to understand how the zero temperature entropy of a system is related to its quantum mechanical ground state degeneracy. Walther Nernst, in 1906, articulated a principle which is sometimes called the Third Law of Thermodynamics,

*The entropy of every system at absolute zero temperature always vanishes.*

Again, this is not quite correct, and quantum mechanics tells us that $S(T=0) = k_{\rm B} \log g$, where $g$ is the ground state degeneracy. Nernst's law holds when $g = 1$. A more accurate version of the Third Law:

> *The entropy of a perfect crystal with a unique equilibrium ground state approaches zero as* $T \to 0$.

We can combine the First and Second laws to write

$$dE + đW = đQ \le T\,dS \quad , \tag{2.98}$$

where the equality holds for reversible processes.

### 2.6.3 Entropy changes in cyclic processes

For a cyclic process, whether reversible or not, the change in entropy around a cycle is zero: $\Delta S_{\rm CYC} = 0$. This is because the entropy $S$ is a state function, with a unique value for every equilibrium state. A cyclical process returns to the same equilibrium state, hence $S$ must return as well to its corresponding value from the previous cycle.

Consider now a general engine, as in fig. 2.12. Let us compute the total entropy change in the entire Universe over one cycle. We have

$$(\Delta S)_{\rm TOTAL} = (\Delta S)_{\rm ENGINE} + (\Delta S)_{\rm HOT} + (\Delta S)_{\rm COLD} \quad , \tag{2.99}$$

written as a sum over entropy changes of the engine itself, the hot reservoir, and the cold reservoir[14]. Clearly $(\Delta S)_{\rm ENGINE} = 0$. The changes in the reservoir entropies are

$$(\Delta S)_{\rm HOT} = \int\limits_{T=T_2} \frac{đQ_{\rm HOT}}{T} = -\frac{Q_2}{T_2} \ < \ 0 \tag{2.100}$$

$$(\Delta S)_{\rm COLD} = \int\limits_{T=T_1} \frac{đQ_{\rm COLD}}{T} = \frac{\mathcal{Q}_1}{T_1} = -\frac{Q_1}{T_1} \ > \ 0 \quad , \tag{2.101}$$

because the hot reservoir loses heat $Q_2 > 0$ to the engine, and the cold reservoir gains heat $\mathcal{Q}_1 = -Q_1 > 0$ from the engine. Therefore,

$$(\Delta S)_{\rm TOTAL} = -\left(\frac{Q_1}{T_1} + \frac{Q_2}{T_2}\right) \ge 0 \quad . \tag{2.102}$$

Thus, for a reversible cycle, the net change in the total entropy of the engine plus reservoirs is zero. For an irreversible cycle, there is an increase in total entropy, due to spontaneous processes.

[14] We neglect any interfacial contributions to the entropy change, which will be small compared with the bulk entropy change in the thermodynamic limit of large system size.

### 2.6.4 Gibbs-Duhem relation

Consider a function $f(x_1, \ldots, x_n)$ which is *homogeneous of degree* $k$, meaning

$$f(\lambda x_1, \ldots, \lambda x_n) = \lambda^k f(x_1, \ldots, x_n) \tag{2.103}$$

for all $\boldsymbol{x}$ and $\lambda$. Thus, $f(x, y) = x^2 + 2xy$ is homogeneous of degree $k = 2$, $f(x, y, z) = x^{-1} \exp(y/z) + \pi xyz^{-3}$ is homogeneous of degree $k = -1$, *etc.* Now differentiate the above equation with respect to $\lambda$ and then set $\lambda = 1$, in which case we obtain

$$\sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i} = kf \quad , \tag{2.104}$$

a result known as *Euler's theorem for homogeneous functions*. Try it out for yourself on the above examples!

Now recall the general expression eqn. 2.12 for differential work,

$$đW = -\sum_{j=1}^{\mathsf{d}} y_j \, dX_j - \sum_{a=1}^{\mathsf{c}} \mu_a \, dN_a \quad , \tag{2.105}$$

where the $(-y_j, X_j)$ are conjugate pairs of generalized forces and generalized displacements, and $(\mu_a, N_a)$ are conjugate pairs of chemical potentials and species numbers. Here each $y_j$ and $\mu_a$ is intensive, while each $X_j$ and $N_a$ is extensive. For reversible systems, we can write

$$dE = đQ - đW = T \, dS + \sum_{j=1}^{\mathsf{d}} y_j \, dX_j + \sum_{a=1}^{\mathsf{c}} \mu_a \, dN_a \quad , \tag{2.106}$$

which says that the energy $E$ is a function of the entropy $S$, the generalized displacements $\{X_j\}$, and the particle numbers $\{N_a\}$:

$$E = E\big(S, \{X_j\}, \{N_a\}\big) \quad . \tag{2.107}$$

Furthermore, we have

$$T = \left(\frac{\partial E}{\partial S}\right)_{\{X_j, N_a\}} \quad , \quad y_j = \left(\frac{\partial E}{\partial X_j}\right)_{S, \{X_{i(\neq j)}, N_a\}} \quad , \quad \mu_a = \left(\frac{\partial E}{\partial N_a}\right)_{S, \{X_j, N_{b(\neq a)}\}} \tag{2.108}$$

Since $E$ and all its arguments are extensive, we have

$$E\big(\lambda S, \{\lambda X_j\}, \{\lambda N_a\}\big) = \lambda E\big(S, \{X_j\}, \{N_a\}\big) \quad , \tag{2.109}$$

which says that the energy is a homogeneous function of all the extensive variables $S$, $\{X_j\}$, and $\{N_a\}$, of degree $k = 1$. We then conclude that

$$\begin{aligned} E &= S \frac{\partial E}{\partial S} + \sum_{j=1}^{\mathsf{d}} X_j \frac{\partial E}{\partial X_j} + \sum_{a=1}^{\mathsf{c}} N_a \frac{\partial E}{\partial N_a} \\ &= TS + \sum_{j=1}^{\mathsf{d}} y_j \, X_j + \sum_{a=1}^{\mathsf{c}} \mu_a \, N_a \quad . \end{aligned} \tag{2.110}$$

Taking the differential of eqn. 2.110, and then subtracting eqn. 2.106, we obtain

$$S\,dT + \sum_{j=1}^{\mathsf{d}} X_j\,dy_j + \sum_{a=1}^{\mathsf{c}} N_a\,d\mu_a = 0 \quad . \tag{2.111}$$

This is called the *Gibbs-Duhem relation*. It says that there is one equation of state which may be written in terms of all the intensive quantities alone. For example, for a single component $(p, V)$ system, we have $E = TS - pV + \mu N$ and $p = p(T, \mu)$, the latter following from

$$S\,dT - V dp + N\,d\mu = 0 \quad . \tag{2.112}$$

### 2.6.5 Entropy for an ideal gas

For an ideal gas, we have $E = \frac{1}{2} f N k_{\rm B} T$, and

$$\begin{aligned} dS &= \frac{1}{T}\,dE + \frac{p}{T}\,dV - \frac{\mu}{T}\,dN \\ &= \tfrac{1}{2} f N k_{\rm B}\,\frac{dT}{T} + \frac{p}{T}\,dV + \left(\tfrac{1}{2} f k_{\rm B} - \frac{\mu}{T}\right) dN \quad . \end{aligned} \tag{2.113}$$

Invoking the ideal gas equation of state $pV = N k_{\rm B} T$, we have

$$dS\big|_N = \tfrac{1}{2} f N k_{\rm B}\,d\log T + N k_{\rm B} d\log V \quad . \tag{2.114}$$

Integrating, we obtain

$$S(T, V, N) = \tfrac{1}{2} f N k_{\rm B}\,\log T + N k_{\rm B}\,\log V + \varphi(N) \quad , \tag{2.115}$$

where $\varphi(N)$ is an arbitrary function. Extensivity of $S$ places restrictions on $\varphi(N)$, so that the most general case is

$$\begin{aligned} S(T, V, N) &= \tfrac{1}{2} f N k_{\rm B}\,\log T + N k_{\rm B}\,\log\left(\frac{V}{N}\right) + Na \\ &= \tfrac{1}{2} f N k_{\rm B}\,\log\left(\frac{E}{N}\right) + N k_{\rm B}\,\log\left(\frac{V}{N}\right) + Nb \quad , \end{aligned} \tag{2.116}$$

where $a$ and $b$ are constants, where $b = a - \frac{1}{2} f k_{\rm B} \log(\frac{1}{2} f k_{\rm B})$. When we study statistical mechanics, we will find that for the monatomic ideal gas the entropy is

$$S(T, V, N) = N k_{\rm B}\left[\tfrac{5}{2} + \log\left(\frac{V}{N\lambda_T^3}\right)\right] \quad , \tag{2.117}$$

where $\lambda_T = \sqrt{2\pi\hbar^2/m k_{\rm B} T}$ is the *thermal wavelength*, which involves Planck's constant $h$ (with $\hbar = h/2\pi$). Let's now contrast two illustrative cases.

- *Adiabatic free expansion* – Suppose the volume freely expands from $V_{\rm i}$ to $V_{\rm f} = r\,V_{\rm i}$, with $r > 1$. Such an expansion can be effected by a removal of a partition between two chambers that are otherwise thermally insulated (see fig. 2.10). We have already seen how this process entails

$$\Delta E = Q = W = 0 \quad , \tag{2.118}$$

and thus for an ideal gas $T_{\rm f} = T_{\rm i}$. But the entropy changes! According to eqn. 2.116, we have

$$\Delta S = S_{\rm f} - S_{\rm i} = N k_{\rm B} \log r \quad . \tag{2.119}$$

- *Reversible adiabatic expansion* – If the gas expands quasistatically and reversibly, then $S = S(E, V, N)$ holds everywhere along the thermodynamic path. We then have, assuming $dN = 0$,

$$\begin{aligned} 0 = dS &= \tfrac{1}{2} f N k_{\rm B} \frac{dE}{E} + N k_{\rm B} \frac{dV}{V} \\ &= N k_{\rm B}\, d\log\left(V E^{f/2}\right) \quad . \end{aligned} \tag{2.120}$$

Integrating, we find

$$\frac{E}{E_0} = \left(\frac{V_0}{V}\right)^{2/f} \quad . \tag{2.121}$$

Thus,

$$E_{\rm f} = r^{-2/f}\, E_{\rm i} \quad \Longleftrightarrow \quad T_{\rm f} = r^{-2/f}\, T_{\rm i} \quad . \tag{2.122}$$

### 2.6.6 Example system

Consider a model thermodynamic system for which

$$E(S, V, N) = \frac{a S^3}{NV} \quad , \tag{2.123}$$

where $a$ is a constant. We have

$$dE = T\,dS - p\,dV + \mu\,dN \quad , \tag{2.124}$$

and therefore

$$T = \left(\frac{\partial E}{\partial S}\right)_{V,N} = \frac{3aS^2}{NV} \tag{2.125}$$

$$p = -\left(\frac{\partial E}{\partial V}\right)_{S,N} = \frac{aS^3}{NV^2} \tag{2.126}$$

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{S,V} = -\frac{aS^3}{N^2 V} \quad . \tag{2.127}$$

Choosing any two of these equations, we can eliminate $S$, which is inconvenient for experimental purposes. This yields three equations of state,

$$\frac{T^3}{p^2} = 27a\,\frac{V}{N} \qquad , \qquad \frac{T^3}{\mu^2} = 27a\,\frac{N}{V} \qquad , \qquad \frac{p}{\mu} = -\frac{N}{V} \quad , \tag{2.128}$$

only two of which are independent.

What about $C_V$ and $C_p$? To find $C_V$, we recast eqn. 2.125 as $S = \sqrt{NVT/3a}$. We then have

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_{V,N} = \frac{1}{2}\left(\frac{NVT}{3a}\right)^{1/2} = \frac{N}{18a}\,\frac{T^2}{p} \quad , \tag{2.129}$$

where the last equality on the RHS follows upon invoking the first of the equations of state in eqn. 2.128. To find $C_p$, we eliminate $V$ from eqns. 2.125 and 2.126, obtaining $T^2/p = 9aS/N$. From this we obtain

$$C_p = T\left(\frac{\partial S}{\partial T}\right)_{p,N} = \frac{2N}{9a}\,\frac{T^2}{p} \quad . \tag{2.130}$$

Thus, $C_p/C_V = 4$.

We can derive still more. To find the isothermal compressibility $\kappa_T = -V^{-1}(\partial V/\partial p)_{T,N}$, use the first of the equations of state in eqn. 2.128. To derive the adiabatic compressibility $\kappa_S = -V^{-1}(\partial V/\partial p)_{S,N}$, use eqn. 2.126, and then eliminate the inconvenient variable $S$.

Suppose we use this system as the working substance for a Carnot engine. Let's compute the work done and the engine efficiency. To do this, it is helpful to eliminate $S$ in the expression for the energy, and to rewrite the equation of state:

$$E = pV = \sqrt{\frac{N}{27a}}\,V^{1/2}\,T^{3/2} \qquad , \qquad p = \sqrt{\frac{N}{27a}}\,\frac{T^{3/2}}{V^{1/2}} \quad . \tag{2.131}$$

We assume $dN = 0$ throughout. We now see that for isotherms,

$$dT = 0 \;:\; \frac{E}{\sqrt{V}} = \text{constant} \tag{2.132}$$

Furthermore, since

$$đW\big|_T = \sqrt{\frac{N}{27a}}\,T^{3/2}\,\frac{dV}{V^{1/2}} = 2\,dE\big|_T \quad , \tag{2.133}$$

we conclude that

$$dT = 0 \;:\; W_{\rm if} = 2(E_{\rm f} - E_{\rm i}) \quad , \quad Q_{\rm if} = E_{\rm f} - E_{\rm i} + W_{\rm if} = 3(E_{\rm f} - E_{\rm i}) \quad . \tag{2.134}$$

For adiabats, eqn. 2.125 says $d(TV) = 0$, and therefore

$$đQ = 0 \;:\; TV = \text{constant} \quad , \quad \frac{E}{T} = \text{constant} \quad , \quad EV = \text{constant} \tag{2.135}$$

as well as $W_{\rm if} = E_{\rm i} - E_{\rm f}$. We can use these relations to derive the following:

$$E_{\rm B} = \sqrt{\frac{V_{\rm B}}{V_{\rm A}}}\,E_{\rm A} \quad , \quad E_{\rm C} = \frac{T_1}{T_2}\sqrt{\frac{V_{\rm B}}{V_{\rm A}}}\,E_{\rm A} \quad , \quad E_{\rm D} = \frac{T_1}{T_2}\,E_{\rm A} \quad . \tag{2.136}$$

Now we can write

$$\begin{aligned} W_{\rm AB} &= 2(E_{\rm B} - E_{\rm A}) = 2\left(\sqrt{\frac{V_{\rm B}}{V_{\rm A}}} - 1\right) E_{\rm A} \\ W_{\rm BC} &= (E_{\rm B} - E_{\rm C}) = \sqrt{\frac{V_{\rm B}}{V_{\rm A}}}\left(1 - \frac{T_1}{T_2}\right) E_{\rm A} \\ W_{\rm CD} &= 2(E_{\rm D} - E_{\rm C}) = \frac{2T_1}{T_2}\left(1 - \sqrt{\frac{V_{\rm B}}{V_{\rm A}}}\right) E_{\rm A} \\ W_{\rm DA} &= (E_{\rm D} - E_{\rm A}) = \left(\frac{T_1}{T_2} - 1\right) E_{\rm A} \end{aligned} \tag{2.137}$$

Adding up all the work, we obtain

$$\begin{aligned} W &= W_{\rm AB} + W_{\rm BC} + W_{\rm CD} + W_{\rm DA} \\ &= 3\left(\sqrt{\frac{V_{\rm B}}{V_{\rm A}}} - 1\right)\left(1 - \frac{T_1}{T_2}\right) E_{\rm A} \quad . \end{aligned} \tag{2.138}$$

Since

$$Q_{\rm AB} = 3(E_{\rm B} - E_{\rm A}) = \tfrac{3}{2} W_{\rm AB} = 3\left(\sqrt{\frac{V_{\rm B}}{V_{\rm A}}} - 1\right) E_{\rm A} \quad , \tag{2.139}$$

we find once again

$$\eta = \frac{W}{Q_{\rm AB}} = 1 - \frac{T_1}{T_2} \quad . \tag{2.140}$$

### 2.6.7 Measuring the entropy of a substance

If we can measure the heat capacity $C_V(T)$ or $C_p(T)$ of a substance as a function of temperature down to the lowest temperatures, then we can measure the entropy. At constant pressure, for example, we have $T\,dS = C_p\,dT$, hence

$$S(p,T) = S(p,T=0) + \int_0^T dT'\,\frac{C_p(T')}{T'} \quad . \tag{2.141}$$

The zero temperature entropy is $S(p,T=0) = k_{\rm B}\log g$ where $g$ is the quantum ground state degeneracy at pressure $p$. In cases where $g = 1$ we have $S(p,T=0) = 0$. In some systems, the presence of certain symmetries entails $g > 1$.

## 2.7 Thermodynamic Potentials

Thermodynamic systems may do work on their environments. Under certain constraints, the work done may be bounded from above by the change in an appropriately defined *thermodynamic potential*.

### 2.7.1 Energy $E$

Suppose we wish to create a thermodynamic system from scratch. Let's imagine that we create it from scratch in a thermally insulated box of volume $V$. The work we must do to assemble the system is then $\mathcal{W} = E$. After we bring all the constituent particles together, pulling them in from infinity (say), the system will have total energy $E$. After we finish, the system may not be in thermal equilibrium. Spontaneous processes will then occur so as to maximize the system's entropy, but the internal energy remains at $E$.

We have, from the First Law, $dE = đQ - đW$ and combining this with the Second Law in the form $đQ \le T\,dS$ yields

$$dE \le T\,dS - đW \quad . \tag{2.142}$$

Rearranging terms, we have $đW \le T\,dS - dE$. Hence, *the work done by a thermodynamic system under conditions of constant entropy is bounded above by $-dE$, and the maximum $đW$ is achieved for a reversible process.* It is sometimes useful to define the quantity

$$đW_{\rm free} \equiv đW - p\,dV \quad , \tag{2.143}$$

which is the differential work done by the system other than that required to change its volume. Then we have

$$đW_{\rm free} \le T\,dS - p\,dV - dE \quad , \tag{2.144}$$

and we conclude for systems at fixed $(S, V)$ that $đW_{\rm free} \le -dE$. Recall that in general we have $đW = -\boldsymbol{y} \cdot d\boldsymbol{X} - \boldsymbol{\mu} \cdot d\boldsymbol{N}$, with $\boldsymbol{X} = \{X_1, \ldots, X_{\mathsf{d}}\}$ and $\boldsymbol{N} = \{N_1, \ldots, N_{\mathsf{c}}\}$. For what we may call $(p, V)$ systems, there is only $\mathsf{d} = 1$ force/displacement pair, which is $(y, X) = (-p, V)$.

For single component $(p, V)$ systems in equilibrium, eqn. 2.142 says

$$dE = T\,dS - p\,dV + \mu\,dN \quad . \tag{2.145}$$

Therefore $E = E(S, V, N)$ with

$$T = \left(\frac{\partial E}{\partial S}\right)_{V,N} \quad , \quad -p = \left(\frac{\partial E}{\partial V}\right)_{S,N} \quad , \quad \mu = \left(\frac{\partial E}{\partial N}\right)_{S,V} \quad . \tag{2.146}$$

These expressions are easily generalized to multicomponent systems, magnetic systems, *etc.*

Now consider a single component $(p, V)$ system at fixed $(S, V, N)$. We conclude that $dE \le 0$, which says that spontaneous processes in a system with $dS = dV = dN = 0$ always lead to a reduction in the internal energy $E$. Therefore, *spontaneous processes drive the internal energy $E$ to a minimum in systems at fixed $(S, V, N)$.*

### 2.7.2 Helmholtz free energy $F$

Suppose we spontaneously create our system while it is in constant contact with a thermal reservoir at temperature $T$. Then as we create our system, it will absorb heat from the reservoir. Therefore, we don't have to supply the full internal energy $E$, but rather only $E - Q$, since the system receives heat energy

$Q$ from the reservoir. In other words, we must perform work $\mathcal{W} = E - TS$ to create our system, if it is constantly in equilibrium at temperature $T$. The quantity $E - TS$ is known as the *Helmholtz free energy*, $F$, which is related to the energy $E$ by a *Legendre transformation*,

$$F = E - TS \quad . \tag{2.147}$$

The general properties of Legendre transformations are discussed in Appendix III, §2.16.

Again invoking the Second Law, we have

$$dF \leq -S\,dT - đW \quad . \tag{2.148}$$

Rearranging terms, we have $đW \leq -S\,dT - dF$ , which says that *the work done by a thermodynamic system under conditions of constant temperature is bounded above by $-dF$, and the maximum $đW$ is achieved for a reversible process.* We also have the general result

$$đW_{\rm free} \leq -S\,dT - p\,dV - dF \quad , \tag{2.149}$$

and we conclude, for systems at fixed $(T, V)$, that $đW_{\rm free} \leq -dF$.

Under equilibrium conditions, the equality in eqn. 2.148 holds, and for single component $(p, V)$ systems where $đW = p\,dV - \mu\,dN$ we have $dF = -S\,dT - p\,dV + \mu\,dN$ . This says that $F = F(T, V, N)$ with

$$-S = \left(\frac{\partial F}{\partial T}\right)_{V,N} \quad , \quad -p = \left(\frac{\partial F}{\partial V}\right)_{T,N} \quad , \quad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \quad . \tag{2.150}$$

For spontaneous processes, $dF \leq -S\,dT - p\,dV + \mu\,dN$ says that *spontaneous processes drive the Helmholtz free energy $F$ to a minimum in systems at fixed $(T, V, N)$.*

### 2.7.3 Enthalpy H

Suppose that when we spontaneously create our system while it is thermally insulated, but in constant mechanical contact with a 'volume bath' at pressure $p$. For example, we could create our system inside a thermally insulated chamber with one movable wall where the external pressure is fixed at $p$. Thus, when creating the system, in addition to the system's internal energy $E$, we must also perform work $pV$ in order to make room for it. In other words, we must perform work $\mathcal{W} = E + pV$. The quantity $E + pV$ is known as the *enthalpy*, H.[15] The enthalpy is obtained from the energy via a different Legendre transformation than that used to obtain the Helmholtz free energy $F$, *i.e.*

$$\mathsf{H} = E + pV \quad . \tag{2.151}$$

Again invoking the Second Law, we have

$$d\mathsf{H} \leq T\,dS - đW + p\,dV + V\,dp \quad , \tag{2.152}$$

[15] We use a different font to distinguish enthalpy H from magnetic field $H$.

hence with $đW_{\rm free} = đW - p\,dV$, we have in general

$$đW_{\rm free} \le T\,dS + V dp - d\mathsf{H} \quad , \tag{2.153}$$

and we conclude, for systems at fixed $(S,p)$, that $đW_{\rm free} \le -d\mathsf{H}$.

In equilibrium, for single component $(p,V)$ systems,

$$d\mathsf{H} = T\,dS + V dp + \mu\,dN \quad , \tag{2.154}$$

which says $\mathsf{H} = \mathsf{H}(S,p,N)$, with

$$T = \left(\frac{\partial \mathsf{H}}{\partial S}\right)_{p,N} \quad , \quad V = \left(\frac{\partial \mathsf{H}}{\partial p}\right)_{S,N} \quad , \quad \mu = \left(\frac{\partial \mathsf{H}}{\partial N}\right)_{S,p} \quad . \tag{2.155}$$

For spontaneous processes, $d\mathsf{H} \le T\,dS + V dp + \mu\,dN$, which says that *spontaneous processes drive the enthalpy $H$ to a minimum in systems at fixed $(S,p,N)$.*

### 2.7.4 Gibbs free energy $G$

If we create a thermodynamic system at conditions of constant temperature $T$ and constant pressure $p$, then it absorbs heat energy $Q = TS$ from the reservoir and we must expend work energy $pV$ in order to make room for it. Thus, the total amount of work we must do in assembling our system is $\mathcal{W} = E - TS + pV$. This is the *Gibbs free energy*, $G$. The Gibbs free energy is obtained from $E$ after two Legendre transformations, *viz.*

$$G = E - TS + pV \tag{2.156}$$

Note that $G = F + pV = \mathsf{H} - TS$. The Second Law says that

$$dG \le -S\,dT + V dp + p\,dV - đW \quad , \tag{2.157}$$

which we may rearrange as $đW_{\rm free} \le -S\,dT + V dp - dG$. Accordingly, we conclude, for systems at fixed $(T,p)$, that $đW_{\rm free} \le -dG$.

For equilibrium single component $(p,V)$ systems, the differential of $G$ is

$$dG = -S\,dT + V dp + \mu\,dN \quad , \tag{2.158}$$

therefore $G = G(T,p,N)$, with

$$-S = \left(\frac{\partial G}{\partial T}\right)_{p,N} \quad , \quad V = \left(\frac{\partial G}{\partial p}\right)_{T,N} \quad , \quad \mu = \left(\frac{\partial G}{\partial N}\right)_{T,p} \quad . \tag{2.159}$$

Recall that Euler's theorem for single component systems requires $E = TS - pV + \mu N$ which says $G = \mu N$. Thus, the chemical potential $\mu$ is the Gibbs free energy per particle. For spontaneous processes, $dG \le -S\,dT + V dp + \mu\,dN$, hence *spontaneous processes drive the Gibbs free energy $G$ to a minimum in systems at fixed $(T,p,N)$.*

### 2.7.5 Grand potential $\Omega$

The grand potential, sometimes called the Landau free energy, is defined by

$$\Omega = E - TS - \sum_{a=1}^{\mathsf{c}} \mu_a N_a \quad . \tag{2.160}$$

The Second Law then requires

$$d\Omega \le -S\,dT - p\,dV - \sum_{a=1}^{\mathsf{c}} N_a\,d\mu_a - đ\widetilde{W}_{\rm free} \quad , \tag{2.161}$$

where we formally define

$$đ\widetilde{W}_{\rm free} \equiv đW - p\,dV + \sum_{a=1}^{\mathsf{c}} \mu_a\,dN_a \quad . \tag{2.162}$$

We then have $đW_{\rm free} \le -S\,dT - p\,dV - \boldsymbol{N}\cdot d\boldsymbol{\mu} - d\Omega$. We conclude that for systems at fixed $(T, V, \boldsymbol{\mu})$ that $đW_{\rm free} \le -d\Omega$.

Under equilibrium conditions for single component $(p, V)$ systems, we have

$$d\Omega = -S\,dT - p\,dV - N\,d\mu \quad , \tag{2.163}$$

hence

$$-S = \left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu} \quad , \quad -p = \left(\frac{\partial\Omega}{\partial V}\right)_{T,\mu} \quad , \quad -N = \left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} \quad . \tag{2.164}$$

Again invoking eqn. 2.110, we find $\Omega = -pV$, which says that the pressure is the negative of the grand potential per unit volume. For spontaneous processes, $d\Omega \le -S\,dT - p\,dV - Nd\mu$, hence *spontaneous processes drive the grand potential $\Omega$ to a minimum in systems at fixed $(T, V, \mu)$.*

## 2.8 Maxwell Relations

Maxwell relations are conditions equating certain derivatives of state variables which follow from the exactness of the differentials of the various state functions. For a general function $\Phi(u_1, u_2, \ldots, u_n)$, if $\Phi$ is of class $C^2$ or smoother (meaning all its second derivatives exist and are continuous), then

$$\frac{\partial^2\Phi}{\partial u_i\,\partial u_j} = \frac{\partial}{\partial u_i}\left(\frac{\partial\Phi}{\partial u_j}\right) = \frac{\partial}{\partial u_j}\left(\frac{\partial\Phi}{\partial u_i}\right) \quad .$$

Let us write

$$d\Phi = R_1\,du_1 + R_2\,du_2 + \ldots + R_n\,du_n \quad ,$$

with $R_j \equiv \partial\Phi/\partial u_j$ , where the derivative is computed holding all $u_i$ for $i \ne j$ constant. Then the equality of the mixed second partial derivatives is equivalent to the conditions

$$\left(\frac{\partial R_i}{\partial u_j}\right)_{u_{k(\ne j)}} = \left(\frac{\partial R_j}{\partial u_i}\right)_{u_{k(\ne i)}} \quad .$$

### 2.8.1 Relations deriving from $E(S, V, N)$

The energy $E(S, V, N)$ is a state function, with

$$dE = T\,dS - p\,dV + \mu\,dN \quad , \tag{2.165}$$

and therefore

$$T = \left(\frac{\partial E}{\partial S}\right)_{V,N} \qquad , \qquad -p = \left(\frac{\partial E}{\partial V}\right)_{S,N} \qquad , \qquad \mu = \left(\frac{\partial E}{\partial N}\right)_{S,V} \quad . \tag{2.166}$$

Taking the mixed second derivatives, we find

$$\frac{\partial^2 E}{\partial S\,\partial V} = \left(\frac{\partial T}{\partial V}\right)_{S,N} = -\left(\frac{\partial p}{\partial S}\right)_{V,N} \tag{2.167}$$

$$\frac{\partial^2 E}{\partial S\,\partial N} = \left(\frac{\partial T}{\partial N}\right)_{S,V} = \left(\frac{\partial \mu}{\partial S}\right)_{V,N} \tag{2.168}$$

$$\frac{\partial^2 E}{\partial V\,\partial N} = -\left(\frac{\partial p}{\partial N}\right)_{S,V} = \left(\frac{\partial \mu}{\partial V}\right)_{S,N} \quad . \tag{2.169}$$

### 2.8.2 Relations deriving from $F(T, V, N)$

The energy $F(T, V, N)$ is a state function, with

$$dF = -S\,dT - p\,dV + \mu\,dN \quad , \tag{2.170}$$

and therefore

$$-S = \left(\frac{\partial F}{\partial T}\right)_{V,N} \qquad , \qquad -p = \left(\frac{\partial F}{\partial V}\right)_{T,N} \qquad , \qquad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \quad . \tag{2.171}$$

Taking the mixed second derivatives, we find

$$\frac{\partial^2 F}{\partial T\,\partial V} = -\left(\frac{\partial S}{\partial V}\right)_{T,N} = -\left(\frac{\partial p}{\partial T}\right)_{V,N} \tag{2.172}$$

$$\frac{\partial^2 F}{\partial T\,\partial N} = -\left(\frac{\partial S}{\partial N}\right)_{T,V} = \left(\frac{\partial \mu}{\partial T}\right)_{V,N} \tag{2.173}$$

$$\frac{\partial^2 F}{\partial V\,\partial N} = -\left(\frac{\partial p}{\partial N}\right)_{T,V} = \left(\frac{\partial \mu}{\partial V}\right)_{T,N} \quad . \tag{2.174}$$

From Gibbs-Duhem, we know that for a single component system any three intensive quantities are related by an equation of state. Thus from the first of the above equations we may write

$$\left(\frac{\partial (S/N)}{\partial (V/N)}\right)_{T,N} = \left(\frac{\partial \mathsf{s}}{\partial \mathsf{v}}\right)_{T} = \left(\frac{\partial p}{\partial T}\right)_{\mathsf{v}} \quad , \tag{2.175}$$

since $\mathsf{s} = S/N$ may be expressed as a function of $T$ and $\mathsf{v} = V/N$ alone[16]. From the second equation, we have, similarly,

$$-\left(\frac{\partial(S/V)}{\partial(N/V)}\right)_{T,V} = \mathsf{v}\left(\frac{\partial \mathsf{s}}{\partial \mathsf{v}}\right)_T - \mathsf{s} = \left(\frac{\partial \mu}{\partial T}\right)_{\mathsf{v}} \quad . \tag{2.176}$$

### 2.8.3 Relations deriving from $\mathsf{H}(S, p, N)$

The enthalpy $\mathsf{H}(S, p, N)$ satisfies

$$d\mathsf{H} = T\,dS + V\,dp + \mu\,dN \quad , \tag{2.177}$$

which says $\mathsf{H} = \mathsf{H}(S, p, N)$, with

$$T = \left(\frac{\partial \mathsf{H}}{\partial S}\right)_{p,N} \quad , \quad V = \left(\frac{\partial \mathsf{H}}{\partial p}\right)_{S,N} \quad , \quad \mu = \left(\frac{\partial \mathsf{H}}{\partial N}\right)_{S,p} \quad . \tag{2.178}$$

Taking the mixed second derivatives, we find

$$\frac{\partial^2 \mathsf{H}}{\partial S\,\partial p} = \left(\frac{\partial T}{\partial p}\right)_{S,N} = \left(\frac{\partial V}{\partial S}\right)_{p,N} \tag{2.179}$$

$$\frac{\partial^2 \mathsf{H}}{\partial S\,\partial N} = \left(\frac{\partial T}{\partial N}\right)_{S,p} = \left(\frac{\partial \mu}{\partial S}\right)_{p,N} \tag{2.180}$$

$$\frac{\partial^2 \mathsf{H}}{\partial p\,\partial N} = \left(\frac{\partial V}{\partial N}\right)_{S,p} = \left(\frac{\partial \mu}{\partial p}\right)_{S,N} \quad . \tag{2.181}$$

The first and third of these entail

$$\mathsf{v} - \mathsf{s}\left(\frac{\partial \mathsf{v}}{\partial \mathsf{s}}\right)_p = \left(\frac{\partial \mu}{\partial p}\right)_{\mathsf{s}} \quad , \quad \left(\frac{\partial T}{\partial p}\right)_{\mathsf{s}} = \left(\frac{\partial \mathsf{v}}{\partial \mathsf{s}}\right)_p \quad . \tag{2.182}$$

### 2.8.4 Relations deriving from $G(T, p, N)$

The Gibbs free energy $G(T, p, N)$ satisfies

$$dG = -S\,dT + V\,dp + \mu\,dN \quad , \tag{2.183}$$

therefore $G = G(T, p, N)$, with

$$-S = \left(\frac{\partial G}{\partial T}\right)_{p,N} \quad , \quad V = \left(\frac{\partial G}{\partial p}\right)_{T,N} \quad , \quad \mu = \left(\frac{\partial G}{\partial N}\right)_{T,p} \quad . \tag{2.184}$$

[16] Recall that $v = N_{\rm A}\mathsf{v}$ is the molar volume and $s = N_{\rm A}\mathsf{s}$ the molar entropy.

Taking the mixed second derivatives, we find

$$\frac{\partial^2 G}{\partial T\,\partial p} = -\left(\frac{\partial S}{\partial p}\right)_{T,N} = \left(\frac{\partial V}{\partial T}\right)_{p,N} \tag{2.185}$$

$$\frac{\partial^2 G}{\partial T\,\partial N} = -\left(\frac{\partial S}{\partial N}\right)_{T,p} = \left(\frac{\partial \mu}{\partial T}\right)_{p,N} \tag{2.186}$$

$$\frac{\partial^2 G}{\partial p\,\partial N} = \left(\frac{\partial V}{\partial N}\right)_{T,p} = \left(\frac{\partial \mu}{\partial p}\right)_{T,N} \quad . \tag{2.187}$$

The second and third of these entail

$$\left(\frac{\partial \mu}{\partial T}\right)_p = -\mathsf{s} \qquad , \qquad \left(\frac{\partial \mu}{\partial p}\right)_T = \mathsf{v} \quad . \tag{2.188}$$

### 2.8.5 Relations deriving from $\Omega(T, V, \mu)$

The grand potential $\Omega(T, V, \mu)$ satisfies

$$d\Omega = -S\,dT - p\,dV - N\,d\mu \quad , \tag{2.189}$$

hence

$$-S = \left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} \qquad , \qquad -p = \left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} \qquad , \qquad -N = \left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} \quad . \tag{2.190}$$

Taking the mixed second derivatives, we find

$$\frac{\partial^2 \Omega}{\partial T\,\partial V} = -\left(\frac{\partial S}{\partial V}\right)_{T,\mu} = -\left(\frac{\partial p}{\partial T}\right)_{V,\mu} \tag{2.191}$$

$$\frac{\partial^2 \Omega}{\partial T\,\partial \mu} = -\left(\frac{\partial S}{\partial \mu}\right)_{T,V} = -\left(\frac{\partial N}{\partial T}\right)_{V,\mu} \tag{2.192}$$

$$\frac{\partial^2 \Omega}{\partial V\,\partial \mu} = -\left(\frac{\partial p}{\partial \mu}\right)_{T,V} = -\left(\frac{\partial N}{\partial V}\right)_{T,\mu} \quad . \tag{2.193}$$

The first and third of these entail

$$\left(\frac{\partial p}{\partial T}\right)_\mu = \frac{\mathsf{s}}{\mathsf{v}} \qquad , \qquad \left(\frac{\partial p}{\partial \mu}\right)_T = \frac{1}{\mathsf{v}} \quad . \tag{2.194}$$

#### Relations deriving from $S(E, V, N)$

We can also derive Maxwell relations based on the entropy $S(E, V, N)$ itself. For example, we have

$$dS = \frac{1}{T}\,dE + \frac{p}{T}\,dV - \frac{\mu}{T}\,dN \quad . \tag{2.195}$$

Therefore $S = S(E, V, N)$ and

$$\frac{\partial^2 S}{\partial E\,\partial V} = \left(\frac{\partial(T^{-1})}{\partial V}\right)_{E,N} = \left(\frac{\partial(pT^{-1})}{\partial E}\right)_{V,N} \quad , \tag{2.196}$$

*et cetera*.

### 2.8.6 Generalized thermodynamic potentials

We have up until now assumed a generalized force-displacement pair $(y, X) = (-p, V)$. But the above results also generalize to *e.g.* magnetic systems, where $(y, X) = (H, M)$. In general, we have

| | | |
|---|---|---|
| THIS SPACE AVAILABLE | $dE = T\,dS + y\,dX + \mu\,dN$ | (2.197) |
| $F = E - TS$ | $dF = -S\,dT + y\,dX + \mu\,dN$ | (2.198) |
| $\mathsf{H} = E - yX$ | $d\mathsf{H} = T\,dS - X\,dy + \mu\,dN$ | (2.199) |
| $G = E - TS - yX$ | $dG = -S\,dT - X\,dy + \mu\,dN$ | (2.200) |
| $\Omega = E - TS - \mu N$ | $d\Omega = -S\,dT + y\,dX - N\,d\mu$ . | (2.201) |

Generalizing $(-p, V) \to (y, X)$, we also obtain, *mutatis mutandis*, the following Maxwell relations:

$$\left(\frac{\partial T}{\partial X}\right)_{S,N} = \left(\frac{\partial y}{\partial S}\right)_{X,N} \qquad \left(\frac{\partial T}{\partial N}\right)_{S,X} = \left(\frac{\partial \mu}{\partial S}\right)_{X,N} \qquad \left(\frac{\partial y}{\partial N}\right)_{S,X} = \left(\frac{\partial \mu}{\partial X}\right)_{S,N} \tag{2.202}$$

$$\left(\frac{\partial T}{\partial y}\right)_{S,N} = -\left(\frac{\partial X}{\partial S}\right)_{y,N} \qquad \left(\frac{\partial T}{\partial N}\right)_{S,y} = \left(\frac{\partial \mu}{\partial S}\right)_{y,N} \qquad \left(\frac{\partial X}{\partial N}\right)_{S,y} = -\left(\frac{\partial \mu}{\partial y}\right)_{S,N} \tag{2.203}$$

$$\left(\frac{\partial S}{\partial X}\right)_{T,N} = -\left(\frac{\partial y}{\partial T}\right)_{X,N} \qquad \left(\frac{\partial S}{\partial N}\right)_{T,X} = -\left(\frac{\partial \mu}{\partial T}\right)_{X,N} \qquad \left(\frac{\partial y}{\partial N}\right)_{T,X} = \left(\frac{\partial \mu}{\partial X}\right)_{T,N} \tag{2.204}$$

$$\left(\frac{\partial S}{\partial y}\right)_{T,N} = \left(\frac{\partial X}{\partial T}\right)_{y,N} \qquad \left(\frac{\partial S}{\partial N}\right)_{T,y} = -\left(\frac{\partial \mu}{\partial T}\right)_{y,N} \qquad \left(\frac{\partial X}{\partial N}\right)_{T,y} = -\left(\frac{\partial \mu}{\partial y}\right)_{T,N} \tag{2.205}$$

$$\left(\frac{\partial S}{\partial X}\right)_{T,\mu} = -\left(\frac{\partial y}{\partial T}\right)_{X,\mu} \qquad \left(\frac{\partial S}{\partial \mu}\right)_{T,X} = \left(\frac{\partial N}{\partial T}\right)_{X,\mu} \qquad \left(\frac{\partial y}{\partial \mu}\right)_{T,X} = -\left(\frac{\partial N}{\partial X}\right)_{T,\mu} \quad . \tag{2.206}$$

## 2.9 Equilibrium and Stability

### 2.9.1 Equilibrium

Suppose we have two systems, A and B, which are free to exchange energy, volume, and particle number, subject to overall conservation rules

$$E_{\rm A} + E_{\rm B} = E \quad , \quad V_{\rm A} + V_{\rm B} = V \quad , \quad N_{\rm A} + N_{\rm B} = N \quad , \tag{2.207}$$

where $E$, $V$, and $N$ are fixed. Now let us compute the change in the total entropy of the combined systems when they are allowed to exchange energy, volume, or particle number. We assume that the entropy is additive, *i.e.*

$$\begin{aligned} dS = &\left[\left(\frac{\partial S_{\rm A}}{\partial E_{\rm A}}\right)_{V_{\rm A},N_{\rm A}} - \left(\frac{\partial S_{\rm B}}{\partial E_{\rm B}}\right)_{V_{\rm B},N_{\rm B}}\right] dE_{\rm A} + \left[\left(\frac{\partial S_{\rm A}}{\partial V_{\rm A}}\right)_{E_{\rm A},N_{\rm A}} - \left(\frac{\partial S_{\rm B}}{\partial V_{\rm B}}\right)_{E_{\rm B},N_{\rm B}}\right] dV_{\rm A} \\ &+ \left[\left(\frac{\partial S_{\rm A}}{\partial N_{\rm A}}\right)_{E_{\rm A},V_{\rm A}} - \left(\frac{\partial S_{\rm B}}{\partial N_{\rm B}}\right)_{E_{\rm B},V_{\rm B}}\right] dN_{\rm A} \quad . \end{aligned} \tag{2.208}$$

Note that we have used $dE_{\rm B} = -dE_{\rm A}$, $dV_{\rm B} = -dV_{\rm A}$, and $dN_{\rm B} = -dN_{\rm A}$. Now we know from the Second Law that spontaneous processes result in $T\,dS > 0$, which means that $S$ tends to a maximum. If $S$ is a maximum, it must be that the coefficients of $dE_{\rm A}$, $dV_{\rm A}$, and $dN_{\rm A}$ all vanish, else we could increase the total entropy of the system by a judicious choice of these three differentials. From $T\,dS = dE + p\,dV - \mu\,dN$, we have

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N} \quad , \quad \frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,N} \quad , \quad \frac{\mu}{T} = -\left(\frac{\partial S}{\partial N}\right)_{E,V} \quad . \tag{2.209}$$

Thus, we conclude that in order for the system to be in equilibrium, so that $S$ is maximized and can increase no further under spontaneous processes, we must have

$$T_{\rm A} = T_{\rm B} \qquad \text{(thermal equilibrium)} \tag{2.210}$$

$$\frac{p_{\rm A}}{T_{\rm A}} = \frac{p_{\rm B}}{T_{\rm B}} \qquad \text{(mechanical equilibrium)} \tag{2.211}$$

$$\frac{\mu_{\rm A}}{T_{\rm A}} = \frac{\mu_{\rm B}}{T_{\rm B}} \qquad \text{(chemical equilibrium)} \tag{2.212}$$

In general, with conjugate force and displacement $(y, X)$, where $X$ can be exchanged between subsystems, equilibrium requires $y_{\rm A}/T_{\rm A} = y_{\rm B}/T_{\rm B}$.

### 2.9.2 Stability

Next, consider a uniform system with energy $E' = 2E$, volume $V' = 2V$, and particle number $N' = 2N$. We wish to check that this system is not unstable with respect to spontaneously becoming inhomogeneous. To that end, we imagine dividing the system in half. Each half would have energy $E$, volume $V$, and particle number $N$. But suppose we divided up these quantities differently, so that the left half had

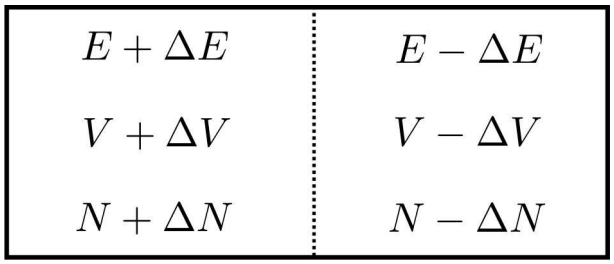



Figure 2.16: To check for an instability, we compare the energy of a system to its total energy when we reapportion its energy, volume, and particle number slightly unequally.

slightly different energy, volume, and particle number than the right, as depicted in fig. 2.16. Does the entropy increase or decrease? We have

$$\begin{aligned}\Delta S &= S(E+\Delta E, V+\Delta V, N+\Delta N) + S(E-\Delta E, V-\Delta V, N-\Delta N) - S(2E, 2V, 2N) \\ &= \frac{\partial^2 S}{\partial E^2}\,(\Delta E)^2 + \frac{\partial^2 S}{\partial V^2}\,(\Delta V)^2 + \frac{\partial^2 S}{\partial N^2}\,(\Delta N)^2 \\ &\qquad + 2\,\frac{\partial^2 S}{\partial E\,\partial V}\,\Delta E\,\Delta V + 2\,\frac{\partial^2 S}{\partial E\,\partial N}\,\Delta E\,\Delta N + 2\,\frac{\partial^2 S}{\partial V\,\partial N}\,\Delta V\,\Delta N \quad .\end{aligned} \tag{2.213}$$

Thus, we can write

$$\Delta S = \sum_{i,j} Q_{ij}\,\Psi_i\,\Psi_j \quad , \tag{2.214}$$

where

$$Q = \begin{pmatrix} \frac{\partial^2 S}{\partial E^2} & \frac{\partial^2 S}{\partial E\,\partial V} & \frac{\partial^2 S}{\partial E\,\partial N} \\ \frac{\partial^2 S}{\partial E\,\partial V} & \frac{\partial^2 S}{\partial V^2} & \frac{\partial^2 S}{\partial V\,\partial N} \\ \frac{\partial^2 S}{\partial E\,\partial N} & \frac{\partial^2 S}{\partial V\,\partial N} & \frac{\partial^2 S}{\partial N^2} \end{pmatrix} \tag{2.215}$$

is the matrix of second derivatives, known in mathematical parlance as the *Hessian*, and $\boldsymbol{\Psi} = (\Delta E, \Delta V, \Delta N)$. Note that $Q$ is a symmetric matrix.

Since $S$ must be a maximum in order for the system to be in equilibrium, we are tempted to conclude that the homogeneous system is stable if and only if all three eigenvalues of $Q$ are negative. If one or more of the eigenvalues is positive, then it is possible to choose a set of variations $\boldsymbol{\Psi}$ such that $\Delta S > 0$, which would contradict the assumption that the homogeneous state is one of maximum entropy. A matrix with this restriction is said to be *negative definite*. While it is true that $Q$ can have no positive eigenvalues, it is clear from homogeneity of $S(E, V, N)$ that one of the three eigenvalues must be zero, corresponding to the eigenvector $\boldsymbol{\Psi} = (E, V, N)$. Homogeneity means $S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N)$. Now let us take $\lambda = 1 + \eta$, where $\eta$ is infinitesimal. Then $\Delta E = \eta E$, $\Delta V = \eta V$, and $\Delta N = \eta N$, and homogeneity says $S(E \pm \Delta E, V \pm \Delta V, N \pm \Delta N) = (1 \pm \eta)\,S(E, V, N)$ and $\Delta S = (1+\eta)S + (1-\eta)S - 2S = 0$. We then have a slightly weaker characterization of $Q$ as *negative semidefinite*.

However, if we fix one of the components of $(\Delta E, \Delta V, \Delta N)$ to be zero, then $\boldsymbol{\Psi}$ must have some component orthogonal to the zero eigenvector, in which case $\Delta S < 0$. Suppose we set $\Delta N = 0$ and we

just examine the stability with respect to inhomogeneities in energy and volume. We then restrict our attention to the upper left $2 \times 2$ submatrix of $Q$. A general symmetric $2 \times 2$ matrix may be written

$$Q = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \tag{2.216}$$

It is easy to solve for the eigenvalues of $Q$. One finds

$$\lambda_\pm = \left( \frac{a+c}{2} \right) \pm \sqrt{\left( \frac{a-c}{2} \right)^2 + b^2} \quad . \tag{2.217}$$

In order for $Q$ to be negative definite, we require $\lambda_+ < 0$ and $\lambda_- < 0$. Thus, $\mathrm{Tr}\, Q = a + c = \lambda_+ + \lambda_- < 0$ and $\det Q = ac - b^2 = \lambda_+ \, \lambda_- > 0$. Taken together, these conditions require

$$a < 0 \qquad , \qquad c < 0 \qquad , \qquad ac > b^2 \quad . \tag{2.218}$$

Going back to thermodynamic variables, this requires

$$\frac{\partial^2 S}{\partial E^2} < 0 \qquad , \qquad \frac{\partial^2 S}{\partial V^2} < 0 \qquad , \qquad \frac{\partial^2 S}{\partial E^2} \cdot \frac{\partial^2 S}{\partial V^2} > \left( \frac{\partial^2 S}{\partial E \, \partial V} \right)^2 \quad . \tag{2.219}$$

Thus *the entropy is a concave function of $E$ and $V$ at fixed $N$*. Had we set $\Delta E = 0$ and considered the lower right $2 \times 2$ block of $Q$, we'd have concluded that $S(V, N)$ is concave at fixed $E$. Since $(\partial S/\partial E)_V = T^{-1}$, we have $\partial^2 S/\partial E^2 = -T^{-2}(\partial T/\partial E)_V = -1/T^2 C_V < 0$, whence $C_V > 0$ for stability.

Many thermodynamic systems are held at fixed $(T, p, N)$, which suggests we examine the stability criteria for $G(T, p, N)$. Suppose our system is in equilibrium with a reservoir at temperature $T_0$ and pressure $p_0$. Then, suppressing $N$ (which is assumed constant), we have

$$G(T_0, p_0) = E - T_0 \, S + p_0 \, V \quad . \tag{2.220}$$

Now suppose there is a fluctuation in the entropy and the volume of our system, which is held at fixed particle number. Going to second order in $\Delta S$ and $\Delta V$, we have

$$\begin{aligned} \Delta G = \left[ \left( \frac{\partial E}{\partial S} \right)_V - T_0 \right] \Delta S + \left[ \left( \frac{\partial E}{\partial V} \right)_S + p_0 \right] \Delta V \\ + \frac{1}{2} \left[ \frac{\partial^2 E}{\partial S^2} \, (\Delta S)^2 + 2 \, \frac{\partial^2 E}{\partial S \, \partial V} \, \Delta S \, \Delta V + \frac{\partial^2 E}{\partial V^2} \, (\Delta V)^2 \right] + \dots \quad . \end{aligned} \tag{2.221}$$

Equilibrium requires that the coefficients of $\Delta S$ and $\Delta V$ both vanish, *i.e.* that $T = (\partial E/\partial S)_{V,N} = T_0$ and $p = -(\partial E/\partial V)_{S,N} = p_0$. The condition for stability is that $\Delta G > 0$ for all $(\Delta S, \Delta V)$. Stability therefore requires that the Hessian matrix $Q$ be positive definite, with

$$Q = \begin{pmatrix} \frac{\partial^2 E}{\partial S^2} & \frac{\partial^2 E}{\partial S \, \partial V} \\ \\ \frac{\partial^2 E}{\partial S \, \partial V} & \frac{\partial^2 E}{\partial V^2} \end{pmatrix} \quad . \tag{2.222}$$

Thus, we have the following three conditions:

$$\frac{\partial^2 E}{\partial S^2} = \left(\frac{\partial T}{\partial S}\right)_V = \frac{T}{C_V} > 0 \tag{2.223}$$

$$\frac{\partial^2 E}{\partial V^2} = -\left(\frac{\partial p}{\partial V}\right)_S = \frac{1}{V\kappa_S} > 0 \tag{2.224}$$

$$\frac{\partial^2 E}{\partial S^2}\cdot\frac{\partial^2 E}{\partial V^2} - \left(\frac{\partial^2 E}{\partial S\,\partial V}\right)^2 = \frac{T}{V\kappa_S\,C_V} - \left(\frac{\partial T}{\partial V}\right)_S^2 > 0 \quad . \tag{2.225}$$

As we shall discuss below, the quantity $\alpha_S \equiv V^{-1}(\partial V/\partial T)_{S,N}$ is the adiabatic thermal expansivity coefficient. We therefore conclude that stability of any thermodynamic system requires

$$\frac{C_V}{T} > 0 \quad , \quad \kappa_S > 0 \quad , \quad |\alpha_S| > \sqrt{\frac{\kappa_S\,C_V}{VT}} \quad . \tag{2.226}$$

## 2.10 Applications of Thermodynamics

A discussion of various useful mathematical relations among partial derivatives may be found in the appendix in §2.17. Some facility with the differential multivariable calculus is extremely useful in the analysis of thermodynamics problems.

### 2.10.1 Adiabatic free expansion revisited

Consider once again the adiabatic free expansion of a gas from initial volume $V_{\rm i}$ to final volume $V_{\rm f} = rV_{\rm i}$. Since the system is not in equilibrium during the free expansion process, the initial and final states do not lie along an adiabat, *i.e.* they do not have the same entropy. Rather, as we found, from $Q = W = 0$, we have that $E_{\rm i} = E_{\rm f}$, which means they have the same *energy*, and, in the case of an ideal gas, the same temperature (assuming $N$ is constant). Thus, the initial and final states lie along an *isotherm*. The situation is depicted in fig. 2.17. Now let us compute the change in entropy $\Delta S = S_{\rm f} - S_{\rm i}$ by integrating along this isotherm. Note that the actual dynamics are *irreversible* and *do not* quasistatically follow any continuous thermodynamic path. However, we can use what is a fictitious thermodynamic path as a means of comparing $S$ in the initial and final states.

We have

$$\Delta S = S_{\rm f} - S_{\rm i} = \int_{V_{\rm i}}^{V_{\rm f}} dV \left(\frac{\partial S}{\partial V}\right)_{T,N} \quad . \tag{2.227}$$

But from a Maxwell equation deriving from $F$, we have

$$\left(\frac{\partial S}{\partial V}\right)_{T,N} = \left(\frac{\partial p}{\partial T}\right)_{V,N} \quad , \tag{2.228}$$

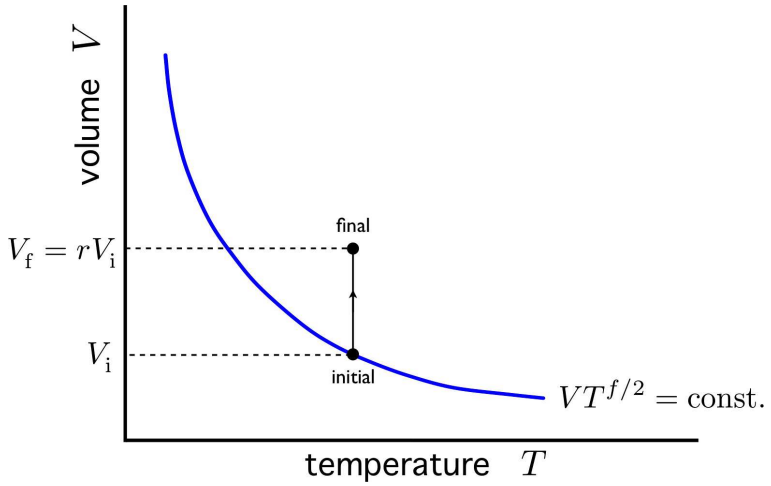



Figure 2.17: Adiabatic free expansion via a thermal path. The initial and final states do not lie along an adiabat! Rather, for an ideal gas, the initial and final states lie along an isotherm.

hence

$$\Delta S = \int_{V_{\rm i}}^{V_{\rm f}} dV \left(\frac{\partial p}{\partial T}\right)_{V,N} \quad . \tag{2.229}$$

For an ideal gas, we can use the equation of state $pV = Nk_{\rm B}T$ to obtain

$$\left(\frac{\partial p}{\partial T}\right)_{V,N} = \frac{Nk_{\rm B}}{V} \quad . \tag{2.230}$$

The integral can now be computed:

$$\Delta S = \int_{V_{\rm i}}^{rV_{\rm i}} dV \, \frac{Nk_{\rm B}}{V} = Nk_{\rm B} \log r \quad , \tag{2.231}$$

as we found before, in eqn. 2.119. What is different about this derivation? Previously, we derived the entropy change from the explicit formula for $S(E, V, N)$. Here, we did not need to know this function. The Maxwell relation allowed us to compute the entropy change using only the equation of state.

### 2.10.2 Energy and volume

We saw how $E(T, V, N) = \frac{1}{2} f N k_{\rm B} T$ for an ideal gas, independent of the volume. In general we should have

$$E(T, V, N) = N \, \phi\big(T, \tfrac{V}{N}\big) \quad . \tag{2.232}$$

For the ideal gas, $\phi\big(T, V/N\big) = \frac{1}{2} f k_{\rm B} T$ is a function of $T$ alone and is independent on the other intensive quantity $V/N$. How does energy vary with volume? At fixed temperature and particle number, we

have, from $E = F + TS$,

$$\left(\frac{\partial E}{\partial V}\right)_{T,N} = \left(\frac{\partial F}{\partial V}\right)_{T,N} + T\left(\frac{\partial S}{\partial V}\right)_{T,N} = -p + T\left(\frac{\partial p}{\partial T}\right)_{V,N} \quad , \tag{2.233}$$

where we have used the Maxwell relation $(\partial S/\partial V)_{T,N} = (\partial p/\partial T)_{V,N}$, derived from the mixed second derivative $\partial^2 F/\partial T\,\partial V$. Another way to derive this result is as follows. Write $dE = T\,dS - p\,dV + \mu\,dN$ and then express $dS$ in terms of $dT$, $dV$, and $dN$, resulting in

$$dE = T\left(\frac{\partial S}{\partial T}\right)_{V,N} dT + \left[T\left(\frac{\partial S}{\partial V}\right)_{T,N} - p\right] dV - \left[T\left(\frac{\partial \mu}{\partial T}\right)_{V,N} + \mu\right] dN \quad . \tag{2.234}$$

Now read off $(\partial E/\partial V)_{T,N}$ and use the same Maxwell relation as before to recover eqn. 2.233. Applying this result to the ideal gas law $pV = Nk_{\rm B}T$ results in the vanishing of the RHS, hence for any substance obeying the ideal gas law we must have

$$E(T, V, N) = \nu\,\varepsilon(T) = N\varepsilon(T)/N_{\rm A} \quad . \tag{2.235}$$

### 2.10.3 van der Waals equation of state

It is clear that the same conclusion follows for any equation of state of the form $p(T, V, N) = T \cdot f(V/N)$, where $f(V/N)$ is an arbitrary function of its argument: the ideal gas law remains valid[17]. This is not true, however, for the van der Waals equation of state,

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT \quad , \tag{2.236}$$

where $v = N_{\rm A}V/N$ is the molar volume. We then find (always assuming constant $N$),

$$\left(\frac{\partial E}{\partial V}\right)_T = \left(\frac{\partial \varepsilon}{\partial v}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_V - p = \frac{a}{v^2} \quad , \tag{2.237}$$

where $E(T, V, N) \equiv \nu\,\varepsilon(T, v)$. We can integrate this to obtain

$$\varepsilon(T, v) = \omega(T) - \frac{a}{v} \quad , \tag{2.238}$$

where $\omega(T)$ is arbitrary. From eqn. 2.31, we immediately have

$$c_V = \left(\frac{\partial \varepsilon}{\partial T}\right)_v = \omega'(T) \quad . \tag{2.239}$$

What about $c_p$? This requires a bit of work. We start with eqn. 2.32,

$$c_p = \left(\frac{\partial \varepsilon}{\partial T}\right)_p + p\left(\frac{\partial v}{\partial T}\right)_p = c_V + \left(p + \frac{a}{v^2}\right)\left(\frac{\partial v}{\partial T}\right)_p \quad . \tag{2.240}$$

[17] Note $V/N = v/N_{\rm A}$.

| gas | $a\ \left(\frac{\mathrm{L}^2\cdot\mathrm{bar}}{\mathrm{mol}^2}\right)$ | $b\ \left(\frac{\mathrm{L}}{\mathrm{mol}}\right)$ | $p_c$ (bar) | $T_c$ (K) | $v_c$ (L/mol) |
|---|---|---|---|---|---|
| Acetone | 14.09 | 0.0994 | 52.82 | 505.1 | 0.2982 |
| Argon | 1.363 | 0.03219 | 48.72 | 150.9 | 0.0966 |
| Carbon dioxide | 3.640 | 0.04267 | 74.04 | 304.0 | 0.1280 |
| Ethanol | 12.18 | 0.08407 | 63.83 | 516.3 | 0.2522 |
| Freon | 10.78 | 0.0998 | 40.09 | 384.9 | 0.2994 |
| Helium | 0.03457 | 0.0237 | 2.279 | 5.198 | 0.0711 |
| Hydrogen | 0.2476 | 0.02661 | 12.95 | 33.16 | 0.0798 |
| Mercury | 8.200 | 0.01696 | 1055 | 1723 | 0.0509 |
| Methane | 2.283 | 0.04278 | 46.20 | 190.2 | 0.1283 |
| Nitrogen | 1.408 | 0.03913 | 34.06 | 128.2 | 0.1174 |
| Oxygen | 1.378 | 0.03183 | 50.37 | 154.3 | 0.0955 |
| Water | 5.536 | 0.03049 | 220.6 | 647.0 | 0.0915 |

Table 2.3: Van der Waals parameters for some common gases. (*Source: Wikipedia.*)

We may rewrite the equation of state as

$$p = \frac{RT}{v-b} - \frac{a}{v^2} \quad . \tag{2.241}$$

Taking the differential of both sides,

$$\begin{aligned} dp &= \left(\frac{\partial p}{\partial T}\right)_v dT + \left(\frac{\partial p}{\partial v}\right)_T dv \\ &= \frac{R}{v-b}\,dT + \left[-\frac{RT}{(v-b)^2} + \frac{2a}{v^3}\right] dv \quad . \end{aligned} \tag{2.242}$$

Setting $dp = 0$, we may read off

$$\left(\frac{\partial v}{\partial T}\right)_p = -\left(\frac{\partial p}{\partial T}\right)_v \Bigg/ \left(\frac{\partial p}{\partial v}\right)_T \quad , \tag{2.243}$$

which follows from eqn. 2.487 in the mathematical appendix §2.17 below. Putting this all together, we obtain

$$\left(\frac{\partial v}{\partial T}\right)_p = \frac{Rv^3(v-b)}{RTv^3 - 2a(v-b)^2} \quad . \tag{2.244}$$

One immediate result is the following expression for the isobaric thermal expansion coefficient,

$$\alpha_p = \frac{1}{v}\left(\frac{\partial v}{\partial T}\right)_p = \frac{Rv^2(v-b)}{RTv^3 - 2a(v-b)^2} \quad . \tag{2.245}$$

Another result is the difference $c_p - c_V$ from eqn. 2.240,

$$c_p - c_V = \left(p + \frac{a}{v^2}\right)\left(\frac{\partial v}{\partial T}\right)_p = \frac{R^2 T v^3}{RTv^3 - 2a(v-b)^2} \quad . \tag{2.246}$$

To fix $\omega(T)$ and thus $c_V = \omega'(T)$, we consider the $v \to \infty$ limit, where the density of the gas vanishes. In this limit, the gas must be ideal, hence eqn. 2.238 says that $\omega(T) = \frac{1}{2} f RT$. Therefore $c_V(T, v) = \frac{1}{2} f R$, just as in the case of an ideal gas. However, rather than $c_p = c_V + R$, which holds for ideal gases, $c_p(T, v)$ is given by eqn. 2.246, and

$$\begin{aligned} c_V^{\rm VDW} &= \tfrac{1}{2} f R \\ c_p^{\rm VDW} &= \tfrac{1}{2} f R + \frac{R^2 T v^3}{RTv^3 - 2a(v-b)^2} \quad . \end{aligned} \tag{2.247}$$

Note that $c_p(a \to 0) = c_V + R$, which is the ideal gas result.

As we shall see in chapter 7, the van der Waals system is unstable throughout a region of parameters, where it undergoes *phase separation* between high density (liquid) and low density (gas) phases. The above results are valid only in the stable regions of the phase diagram.

### 2.10.4 General equation of state

Suppose we can't isolate any of the state variables in the equation of state as we did when writing $p = p(T, v)$ above. Rather, let the equation of state take the form $Z(T, p, v) = 0$. Setting consecutively $dv = 0$, $dT = 0$, and $dp = 0$, we then obtain

$$dZ = \left(\frac{\partial Z}{\partial T}\right)_{p,v} dT + \left(\frac{\partial Z}{\partial p}\right)_{T,v} dp + \left(\frac{\partial Z}{\partial v}\right)_{T,p} dv = 0 \quad . \tag{2.248}$$

We then have

$$\begin{aligned} \left(\frac{\partial T}{\partial p}\right)_v &= -\left(\frac{\partial Z}{\partial p}\right)_{T,v} \Big/ \left(\frac{\partial Z}{\partial T}\right)_{p,v} \\ \left(\frac{\partial p}{\partial v}\right)_T &= -\left(\frac{\partial Z}{\partial v}\right)_{T,p} \Big/ \left(\frac{\partial Z}{\partial p}\right)_{T,v} \\ \left(\frac{\partial v}{\partial T}\right)_p &= -\left(\frac{\partial Z}{\partial T}\right)_{p,v} \Big/ \left(\frac{\partial Z}{\partial v}\right)_{T,p} \quad . \end{aligned} \tag{2.249}$$

### 2.10.5 Thermodynamic response functions

Consider the entropy $S$ expressed as a function of $T$, $V$, and $N$:

$$dS = \left(\frac{\partial S}{\partial T}\right)_{V,N} dT + \left(\frac{\partial S}{\partial V}\right)_{T,N} dV + \left(\frac{\partial S}{\partial N}\right)_{T,V} dN \quad . \tag{2.250}$$

Dividing by $dT$, multiplying by $T$, and assuming $dN = 0$ throughout, we have

$$C_p - C_V = T \left( \frac{\partial S}{\partial V} \right)_T \left( \frac{\partial V}{\partial T} \right)_p \quad . \tag{2.251}$$

Appealing to a Maxwell relation derived from $F(T, V, N)$, and then appealing to eqn. 2.486, we have

$$\left( \frac{\partial S}{\partial V} \right)_T = \left( \frac{\partial p}{\partial T} \right)_V = - \left( \frac{\partial p}{\partial V} \right)_T \left( \frac{\partial V}{\partial T} \right)_p . \tag{2.252}$$

This allows us to write

$$C_p - C_V = -T \left( \frac{\partial p}{\partial V} \right)_T \left( \frac{\partial V}{\partial T} \right)_p^2 \quad . \tag{2.253}$$

We define the *response functions*,

$$\text{isothermal compressibility:} \quad \kappa_T = -\frac{1}{V} \left( \frac{\partial V}{\partial p} \right)_T = -\frac{1}{V} \frac{\partial^2 G}{\partial p^2} \tag{2.254}$$

$$\text{adiabatic compressibility:} \quad \kappa_S = -\frac{1}{V} \left( \frac{\partial V}{\partial p} \right)_S = -\frac{1}{V} \frac{\partial^2 H}{\partial p^2} \tag{2.255}$$

$$\text{thermal expansivity:} \quad \alpha_p = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_p \quad . \tag{2.256}$$

Thus,

$$C_p - C_V = V \frac{T \alpha_p^2}{\kappa_T} \quad , \tag{2.257}$$

or, in terms of intensive quantities,

$$c_p - c_V = \frac{v \, T \alpha_p^2}{\kappa_T} \quad , \tag{2.258}$$

where, as always, $v = V N_{\rm A}/N$ is the molar volume.

This above relation generalizes to any conjugate force-displacement pair $(-p, V) \to (y, X)$:

$$C_y - C_X = -T \left( \frac{\partial y}{\partial T} \right)_X \left( \frac{\partial X}{\partial T} \right)_y = T \left( \frac{\partial y}{\partial X} \right)_T \left( \frac{\partial X}{\partial T} \right)_y^2 \quad . \tag{2.259}$$

For example, we could have $(y, X) = (H^\alpha, M^\alpha)$.

A similar relationship can be derived between the compressibilities $\kappa_T$ and $\kappa_S$. We then clearly must start with the volume, writing

$$dV = \left( \frac{\partial V}{\partial p} \right)_{S,N} dp + \left( \frac{\partial V}{\partial S} \right)_{p,N} dS + \left( \frac{\partial V}{\partial N} \right)_{S,p} dN \quad . \tag{2.260}$$

Dividing by $dp$, multiplying by $-V^{-1}$, and keeping $N$ constant, we have

$$\kappa_T - \kappa_S = -\frac{1}{V} \left( \frac{\partial V}{\partial S} \right)_p \left( \frac{\partial S}{\partial p} \right)_T \quad . \tag{2.261}$$

Again we appeal to a Maxwell relation, writing

$$\left(\frac{\partial S}{\partial p}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_p \quad , \tag{2.262}$$

and after invoking the chain rule,

$$\left(\frac{\partial V}{\partial S}\right)_p = \left(\frac{\partial V}{\partial T}\right)_p \left(\frac{\partial T}{\partial S}\right)_p = \frac{T}{C_p}\left(\frac{\partial V}{\partial T}\right)_p \quad , \tag{2.263}$$

we obtain

$$\kappa_T - \kappa_S = \frac{v\,T\alpha_p^2}{c_p} \quad . \tag{2.264}$$

Comparing eqns. 2.258 and 2.264, we find

$$(c_p - c_V)\,\kappa_T = (\kappa_T - \kappa_S)\,c_p = v\,T\alpha_p^2 \quad . \tag{2.265}$$

This result entails

$$\frac{c_p}{c_V} = \frac{\kappa_T}{\kappa_S} \quad . \tag{2.266}$$

The corresponding result for magnetic systems is[18]

$$(c_H - c_M)\,\chi_T = (\chi_T - \chi_S)\,c_H = T\left(\frac{\partial m}{\partial T}\right)_H^2 \quad , \tag{2.267}$$

where $m = M/\nu$ is the magnetization per mole of substance, and

$$\text{isothermal susceptibility:} \quad \chi_T = \left(\frac{\partial m}{\partial H}\right)_T = -\frac{1}{\nu}\frac{\partial^2 G}{\partial H^2} \tag{2.268}$$

$$\text{adiabatic susceptibility:} \quad \chi_S = \left(\frac{\partial m}{\partial H}\right)_S = -\frac{1}{\nu}\frac{\partial^2 \mathsf{H}}{\partial H^2} \quad . \tag{2.269}$$

Here the enthalpy and Gibbs free energy are

$$\mathsf{H} = E - HM \qquad d\mathsf{H} = T\,dS - M\,dH \tag{2.270}$$

$$G = E - TS - HM \qquad dG = -S\,dT - M\,dH \quad . \tag{2.271}$$

The previous discussion has assumed an isotropic magnetic system where $\boldsymbol{M}$ and $\boldsymbol{H}$ are collinear, hence $\boldsymbol{H}\cdot\boldsymbol{M} = HM$. In general, we have

$$\chi_T^{\alpha\beta} = \left(\frac{\partial m^\alpha}{\partial H^\beta}\right)_T = -\frac{1}{\nu}\frac{\partial^2 G}{\partial H^\alpha\,\partial H^\beta} \tag{2.272}$$

$$\chi_S^{\alpha\beta} = \left(\frac{\partial m^\alpha}{\partial H^\beta}\right)_S = -\frac{1}{\nu}\frac{\partial^2 H}{\partial H^\alpha\,\partial H^\beta} \quad . \tag{2.273}$$

In this case, the enthalpy and Gibbs free energy are

$$\mathsf{H} = E - \boldsymbol{H}\cdot\boldsymbol{M} \qquad d\mathsf{H} = T\,dS - \boldsymbol{M}\cdot d\boldsymbol{H} \tag{2.274}$$

$$G = E - TS - \boldsymbol{H}\cdot\boldsymbol{M} \qquad dG = -S\,dT - \boldsymbol{M}\cdot d\boldsymbol{H} \quad . \tag{2.275}$$

[18] Recall the subtle font difference in notation: $\mathsf{H}$ is enthalpy while $H$ is magnetic field.

### 2.10.6 Joule effect: free expansion of a gas

Previously we considered the adiabatic free expansion of an ideal gas. We found that $Q = W = 0$ hence $\Delta E = 0$, which means the process is isothermal, since $E = \nu\varepsilon(T)$ is volume-independent. The entropy changes, however, since $S(E,V,N) = Nk_{\rm B}\log(V/N) + \frac{1}{2}fNk_{\rm B}\log(E/N) + Ns_0$. Thus,

$$S_{\rm f} = S_{\rm i} + Nk_{\rm B}\log\left(\frac{V_{\rm f}}{V_{\rm i}}\right) \quad . \tag{2.276}$$

What happens if the gas is nonideal?

We integrate along a fictitious thermodynamic path connecting initial and final states, where $dE = 0$ along the path. We have

$$0 = dE = \left(\frac{\partial E}{\partial V}\right)_T dV + \left(\frac{\partial E}{\partial T}\right)_V dT \tag{2.277}$$

hence

$$\left(\frac{\partial T}{\partial V}\right)_E = -\frac{(\partial E/\partial V)_T}{(\partial E/\partial T)_V} = -\frac{1}{C_V}\left(\frac{\partial E}{\partial V}\right)_T \quad . \tag{2.278}$$

We also have

$$\left(\frac{\partial E}{\partial V}\right)_T = T\left(\frac{\partial S}{\partial V}\right)_T - p = T\left(\frac{\partial p}{\partial T}\right)_V - p \quad . \tag{2.279}$$

Thus,

$$\left(\frac{\partial T}{\partial V}\right)_E = \frac{1}{C_V}\left[p - T\left(\frac{\partial p}{\partial T}\right)_V\right] \quad . \tag{2.280}$$

Note that the term on the RHS vanishes for any system obeying the ideal gas law. For a nonideal gas,

$$\Delta T = \int\limits_{V_{\rm i}}^{V_{\rm f}} dV \left(\frac{\partial T}{\partial V}\right)_E \quad , \tag{2.281}$$

which is in general nonzero.

Now consider a van der Waals gas, for which

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT \quad .$$

We then have

$$p - T\left(\frac{\partial p}{\partial T}\right)_V = -\frac{a}{v^2} = -\frac{a\nu^2}{V^2} \quad . \tag{2.282}$$

In §2.10.3 we concluded that $C_V = \frac{1}{2}f\nu R$ for the van der Waals gas, hence

$$\Delta T = -\frac{2a\nu}{fR}\int\limits_{V_{\rm i}}^{V_{\rm f}} \frac{dV}{V^2} = \frac{2a}{fR}\left(\frac{1}{v_{\rm f}} - \frac{1}{v_{\rm i}}\right) \quad . \tag{2.283}$$

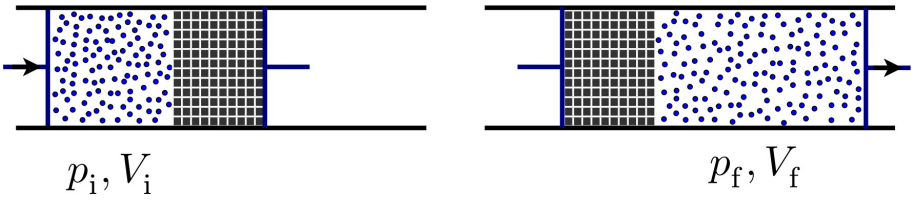



Figure 2.18: In a throttle, a gas is pushed through a porous plug separating regions of different pressure. The change in energy is the work done, hence enthalpy is conserved during the throttling process.

Thus, if $V_\mathrm{f} > V_\mathrm{i}$, we have $T_\mathrm{f} < T_\mathrm{i}$ and the gas cools upon expansion.

Consider $O_2$ gas with an initial specific volume of $v_\mathrm{i} = 22.4\,\mathrm{L/mol}$, which is the STP value for an ideal gas, freely expanding to a volume $v_\mathrm{f} = \infty$ for maximum cooling. According to table 2.3, $a = 1.378\,\mathrm{L}^2 \cdot \mathrm{bar/mol}^2$, and we have $\Delta T = -2a/fRv_i = -0.296\,\mathrm{K}$, which is a pitifully small amount of cooling. Adiabatic free expansion is a very inefficient way to cool a gas.

### 2.10.7 Throttling: the Joule-Thomson effect

In a throttle, depicted in fig. 2.18, a gas is forced through a porous plug which separates regions of different pressures. According to the figure, the work done on a given element of gas is

$$W = \int_0^{V_\mathrm{f}} dV\, p_\mathrm{f} \; - \int_0^{V_\mathrm{i}} dV\, p_\mathrm{i} = p_\mathrm{f} V_\mathrm{f} - p_\mathrm{i} V_\mathrm{i} \quad . \tag{2.284}$$

Now we assume that the system is thermally isolated so that the gas exchanges no heat with its environment, nor with the plug. Then $Q = 0$ so $\Delta E = -W$, and

$$\begin{aligned} E_\mathrm{i} + p_\mathrm{i} V_\mathrm{i} &= E_\mathrm{f} + p_\mathrm{f} V_\mathrm{f} \\ \mathsf{H}_\mathrm{i} &= \mathsf{H}_\mathrm{f} \quad , \end{aligned} \tag{2.285}$$

where $\mathsf{H}$ is enthalpy. Thus, the throttling process is *isenthalpic*. We can therefore study it by defining a fictitious thermodynamic path along which $d\mathsf{H} = 0$. Then, choosing $T$ and $p$ as state variables,

$$0 = d\mathsf{H} = \left(\frac{\partial \mathsf{H}}{\partial T}\right)_p dT \; + \; \left(\frac{\partial \mathsf{H}}{\partial p}\right)_T dp \tag{2.286}$$

hence

$$\left(\frac{\partial T}{\partial p}\right)_\mathsf{H} = -\frac{(\partial \mathsf{H}/\partial p)_T}{(\partial \mathsf{H}/\partial T)_p} \quad . \tag{2.287}$$

The numerator on the RHS is computed by writing $d\mathsf{H} = T\,dS + V\,dp$ and then dividing by $dp$, to obtain

$$\left(\frac{\partial \mathsf{H}}{\partial p}\right)_T = V + T\left(\frac{\partial S}{\partial p}\right)_T = V - T\left(\frac{\partial V}{\partial T}\right)_p \quad . \tag{2.288}$$

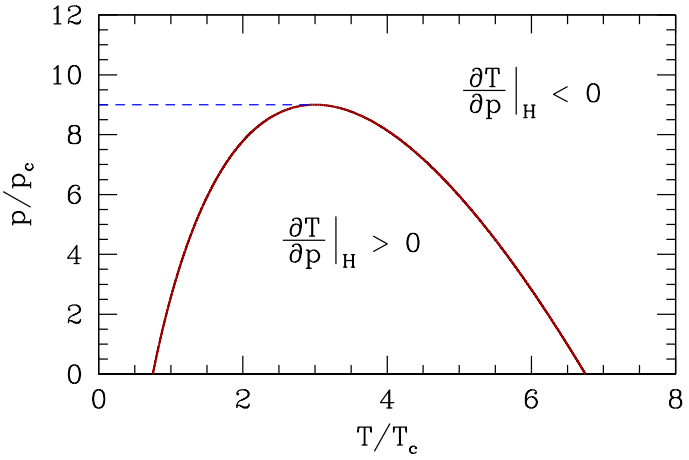



Figure 2.19: Inversion temperature $T^*(p)$ for the van der Waals gas. Pressure and temperature are given in terms of $p_c = a/27b^2$ and $T_c = 8a/27bR$, respectively.

The denominator is

$$\left(\frac{\partial H}{\partial T}\right)_p = \left(\frac{\partial H}{\partial S}\right)_p \left(\frac{\partial S}{\partial T}\right)_p = T\left(\frac{\partial S}{\partial T}\right)_p = C_p \quad . \tag{2.289}$$

Thus,

$$\left(\frac{\partial T}{\partial p}\right)_H = \frac{1}{c_p}\left[T\left(\frac{\partial v}{\partial T}\right)_p - v\right] = \frac{v}{c_p}(T\alpha_p - 1) \quad , \tag{2.290}$$

where $\alpha_p = V^{-1}(\partial V/\partial T)_p$ is the volume expansion coefficient at constant pressure.

From the van der Waals equation of state, we obtain, from eqn. 2.245,

$$T\alpha_p = \frac{T}{v}\left(\frac{\partial v}{\partial T}\right)_p = \frac{RT/v}{p - \frac{a}{v^2} + \frac{2ab}{v^3}} = \frac{v-b}{v - \frac{2a}{RT}\left(\frac{v-b}{v}\right)^2} \quad . \tag{2.291}$$

Assuming $v \gg a/RT$ and $v \gg b$, we have

$$\left(\frac{\partial T}{\partial p}\right)_H = \frac{1}{c_p}\left(\frac{2a}{RT} - b\right) \quad . \tag{2.292}$$

Thus, for $T > T^* = 2a/Rb$, we have $(\partial T/\partial p)_H < 0$ and the gas heats up upon an isenthalpic pressure decrease. For $T < T^*$, the gas cools under such conditions.

In fact, there are two inversion temperatures $T^*_{1,2}$ for the van der Waals gas. To see this, we set $T\alpha_p = 1$, which is the criterion for inversion. From eqn. 2.291 it is easy to derive

$$\frac{b}{v} = 1 - \sqrt{\frac{bRT}{2a}} \quad . \tag{2.293}$$

We insert this into the van der Waals equation of state to derive a relationship $T = T^*(p)$ at which $T\alpha_p = 1$ holds. After a little work, we find

$$p = -\frac{3RT}{2b} + \sqrt{\frac{8aRT}{b^3}} - \frac{a}{b^2} \quad . \tag{2.294}$$

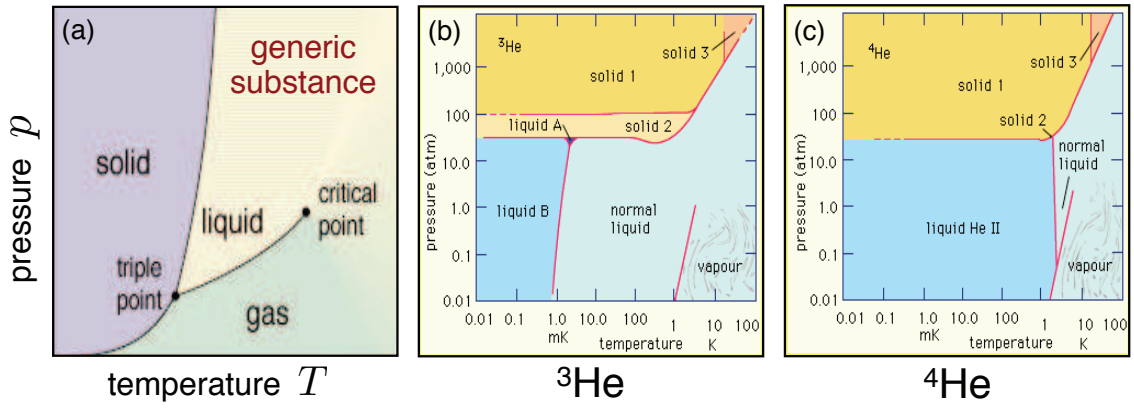



Figure 2.20: (a) Typical thermodynamic phase diagram of a single component $p$-$v$-$T$ system, showing triple point (three phase coexistence) and critical point. (*Credit: Univ. of Helsinki.*) Also shown: phase diagrams for $^3$He (b) and $^4$He (c). What a difference a neutron makes! (*Credit: Britannica.*)

This is a quadratic equation for $\sqrt{T}$, the solution of which is

$$T^*(p) = \frac{2a}{9\,bR}\left(2 \pm \sqrt{1 - \frac{3b^2 p}{a}}\right)^2 \quad . \tag{2.295}$$

In fig. 2.19 we plot pressure *versus* temperature in scaled units, showing the curve along which the derivative $(\partial T/\partial p)_H = 0$. The volume, pressure, and temperature scales defined are

$$v_{\rm c} = 3b \qquad , \qquad p_{\rm c} = \frac{a}{27\,b^2} \qquad , \qquad T_{\rm c} = \frac{8a}{27\,bR} \quad . \tag{2.296}$$

Values for $p_{\rm c}$, $T_{\rm c}$, and $v_{\rm c}$ are provided in table 2.3. If we define $v = v/v_{\rm c}$, $p = p/p_{\rm c}$, and $T = T/T_{\rm c}$, then the van der Waals equation of state may be written in dimensionless form:

$$\left(p + \frac{3}{v^2}\right)(3v - 1) = 8T \quad . \tag{2.297}$$

In terms of the scaled parameters, the equation for the inversion curve $(\partial T/\partial p)_H = 0$ becomes

$$p = 9 - 36\left(1 - \sqrt{\tfrac{1}{3}T}\,\right)^2 \qquad \Longleftrightarrow \qquad T = 3\Big(1 \pm \tfrac{1}{2}\sqrt{1 - \tfrac{1}{9}\,p}\,\Big)^2 \quad . \tag{2.298}$$

Thus, there is no inversion for $p > 9\,p_{\rm c}$. We are usually interested in the upper inversion temperature, $T_2^*$, corresponding to the upper sign in eqn. 2.295. The maximum inversion temperature occurs for $p = 0$, where $T^*_{\rm max} = 2a/bR = 27T_{\rm c}/4$. For $H_2$, from the data in table 2.3, we find $T^*_{\rm max}({\rm H_2}) = 224\,{\rm K}$, which is within 10% of the experimentally measured value of $205\,{\rm K}$.

What happens when $H_2$ gas leaks from a container with $T > T_2^*$? Since $(\partial T/\partial p)_H < 0$ and $\Delta p < 0$, we have $\Delta T > 0$. The gas warms up, and the heat facilitates the reaction $2\,{\rm H_2} + {\rm O_2} \longrightarrow 2\,{\rm H_2O}$, which releases energy, and we have a nice explosion.

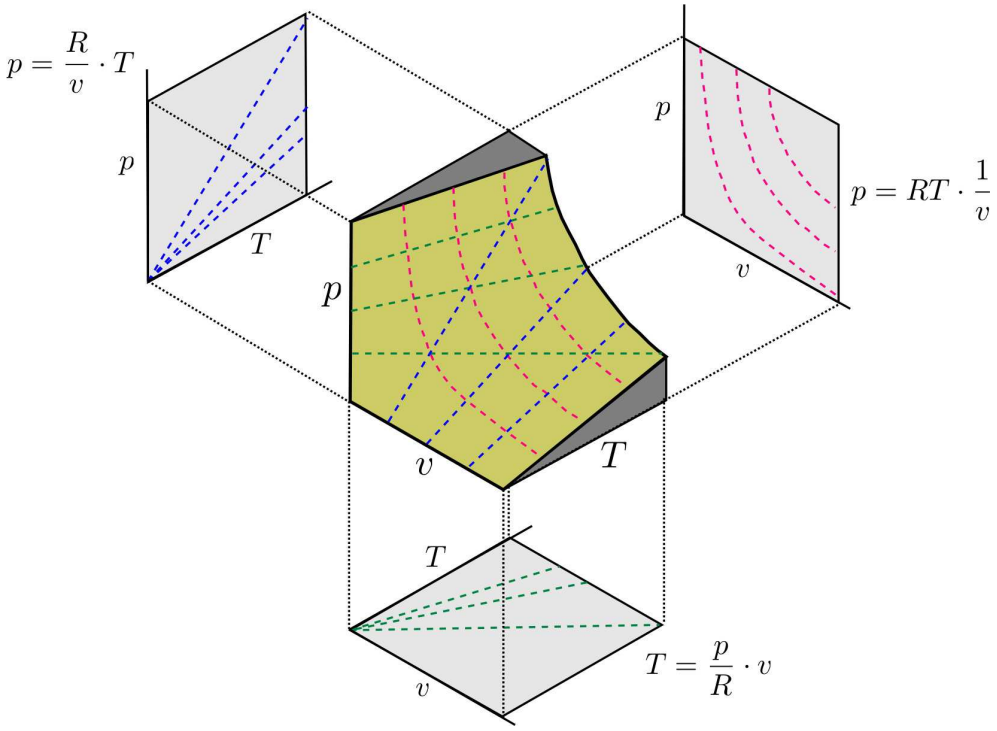



Figure 2.21: The surface $p(v,T) = RT/v$ corresponding to the ideal gas equation of state, and its projections onto the $(p,T)$, $(p,v)$, and $(T,v)$ planes.

## 2.11 Phase Transitions and Phase Equilibria

A typical phase diagram of a $p$-$v$-$T$ system is shown in fig. 2.20(a). The solid lines delineate boundaries between distinct thermodynamic phases. These lines are called *coexistence curves*. Along these curves, we can have coexistence of two phases, and at least one of the thermodynamic potentials is singular. The order of the singularity is often taken as a classification of the phase transition. That is, if one or more of the thermodynamic potentials $E$, $F$, $G$, and $H$ have discontinuous or divergent derivatives of $m^{\rm th}$ and higher, the transition between the respective phases is said to be $m^{\rm th}$ *order*. Modern theories of phase transitions generally recognize only two possibilities: *first order transitions*, where the *order parameter* changes *discontinuously* through the transition, and *continuous transitions* (often called *second order transitions*), where the order parameter vanishes continuously at the boundary from ordered to disordered phases[19]. We'll discuss order parameters when we get to chapter 7.

For a more interesting phase diagram, see fig. 2.20(b,c) for $^3$He and $^4$He. The only difference between these two atoms is that the former has one fewer neutron: (2p + 1n + 2e) in $^3$He *versus* (2p + 2n + 2e) in $^4$He. As we shall learn when we study quantum statistics, this extra neutron makes all the difference, because $^3$He is a fermion while $^4$He is a boson.

[19] Some exotic phase transitions in quantum matter, which do not quite fit the usual classification schemes, have recently been proposed.

### 2.11.1 $p$-$v$-$T$ surfaces

The equation of state for a single component system may be written as

$$f(p, v, T) = 0 \quad . \tag{2.299}$$

This may in principle be inverted to yield $p = p(v, T)$ or $v = v(T, p)$ or $T = T(p, v)$. The single constraint $f(p, v, T)$ on the three state variables defines a *surface* in $\{p, v, T\}$ space. An example of such a surface is shown in fig. 2.21, for the ideal gas.

Real $p$-$v$-$T$ surfaces are much richer than that for the ideal gas, because real systems undergo *phase transitions* in which thermodynamic properties are singular or discontinuous along certain curves on the $p$-$v$-$T$ surface. An example is shown in fig. 2.22. The high temperature isotherms resemble those of the ideal gas, but as one cools below the *critical temperature* $T_{\rm c}$, the isotherms become singular. Precisely at $T = T_{\rm c}$, the isotherm $p = p(v, T_{\rm c})$ becomes perfectly horizontal at $v = v_{\rm c}$, which is the *critical molar volume*. This means that the isothermal compressibility, $\kappa_T = -v^{-1}(\partial v/\partial p)_T$ *diverges* at $T = T_{\rm c}$. Below $T_{\rm c}$, the isotherms have a flat portion, as shown in fig. 2.25, corresponding to a *two-phase region* where liquid and vapor coexist. In the $(p, T)$ plane, sketched for $H_2O$ in fig. 2.5 and shown for $CO_2$ in fig. 2.24, this liquid-vapor phase coexistence occurs along a curve, called the vaporization (or boiling) curve. The density changes discontinuously across this curve; for $H_2O$, the liquid is approximately 1000 times denser than the vapor at atmospheric pressure. The density discontinuity vanishes at the critical point. Note that one can continuously transform between liquid and vapor phases, without encountering any phase transitions, by going around the critical point and avoiding the two-phase region.

In addition to liquid-vapor coexistence, solid-liquid and solid-vapor coexistence also occur, as shown in fig. 2.22. The *triple point* $(\mathrm{T_t}\,,\ p_{\rm t})$ lies at the confluence of these three coexistence regions. For $H_2O$, the location of the triple point and critical point are given by

$$\mathrm{T_t} = 273.16\,\mathrm{K} \qquad\qquad T_{\rm c} = 647\,\mathrm{K}$$
$$\mathrm{p_t} = 611.7\,\mathrm{Pa} = 6.037 \times 10^{-3}\,\mathrm{atm} \qquad\qquad \mathrm{p_c} = 22.06\,\mathrm{MPa} = 217.7\,\mathrm{atm}$$

### 2.11.2 The Clausius-Clapeyron relation

Recall that the homogeneity of $E(S, V, N)$ guaranteed $E = TS - pV + \mu N$, from Euler's theorem. It also guarantees a relation between the intensive variables $T$, $p$, and $\mu$, according to eqn. 2.112. Let us define $g \equiv G/\nu = N_{\rm A}\mu$, the Gibbs free energy per mole. Then

$$dg = -s\,dT + v\,dp \quad , \tag{2.300}$$

where $s = S/\nu$ and $v = V/\nu$ are the molar entropy and molar volume, respectively. Along a coexistence curve between phase #1 and phase #2, we must have $g_1 = g_2$, since the phases are free to exchange energy and particle number, *i.e.* they are in thermal and chemical equilibrium. This means

$$dg_1 = -s_1\,dT + v_1\,dp = -s_2\,dT + v_2\,dp = dg_2 \quad . \tag{2.301}$$

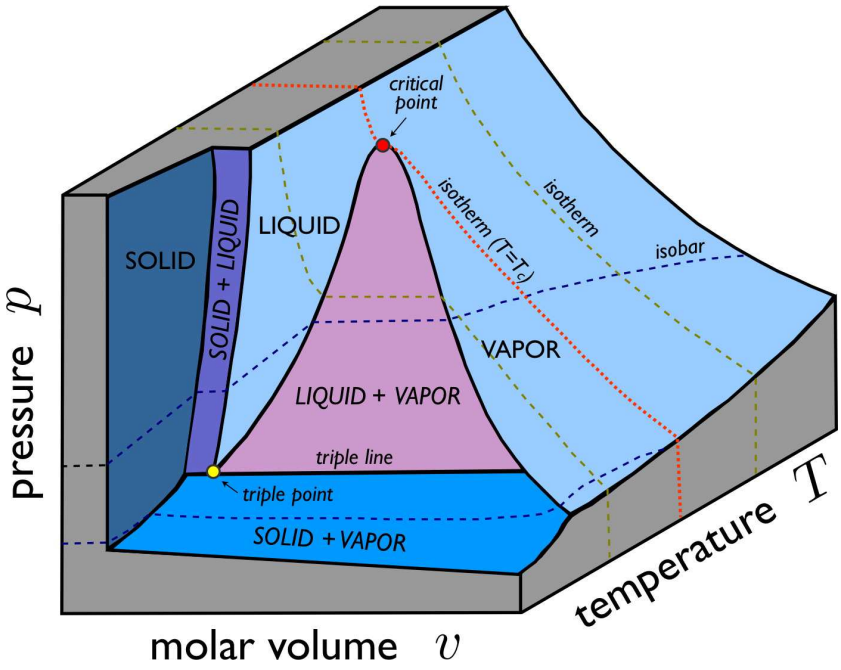



Figure 2.22: A $p$-$v$-$T$ surface for a substance which contracts upon freezing. The red dot is the *critical point* and the red dashed line is the *critical isotherm*. The yellow dot is the *triple point* at which there is three phase coexistence of solid, liquid, and vapor.

Therefore, along the coexistence curve we must have

$$\left(\frac{dp}{dT}\right)_{\rm coex} = \frac{s_2 - s_1}{v_2 - v_1} = \frac{\ell}{T\,\Delta v} \quad , \tag{2.302}$$

where $\ell \equiv T\,\Delta s = T\,(s_2 - s_1)$ is the *molar latent heat of transition*. This is known as the *Clapeyron relation*. Heat energy $\ell$ per mole must be supplied in order to change from phase #1 to phase #2, even without changing $p$ or $T$. If $\ell$ is the latent heat per mole, then we write $\tilde{\ell}$ as the latent heat per gram: $\tilde{\ell} = \ell/M$, where $M$ is the molar mass.

Along the liquid-gas coexistence curve, we typically have $v_{\rm gas} \gg v_{\rm liquid}$, and assuming the vapor is ideal, we may write $\Delta v \approx v_{\rm gas} \approx RT/p$, resulting in the *Clausius-Clapeyron relation*,

$$\left(\frac{dp}{dT}\right)_{\rm liq-gas} = \frac{\ell}{T\,\Delta v} \approx \frac{p\,\ell}{RT^2} \quad . \tag{2.303}$$

If $\ell$ remains constant throughout a section of the liquid-gas coexistence curve, we may integrate the above equation to get

$$\frac{dp}{p} = \frac{\ell}{R}\,\frac{dT}{T^2} \qquad \Rightarrow \qquad p(T) = p(T_0)\,e^{\ell/RT_0}\,e^{-\ell/RT} \quad . \tag{2.304}$$

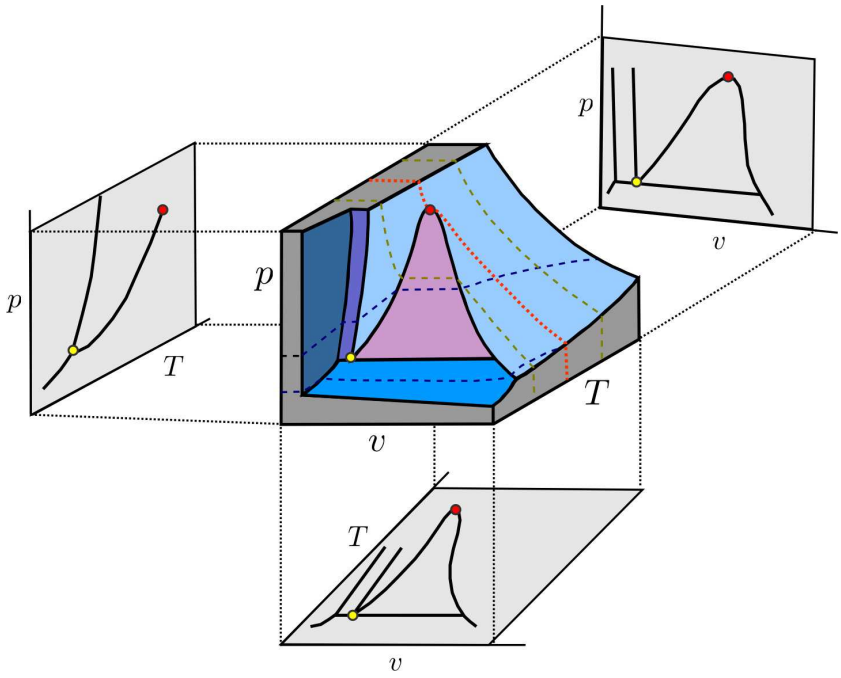



Figure 2.23: Equation of state for a substance which expands upon freezing, projected to the $(v, T)$ and $(v, p)$ and $(T, p)$ planes.

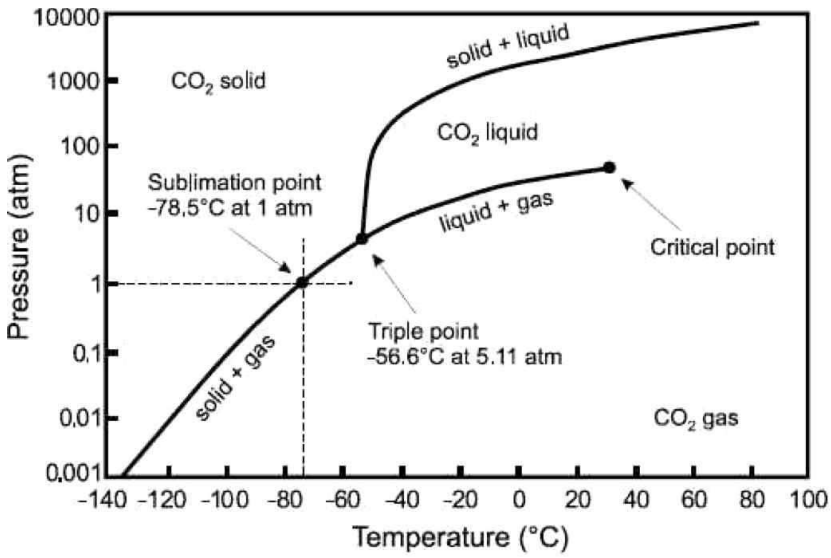



Figure 2.24: Phase diagram for $CO_2$ in the $(p, T)$ plane. (*Source: www.scifun.org.*)

### 2.11.3 Liquid-solid line in $H_2O$

Life on planet earth owes much of its existence to a peculiar property of water: the solid is less dense than the liquid along the coexistence curve[20]. For example at $T = 273.1\,\mathrm{K}$ and $p = 1\,\mathrm{atm}$,

$$\tilde{v}_{\rm water} = 1.00013\,\mathrm{cm}^3/\mathrm{g} \quad , \quad \tilde{v}_{\rm ice} = 1.0907\,\mathrm{cm}^3/\mathrm{g} \quad . \tag{2.305}$$

The latent heat of the transition is $\tilde{\ell} = 333\,\mathrm{J/g} = 79.5\,\mathrm{cal/g}$. Thus,

$$\begin{aligned}\left(\frac{dp}{dT}\right)_{\rm liq-sol} &= \frac{\tilde{\ell}}{T\,\Delta\tilde{v}} = \frac{333\,\mathrm{J/g}}{(273.1\,\mathrm{K})\,(-9.05\times10^{-2}\,\mathrm{cm}^3/\mathrm{g})} \\ &= -1.35\times10^8\,\frac{\mathrm{dyn}}{\mathrm{cm}^2\,\mathrm{K}} = -134\,\frac{\mathrm{atm}}{^\circ\mathrm{C}} \quad .\end{aligned} \tag{2.306}$$

The negative slope of the melting curve is invoked to explain the movement of glaciers: as glaciers slide down a rocky slope, they generate enormous pressure at obstacles[21] Due to this pressure, the story goes, the melting temperature decreases, and the glacier melts around the obstacle, so it can flow past it, after which it refreezes. But it is not the case that the bottom of the glacier melts under the pressure, for consider a glacier of height $h = 1\,\mathrm{km}$. The pressure at the bottom is $p \sim gh/\tilde{v} \sim 10^7\,\mathrm{Pa}$, which is only about 100 atmospheres. Such a pressure can produce only a small shift in the melting temperature of about $\Delta T_{\rm melt} = -0.75^\circ\,\mathrm{C}$.

Does the Clausius-Clapeyron relation explain how we can skate on ice? When my daughter was seven years old, she had a mass of about $M = 20\,\mathrm{kg}$. Her ice skates had blades of width about $5\,\mathrm{mm}$ and length about $10\,\mathrm{cm}$. Thus, even on one foot, she imparted an additional pressure of only

$$\Delta p = \frac{Mg}{A} \approx \frac{20\,\mathrm{kg}\times9.8\,\mathrm{m/s}^2}{(5\times10^{-3}\,\mathrm{m})\times(10^{-1}\,\mathrm{m})} = 3.9\times10^5\,\mathrm{Pa} = 3.9\,\mathrm{atm} \quad . \tag{2.307}$$

The corresponding change in the melting temperature is thus minuscule: $\Delta T_{\rm melt} \approx -0.03^\circ\,\mathrm{C}$.

So why could my daughter skate so nicely? The answer isn't so clear![22] There seem to be two relevant issues in play. First, friction generates heat which can locally melt the surface of the ice. Second, the surface of ice, and of many solids, is naturally slippery. Indeed, this is the case for ice even if one is standing still, generating no frictional forces. Why is this so? It turns out that the Gibbs free energy of the ice-air interface is larger than the sum of free energies of ice-water and water-air interfaces. That is to say, ice, as well as many simple solids, prefers to have a thin layer of liquid on its surface, even at temperatures well below its bulk melting point. If the intermolecular interactions are not short-ranged[23], theory predicts a surface melt thickness $d \propto (T_{\rm m} - T)^{-1/3}$. In fig. 2.26 we show measurements by Gilpin (1980) of the surface melt on ice, down to about $-50^\circ\,\mathrm{C}$. Near $0^\circ\,\mathrm{C}$ the melt layer thickness is about $40\,\mathrm{nm}$,

[20] Were the solid more dense, ponds and lakes would freeze from the bottom, with unfortunate consequences for life on earth.

[21] The melting curve has a negative slope at relatively low pressures, where the solid has the so-called Ih hexagonal crystal structure. At pressures above about 2500 atmospheres, the crystal structure changes, and the slope of the melting curve becomes positive.

[22] For a recent discussion, see R. Rosenberg, *Physics Today* **58**, 50 (2005).

[23] For example, they could be of the van der Waals form, due to virtual dipole fluctuations, with an attractive $1/r^6$ tail.

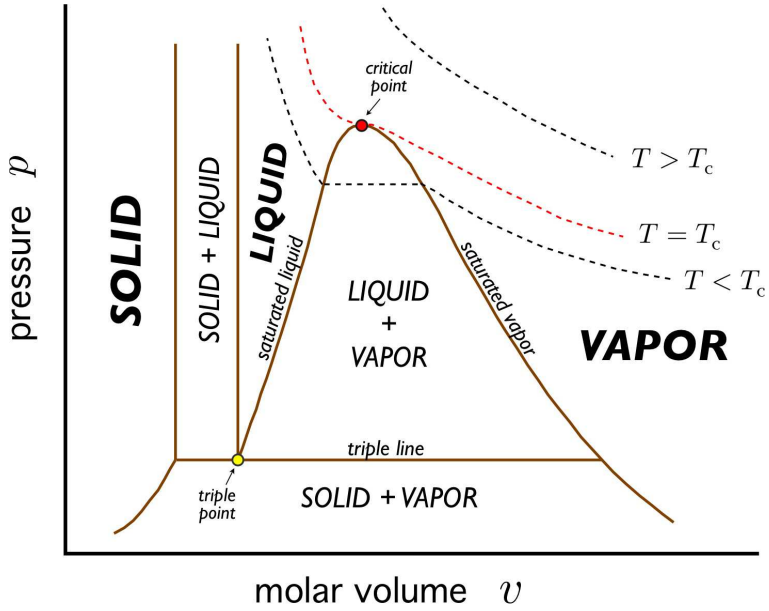



Figure 2.25: Projection of the $p$-$v$-$T$ surface of fig. 2.22 onto the $(v, p)$ plane.

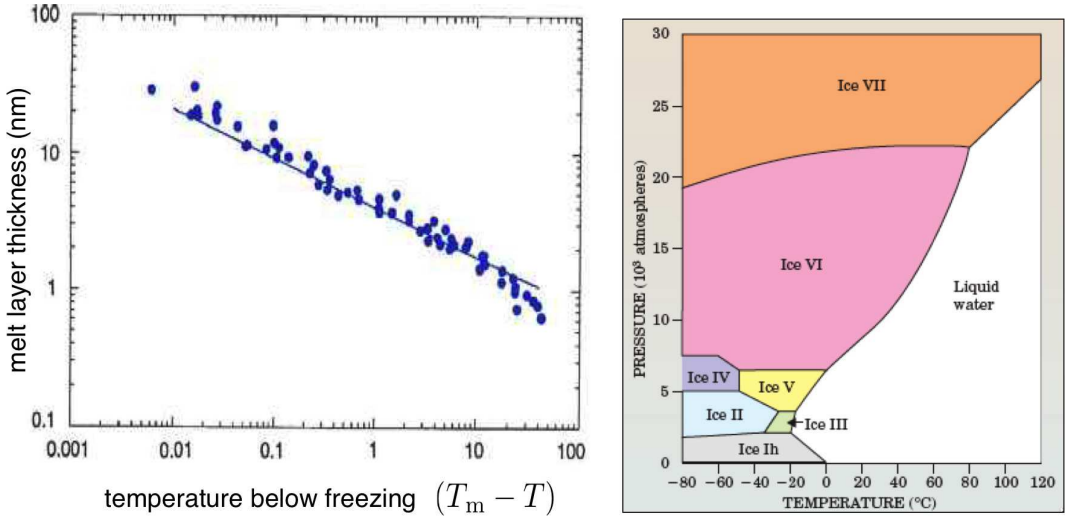



Figure 2.26: Left panel: data from R. R. Gilpin, *J. Colloid Interface Sci.* **77**, 435 (1980) showing measured thickness of the surface melt on ice at temperatures below 0°C. The straight line has slope $-\frac{1}{3}$, as predicted by theory. Right panel: phase diagram of $H_2O$, showing various high pressure solid phases. (*Source* : *Physics Today*, December 2005).

but this decreases to $\sim 1\,\text{nm}$ at $T = -35^\circ\,\text{C}$. At very low temperatures, skates stick rather than glide. Of course, the skate material is also important, since that will affect the energetics of the second interface. The 19th century novel, *Hans Brinker, or The Silver Skates* by Mary Mapes Dodge tells the story of the poor but stereotypically decent and hardworking Dutch boy Hans Brinker, who dreams of winning an upcoming ice skating race, along with the top prize: a pair of silver skates. All he has are some lousy

wooden skates, which won't do him any good in the race. He has money saved to buy steel skates, but of course his father desperately needs an operation because – I am not making this up – he fell off a dike and lost his mind. The family has no other way to pay for the doctor. What a story! At this point, I imagine the suspense must be too much for you to bear, but this isn't an American Literature class, so you can use Google to find out what happens (or rent the 1958 movie, directed by Sidney Lumet). My point here is that Hans' crappy wooden skates can't compare to the metal ones, even though the surface melt between the ice and the air is the same. The skate blade material also makes a difference, both for the interface energy and, perhaps more importantly, for the generation of friction as well.

### 2.11.4 Slow melting of ice : a quasistatic but irreversible process

Suppose we have an ice cube initially at temperature $T_0 < \Theta \equiv 273.15\,\mathrm{K}$ (*i.e.* $\Theta = 0^\circ\,\mathrm{C}$) and we toss it into a pond of water. We regard the pond as a heat bath at some temperature $T_1 > \Theta$. Let the mass of the ice be $M$. How much heat $Q$ is absorbed by the ice in order to raise its temperature to $T_1$? Clearly

$$Q = M\tilde{c}_{\rm S}(\Theta - T_0) + M\tilde{\ell} + M\tilde{c}_{\rm L}(T_1 - \Theta) \quad , \tag{2.308}$$

where $\tilde{c}_{\rm S}$ and $\tilde{c}_{\rm L}$ are the specific heats of ice (solid) and water (liquid), respectively[24], and $\tilde{\ell}$ is the latent heat of melting per unit mass. The pond must give up this much heat to the ice, hence the entropy of the pond, discounting the new water which will come from the melted ice, must decrease:

$$\Delta S_{\rm pond} = -\frac{Q}{T_1} \quad . \tag{2.309}$$

Now we ask what is the entropy change of the $H_2O$ in the ice. We have

$$\begin{aligned} \Delta S_{\rm ice} &= \int \frac{đQ}{T} = \int\limits_{T_0}^{\Theta} dT\,\frac{M\tilde{c}_{\rm S}}{T} + \frac{M\tilde{\ell}}{\Theta} + \int\limits_{\Theta}^{T_1} dT\,\frac{M\tilde{c}_{\rm L}}{T} \\ &= M\tilde{c}_{\rm S}\log\left(\frac{\Theta}{T_0}\right) + \frac{M\tilde{\ell}}{\Theta} + M\tilde{c}_{\rm L}\log\left(\frac{T_1}{\Theta}\right) \quad . \end{aligned} \tag{2.310}$$

The total entropy change of the system is then

$$\begin{aligned} \Delta S_{\rm total} &= \Delta S_{\rm pond} + \Delta S_{\rm ice} \\ &= M\tilde{c}_{\rm S}\log\left(\frac{\Theta}{T_0}\right) - M\tilde{c}_{\rm S}\left(\frac{\Theta - T_0}{T_1}\right) + M\tilde{\ell}\left(\frac{1}{\Theta} - \frac{1}{T_1}\right) + M\tilde{c}_{\rm L}\log\left(\frac{T_1}{\Theta}\right) - M\tilde{c}_{\rm L}\left(\frac{T_1 - \Theta}{T_1}\right) \quad . \end{aligned} \tag{2.311}$$

Now since $T_0 < \Theta < T_1$, we have

$$M\tilde{c}_{\rm S}\left(\frac{\Theta - T_0}{T_1}\right) < M\tilde{c}_{\rm S}\left(\frac{\Theta - T_0}{\Theta}\right) \quad . \tag{2.312}$$

[24] We assume $\tilde{c}_{\rm S}(T)$ and $\tilde{c}_{\rm L}(T)$ have no appreciable temperature dependence, and we regard them both as constants.

Therefore,

$$\Delta S > M\tilde{\ell}\left(\frac{1}{\Theta}-\frac{1}{T_1}\right)+M\tilde{c}_{\rm S}\,f\big(T_0/\Theta\big)+M\tilde{c}_{\rm L}\,f\big(\Theta/T_1\big) \quad , \tag{2.313}$$

where $f(x)=x-1-\log x$. Clearly $f'(x)=1-x^{-1}$ is negative on the interval $(0,1)$, which means that the maximum of $f(x)$ occurs at $x=0$ and the minimum at $x=1$. But $f(0)=\infty$ and $f(1)=0$, which means that $f(x)\geq 0$ for $x\in[0,1]$. Since $T_0<\Theta<T_1$, we conclude $\Delta S_{\rm total}>0$.

Note that by 'slow' we mean that the process is quasistatic and continuous, so that we may use results from equilibrium thermodynamics in our computations. We know that as a practical matter the process is not quasistatic – an ice cube tossed into a pool of water will develop internal temperature variations and will get progressively smaller as it melts from its instantaneous surface. In our idealized example we assume that the temperature is always uniform throughout the ice, which finally begins to melt (*i.e.* change phase) when its temperature throughout reaches $\Theta$. When $T_0=\Theta^-$ and $T_1=\Theta^+$ are infinitesimally below and above the melting temperature, the entropy change in eqn. 2.312 is a positive infinitesimal, and when $T_0=\Theta=T_1$ we have $\Delta S=0$ and the process is reversible.

### 2.11.5 Generalized Gibbs phase rule

Equilibrium between two phases means that $p$, $T$, and $\mu(p,T)$ are identical. From

$$\mu_1(p,T)=\mu_2(p,T) \quad , \tag{2.314}$$

we derive an equation for the slope of the coexistence curve, the Clausius-Clapeyron relation. Note that we have one equation in two unknowns $(T,p)$, so the solution set is a curve. For three phase coexistence, we have

$$\mu_1(p,T)=\mu_2(p,T)=\mu_3(p,T) \quad , \tag{2.315}$$

which gives us two equations in two unknowns. The solution is then a point (or a set of points). A *critical point* also is a solution of two simultaneous equations:

$$\text{critical point} \quad \Longrightarrow \quad v_1(p,T)=v_2(p,T) \quad , \quad \mu_1(p,T)=\mu_2(p,T) \quad . \tag{2.316}$$

Recall $v=N_{\rm A}\big(\frac{\partial\mu}{\partial p}\big)_T$. Note that there can be no four phase coexistence for a simple $p$-$v$-$T$ system.

Now for the general result. Suppose we have $\mathsf{c}$ species, with particle numbers $N_a$, where $a=1,\ldots,\mathsf{c}$. It is useful to briefly recapitulate the derivation of the Gibbs-Duhem relation. The internal energy $E(S,V,N_1,\ldots,N_{\mathsf{c}})$ is a homogeneous function of degree one:

$$E(\lambda S,\lambda V,\lambda N_1,\ldots,\lambda N_{\mathsf{c}})=\lambda E(S,V,N_1,\ldots,N_{\mathsf{c}}) \quad . \tag{2.317}$$

From Euler's theorem for homogeneous functions (just differentiate with respect to $\lambda$ and then set $\lambda=1$), we have

$$E=TS-pV+\sum_{a=1}^{\mathsf{c}}\mu_a N_a \quad . \tag{2.318}$$

Taking the differential, and invoking the First Law,

$$dE=T\,dS-p\,dV+\sum_{a=1}^{\mathsf{c}}\mu_a\,dN_a \quad , \tag{2.319}$$

we arrive at the relation

$$S\,dT - V\,dp + \sum_{a=1}^{\mathsf{c}} N_a\,d\mu_a = 0 \quad , \tag{2.320}$$

of which eqn. 2.111 is a generalization to additional internal 'work' variables. This says that the $\mathsf{c}+2$ quantities $(T, p, \mu_1, \ldots, \mu_{\mathsf{c}})$ are not all independent. We can therefore write

$$\mu_{\mathsf{c}} = \mu_{\mathsf{c}}\big(T, p, \mu_1, \ldots, \mu_{\mathsf{c}-1}\big) \quad . \tag{2.321}$$

If there are $\varphi$ different phases, then in each phase $j$, with $j = 1, \ldots, \varphi$, there is a chemical potential $\mu_a^{(j)}$ for each species $a$. We then have

$$\mu_{\mathsf{c}}^{(j)} = \mu_{\mathsf{c}}^{(j)}\Big(T, p, \mu_1^{(j)}, \ldots, \mu_{\mathsf{c}-1}^{(j)}\Big) \quad . \tag{2.322}$$

Here $\mu_a^{(j)}$ is the chemical potential of the $a^{\rm th}$ species in the $j^{\rm th}$ phase. Thus, there are $\varphi$ such equations relating the $2 + \mathsf{c}\,\varphi$ variables $\big(T, p, \{\mu_a^{(j)}\}\big)$, meaning that only $2 + (\mathsf{c}-1)\varphi$ of them may be chosen as independent. This, then, is the dimension of 'thermodynamic space' containing a maximal number of intensive variables:

$$d(\mathsf{c}, \varphi) = 2 + \varphi\,(\mathsf{c}-1) \quad . \tag{2.323}$$

To completely specify the state of our system, we of course introduce a single extensive variable, such as the total volume $V$. Note that the total particle number $N = \sum_{a=1}^{\mathsf{c}} N_a$ may not be conserved in the presence of chemical reactions!

Now suppose we have equilibrium among $\varphi$ phases. We have implicitly assumed thermal and mechanical equilibrium among all the phases, meaning that $p$ and $T$ are constant. Chemical equilibrium applies on a species-by-species basis. This means

$$\mu_a^{(j)} = \mu_a^{(j')} \tag{2.324}$$

where $j, j' \in \{1, \ldots, \varphi\}$. This gives $\mathsf{c}(\varphi-1)$ independent equations[25]. Thus, we can have *phase equilibrium* among the $\varphi$ phases of $\mathsf{c}$ species over a region of dimension

$$d_{\rm PE}(\mathsf{c}, \varphi) = 2 + \varphi\,(\mathsf{c}-1) - \mathsf{c}\,(\varphi-1) = 2 + \mathsf{c} - \varphi \quad . \tag{2.325}$$

Since $d_{\rm PE} \geq 0$, we must have $\varphi \leq \mathsf{c} + 2$. Thus, with two species ($\mathsf{c} = 2$), we could have at most four phase coexistence.

If the various species can undergo $\mathsf{r}$ distinct chemical reactions of the form

$$\zeta_1^{(l)}\,\mathrm{A}_1 + \zeta_2^{(l)}\,\mathrm{A}_2 + \cdots + \zeta_{\mathsf{c}}^{(l)}\,\mathrm{A}_{\mathsf{c}} = 0 \quad , \tag{2.326}$$

where $\mathrm{A}_a$ is the chemical formula for species $a$, and $\zeta_a^{(l)}$ is the stoichiometric coefficient for the $a^{\rm th}$ species in the $l^{\rm th}$ reaction, with $l \in \{1, \ldots, \mathsf{r}\}$, then we have an additional $\mathsf{r}$ constraints of the form

$$\sum_{a=1}^{\mathsf{c}} \zeta_a^{(l)}\,\mu_a^{(j)} = 0 \quad . \tag{2.327}$$

[25] Set $j = 1$ and let $j'$ range over the $\varphi - 1$ values $2, \ldots, \varphi$.

Therefore,

$$d_{\rm PE}(\mathsf{c},\varphi,\mathsf{r}) = 2 + \mathsf{c} - \varphi - \mathsf{r} \quad . \tag{2.328}$$

One might ask what value of $j$ are we to use in eqn. 2.327, or do we in fact have $\varphi$ such equations for each $l$? The answer is that eqn. 2.324 guarantees that the chemical potential of species $a$ is the same in all the phases, hence it doesn't matter what value one chooses for $j$ in eqn. 2.327.

Let us assume that no reactions take place, *i.e.* $\mathsf{r} = 0$, so the total number of particles $\sum_{a=1}^{\mathsf{c}} N_a$ is conserved. Instead of choosing $(T, p, \mu_1, \ldots, \mu_{\mathsf{c}-1}^{(j)})$ as $d_{\rm TD}$ intensive variables, we could have chosen $(T, p, \mu_1, \ldots, x_{\mathsf{c}-1}^{(j)})$, where $x_a = N_a/N$ is the concentration of species $a$.

With $\mathsf{d}$ independent generalized displacements, the dimensions of the thermodynamic space and regions of phase equilibria are

$$d(\mathsf{c},\varphi) = 1 + \mathsf{d} + \varphi\,(\mathsf{c}-1) \qquad , \qquad d_{\rm PE}(\mathsf{c},\varphi,\mathsf{r}) = 1 + \mathsf{d} + \mathsf{c} - \varphi - \mathsf{r} \quad , \tag{2.329}$$

whence $\varphi \le 1 + \mathsf{d} + \mathsf{c} - \mathsf{r}$.

Why do phase diagrams in the $(p, v)$ and $(T, v)$ plane look different than those in the $(p, T)$ plane?[26] For example, fig. 2.23 shows projections of the $p$-$v$-$T$ surface of a typical single component substance onto the $(T, v)$, $(p, v)$, and $(p, T)$ planes. Coexistence takes place along *curves* in the $(p, T)$ plane, but in extended two-dimensional regions in the $(T, v)$ and $(p, v)$ planes. The reason that $p$ and $T$ are special is that temperature, pressure, and chemical potential must be equal throughout an equilibrium phase if it is truly in thermal, mechanical, and chemical equilibrium. This is not the case for an intensive variable such as specific volume $v = N_{\rm A}V/N$ or chemical concentration $x_a = N_a/N$.

## 2.12 Entropy of Mixing and the Gibbs Paradox

### 2.12.1 Computing the entropy of mixing

Entropy is widely understood as a measure of disorder. Of course, such a definition should be supplemented by a more precise definition of disorder – after all, one man's trash is another man's treasure. To gain some intuition about entropy, let us explore the mixing of a multicomponent ideal gas with $\mathsf{c}$ component species. Let $N = \sum_{a=1}^{c} N_a$ be the total number of particles of all species, and let $x_a = N_a/N$ be the concentration of species $a$. Note that $\sum_{a=1}^{c} x_a = 1$.

For any substance obeying the ideal gas law $pV = Nk_{\rm B}T$, the entropy is

$$S(T, V, N) = Nk_{\rm B}\log(V/N) + N\phi(T) \quad , \tag{2.330}$$

since $(\partial S/\partial V)_{T,N} = (\partial p/\partial T)_{V,N} = Nk_{\rm B}/V$. Note that in eqn. 2.330 we have divided $V$ by $N$ before taking the logarithm. This is essential if the entropy is to be an extensive function (see §2.6.5). If, as we shall discuss below, the entropy is proportional to the number of ways a system can be 'configured', one

[26] The same can be said for multicomponent systems: the phase diagram in the $(T, x)$ plane at constant $p$ looks different than the phase diagram in the $(T, \mu)$ plane at constant $p$.

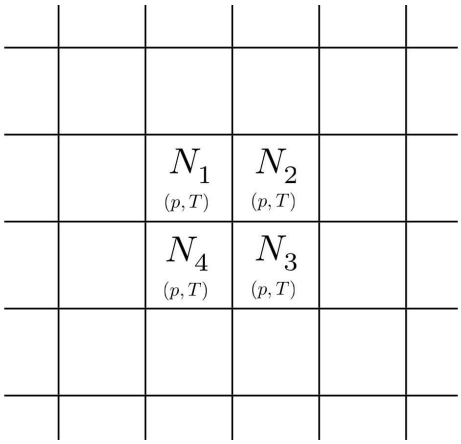



Figure 2.27: A multicomponent system consisting of isolated gases, each at temperature $T$ and pressure $p$. Then system entropy increases when all the walls between the different subsystems are removed.

might think that the entropy of an ideal gas should scale as $\log(V^N) = N \log V$, since each particle can be anywhere in the volume $V$. However, if the particles are *indistinguishable*, then permuting the particle labels does not result in a distinct configuration, and so the configurational entropy is proportional to $\log(V^N/N!) \sim N \log(V/N) + N$. The origin of this indistinguishability factor will become clear when we discuss the quantum mechanical formulation of statistical mechanics. For now, note that such a correction is necessary in order that the entropy be an extensive function.

If we did not include this factor and instead wrote $S^*(T, V, N) = Nk_{\rm B} \log V + N\phi(T)$, then we would find $S^*(T, V, N) - 2S^*(T, \frac{1}{2}V, \frac{1}{2}N) = Nk_{\rm B} \log 2$, *i.e.* the total entropy of two identical systems of particles separated by a barrier will increase if the barrier is removed and they are allowed to mix. This seems absurd, though, because we could just as well regard the barriers as invisible. This is known as the *Gibbs paradox*. The resolution of the Gibbs paradox is to include the indistinguishability correction, which renders $S$ extensive, in which case $S(T, V, N) = 2S(T, \frac{1}{2}V, \frac{1}{2}N)$.

Consider now the situation in fig. 2.27, where we have separated the different components into their own volumes $V_a$. Let the pressure and temperature be the same everywhere, so $pV_a = N_a k_{\rm B} T$. The entropy of the unmixed system is then

$$S_{\rm unmixed} = \sum_a S_a = \sum_a \Big[ N_a \, k_{\rm B} \log(V_a/N_a) + N_a \, \phi_a(T) \Big] \quad . \tag{2.331}$$

Now let us imagine removing all the barriers separating the different gases and letting the particles mix thoroughly. The result is that each component gas occupies the full volume $V$, so the entropy is

$$S_{\rm mixed} = \sum_a S_a = \sum_a \Big[ N_a \, k_{\rm B} \log(V/N_a) + N_a \, \phi_a(T) \Big] \quad . \tag{2.332}$$

Thus, the *entropy of mixing* is

$$\begin{aligned} \Delta S_{\rm mix} &= S_{\rm mixed} - S_{\rm unmixed} \\ &= \sum_a N_a \, k_{\rm B} \, \log(V/V_a) = -Nk_{\rm B} \sum_a f_a \log f_a \quad , \end{aligned} \tag{2.333}$$

where $f_a = N_a/N = V_a/V$ is the fraction of species $a$. Note that $\Delta S_{\rm mix} \geq 0$.

What if all the components were initially identical? It seems absurd that the entropy should increase simply by removing some invisible barriers. This is again the Gibbs paradox. In this case, the resolution of the paradox is to note that the sum in the expression for $S_{\rm mixed}$ is a sum over distinct species. Hence if the particles are all identical, we have $S_{\rm mixed} = Nk_{\rm B}\log(V/N) + N\phi(T) = S_{\rm unmixed}$, hence $\Delta S_{\rm mix} = 0$.

### 2.12.2 Entropy and combinatorics

As we shall learn when we study statistical mechanics, the entropy may be interpreted in terms of the number of ways $W(E,V,N)$ a system at fixed energy and volume can arrange itself. One has

$$S(E,V,N) = k_{\rm B}\log W(E,V,N) \quad . \tag{2.334}$$

Consider a system consisting of $\mathsf{c}$ different species of particles. Initially, each species $a \in \{1,\ldots,\mathsf{c}\}$ is confined to its own region consisting of $B_a$ little boxes, with each little box containing either 0 or 1 of the $a$-particles. Adding up the total number of boxes over all of the $\mathsf{c}$ regions yields $B = \sum_{a=1}^{\mathsf{c}} B_a$ (see fig. 2.28). How many ways $W_a$ are there to configure $N_a$ identical particles among $B_a$ little boxes? Clearly

$$W_a = \binom{B_a}{N_a} = \frac{B_a!}{N_a!\,(B_a - N_a)!} \quad . \tag{2.335}$$

Were the particles distinct, we'd have had $W_a^{\rm distinct} = B_a!/(B_a - N_a)!$, which is $N_a!$ times greater. This is because permuting distinct particles results in a different configuration, and there are $N_a!$ ways to permute $N_a$ particles.

The entropy for species $a$ is then $S_a = k_{\rm B}\log W_a = k_{\rm B}\log\binom{B_a}{N_a}$. We then use Stirling's approximation,

$$\log(K!) = K\log K - K + \tfrac{1}{2}\log K + \tfrac{1}{2}\log(2\pi) + \mathcal{O}(K^{-1}) \quad , \tag{2.336}$$

which is an asymptotic expansion valid for $K \gg 1$. One then finds for $B, N \gg 1$, with $x = N/B \in [0,1]$,

$$\begin{aligned}\log\binom{B}{N} &= \Big(B\log B - B\Big) - \Big(xB\log(xB) - xB\Big) - \Big((1-x)B\log\big((1-x)B\big) - (1-x)B\Big) \\ &= -B\Big[x\log x + (1-x)\log(1-x)\Big] \quad .\end{aligned} \tag{2.337}$$

This is valid up to terms of order $B$ in Stirling's expansion. Since $\log B \ll B$, the next term is small and we are safe to stop here. Summing up the contributions from all the species, we get

$$S_{\rm unmixed} = k_{\rm B}\sum_{a=1}^{\mathsf{c}}\log W_a = -k_{\rm B}\sum_{a=1}^{\mathsf{c}} B_a\Big[x_a\log x_a + (1-x_a)\log(1-x_a)\Big] \quad , \tag{2.338}$$

where $x_a = N_a/B_a$ is the initial dimensionless density of species $a$.

Now let's remove all the partitions between the different species so that each of the particles is free to explore all of the boxes. The total number of ways of placing $N_1$ particles of species $a = 1$ through $N_{\mathsf{c}}$ particles of species $\mathsf{c}$ is

$$W_{\rm mixed} = \frac{B!}{N_0!\,N_1!\cdots N_{\mathsf{c}}!} \quad , \tag{2.339}$$

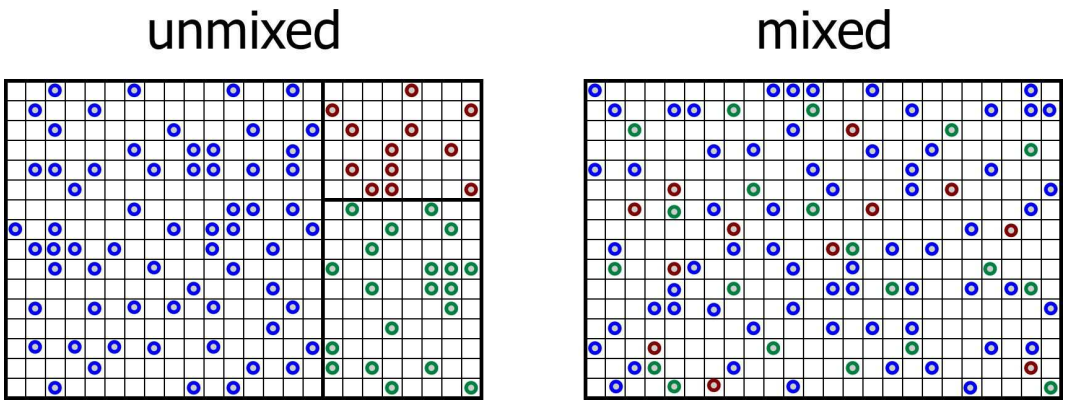



Figure 2.28: Mixing among three different species of particles. The mixed configuration has an additional entropy, $\Delta S_{\rm mix}$.

where $N_0 = B - \sum_{a=1}^{\mathsf{c}} N_a$ is the number of vacant boxes. Again using Stirling's rule, we find

$$S_{\rm mixed} = -k_{\rm B} B \sum_{a=0}^{\mathsf{c}} y_a \log y_a \quad , \tag{2.340}$$

where $y_a = N_a/B$ is the fraction of *all* boxes containing a particle of species $a$, where we include the $a = 0$ term in the sum to account for the vacant boxes. Note that

$$y_a = \frac{N_a}{B} = \frac{N_a}{B_a} \cdot \frac{B_a}{B} = x_a \, f_a \quad , \tag{2.341}$$

where $f_a \equiv B_a/B$. Note that $\sum_{a=1}^{\mathsf{c}} f_a = 1$ because $B = \sum_{a=1}^{\mathsf{c}} B_a$ .

We may now write

$$S_{\rm mixed} = -k_{\rm B} B \sum_{a=1}^{\mathsf{c}} \Big[ x_a \, f_a \log(f_a) + f_a \, x_a \log x_a \Big] - k_{\rm B} B \overbrace{\Big( 1 - \sum_{a=1}^{\mathsf{c}} x_a \, f_a \Big)}^{y_0} \log\Big( 1 - \sum_{a=1}^{\mathsf{c}} x_a f_a \Big) \tag{2.342}$$

If $x_a = x$ for all $a \in \{1, \ldots, \mathsf{c}\}$, meaning that the dimensionless number density is the same in each of the $\mathsf{c}$ initial regions, then

$$\begin{aligned} S_{\rm unmixed} &= -k_{\rm B} B \Big[ x \log x + (1-x) \log(1-x) \Big] \\ S_{\rm mixed} &= -k_{\rm B} B x \sum_{a=1}^{\mathsf{c}} f_a \log f_a - k_{\rm B} B \Big[ x \log x + (1-x) \log(1-x) \Big] \end{aligned} \tag{2.343}$$

and thus, after identifying $N = xB$ is the total number of particles (*i.e.* occupied little boxes), the entropy of mixing is

$$\Delta S_{\rm mix} = -N k_{\rm B} \sum_{a=1}^{\mathsf{c}} f_a \log f_a \quad , \tag{2.344}$$

where $N = \sum_{a=1}^{\mathsf{c}} N_a$ is the total number of particles of all species (excluding vacancies) and $f_a = N_a/N$ is the fraction of particles with species $a$.

### 2.12.3 Weak solutions and osmotic pressure

Suppose one of the species is much more plentiful than all the others, and label it with $a = 0$. We will call this the *solvent*. The remaining species are the *solutes*. The entropy of mixing is then

$$\Delta S_{\rm mix} = -k_{\rm B}\left[N_0 \log\left(\frac{N_0}{N_0+\widetilde{N}}\right) + \sum_{a=1}^{\tilde{\mathsf{c}}} N_a \log\left(\frac{N_a}{N_0+\widetilde{N}}\right)\right] \quad , \tag{2.345}$$

where $\widetilde{N} = \sum_{a=1}^{\tilde{\mathsf{c}}} N_a$ is the total number of solute molecules, summed over all solute species. The total number of solute species is defined to be $\tilde{\mathsf{c}}$, with $\mathsf{c} = 1 + \tilde{\mathsf{c}}$ the total number of all species including the solvent. We assume the solution is *weak*, which means $N_a \le \widetilde{N} \ll N_0$. Expanding in powers of $\widetilde{N}/N_0$ and $N_a/N_0$, we find

$$\Delta S_{\rm mix} = -k_{\rm B}\sum_{a=1}^{\tilde{\mathsf{c}}}\left[N_a \log\left(\frac{N_a}{N_0}\right) - N_a\right] + \frac{k_{\rm B}}{2N_0}\sum_{a,b=1}^{\tilde{\mathsf{c}}} N_a N_b + \mathcal{O}\left(\frac{\widetilde{N}^3}{N_0^2}\right) \quad . \tag{2.346}$$

Consider now a solution consisting of $N_0$ molecules of a solvent and $N_a$ molecules of species $a$ of solute, where $a = 1, \ldots, \tilde{\mathsf{c}}$, with $\tilde{\mathsf{c}}$ the total number of solute species. Including the solvent, the total number of all species is $\mathsf{c} = 1 + \tilde{\mathsf{c}}$. We begin by expanding the Gibbs free energy $G(T, p, N_0, N_1, \ldots, N_{\tilde{\mathsf{c}}})$ in powers of the solute populations. Recall $G = E - TS + pV$, so we should be careful to include a contribution $\Delta G = -T\Delta S_{\rm mix}$ due to the entropy of mixing. Thus,

$$\begin{aligned} G\big(T, p, N_0, \{N_a\}\big) = N_0\, g_0(T,p) &+ k_{\rm B}T\sum_{a=1}^{\tilde{\mathsf{c}}} N_a \log\left(\frac{N_a}{eN_0}\right) \\ &+ \sum_{a=1}^{\tilde{\mathsf{c}}} N_a\, \omega_a(T,p) + \frac{1}{2N_0}\sum_{a,b=1}^{\tilde{\mathsf{c}}} A_{ab}(T,p)\, N_a\, N_b \quad . \end{aligned} \tag{2.347}$$

The first term on the RHS corresponds to the Gibbs free energy of the solvent. The second term is due to the entropy of mixing. The third term is the contribution to the total free energy from the individual species. Note the factor of $e$ in the denominator inside the logarithm, which accounts for the second term in the brackets on the RHS of eqn. 2.346. The last term is due to interactions between the species plus a contribution from $\Delta S_{\rm mix}$. It is truncated at second order in the solute populations.

The chemical potential for the solvent is

$$\mu_0(T,p) = \frac{\partial G}{\partial N_0} = g_0(T,p) - k_{\rm B}T\sum_{a=1}^{\tilde{\mathsf{c}}} x_a - \tfrac{1}{2}\sum_{a,b=1}^{\tilde{\mathsf{c}}} A_{ab}(T,p)\, x_a\, x_b \quad , \tag{2.348}$$

and the chemical potential for species $a$ is

$$\mu_a(T,p) = \frac{\partial G}{\partial N_a} = k_{\rm B}T\log x_a + \omega_a(T,p) + \sum_{b=1}^{\tilde{\mathsf{c}}} A_{ab}(T,p)\, x_b \quad , \tag{2.349}$$

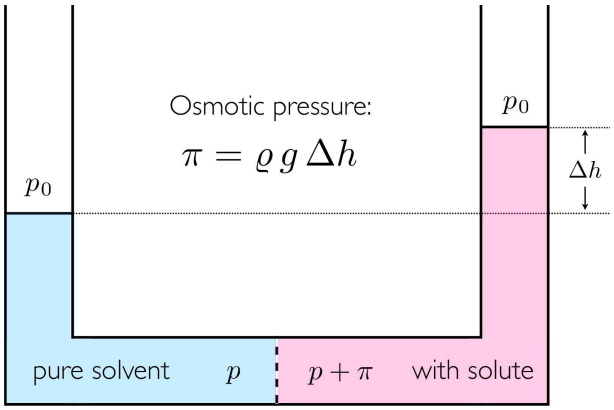



Figure 2.29: Osmotic pressure causes the column on the right side of the U-tube to rise higher than the column on the left by an amount $\Delta h = \pi/\varrho\, g$.

where $x_a = N_a/N_0$ is the concentration of solute species $a$. By assumption, the last term on the RHS of each of these equations is small, since $\widetilde{N} \ll N_0$, where $\widetilde{N} = \sum_{a=1}^{\tilde{c}} N_a$ is the total number of solute molecules. To lowest order, then, we have

$$\begin{aligned} \mu_0(T,p) &= g_0(T,p) - x\, k_{\rm B} T \\ \mu_a(T,p) &= k_{\rm B} T \log x_a + \omega_a(T,p) \quad , \end{aligned} \tag{2.350}$$

where $x = \sum_{a=1}^{\tilde{c}} x_a$ is the total solute concentration.

If we add sugar to a solution confined by a semipermeable membrane[27], the pressure increases! To see why, consider a situation where a rigid semipermeable membrane separates a solution (solvent plus solutes) from a pure solvent. There is energy exchange through the membrane, so the temperature is $T$ throughout. There is no volume exchange, however: $dV = dV' = 0$, hence the pressure need not be the same. Since the membrane is permeable to the solvent, we have that the chemical potential $\mu_0$ is the same on each side. This means

$$g_0(T, p_{\rm R}) - x k_{\rm B} T = g_0(T, p_{\rm L}) \quad , \tag{2.351}$$

where $p_{\rm L,R}$ is the pressure on the left and right sides of the membrane, and $x = \widetilde{N}/N_0$ is again the total solute concentration. This equation once again tells us that the pressure $p$ cannot be the same on both sides of the membrane. If the pressure difference is small, we can expand in powers of the *osmotic pressure*, $\pi \equiv p_{\rm R} - p_{\rm L}$, and we find

$$\pi = x\, k_{\rm B} T \Big/ \left(\frac{\partial \mu_0}{\partial p}\right)_T \quad . \tag{2.352}$$

But a Maxwell relation (§2.8) guarantees

$$\left(\frac{\partial \mu}{\partial p}\right)_{T,N} = \left(\frac{\partial V}{\partial N}\right)_{T,p} = v(T,p)/N_{\rm A} \quad , \tag{2.353}$$

where $v(T,p)$ is the molar volume of the solvent.

$$\pi v = xRT \quad , \tag{2.354}$$

[27] 'Semipermeable' in this context means permeable to the solvent but not the solute(s).

which looks very much like the ideal gas law, even though we are talking about dense (but 'weak') solutions! The resulting pressure has a demonstrable effect, as sketched in fig. 2.29. Consider a solution containing $\nu$ moles of sucrose ($C_{12}H_{22}O_{11}$) per kilogram ($55.52\,\mathrm{mol}$) of water at $30^\circ\,\mathrm{C}$. We then find $\pi = 2.5\,\mathrm{atm}$ when $\nu = 0.1$.

One might worry about the expansion in powers of $\pi$ when $\pi$ is much larger than the ambient pressure. But in fact the next term in the expansion is smaller than the first term by a factor of $\pi\kappa_T$, where $\kappa_T$ is the isothermal compressibility. For water one has $\kappa_T \approx 4.4 \times 10^{-5}\,(\mathrm{atm})^{-1}$, hence we can safely ignore the higher order terms in the Taylor expansion.

### 2.12.4 Effect of impurities on boiling and freezing points

Along the coexistence curve separating liquid and vapor phases, the chemical potentials of the two phases are identical:

$$\mu^0_{\rm L}(T,p) = \mu^0_{\rm V}(T,p) \quad . \tag{2.355}$$

Here we write $\mu^0$ for $\mu$ to emphasize that we are talking about a phase with no impurities present. This equation provides a single constraint on the two variables $T$ and $p$, hence one can, in principle, solve to obtain $T = T^*_0(p)$, which is the equation of the liquid-vapor coexistence curve in the $(T,p)$ plane. Now suppose there is a solute present in the liquid. We then have

$$\mu_{\rm L}(T,p,x) = \mu^0_{\rm L}(T,p) - xk_{\rm B}T \quad , \tag{2.356}$$

where $x$ is the dimensionless solute concentration, summed over all species. The condition for liquid-vapor coexistence now becomes

$$\mu^0_{\rm L}(T,p) - xk_{\rm B}T = \mu^0_{\rm V}(T,p) \quad . \tag{2.357}$$

This will lead to a shift in the boiling temperature at fixed $p$. Assuming this shift is small, let us expand to lowest order in $\big(T - T^*_0(p)\big)$, writing

$$\mu^0_{\rm L}(T^*_0,p) + \left(\frac{\partial\mu^0_{\rm L}}{\partial T}\right)_p \big(T - T^*_0\big) - xk_{\rm B}T = \mu^0_{\rm V}(T^*_0,p) + \left(\frac{\partial\mu^0_{\rm V}}{\partial T}\right)_p \big(T - T^*_0\big) \quad . \tag{2.358}$$

Note that

$$\left(\frac{\partial\mu}{\partial T}\right)_{p,N} = -\left(\frac{\partial S}{\partial N}\right)_{T,p} \tag{2.359}$$

from a Maxwell relation deriving from exactness of $dG$. Since $S$ is extensive, we can write $S = (N/N_{\rm A})\,s(T,p)$, where $s(T,p)$ is the molar entropy. Solving for $T$, we obtain

$$T^*(p,x) = T^*_0(p) + \frac{xR\big[T^*_0(p)\big]^2}{\ell_{\rm v}(p)} \quad , \tag{2.360}$$

where $\ell_{\rm v} = T^*_0 \cdot (s_{\rm V} - s_{\rm L})$ is the *latent heat* of the liquid-vapor transition (see §2.11.2). The shift $\Delta T^* = T^* - T^*_0$ is called the *boiling point elevation*.

| Substance | Latent Heat of Fusion $\tilde{\ell}_{\rm f}$ J/g | Melting Point °C | Latent Heat of Vaporization $\tilde{\ell}_{\rm v}$ J/g | Boiling Point °C |
|---|---|---|---|---|
| $C_2H_5OH$ | 108 | -114 | 855 | 78.3 |
| $NH_3$ | 339 | -75 | 1369 | -33.34 |
| He | – | – | 21 | -268.93 |
| $H_2$ | 58 | -259 | 455 | -253 |
| Pb | 24.5 | 327.5 | 871 | 1750 |
| $N_2$ | 25.7 | -210 | 200 | -196 |
| $O_2$ | 13.9 | -219 | 213 | -183 |
| $H_2O$ | 334 | 0 | 2270 | 100 |

Table 2.4: Latent heats of fusion and vaporization at $p = 1$ atm.

As an example, consider seawater, which contains approximately $35\,{\rm g}$ of dissolved $\rm Na^+Cl^-$ per kilogram of $H_2O$. The atomic masses of Na and Cl are $23.0$ and $35.4$, respectively, hence the total ionic concentration in seawater (neglecting everything but sodium and chlorine) is given by

$$x = \frac{2\cdot 35}{23.0+35.4}\Big/\frac{1000}{18} \approx 0.022 \quad . \tag{2.361}$$

The latent heat of vaporization of $H_2O$ at atmospheric pressure is $\ell = 40.7\,{\rm kJ/mol}$, hence

$$\Delta T^* = \frac{(0.022)(8.3\,{\rm J/mol\,K})(373\,{\rm K})^2}{4.1\times 10^4\,{\rm J/mol}} \approx 0.6\,{\rm K} \quad . \tag{2.362}$$

Put another way, the boiling point elevation of $H_2O$ at atmospheric pressure is about $0.28^\circ{\rm C}$ per percent solute. We can express this as $\Delta T^* = Km$, where the *molality* $m$ is the number of moles of solute per kilogram of solvent. For $H_2O$, we find $K = 0.51^\circ{\rm C\,kg/mol}$.

Similar considerations apply at the freezing point, when we equate the chemical potential of the solvent plus solute to that of the pure solid. The *latent heat of fusion* for $H_2O$ is about $\ell_{\rm f} = T_{\rm f}^0\cdot(s_{\rm LIQUID} - s_{\rm SOLID}) = 6.01\,{\rm kJ/mol}$[28]. We thus predict a *freezing point depression* of $\Delta T^* = -xR\big[T_0^*\big]^2/\ell_{\rm f} = -1.03^\circ{\rm C}\cdot x[\%]$. This can be expressed once again as $\Delta T^* = -Km$, with $K = 1.86^\circ{\rm C\,kg/mol}$[29].

[28] See table 2.4, and recall $M = 18\,{\rm g}$ is the molar mass of $H_2O$.

[29] It is more customary to write $\Delta T^* = T^*_{\text{pure solvent}} - T^*_{\text{solution}}$ in the case of the freezing point depression, in which case $\Delta T^*$ is positive.

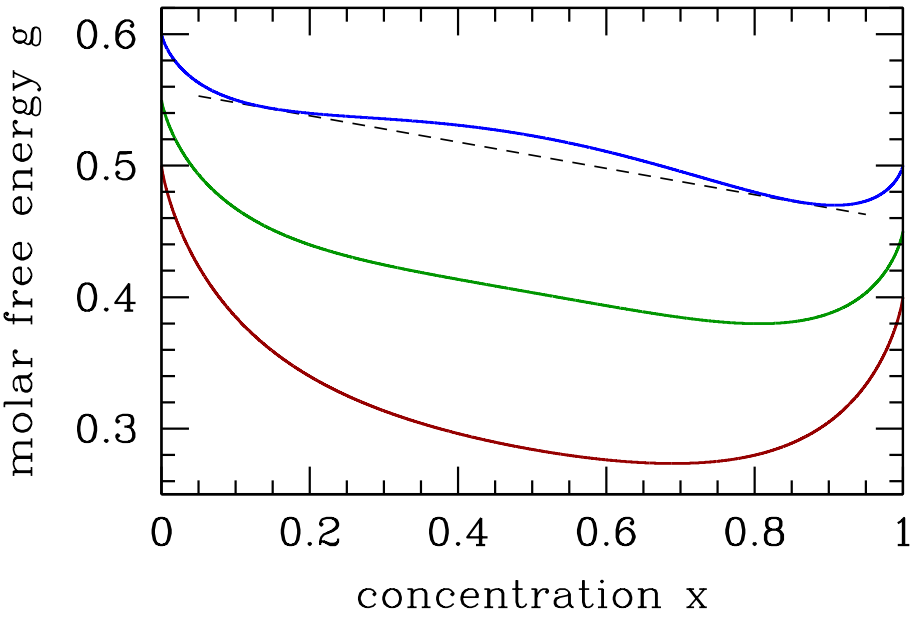



Figure 2.30: Gibbs free energy per particle for a binary solution as a function of concentration $x = x_{\rm B}$ of the B species (pure A at the left end $x = 0$ ; pure B at the right end $x = 1$), in units of the interaction parameter $\lambda$. Dark red curve: $T = 0.65\,\lambda/k_{\rm B} > T_{\rm c}$ ; green curve: $T = \lambda/2k_{\rm B} = T_{\rm c}$ ; blue curve: $T = 0.40\,\lambda/k_{\rm B} < T_{\rm c}$. We have chosen $\mu^0_{\rm A} = 0.60\,\lambda - 0.50\,k_{\rm B}T$ and $\mu^0_{\rm B} = 0.50\,\lambda - 0.\,50\,k_{\rm B}T$. Note that the free energy $g(T,p,x)$ is not convex in $x$ for $T < T_{\rm c}$, indicating an instability and necessitating a Maxwell construction.

### 2.12.5 Binary solutions

Consider a binary solution, and write the Gibbs free energy $G(T,p,N_{\rm A},N_{\rm B})$ as

$$G(T,p,N_{\rm A},N_{\rm B}) = N_{\rm A}\,\mu^0_{\rm A}(T,p) + N_{\rm B}\,\mu^0_{\rm B}(T,p) + N_{\rm A}k_{\rm B}T\,\log\left(\frac{N_{\rm A}}{N_{\rm A}+N_{\rm B}}\right) + N_{\rm B}\,k_{\rm B}T\,\log\left(\frac{N_{\rm B}}{N_{\rm A}+N_{\rm B}}\right) + \lambda\,\frac{N_{\rm A}N_{\rm B}}{N_{\rm A}+N_{\rm B}} \quad . \tag{2.363}$$

The first four terms on the RHS represent the free energy of the individual component fluids and the entropy of mixing. The last term is an interaction contribution. With $\lambda > 0$, the interaction term prefers that the system be either fully A or fully B. The entropy contribution prefers a mixture, so there is a competition. What is the stable thermodynamic state?

It is useful to write the Gibbs free energy per particle, $g(T,p,x) = G/(N_{\rm A}+N_{\rm B})$, in terms of $T$, $p$, and the concentration $x \equiv x_{\rm B} = N_{\rm B}/(N_{\rm A}+N_{\rm B})$ of species B (hence $x_{\rm A} = 1-x$ is the concentration of species A). Then

$$g(T,p,x) = (1-x)\,\mu^0_{\rm A} + x\,\mu^0_{\rm B} + k_{\rm B}T\Big[x\log x + (1-x)\log(1-x)\Big] + \lambda\,x\,(1-x) \quad . \tag{2.364}$$

In order for the system to be stable against phase separation into relatively A-rich and B-rich regions, we must have that $g(T,p,x)$ be a convex function of $x$. Our first check should be for a local instability,

*i.e.* spinodal decomposition. We have

$$\frac{\partial g}{\partial x} = \mu_{\rm B}^0 - \mu_{\rm A}^0 + k_{\rm B}T\,\log\left(\frac{x}{1-x}\right) + \lambda\,(1-2x) \tag{2.365}$$

and

$$\frac{\partial^2 g}{\partial x^2} = \frac{k_{\rm B}T}{x} + \frac{k_{\rm B}T}{1-x} - 2\lambda \quad . \tag{2.366}$$

The spinodal is given by the solution to the equation $\partial^2 g/\partial x^2 = 0$, which is

$$T^*(x) = \frac{2\lambda}{k_{\rm B}}\,x\,(1-x) \quad . \tag{2.367}$$

Since $x\,(1-x)$ achieves its maximum value of $\frac{1}{4}$ at $x = \frac{1}{2}$, we have $T^* \le \lambda/2k_{\rm B}$.

In fig. 2.30 we sketch the free energy $g(T,p,x)$ *versus* $x$ for three representative temperatures. For $T > \lambda/2k_{\rm B}$, the free energy is everywhere convex in $x$. When $T < \lambda/2k_{\rm B}$, the free energy resembles the blue curve in fig. 2.30, and the system is unstable to phase separation. The two phases are said to be *immiscible*, or, equivalently, there exists a *solubility gap*. To determine the coexistence curve, we perform a Maxwell construction, writing

$$\frac{g(x_2)-g(x_1)}{x_2-x_1} = \left.\frac{\partial g}{\partial x}\right|_{x_1} = \left.\frac{\partial g}{\partial x}\right|_{x_2} \quad . \tag{2.368}$$

Here, $x_1$ and $x_2$ are the boundaries of the two phase region. These equations admit a symmetry of $x \leftrightarrow 1-x$, hence we can set $x = x_1$ and $x_2 = 1-x$. We find

$$g(1-x) - g(x) = (1-2x)\left(\mu_{\rm B}^0 - \mu_{\rm A}^0\right) \quad , \tag{2.369}$$

and invoking eqns. 2.368 and 2.365 we obtain the solution

$$T_{\rm coex}(x) = \frac{\lambda}{k_{\rm B}} \cdot \frac{1-2x}{\log\left(\frac{1-x}{x}\right)} \quad . \tag{2.370}$$

The phase diagram for the binary system is shown in fig. 2.32. For $T < T^*(x)$, the system is unstable, and spinodal decomposition occurs. For $T^*(x) < T < T_{\rm coex}(x)$, the system is *metastable*, just like the van der Waals gas in its corresponding regime. Real binary solutions behave qualitatively like the model discussed here, although the coexistence curve is generally not symmetric under $x \leftrightarrow 1-x$, and the single phase region extends down to low temperatures for $x \approx 0$ and $x \approx 1$. If $\lambda$ itself is temperature-dependent, there can be multiple solutions to eqns. 2.367 and 2.370. For example, one could take

$$\lambda(T) = \frac{\lambda_0\,T^2}{T^2 + T_0^2} \quad . \tag{2.371}$$

In this case, $k_{\rm B}T > \lambda$ at both high and low temperatures, and we expect the single phase region to be *reentrant*. Such a phenomenon occurs in water-nicotine mixtures, for example.

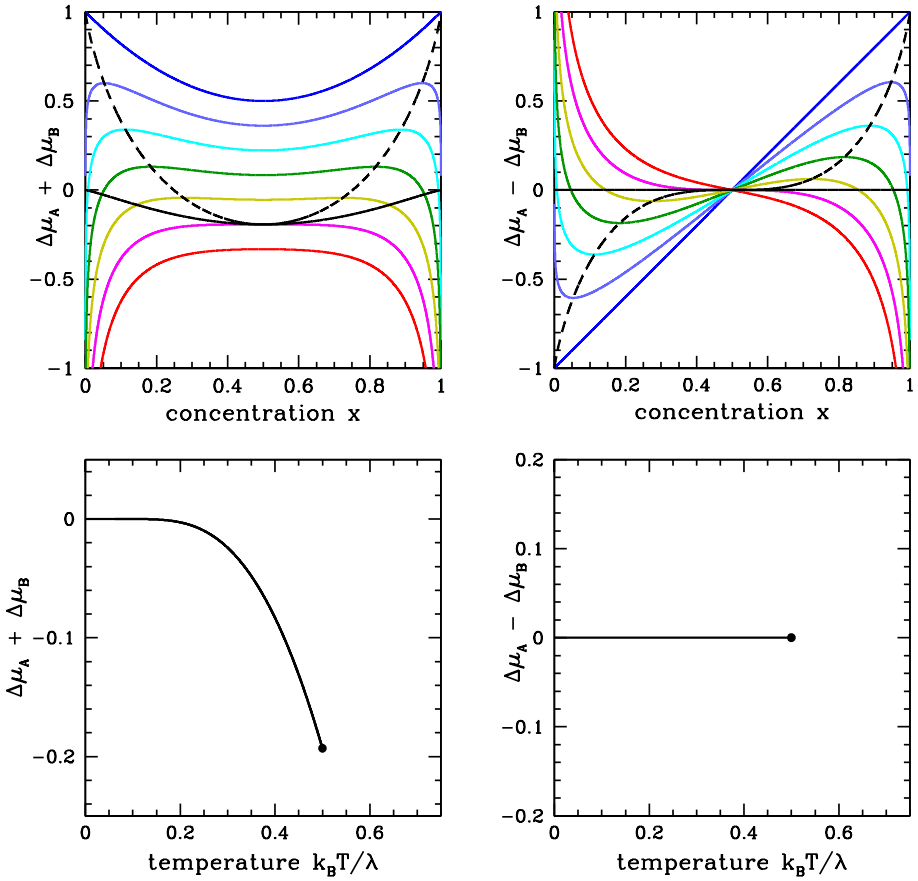



Figure 2.31: Upper panels: chemical potential shifts $\Delta\mu_\pm = \Delta\mu_{\rm A} \pm \Delta\mu_{\rm B}$ *versus* concentration $x = x_{\rm B}$. The dashed black line is the spinodal, and the solid black line the coexistence boundary. Temperatures range from $T = 0$ (dark blue) to $T = 0.6\,\lambda/k_{\rm B}$ (red) in units of $0.1\,\lambda/k_{\rm B}$. Lower panels: phase diagram in the $(T, \Delta\mu_\pm)$ planes. The black dot is the critical point.

It is instructive to consider the phase diagram in the $(T, \mu)$ plane. We define the chemical potential shifts,

$$\Delta\mu_{\rm A} \equiv \mu_{\rm A} - \mu_{\rm A}^0 = k_{\rm B}T\log(1-x) + \lambda\,x^2 \tag{2.372}$$

$$\Delta\mu_{\rm B} \equiv \mu_{\rm B} - \mu_{\rm B}^0 = k_{\rm B}T\log x + \lambda\,(1-x)^2 \quad , \tag{2.373}$$

and their sum and difference, $\Delta\mu_\pm \equiv \Delta\mu_{\rm A} \pm \Delta\mu_{\rm B}$. From the Gibbs-Duhem relation, we know that we can write $\mu_{\rm B}$ as a function of $T$, $p$, and $\mu_{\rm A}$. Alternately, we could write $\Delta\mu_\pm$ in terms of $T$, $p$, and $\Delta\mu_\mp$, so we can choose which among $\Delta\mu_+$ and $\Delta\mu_-$ we wish to use in our phase diagram. The results are plotted in fig. 2.31. It is perhaps easiest to understand the phase diagram in the $(T, \Delta\mu_-)$ plane. At low temperatures, below $T = T_{\rm c} = \lambda/2k_{\rm B}$, there is a first order phase transition at $\Delta\mu_- = 0$. For $T < T_{\rm c} = \lambda/2k_{\rm B}$ and $\Delta\mu_- = 0^+$, *i.e.* infinitesimally positive, the system is in the A-rich phase, but for $\Delta\mu_- = 0^-$, *i.e.* infinitesimally negative, it is B-rich. The concentration $x = x_{\rm B}$ changes discontinuously across the phase boundary. The critical point lies at $(T, \Delta\mu_-) = (\lambda/2k_{\rm B}\,,\,0)$.

If we choose $N = N_{\rm A} + N_{\rm B}$ to be the extensive variable, then fixing $N$ means $dN_{\rm A} + dN_{\rm B} = 0$. So at fixed

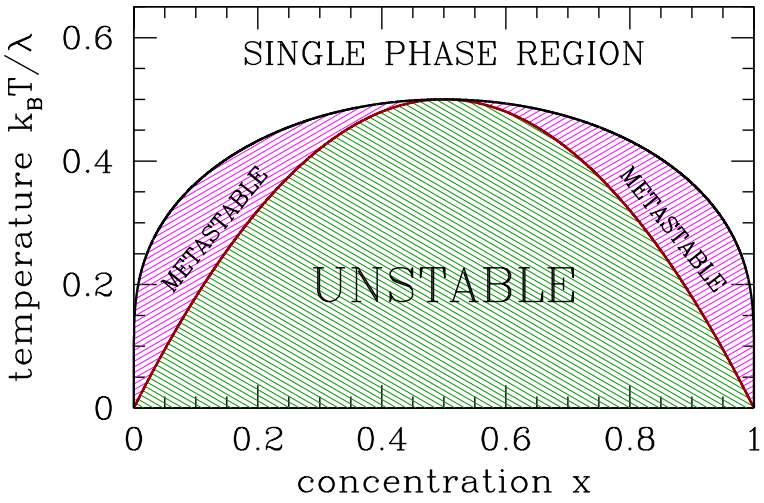



Figure 2.32: Phase diagram for the binary system. The black curve is the coexistence curve, and the dark red curve is the spinodal. A-rich material is to the left and B-rich to the right.

$T$ and $p$,

$$dG\big|_{T,p} = \mu_{\rm A}\,dN_{\rm A} + \mu_{\rm B}\,dN_{\rm B} \qquad \Rightarrow \qquad dg\big|_{T,p} = \big(\mu^0_{\rm B} - \mu^0_{\rm A} - \Delta\mu_-\big)dx \quad . \tag{2.374}$$

Since $\Delta\mu_-(x,T) = \varphi(x,T) - \varphi(1-x,T) = -\Delta\mu_-(1-x,T)$, where $\varphi(x,T) = \lambda x - k_{\rm B}T\log x$, we have that the coexistence boundary in the $(x,\Delta_-)$ plane is simply the line $\Delta\mu_- = 0$, because $\int\limits_x^{1-x} dx'\,\Delta\mu_-(x',T) = 0$.

Note also that there is no two-phase region in the $(T,\Delta\mu)$ plane; the phase boundary in this plane is a curve which terminates at a critical point. As we saw in §2.11, the same situation pertains in single component $(p,v,T)$ systems. That is, the phase diagram in the $(p,v)$ or $(T,v)$ plane contains two-phase regions, but in the $(p,T)$ plane the boundaries between phases are one-dimensional curves. Any two-phase behavior is confined to these curves, where the thermodynamic potentials are singular.

The phase separation can be seen in a number of systems. A popular example involves mixtures of water and ouzo or other anise-based liqueurs, such as arak and absinthe. Starting with the pure liqueur ($x = 1$), and at a temperature below the coexistence curve maximum, the concentration is diluted by adding water. Follow along on fig. 2.32 by starting at the point $(x = 1\,,\ k_{\rm B}T/\lambda = 0.4)$ and move to the left. Eventually, one hits the boundary of the two-phase region. At this point, the mixture turns milky, due to the formation of large droplets of the pure phases on either side of coexistence region which scatter light, a process known as *spontaneous emulsification*[30]. As one continues to dilute the solution with more water, eventually one passes all the way through the coexistence region, at which point the solution becomes clear once again, and described as a single phase.

What happens if $\lambda < 0$ ? In this case, both the entropy and the interaction energy prefer a mixed phase, and there is no instability to phase separation. The two fluids are said to be *completely miscible*. An example would be benzene, $C_6H_6$, and toluene, $C_7H_8$ (*i.e.* $C_6H_5CH_3$). The phase diagram would be blank, with no phase boundaries below the boiling transition, because the fluid could exist as a mixture

[30] An emulsion is a mixture of two or more immiscible liquids.

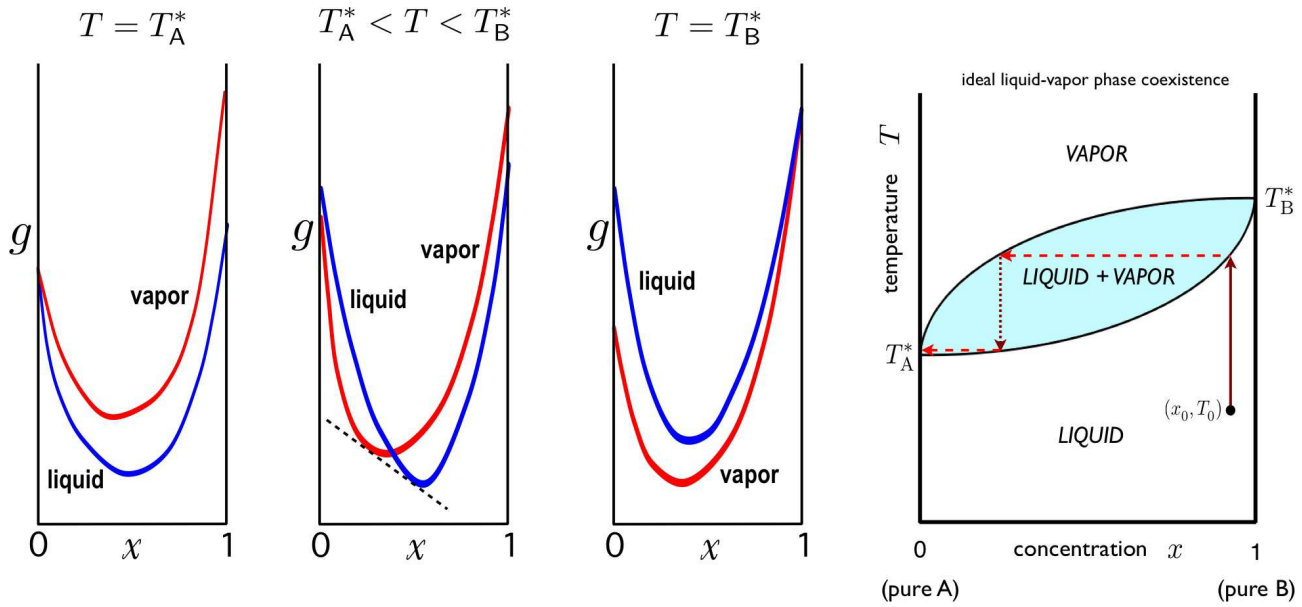



Figure 2.33: Gibbs free energy per particle $g$ for an ideal binary solution for temperatures $T \in [T_A^*, T_B^*]$. The Maxwell construction is shown for the case $T_A^* < T < T_B^*$. Right: phase diagram, showing two-phase region and distillation sequence in $(x, T)$ space.

in any proportion.

Any fluid will eventually boil if the temperature is raised sufficiently high. Let us assume that the boiling points of our A and B fluids are $T_{A,B}^*$, and without loss of generality let us take $T_A^* < T_B^*$ at some given fixed pressure[31]. This means $\mu_A^L(T_A^*, p) = \mu_A^V(T_A^*, p)$ and $\mu_B^L(T_B^*, p) = \mu_B^V(T_B^*, p)$. What happens to the mixture? We begin by writing the free energies of the mixed liquid and mixed vapor phases as

$$g_L(T, p, x) = (1-x)\,\mu_A^L(T, p) + x\,\mu_B^L(T, p) + k_B T\Big[x \log x + (1-x)\log(1-x)\Big] + \lambda_L\, x(1-x) \tag{2.375}$$

$$g_V(T, p, x) = (1-x)\,\mu_A^V(T, p) + x\,\mu_B^V(T, p) + k_B T\Big[x \log x + (1-x)\log(1-x)\Big] + \lambda_V\, x(1-x) \quad . \tag{2.376}$$

Typically $\lambda_V \approx 0$. Consider these two free energies as functions of the concentration $x$, at fixed $T$ and $p$. If the curves never cross, and $g_L(x) < g_V(x)$ for all $x \in [0, 1]$, then the liquid is always the state of lowest free energy. This is the situation in the first panel of fig. 2.33. Similarly, if $g_V(x) < g_L(x)$ over this range, then the mixture is in the vapor phase throughout. What happens if the two curves cross at some value of $x$? This situation is depicted in the second panel of fig. 2.33. In this case, there is always a Maxwell construction which lowers the free energy throughout some range of concentration, *i.e.* the system undergoes phase separation.

In an *ideal fluid*, we have $\lambda_L = \lambda_V = 0$, and setting $g_L = g_V$ requires

$$(1-x)\,\Delta\mu_A(T, p) + x\,\Delta\mu_B(T, p) = 0 \quad , \tag{2.377}$$

[31] We assume the boiling temperatures are not exactly equal.

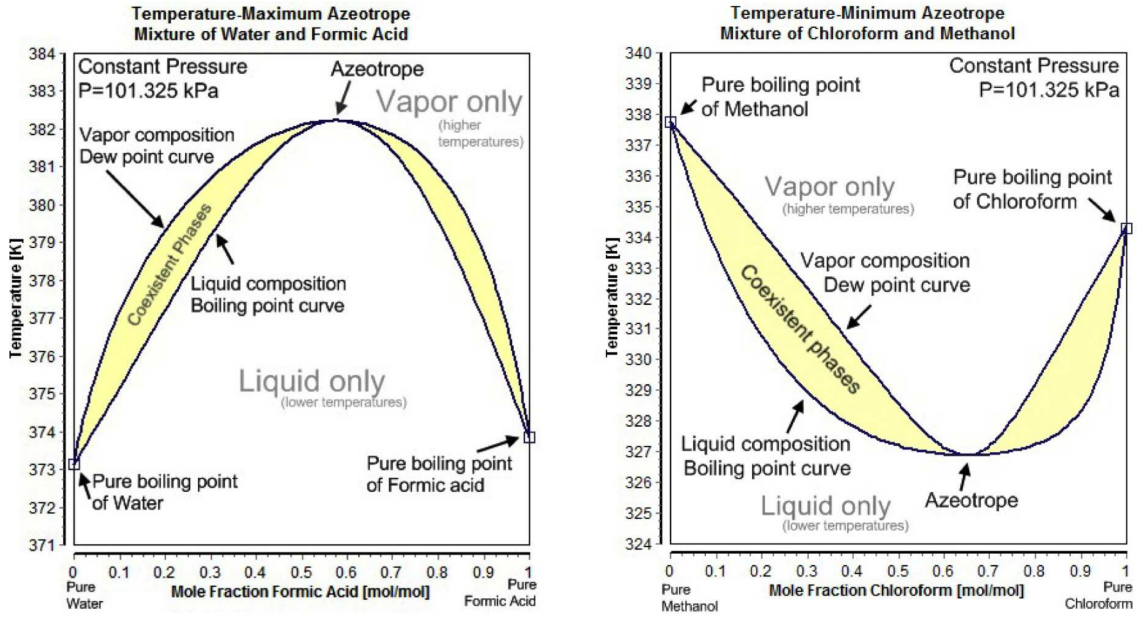



Figure 2.34: Negative (left) and positive (right) azeotrope phase diagrams. From *Wikipedia*.

where $\Delta\mu_{\mathsf{A/B}}(T,p) = \mu^{\rm L}_{\mathsf{A/B}}(T,p) - \mu^{\rm V}_{\mathsf{A/B}}(T,p)$. Expanding the chemical potential about a given temperature $T^*$,

$$\mu(T,p) = \mu(T^*,p) - s(T^*,p)\,(T-T^*) - \frac{c_p(T^*,p)}{2T^*}\,(T-T^*)^2 + \dots \quad , \tag{2.378}$$

where we have used the fact that $(\partial\mu/\partial T)_{p,N} = -\partial S/\partial N_{T,p} = -s(T,p)$ is the entropy per particle, and $(\partial s/\partial T)_{p,N} = c_p/T$. Thus, expanding $\Delta\mu_{\mathsf{A/B}}$ about $T^*_{\mathsf{A/B}}$, we have

$$\begin{aligned}\Delta\mu_{\mathsf{A}} &\equiv \mu^{\rm L}_{\mathsf{A}} - \mu^{\rm V}_{\mathsf{A}} = (s^{\rm V}_{\mathsf{A}} - s^{\rm L}_{\mathsf{A}})(T - T^*_{\mathsf{A}}) + \frac{c^{\rm V}_{p,\mathsf{A}} - c^{\rm L}_{p,\mathsf{A}}}{2T^*_{\mathsf{A}}}\,(T-T^*_{\mathsf{A}})^2 + \dots \\ \Delta\mu_{\mathsf{B}} &\equiv \mu^{\rm L}_{\mathsf{B}} - \mu^{\rm V}_{\mathsf{B}} = (s^{\rm V}_{\mathsf{B}} - s^{\rm L}_{\mathsf{B}})(T - T^*_{\mathsf{B}}) + \frac{c^{\rm V}_{p,\mathsf{B}} - c^{\rm L}_{p,\mathsf{B}}}{2T^*_{\mathsf{B}}}\,(T-T^*_{\mathsf{B}})^2 + \dots\end{aligned} \tag{2.379}$$

We assume $s^{\rm V}_{\mathsf{A/B}} > s^{\rm L}_{\mathsf{A/B}}$, *i.e.* the vapor phase has greater entropy per particle. Thus, $\Delta\mu_{\mathsf{A/B}}(T)$ changes sign from negative to positive as $T$ rises through $T^*_{\mathsf{A/B}}$. If we assume that these are the *only* sign changes for $\Delta\mu_{\mathsf{A/B}}(T)$ at fixed $p$, then eqn. 2.377 can only be solved for $T \in [T^*_{\mathsf{A}}, T^*_{\mathsf{B}}]$. This immediately leads to the phase diagram in the rightmost panel of fig. 2.33.

According to the Gibbs phase rule, with $\mathsf{c} = 2$, two-phase equilibrium ($\varphi = 2$) occurs along a subspace of dimension $d_{\rm PE} = 2 + \mathsf{c} - \varphi = 2$. Thus, if we fix the pressure $p$ and the concentration $x = x_{\rm B}$, liquid-gas equilibrium occurs at a particular temperature $T^*$, known as the boiling point. Since the liquid and the vapor with which it is in equilibrium at $T^*$ may have different composition, *i.e.* different values of $x$, one may *distill* the mixture to separate the two pure substances, as follows. First, given a liquid mixture of A and B, we bring it to boiling, as shown in the rightmost panel of fig. 2.33. The vapor is at a different concentration $x$ than the liquid (a lower value of $x$ if the boiling point of pure A is less than that of pure

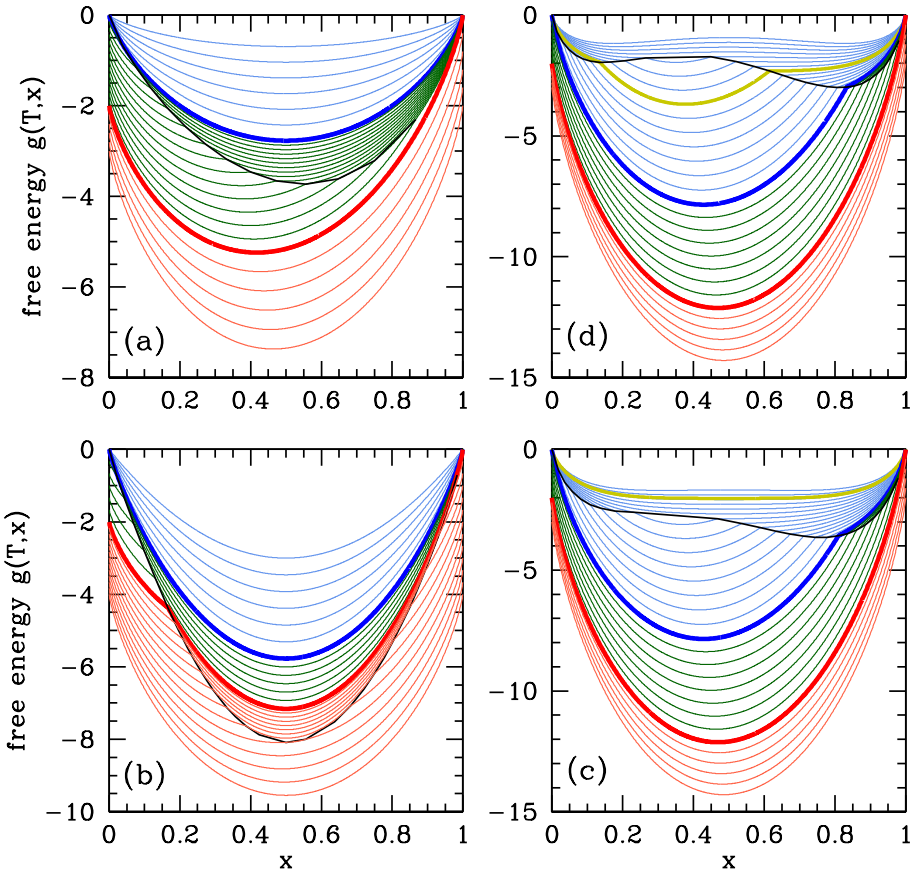



Figure 2.35: Free energies *before Maxwell constructions* for a binary fluid mixture in equilibrium with a vapor ($\lambda_{\rm V} = 0$). Panels show (a) $\lambda_{\rm L} = 0$ (ideal fluid), (b) $\lambda_{\rm L} < 0$ (miscible fluid; negative azeotrope), (c) $\lambda^{\rm L}_{\rm AB} > 0$ (positive azeotrope), (d) $\lambda^{\rm L}_{\rm AB} > 0$ (heteroazeotrope). Thick blue and red lines correspond to temperatures $T^*_{\rm A}$ and $T^*_{\rm B}$, respectively, with $T^*_{\rm A} < T^*_{\rm B}$. Thin blue and red curves are for temperatures outside the range $[T^*_{\rm A}, T^*_{\rm B}]$. The black curves mark the crossings of the liquid and vapor free energy curves, where $g_{\rm L} = g_{\rm V}$. Before the common-tangent construction, the lower envelope of these curves is continuous but generally has a discontinuous slope. The yellow curve in (d) corresponds to the coexistence temperature for the fluid mixture. In this case the azeotrope forms within the coexistence region.

B, as shown). If we collect the vapor, the remaining fluid is at a higher value of $x$. The collected vapor is then captured and then condensed, forming a liquid at the lower $x$ value. This is then brought to a boil, and the resulting vapor is drawn off and condensed, *etc.* The result is a purified A state. The remaining liquid is then at a higher B concentration. By repeated boiling and condensation, A and B can be separated. For liquid-vapor transitions, the upper curve, representing the lowest temperature at a given concentration for which the mixture is a homogeneous vapor, is called the *dew point curve*. The lower curve, representing the highest temperature at a given concentration for which the mixture is a homogeneous liquid, is called the *bubble point curve*. The same phase diagram applies to liquid-solid mixtures where both phases are completely miscible. In that case, the upper curve is called the *liquidus*, and the lower curve the *solidus*.

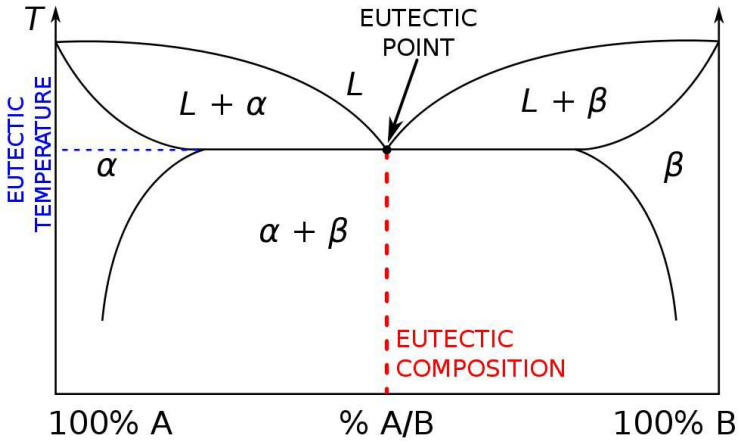



Figure 2.36: Phase diagram for a eutectic mixture in which a liquid $L$ is in equilibrium with two solid phases $\alpha$ and $\beta$. The same phase diagram holds for heteroazeotropes, where a vapor is in equilibrium with two liquid phases.

When a homogeneous liquid or vapor at concentration $x$ is heated or cooled to a temperature $T$ such that $(x, T)$ lies within the two-phase region, the mixture phase separates into the two end components $(x^*_{\rm L}, T)$ and $(x^*_{\rm V}, T)$, which lie on opposite sides of the boundary of the two-phase region, at the same temperature. The locus of points at constant $T$ joining these two points is called the *tie line*. To determine how much of each of these two homogeneous phases separates out, we use particle number conservation. If $\eta_{\rm L,V}$ is the fraction of the homogeneous liquid and homogeneous vapor phases present, then $\eta_{\rm L} x^*_{\rm L} + \eta_{\rm V} x^*_{\rm V} = x$, which says $\eta_{\rm L} = (x - x^*_{\rm V})/(x^*_{\rm L} - x^*_{\rm V})$ and $\eta_{\rm V} = (x - x^*_{\rm L})/(x^*_{\rm V} - x^*_{\rm L})$. This is known as the *lever rule*.

For many binary mixtures, the boiling point curve is as shown in fig. 2.34. Such cases are called *azeotropes*. For *negative azeotropes*, the maximum of the boiling curve lies above both $T^*_{\rm A,B}$. The free energy curves for this case are shown in panel (b) of fig. 2.35. For $x < x^*$, where $x^*$ is the azeotropic composition, one can distill A but not B. Similarly, for $x > x^*$ one can distill B but not A. The situation is different for *positive azeotropes*, where the minimum of the boiling curve lies below both $T^*_{\rm A,B}$, corresponding to the free energy curves in panel (c) of fig. 2.35. In this case, distillation (*i.e.* condensing and reboiling the collected vapor) from either side of $x^*$ results in the azeotrope. One can of course collect the fluid instead of the vapor. In general, for both positive and negative azeotropes, starting from a given concentration $x$, one can only arrive at pure A plus azeotrope (if $x < x^*$) or pure B plus azeotrope (if $x > x^*$). Ethanol ($C_2H_5OH$) and water ($H_2O$) form a positive azeotrope which is 95.6% ethanol and 4.4% water by weight. The individual boiling points are $T^*_{\rm C_2H_5OH} = 78.4^\circ{\rm C}$, $T^*_{\rm H_2O} = 100^\circ{\rm C}$, while the azeotrope boils at $T^*_{\rm AZ} = 78.2^\circ{\rm C}$. No amount of distillation of this mixture can purify ethanol beyond the 95.6% level. To go beyond this level of purity, one must resort to *azeotropic distillation*, which involves introducing another component, such as benzene (or a less carcinogenic additive), which alters the molecular interactions.

To model the azeotrope system, we need to take $\lambda_{\rm L} \neq 0$, in which case one can find *two* solutions to the energy crossing condition $g_{\rm V}(x) = g_{\rm L}(x)$. With two such crossings come two Maxwell constructions, hence the phase diagrams in fig. 2.34. *Generally*, negative azeotropes are found in systems with $\lambda_{\rm L} < 0$,

whereas positive azeotropes are found when $\lambda_{\rm L} > 0$. As we've seen, such repulsive interactions between the A and B components in general lead to a phase separation below a coexistence temperature $T_{\rm COEX}(x)$ given by eqn. 2.370. What happens if the minimum boiling point lies within the coexistence region? This is the situation depicted in panel (d) of fig. 2.35. The system is then a liquid/vapor version of the solid/liquid *eutectic* (see fig. 2.36), and the minimum boiling point mixture is called a *heteroazeotrope*.

## 2.13 Some Concepts in Thermochemistry

### 2.13.1 Chemical reactions and the law of mass action

Suppose we have a chemical reaction among c species, written as

$$\zeta_1 \, {\rm A}_1 + \zeta_2 \, {\rm A}_2 + \cdots + \zeta_{\rm c} \, {\rm A}_{\rm c} = 0 \quad , \tag{2.380}$$

where ${\rm A}_a$ represents a chemical formula and $\zeta_a$ the corresponding stoichiometric coefficient. For example, we could have

$$-3\,{\rm H}_2 - {\rm N}_2 + 2\,{\rm NH}_3 = 0 \qquad (3\,{\rm H}_2 + {\rm N}_2 \rightleftharpoons 2\,{\rm NH}_3) \tag{2.381}$$

for which

$$\zeta({\rm H}_2) = -3 \quad , \quad \zeta({\rm N}_2) = -1 \quad , \quad \zeta({\rm NH}_3) = 2 \quad . \tag{2.382}$$

When $\zeta_a > 0$, the corresponding ${\rm A}_a$ is a *product*; when $\zeta_a < 0$, the corresponding ${\rm A}_a$ is a *reactant*. The bookkeeping of the coefficients $\zeta_a$ which ensures conservation of each individual species of atom in the reaction(s) is known as *stoichiometry*[32].

Now we ask: what are the conditions for equilibrium? At constant $T$ and $p$, which is typical for many chemical reactions, the conditions are that $G\big(T, p, \{N_a\}\big)$ be a minimum. Now

$$dG = -S\,dT + V\,dp + \sum_a \mu_a \, dN_a \quad , \tag{2.383}$$

[32] Antoine Lavoisier, the "father of modern chemistry", made pioneering contributions in both chemistry and biology. In particular, he is often credited as the progenitor of stoichiometry. An aristocrat by birth, Lavoisier was an administrator of the *Ferme générale*, an organization in pre-revolutionary France which collected taxes on behalf of the king. At the age of 28, Lavoisier married Marie-Anne Pierette Paulze, the 13-year-old daughter of one of his business partners. She would later join her husband in his research, and she played a role in his disproof of the phlogiston theory of combustion. The phlogiston theory was superseded by Lavoisier's work, where, based on contemporary experiments by Joseph Priestley, he correctly identified the pivotal role played by oxygen in both chemical and biological processes (*i.e.* respiration). Despite his fame as a scientist, Lavoisier succumbed to the Reign of Terror. His association with the *Ferme générale*, which collected taxes from the poor and the downtrodden, was a significant liability in revolutionary France. Furthermore – and let this be a lesson to all of us – Lavoisier had unwisely ridiculed a worthless pseudoscientific pamphlet, ostensibly on the physics of fire, and its author, Jean-Paul Marat. Marat was a journalist with scientific pretensions, but apparently little in the way of scientific talent or acumen. Lavoisier effectively blackballed Marat's candidacy to the French Academy of Sciences, and the time came when Marat sought revenge. Marat was instrumental in getting Lavoisier and other members of the *Ferme générale* arrested on charges of counterrevolutionary activities, and on May 8, 1794, after a trial lasting less than a day, Lavoisier was guillotined. Along with Fourier and Carnot, Lavoisier's name is one of the 72 engraved on the Eiffel Tower.

so if we let the reaction go forward, we have $dN_a = \zeta_a$, and if it runs in reverse we have $dN_a = -\zeta_a$. Thus, setting $dT = dp = 0$, we have the equilibrium condition $\sum_{a=1}^{\mathsf{c}} \zeta_a \, \mu_a = 0$.

Let us investigate the consequences of this relation for ideal gases. From eqn. 2.332 and the relation $\mu_a = -T(\partial S/\partial N_a)$, the chemical potential of the $a^{\rm th}$ species is of the form

$$\mu_a(T,p) = k_{\rm B}T \log p_a + k_{\rm B}T \log \psi_a(T) \quad , \tag{2.384}$$

where each $\psi_a(T)$ has units of inverse pressure[33]. Here $p_a = p x_a$ is the partial pressure of species $a$, where $x_a = N_a / \sum_b N_b$ is its dimensionless concentration[34]. In equilibrium we must have

$$\sum_{a=1}^{\mathsf{c}} \zeta_a \, \mu_a = k_{\rm B}T \sum_a \zeta_a \Big[ \log x_a + \log p + \log \psi_a(T) \Big] = 0 \quad . \tag{2.385}$$

Exponentiating, we obtain the *law of mass action*:

$$\prod_a x_a^{\zeta_a} = \prod_a \big( p \, \psi_a(T) \big)^{-\zeta_a} \equiv \kappa(p,T) \quad . \tag{2.386}$$

The quantity $\kappa(p,T)$ is called the *equilibrium constant*. When $\kappa$ is large, the LHS of the above equation is large. This favors maximal concentration $x_a$ for the products ($\zeta_a > 0$) and minimal concentration $x_a$ for the reactants ($\zeta_a < 0$). This means that the equation $\mathsf{REACTANTS} \rightleftharpoons \mathsf{PRODUCTS}$ is shifted to the right, *i.e.* the products are plentiful and the reactants are scarce. When $\kappa$ is small, the LHS is small and the reaction is shifted to the left, *i.e.* the reactants are plentiful and the products are scarce. Remember we are describing *equilibrium conditions* here. Now we observe that reactions for which $\sum_a \zeta_a > 0$ shift to the left with increasing pressure and shift to the right with decreasing pressure, while reactions for which $\sum_a \zeta_a < 0$ the situation is reversed: they shift to the right with increasing pressure and to the left with decreasing pressure. When $\sum_a \zeta_a = 0$ there is no shift upon increasing or decreasing pressure.

The rate at which the equilibrium constant changes with temperature is given by

$$\left( \frac{\partial \log \kappa}{\partial T} \right)_p = -\sum_a \zeta_a \, \frac{\partial \log \psi_a(T)}{\partial T} \quad . \tag{2.387}$$

Now from eqn. 2.384 we have that the enthalpy per particle for species $a$ is

$$\mathsf{h}_a = \mu_a - T \left( \frac{\partial \mu_a}{\partial T} \right)_{p_a} \quad , \tag{2.388}$$

since $\mathsf{H} = G + TS$ and $S = -(\partial G/\partial T)_p$ . We find

$$\mathsf{h}_a = -k_{\rm B}T^2 \, \frac{\partial \log \psi_a(T)}{\partial T} \quad , \tag{2.389}$$

and thus

$$\left( \frac{\partial \log \kappa}{\partial T} \right)_p = \frac{\sum_i \zeta_a \, \mathsf{h}_a}{k_{\rm B}T^2} = \frac{\Delta \mathsf{h}}{k_{\rm B}T^2} \quad , \tag{2.390}$$

[33] In terms of the functions $\phi_a(T)$ from eqn. 2.332, one has $\log \psi_a(T) = -\phi_a(T)/k_{\rm B} - \log(k_{\rm B}T/e)$.

[34] Chemists sometimes write $x_a = [{\rm A}_a]$ for the concentration of species $a$.

where $\Delta\mathsf{h}$ is the enthalpy of the reaction, which is the heat absorbed or emitted as a result of the reaction.

When $\Delta\mathsf{h} > 0$ the reaction is *endothermic* and the yield increases with increasing $T$. When $\Delta\mathsf{h} < 0$ the reaction is *exothermic* and the yield decreases with increasing $T$.

As an example, consider the gas phase reaction $\mathrm{H_2} + \mathrm{I_2} \rightleftharpoons 2\,\mathrm{HI}$. We have

$$\zeta(\mathrm{H_2}) = -1 \quad , \quad \zeta(\mathrm{I_2}) = -1 \qquad\qquad \zeta(\mathrm{HI}) = 2 \quad . \tag{2.391}$$

Suppose our initial system consists of $\nu_1^0$ moles of $\mathrm{H_2}$, $\nu_2^0 = 0$ moles of $\mathrm{I_2}$, and $\nu_3^0$ moles of undissociated HI . These mole numbers determine the initial concentrations $x_a^0$, where $x_a = \nu_a / \sum_b \nu_b$ . Define

$$\alpha \equiv \frac{x_3^0 - x_3}{x_3^0} \quad , \tag{2.392}$$

in which case we have

$$x_1 = x_1^0 + \tfrac{1}{2}\alpha x_3^0 \qquad , \qquad x_2 = \tfrac{1}{2}\alpha x_3^0 \qquad , \qquad x_3 = \frac{x_3^0}{1+\alpha} \quad . \tag{2.393}$$

Then the law of mass action gives

$$\frac{x_3^2}{x_1 x_2} = \frac{4\,(1-\alpha)^2}{\alpha(\alpha + 2r)} = \kappa \quad . \tag{2.394}$$

where $r \equiv x_1^0/x_3^0 = \nu_1^0/\nu_3^0$. This yields a quadratic equation, which can be solved to find $\alpha(\kappa, r)$. Note that $\kappa = \kappa(T)$ for this reaction since $\sum_a \zeta_a = 0$. The formation of HI in this gas phase reaction is exothermic: $\Delta\mathsf{h} < 0$.

### 2.13.2 Enthalpy of formation

Most chemical reactions take place under constant pressure. The heat $Q_{\mathrm{if}}$ associated with a given isobaric process is

$$Q_{\mathrm{if}} = \int_{\mathrm{i}}^{\mathrm{f}} dE + \int_{\mathrm{i}}^{\mathrm{f}} p\,dV = (E_{\mathrm{f}} - E_{\mathrm{i}}) + p\,(V_{\mathrm{f}} - V_{\mathrm{i}}) = \mathsf{H}_{\mathrm{f}} - \mathsf{H}_{\mathrm{i}} \quad , \tag{2.395}$$

where $\mathsf{H}$ is the *enthalpy*,

$$\mathsf{H} = E + pV \quad . \tag{2.396}$$

Note that the enthalpy $\mathsf{H}$ is a state function, since $E$ is a state function and $p$ and $V$ are state variables. Hence, we can meaningfully speak of changes in enthalpy: $\Delta\mathsf{H} = \mathsf{H}_{\mathrm{f}} - \mathsf{H}_{\mathrm{i}}$. If $\Delta\mathsf{H} < 0$ for a given reaction, we call it *exothermic* – this is the case when $Q_{\mathrm{if}} < 0$ and thus heat is transferred to the surroundings. Such reactions can occur spontaneously, and, in really fun cases, can produce explosions. The combustion of fuels is always exothermic. If $\Delta\mathsf{H} > 0$, the reaction is called *endothermic*, and the system absorbs heat under conditions of constant pressure. Photosynthesis is an example of an endothermic reaction. In general, a reaction may proceed spontaneously provided that $\Delta G = \Delta\mathsf{H} - T\Delta S < 0$.

Suppose we have two reactions

$$A + B \xrightarrow{(\Delta\mathsf{H})_1} C \tag{2.397}$$

| Formula | Name | State | $\Delta H_f^0$ kJ/mol | Formula | Name | State | $\Delta H_f^0$ kJ/mol |
|---|---|---|---|---|---|---|---|
| Ag | Silver | crystal | 0.0 | $NiSO_4$ | Nickel sulfate | crystal | -872.9 |
| C | Graphite | crystal | 0.0 | $Al_2O_3$ | Aluminum oxide | crystal | -1657.7 |
| C | Diamond | crystal | 1.9 | $Ca_3(PO_4)_2$ | Calcium phosphate | solid | -4120.8 |
| $O_3$ | Ozone | gas | 142.7 | HCN | Hydrogen cyanide | liquid | 108.9 |
| $H_2O$ | Water | liquid | -285.8 | $SF_6$ | Sulfur hexafluoride | gas | -1220.5 |
| $H_3BO_3$ | Boric acid | crystal | -1094.3 | $CaF_2$ | Calcium fluoride | crystal | -1228.0 |
| $ZnSO_4$ | Zinc sulfate | crystal | -982.8 | $CaCl_2$ | Calcium chloride | crystal | -795.4 |

Table 2.5: Enthalpies of formation of some common substances.

and

$$C + D \xrightarrow{(\Delta H)_2} E \quad . \tag{2.398}$$

Then we may write

$$A + B + D \xrightarrow{(\Delta H)_3} E \quad , \tag{2.399}$$

with

$$(\Delta H)_1 + (\Delta H)_2 = (\Delta H)_3 \quad . \tag{2.400}$$

We can use this additivity of reaction enthalpies to define a *standard molar enthalpy of formation*. We first define the *reference standard state* of a pure substance at a given temperature to be its state (gas, liquid, or solid) at a pressure $p = 1\,\text{bar}$. The *standard reaction enthalpies* at a given temperature are then defined to be the reaction enthalpies when the reactants and products are all in their standard states. Finally, we define the *standard molar enthalpy of formation* $\Delta H_f^0(X)$ of a compound $X$ at temperature $T$ as the reaction enthalpy for the compound $X$ to be produced by its constituents when they are in their standard state. For example, if $X = SO_2$, then we write

$$S + O_2 \xrightarrow{\Delta H_f^0[SO_2]} SO_2 \quad . \tag{2.401}$$

The enthalpy of formation of any *element* in its reference standard state is zero at all temperatures, by definition: $\Delta H_f^0[\text{He}] = \Delta H_f^0[\text{K}] = \Delta H_f^0[\text{Mn}] = 0$, *etc.*

Suppose now we have a reaction

$$a\,A + b\,B \xrightarrow{\Delta H} c\,C + d\,D \quad . \tag{2.402}$$

To compute the reaction enthalpy $\Delta H$, we can imagine forming the components $A$ and $B$ from their standard state constituents. Similarly, we can imagine doing the same for $C$ and $D$. Since the number of atoms of a given kind is conserved in the process, the constituents of the reactants must be the same as those of the products, we have

$$\Delta H = -a\,\Delta H_f^0(A) - b\,\Delta H_f^0(B) + c\,\Delta H_f^0(C) + d\,\Delta H_f^0(D) \quad . \tag{2.403}$$

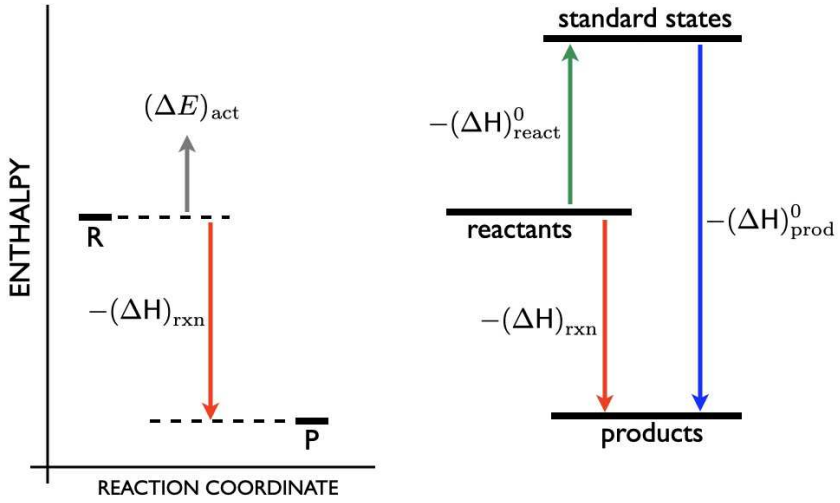



Figure 2.37: Left panel: reaction enthalpy and activation energy (exothermic case shown). Right panel: reaction enthalpy as a difference between enthalpy of formation of reactants and products.

A list of a few enthalpies of formation is provided in table 2.5. Note that the reaction enthalpy is independent of the actual reaction path. That is, the difference in enthalpy between $A$ and $B$ is the same whether the reaction is $A \longrightarrow B$ or $A \longrightarrow X \longrightarrow (Y + Z) \longrightarrow B$. This statement is known as *Hess's Law*.

Note that

$$d{\rm H} = dE + p\,dV + V\,dp = đQ + V\,dp \quad , \tag{2.404}$$

hence

$$C_p = \left(\frac{đQ}{dT}\right)_p = \left(\frac{\partial {\rm H}}{\partial T}\right)_p \quad . \tag{2.405}$$

We therefore have

$${\rm H}(T,p,\nu) = {\rm H}(T_0,p,\nu) + \nu \int\limits_{T_0}^{T} dT'\, c_p(T') \quad . \tag{2.406}$$

For ideal gases, we have $c_p(T) = (1+\frac{1}{2}f)\,R$. For real gases, over a range of temperatures, there are small variations:

$$c_p(T) = \alpha + \beta\,T + \gamma\,T^2 \quad . \tag{2.407}$$

Two examples ($300\,{\rm K} < T < 1500\,{\rm K}$, $p = 1\,{\rm atm}$):

$$\begin{aligned} {\rm O_2}: \quad & \alpha = 25.503\,\frac{\rm J}{\rm mol\,K} \quad , \quad \beta = 13.612\times 10^{-3}\,\frac{\rm J}{\rm mol\,K^2} \quad , \quad \gamma = -42.553\times 10^{-7}\,\frac{\rm J}{\rm mol\,K^3} \\ {\rm H_2O}: \quad & \alpha = 30.206\,\frac{\rm J}{\rm mol\,K} \quad , \quad \beta = 9.936\times 10^{-3}\,\frac{\rm J}{\rm mol\,K^2} \quad , \quad \gamma = 11.14\times 10^{-7}\,\frac{\rm J}{\rm mol\,K^3} \end{aligned}$$

If all the gaseous components in a reaction can be approximated as ideal, then we may write

$$(\Delta {\rm H})_{\rm rxn} = (\Delta E)_{\rm rxn} + \sum_a \zeta_a\,RT \quad , \tag{2.408}$$

| bond | enthalpy (kJ/mol) | bond | enthalpy (kJ/mol) | bond | enthalpy (kJ/mol) | bond | enthalpy (kJ/mol) |
|---|---|---|---|---|---|---|---|
| H − H | 436 | C − C | 348 | C − S | 259 | F − F | 155 |
| H − C | 412 | C = C | 612 | N − N | 163 | F − Cl | 254 |
| H − N | 388 | C ≡ C | 811 | N = N | 409 | Cl − Br | 219 |
| H − O | 463 | C − N | 305 | N ≡ N | 945 | Cl − I | 210 |
| H − F | 565 | C = N | 613 | N − O | 157 | Cl − S | 250 |
| H − Cl | 431 | C ≡ N | 890 | N − F | 270 | Br − Br | 193 |
| H − Br | 366 | C − O | 360 | N − Cl | 200 | Br − I | 178 |
| H − I | 299 | C = O | 743 | N − Si | 374 | Br − S | 212 |
| H − S | 338 | C − F | 484 | O − O | 146 | I − I | 151 |
| H − P | 322 | C − Cl | 338 | O = O | 497 | S − S | 264 |
| H − Si | 318 | C − Br | 276 | O − F | 185 | P − P | 172 |
| | | C − I | 238 | O − Cl | 203 | Si − Si | 176 |

Table 2.6: Average bond enthalpies for some common bonds. (*Credit: L. Pauling, The Nature of the Chemical Bond (Cornell Univ. Press, NY, 1960).)*

where the subscript 'rxn' stands for 'reaction'. Here $(\Delta E)_{\rm rxn}$ is the change in energy from reactants to products.

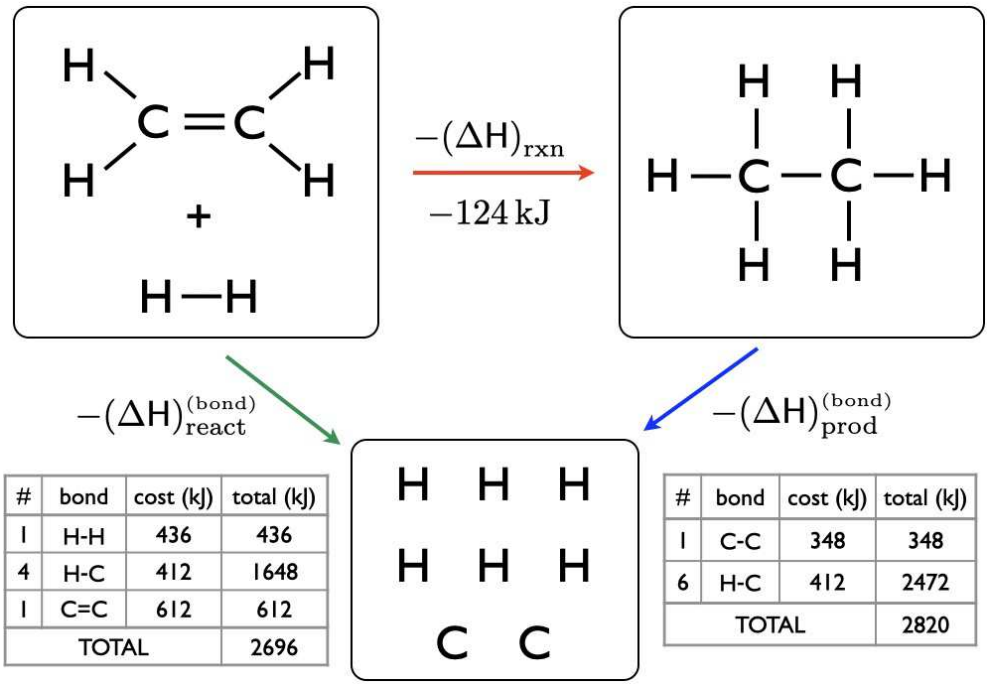
### 2.13.3 Bond enthalpies

The enthalpy needed to break a chemical bond is called the *bond enthalpy*, $\mathsf{h}[\bullet]$. The bond enthalpy is the energy required to dissociate one mole of gaseous bonds to form gaseous atoms. A table of bond enthalpies is given in Tab. 2.6. Bond enthalpies are endothermic, since energy is required to break chemical bonds. Of course, the actual bond energies can depend on the location of a bond in a given molecule, and the values listed in the table reflect averages over the possible bond environment.

The bond enthalpies in Tab. 2.6 may be used to compute reaction enthalpies. Consider, for example, the reaction $2\,{\rm H_2(g)} + {\rm O_2(g)} \longrightarrow 2\,{\rm H_2O(l)}$. We then have, from the table,

$$\begin{aligned}(\Delta\mathsf{H})_{\rm rxn} &= 2\,\mathsf{h}[{\rm H-H}] + \mathsf{h}[{\rm O=O}] - 4\,\mathsf{h}[{\rm H-O}] \\ &= -483\,{\rm kJ/mol\ O_2} \quad .\end{aligned} \tag{2.409}$$

Thus, 483 kJ of heat would be released for every two moles of $\rm H_2O$ produced, if the $\rm H_2O$ were in the gaseous phase. Since $\rm H_2O$ is liquid at STP, we should also include the condensation energy of the gaseous water vapor into liquid water. At $T = 100^\circ{\rm C}$ the latent heat of vaporization is $\tilde\ell = 2270\,{\rm J/g}$, but at $T = 20^\circ{\rm C}$, one has $\tilde\ell = 2450\,{\rm J/g}$, hence with $M = 18$ we have $\ell = 44.1\,{\rm kJ/mol}$. Therefore, the heat produced by the reaction $2\,{\rm H_2(g)} + {\rm O_2(g)} \rightleftharpoons 2\,{\rm H_2O(l)}$ is $(\Delta H)_{\rm rxn} = -571.2\,{\rm kJ\,/\,mol\ O_2}$. Since the




| # | bond | cost (kJ) | total (kJ) |
|---|---|---|---|
| 1 | H-H | 436 | 436 |
| 4 | H-C | 412 | 1648 |
| 1 | C=C | 612 | 612 |
| TOTAL | | | 2696 |

| # | bond | cost (kJ) | total (kJ) |
|---|---|---|---|
| 1 | C-C | 348 | 348 |
| 6 | H-C | 412 | 2472 |
| TOTAL | | | 2820 |

Figure 2.38: Calculation of reaction enthalpy for the hydrogenation of ethene (ethylene), $C_2H_4$.

reaction produces two moles of water, we conclude that the enthalpy of formation of liquid water at STP is half this value: $\Delta H_f^0[H_2O] = -285.6\,\mathrm{kJ/mol}$.

### Hydrogenation of ethene

*Hydrogenation* is the adding of hydrogen to a carbon-carbon double bond. Consider the hydrogenation of ethene (ethylene): $C_2H_4 + H_2 \rightleftharpoons C_2H_6$. The product is known as ethane. The energy accounting is shown in fig. 2.38. To compute the enthalpies of formation of ethene and ethane from the bond enthalpies, we need one more bit of information, which is the standard enthalpy of formation of $C(g)$ from $C(s)$, since the solid is the standard state at STP. This value is $\Delta H_f^0[C(g)] = 718\,\mathrm{kJ/mol}$. We may now write

$$\begin{aligned} 2\,C(g) + 4\,H(g) &\xrightarrow{-2260\,\mathrm{kJ}} C_2H_4(g) \\ 2\,C(s) &\xrightarrow{1436\,\mathrm{kJ}} 2\,C(g) \\ 2\,H_2(g) &\xrightarrow{872\,\mathrm{kJ}} 4\,H(g) \quad . \end{aligned}$$

Thus, using Hess's law, *i.e.* adding up these reaction equations, we have

$$2\,C(s) + 2\,H_2(g) \xrightarrow{48\,\mathrm{kJ}} C_2H_4(g) \quad .$$

Thus, the formation of ethene is endothermic. For ethane,

$$\begin{aligned} 2\,\mathrm{C(g)} + 6\,\mathrm{H(g)} &\xrightarrow{-2820\,\mathrm{kJ}} \mathrm{C_2H_6(g)} \\ 2\,\mathrm{C(s)} &\xrightarrow{1436\,\mathrm{kJ}} 2\,\mathrm{C(g)} \\ 3\,\mathrm{H_2(g)} &\xrightarrow{1308\,\mathrm{kJ}} 6\,\mathrm{H(g)} \end{aligned}$$

Add 'em up:

$$2\,\mathrm{C(s)} + 3\,\mathrm{H_2(g)} \xrightarrow{-76\,\mathrm{kJ}} \mathrm{C_2H_6(g)} \quad ,$$

which is exothermic. Thus,

$$\mathrm{C_2H_4(g)} + \mathrm{H_2(g)} \xrightarrow{-48\,\mathrm{kJ}} 2\,\mathrm{C(s)} + 3\,\mathrm{H_2(g)} \xrightarrow{-76\,\mathrm{kJ}} \mathrm{C_2H_6(g)} \quad , \tag{2.410}$$

which says

$$\mathrm{C_2H_4(g)} + \mathrm{H_2(g)} \xrightarrow{-124\,\mathrm{kJ}} \mathrm{C_2H_6(g)} \quad . \tag{2.411}$$

The hydrogenation of ethene is thus exothermic. However, there is a high activation energy associated with the intermediate state which prevents the reaction from taking place under normal conditions. This may be circumvented, however, through use of a catalyst, such as a Pt, Pd, or Ni surface.

## 2.14 Appendix I: Integrating Factors

Suppose we have an inexact differential

$$đW = A_i\,dx_i \quad . \tag{2.412}$$

Here I am adopting the 'Einstein convention' where we sum over repeated indices unless otherwise explicitly stated; $A_i\,dx_i = \sum_i A_i\,dx_i$. An *integrating factor* $\exp\bigl[L(\boldsymbol{x})\bigr]$ is a function which, when divided into $đW$, yields an exact differential:

$$dU = e^{-L}\,đW = \frac{\partial U}{\partial x_i}\,dx_i \quad . \tag{2.413}$$

Clearly we must have

$$\frac{\partial^2 U}{\partial x_i\,\partial x_j} = \frac{\partial}{\partial x_i}\left(e^{-L}\,A_j\right) = \frac{\partial}{\partial x_j}\left(e^{-L}\,A_i\right) \quad . \tag{2.414}$$

Applying the Leibniz rule and then multiplying by $e^L$ yields

$$\frac{\partial A_j}{\partial x_i} - A_j\,\frac{\partial L}{\partial x_i} = \frac{\partial A_i}{\partial x_j} - A_i\,\frac{\partial L}{\partial x_j} \quad . \tag{2.415}$$

If there are $K$ independent variables $\{x_1, \ldots, x_K\}$, then there are $\frac{1}{2}K(K-1)$ independent equations of the above form – one for each distinct $(i,j)$ pair. These equations can be written compactly as

$$\Omega_{ijk}\,\frac{\partial L}{\partial x_k} = F_{ij} \quad , \tag{2.416}$$

where

$$\Omega_{ijk} = A_j\,\delta_{ik} - A_i\,\delta_{jk} \tag{2.417}$$

$$F_{ij} = \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \quad . \tag{2.418}$$

Note that $F_{ij}$ is antisymmetric, and resembles a field strength tensor, and that $\Omega_{ijk} = -\Omega_{jik}$ is antisymmetric in the first two indices (but is not totally antisymmetric in all three).

Can we solve these $\frac{1}{2}K(K-1)$ coupled equations to find an integrating factor $L$? In general the answer is *no*. However, when $K = 2$ we can always find an integrating factor. To see why, let's call $x \equiv x_1$ and $y \equiv x_2$. Consider now the ODE

$$\frac{dy}{dx} = -\frac{A_x(x,y)}{A_y(x,y)} \quad . \tag{2.419}$$

This equation can be integrated to yield a one-parameter set of integral curves, indexed by an initial condition. The equation for these curves may be written as $U_c(x,y) = 0$, where $c$ labels the curves. Then along each curve we have

$$\begin{aligned} 0 = \frac{dU_c}{dx} &= \frac{\partial U_c}{\partial x} + \frac{\partial U_c}{\partial y}\,\frac{dy}{dx} \\ &= \frac{\partial U_c}{\partial x} - \frac{A_x}{A_y}\,\frac{\partial U_c}{\partial y} \quad . \end{aligned} \tag{2.420}$$

Thus,

$$\frac{\partial U_c}{\partial x}\,A_y = \frac{\partial U_c}{\partial y}\,A_x \equiv e^{-L} A_x\,A_y \quad . \tag{2.421}$$

This equation defines the integrating factor $L$:

$$L = -\log\left(\frac{1}{A_x}\,\frac{\partial U_c}{\partial x}\right) = -\log\left(\frac{1}{A_y}\,\frac{\partial U_c}{\partial y}\right) \quad . \tag{2.422}$$

We now have that

$$A_x = e^L\,\frac{\partial U_c}{\partial x} \qquad , \qquad A_y = e^L\,\frac{\partial U_c}{\partial y} \quad , \tag{2.423}$$

and hence

$$e^{-L}\,đW = \frac{\partial U_c}{\partial x}\,dx + \frac{\partial U_c}{\partial y}\,dy = dU_c \quad . \tag{2.424}$$

## 2.15 Appendix II: Other Thermodynamic Cycles

### 2.15.1 The Stirling cycle

In §2.5.4 we analyzed the famous Carnot cycle. Many other thermodynamic cycles are commonly used. The Stirling cycle, depicted in fig. 2.39, consists of two isotherms and two isochores. Recall the isothermal ideal gas equation of state, $d(pV) = 0$. Thus, for an ideal gas Stirling cycle, we have

$$p_{\rm A} V_1 = p_{\rm B} V_2 \qquad , \qquad p_{\rm D} V_1 = p_{\rm C} V_2 \quad , \tag{2.425}$$

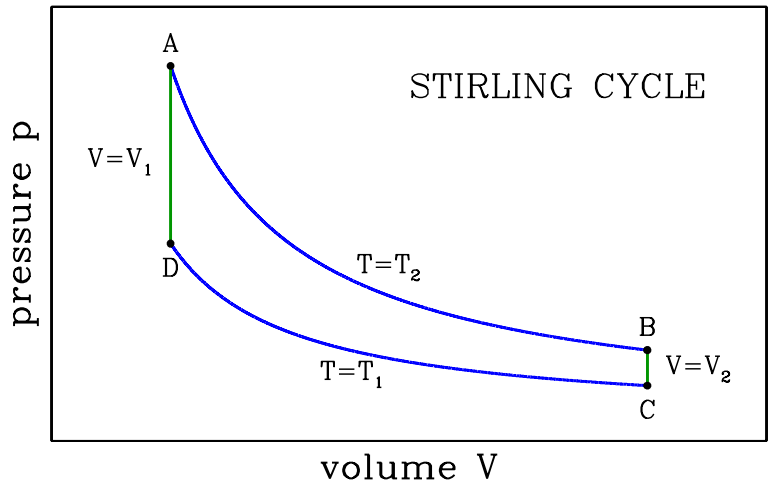



Figure 2.39: A Stirling cycle consists of two isotherms (blue) and two isochores (green).

which says

$$\frac{p_{\rm B}}{p_{\rm A}} = \frac{p_{\rm C}}{p_{\rm D}} = \frac{V_1}{V_2} \quad . \tag{2.426}$$

AB: This isothermal expansion is the power stroke. Assuming $\nu$ moles of ideal gas throughout, we have $pV = \nu RT_2 = p_{\rm A}V_1$, hence

$$W_{\rm AB} = \int_{V_1}^{V_2} dV\, \frac{\nu RT_2}{V} = \nu RT_2\, \log\left(\frac{V_2}{V_1}\right) \quad . \tag{2.427}$$

Since AB is an isotherm, we have $E_{\rm A} = E_{\rm B}$, and from $\Delta E_{\rm AB} = 0$ we conclude $Q_{\rm AB} = W_{\rm AB}$.

BC: Isochoric cooling. Since $dV = 0$ we have $W_{\rm BC} = 0$. The energy change is given by

$$\Delta E_{\rm BC} = E_{\rm C} - E_{\rm B} = \frac{\nu R(T_1 - T_2)}{\gamma - 1}, \tag{2.428}$$

which is negative. Since $W_{\rm BC} = 0$, we have $Q_{\rm BC} = \Delta E_{\rm BC}$.

CD: Isothermal compression. Clearly

$$W_{\rm CD} = \int_{V_2}^{V_1} dV\, \frac{\nu RT_1}{V} = -\nu RT_1\, \log\left(\frac{V_2}{V_1}\right) \quad . \tag{2.429}$$

Since CD is an isotherm, we have $E_{\rm C} = E_{\rm D}$, and from $\Delta E_{\rm CD} = 0$ we conclude $Q_{\rm CD} = W_{\rm CD}$.

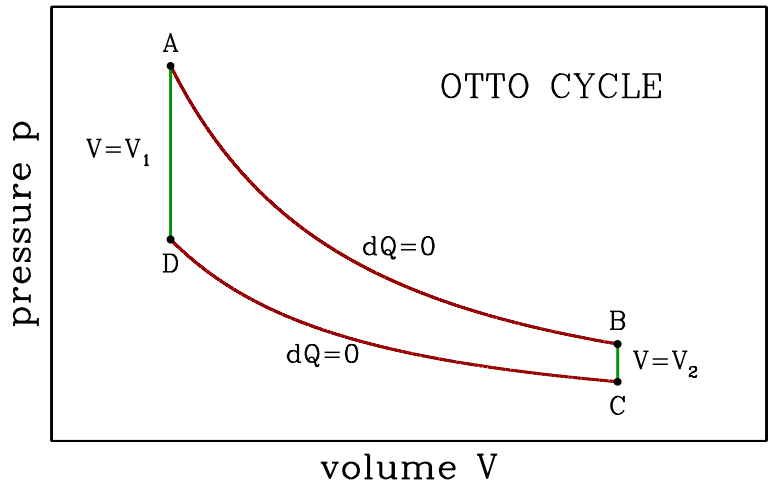



Figure 2.40: An Otto cycle consists of two adiabats (dark red) and two isochores (green).

DA: Isochoric heating. Since $dV = 0$ we have $W_{\rm DA} = 0$. The energy change is given by

$$\Delta E_{\rm DA} = E_{\rm A} - E_{\rm D} = \frac{\nu R(T_2 - T_1)}{\gamma - 1} \quad , \tag{2.430}$$

which is positive, and opposite to $\Delta E_{\rm BC}$. Since $W_{\rm DA} = 0$, we have $Q_{\rm DA} = \Delta E_{\rm DA}$.

We now add up all the work contributions to obtain

$$\begin{aligned} W &= W_{\rm AB} + W_{\rm BC} + W_{\rm CD} + W_{\rm DA} \\ &= \nu R(T_2 - T_1)\, \log\left(\frac{V_2}{V_1}\right) \quad . \end{aligned} \tag{2.431}$$

The cycle efficiency is once again

$$\eta = \frac{W}{Q_{\rm AB}} = 1 - \frac{T_1}{T_2} \quad . \tag{2.432}$$

### 2.15.2 The Otto and Diesel cycles

The Otto cycle is a rough approximation to the physics of a gasoline engine. It consists of two adiabats and two isochores, and is depicted in fig. 2.40. Assuming an ideal gas, along the adiabats we have $d(pV^\gamma) = 0$. Thus,

$$p_{\rm A}\, V_1^\gamma = p_{\rm B}\, V_2^\gamma \qquad , \qquad p_{\rm D}\, V_1^\gamma = p_{\rm C}\, V_2^\gamma \quad , \tag{2.433}$$

which says

$$\frac{p_{\rm B}}{p_{\rm A}} = \frac{p_{\rm C}}{p_{\rm D}} = \left(\frac{V_1}{V_2}\right)^\gamma \quad . \tag{2.434}$$

AB: Adiabatic expansion, the power stroke. The heat transfer is $Q_{\rm AB} = 0$, so from the First Law we have $W_{\rm AB} = -\Delta E_{\rm AB} = E_{\rm A} - E_{\rm B}$, thus

$$W_{\rm AB} = \frac{p_{\rm A}V_1 - p_{\rm B}V_2}{\gamma - 1} = \frac{p_{\rm A}V_1}{\gamma - 1}\left[1 - \left(\frac{V_1}{V_2}\right)^{\gamma-1}\right] \quad . \tag{2.435}$$

Note that this result can also be obtained from the adiabatic equation of state $pV^\gamma = p_{\rm A}V_1^\gamma$:

$$W_{\rm AB} = \int_{V_1}^{V_2} p\,dV = p_{\rm A}V_1^\gamma \int_{V_1}^{V_2} dV\,V^{-\gamma} = \frac{p_{\rm A}V_1}{\gamma - 1}\left[1 - \left(\frac{V_1}{V_2}\right)^{\gamma-1}\right] \quad . \tag{2.436}$$

BC: Isochoric cooling (exhaust); $dV = 0$ hence $W_{\rm BC} = 0$. The heat $Q_{\rm BC}$ absorbed is then

$$Q_{\rm BC} = E_{\rm C} - E_{\rm B} = \frac{V_2}{\gamma - 1}\left(p_{\rm C} - p_{\rm B}\right) \quad . \tag{2.437}$$

In a realistic engine, this is the stage in which the old burned gas is ejected and new gas is inserted.

CD: Adiabatic compression; $Q_{\rm CD} = 0$ and $W_{\rm CD} = E_{\rm C} - E_{\rm D}$:

$$W_{\rm CD} = \frac{p_{\rm C}V_2 - p_{\rm D}V_1}{\gamma - 1} = -\frac{p_{\rm D}V_1}{\gamma - 1}\left[1 - \left(\frac{V_1}{V_2}\right)^{\gamma-1}\right] \quad . \tag{2.438}$$

DA: Isochoric heating, *i.e.* the combustion of the gas. As with BC we have $dV = 0$, and thus $W_{\rm DA} = 0$. The heat $Q_{\rm DA}$ absorbed by the gas is then

$$Q_{\rm DA} = E_{\rm A} - E_{\rm D} = \frac{V_1}{\gamma - 1}\left(p_{\rm A} - p_{\rm D}\right) \quad . \tag{2.439}$$

The total work done per cycle is then

$$\begin{aligned} W &= W_{\rm AB} + W_{\rm BC} + W_{\rm CD} + W_{\rm DA} \\ &= \frac{(p_{\rm A} - p_{\rm D})V_1}{\gamma - 1}\left[1 - \left(\frac{V_1}{V_2}\right)^{\gamma-1}\right] \quad , \end{aligned} \tag{2.440}$$

and the efficiency is defined to be

$$\eta \equiv \frac{W}{Q_{\rm DA}} = 1 - \left(\frac{V_1}{V_2}\right)^{\gamma-1} \quad . \tag{2.441}$$

The ratio $V_2/V_1$ is called the *compression ratio*. We can make our Otto cycle more efficient simply by increasing the compression ratio. The problem with this scheme is that if the fuel mixture becomes too hot, it will spontaneously 'preignite', and the pressure will jump up before point D in the cycle is reached. A Diesel engine avoids preignition by compressing the air only, and then later spraying the fuel

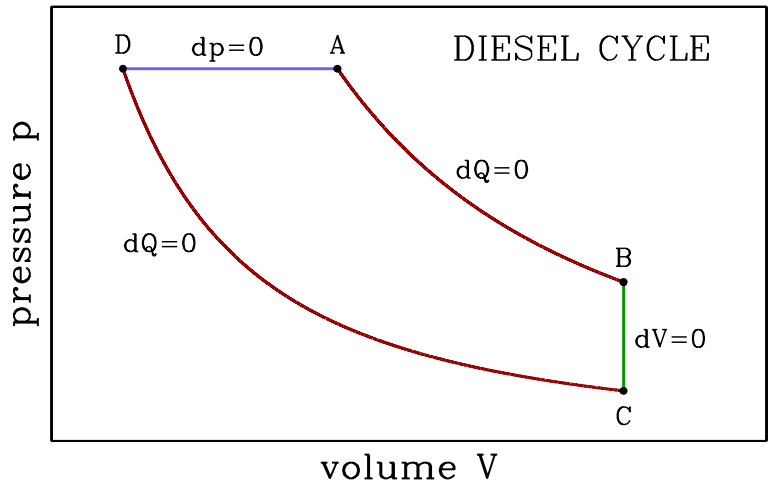



Figure 2.41: A Diesel cycle consists of two adiabats (dark red), one isobar (light blue), and one isochore (green).

into the cylinder when the air temperature is sufficient for fuel ignition. The rate at which fuel is injected is adjusted so that the ignition process takes place at constant pressure. Thus, in a Diesel engine, step DA is an isobar. The compression ratio is $r \equiv V_{\rm B}/V_{\rm D}$, and the *cutoff ratio* is $s \equiv V_{\rm A}/V_{\rm D}$. This refinement of the Otto cycle allows for higher compression ratios (of about 20) in practice, and greater engine efficiency.

For the Diesel cycle, we have, briefly,

$$\begin{aligned} W &= p_{\rm A}(V_{\rm A} - V_{\rm D}) + \frac{p_{\rm A}V_{\rm A} - p_{\rm B}V_{\rm B}}{\gamma - 1} + \frac{p_{\rm C}V_{\rm C} - p_{\rm D}V_{\rm D}}{\gamma - 1} \\ &= \frac{\gamma\, p_{\rm A}(V_{\rm A} - V_{\rm D})}{\gamma - 1} - \frac{(p_{\rm B} - p_{\rm C})V_{\rm B}}{\gamma - 1} \end{aligned} \tag{2.442}$$

and

$$Q_{\rm DA} = \frac{\gamma\, p_{\rm A}(V_{\rm A} - V_{\rm D})}{\gamma - 1} \quad . \tag{2.443}$$

To find the efficiency, we will need to eliminate $p_{\rm B}$ and $p_{\rm C}$ in favor of $p_{\rm A}$ using the adiabatic equation of state $d(pV^\gamma) = 0$. Thus,

$$p_{\rm B} = p_{\rm A} \cdot \left(\frac{V_{\rm A}}{V_{\rm B}}\right)^\gamma \quad , \quad p_{\rm C} = p_{\rm A} \cdot \left(\frac{V_{\rm D}}{V_{\rm B}}\right)^\gamma \quad , \tag{2.444}$$

where we've used $p_{\rm D} = p_{\rm A}$ and $V_{\rm C} = V_{\rm B}$. Putting it all together, the efficiency of the Diesel cycle is

$$\eta = \frac{W}{Q_{\rm DA}} = 1 - \frac{1}{\gamma}\,\frac{r^{1-\gamma}(s^\gamma - 1)}{s - 1} \quad . \tag{2.445}$$

### 2.15.3 The Joule-Brayton cycle

Our final example is the Joule-Brayton cycle, depicted in fig. 2.42, consisting of two adiabats and two isobars. Along the adiabats we have

$$p_2\, V_{\rm A}^\gamma = p_1\, V_{\rm D}^\gamma \quad , \quad p_2\, V_{\rm B}^\gamma = p_1\, V_{\rm C}^\gamma \quad , \tag{2.446}$$

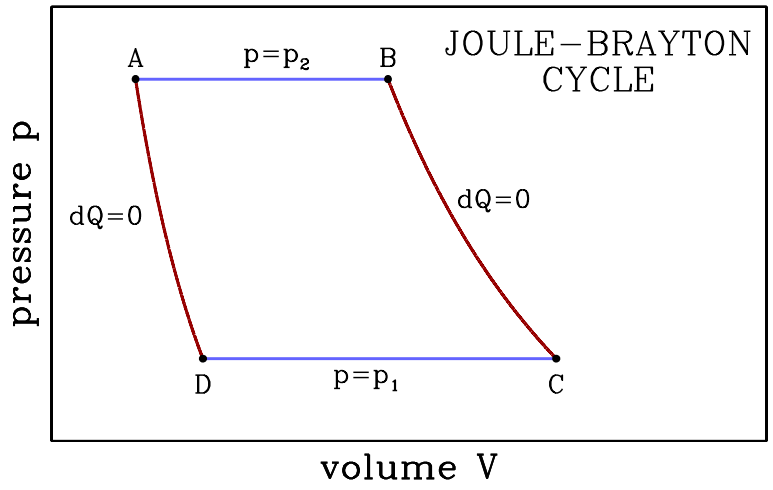



Figure 2.42: A Joule-Brayton cycle consists of two adiabats (dark red) and two isobars (light blue).

which says

$$\frac{V_{\rm D}}{V_{\rm A}} = \frac{V_{\rm C}}{V_{\rm B}} = \left(\frac{p_2}{p_1}\right)^{1/\gamma} \quad . \tag{2.447}$$

AB: This isobaric expansion at $p = p_2$ is the power stroke. We have

$$W_{\rm AB} = \int_{V_{\rm A}}^{V_{\rm B}} dV\, p_2 = p_2\,(V_{\rm B} - V_{\rm A}) \tag{2.448}$$

$$\Delta E_{\rm AB} = E_{\rm B} - E_{\rm A} = \frac{p_2\,(V_{\rm B} - V_{\rm A})}{\gamma - 1} \tag{2.449}$$

$$Q_{\rm AB} = \Delta E_{\rm AB} + W_{\rm AB} = \frac{\gamma\, p_2\,(V_{\rm B} - V_{\rm A})}{\gamma - 1} \quad . \tag{2.450}$$

BC: Adiabatic expansion; $Q_{\rm BC} = 0$ and $W_{\rm BC} = E_{\rm B} - E_{\rm C}$. The work done by the gas is

$$\begin{aligned} W_{\rm BC} &= \frac{p_2 V_{\rm B} - p_1 V_{\rm C}}{\gamma - 1} = \frac{p_2 V_{\rm B}}{\gamma - 1}\left(1 - \frac{p_1}{p_2}\cdot\frac{V_{\rm C}}{V_{\rm B}}\right) \\ &= \frac{p_2\, V_{\rm B}}{\gamma - 1}\left[1 - \left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}}\right] \quad . \end{aligned} \tag{2.451}$$

CD: Isobaric compression at $p = p_1$.

$$W_{\rm CD} = \int_{V_{\rm C}}^{V_{\rm D}} dV\, p_1 = p_1\,(V_{\rm D} - V_{\rm C}) = -p_2\,(V_{\rm B} - V_{\rm A})\left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}} \tag{2.452}$$

$$\Delta E_{\rm CD} = E_{\rm D} - E_{\rm C} = \frac{p_1\,(V_{\rm D} - V_{\rm C})}{\gamma - 1} \tag{2.453}$$

$$Q_{\rm CD} = \Delta E_{\rm CD} + W_{\rm CD} = -\frac{\gamma\, p_2}{\gamma - 1}\,(V_{\rm B} - V_{\rm A})\left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}} \quad . \tag{2.454}$$

DA: Adiabatic compression; $Q_{\rm DA} = 0$ and $W_{\rm DA} = E_{\rm D} - E_{\rm A}$. The work done by the gas is

$$\begin{aligned} W_{\rm DA} &= \frac{p_1 V_{\rm D} - p_2 V_{\rm A}}{\gamma - 1} = -\frac{p_2 V_{\rm A}}{\gamma - 1}\left(1 - \frac{p_1}{p_2}\cdot\frac{V_{\rm D}}{V_{\rm A}}\right) \\ &= -\frac{p_2\, V_{\rm A}}{\gamma - 1}\left[1 - \left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}}\right] \quad . \end{aligned} \tag{2.455}$$

The total work done per cycle is then

$$\begin{aligned} W &= W_{\rm AB} + W_{\rm BC} + W_{\rm CD} + W_{\rm DA} \\ &= \frac{\gamma\, p_2\,(V_{\rm B} - V_{\rm A})}{\gamma - 1}\left[1 - \left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}}\right] \end{aligned} \tag{2.456}$$

and the efficiency is defined to be

$$\eta \equiv \frac{W}{Q_{\rm AB}} = 1 - \left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}} \quad . \tag{2.457}$$

## 2.16 Appendix III: Legendre Transformations

A *strictly convex function* of a single variable $f(x)$ is one for which $f''(x) > 0$ everywhere. The *Legendre transform* of a strictly convex function $f(x)$ is a function $g(p)$ defined as follows. Let $p$ be a real number, and consider the line $y = px$, as shown in fig. 2.43. We define the point $x(p)$ as the value of $x$ for which the difference $F(x,p) = px - f(x)$ is greatest. Then define $g(p) = F\big(x(p), p\big)$.[35] The value $x(p)$ is unique if $f(x)$ is strictly convex, since $x(p)$ is determined by the equation

$$f'\big(x(p)\big) = p \quad . \tag{2.458}$$

Note that from $p = f'\big(x(p)\big)$ we have, according to the chain rule,

$$1 = \frac{d}{dp}\, f'\big(x(p)\big) = f''\big(x(p)\big)\, x'(p) \qquad \Longrightarrow \qquad x'(p) = \Big[f''\big(x(p)\big)\Big]^{-1} \quad . \tag{2.459}$$

[35] Note that $g(p)$ may be a negative number, if the line $y = px$ lies everywhere below $f(x)$.

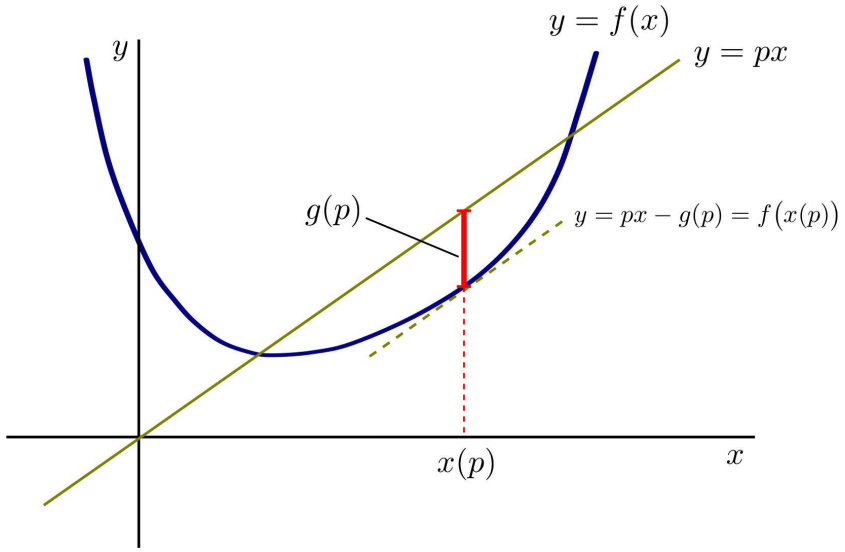



Figure 2.43: Construction for the Legendre transformation of a function $f(x)$.

From this, we can prove that $g(p)$ is itself strictly convex:

$$\begin{aligned} g'(p) &= \frac{d}{dp}\Big[p\,x(p) - f\big(x(p)\big)\Big] \\ &= p\,x'(p) + x(p) - f'\big(x(p)\big)\,x'(p) = x(p) \quad , \end{aligned} \tag{2.460}$$

hence

$$g''(p) = x'(p) = \Big[f''\big(x(p)\big)\Big]^{-1} > 0 \quad . \tag{2.461}$$

In higher dimensions, the generalization of the definition $f''(x) > 0$ is that a function $F(x_1, \ldots, x_n)$ is strictly convex if the matrix of second derivatives, called the *Hessian*,

$$H_{ij}(\boldsymbol{x}) = \frac{\partial^2 F}{\partial x_i\,\partial x_j} \tag{2.462}$$

is positive definite. That is, all the eigenvalues of $H_{ij}(\boldsymbol{x})$ must be positive for every $\boldsymbol{x}$. We then define the Legendre transform $\boldsymbol{G}(\boldsymbol{p})$ as

$$\boldsymbol{G}(\boldsymbol{p}) = \boldsymbol{p}\cdot\boldsymbol{x} - F(\boldsymbol{x}) \tag{2.463}$$

where $\boldsymbol{p} = \boldsymbol{\nabla} F$. Note that

$$dG = \boldsymbol{x}\cdot d\boldsymbol{p} + \boldsymbol{p}\cdot d\boldsymbol{x} - \boldsymbol{\nabla}F\cdot d\boldsymbol{x} = \boldsymbol{x}\cdot d\boldsymbol{p} \quad , \tag{2.464}$$

which establishes that $G$ is a function of $\boldsymbol{p}$ and that

$$\frac{\partial G}{\partial p_j} = x_j \quad . \tag{2.465}$$

Note also that the Legendre transformation is *self-dual*, which is to say that the Legendre transform of $G(\boldsymbol{p})$ is $F(\boldsymbol{x})$: $F \to G \to F$ under successive Legendre transformations.

### 2.16.1 Partial Legendre transformation

We can also define a *partial Legendre transformation* as follows. Consider a function of $q$ variables $F(\boldsymbol{x}, \boldsymbol{y})$, where $\boldsymbol{x} = \{x_1, \ldots, x_m\}$ and $\boldsymbol{y} = \{y_1, \ldots, y_n\}$, with $q = m + n$. Define $\boldsymbol{p} = \{p_1, \ldots, p_m\}$, and

$$G(\boldsymbol{p}, \boldsymbol{y}) = \boldsymbol{p} \cdot \boldsymbol{x} - F(\boldsymbol{x}, \boldsymbol{y}) \quad , \tag{2.466}$$

where

$$p_a = p_a(\boldsymbol{x}, \boldsymbol{y}) = \left( \frac{\partial F}{\partial x_a} \right)_{\boldsymbol{y}} \quad . \tag{2.467}$$

Note that we may also write

$$x_a = x_a(\boldsymbol{p}, \boldsymbol{y}) = \left( \frac{\partial G}{\partial p_a} \right)_{\boldsymbol{y}} \quad . \tag{2.468}$$

For the full Hessian of $F$, write

$$H = \begin{pmatrix} K & B \\ B^{\mathsf{T}} & L \end{pmatrix} \quad , \tag{2.469}$$

where

$$K_{ab} = \frac{\partial^2 F}{\partial x_a \partial x_b} \quad , \quad B_{a\nu} = \frac{\partial^2 F}{\partial x_a \partial y_\nu} \quad , \quad L_{\mu\nu} = \frac{\partial^2 F}{\partial y_\mu \partial y_\nu} \quad . \tag{2.470}$$

The variation of $x^a$ at fixed $\boldsymbol{p}$ is then given by setting

$$0 = \left( \frac{\partial p_a}{\partial y_\nu} \right)_{\boldsymbol{p}} = K_{ab} \left( \frac{\partial x_b}{\partial y_\nu} \right)_{\boldsymbol{p}} + B_{a\nu} \quad \Rightarrow \quad \left( \frac{\partial x_a}{\partial y_\nu} \right)_{\boldsymbol{p}} = -K^{-1}_{ab} B_{b\nu} \quad . \tag{2.471}$$

Now from eqn. 2.466 we have

$$dG = x_a \, dp_a + \overbrace{p_a \, dx_a - \frac{\partial F}{\partial x_a} \, dx_a}^{\text{these cancel}} - \frac{\partial F}{\partial y_\nu} \, dy_\nu = x_a \, dp_a - \frac{\partial F}{\partial y_\nu} \, dy_\nu \quad , \tag{2.472}$$

and setting $d\boldsymbol{p} = 0$ we obtain

$$\frac{\partial G(\boldsymbol{p}, \boldsymbol{y})}{\partial y_\nu} = - \frac{\partial F(\boldsymbol{x}, \boldsymbol{y})}{\partial y_\nu} \bigg|_{\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{p}, \boldsymbol{y})} \quad . \tag{2.473}$$

We then have

$$\frac{\partial^2 G}{\partial y_\mu \partial y_\nu} = - \frac{\partial^2 F}{\partial y_\mu \partial y_\nu} - \frac{\partial^2 F}{\partial x_a \partial y_\mu} \left( \frac{\partial x_a}{\partial y_\nu} \right)_{\boldsymbol{p}} \tag{2.474}$$

and thus

$$G_{\mu\nu} = -L_{\mu\nu} - B^{\mathsf{T}}_{\mu a} \left( \frac{\partial x_a}{\partial y_\nu} \right)_{\boldsymbol{p}} \quad . \tag{2.475}$$

Now invoking eqn. 2.471 we arrive at our final result relating the Hessian blocks,

$$\frac{\partial^2 G}{\partial y_\mu \partial y_\nu} = \left( B^{\mathsf{T}} K^{-1} B - L \right)_{\mu\nu} \quad . \tag{2.476}$$

Note the dimensions: $K^{m\times m}$, $L^{n\times n}$, and $B^{m\times n}$.

It is easy to see that if the full $q \times q$ Hessian matrix $H_{ij}$ is positive definite, then any diagonal subblock such as $K_{ab}$ or $L_{\mu\nu}$ must also be positive definite. In this case, the partial Legendre transform is convex in $\{p_1, \ldots, p_m\}$ and concave in $\{y_1, \ldots, y_n\}$.

## 2.17 Appendix IV: Useful Mathematical Relations

Consider a set of $n$ independent variables $\{x_1, \ldots, x_n\}$, which can be thought of as a point in $n$-dimensional space. Let $\{y_1, \ldots, y_n\}$ and $\{z_1, \ldots, z_n\}$ be other choices of coordinates. Then

$$\frac{\partial x_i}{\partial z_k} = \frac{\partial x_i}{\partial y_j}\,\frac{\partial y_j}{\partial z_k} \quad . \tag{2.477}$$

Note that this entails a matrix multiplication: $A_{ik} = B_{ij}\,C_{jk}$, where $A_{ik} = \partial x_i/\partial z_k$, $B_{ij} = \partial x_i/\partial y_j$, and $C_{jk} = \partial y_j/\partial z_k$. We define the determinant

$$\det\left(\frac{\partial x_i}{\partial z_k}\right) \equiv \frac{\partial(x_1, \ldots, x_n)}{\partial(z_1, \ldots, z_n)} \quad . \tag{2.478}$$

Such a determinant is called a Jacobian. Now if $A = BC$, then $\det(A) = \det(B) \cdot \det(C)$. Thus,

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(z_1, \ldots, z_n)} = \frac{\partial(x_1, \ldots, x_n)}{\partial(y_1, \ldots, y_n)} \cdot \frac{\partial(y_1, \ldots, y_n)}{\partial(z_1, \ldots, z_n)} \quad . \tag{2.479}$$

Recall also that

$$\frac{\partial x_i}{\partial x_k} = \delta_{ik} \quad . \tag{2.480}$$

Consider the case $n = 2$. We have

$$\frac{\partial(x,y)}{\partial(u,v)} = \det\begin{pmatrix} \left(\frac{\partial x}{\partial u}\right)_v & \left(\frac{\partial x}{\partial v}\right)_u \\ \left(\frac{\partial y}{\partial u}\right)_v & \left(\frac{\partial y}{\partial v}\right)_u \end{pmatrix} = \left(\frac{\partial x}{\partial u}\right)_v \left(\frac{\partial y}{\partial v}\right)_u - \left(\frac{\partial x}{\partial v}\right)_u \left(\frac{\partial y}{\partial u}\right)_v \quad . \tag{2.481}$$

We also have

$$\frac{\partial(x,y)}{\partial(u,v)} \cdot \frac{\partial(u,v)}{\partial(r,s)} = \frac{\partial(x,y)}{\partial(r,s)} \quad . \tag{2.482}$$

From this simple mathematics follows several very useful results.

1) First, write

$$\frac{\partial(x,y)}{\partial(u,v)} = \left[\frac{\partial(u,v)}{\partial(x,y)}\right]^{-1} \quad . \tag{2.483}$$

Now let $v = y$:

$$\frac{\partial(x,y)}{\partial(u,y)} = \left(\frac{\partial x}{\partial u}\right)_y = 1 \Big/ \left(\frac{\partial u}{\partial x}\right)_y \quad . \tag{2.484}$$

Thus,

$$\left(\frac{\partial x}{\partial u}\right)_y = 1 \Big/ \left(\frac{\partial u}{\partial x}\right)_y \tag{2.485}$$

2) Second, we have

$$\frac{\partial(x,y)}{\partial(u,y)} = \left(\frac{\partial x}{\partial u}\right)_y = \frac{\partial(x,y)}{\partial(x,u)} \cdot \frac{\partial(x,u)}{\partial(u,y)} = -\left(\frac{\partial y}{\partial u}\right)_x \left(\frac{\partial x}{\partial y}\right)_u \quad ,$$

which is to say

$$\left(\frac{\partial x}{\partial y}\right)_u \left(\frac{\partial y}{\partial u}\right)_x = -\left(\frac{\partial x}{\partial u}\right)_y \quad . \tag{2.486}$$

Invoking eqn. 2.485, we conclude that

$$\left(\frac{\partial x}{\partial y}\right)_u \left(\frac{\partial y}{\partial u}\right)_x \left(\frac{\partial u}{\partial x}\right)_y = -1 \quad . \tag{2.487}$$

3) Third, we have

$$\frac{\partial(x,v)}{\partial(u,v)} = \frac{\partial(x,v)}{\partial(y,v)} \cdot \frac{\partial(y,v)}{\partial(u,v)} \quad , \tag{2.488}$$

which says

$$\left(\frac{\partial x}{\partial u}\right)_v = \left(\frac{\partial x}{\partial y}\right)_v \left(\frac{\partial y}{\partial u}\right)_v \tag{2.489}$$

This is simply the chain rule of partial differentiation.

4) Fourth, we have

$$\begin{aligned} \frac{\partial(x,y)}{\partial(u,y)} &= \frac{\partial(x,y)}{\partial(u,v)} \cdot \frac{\partial(u,v)}{\partial(u,y)} \\ &= \left(\frac{\partial x}{\partial u}\right)_v \left(\frac{\partial y}{\partial v}\right)_u \left(\frac{\partial v}{\partial y}\right)_u - \left(\frac{\partial x}{\partial v}\right)_u \left(\frac{\partial y}{\partial u}\right)_v \left(\frac{\partial v}{\partial y}\right)_u \quad , \end{aligned} \tag{2.490}$$

which says

$$\left(\frac{\partial x}{\partial u}\right)_y = \left(\frac{\partial x}{\partial u}\right)_v - \left(\frac{\partial x}{\partial y}\right)_u \left(\frac{\partial y}{\partial u}\right)_v \tag{2.491}$$

5) Fifth, whenever we differentiate one extensive quantity with respect to another, holding only intensive quantities constant, the result is simply the ratio of those extensive quantities. For example,

$$\left(\frac{\partial S}{\partial V}\right)_{p,T} = \frac{S}{V} \quad . \tag{2.492}$$

The reason should be obvious. In the above example, $S(p,V,T) = V\phi(p,T)$, where $\phi$ is a function of the two intensive quantities $p$ and $T$. Hence differentiating $S$ with respect to $V$ holding $p$ and $T$ constant is the same as dividing $S$ by $V$. Note that this implies

$$\left(\frac{\partial S}{\partial V}\right)_{p,T} = \left(\frac{\partial S}{\partial V}\right)_{p,\mu} = \left(\frac{\partial S}{\partial V}\right)_{n,T} = \frac{S}{V} \quad , \tag{2.493}$$

where $n = N/V$ is the particle density.

6) Sixth, suppose we have a function $\Phi(y, v)$ and we write

$$d\Phi = x\,dy + u\,dv \quad . \tag{2.494}$$

That is,

$$x = \left(\frac{\partial \Phi}{\partial y}\right)_v \equiv \Phi_y \qquad , \qquad u = \left(\frac{\partial \Phi}{\partial v}\right)_y \equiv \Phi_v \quad . \tag{2.495}$$

Now we may write

$$dx = \Phi_{yy}\,dy + \Phi_{yv}\,dv \tag{2.496}$$

$$du = \Phi_{vy}\,dy + \Phi_{vv}\,dv \quad . \tag{2.497}$$

If we demand $dv = 0$, this yields

$$\left(\frac{\partial x}{\partial u}\right)_v = \frac{\Phi_{yy}}{\Phi_{vy}} \quad . \tag{2.498}$$

Note that $\Phi_{vy} = \Phi_{yv}$. From the equation $du = 0$ we also derive

$$\left(\frac{\partial y}{\partial v}\right)_u = -\frac{\Phi_{vv}}{\Phi_{vy}} \quad . \tag{2.499}$$

Next, we use eqn. 2.497 with $du = 0$ to eliminate $dy$ in favor of $dv$, and then substitute into eqn. 2.496. This yields

$$\left(\frac{\partial x}{\partial v}\right)_u = \Phi_{yv} - \frac{\Phi_{yy}\,\Phi_{vv}}{\Phi_{vy}} \quad . \tag{2.500}$$

Finally, eqn. 2.497 with $dv = 0$ yields

$$\left(\frac{\partial y}{\partial u}\right)_v = \frac{1}{\Phi_{vy}} \quad . \tag{2.501}$$

Combining the results of eqns. 2.498, 2.499, 2.500, and 2.501, we have

$$\begin{aligned} \frac{\partial(x,y)}{\partial(u,v)} &= \left(\frac{\partial x}{\partial u}\right)_v \left(\frac{\partial y}{\partial v}\right)_u - \left(\frac{\partial x}{\partial v}\right)_u \left(\frac{\partial y}{\partial u}\right)_v \\ &= \left(\frac{\Phi_{yy}}{\Phi_{vy}}\right)\left(-\frac{\Phi_{vv}}{\Phi_{vy}}\right) - \left(\Phi_{yv} - \frac{\Phi_{yy}\,\Phi_{vv}}{\Phi_{vy}}\right)\left(\frac{1}{\Phi_{vy}}\right) = -1 \quad . \end{aligned} \tag{2.502}$$

Thus, if $\Phi = E(S, V)$, then $(x, y) = (T, S)$ and $(u, v) = (-p, V)$, we have

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{\partial(T,S)}{\partial(-p,V)} = -1 \quad . \tag{2.503}$$

*Nota bene:* It is important to understand what other quantities are kept constant, else problems may arise. For example, take $\Phi = G(p, N)$, whence $(x, y) = (V, p)$ and $(u, v) = (\mu, N)$, hence eqn. 2.502 yields

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{\partial(V,p)}{\partial(\mu,N)} = \frac{\partial(-p,V)}{\partial(\mu,N)} = -1 \quad . \tag{2.504}$$

But then it seems that we should have

$$\frac{\partial(T,S)}{\partial(\mu,N)} = \frac{\partial(T,S)}{\partial(-p,V)} \cdot \frac{\partial(-p,V)}{\partial(\mu,N)} = +1 \qquad \textbf{(WRONG!)}$$

when it should be $-1$: if $\Phi = F(T,N)$, then $(x,y) = (-S,T)$ and $(u,v) = (\mu,N)$, and again from eqn. 2.502 we have

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{\partial(-S,T)}{\partial(\mu,N)} = \frac{\partial(T,S)}{\partial(\mu,N)} = -1 \quad . \tag{2.505}$$

What has gone wrong?

The problem is that we haven't been consistent in what is being held constant. In eqn. 2.503 we are implicitly taking $dN = 0$ since $(y,v) = (S,V)$ and $E = E(S,V,N)$. In eqn. 2.504 we take $dT = 0$ since $(y,v) = (p,N)$ and $G = G(T,p,N)$. And in eqn. 2.505 we take $dV = 0$ since $(y,v) = (T,N)$ and $F = F(T,p,N)$. Thus a naïve and careless application of the chain rule for determinants yields the wrong result.

Let's be more careful. Applying the same derivation to $dE = x\,dy + u\,dv + r\,ds$ and holding $s$ constant, we conclude

$$\frac{\partial(x,y,s)}{\partial(u,v,s)} = \left(\frac{\partial x}{\partial u}\right)_{v,s}\left(\frac{\partial y}{\partial v}\right)_{u,s} - \left(\frac{\partial x}{\partial v}\right)_{u,s}\left(\frac{\partial y}{\partial u}\right)_{v,s} = -1 \quad . \tag{2.506}$$

Thus, if

$$dE = T\,dS + y\,dX + \mu\,dN \quad , \tag{2.507}$$

where $(y,X) = (-p,V)$ or $(H^\alpha, M^\alpha)$ or $(E^\alpha, P^\alpha)$, the appropriate thermodynamic relations are

$$\begin{aligned}
\frac{\partial(T,S,N)}{\partial(y,X,N)} &= -1 & \qquad \frac{\partial(T,S,\mu)}{\partial(y,X,\mu)} &= -1 \\
\frac{\partial(\mu,N,X)}{\partial(T,S,X)} &= -1 & \qquad \frac{\partial(\mu,N,y)}{\partial(T,S,y)} &= -1 \\
\frac{\partial(y,X,S)}{\partial(\mu,N,S)} &= -1 & \qquad \frac{\partial(y,X,T)}{\partial(\mu,N,T)} &= -1
\end{aligned} \tag{2.508}$$

For example,

$$\frac{\partial(T,S,N)}{\partial(-p,V,N)} = \frac{\partial(-p,V,S)}{\partial(\mu,N,S)} = \frac{\partial(\mu,N,V)}{\partial(T,S,V)} = -1 \tag{2.509}$$

and

$$\frac{\partial(T,S,\mu)}{\partial(-p,V,\mu)} = \frac{\partial(-p,V,T)}{\partial(\mu,N,T)} = \frac{\partial(\mu,N,-p)}{\partial(T,S,-p)} = -1 \quad . \tag{2.510}$$

If we are careful, then the results in eq. 2.508 can be quite handy, especially when used in conjunction with eqn. 2.479. For example, we have

$$\left(\frac{\partial S}{\partial V}\right)_{T,N} = \frac{\partial(T,S,N)}{\partial(T,V,N)} = \overbrace{\frac{\partial(T,S,N)}{\partial(p,V,N)}}^{=1} \cdot \frac{\partial(p,V,N)}{\partial(T,V,N)} = \left(\frac{\partial p}{\partial T}\right)_{V,N} \quad , \tag{2.511}$$

which is one of the Maxwell relations derived from the exactness of $dF(T,V,N)$. Some other examples include

$$\left(\frac{\partial V}{\partial S}\right)_{p,N} = \frac{\partial(V,p,N)}{\partial(S,p,N)} = \overbrace{\frac{\partial(V,p,N)}{\partial(S,T,N)}}^{=1} \cdot \frac{\partial(S,T,N)}{\partial(S,p,N)} = \left(\frac{\partial T}{\partial p}\right)_{S,N} \tag{2.512}$$

$$\left(\frac{\partial S}{\partial N}\right)_{T,p} = \frac{\partial(S,T,p)}{\partial(N,T,p)} = \overbrace{\frac{\partial(S,T,p)}{\partial(\mu,N,p)}}^{=1} \cdot \frac{\partial(\mu,N,p)}{\partial(N,T,p)} = -\left(\frac{\partial \mu}{\partial T}\right)_{p,N} \quad , \tag{2.513}$$

which are Maxwell relations deriving from $dH(S,p,N)$ and $dG(T,p,N)$, respectively. Note that due to the alternating nature of the determinant – it is antisymmetric under interchange of any two rows or columns – we have

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = -\frac{\partial(y,x,z)}{\partial(u,v,w)} = \frac{\partial(y,x,z)}{\partial(w,v,u)} = \ldots \quad . \tag{2.514}$$

In general, it is usually advisable to eliminate $S$ from a Jacobian. If we have a Jacobian involving $T$, $S$, and $N$, we can write

$$\frac{\partial(T,S,N)}{\partial(\bullet\,,\bullet\,,N)} = \frac{\partial(T,S,N)}{\partial(p,V,N)}\,\frac{\partial(p,V,N)}{\partial(\bullet\,,\bullet\,,N)} = \frac{\partial(p,V,N)}{\partial(\bullet\,,\bullet\,,N)} \quad , \tag{2.515}$$

where each $\bullet$ is a distinct arbitrary state variable other than $N$.

If our Jacobian involves $S$, $V$, and $N$, we write

$$\frac{\partial(S,V,N)}{\partial(\bullet\,,\bullet\,,N)} = \frac{\partial(S,V,N)}{\partial(T,V,N)} \cdot \frac{\partial(T,V,N)}{\partial(\bullet\,,\bullet\,,N)} = \frac{C_V}{T} \cdot \frac{\partial(T,V,N)}{\partial(\bullet\,,\bullet\,,N)} \quad . \tag{2.516}$$

If our Jacobian involves $S$, $p$, and $N$, we write

$$\frac{\partial(S,p,N)}{\partial(\bullet\,,\bullet\,,N)} = \frac{\partial(S,p,N)}{\partial(T,p,N)} \cdot \frac{\partial(T,p,N)}{\partial(\bullet\,,\bullet\,,N)} = \frac{C_p}{T} \cdot \frac{\partial(T,p,N)}{\partial(\bullet\,,\bullet\,,N)} \quad . \tag{2.517}$$

For example,

$$\left(\frac{\partial T}{\partial p}\right)_{S,N} = \frac{\partial(T,S,N)}{\partial(p,S,N)} = \overbrace{\frac{\partial(T,S,N)}{\partial(p,V,N)}}^{=1} \cdot \frac{\partial(p,V,N)}{\partial(p,T,N)} \cdot \frac{\partial(p,T,N)}{\partial(p,S,N)} = \frac{T}{C_p}\left(\frac{\partial V}{\partial T}\right)_{p,N} \tag{2.518}$$

$$\left(\frac{\partial V}{\partial p}\right)_{S,N} = \frac{\partial(V,S,N)}{\partial(p,S,N)} = \frac{\partial(V,S,N)}{\partial(V,T,N)} \cdot \frac{\partial(V,T,N)}{\partial(p,T,N)} \cdot \frac{\partial(p,T,N)}{\partial(p,S,N)} = \frac{C_V}{C_p}\left(\frac{\partial V}{\partial p}\right)_{T,N} \quad . \tag{2.519}$$

With $\kappa \equiv -\frac{1}{V}\frac{\partial V}{\partial p}$ the compressibility, we see that the second of these equations says $\kappa_T\, c_V = \kappa_S\, c_p$, relating the isothermal and adiabatic compressibilities and the molar heat capacities at constant volume and constant pressure. This relation was previously established in eqn. 2.266.

# Chapter 3

# Ergodicity and the Approach to Equilibrium

## 3.1 References

- R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, 1975)
  An advanced text with an emphasis on fluids and kinetics.

- R. Balian, *From Macrophysics to Microphysics* (2 vols., Springer-Verlag, 2006)
  A very detailed discussion of the fundamental postulates of statistical mechanics and their implications.

## 3.2 The Master Equation and the Approach to Equilibrium

### 3.2.1 Equilibrium

A thermodynamic system typically consists of an enormously large number of constituent particles, a typical 'large number' being Avogadro's number, $N_{\rm A} = 6.02 \times 10^{23}$. Nevertheless, in *equilibrium*, such a system is characterized by a relatively small number of thermodynamic state variables. Thus, while a complete description of a (classical) system would require us to account for $\mathcal{O}(10^{23})$ evolving degrees of freedom, with respect to the physical quantities in which we are interested, the details of the initial conditions are effectively forgotten over some microscopic time scale $\tau$, called the collision time, and over some microscopic distance scale, $\ell$, called the mean free path[1]. The equilibrium state is time-independent.

### 3.2.2 The master equation

Relaxation to equilibrium is often modeled with something called the *master equation*. Let $P_i(t)$ be the probability that the system is in a quantum or classical state $i$ at time $t$. Then write

$$\frac{dP_i}{dt} = \sum_j \big( W_{ij}\, P_j - W_{ji}\, P_i \big) \quad . \tag{3.1}$$

Here, $W_{ij}$ is the rate at which $j$ makes a transition to $i$. We assume the state space is finite and of size $N$, so that $i, j \in \{1, \ldots, N\}$. Note that the diagonal elements $W_{ii}$ (no sum on $i$) all cancel from the above equation and we may without loss of generality set $W_{ii} = 0 \;\forall\; i$. We may recast the master equation in the form

$$\frac{dP_i}{dt} = -\sum_j \Gamma_{ij}\, P_j \quad , \tag{3.2}$$

where

$$\Gamma_{ij} = \begin{cases} -W_{ij} & \text{if } i \neq j \\ \sum_k{}' W_{kj} & \text{if } i = j \end{cases} \quad , \tag{3.3}$$

where the prime on the sum indicates that $k = j$ is to be excluded. The constraints on the $W_{ij}$ are that $W_{ij} \geq 0$ for all $i, j$, and we set each $W_{ii} = 0$. Fermi's Golden Rule of quantum mechanics says that

$$W_{ij} = \frac{2\pi}{\hbar} \left| \langle\, i \,|\, \hat{V} \,|\, j \,\rangle \right|^2 \varrho(E_i) \quad , \tag{3.4}$$

with $i \neq j$, where $\hat{H}_0 \,|\, i \,\rangle = E_i \,|\, i \,\rangle$, $\hat{V}$ is an additional potential which leads to transitions, and $\varrho(E_i)$ is the density of final states at energy $E_i$. The fact that $W_{ij} \geq 0$ means that if each $P_i(t = 0) \geq 0$, then $P_i(t) \geq 0$ for all $t \geq 0$. To see this, suppose that at some time $t > 0$ one of the probabilities $P_i$ is crossing zero and about to become negative. But then eqn. 3.1 says that $\dot{P}_i(t) = \sum_j W_{ij} P_j(t) \geq 0$. So $P_i(t)$ can never become negative.

[1] Exceptions involve quantities which are conserved by collisions, such as overall particle number, momentum, and energy. These quantities relax to equilibrium in a special way called *hydrodynamics*.

### 3.2.3 Solution to the master equation

If the transition rates $W_{ij}$ are themselves time-independent, then we may formally write

$$P_i(t) = \left(e^{-\Gamma t}\right)_{ij} P_j(0) \quad . \tag{3.5}$$

Here we have used the Einstein 'summation convention' in which repeated indices are summed over (in this case, the $j$ index). Note that $\sum_i \Gamma_{ij} = 0$, which says that the total probability $\sum_i P_i$ is conserved:

$$\frac{d}{dt}\sum_i P_i = -\sum_{i,j} \Gamma_{ij}\, P_j = -\sum_j \left( P_j \sum_i \Gamma_{ij} \right) = 0 \quad . \tag{3.6}$$

We conclude that $\boldsymbol{L}^1 = (1\ ,\ 1\ ,\ \ldots\ ,\ 1)$ is a left eigenvector of $\Gamma$ with eigenvalue $\lambda = 0$. The corresponding right eigenvector, which we write as $\boldsymbol{R}^1$ with $R^1_i = P^{\rm eq}_i$, satisfies $\sum_j \Gamma_{ij} R^1_j = 0$, and is a stationary (*i.e.* time independent) solution to the master equation. The $N \times N$ matrix $\Gamma_{ij}$ is real but not necessarily symmetric. For such a matrix, the left eigenvectors $L^\alpha_i$ and the right eigenvectors $R^\beta_j$ satisfy

$$L^\alpha_i\, \Gamma_{ij} = \lambda_\alpha\, L^\alpha_j \qquad , \qquad \Gamma_{ij}\, R^\beta_j = \lambda_\beta\, R^\beta_i \quad , \tag{3.7}$$

but in general are not related by a simple transpose operation as they are for real symmetric matrices. We denote the right and left eigenvectors by $|\, R^\alpha \,\rangle$ and $\langle\, L^\alpha \,|$, which are column vectors and row vectors, respectively. Thus $\Gamma\,|\, R^\alpha\,\rangle = \lambda_\alpha |\, R^\alpha\,\rangle$ and $\langle\, L^\alpha\,|\,\Gamma = \langle\, L^\alpha\,|\,\lambda_\alpha$, the second of which may be written as $\Gamma^{\mathsf T}\,|\, L^\alpha\,\rangle = \lambda_\alpha\,|\, L^\alpha\,\rangle$, where the column vector $|\, L^\alpha\,\rangle$ is the transpose of the row vector $\langle\, L^\alpha\,|$ and $\Gamma^{\mathsf T}$ is the matrix transpose of $\Gamma$. The characteristic polynomial is the same in both cases:

$$F(\lambda) \equiv \det\,(\lambda - \Gamma) = \det\,(\lambda - \Gamma^{\mathsf T}) \quad , \tag{3.8}$$

which means that the left and right eigenvalues are the same. Note also that $\big[F(\lambda)\big]^* = F(\lambda^*)$, hence the eigenvalues are either real or appear in complex conjugate pairs. Multiplying the eigenvector equation for $L^\alpha$ on the right by $R^\beta_j$ and summing over $j$, and multiplying the eigenvector equation for $R^\beta$ on the left by $L^\alpha_i$ and summing over $i$, and subtracting the two results yields $(\lambda_\alpha - \lambda_\beta)\,\langle\, L^\alpha\,|\,R^\beta\,\rangle = 0$, where the inner product is now $\langle\, L^\alpha\,|\,R^\beta\,\rangle = \sum_i L^\alpha_i\, R^\beta_i$ with no complex conjugation on the bra vector. We may now demand $\langle\, L^\alpha\,|\,R^\beta\,\rangle = \delta^{\alpha\beta}$, which is our eigenvector normalization condition. As discussed above in §3.7, in the event that $\Gamma$ contains nontrivial Jordan blocks, its eigenvectors do not span. However, this is a nongeneric state of affairs, and here we assume that $\Gamma$ contains no nontrivial Jordan blocks. In the generic case, there are no Jordan blocks and we have the spectral decomposition

$$\Gamma = \sum_{a=1}^N \lambda_a\, |\, R^a\,\rangle\langle\, L^a\,| \qquad \Rightarrow \qquad \Gamma_{ij} = \sum_{a=1}^N \lambda_a\, R^a_i\, L^a_j \quad . \tag{3.9}$$

We have seen that $\langle\, L\,| = (1\ ,\ 1\ ,\ \ldots\ ,\ 1)$ is a left eigenvector of the matrix $\Gamma$ with eigenvalue $\lambda = 0$, since $\sum_i \Gamma_{ij} = 0$. We do not know *a priori* the corresponding right eigenvector, which depends on other details of $\Gamma_{ij}$. We reiterate that *generically* a matrix which is not normal has spanning eigenvectors, *i.e.*

the existence of nontrivial Jordan blocks is nongeneric. Assuming that the eigenvectors of $\Gamma$ span, then, let's expand the probability distribution $P_i(t)$ in the right eigenvectors of $\Gamma$, writing

$$P_i(t) = \sum_\alpha C_\alpha(t)\, R_i^\alpha \quad , \tag{3.10}$$

where $\alpha \in \{1, \ldots, N\}$. Then

$$\frac{dP_i}{dt} = \sum_\alpha \frac{dC_\alpha}{dt}\, R_i^\alpha = -\Gamma_{ij}\, P_j = -\sum_\alpha C_\alpha\, \Gamma_{ij}\, R_j^\alpha = -\sum_\alpha \lambda_\alpha\, C_\alpha\, R_i^\alpha \quad . \tag{3.11}$$

This allows us to write

$$\frac{dC_\alpha}{dt} = -\lambda_\alpha\, C_\alpha \qquad \Longrightarrow \qquad C_\alpha(t) = C_\alpha(0)\, e^{-\lambda_\alpha t} \quad . \tag{3.12}$$

Hence, we have the formal solution

$$P_i(t) = \sum_\alpha C_\alpha(0)\, e^{-\lambda_\alpha t}\, R_i^\alpha \quad . \tag{3.13}$$

Let $\alpha = 1$ correspond to the left eigenvector $\langle\, L^1 \,| = (1\ ,\ 1\ ,\ \ldots\ ,\ 1)$. The corresponding eigenvalue is $\lambda_1 = 0$. It is now easy to see that $\mathsf{Re}\,(\lambda_\alpha) > 0$ for all $\alpha > 1$, or else the probabilities will become negative[2]. For suppose $\mathsf{Re}\,(\lambda_\alpha) < 0$ for some $\alpha$. Then as $t \to \infty$, the sum in eqn. 3.13 will be dominated by the term for which $\lambda_\alpha$ has the largest negative real part; all other contributions will be subleading. But we must have $\sum_i R_i^\alpha = 0$ since $\|\, R^\alpha \,\rangle\!\rangle$ must be orthogonal to the left eigenvector $\langle\, L^1 \,|$. Therefore, at least one component of $R_i^\alpha$ (*i.e.* for some value of $i$) must have a negative real part, which means a negative probability![3] As we have already proven that an initial nonnegative distribution $\{P_i(t=0)\}$ will remain nonnegative under the evolution of the master equation, we conclude that $P_i(t) \to P_i^{\rm eq}$ as $t \to \infty$, relaxing to the $\lambda = 0$ right eigenvector, with $\mathsf{Re}\,(\lambda_\alpha) \geq 0$ for all $\alpha$.

### Detailed balance

In equilibrium, the net *flux* of transitions into a state $|i\rangle$ is equal to the flux of transitions out of $|i\rangle$. The flux of transitions from $|j\rangle$ to $|i\rangle$ is given by the product of the probability $\pi_j$ for the system to be in the state $|j\rangle$ multiplied by the bare rate $W_{ij}$. If, for each state $|j\rangle$, the flux from $|i\rangle$ to $|j\rangle$ is equal to the flux from $|j\rangle$ to $|i\rangle$, we say that the rates satisfy the condition of *detailed balance*. In other words,

$$W_{ij}\, \pi_j = W_{ji}\, \pi_i \quad , \tag{3.14}$$

where $\pi_i \equiv P_i^{\rm eq}$ is the equilibrium distribution. Note that this does not say that the bare rates are equal unless $\pi_i = \pi_j$. Assuming $W_{ij} \neq 0$ and $\pi_j \neq 0$, we may divide to obtain

$$\frac{W_{ji}}{W_{ij}} = \frac{\pi_j}{\pi_i} \quad . \tag{3.15}$$

[2]We presume that the eigenvalue $\lambda = 0$ is nondegenerate.

[3]Since the probability $P_i(t)$ is real, if the eigenvalue with the smallest (*i.e.* largest negative) real part is complex, there will be a corresponding complex conjugate eigenvalue, and summing over all eigenvectors will result in a real value for $P_i(t)$.

Note that detailed balance is a stronger condition than that required for a stationary solution to the master equation. Recall that $\Gamma_{ij} = -W_{ij}$ for all $i \neq j$. When detailed balance holds, we then have $\Gamma_{ji}\pi_i = \Gamma_{ij}\pi_j \; \forall \; i, j$. But this entails

$$\widetilde{\Gamma}_{ij} \equiv \pi_i^{-1/2}\Gamma_{ij}\pi_j^{1/2} = \widetilde{\Gamma}_{ji} \quad . \tag{3.16}$$

Note that $\widetilde{\Gamma}_{ii} = \Gamma_{ii}$ for the diagonal elements. The real symmetric matrix $\widetilde{\Gamma}$ has real eigenvalues $\lambda_a$ and a complete and orthonormal set of left/right eigenvectors $\phi_i^a$. For the right and left eigenvectors of $\Gamma$, we then have $R_j^a = \pi_j^{1/2}\phi_j^a$ and $L_i^a = \pi_i^{-1/2}\phi_i^a$. Since $\sum_i \pi_i^{1/2}\widetilde{\Gamma}_{ij} = 0$, we have that $\pi_i^{1/2}$ is a left eigenvector of $\widetilde{\Gamma}$ with eigenvalue zero. Therefore $R_j^1 = \pi_j$, which is the equilibrium distribution.

Consider an arbitrary nonnegative real upper triangular matrix $T$ with $T_{ij} \geq 0$ for all $1 \leq i < j \leq N$ and where each nonzero element has dimensions of rate (*i.e.* inverse time). Let $\pi_i$ be a normalized distribution, *i.e.* $\pi_i \geq 0$ for all $i \in \{1, \ldots, n\}$ with $\sum_i \pi_i = 1$. Now define the nonnegative matrix

$$W_{ij} = \begin{cases} \pi_i \, T_{ij} & \text{if} \quad i < j \\ \pi_i \, T_{ji} & \text{if} \quad i > j \end{cases} \quad . \tag{3.17}$$

and take this to be the matrix of transition rates, so that the master equation is as in Eqn. 3.1, *viz.*

$$\frac{dP_i}{dt} = \sum_j \left( W_{ij} \, P_j - W_{ji} \, P_i \right) \quad . \tag{3.18}$$

Since $W_{ij}/W_{ji} = \pi_i/\pi_j$, the matrix $W$ satisfies detailed balance relative to the distribution $\pi$. With $\Gamma_{ij} \equiv -W_{ij}$ for $i \neq j$ and $\Gamma_{ii} \equiv \sum_k{}' W_{ki}$ (with $k = i$ excluded from the sum) as before, we recover the form of the master equation $\dot{P}_i = -\sum_j \Gamma_{ij} \, P_j$ .

How many parameters does it take to describe a general $N \times N$ transition matrix $W_{ij}$ satisfying detailed balance? Since there are $\frac{1}{2}N(N-1)$ freedoms in $T$ and $N-1$ freedoms in $\pi$, we conclude that there are $\frac{1}{2}(N-1)(N+2)$ parameters required to specify $W_{ij}$ . But if we drop the constraint of detailed balance, then all the off-diagonal elements of $W_{ij}$ are independent, corresponding to $N(N-1)$ parameters. Note that we may set $W_{ii} = 0$ for all $i$.

For the Boltzmann distribution at temperature $T$, we have $\pi_i = Z^{-1}\exp(-\beta E_i)$ where $\beta = 1/k_{\rm B}T$ and $Z(\beta) = \sum_i \exp(-\beta E_i)$ is the ordinary canonical partition function. To satisfy detailed balance, we could take $W_{ij} = \gamma \, e^{\beta(E_j - E_i)/2}$, with $\gamma$ a constant having dimensions of inverse time. Or we could choose

$$W_{ij} = \begin{cases} \gamma & \text{if} \quad E_i < E_j \\ \gamma \, e^{\beta(E_j - E_i)} & \text{if} \quad E_i \geq E_j \end{cases} \quad . \tag{3.19}$$

The latter choice is common in Monte Carlo simulations.

If $\Gamma = \Gamma^{\intercal}$ is symmetric, then, up to normalization, the right eigenvectors and left eigenvectors are transposes of each other, hence $P^{\rm eq} = 1/N$, where $N$ is the matrix dimension of $\Gamma$. The system then satisfies the conditions of detailed balance. See §3.8 below for an example of this formalism applied to Poisson processes and to a model of radioactive decay.

### 3.2.4 Boltzmann's $\mathcal{H}$-theorem

Suppose for the moment that $\Gamma$ is a symmetric matrix, *i.e.* $\Gamma_{ij} = \Gamma_{ji}$. Then construct the function

$$\mathcal{H}(t) = \sum_i P_i(t)\ \log P_i(t) \quad . \tag{3.20}$$

Then

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= \sum_i \frac{dP_i}{dt}\left(1 + \log P_i\right) = \sum_i \frac{dP_i}{dt}\ \log P_i \\ &= -\sum_{i,j} \Gamma_{ij}\, P_j \log P_i = \sum_{i,j} \Gamma_{ij}\, P_j \big(\log P_j - \log P_i\big) \quad , \end{aligned} \tag{3.21}$$

where we have used $\sum_i \Gamma_{ij} = 0$. Now switch $i \leftrightarrow j$ in the above sum and add the terms to get

$$\frac{d\mathcal{H}}{dt} = \frac{1}{2}\sum_{i,j} \Gamma_{ij}\big(P_i - P_j\big)\,\big(\log P_i - \log P_j\big) \quad . \tag{3.22}$$

Note that the $i = j$ term does not contribute to the sum. For $i \neq j$ we have $\Gamma_{ij} = -W_{ij} \leq 0$, and using the result

$$(x - y)\,(\log x - \log y) \geq 0 \quad , \tag{3.23}$$

we conclude

$$\frac{d\mathcal{H}}{dt} \leq 0 \quad . \tag{3.24}$$

In equilibrium, $P_i^{\rm eq}$ is a constant, independent of $i$. We write

$$P_i^{\rm eq} = \frac{1}{\Omega} \quad , \quad \Omega = \sum_i 1 \quad \Longrightarrow \quad \mathcal{H} = -\log\Omega \quad , \tag{3.25}$$

with $\Omega = N$, the matrix dimension of $\Gamma$.

If $\Gamma_{ij} \neq \Gamma_{ji}$, we can still prove a version of the $\mathcal{H}$-theorem *when there is detailed balance*. Define a new symmetric matrix

$$\overline{W}_{ij} \equiv W_{ij}\, P_j^{\rm eq} = W_{ji}\, P_i^{\rm eq} = \overline{W}_{ji} \quad , \tag{3.26}$$

and the generalized $\mathcal{H}$-function,

$$\mathcal{H}(t) \equiv \sum_i P_i(t)\ \log\!\left(\frac{P_i(t)}{P_i^{\rm eq}}\right) \quad . \tag{3.27}$$

Then

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= -\frac{1}{2}\sum_{i,j}\left(W_{ji}P_i - W_{ij}P_j\right)\,\log\!\left(\frac{W_{ji}P_i}{W_{ij}P_j}\right) \\ &= -\frac{1}{2}\sum_{i,j}\overline{W}_{ij}\left(\frac{P_i}{P_i^{\rm eq}} - \frac{P_j}{P_j^{\rm eq}}\right)\left[\log\!\left(\frac{P_i}{P_i^{\rm eq}}\right) - \log\!\left(\frac{P_j}{P_j^{\rm eq}}\right)\right] \leq 0 \quad . \end{aligned} \tag{3.28}$$

### 3.2.5 Markov chains

A *Markov chain* is a process which describes transitions of a stochastic variable in time. There are four types to consider, depending on whether the state space is discrete (*i.e.* countable) or continuous (and measurable), and whether the process occurs in discrete time or continuous time. Here we will consider a discrete time Markov chain of a discrete state space. Let $P_i(t)$ be the probability that the system is in state $i$ at time $t$. The evolution equation is

$$P_i(t+1) = \sum_j Q_{ij}\, P_j(t) \quad . \tag{3.29}$$

Note the similarity to the (continuous time) master equation in eqn. 3.2. Indeed, the master equation is an example of a *continuous time Markov process*. All the elements of the *transition matrix* $Q_{ij}$ are real and nonnegative, and $Q_{ij}$ satisfies $\sum_i Q_{ij} = 1$ so that the total probability $\sum_i P_i(t)$ is conserved. The element $Q_{ij}$ is the *conditional probability* that for the system to evolve to state $i$ at time $t+1$ given that it is in state $j$ at time $t$, with $Q_{ij}$ independent of $t$. Iterating eqn. 3.29, we have

$$P_i(t+n) = \left[Q^n\right]_{ij} P_j(t) \quad . \tag{3.30}$$

The transition matrix $Q$ is in general non-normal, meaning $Q\,Q^{\mathsf{T}} \neq Q^{\mathsf{T}}Q$. In such cases the left eigenvectors are in general not simply related to the right eigenvectors at the corresponding eigenvalues, and the eigenvalues are not necessarily real. Under generic circumstances, we may write

$$Q = \sum_{a=1}^{n} \lambda_a \,|\, R^a \,\rangle\langle\, L^a \,| \quad , \tag{3.31}$$

with $|\lambda_a| \leq 1$ for all $a$. In nongeneric cases, there may be Jordan blocks in the canonical form of $Q$, which means the eigenvectors do not span – again see §3.7. From the condition $\sum_i Q_{ij} = 1$, we have that $\boldsymbol{L}^1 = (1, 1, \ldots, 1)$ is a left eigenvector corresponding to eigenvalue $\lambda = 1$, and the corresponding right eigenvector $\boldsymbol{R}^1$ is the asymptotic steady state, with $P_j^{\rm eq} = \langle\, j \,|\, R^1 \,\rangle = R_j^1$., here presumed unique. Since $Q \in M_N(\mathbb{R})$, where $N$ is the matrix dimension (which we take here to be finite), if $\lambda \in \mathbb{C}$ is an eigenvalue, then $\lambda^*$ is also an eigenvalue.

#### Example

Consider a group of Physics graduate students consisting of three theorists and four experimentalists. Within each group, the students are to be regarded as indistinguishable. Together, the students rent two apartments, A and B. Initially the three theorists live in A and the four experimentalists live in B. Each month, a random occupant of A and a random occupant of B exchange domiciles. Compute the transition matrix $Q_{ij}$ for this Markov chain, and compute the average fraction of the time that B contains two theorists and two experimentalists, averaged over the effectively infinite time it takes the students to get their degrees. *Hint:* $Q$ is a $4 \times 4$ matrix.

There are four states available (see Tab. 3.1). Let's compute the transition probabilities. First, we compute the transition probabilities out of state $|\,1\,\rangle$, *i.e.* the matrix elements $Q_{j1}$. Clearly $Q_{21} = 1$

| $\vert\, j\,\rangle$ | room A | room B | $g_j^{\rm A}$ | $g_j^{\rm B}$ | $g_j^{\rm TOT}$ |
|---|---|---|---|---|---|
| $\vert\, 1\,\rangle$ | TTT | EEEE | 1 | 1 | 1 |
| $\vert\, 2\,\rangle$ | TTE | EEET | 3 | 4 | 12 |
| $\vert\, 3\,\rangle$ | TEE | EETT | 3 | 6 | 18 |
| $\vert\, 4\,\rangle$ | EEE | ETTT | 1 | 4 | 4 |

Table 3.1: States and their degeneracies.

since we must exchange a theorist (T) for an experimentalist (E). All the other probabilities are zero: $Q_{11} = Q_{31} = Q_{41} = 0$. For transitions out of state $\vert\, 2\,\rangle$, the nonzero elements are

$$Q_{12} = \tfrac{1}{4} \times \tfrac{1}{3} = \tfrac{1}{12} \qquad , \qquad Q_{22} = \tfrac{3}{4} \times \tfrac{1}{3} + \tfrac{1}{4} \times \tfrac{2}{3} = \tfrac{5}{12} \qquad , \qquad Q_{32} = \tfrac{1}{2} \quad . \tag{3.32}$$

To compute $Q_{12}$, we must choose the experimentalist from room A (probability $\frac{1}{3}$) with the theorist from room B (probability $\frac{1}{4}$). For $Q_{22}$, we can either choose E from A and one of the E's from B, or one of the T's from A and the T from B. This explains the intermediate steps written above. For transitions out of state $\vert\, 3\,\rangle$, the nonzero elements are then $Q_{23} = \frac{1}{3}$, $Q_{33} = \frac{1}{2}$, and $Q_{43} = \frac{1}{6}$. Finally, for transitions out of state $\vert\, 4\,\rangle$, the nonzero elements are $Q_{34} = \frac{3}{4}$ and $Q_{44} = \frac{1}{4}$. The full transition matrix is then

$$Q = \begin{pmatrix} 0 & \frac{1}{12} & 0 & 0 \\ 1 & \frac{5}{12} & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{3}{4} \\ 0 & 0 & \frac{1}{6} & \frac{1}{4} \end{pmatrix} \quad . \tag{3.33}$$

Note that $\sum_i Q_{ij} = 1$ for all $j = 1, 2, 3, 4$. This guarantees that $\boldsymbol{L}^1 = (1\,,\,1\,,\,1\,,\,1)$ is a left eigenvector of $Q$ with eigenvalue 1. The corresponding right eigenvector is obtained by setting $Q_{ij}\, R_j^1 = R_i^1$. Simultaneously solving these four equations and normalizing so that $\sum_j R_j^1 = 1$, we easily obtain

$$\boldsymbol{R}^1 = \frac{1}{35} \begin{pmatrix} 1 \\ 12 \\ 18 \\ 4 \end{pmatrix} \quad . \tag{3.34}$$

This is the state we converge to after repeated application of the transition matrix $Q$. If we decompose $Q = \sum_{\alpha=1}^4 \lambda_\alpha \,\vert\, R^\alpha \,\rangle\langle\, L^\alpha \,\vert$, then in the limit $t \to \infty$ we have $Q^t \approx \vert\, R^1 \,\rangle\langle\, L^1 \,\vert$, where $\lambda_1 = 1$, since the remaining eigenvalues are all less than 1 in magnitude[4]. Thus, $Q^t\big\vert_{t\to\infty}$ acts as a *projector* onto the state $\vert\, R^1 \,\rangle$. Whatever the initial set of probabilities $P_j(t=0)$, we must have $\langle\, L^1 \,\vert\, P(0) \,\rangle = \sum_j P_j(0) = 1$. Therefore, $\lim_{t\to\infty} P_j(t) = R_j^1$, from which we have $P_3(\infty) = \frac{18}{35}$. Note that the equilibrium distribution satisfies detailed balance:

$$R_j^1 = \frac{g_j^{\rm TOT}}{\sum_l g_l^{\rm TOT}} \quad . \tag{3.35}$$

[4]One can check that $\lambda_1 = 1$, $\lambda_2 = \frac{5}{12}$, $\lambda_3 = 0$, and $\lambda_4 = -\frac{1}{4}$.

## 3.3 Phase Flows in Classical Mechanics

### 3.3.1 Hamiltonian evolution

The master equation provides us with a semi-phenomenological description of a dynamical system's relaxation to equilibrium. It explicitly breaks time reversal symmetry. Yet the microscopic laws of Nature are (approximately) time-reversal symmetric. How can a system which obeys Hamilton's equations of motion come to equilibrium?

Let's start our investigation by reviewing the basics of Hamiltonian dynamics. Recall the Lagrangian $L = L(q, \dot{q}, t) = T - V$. The Euler-Lagrange equations of motion for the action $S\big[q(t)\big] = \int dt\, L$ are

$$\dot{p}_\sigma = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) = \frac{\partial L}{\partial q_\sigma} \quad , \tag{3.36}$$

where $p_\sigma$ is the canonical momentum conjugate to the generalized coordinate $q_\sigma$, *i.e.* $p_\sigma = \partial L/\partial \dot{q}_\sigma$. Here $N$ is the number of degrees of freedom of the system, which is the total number of generalized coordinates.

The Hamiltonian, $H(q, p)$ is obtained by a Legendre transformation,

$$H(q,p) = \sum_{\sigma=1}^{N} p_\sigma\, \dot{q}_\sigma - L \quad . \tag{3.37}$$

Note that

$$\begin{aligned} dH &= \sum_{\sigma=1}^{N}\left(p_\sigma\, d\dot{q}_\sigma + \dot{q}_\sigma\, dp_\sigma - \frac{\partial L}{\partial q_\sigma}\, dq_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma}\, d\dot{q}_\sigma\right) - \frac{\partial L}{\partial t}\, dt \\ &= \sum_{\sigma=1}^{N}\left(\dot{q}_\sigma\, dp_\sigma - \frac{\partial L}{\partial q_\sigma}\, dq_\sigma\right) - \frac{\partial L}{\partial t}\, dt \quad . \end{aligned} \tag{3.38}$$

Thus, we obtain Hamilton's equations of motion,

$$\frac{\partial H}{\partial p_\sigma} = \dot{q}_\sigma \quad , \quad \frac{\partial H}{\partial q_\sigma} = -\frac{\partial L}{\partial q_\sigma} = -\dot{p}_\sigma \tag{3.39}$$

and

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \quad . \tag{3.40}$$

Define the rank $2N$ vector $\boldsymbol{\varphi}$ by its components,

$$\varphi_i = \begin{cases} q_i & \text{if } 1 \le i \le N \\ p_{i-N} & \text{if } N < i \le 2N \end{cases} \quad . \tag{3.41}$$

Then we may write Hamilton's equations compactly as

$$\dot{\varphi}_i = J_{ij}\, \frac{\partial H}{\partial \varphi_j} \quad , \tag{3.42}$$

where

$$J = \begin{pmatrix} 0_{N\times N} & 1_{N\times N} \\ -1_{N\times N} & 0_{N\times N} \end{pmatrix} \tag{3.43}$$

is a $2N \times 2N$ matrix. Note that $J^{\mathsf{T}} = -J$, *i.e.* $J$ is antisymmetric, and that $J^2 = -1_{2N\times 2N}$.

For any function $F(q, p, t)$, the total time derivative is given by

$$\begin{aligned} \frac{dF}{dt} &= \frac{\partial F}{\partial t} + \sum_{\sigma=1}^{N}\left(\frac{\partial F}{\partial q_\sigma}\frac{dq_\sigma}{dt} + \frac{\partial F}{\partial p_\sigma}\frac{dp_\sigma}{dt}\right) \\ &= \frac{\partial F}{\partial t} + \big\{F, H\big\} \quad , \end{aligned} \tag{3.44}$$

where $\big\{\bullet, \bullet\big\}$ is the *Poisson bracket*,

$$\big\{A, B\big\} = \sum_{\sigma=1}^{N}\left(\frac{\partial A}{\partial q_\sigma}\frac{\partial B}{\partial p_\sigma} - \frac{\partial A}{\partial p_\sigma}\frac{\partial B}{\partial q_\sigma}\right) \quad . \tag{3.45}$$

### 3.3.2 Dynamical systems and the evolution of phase space volumes

Consider a general dynamical system,

$$\frac{d\boldsymbol{\varphi}}{dt} = \boldsymbol{V}(\boldsymbol{\varphi}) \quad , \tag{3.46}$$

where $\boldsymbol{\varphi}(t)$ is a point in an $n$-dimensional phase space. Consider now a compact[5] region $\mathcal{R}_0$ in phase space, and consider its evolution under the dynamics. That is, $\mathcal{R}_0$ consists of a set of points $\big\{\boldsymbol{\varphi} \,|\, \boldsymbol{\varphi} \in \mathcal{R}_0\big\}$, and if we regard each $\boldsymbol{\varphi} \in \mathcal{R}_0$ as an initial condition, we can define the time-dependent set $\mathcal{R}(t)$ as the set of points $\boldsymbol{\varphi}(t)$ that *were* in $\mathcal{R}_0$ at time $t = 0$:

$$\mathcal{R}(t) = \big\{\boldsymbol{\varphi}(t) \,\big|\, \boldsymbol{\varphi}(0) \in \mathcal{R}_0\big\} \quad . \tag{3.47}$$

Now consider the volume $\Omega(t)$ of the set $\mathcal{R}(t)$. We have

$$\Omega(t) = \int\limits_{\mathcal{R}(t)} d\mu \tag{3.48}$$

where $d\mu = d\varphi_1\, d\varphi_2 \cdots d\varphi_n$ for an $n$-dimensional phase space. For a Hamiltonian system, $n = 2N$. We then have

$$\Omega(t+dt) = \int\limits_{\mathcal{R}(t+dt)} d\mu' = \int\limits_{\mathcal{R}(t)} d\mu \left|\frac{\partial\varphi_i(t+dt)}{\partial\varphi_j(t)}\right| \quad , \tag{3.49}$$

where

$$\left|\frac{\partial\varphi_i(t+dt)}{\partial\varphi_j(t)}\right| \equiv \frac{\partial(\varphi_1', \ldots, \varphi_n')}{\partial(\varphi_1, \ldots, \varphi_n)} \tag{3.50}$$

[5] 'Compact' in the parlance of mathematical analysis means 'closed and bounded'.

is a determinant, which is the Jacobian of the transformation from the set of coordinates $\{\varphi_i = \varphi_i(t)\}$ to the coordinates $\{\varphi_i' = \varphi_i(t+dt)\}$. But according to the dynamics, we have

$$\varphi_i(t+dt) = \varphi_i(t) + V_i\big(\boldsymbol{\varphi}(t)\big)\,dt + \mathcal{O}(dt^2) \tag{3.51}$$

and therefore

$$\frac{\partial\varphi_i(t+dt)}{\partial\varphi_j(t)} = \delta_{ij} + \frac{\partial V_i}{\partial\varphi_j}\,dt + \mathcal{O}(dt^2) \quad . \tag{3.52}$$

We now make use of the matrix equality $\log\det M = \mathsf{Tr}\,\log M$, which gives us[6], for small $\varepsilon$,

$$\det\big(1+\varepsilon A\big) = \exp\mathsf{Tr}\,\log\big(1+\varepsilon A\big) = 1 + \varepsilon\;\mathsf{Tr}\,A + \tfrac{1}{2}\,\varepsilon^2\Big(\big(\mathsf{Tr}\,A\big)^2 - \mathsf{Tr}\,(A^2)\Big) + \ldots \tag{3.53}$$

Thus,

$$\Omega(t+dt) = \Omega(t) + \int\limits_{\mathcal{R}(t)} d\mu\;\boldsymbol{\nabla}\cdot\boldsymbol{V}\;dt + \mathcal{O}(dt^2) \quad , \tag{3.54}$$

which says

$$\frac{d\Omega}{dt} = \int\limits_{\mathcal{R}(t)} d\mu\;\boldsymbol{\nabla}\cdot\boldsymbol{V} = \int\limits_{\partial\mathcal{R}(t)} dS\;\hat{\boldsymbol{n}}\cdot\boldsymbol{V} \tag{3.55}$$

Here, the divergence is the *phase space divergence*,

$$\boldsymbol{\nabla}\cdot\boldsymbol{V} = \sum_{i=1}^{n}\frac{\partial V_i}{\partial\varphi_i} \quad , \tag{3.56}$$

and we have used the divergence theorem to convert the volume integral of the divergence to a surface integral of $\hat{\boldsymbol{n}}\cdot\boldsymbol{V}$, where $\hat{\boldsymbol{n}}$ is the surface normal and $dS$ is the differential element of surface area, and $\partial\mathcal{R}$ denotes the boundary of the region $\mathcal{R}$. We see that if $\boldsymbol{\nabla}\cdot\boldsymbol{V} = 0$ everywhere in phase space, then $\Omega(t)$ is a constant, and phase space volumes are *preserved* by the evolution of the system.

For an alternative derivation, consider a function $\varrho(\boldsymbol{\varphi}, t)$ which is defined to be the *density* of some collection of points in phase space at phase space position $\boldsymbol{\varphi}$ and time $t$. This must satisfy the continuity equation,

$$\frac{\partial\varrho}{\partial t} + \boldsymbol{\nabla}\cdot(\varrho\boldsymbol{V}) = 0 \quad . \tag{3.57}$$

This is called the *continuity equation*. It says that 'nobody gets lost'. If we integrate it over a region of phase space $\mathcal{R}$, we have

$$\frac{d}{dt}\int\limits_{\mathcal{R}} d\mu\;\varrho = -\int\limits_{\mathcal{R}} d\mu\;\boldsymbol{\nabla}\cdot(\varrho\boldsymbol{V}) = -\int\limits_{\partial\mathcal{R}} dS\;\hat{\boldsymbol{n}}\cdot(\varrho\boldsymbol{V}) \quad . \tag{3.58}$$

It is perhaps helpful to think of $\varrho$ as a charge density, in which case $\boldsymbol{J} = \varrho\boldsymbol{V}$ is the current density. The above equation then says

$$\frac{dQ_{\mathcal{R}}}{dt} = -\int\limits_{\partial\mathcal{R}} dS\;\hat{\boldsymbol{n}}\cdot\boldsymbol{J} \quad , \tag{3.59}$$

[6]The equality $\log\det M = \mathsf{Tr}\,\log M$ is most easily proven by bringing the matrix to diagonal form via a similarity transformation, and proving the equality for diagonal matrices.

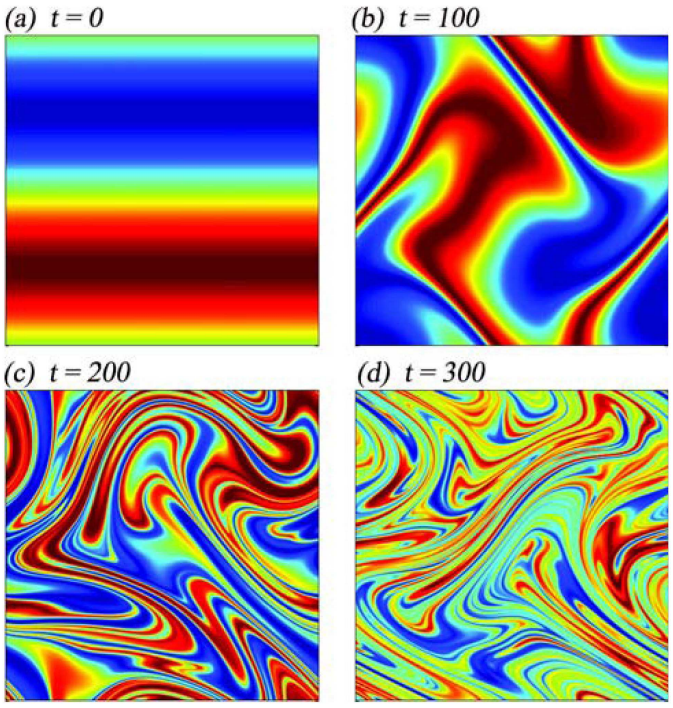



Figure 3.1: Hamiltonian time evolution of a density distribution. Dark red corresponds to $\varrho = 0$ and dark blue to $\varrho = 1$. The local density remains constant in the comoving frame. From D. Saintillan and M. Shelley, *Phys. Fluids* **20**, 123304 (2008).

where $Q_{\mathcal{R}}$ is the total charge contained inside the region $\mathcal{R}$. In other words, the rate of increase or decrease of the charge within the region $\mathcal{R}$ is equal to the total integrated current flowing in or out of $\mathcal{R}$ at its boundary.

The Leibniz rule lets us write the continuity equation as

$$\frac{\partial \varrho}{\partial t} + \boldsymbol{V}\cdot\boldsymbol{\nabla}\varrho + \varrho\,\boldsymbol{\nabla}\cdot\boldsymbol{V} = 0 \quad . \tag{3.60}$$

But now suppose that the phase flow is divergenceless, *i.e.* $\boldsymbol{\nabla}\cdot\boldsymbol{V} = 0$. Then we have

$$\frac{D\varrho}{Dt} \equiv \left(\frac{\partial}{\partial t} + \boldsymbol{V}\cdot\boldsymbol{\nabla}\right)\varrho = 0 \quad . \tag{3.61}$$

The combination inside the brackets above is known as the *convective derivative*. It tells us the total rate of change of $\varrho$ for an observer *moving with the phase flow*. That is

$$\begin{aligned}\frac{d}{dt}\,\varrho\big(\boldsymbol{\varphi}(t),t\big) &= \frac{\partial\varrho}{\partial\varphi_i}\,\frac{d\varphi_i}{dt} + \frac{\partial\varrho}{\partial t} \\ &= \sum_{i=1}^{n} V_i\,\frac{\partial\varrho}{\partial\varphi_i} + \frac{\partial\varrho}{\partial t} = \frac{D\varrho}{Dt} \quad .\end{aligned} \tag{3.62}$$

If $D\varrho/Dt = 0$, the local density remains the same during the evolution of the system. If we consider the 'characteristic function'

$$\varrho(\boldsymbol{\varphi}, t=0) = \begin{cases} 1 & \text{if } \boldsymbol{\varphi} \in \mathcal{R}_0 \\ 0 & \text{otherwise} \end{cases} \tag{3.63}$$

then the vanishing of the convective derivative means that the image of the set $\mathcal{R}_0$ under time evolution will always have the same volume.

Hamiltonian evolution in classical mechanics is volume preserving. The equations of motion are

$$\dot{q}_\sigma = +\frac{\partial H}{\partial p_\sigma} \quad , \quad \dot{p}_\sigma = -\frac{\partial H}{\partial q_\sigma} \tag{3.64}$$

A point in phase space is specified by $N$ positions $q_\sigma$ and $N$ momenta $p_\sigma$, hence the dimension of phase space is $n = 2N$:

$$\boldsymbol{\varphi} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} \quad , \quad \boldsymbol{V} = \begin{pmatrix} \dot{\boldsymbol{q}} \\ \dot{\boldsymbol{p}} \end{pmatrix} = \begin{pmatrix} +\partial H/\partial \boldsymbol{p} \\ -\partial H/\partial \boldsymbol{q} \end{pmatrix} \quad . \tag{3.65}$$

Hamilton's equations of motion guarantee that the phase space flow is divergenceless:

$$\begin{aligned} \nabla \cdot \boldsymbol{V} &= \sum_{\sigma=1}^{N} \left\{ \frac{\partial \dot{q}_\sigma}{\partial q_\sigma} + \frac{\partial \dot{p}\sigma}{\partial p_\sigma} \right\} \\ &= \sum_{\sigma=1}^{N} \left\{ \frac{\partial}{\partial q_\sigma}\left(\frac{\partial H}{\partial p_\sigma}\right) + \frac{\partial}{\partial p_\sigma}\left(-\frac{\partial H}{\partial q_\sigma}\right) \right\} = 0 \quad . \end{aligned} \tag{3.66}$$

Thus, we have that the convective derivative vanishes, *viz.*

$$\frac{D\varrho}{Dt} \equiv \frac{\partial \varrho}{\partial t} + \boldsymbol{V} \cdot \boldsymbol{\nabla} \varrho = 0 \quad , \tag{3.67}$$

for any distribution $\varrho(\boldsymbol{\varphi}, t)$ on phase space advected by a Hamiltonian flow. Thus, the value of the density $\varrho(\boldsymbol{\varphi}(t), t)$ is constant, which tells us that the phase flow is *incompressible*. In particular, phase space volumes are preserved.

### 3.3.3 Liouville's equation and the microcanonical distribution

Let $\varrho(\boldsymbol{\varphi}, t) = \varrho(\boldsymbol{q}, \boldsymbol{p}, t)$ be a distribution on phase space. Assuming the evolution is Hamiltonian, we can write

$$\frac{\partial \varrho}{\partial t} = -\dot{\boldsymbol{\varphi}} \cdot \boldsymbol{\nabla} \varrho = -\big\{ \varrho, H \big\} \quad . \tag{3.68}$$

We may also write this as $\partial_t \varrho + i\hat{L}\varrho = 0$, where $\hat{L}$ is a differential operator known as the *Liouvillian*:

$$i\hat{L} = \sum_{\sigma=1}^{N} \left\{ \frac{\partial H}{\partial p_\sigma}\frac{\partial}{\partial q_\sigma} - \frac{\partial H}{\partial q_\sigma}\frac{\partial}{\partial p_\sigma} \right\} \quad . \tag{3.69}$$

Eqn. 3.68 is known as *Liouville's equation*. Note $i\hat{L}F = \big\{F, H\big\}$ for any function $F(\boldsymbol{\varphi}, t)$.

Recall that the evolution of quantum mechanical density matrices satisfies

$$\frac{\partial\hat{\varrho}}{\partial t} = \frac{i}{\hbar}\,\big[\,\hat{\varrho},\hat{H}\,\big] \quad , \tag{3.70}$$

whence we infer the correspondence

$$\big\{\varrho, H\big\} \to \frac{1}{i\hbar}\,\big[\,\hat{\varrho},\hat{H}\,\big] \quad . \tag{3.71}$$

Suppose that there is a family of conserved quantities $\Lambda_a(\boldsymbol{\varphi})$, with $a \in \{1,\ldots,k\}$, each of which is conserved by the dynamics of the system. Such conserved quantities might include the components of the total linear momentum (if there is translational invariance), the components of the total angular momentum (if there is rotational invariance), and the Hamiltonian itself (if it is not explicitly time-dependent). Now consider a distribution $\varrho(\boldsymbol{\varphi}) = \varrho(\Lambda_1,\Lambda_2,\ldots,\Lambda_k)$ which is a function only of these various conserved quantities. That $\Lambda_a$ is conserved entails $\dot{\boldsymbol{\varphi}}\cdot\boldsymbol{\nabla}\Lambda_a = \big\{\Lambda_a, H\big\} = 0$. Then from the chain rule, we have

$$\dot{\boldsymbol{\varphi}}\cdot\boldsymbol{\nabla}\varrho = \sum_{a=1}^{k}\frac{\partial\varrho}{\partial\Lambda_a}\,\dot{\boldsymbol{\varphi}}\cdot\boldsymbol{\nabla}\Lambda_a = 0 \quad . \tag{3.72}$$

We conclude that any distribution $\varrho(\boldsymbol{\varphi}) = \varrho(\Lambda_1,\Lambda_2,\ldots,\Lambda_k)$ which is a function solely of conserved dynamical quantities is a stationary solution to Liouville's equation.

Clearly the microcanonical distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta\big(E-H(\boldsymbol{\varphi})\big)}{D(E)} = \frac{\delta\big(E-H(\boldsymbol{\varphi})\big)}{\int d\mu\,\delta\big(E-H(\boldsymbol{\varphi})\big)} \quad , \tag{3.73}$$

is a fixed point solution of Liouville's equation. If there were a second conserved quantity, $\Lambda(\boldsymbol{\varphi})$, the generalized microcanonical distribution,

$$\varrho_{E,Q}(\boldsymbol{\varphi}) = \frac{\delta\big(E-H(\boldsymbol{\varphi})\big)\,\delta\big(Q-\Lambda(\boldsymbol{\varphi})\big)}{D(E,Q)} = \frac{\delta\big(E-H(\boldsymbol{\varphi})\big)\,\delta\big(Q-\Lambda(\boldsymbol{\varphi})\big)}{\int d\mu\,\delta\big(E-H(\boldsymbol{\varphi})\,\delta\big(Q-\Lambda(\boldsymbol{\varphi})\big)} \quad , \tag{3.74}$$

would be a solution to Liouville's equation for arbitrary $E$ and $Q$.

Similarly, the Gibbs distribution,

$$\varrho_\beta(\boldsymbol{\varphi}) = \frac{1}{Z(\beta)}\,e^{-\beta H(\boldsymbol{\varphi})} \quad , \tag{3.75}$$

where $Z(\beta) = \mathsf{Tr}\;e^{-\beta H(\boldsymbol{\varphi})} = \int d\mu(\boldsymbol{\varphi})\,e^{-\beta H(\boldsymbol{\varphi})}$ is the partition function, satisfies $\big\{\varrho_\beta, H\big\} = 0$. In the presence of multiple conserved quantities, one defines the *generalized Gibbs distribution*,

$$\varrho_{\boldsymbol{\beta}}(\boldsymbol{\varphi}) = \frac{1}{Z(\boldsymbol{\beta})}\,\exp\Big(-\sum_{a=1}^{k}\beta_a\,\Lambda_a(\boldsymbol{\varphi})\Big) \quad , \tag{3.76}$$

where the $\{\Lambda_a(\boldsymbol{\varphi})\}$, with $a \in \{1,\ldots,k\}$, are the conserved quantities, including among them $H$ itself. The coefficients $\{\beta_a\}$ are $k$ Lagrange multipliers enforcing the $k$ conservation constraints $\Lambda_a(\boldsymbol{\varphi}) = Q_a$.

## 3.4 Irreversibility and Poincaré Recurrence

The dynamics of the master equation describe an approach to equilibrium. These dynamics are irreversible: $dH/dt \leq 0$, where $H$ is Boltzmann's $H$-function. However, the microscopic laws of physics are (almost) time-reversal invariant[7], so how can we understand the emergence of irreversibility? Furthermore, any dynamics which are deterministic and volume-preserving in a finite phase space exhibit the phenomenon of *Poincaré recurrence*, which guarantees that phase space trajectories are arbitrarily close to periodic if one waits long enough.

### 3.4.1 Poincaré recurrence theorem

The proof of the recurrence theorem is simple. Let $g_\tau$ be the '$\tau$-advance mapping' which evolves points in phase space according to Hamilton's equations. Assume that $g_\tau$ is invertible and volume-preserving, as is the case for Hamiltonian flow. Further assume that phase space volume is finite. Since energy is preserved in the case of time-independent Hamiltonians, we simply ask that the volume of phase space *at fixed total energy* $E$ be finite, *i.e.*

$$\int d\mu\, \delta\big(E - H(\boldsymbol{q},\boldsymbol{p})\big) < \infty \quad , \tag{3.77}$$

where $d\mu = d\boldsymbol{q}\, d\boldsymbol{p}$ is the phase space uniform integration measure.

**Theorem:** In any finite neighborhood $\mathcal{R}_0$ of phase space there exists a point $\boldsymbol{\varphi}_0$ which will return to $\mathcal{R}_0$ after $m$ applications of $g_\tau$, where $m$ is finite.

**Proof:** Assume the theorem fails; we will show this assumption results in a contradiction. Consider the set $\Upsilon$ formed from the union of all sets $g_\tau^k\,\mathcal{R}$ over all $k$:

$$\Upsilon = \bigcup_{k=0}^{\infty} g_\tau^k\,\mathcal{R}_0 \tag{3.78}$$

We assume that the set $\{g_\tau^k\,\mathcal{R}_0 \,|\, k \in \mathbb{N}_0\}$ is disjoint[8]. The volume of a union of disjoint sets is the sum of the individual volumes. Thus,

$$\mathrm{vol}(\Upsilon) = \sum_{k=0}^{\infty} \mathrm{vol}\big(g_\tau^k\,\mathcal{R}_0\big) = \mathrm{vol}(\mathcal{R}_0)\cdot\sum_{k=0}^{\infty} 1 = \infty \quad , \tag{3.79}$$

since $\mathrm{vol}\big(g_\tau^k\,\mathcal{R}_0\big) = \mathrm{vol}\big(\mathcal{R}_0\big)$ from volume preservation. But clearly $\Upsilon$ is a subset of the entire phase space, hence we have a contradiction, because by assumption phase space is of finite volume.

Thus, the assumption that the set $\{g_\tau^k\,\mathcal{R}_0 \,|\, k \in \mathbb{N}_0\}$ is disjoint fails. This means that there exists some pair of integers $k$ and $l$, with $k \neq l$, such that $g_\tau^k\,\mathcal{R}_0 \cap g_\tau^l\,\mathcal{R}_0 \neq \emptyset$. Without loss of generality we may assume

[7] Actually, the microscopic laws of physics are *not* time-reversal invariant, but rather are invariant under the product $CPT$, where $C$ is charge conjugation, $P$ is parity, and $T$ is time reversal.

[8] The natural numbers $\mathbb{N}_0$ is the set of non-negative integers $\{0, 1, 2, \ldots\}$.

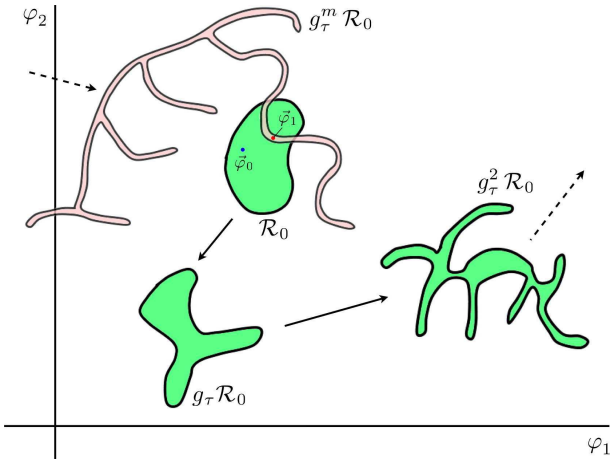



Figure 3.2: Successive images of a set $\mathcal{R}_0$ under the $\tau$-advance mapping $g_\tau$, projected onto a two-dimensional phase plane. The Poincaré recurrence theorem guarantees that if phase space has finite volume, and $g_\tau$ is invertible and volume preserving, then for any set $\mathcal{R}_0$ there exists an integer $m$ such that $\mathcal{R}_0 \cap g_\tau^m\,\mathcal{R}_0 \neq \emptyset$.

$k < l$. Apply the inverse $g_\tau^{-1}$ to this relation $k$ times to get $g_\tau^{l-k}\,\mathcal{R}_0 \cap \mathcal{R}_0 \neq \emptyset$. Now choose any point $\boldsymbol{\varphi}_1 \in g_\tau^m\,\mathcal{R}_0 \cap \mathcal{R}_0$, where $m = l - k$, and define $\boldsymbol{\varphi}_0 = g_\tau^{-m}\boldsymbol{\varphi}_1$. Then by construction both $\boldsymbol{\varphi}_0$ and $g_\tau^m\,\boldsymbol{\varphi}_0$ lie within $\mathcal{R}_0$ and the theorem is proven.

Poincaré recurrence has remarkable implications. Consider a bottle of perfume which is opened in an otherwise evacuated room, as depicted in fig. 3.3. The perfume molecules evolve according to Hamiltonian evolution. The positions are bounded because physical space is finite. The momenta are bounded because the total energy is conserved, hence no single particle can have a momentum such that $T(\boldsymbol{p}) > E_{\rm TOT}$, where $T(\boldsymbol{p})$ is the single particle kinetic energy function[9]. Thus, phase space, however large, is still bounded. Hamiltonian evolution, as we have seen, is invertible and volume preserving, therefore the system is recurrent. All the molecules must eventually return to the bottle. What's more, they all must return with momenta arbitrarily close to their initial momenta![10] In this case, we could define the region $\mathcal{R}_0$ as

$$\mathcal{R}_0 = \big\{(q_1,\ldots,q_r,p_1,\ldots,p_r)\;\big|\;|q_i - q_i^0| \le \Delta q \text{ and } |p_j - p_j^0| \le \Delta p\;\forall\;i,j\big\} \quad , \tag{3.80}$$

which specifies a hypercube in phase space centered about the point $(\boldsymbol{q}^0, \boldsymbol{p}^0)$.

Each of the three central assumptions – finite phase space, invertibility, and volume preservation – is crucial. If any one of these assumptions does not hold, the proof fails. Obviously if phase space is infinite the flow needn't be recurrent since it can keep moving off in a particular direction. Consider next a volume-preserving map which is not invertible. An example might be a mapping $f\colon \mathbb{R} \to \mathbb{R}$ which takes any real number to its fractional part. Thus, $f(\pi) = 0.14159265\ldots$. Let us restrict our attention to intervals of width less than unity. Clearly $f$ is then volume preserving. The action of $f$ on

[9] In the nonrelativistic limit, $T = p^2/2m$. For relativistic particles, we have $T = (p^2c^2 + m^2c^4)^{1/2} - mc^2$.

[10] Actually, what the recurrence theorem guarantees is that there is a configuration arbitrarily close to the initial one which recurs, to within the same degree of closeness.

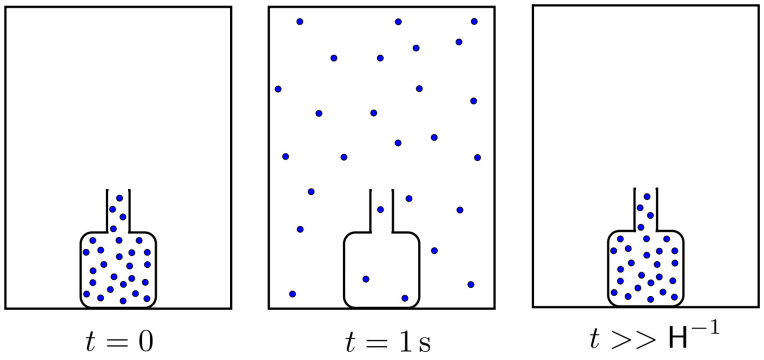



Figure 3.3: Poincaré recurrence guarantees that if we remove the cap from a bottle of perfume in an otherwise evacuated room, all the perfume molecules will eventually return to the bottle! (Here $\mathsf{H}$ is the Hubble constant.)

the interval $[2, 3)$ is to map it to the interval $[0, 1)$. But $[0, 1)$ remains fixed under the action of $f$, so no point within the interval $[2, 3)$ will ever return under repeated iterations of $f$. Thus, $f$ does not exhibit Poincaré recurrence.

Consider next the case of the damped harmonic oscillator. In this case, phase space volumes contract. For a one-dimensional oscillator obeying $\ddot{x} + 2\beta\dot{x} + \Omega_0^2\, x = 0$ one has $\nabla \cdot \boldsymbol{V} = -2\beta < 0$, since $\beta > 0$ for physical damping. Thus the convective derivative is $D_t\varrho = -(\nabla \cdot \boldsymbol{V})\varrho = 2\beta\varrho$ which says that the density increases exponentially in the comoving frame, as $\varrho(t) = e^{2\beta t}\,\varrho(0)$. Thus, phase space volumes collapse: $\Omega(t) = e^{-2\beta t}\,\Omega(0)$, and are not preserved by the dynamics. The proof of recurrence therefore fails. In this case, it is possible for the set $\Upsilon$ to be of finite volume, even if it is the union of an infinite number of sets $g_\tau^k\,\mathcal{R}_0$, because the volumes of these component sets themselves decrease exponentially, as $\mathrm{vol}(g_\tau^n\,\mathcal{R}_0) = e^{-2n\beta\tau}\,\mathrm{vol}(\mathcal{R}_0)$. A damped pendulum, released from rest at some small angle $\theta_0$, will not return arbitrarily close to these initial conditions.

### 3.4.2 Kac ring model

The implications of the Poincaré recurrence theorem are surprising – even shocking. If one takes a bottle of perfume in a sealed, evacuated room and opens it, the perfume molecules will diffuse throughout the room. The recurrence theorem guarantees that after some finite time $T$ all the molecules will go back inside the bottle (and arbitrarily close to their initial velocities as well). The hitch is that this could take a very long time, *e.g.* much much longer than the age of the Universe.

On less absurd time scales, we know that most systems come to thermodynamic equilibrium. But how can a system both exhibit equilibration *and* Poincaré recurrence? The two concepts seem utterly incompatible!

A beautifully simple model due to Kac shows how a recurrent system can exhibit the phenomenon of equilibration. Consider a ring with $N$ sites. On each site, place a 'spin' which can be in one of two states: up or down. Along the $N$ links of the system, $F$ of them contain 'flippers'. The configuration of the flippers is set at the outset and never changes. The dynamics of the system are as follows: during

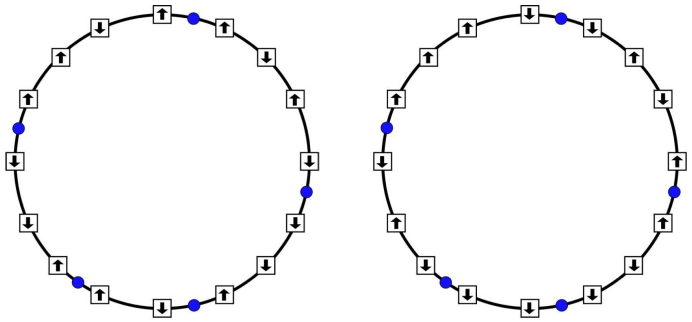



Figure 3.4: Left: A configuration of the Kac ring with $N = 16$ sites and $F = 4$ flippers. The flippers, which live on the links, are represented by blue dots. Right: The ring system after one time step. Evolution proceeds by clockwise rotation. Spins passing through flippers are flipped.

each time step, every spin moves clockwise a distance of one lattice spacing. Spins which pass through flippers reverse their orientation: up becomes down, and down becomes up.

The 'phase space' for this system consists of $2^N$ discrete configurations. Since each configuration maps onto a unique image under the evolution of the system, phase space 'volume' is preserved. The evolution is invertible; the inverse is obtained simply by rotating the spins counterclockwise. Figure 3.4 depicts an example configuration for the system, and its first iteration under the dynamics.

Suppose the flippers were not fixed, but moved about randomly. In this case, we could focus on a single spin and determine its configuration probabilistically. Let $p_n$ be the probability that a given spin is in the up configuration at time $n$. The probability that it is up at time $(n+1)$ is then

$$p_{n+1} = (1-x)\,p_n + x\,(1-p_n) \quad , \tag{3.81}$$

where $x = F/N$ is the fraction of flippers in the system. In words: a spin will be up at time $(n+1)$ if it was up at time $n$ and did not pass through a flipper, or if it was down at time $n$ and did pass through a flipper. If the flipper locations are randomized at each time step, then the probability of flipping is simply $x = F/N$. Equation 3.81 can be solved immediately:

$$p_n = \tfrac{1}{2} + (1-2x)^n\,(p_0 - \tfrac{1}{2}) \quad , \tag{3.82}$$

which decays exponentially to the equilibrium value of $p_{\rm eq} = \frac{1}{2}$ with time scale

$$\tau(x) = -\frac{1}{\log|1-2x|} \quad . \tag{3.83}$$

We identify $\tau(x)$ as the microscopic relaxation time over which local equilibrium is established. If we define the magnetization $m \equiv (N_\uparrow - N_\downarrow)/N$, then $m = 2p - 1$, so $m_n = (1-2x)^n\,m_0$. The equilibrium magnetization is $m_{\rm eq} = 0$. Note that for $\frac{1}{2} < x < 1$ the magnetization reverses sign each time step, as well as decreasing exponentially in magnitude.

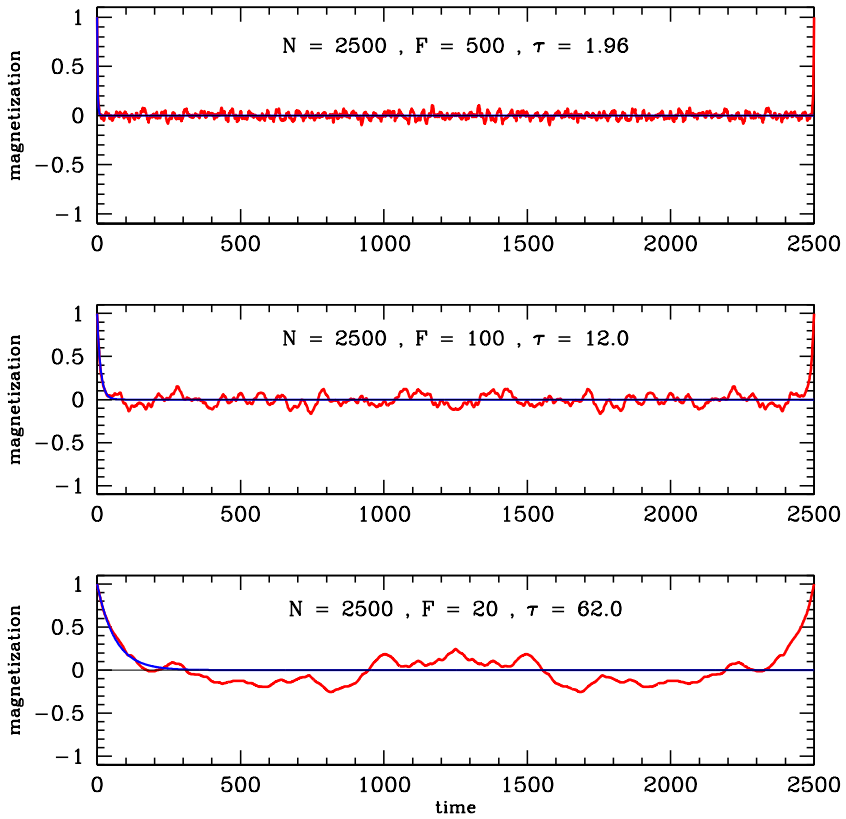



Figure 3.5: Three simulations of the Kac ring model with $N = 2500$ sites and three different concentrations of flippers. The red line shows the magnetization as a function of time, starting from an initial configuration in which 100% of the spins are up. The blue line shows the prediction of the *Stosszahlansatz*, which yields an exponentially decaying magnetization with time constant $\tau$.

The assumption that leads to equation 3.81 is called the *Stosszahlansatz*[11], a long German word meaning, approximately, 'assumption on the counting of hits'. The resulting dynamics are irreversible: the magnetization inexorably decays to zero. However, the Kac ring model is purely deterministic, and the *Stosszahlansatz* can at best be an approximation to the true dynamics. Clearly the *Stosszahlansatz* fails to account for correlations such as the following: if spin $i$ is flipped at time $n$, then spin $i+1$ will have been flipped at time $n-1$. Also if spin $i$ is flipped at time $n$, then it also will be flipped at time $n+N$. Indeed, since the dynamics of the Kac ring model are invertible and volume preserving, it must exhibit Poincaré recurrence. We see this most vividly in figs. 3.5 and 3.6.

The model is trivial to simulate. The results of such a simulation are shown in figure 3.5 for a ring of $N = 2500$ sites, with $F = 500$, 100, and 20 flippers. Note how the magnetization decays and fluctuates about the equilibrium value $m_{\rm eq} = 0$, but that after $N$ iterations $m$ recovers its initial value: $m_N = m_0$. The recurrence time for this system is simply $N$ if $F$ is even, and $2N$ if $F$ is odd, since every spin will then have flipped an even number of times.

[11] Unfortunately, many important physicists were German and we have to put up with a legacy of long German words like *Gedankenexperiment*, *Zitterbewegung*, *Bremsstrahlung*, *Stosszahlansatz*, *Kartoffelsalat*, *etc.*

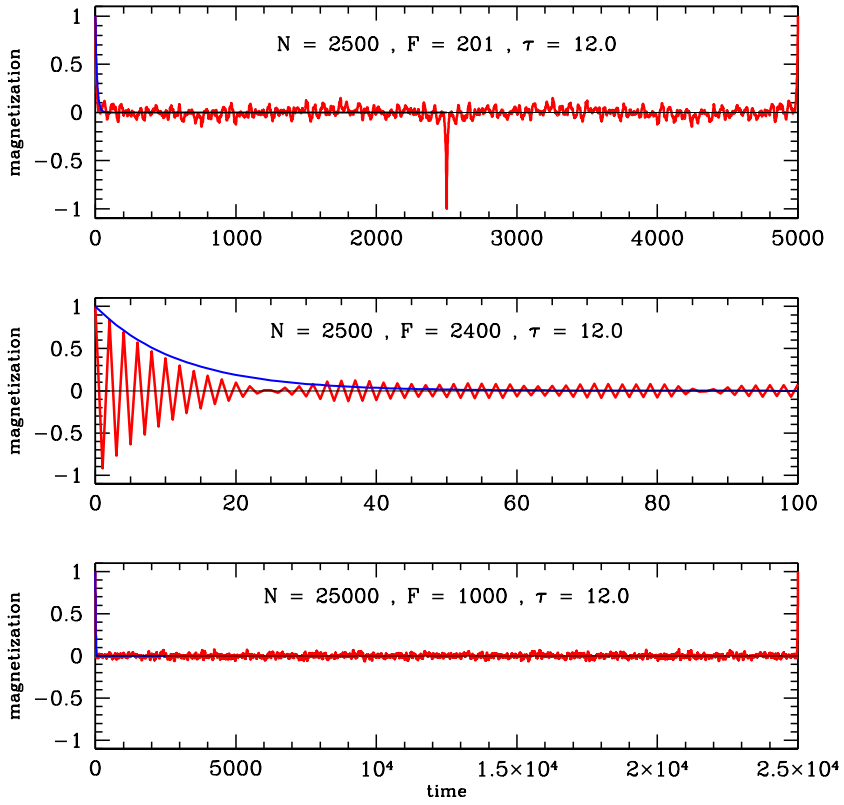



Figure 3.6: Simulations of the Kac ring model. Top: $N = 2500$ sites with $F = 201$ flippers. After 2500 iterations, each spin has flipped an odd number of times, so the recurrence time is $2N$. Middle: $N = 2500$ with $F = 2400$, resulting in a near-complete reversal of the population with every iteration. Bottom: $N = 25000$ with $F = 1000$, showing long time equilibration and dramatic resurgence of the spin population.

In figure 3.6 we plot two other simulations. The middle panel shows what happens when $x > \frac{1}{2}$, so that the magnetization wants to reverse its sign with every iteration. The bottom panel shows a simulation for a larger ring, with $N = 25000$ sites. Note that the fluctuations in $m$ about equilibrium are smaller than in the cases with $N = 1000$ sites. Why?

## 3.5 Remarks on Ergodic Theory

### 3.5.1 Definition of ergodicity

A mechanical system evolves according to Hamilton's equations of motion. We have seen how such a system is *recurrent* in the sense of Poincaré.

There is a level beyond recurrence called *ergodicity*. In an ergodic system, time averages over intervals $[0, T]$ with $T \to \infty$ may be replaced by phase space averages. The time average of a function $f(\varphi)$ is

defined as

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_t = \lim_{T\to\infty}\frac{1}{T}\int\limits_0^T dt\, f\big(\boldsymbol{\varphi}(t)\big) \quad . \tag{3.84}$$

For a Hamiltonian system, the *phase space average* of the same function is defined by

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_E = \int d\mu\, f(\boldsymbol{\varphi})\,\delta\big(E-H(\boldsymbol{\varphi})\big)\Big/\int d\mu\,\delta\big(E-H(\boldsymbol{\varphi})\big) \quad , \tag{3.85}$$

where $H(\boldsymbol{\varphi}) = H(\boldsymbol{q},\boldsymbol{p})$ is the Hamiltonian, and where $\delta(x)$ is the Dirac $\delta$-function. The energy is fixed to be $E = H\big(\boldsymbol{\varphi}(t=0)\big)$. Thus,

$$\text{ergodicity} \quad\Longleftrightarrow\quad \big\langle f(\boldsymbol{\varphi})\big\rangle_t = \big\langle f(\boldsymbol{\varphi})\big\rangle_E \quad , \tag{3.86}$$

for all smooth functions $f(\boldsymbol{\varphi})$ for which $\big\langle f(\boldsymbol{\varphi})\big\rangle_E$ exists and is finite. Note that we do not average over *all* of phase space. Rather, we average only over a hypersurface along which $H(\boldsymbol{\varphi}) = E$ is fixed, *i.e.* over one of the *level sets* of the Hamiltonian function. This is because the dynamics *preserves the energy*. Ergodicity means that almost all points $\boldsymbol{\varphi}$ will, upon Hamiltonian evolution, move in such a way as to eventually pass through every finite neighborhood on the energy surface, and will spend equal time in equal regions of phase space.

Let $\chi_{\mathcal{R}}(\boldsymbol{\varphi})$ be the characteristic function of a region $\mathcal{R}$:

$$\chi_{\mathcal{R}}(\boldsymbol{\varphi}) = \begin{cases} 1 & \text{if } \boldsymbol{\varphi}\in\mathcal{R} \\ 0 & \text{otherwise,}\end{cases} \tag{3.87}$$

where $H(\boldsymbol{\varphi}) = E$ for all $\boldsymbol{\varphi}\in\mathcal{R}$, so $\dim\mathcal{R} = n-1$ , where $n = 2dN$ is the total dimension of phase space for $N$ point particles in $d$ space dimensions. Then

$$\big\langle \chi_{\mathcal{R}}(\boldsymbol{\varphi})\big\rangle_t = \lim_{T\to\infty}\left(\frac{\text{time spent in }\mathcal{R}}{T}\right) \quad . \tag{3.88}$$

If the system is ergodic, then

$$\big\langle \chi_{\mathcal{R}}(\boldsymbol{\varphi})\big\rangle_t = P(\mathcal{R}) = \frac{D_{\mathcal{R}}(E)}{D(E)} \quad , \tag{3.89}$$

where $P(\mathcal{R})$ is the *a priori* probability to find $\boldsymbol{\varphi}\in\mathcal{R}$, based solely on the relative volumes of $\mathcal{R}$ and of the entire energy-restricted phase space. Here,

$$D(E) = \int d\mu\,\delta\big(E-H(\boldsymbol{\varphi})\big) = \int\limits_{\mathcal{S}_E} d\Sigma_E \quad , \tag{3.90}$$

called the *density of states*, is the surface area of phase space at energy $E$, and $D_{\mathcal{R}}(E)$ is defined below in eqn. 3.93. The hypersurface $\mathcal{S}_E$ is the set of points $\boldsymbol{\varphi}$ satisfying $H(\boldsymbol{\varphi}) = E$, and the *invariant differential surface element* $d\Sigma_E$ is defined as follows. We can write the differential phase space volume $d\mu$ as the

product $d\mu = dS_E\, d\zeta_E$, where $dS_E$ is the differential surface element for the level set $\mathcal{S}_E$ and $\zeta_E$ is a phase space coordinate locally perpendicular to $\mathcal{S}_E$. We then define[12]

$$d\Sigma_E \equiv \frac{dS_E}{|\boldsymbol{\nabla} H|}\bigg|_{H(\boldsymbol{\varphi})=E} \quad , \tag{3.91}$$

and we may now write $d\mu = dE\, d\Sigma_E$. Note that we may also express $D(E)$ as

$$D(E) = \frac{d}{dE}\int d\mu\, \Theta\big(E - H(\boldsymbol{\varphi})\big) \equiv \frac{d\Omega(E)}{dE} \quad , \tag{3.92}$$

where $\Omega(E) = \int d\mu\, \Theta\big(E - H(\boldsymbol{\varphi})\big)$ is the volume of phase space over which $H(\boldsymbol{\varphi}) < E$. The density of states for the subset $\mathcal{R}$ is defined as

$$D_{\mathcal{R}}(E) = \int\limits_{\mathcal{R}} d\Sigma_E \quad . \tag{3.93}$$

Note that $\mathcal{R} \subset \mathcal{S}_E$.

### 3.5.2 The microcanonical ensemble

The distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{D(E)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{\int d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\big)} \quad , \tag{3.94}$$

defines the *microcanonical ensemble* ($\mu$CE) of Gibbs. We could also write

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_E = \frac{1}{D(E)}\int\limits_{\mathcal{S}_E} d\Sigma_E\, f(\boldsymbol{\varphi}) \quad , \tag{3.95}$$

integrating over the hypersurface $\mathcal{S}_E$ rather than the entire phase space.

### 3.5.3 Ergodicity and mixing

Just because a system is ergodic, it doesn't necessarily mean that $\varrho(\varphi, t) \to \varrho^{\rm eq}(\boldsymbol{\varphi})$, for consider the following motion on the toroidal space $\big\{\boldsymbol{\varphi} = (q,p)\,\big|\, 0 \le q < 1 \quad , \; 0 \le p < 1\big\}$, where we identify opposite edges, *i.e.* we impose periodic boundary conditions. We also take $q$ and $p$ to be dimensionless, for simplicity of notation. Let the dynamics be given by $\dot q = 1$ and $\dot p = \alpha$. The motion is then $q(t) = q_0 + t$ and $p(t) = p_0 + \alpha t$. Thus the phase curves are given by $p = p_0 + \alpha(q - q_0)$.

[12] Recall that the phase space coordinates don't all have the same units! $N$ of the coordinates have units of position and $N$ have units of momentum. Furthermore, some may be angles and some angular momenta. However in any case $d\mu$ has units $\mathrm{A}^N$, where A stands for action, *i.e.* $[d\mu] = (\mathrm{ML}^2/\mathrm{T})^N$. Thus while the product $d\mu = dS_E\, d\zeta_E$ has units of $\mathrm{A}^N$, individually the units of $dS_E$ and $d\zeta_E$ vary along the hypersurface $\mathcal{S}_E$! However, the invariant differential surface element $d\Sigma_E$ always has units of $\mathrm{A}^N/\mathrm{E}$. To resolve any confusion, one may choose to rescale so that all phase space coordinates are dimensionless.

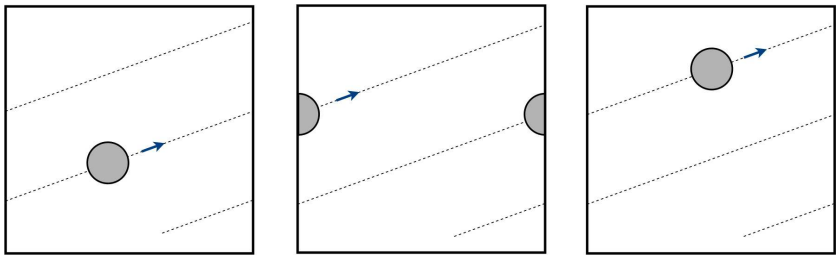


Figure 3.7: Constant phase space velocity at an irrational angle over a toroidal phase space is ergodic, but not mixing. A circle remains a circle, and a blob remains a blob.

Now consider the average of some function $f(q,p)$. We can write $f(q,p)$ in terms of its Fourier transform,

$$f(q,p) = \sum_{m,n} \hat{f}_{m,n}\, e^{2\pi i(mq+np)} \quad . \tag{3.96}$$

We have, then,

$$f\big(q(t),p(t)\big) = \sum_{m,n} \hat{f}_{m,n}\, e^{2\pi i(mq_0+np_0)}\, e^{2\pi i(m+\alpha n)t} \quad . \tag{3.97}$$

We can now perform the time average of $f$:

$$\begin{aligned} \big\langle f(q,p)\big\rangle_t &= \hat{f}_{0,0} + \lim_{T\to\infty} \frac{1}{T} {\sum_{m,n}}' \hat{f}_{m,n}\, e^{2\pi i(mq_0+np_0)}\, \frac{e^{2\pi i(m+\alpha n)T}-1}{2\pi i(m+\alpha n)} \\ &= \hat{f}_{0,0} \quad \text{if } \alpha \notin \mathbb{Q} \quad . \end{aligned} \tag{3.98}$$

Clearly,

$$\big\langle f(q,p)\big\rangle_E = \int_0^1 dq \int_0^1 dp\, f(q,p) = \hat{f}_{0,0} = \big\langle f(q,p)\big\rangle_t \quad , \tag{3.99}$$

so the system is ergodic. However, if $\alpha = r/s$ with $r,s \in \mathbb{Z}$ and $\gcd(r,s) = 1$ (*i.e.* $r$ and $s$ are relatively prime), then $\exp\big\{2\pi i(m+\alpha n)t\big\} = 1$ whenever $(m,n) = (kr,-ks)$ for any $k \in \mathbb{Z}$. Thus,

$$\big\langle f(q,p)\big\rangle_t = \sum_{k=-\infty}^{\infty} \hat{f}_{kr,-ks}\, e^{2\pi ik(rq_0-sp_0)} \quad , \tag{3.100}$$

which is not the same as $\big\langle f(q,p)\big\rangle_{\mu\text{CE}} = \hat{f}_{0,0}$.

The situation is depicted in fig. 3.7. If we start with the characteristic function of a disc,

$$\varrho(q,p,t=0) = \Theta\big(a^2 - (q-q_0)^2 - (p-p_0)^2\big) \quad , \tag{3.101}$$

then it remains the characteristic function of a disc:

$$\varrho(q,p,t) = \Theta\big(a^2 - (q-q_0-t)^2 - (p-p_0-\alpha t)^2\big) \quad , \tag{3.102}$$

For an example of a transition to ergodicity in a simple dynamical Hamiltonian model, see §3.10.

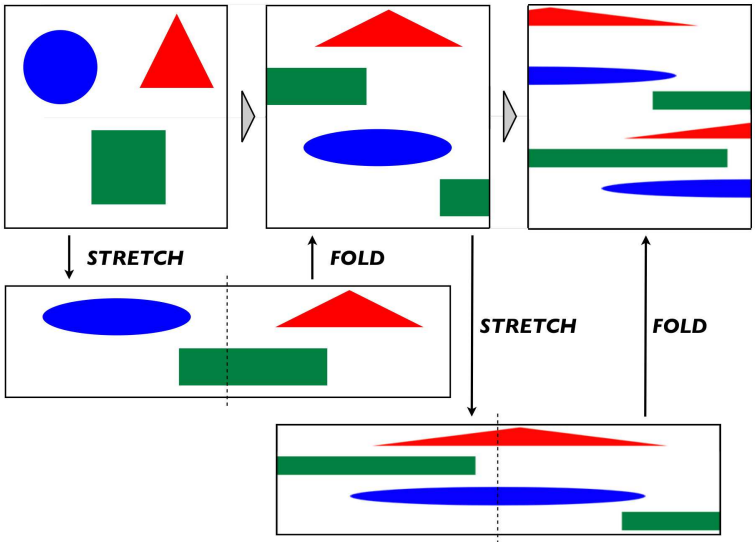



Figure 3.8: The baker's transformation is a successive stretching, cutting, and restacking.

A stronger condition one could impose is the following. Let $A$ and $B$ be subsets of $\mathcal{S}_E$. Define the *measure*

$$\nu(A) = \frac{\int d\Sigma_E\, \chi_A(\boldsymbol{\varphi})}{\int d\Sigma_E} = \frac{D_A(E)}{D(E)} \quad , \tag{3.103}$$

where $\chi_A(\boldsymbol{\varphi})$ is the characteristic function of $A$. The measure of a set $A$ is the fraction of the energy surface $\mathcal{S}_E$ covered by $A$. This means $\nu(\mathcal{S}_E) = 1$, since $\mathcal{S}_E$ is the entire phase space at energy $E$. Now let $g$ be a volume-preserving map on phase space. Given two measurable sets $A$ and $B$, we say that a system is *mixing* if

$$\text{mixing} \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} \nu\Big(g^n A \cap B\Big) = \nu(A)\,\nu(B) \quad . \tag{3.104}$$

In other words, the fraction of $B$ covered by the $n^{\rm th}$ iterate of $A$, *i.e.* $g^n A$, is, as $n \to \infty$, simply the fraction of $\mathcal{S}_E$ covered by $A$. The iterated map $g^n$ distorts the region $A$ so severely that it eventually spreads out 'evenly' over the entire energy hypersurface. Of course by 'evenly' we mean 'with respect to any finite length scale', because at the very smallest scales, the phase space density is still locally constant as one evolves with the dynamics.

Mixing means that

$$\begin{aligned} \big\langle f(\boldsymbol{\varphi})\big\rangle &= \int d\mu\, \varrho(\boldsymbol{\varphi},t)\, f(\boldsymbol{\varphi}) \xrightarrow[t\to\infty]{} \int d\mu\, f(\boldsymbol{\varphi})\, \delta\big(E - H(\boldsymbol{\varphi})\big) \Big/ \int d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\big) \\ &\equiv \mathsf{Tr}\Big[f(\boldsymbol{\varphi})\, \delta\big(E - H(\boldsymbol{\varphi})\big)\Big] \Big/ \mathsf{Tr}\Big[\delta\big(E - H(\boldsymbol{\varphi})\big)\Big] \quad . \end{aligned} \tag{3.105}$$

Physically, we can imagine regions of phase space being successively stretched and folded. During the stretching process, the volume is preserved, so the successive stretch and fold operations map phase space back onto itself.

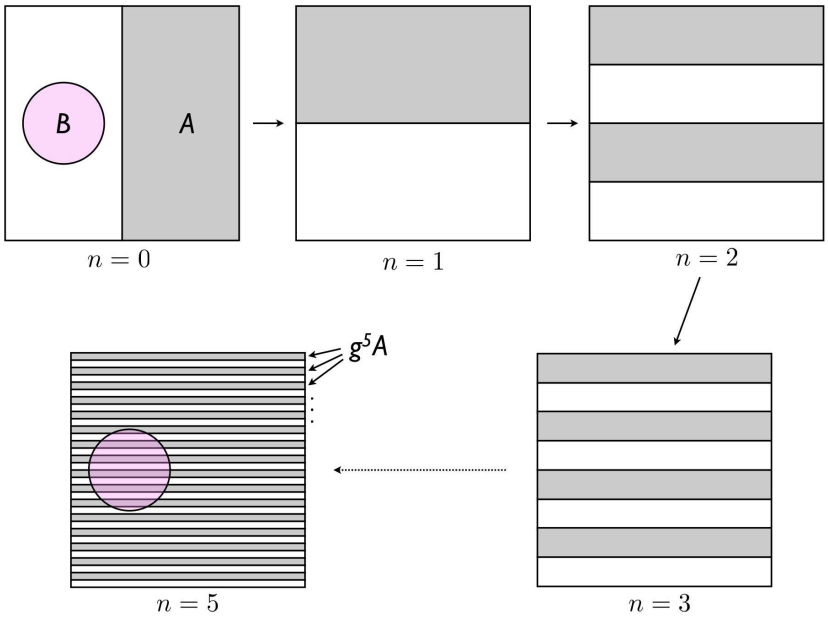



Figure 3.9: The multiply iterated baker's transformation. The set A covers half the phase space and its area is preserved under the map. Initially, the fraction of B covered by A is zero. After many iterations, the fraction of B covered by $g^n$A approaches $\frac{1}{2}$.

An example of a mixing system is the *baker's transformation*, depicted in fig. 3.8, and defined by

$$g(q,p) = \begin{cases} (2q\,,\,\tfrac{1}{2}p) & \text{if} \quad 0 \le q < \tfrac{1}{2} \\ (2q-1\,,\,\tfrac{1}{2}p + \tfrac{1}{2}) & \text{if} \quad \tfrac{1}{2} \le q < 1 \end{cases} \quad . \tag{3.106}$$

Note that $g$ is invertible and volume-preserving. The baker's transformation consists of an initial stretch in which $q$ is expanded by a factor of two and $p$ is contracted by a factor of two, which preserves the total volume. The system is then mapped back onto the original area by cutting and restacking, which we can call a 'fold'. The inverse transformation is accomplished by stretching first in the vertical ($p$) direction and squashing in the horizontal ($q$) direction, followed by a slicing and restacking. Explicitly,

$$g^{-1}(q,p) = \begin{cases} (\tfrac{1}{2}q\,,\,2p) & \text{if} \quad 0 \le p < \tfrac{1}{2} \\ (\tfrac{1}{2}q + \tfrac{1}{2}\,,\,2p-1) & \text{if} \quad \tfrac{1}{2} \le p < 1 \end{cases} \quad . \tag{3.107}$$

Another example of a mixing system is Arnold's 'cat map'[13]

$$g(q,p) = \big(\,[q+p]\,,\,[q+2p]\,\big) \quad , \tag{3.108}$$

where $[x]$ denotes the fractional part of $x$. One can write this in matrix form as

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}}^{M} \begin{pmatrix} q \\ p \end{pmatrix} \mod \mathbb{Z}^2 \quad . \tag{3.109}$$

[13]The cat map gets its name from its initial application, by Arnold, to the image of a cat's face.

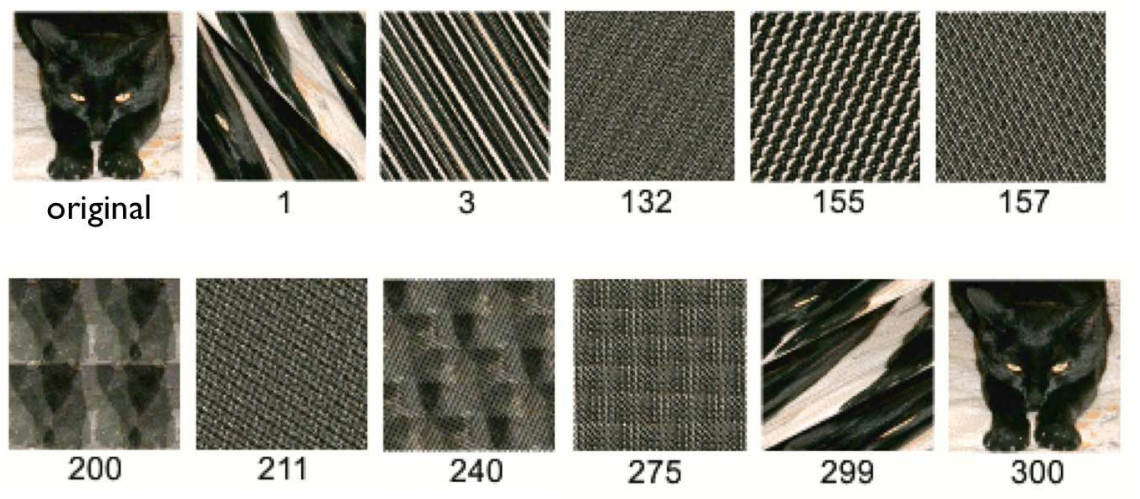



Figure 3.10: The Arnold cat map applied to an image of $150 \times 150$ pixels. After 300 iterations, the image repeats itself. (*Source: Wikipedia*)

The matrix $M$ is very special because it has integer entries and its determinant is $\det M = 1$. This means that the inverse also has integer entries. The inverse transformation is then

$$\begin{pmatrix} q \\ p \end{pmatrix} = \overbrace{\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}}^{M^{-1}} \begin{pmatrix} q' \\ p' \end{pmatrix} \mod \mathbb{Z}^2 \quad . \tag{3.110}$$

Now for something cool. Suppose that our image consists of a set of discrete points at $(n_1/k\,,\, n_2/k)$, where the denominator $k \in \mathbb{Z}$ is fixed, and where $n_1$ and $n_2$ range over the set $\{1, \ldots, k\}$. Clearly $g$ and its inverse preserve this set, since the entries of $M$ and $M^{-1}$ are integers. If there are two possibilities for each pixel (say off and on, or black and white), then there are $2^{(k^2)}$ possible images, and the cat map will map us invertibly from one image to another. Therefore it must exhibit Poincaré recurrence! This phenomenon is demonstrated vividly in fig. 3.10, which shows a $k = 150$ pixel (square) image of a cat subjected to the iterated cat map. The image is stretched and folded with each successive application of the cat map, but after 300 iterations the image is restored! How can this be if the cat map is mixing? The point is that only the discrete set of points $(n_1/k\,,\, n_2/k)$ is periodic. Points with different denominators will exhibit a different periodicity, and points with irrational coordinates will in general never return to their exact initial conditions, although recurrence says they will come arbitrarily close, given enough iterations. The baker's transformation is also different in this respect, since the denominator of the $p$ coordinate is doubled upon each successive iteration.

The student should now contemplate the hierarchy of dynamical systems depicted in fig. 3.11, understanding the characteristic features of each successive refinement[14].

[14]There is something beyond mixing, called a *$K$-system*. A $K$-system has positive Kolmogorov-Sinai entropy. For such a system, closed orbits separate exponentially in time, and consequently the Liouvillian $L$ has a Lebesgue spectrum with denumerably infinite multiplicity.

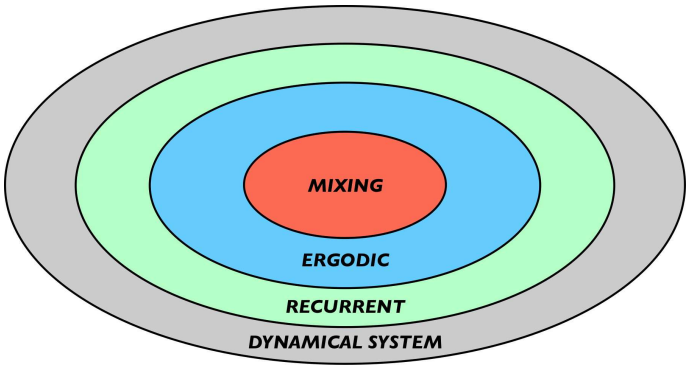



Figure 3.11: The hierarchy of dynamical systems.

## 3.6 Thermalization of Quantum Systems

### 3.6.1 Quantum dephasing

Thermalization of quantum systems is fundamentally different from that of classical systems. Whereas time evolution in classical mechanics is in general a nonlinear dynamical system, the Schrödinger equation for time evolution in quantum mechanics is linear: $i\hbar\,\partial\Psi/\partial t = \hat{H}\Psi$, where $\hat{H}$ is a many-body Hamiltonian. In classical mechanics, the thermal state is constructed by time evolution – this is the content of the ergodic theorem. In quantum mechanics, as we shall see, the thermal distribution must be encoded in the eigenstates themselves.

Let us assume an initial condition at $t = 0$ with $|\Psi(0)\rangle = \sum_\alpha C_\alpha\,|\Psi_\alpha\rangle$, where $\{|\,\Psi_\alpha\,\rangle\}$ is an orthonormal eigenbasis for $\hat{H}$ satisfying $\hat{H}\,|\Psi_\alpha\rangle = E_\alpha\,|\Psi_\alpha\rangle$. The expansion coefficients satisfy $C_\alpha = \langle\Psi_\alpha|\Psi(0)\rangle$ and $\sum_\alpha |C_\alpha|^2 = 1$. Normalization requires $\langle\,\Psi(0)\,|\,\Psi(0)\,\rangle = \sum_\alpha |C_\alpha|^2 = 1$.

The time evolution of $|\Psi\rangle$ is then given by

$$|\Psi(t)\rangle = \sum_\alpha C_\alpha\, e^{-iE_\alpha t/\hbar}\,|\Psi_\alpha\rangle \quad . \tag{3.111}$$

The energy is distributed according to the time-independent function

$$P(E) = \langle\,\Psi(t)\,|\,\delta(E-\hat{H})\,|\,\Psi(t)\,\rangle = \sum_\alpha |C_\alpha|^2\,\delta(E-E_\alpha) \quad . \tag{3.112}$$

Thus, the average energy is time-independent and is given by

$$\langle E\rangle = \langle\,\Psi(t)\,|\,\hat{H}\,|\,\Psi(t)\,\rangle = \int\limits_{-\infty}^{\infty} dE\,P(E)\,E = \sum_\alpha |C_\alpha|^2\,E_\alpha \quad . \tag{3.113}$$

The root mean square fluctuations of the energy are given by

$$(\Delta E)_{\rm rms} = \Big\langle \big(E-\langle E\rangle\big)^2 \Big\rangle^{1/2} = \sqrt{\sum_\alpha |C_\alpha|^2\,E_\alpha^2 - \Big(\sum_\alpha |C_\alpha|^2\,E_\alpha\Big)^2} \quad . \tag{3.114}$$

Typically we assume that the distribution $P(E)$ is narrowly peaked about $\langle E \rangle$, such that $(\Delta E)_{\rm rms} \ll \langle E \rangle - E_0$ , where $E_0$ is the ground state energy. Note that $P(E) = 0$ for $E < E_0$, *i.e.* the eigenspectrum of $\hat{H}$ is bounded from below.

Now consider a general quantum observable described by an operator $\mathcal{A}$. We have

$$\langle \mathcal{A}(t) \rangle = \langle\, \Psi(t) \,|\, \mathcal{A} \,|\, \Psi(t) \,\rangle = \sum_{\alpha,\beta} C_\alpha^* \, C_\beta \, e^{i(E_\alpha - E_\beta)t/\hbar} \, \mathcal{A}_{\alpha\beta} \quad , \tag{3.115}$$

where $\mathcal{A}_{\alpha\beta} = \langle\, \Psi_\alpha \,|\, \mathcal{A} \,|\, \Psi_\beta \,\rangle$. In the limit of large times, we have

$$\langle \mathcal{A} \rangle_t \equiv \lim_{T\to\infty} \frac{1}{T} \int\limits_0^T dt \, \langle \mathcal{A}(t) \rangle = \sum_\alpha |C_\alpha|^2 \, \mathcal{A}_{\alpha\alpha} \quad . \tag{3.116}$$

This implies that all coherence between different eigenstates is lost in the long time limit, due to *dephasing*. Note though that there is a subtlety here associated with the order of limits, for if we take the thermodynamic limit first, then for systems with a continuous spectrum the energy gaps $|E_\alpha - E_\beta|$ become arbitrarily small. In many-body systems, the smallest energy gaps typically are exponentially small in the system size, which means that we must average over absurdly long times in order for the coherences between different eigenstates separated by an exponentially small gap to cancel.

### 3.6.2 Eigenstate thermalization hypothesis

The essential ideas behind the *eigenstate thermalization hypothesis* (ETH) were described independently by J. Deutsch (1991) and by M. Srednicki (1994). The argument goes as follows. If the total energy is the only conserved quantity, and if $\mathcal{A}$ is a local, translationally-invariant, few-body operator, then the time average $\langle \mathcal{A} \rangle$ is given by its microcanonical value,

$$\langle \mathcal{A} \rangle_t = \sum_\alpha |C_\alpha|^2 \, \mathcal{A}_{\alpha\alpha} = \frac{\sum_\alpha \mathcal{A}_{\alpha\alpha} \, \Theta(E_\alpha \in I)}{\sum_\alpha \Theta(E_\alpha \in I)} \equiv \langle \mathcal{A} \rangle_E, \tag{3.117}$$

where $I = \big[E, E + \Delta E\big]$ is an energy interval of width $\Delta E$. So once again, time averages are micro canonical averages[15].

But how is it that this is the case? The hypothesis of Deutsch and of Srednicki is that thermalization in isolated and bounded quantum systems occurs *at the level of individual eigenstates*. That is, for all eigenstates $|\Psi_\alpha\rangle$ with $E_\alpha \in I$, one has $\mathcal{A}_{\alpha\alpha} = \langle \mathcal{A} \rangle_{E_\alpha}$. This means that *thermal information is encoded in each eigenstate*. This is called the *eigenstate thermalization hypothesis* (ETH).

An equivalent version of the ETH is the following scenario. Suppose we have an infinite or extremely large quantum system $U$ (the 'universe') fixed in an eigenstate $|\Psi_\alpha\rangle$. Then form the projection operator $P_\alpha = |\Psi_\alpha\rangle\langle\Psi_\alpha|$. Projection operators satisfy $P^2 = P$ and their eigenspectrum consists of one unit eigenvalue $\lambda = 1$ and the rest of the eigenvalues are zero[16]. Now consider a partition of $U = W \cup S$, where

[15] This also presumes a sufficiently narrow energy distribution and smooth, approximately constant diagonal matrix elements.

[16] More generally, we could project onto a $K$-dimensional subspace, in which case there would be $K$ eigenvalues of $+1$ and $N - K$ eigenvalues of $0$, where $N$ is the dimension of the entire vector space.

$W \gg S$. We imagine $S$ to be the 'system' and $W$ the 'world'. We can always decompose the state $|\Psi_\alpha\rangle$ in a complete product basis for $W$ and $S$, *viz.*

$$|\Psi_\alpha\rangle = \sum_{i=1}^{N_W} \sum_{j=1}^{N_S} \Phi_{ij}^\alpha \, |\psi_i^W\rangle \otimes |\psi_j^S\rangle \quad . \tag{3.118}$$

Here $N_{W/S}$ is the size of the basis for $W/S$. The *reduced density matrix* for $S$ is defined as

$$\varrho_S = \mathop{\mathsf{Tr}}_W P_\alpha = \sum_{j,j'=1}^{N_S} \left( \sum_{i=1}^{N_W} \Phi_{ij}^\alpha \, \Phi_{ij'}^\alpha \right) |\psi_j^S\rangle\langle\psi_{j'}^S| \quad . \tag{3.119}$$

The claim is that $\varrho_S$ approximates a *thermal density matrix* on $S$, *i.e.*

$$\varrho_S \approx \frac{1}{Z_S} \, e^{-\beta \hat{H}_S} \quad , \tag{3.120}$$

where $\hat{H}_S$ is some local Hamiltonian on $S$ and $Z_S = \mathsf{Tr}\; e^{-\beta \hat{H}_S}$, so that $\varrho_S$ is normalized with $\mathsf{Tr}\, \varrho_S = 1$. A number of issues remain to be clarified:

(i) What do we mean by "approximates"?

(ii) What do we mean by $\hat{H}_S$?

(iii) What do we mean by the temperature $T$?

We address these in reverse order. The temperature $T$ of an eigenstate $|\Psi_\alpha\rangle$ of a Hamiltonian $\hat{H}$ is defined by setting its energy density $E_\alpha/V_U$ to the thermal energy density, *i.e.*

$$\frac{E_\alpha}{V} = \frac{1}{V} \, \frac{\mathsf{Tr}\; \hat{H} \, e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} \quad . \tag{3.121}$$

Here, $\hat{H} = \hat{H}_U$ is the full Hamiltonian of the universe $U = W \cup S$, and $V = V_U$. Our intuition is that $\hat{H}_S$ should reflect a restriction of the original Hamiltonian $\hat{H}_U$ to the system $S$. What should be done, though, about the interface parts of $\hat{H}_U$ which link $S$ and $W$? For lattice Hamiltonians, we can simply but somewhat arbitrarily cut all the bonds coupling $S$ and $W$. But we could easily imagine some other prescription, such as halving the coupling strength along all such interface bonds. Indeed, the definition of $H_S$ is somewhat arbitrary. However, so long as we use $\varrho_S$ to compute averages of *local operators* which lie sufficiently far from the boundary of $S$, the precise details of how we truncate $\hat{H}_U$ to $\hat{H}_S$ are unimportant. This brings us to the first issue: the approximation of $\varrho_S$ by its Gibbs form in eqn. 3.120 is only valid when we consider averages of local operators lying within the bulk of $S$. This means that we must only examine operators whose support is confined to regions greater than some distance $\xi_T$ from $\partial S$, where $\xi_T$ is a *thermal correlation length*. This, in turn, requires that $L_S \gg \xi_T$, *i.e.* the region $S$ is very large on the scale of $\xi_T$. How do we define $\xi_T$? For a model such as the Ising model, it can be taken to be the usual correlation length obtained from the spin-spin correlation function $\langle \sigma_{\boldsymbol{r}} \, \sigma_{\boldsymbol{r}'} \rangle_T$. More generally,

we may choose the largest correlation length from among the correlators of all the independent local operators in our system. Again, the requirement is that $\exp(-d_\partial(\boldsymbol{r})/\xi_T) \ll 1$, where $d_\partial(\boldsymbol{r})$ is the shortest distance from the location of our local operator $\mathcal{O}_{\boldsymbol{r}}$ to the boundary of $S$. At criticality, the exponential is replaced by a power law $(d_\partial(\boldsymbol{r})/a)^{-p}$, where $p$ is a critical exponent and $a$ is a microscopic scale, with $d_\partial(\boldsymbol{r}) \gg a$. Another implicit assumption here is that $V_S \ll V_W$. For a general $|\,\Psi\,\rangle$, the reduced density matrix and entanglement entropy can be expressed using the machinery of the singular value decomposition theorem, and are discussed in the appendix §3.6.4.

### 3.6.3 Entanglement and entanglement entropy

The *von Neumann entanglement entropy* between $S$ and $W$ is defined to be

$$S_{\rm vN} = -\,\mathrm{Tr}\;\varrho_S \log \varrho_S = -\sum_j \lambda_j \log \lambda_j \quad , \tag{3.122}$$

where $\{\lambda_j\}$ are the eigenvalues of the reduced density matrix $\varrho_S$. Each eigenvalue $\lambda_j$ lies on the unit interval $[0,1]$, subject to normalization $\sum_j \lambda_j = 1$. For the special case of a *pure state*, the density matrix is a projection operator, *i.e.* $\varrho = |\Psi\rangle\langle\Psi|$, as we have assumed for the universe $U = W \cup S$. Pure state density matrices have a single nonzero eigenvalue $\lambda_1 = 1$, corresponding to the eigenvector $|\Psi\rangle$, with all other eigenvalues $\lambda_{j>1} = 0$. In general, even when the full density matrix of the system is a projector, *i.e.* when $\varrho = |\Psi\rangle\langle\Psi|$ where $|\Psi\rangle$ lies in the Hilbert space for $U$, the reduced density matrix $\varrho_S = \mathrm{Tr}_W \varrho$ does not correspond to a pure state. *I.e.*, as stated above, $\varrho_S$ has a spectrum of eigenvalues $\{\lambda_j\}$ each of which lies on the interval $[0,1]$, subject to normalization $\sum_j \lambda_j = 1$. If $|\Psi\rangle = |\phi_W\rangle \otimes |\chi_S\rangle$ factors, then the reduced density matrix $\varrho_S = \mathrm{Tr}_W |\Psi\rangle\langle\Psi| = |\chi_S\rangle\langle\chi_S|$ is pure, and $S_{\rm vN} = 0$. One then has that $W$ and $S$ are *unentangled*.

For a simple example, consider a system consisting of two spin-$\frac{1}{2}$ objects, with wavefunction

$$|\,\Psi\,\rangle = \alpha\,|\uparrow\downarrow\rangle + \beta\,|\downarrow\uparrow\rangle \tag{3.123}$$

and with $|\alpha|^2 + |\beta|^2 = 1$. Then the two spin density matrix is

$$\varrho = |\alpha|^2\,|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\beta|^2\,|\downarrow\uparrow\rangle\langle\downarrow\uparrow| + \alpha\beta^*\,|\uparrow\downarrow\rangle\langle\downarrow\uparrow| + \alpha^*\beta\,|\downarrow\uparrow\rangle\langle\uparrow\downarrow| \quad . \tag{3.124}$$

Taking the trace over the first spin, the reduced density matrix $\tilde{\varrho}$ is given by

$$\tilde{\varrho} = \mathop{\mathrm{Tr}}_1 \varrho = |\alpha|^2\,|\downarrow\rangle\langle\downarrow| + |\beta|^2\,|\uparrow\rangle\langle\uparrow| = \begin{pmatrix} |\beta|^2 & 0 \\ 0 & |\alpha|^2 \end{pmatrix} \quad . \tag{3.125}$$

The entanglement entropy is then

$$S_{\rm vN} = -|\alpha|^2 \log|\alpha|^2 - |\beta|^2 \log|\beta|^2 \quad , \tag{3.126}$$

and if our initial state is a singlet with $\alpha = \frac{1}{\sqrt{2}}$ and $\beta = -\frac{1}{\sqrt{2}}$, we obtain $S_{\rm vN} = \log 2$.[17]

[17] In fact, any initial state with $\alpha = \frac{1}{\sqrt{2}}$ and $\beta = \frac{1}{\sqrt{2}}\,e^{i\delta}$ will have entanglement entropy $\log 2$, independent of $\delta$.

### 3.6.4 Singular value decomposition and entanglement

Consider a quantum system whose wavefunction may be expressed as a vector in a Hilbert space $\mathcal{H}$ where $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ is a direct product of Hilbert subspaces. For example, $\mathcal{H}$ could be the Hilbert space for $K$ spins, and $\mathcal{H}_{A,B}$ disjoint Hilbert spaces for $K_A$ and $K_B$ spins, respectively, with $K = K_A + K_B$. Now consider orthonormal basis states $|\,\psi_a^A\,\rangle$ and $|\,\psi_b^B\,\rangle$ on $\mathcal{H}_{A,B}$, where $a \in \{1, \ldots, N_A\}$ and $b \in \{1, \ldots, N_B\}$, where $N_{A,B}$ are the dimensions of $\mathcal{H}_{A,B}$. Any wavefunction $|\,\Psi\,\rangle \in \mathcal{H}$ can then be expressed as

$$|\,\Psi\,\rangle = \sum_{a=1}^{N_A} \sum_{b=1}^{N_B} \Psi_{ab}\, |\,\psi_a^A\,\rangle \otimes |\,\psi_b^B\,\rangle \quad , \tag{3.127}$$

where $\Psi_{ab}$ is a complex $N_A \times N_B$ matrix. Any such matrix has a *singular value decomposition* [18] given by $\Psi = UDV^\dagger$, where (i) $U$ is a complex rectangular matrix of dimensions $N_A \times N_{\rm s}$ satisfying $U^\dagger U = \mathbb{I}_{N_{\rm s}\times N_{\rm s}}$, (ii) $D$ is a square diagonal matrix of dimensions $N_{\rm s} \times N_{\rm s}$ with positive real entries, (iii) $V$ is a complex rectangular matrix of dimensions $N_B \times N_{\rm s}$ satisfying $V^\dagger V = \mathbb{I}_{N_{\rm s}\times N_{\rm s}}$, and (iv) $N_{\rm s} \le \min(N_A, N_B)$. Writing $D = \mathrm{diag}\{d_1, \ldots, d_{N_{\rm s}}\}$ with $p \in \{1, \ldots, N_{\rm s}\}$ the *singular values*, we obtain

$$|\,\Psi\,\rangle = \sum_{p=1}^{N_{\rm s}} d_p\, |\,\chi_p^A\,\rangle \otimes |\,\chi_p^B\,\rangle \quad , \tag{3.128}$$

where

$$|\,\chi_p^A\,\rangle = \sum_{a=1}^{N_A} U_{ap}\, |\,\psi_a^A\,\rangle \qquad , \qquad |\,\chi_p^B\,\rangle = \sum_{b=1}^{N_B} V_{bp}\, |\,\psi_b^B\,\rangle \quad . \tag{3.129}$$

Note that $\langle\,\chi_a^A\,|\,\chi_{a'}^A\,\rangle = \delta_{aa'}$ and $\langle\,\chi_b^B\,|\,\chi_{b'}^B\,\rangle = \delta_{bb'}$.

If we trace out over the $B$ degrees of freedom, the reduced density matrices on $A$ and $B$ are

$$\begin{aligned} \varrho_A &= \mathop{\mathrm{Tr}}_B |\,\Psi\,\rangle\langle\,\Psi\,| = \sum_{p=1}^{N_{\rm s}} d_p^2\, |\,\chi_p^A\,\rangle\langle\,\chi_p^A\,| \\ \varrho_B &= \mathop{\mathrm{Tr}}_A |\,\Psi\,\rangle\langle\,\Psi\,| = \sum_{p=1}^{N_{\rm s}} d_p^2\, |\,\chi_p^B\,\rangle\langle\,\chi_p^B\,| \end{aligned} \tag{3.130}$$

Since $\mathrm{Tr}\,\varrho_A = \mathrm{Tr}\,\varrho_B = 1$, we must have that the squares of the singular values sum to unity: $\sum_{p=1}^{N_{\rm s}} d_p^2 = 1$. The entanglement entropy is then

$$S_{\rm vN} \equiv S_A = S_B = -\sum_{p=1}^{N_{\rm s}} d_p^2 \log(d_p^2) \quad . \tag{3.131}$$

Writing $d_p \equiv \exp(-\xi_p/2)$, we have

$$\varrho_A = \sum_{p=1}^{N_{\rm s}} e^{-\xi_p}\, |\,\chi_p^A\,\rangle\langle\,\chi_p^A\,| \equiv e^{-\hat{K}} \quad , \tag{3.132}$$

[18] See §3.7.2 below as well.

and we identify the *entanglement Hamiltonian* as

$$\hat{K} \equiv \sum_{p=1}^{N_{\rm s}} \xi_p \, |\, \chi_p^A \,\rangle\langle\, \chi_p^A \,| \quad , \tag{3.133}$$

with $\sum_{p=1}^{N_{\rm s}} \exp(-\xi_p) = 1$. Thus, $\mathsf{spec}(\hat{K}) = \{\xi_1, \ldots, \xi_{N_{\rm s}}\}$ is the spectrum of eigenvalues of the entanglement Hamiltonian. In the language of eqn. 3.120, we have $\hat{K} = \beta \hat{H}_S + k_{\rm B} T \log Z_S$, with $A = S$ and $B = W$. In general there is no guarantee that the entanglement Hamiltonian $\hat{K}$ is local. For thermalizing local systems, however, sufficiently small subsystems are expected to have reduced density matrices that approach the corresponding thermal reduced density matrices.

### 3.6.5 Digression on random matrices

Random matrices are elements of a particular *ensemble* which is specified by a *symmetry class*. The probability distribution for a random matrix $H$ is

$$P_\beta(H) = C_\beta \, e^{-\beta \, \mathsf{Tr}\,(H^2)/2} \quad , \tag{3.134}$$

with $C_\beta$ a normalization constant, and where $\beta = 1$ for real symmetric, $\beta = 2$ for complex Hermitian, and $\beta = 4$ for quaternion real Hermitian Gaussian ensembles. These are respectively called the Gaussian Orthogonal Ensemble (GOE), Gaussian Unitary Ensemble (GUE), and Gaussian Symplectic Ensemble (GSE), and comprise the 'classical' random matrix ensembles (there are others). They are so named because the ensembles remain invariant under orthogonal $\mathsf{O}(N)$, unitary $\mathsf{U}(N)$, and symplectic $\mathsf{Sp}(2N)$ transformations, respectively, where $N$ is the dimension[19]. The Gaussian probability distributions in eqn. 3.134 possess these invariances, although the invariances themselves do not require a Gaussian distribution[20]. The three ensembles correspond, respectively, to time-reversal invariant systems with integer spin (GOE), systems with no time-reversal invariance (GUE), and time-reversal invariant systems with half-odd-integer spin (GSE). The matrix elements of $H$ in each of these ensembles take the form

$$H_{mn} = \begin{cases} H_{mn}^{(0)} & \text{(GOE)} \\ H_{mn}^{(0)} + i H_{mn}^{(1)} & \text{(GUE)} \\ H_{mn}^{(0)} \, \mathbb{1} + H_{mn}^{(1)} \, \mathbb{i} + H_{mn}^{(2)} \, \mathbb{j} + H_{mn}^{(3)} \, \mathbb{k} & \text{(GSE)} \end{cases} , \tag{3.135}$$

with $m \le n$, where $H_{mn}^{(\mu)} \in \mathbb{R}$ for $\mu \in \{0, 1, 2, 3\}$, $H_{nm}^{(0)} = H_{mn}^{(0)}$, and $H_{nm}^{(r)} = -H_{mn}^{(r)}$ for $r \in \{1, 2, 3\}$. Here, $\{\mathbb{i}, \mathbb{j}, \mathbb{k}\}$ are the quaternions (see below). The measure $dH$ in these three ensembles is defined as

$$dH = \begin{cases} \prod_{j \le k} dH_{jk}^{(0)} & \text{(GOE)} \\ \prod_{j \le k} dH_{jk}^{(0)} \prod_{m<n} dH_{mn}^{(1)} & \text{(GUE)} \\ \prod_{j \le k} dH_{jk}^{(0)} \prod_{m<n} \prod_{r=1}^{3} dH_{mn}^{(r)} & \text{(GSE)} \end{cases} . \tag{3.136}$$

[19] For quaternionic matrices, expressing each quaternionic number as a $2 \times 2$ matrix (see below), the matrix rank becomes $2N$.

[20] More general distributions respecting these invariances are of the form $P(H) \propto e^{\mathsf{Tr}\, V(H)}$ where $V(H) = V(S^{-1} H S)$ and $S$ an element of $\mathsf{O}(N)$ or $\mathsf{U}(N)$ or $\mathsf{Sp}(2N)$ as the case requires.

The standard reference on the subject is M. L. Mehta, *Random Matrices*, 3rd ed. (Academic Press, 2004)[21]. One can then compute

$$\langle H_{ij} H_{mn} \rangle = \begin{cases} \frac{1}{2}\delta_{in}\delta_{jm} + \frac{1}{2}\delta_{im}\delta_{jn} & \text{(GOE)} \\ \frac{1}{2}\delta_{in}\delta_{jm} & \text{(GUE)} \\ \frac{1}{2}\delta_{in}\delta_{jm} - \frac{1}{4}\delta_{im}\delta_{jn} & \text{(GSE)} \end{cases} \quad . \tag{3.137}$$

For each ensemble, the representative matrices have real eigenvalues $\lambda_j$ whose joint probability distribution is given by (see Mehta)[22]

$$\mathcal{P}_\beta(\lambda_1, \ldots, \lambda_N) = \mathcal{C}_\beta \prod_{j<k}^{N} |\lambda_j - \lambda_k|^\beta \prod_{l=1}^{N} e^{-\beta\lambda_l^2/2} \quad . \tag{3.138}$$

Remarkably, the density of states $\sigma(\lambda)$ of the eigenvalues for each ensemble follows the Wigner semicircle law in the large-$N$ limit,

$$\sigma(\lambda) = \frac{1}{\pi}\sqrt{2N - \lambda^2}\,\Theta(2N - \lambda^2) \quad . \tag{3.139}$$

Note that $\int_{-\infty}^{\infty} d\lambda\,\sigma(\lambda) = N$.

**A few words about quaternionic numbers**

A *complex quaternionic number* $q$ is written as $q = a\,\mathbb{1} + b\,\mathbb{i} + c\,\mathbb{j} + d\,\mathbb{k}$, where $a, b, c, d \in \mathbb{C}$ and $\{\mathbb{i}, \mathbb{j}, \mathbb{k}\}$ are the quaternions: $\mathbb{i}^2 = \mathbb{j}^2 = \mathbb{k}^2 = -\mathbb{1}$ and $\mathbb{ij} = -\mathbb{ji} = \mathbb{k}$, $\mathbb{jk} = -\mathbb{kj} = \mathbb{i}$, and $\mathbb{ki} = -\mathbb{ik} = \mathbb{j}$. In a *real quaternionic number* we have $a, b, c, d \in \mathbb{R}$. The quaternionic algebra is satisfied if we take $\mathbb{1}$ as the $2 \times 2$ unit matrix and $(\mathbb{i}, \mathbb{j}, \mathbb{k}) = (iZ, iY, iX)$, where $\{X, Y, Z\}$ are the Pauli matrices. We may then express each quaternionic number $q$ as a $2 \times 2$ matrix, *viz.*

$$q = a\mathbb{1} + b\,\mathbb{i} + c\,\mathbb{j} + d\,\mathbb{k} = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix} \quad . \tag{3.140}$$

There are three conjugacy operations defined for quaternions: complex conjugation $q^*$, quaternion conjugation $\bar{q}$, and Hermitian conjugation $q^\dagger$:

$$\begin{aligned} q^* &= a^*\mathbb{1} + b^*\,\mathbb{i} + c^*\,\mathbb{j} + d^*\,\mathbb{k} = \begin{pmatrix} a^* + ib^* & c^* + id^* \\ -c^* + id^* & a^* - ib^* \end{pmatrix} \\ \bar{q} &= a\mathbb{1} - b\,\mathbb{i} - c\,\mathbb{j} - d\,\mathbb{k} = \begin{pmatrix} a - ib & -c - id \\ c - id & a + ib \end{pmatrix} \\ q^\dagger &= a^*\mathbb{1} - b^*\,\mathbb{i} - c^*\,\mathbb{j} - d^*\,\mathbb{k} = \begin{pmatrix} a^* - ib^* & -c^* - id^* \\ c^* - id^* & a^* + ib^* \end{pmatrix} \quad ; \end{aligned} \tag{3.141}$$

[21] I once saw a review which said that Mehta's book has gotten worse with each subsequent edition. In my opinion there is some truth to this! The first edition was less detailed than others (259 → 562 → 706 pp.), but what has been gained in completeness seems to have been lost in readability.

[22] For the GSE, each eigenvalue is doubly degenerate due to Kramers' theorem, a consequence of time-reversal invariance and half-odd-integer spin.

note that $q^\dagger = \bar{q}^*$. If $q^* = q$ then $q$ is said to be *real*; if $q = \bar{q}$ then $q$ is said to be a *scalar*. Note that if $q = q^\dagger$ then $q$ is an ordinary Hermitian $2 \times 2$ matrix. The set of real quaternionic numbers $q$ is denoted by the symbol $\mathbb{H}$. Unlike the reals $\mathbb{R}$ and complex numbers $\mathbb{C}$, $\mathbb{H}$ is not a field because multiplication within $\mathbb{H}$ is not commutative. Rather, we say that $\mathbb{H}$ is *an associative division algebra* over the real numbers. One may now check that for $q \in \mathbb{H}$,

$$|q|^2 \equiv q\bar{q} = \bar{q}q = (a^2 + b^2 + c^2 + d^2)\,\mathbb{1} = \begin{pmatrix} a^2 + b^2 + c^2 + d^2 & 0 \\ 0 & a^2 + b^2 + c^2 + d^2 \end{pmatrix} \quad . \tag{3.142}$$

Let a quaternionic matrix $Q$ have elements $q_{mn} = a_{mn}\,\mathbb{1} + b_{mn}\,\mathbb{i} + c_{mn}\,\mathbb{j} + d_{mn}\,\mathbb{k}$. If $Q$ is Hermitian and quaternion real, then $q_{nm} = \bar{q}_{mn}$. This entails that the real matrices $a, b, c, d$ satisfy $a = a^{\mathsf{T}}$, $b = -b^{\mathsf{T}}$, $c = -c^{\mathsf{T}}$, and $d = -d^{\mathsf{T}}$. Take care, however, when computing the quaternionic trace, since $\mathsf{Tr}\,\mathbb{1} \equiv 1$ even though the $\mathbb{1}$ on the LHS may be regarded as a $2 \times 2$ matrix.

**Eigenvector statistics**

In the limit $N \to \infty$, the eigenvectors of random matrices are independent random vectors in any basis. Let $\psi_j^\alpha$ be the $\alpha$ component of the $j^{\rm th}$ eigenvector of a given random matrix. The probability distribution for a given eigenvector is given by

$$P\big(\{\psi_j^\alpha, \bar{\psi}_j^\alpha\}\big) \propto \delta\left(\sum_{\alpha=1}^{N} |\psi_j^\alpha|^2 - 1\right) \quad , \tag{3.143}$$

where $\bar{\psi}_j^\alpha = \psi_j^\alpha$ for GOE, $\bar{\psi}_j^\alpha = \psi_j^{\alpha *}$ for GUE, and $\bar{\psi}_j^\alpha$ is the quaternion conjugate of $\psi_j^\alpha$ for GSE. For the full joint distribution, mutual orthogonality must also be respected:

$$P\big(\{\psi_j^\alpha, \bar{\psi}_j^\alpha\}\big) \propto \prod_{j=1}^{N} \delta\left(\sum_{\alpha=1}^{N} |\psi_j^\alpha|^2 - 1\right) \prod_{k<l}^{N} \delta\left(\sum_{\alpha=1}^{N} \bar{\psi}_k^\beta \psi_l^\beta\right) \quad . \tag{3.144}$$

It then follows that

$$\begin{aligned} &\text{GOE}: \quad \langle \psi_i^\alpha \psi_j^\beta \rangle = \frac{1}{N}\,\delta_{ij}\,\delta^{\alpha\beta} \\ &\text{GUE}: \quad \langle \psi_i^\alpha \bar{\psi}_j^\beta \rangle = \frac{1}{N}\,\delta_{ij}\,\delta^{\alpha\beta} \quad , \quad \langle \psi_i^\alpha \psi_j^\beta \rangle = \langle \bar{\psi}_i^\alpha \bar{\psi}_j^\beta \rangle = 0 \\ &\text{GSE}: \quad \langle \psi_i^\alpha \bar{\psi}_j^\beta \rangle = \frac{1}{N}\,\delta_{ij}\,\delta^{\alpha\beta} \quad , \quad \langle \psi_i^\alpha \psi_j^\beta \rangle = \langle \bar{\psi}_i^\alpha \bar{\psi}_j^\beta \rangle = -\frac{1}{2N}\,\delta_{ij}\,\delta^{\alpha\beta} \quad . \end{aligned} \tag{3.145}$$

Now consider an operator $\hat{A} = \sum_\alpha A_\alpha\,|\,\alpha\,\rangle\langle\,\alpha\,|$. Let us compute the matrix element of $\hat{A}$ in the random matrix eigenstates $|\,j\,\rangle$ and $|\,k\,\rangle$:

$$A_{jk} = \langle\, j\,|\,\hat{A}\,|\,k\,\rangle = \sum_\alpha A_\alpha\,\langle\, j\,|\,\alpha\,\rangle\langle\,\alpha\,|\,k\,\rangle = \sum_\alpha A_\alpha\,\bar{\psi}_j^\alpha\,\psi_k^\alpha \quad . \tag{3.146}$$

Then

$$\langle A_{jk} \rangle = \sum_\alpha A_\alpha\,\langle \bar{\psi}_j^\alpha\,\psi_k^\alpha \rangle = \frac{1}{N} \sum_\alpha A_\alpha\,\delta_{jk} = \frac{\mathsf{Tr}\,(\hat{A})}{N}\,\delta_{jk} \quad . \tag{3.147}$$

We may also compute

$$\langle A_{jk}\,\bar{A}_{mn}\rangle = \sum_{\alpha,\beta} A_\alpha\,A_\beta\,\langle\bar{\psi}_j^\alpha\,\psi_k^\alpha\,\psi_m^\beta\,\bar{\psi}_n^\beta\rangle \tag{3.148}$$

using Wick's theorem, *viz.*

$$\langle\bar{\psi}_j^\alpha\,\psi_k^\alpha\,\psi_m^\beta\,\bar{\psi}_n^\beta\rangle = \langle\bar{\psi}_j^\alpha\,\psi_k^\alpha\rangle\langle\psi_m^\beta\,\bar{\psi}_n^\beta\rangle + \langle\bar{\psi}_j^\alpha\,\psi_m^\beta\rangle\langle\psi_k^\alpha\,\bar{\psi}_n^\beta\rangle + \langle\bar{\psi}_j^\alpha\,\bar{\psi}_n^\beta\rangle\langle\psi_m^\beta\,\psi_k^\alpha\rangle \quad . \tag{3.149}$$

Thus we obtain

$$\begin{aligned}
\text{GOE}: \quad &\langle\bar{\psi}_j^\alpha\,\psi_k^\alpha\,\psi_m^\beta\,\bar{\psi}_n^\beta\rangle = \frac{1}{N^2}\left(\delta_{jk}\,\delta_{mn} + \delta_{jm}\,\delta_{kn}\,\delta^{\alpha\beta} + \delta_{jn}\,\delta_{km}\,\delta^{\alpha\beta}\right)\\
\text{GUE}: \quad &\langle\bar{\psi}_j^\alpha\,\psi_k^\alpha\,\psi_m^\beta\,\bar{\psi}_n^\beta\rangle = \frac{1}{N^2}\left(\delta_{jk}\,\delta_{mn} + \delta_{jm}\,\delta_{kn}\,\delta^{\alpha\beta}\right)\\
\text{GSE}: \quad &\langle\bar{\psi}_j^\alpha\,\psi_k^\alpha\,\psi_m^\beta\,\bar{\psi}_n^\beta\rangle = \frac{1}{N^2}\left(\delta_{jk}\,\delta_{mn} + \delta_{jm}\,\delta_{kn}\,\delta^{\alpha\beta} + \tfrac{1}{4}\,\delta_{jn}\,\delta_{km}\,\delta^{\alpha\beta}\right)
\end{aligned} \tag{3.150}$$

We use these results to obtain the following:

$$\langle|A_{mn}|^2\rangle - |\langle A_{mn}\rangle|^2 = \frac{\mathsf{Tr}\,(\hat{A}^2)}{N^2}\cdot\begin{cases}1+\delta_{mn} & \text{(GOE)}\\ 1 & \text{(GUE)}\\ 1+\frac{1}{4}\,\delta_{mn} & \text{(GSE)}\end{cases} . \tag{3.151}$$

Thus for off diagonal elements of $A$ the relative ratios GOE : GUE : GSE are $1:1:1$, while for diagonal elements the relative ratios are $2:1:\frac{5}{4}$.

### 3.6.6 More precise formulation of ETH

More precisely (Srednicki, 1999), ETH is formulated in terms of general matrix elements of local observables in the energy eigenbasis, *viz.*[23]

$$A_{mn} = \langle\,m\,|\,\hat{A}\,|\,n\,\rangle = A(E)\,\delta_{mn} + e^{-S(E)/2}\,f_A(E,\omega)\,R_{mn} \tag{3.152}$$

where $E = \frac{1}{2}(E_m + E_n)$ is an average of the energy eigenvalues, $\omega = E_m - E_n$ is their difference, $R_{mn}$ is a *random matrix* with $\langle R_{mn}\rangle = 0$ and $\mathsf{var}(R_{mn}) \equiv \langle|R_{mn}|^2\rangle = 1$ (no sum), $S(E)$ is the thermodynamic entropy. Toward the middle of the spectrum, $S(E) \sim \log\mathcal{D}$, where $\mathcal{D}$ is the Hilbert space dimension, *i.e.* the dimension of $A$. We further assume that $A(E)$ and $f(E,\omega)$ are smooth functions of their arguments. Additionally, one has

$$\begin{aligned}
A_{mn}\in\mathbb{R} \quad &: \quad R_{nm} = R_{mn} \quad , \quad f_A(E,-\omega) = f_A(E,\omega)\\
A_{mn}\in\mathbb{C} \quad &: \quad R_{nm} = \bar{R}_{mn} \quad , \quad f_A(E,-\omega) = f_A^*(E,\omega) \quad ,
\end{aligned} \tag{3.153}$$

for systems with and without time-reversal symmetry, respectively[24].

[23] See the review article by L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *Adv. Phys.*, **65**, 239 (2016) and references therein.

[24] For systems with time reversal symmetry, a real basis is possible for the $T^2 = +1$ class, where $T$ is the eigenvalue of the time reversal operator. The symplectic/Kramers class with $T^2 = -1$ requires a separate treatment.

**ETH *versus* random matrix theory**

ETH reduces to random matrix theory within a given small energy window. Eqn. 3.152 is to be contrasted with the corresponding result for random matrix theory,

$$A_{mn} = \langle A \rangle \, \delta_{mn} + \sqrt{\frac{\langle A^2 \rangle - \langle A \rangle^2}{\mathcal{D}}} \, R_{mn} \quad , \tag{3.154}$$

where $\langle A \rangle \equiv \mathcal{D}^{-1} \, \mathrm{Tr} \, \hat{A}$ and $\langle A^2 \rangle \equiv \mathcal{D}^{-1} \, \mathrm{Tr} \, (\hat{A}^2)$ so that the averages are over the entire spectrum. Again, $\langle |R_{mn}|^2 \rangle = 1$, with the necessary ensemble-specific adjustments for the diagonal elements as indicated in eqn. 3.151. ETH thus reduces to RMT within any small window which contains $\mathcal{O}(\mathcal{D})$ states.

**A consequence of ETH**

From ETH it follows that the expectation value of an operator in a Gibbs state $\varrho = Z^{-1} \exp(-\beta H)$ can be replaced by its expectation value in any eigenstate whose energy eigenvalue corresponds to the peak in $D(E) \exp(-\beta E)$, where $D(E)$ is the (many-body) density of states. Such an eigenstate is then called a *representative pure state*[25]. To see this, note

$$\mathrm{Tr} \, (\varrho \hat{A}) \approx \frac{1}{Z} \int dE \, D(E) \, e^{-\beta E} A(E) \approx A(E^*) \quad . \tag{3.155}$$

Here it is assumed that the function $A(E)$ in Eqn. 3.152 is a smooth function of $E$, and that the distribution $D(E) \exp(-\beta E)$ has a narrow peak centered at $E = E^*$. For example, if

$$D(E) \sim \exp\left( - \frac{(E - E_{\rm c})^2}{2\sigma^2} \right) \quad , \tag{3.156}$$

where the energy variance is extensive, *i.e.* $\sigma^2 \approx wN$ with $w$ a constant and $N$ the total number of particles or the system volume in microscopic units, then $E^*(\beta) \approx E_{\rm c} - \beta\sigma^2$ and the energy density is

$$\varepsilon^*(\beta) \equiv \frac{E^*}{N} \approx \varepsilon_{\rm c} - \beta w \quad , \tag{3.157}$$

where $\varepsilon_{\rm c}$ is the energy density at the center of the spectrum. Eigenstates of $H$ in the vicinity of energy density $\varepsilon^*(\beta)$ are thus representative pure states of the Gibbs density matrix which reproduce expectation values of few body operators. Note that we may obtain $\varepsilon^*(\beta)$ from the expression

$$\varepsilon^*(\beta) = \frac{1}{NZ} \, \mathrm{Tr} \, (\hat{H} e^{-\beta \hat{H}}) \quad . \tag{3.158}$$

### 3.6.7 When is the ETH true?

There is no rigorous general proof of the ETH. Deutsch showed that the ETH holds for the case of an integrable Hamiltonian weakly perturbed by a single Gaussian random matrix. Horoi *et al.* (1995) showed

[25] See V. Khemani, A. Chandran, H. Kim, and S. L. Sondhi, *Phys. Rev. E* **90**, 052133 (2014).

that nuclear shell model wavefunctions reproduce thermodynamic predictions. Recent numerical work by M. Rigol and collaborators has verified the applicability of the ETH in small interacting boson systems. ETH fails for so-called integrable models, where there are a large number of conserved quantities, which commute with the Hamiltonian. Integrable models are, however, quite special, and as Deutsch showed, integrability is spoiled by weak perturbations, in which case ETH then applies.

ETH also fails in the case of noninteracting disordered systems which exhibit *Anderson localization*. Single particle energy eigenstates $\psi_j$ whose energies $\varepsilon_j$ lie within the localized portion of the eigenspectrum decay exponentially, as $|\psi_j(\boldsymbol{r})|^2 \sim \exp\bigl(-|\boldsymbol{r}-\boldsymbol{r}_j|/\xi(\varepsilon_j)\bigr)$, where $\boldsymbol{r}_j$ is some position in space associated with $\psi_j$ and $\xi(\varepsilon_j)$ is the *localization length*. Within the localized portion of the spectrum, $\xi(\varepsilon)$ is finite. As $\varepsilon$ approaches a *mobility edge*, $\xi(\varepsilon)$ diverges as a power law. In the *delocalized regime*, eigenstates are spatially *extended* and typically decay at worst as a power law[26]. Exponentially localized states are unable to thermalize with other distantly removed localized states. Of course, all noninteracting systems will violate ETH, because they are integrable. The interacting version of this phenomenon, *many-body localization* (MBL), is a topic of current interest and some controversy in condensed matter and statistical physics. MBL systems also exhibit a large number of conserved quantities, but in contrast to the case of integrable systems, where each conserved quantity is in general expressed in terms of an integral of a local density, in MBL systems the conserved quantities are themselves local, although *emergent*. The emergent nature of locally conserved quantities in MBL systems means that they are not simply expressed in terms of the original local operators of the system, but rather are arrived at via a sequence of local unitary transformations.

Note again that in contrast to the classical case, time evolution of a quantum state does not create the thermal state. Rather, it *reveals* the thermal distribution which is encoded in all eigenstates after sufficient time for dephasing to occur, so that correlations between all the wavefunction expansion coefficients $\{C_\alpha\}$ in $|\Psi(t)\rangle = \sum_\alpha C_\alpha\, e^{-iE_\alpha t/\hbar}\,|\Psi_\alpha\rangle$ for $\alpha \neq \alpha'$ are all lost[27].

## 3.7 Appendix I: Matrices and Jordan Canonical Form

### 3.7.1 Normal and non-normal matrices

If a matrix $A$ is normal, meaning $\bigl[A, A^\dagger\bigr] = 0$, then it may be diagonalized by a unitary transformation. Indeed, any $n \times n$ matrix $A$ is diagonalizable by a unitary transformation if and only if $A$ is normal. When $A$ is normal, the eigenspaces span, and we may choose $\langle\, L^\alpha \,|\, R^\beta \,\rangle = \delta^{\alpha\beta}$, using Gram-Schmidt in the case of degeneracies.

When $A$ is not normal, while the sum of the dimensions of its eigenspaces generically is equal to its dimension $\dim(A) = n$, this is not guaranteed, and it may be less than $n$. What is true is that any non-normal complex matrix $A$ can be brought to Jordan canonical form by a similarity transformation

[26] Recall that in systems with no disorder, eigenstates exhibit Bloch periodicity in space.

[27] To be more precise, "strong ETH" proposes that each bulk eigenstate encodes thermal information, but "weak ETH" allows for exceptional states, such as quantum scars.

$\tilde{A} = Q^{-1} A\, Q$, where $Q$ is invertible and

$$\tilde{A} = \begin{pmatrix} J_1 & 0 & \\ 0 & \ddots & 0 \\ & 0 & J_b \end{pmatrix} \tag{3.159}$$

is block diagonal with $b$ Jordan blocks. Each Jordan block $J_\alpha$ with $\alpha \in \{1, \ldots, b\}$ is of the form

$$J_\alpha = \begin{pmatrix} \lambda_\alpha & 1 & 0 & \\ 0 & \lambda_\alpha & \ddots & 0 \\ & & \ddots & 1 \\ & & 0 & \lambda_\alpha \end{pmatrix} \quad . \tag{3.160}$$

Thus each $J_\alpha$ is tridiagonal, with diagonal elements all given by $\lambda_\alpha$ and each element directly above the diagonal equal to one. We denote the right and left eigenvectors of $\tilde{A}$ as $|\, \widetilde{R}^\alpha \,\rangle = Q^{-1} \,|\, R^\alpha \,\rangle$ and $\langle\, \widetilde{L}^\alpha \,| = \langle\, L^\alpha \,|\, Q$, respectively. Each $J_\alpha$ has only one right eigenvector, $\psi^\alpha_j = \delta_{j,1}$, whose corresponding left eigenvector is $\chi^\alpha_j = \delta_{j,n_\alpha}$, where $n_\alpha = \mathsf{dim}(J_\alpha)$. Note $n = \mathsf{dim}(A)$ is the sum of the dimensions of the Jordan blocks, *i.e.* $n = \sum_{\alpha=1}^b n_\alpha$. When $n_\alpha = 1$, the Jordan block is the $1 \times 1$ matrix $\lambda_\alpha$. For a non-normal matrix $A$, its eigenvalues $\lambda_\alpha$ may be complex. However, if all the elements of $A$ are real, then any complex eigenvalues must occur in complex conjugate pairs, because the characteristic polynomial $P(\lambda) = \mathsf{det}(\lambda - A)$ satisfies $\big[P(\lambda)\big]^* = P(\lambda^*)$.

When $\lambda_\alpha \neq \lambda_\beta$, we have $\langle\, L^\alpha \,|\, R^\beta \,\rangle = \langle\, \widetilde{L}^\alpha \,|\, \widetilde{R}^\beta \,\rangle = 0$. For eigenspaces with $n_\alpha = 1$, we may choose $\langle\, L^\alpha \,|\, R^\beta \,\rangle = \langle\, \widetilde{L}^\alpha \,|\, \widetilde{R}^\beta \,\rangle = \delta^{\alpha\beta}$, but for the nontrivial Jordan blocks with $n_\alpha > 1$ we have $\langle\, \widetilde{L}^\alpha \,|\, \widetilde{R}^\alpha \,\rangle = 0$, as we have seen in the previous paragraph, and therefore $\langle\, L^\alpha \,|\, R^\alpha \,\rangle = 0$. Real symmetric matrices are all normal, with no Jordan blocks. For complex symmetric matrices, we may have nontrivial Jordan blocks.

Since $\langle\, L \,|\, R \,\rangle = \langle\, \widetilde{L} \,|\, \widetilde{R} \,\rangle = 0$, we may write

$$A = \bigoplus_{n_\alpha = 1} \lambda_\alpha \,|\, R^\alpha \,\rangle\langle\, L^\alpha \,| \bigoplus_{n_\beta > 1} Q J_\beta\, Q^{-1} \tag{3.161}$$

and raising $A$ to the $k$ power yields

$$A^k = \bigoplus_{n_\alpha = 1} \lambda_\alpha^k \,|\, R^\alpha \,\rangle\langle\, L^\alpha \,| \bigoplus_{n_\beta > 1} Q J_\beta^k\, Q^{-1} \quad . \tag{3.162}$$

Note that $J^k_\beta$ is upper triangular with all diagonal elements given by $\lambda^k_\beta$. Note that for complex symmetric matrices, the left and right eigenvectors are identical and we may write

$$L^\alpha_j = R^\alpha_j \equiv \Psi^\alpha_j \qquad \Rightarrow \qquad \langle\, L^\alpha \,| = \langle\, \Psi^\alpha \,| \quad , \quad |\, R^\alpha \,\rangle = |\, \Psi^\alpha \,\rangle \quad , \tag{3.163}$$

with no complex conjugation, *i.e.* $\langle\, L^\alpha \,|\, j \,\rangle = \langle\, j \,|\, R^\alpha \,\rangle \equiv \Psi^\alpha_j$.

### 3.7.2 Contrast with singular value decomposition

We now remark upon the difference between the decomposition into Jordan canonical form and the singular value decomposition (SVD), in which we write an $m \times n$ matrix $A$ as $A = U D V^\dagger$, where $U$ is $m \times k$, $V$ is $n \times k$ (hence $V^\dagger$ is $k \times n$), $U^\dagger U = V^\dagger V = \mathbb{I}_{k\times k}$, and $D = \mathsf{diag}(d_1, \ldots, d_k)$ is $k \times k$ with $k \le \mathsf{min}(m, n)$ and each $d_j > 0$. The elements $d_j$ are the singular values and the columns of $U$ and $V$ are the singular vectors. Note that $A^\dagger A = V D^2 V^\dagger$ is $n \times n$ and $AA^\dagger = U D^2 U^\dagger$ is $m \times m$. If we define

$$\widetilde{P}(\lambda) = \prod_{j=1}^{k} \left(\lambda - d_j^2\right) \quad , \tag{3.164}$$

Then

$$P(\lambda) \equiv \mathsf{det}(\lambda - A^\dagger A) = \lambda^{n-k} \widetilde{P}(\lambda) \qquad , \qquad Q(\lambda) \equiv \mathsf{det}(\lambda - AA^\dagger) = \lambda^{m-k} \widetilde{P}(\lambda) \quad . \tag{3.165}$$

For any square $n \times n$ complex matrix $A$ we therefore have two decompositions, via JCF and SVD, *viz.*

$$A = Q J Q^{-1} = U D V^\dagger \quad , \tag{3.166}$$

where $J$ is the Jordan canonical form of $A$. When $A$ is normal, its Jordan canonical form is diagonal. Its singular values are $d_j = |\lambda_j|$, and its left and right singular vectors can be chosen to differ only by the phases of the corresponding eigenvalues. In a reduced SVD, $k = \mathsf{rank}(A)$, which need not equal $n$. Only when $A$ is positive semidefinite Hermitian may one generally take $U = V = Q$.

### 3.7.3 Examples

Consider the real asymmetric matrix

$$A = \begin{pmatrix} 2a & -4a^2 \\ \frac{1}{4} & 0 \end{pmatrix} \quad , \tag{3.167}$$

where $a \in \mathbb{R}$ is any real number. The characteristic polynomial is $P(\lambda) = \mathsf{det}(\lambda - A) = (\lambda - a)^2$ and there is a single eigenvalue, $\lambda = a$. The right and left eigenvectors are found to be

$$|\, R \,\rangle = \begin{pmatrix} 4a \\ 1 \end{pmatrix} \qquad , \qquad \langle\, L \,| = \begin{pmatrix} \frac{1}{4} & -a \end{pmatrix} \quad , \tag{3.168}$$

where the normalization is arbitrary. Note $\langle\, L \,|\, R \,\rangle = 0$. The matrix $A$ is brought to JCF by the similarity transformation $\tilde{A} = Q^{-1} A Q$ with

$$Q = \begin{pmatrix} 4a & 4 \\ 1 & 0 \end{pmatrix} \qquad , \qquad Q^{-1} = \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & -a \end{pmatrix} \qquad , \qquad \tilde{A} = Q^{-1} A Q = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \quad . \tag{3.169}$$

Note that

$$|\, \widetilde{R} \,\rangle = Q^{-1} \,|\, R \,\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad , \qquad \langle\, \widetilde{L} \,| = \langle\, L \,| \, Q = \begin{pmatrix} 0 & 1 \end{pmatrix} \quad , \tag{3.170}$$

and that $\langle\, \widetilde{L} \,|\, \widetilde{R} \,\rangle = 0$.

Adding another row and column to our matrix $A$, consider the matrix

$$B = \begin{pmatrix} 2a & -4a^2 & 0 \\ \frac{1}{4} & 0 & 0 \\ 0 & 0 & b \end{pmatrix} \quad , \tag{3.171}$$

where both $a$ and $b$ are arbitrary real numbers. Since $\det(\lambda - B) = (\lambda - a)^2 (\lambda - b)$, $B$ thus has two eigenvalues: $\lambda_1 = a$ and $\lambda_2 = b$. The decomposition of $B$ is then

$$B = Q J_1 Q^{-1} \bigoplus \lambda_2 \, | \, R^2 \, \rangle \langle \, L^2 \, | \quad , \tag{3.172}$$

where $\langle \, L^2 \, | = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ and $| \, R^2 \, \rangle = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^{\mathsf{T}}$. Of course, we could mix up the various elements of $B$ by applying a general similarity transformation $B \to B' \equiv SBS^{-1}$, but the JCF of $B'$ would be the same.

### 3.7.4 Exponentiating Jordan blocks

For a diagonal matrix $M = \mathsf{diag}\{\lambda_1, \ldots, \lambda_n\}$, its exponential is $\exp(Mt) = \mathsf{diag}\{e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\}$. Consider instead the $n \times n$ Jordan block

$$J = \begin{pmatrix} \lambda & 1 & 0 & \\ 0 & \lambda & \ddots & 0 \\ & & \ddots & 1 \\ & & 0 & \lambda \end{pmatrix} \quad . \tag{3.173}$$

We can write

$$J = \lambda I + N \tag{3.174}$$

where $I_{jk} = \delta_{jk}$ is the $n \times n$ identity matrix and $N_{jk} = \delta_{k,j+1}$ has ones along its first superdiagonal (*i.e.* just above its main diagonal) and zeros elsewhere. The matrix $N$ is *nilpotent*, with $N^n = 0$, and thus the exponential series for $\exp(Nt)$ terminates after $n$ terms, *viz.*

$$\exp(Nt) = I + tN + \frac{t^2}{2!} N^2 + \ldots + \frac{t^{n-1}}{(n-1)!} N^{n-1} \quad . \tag{3.175}$$

Since $\lambda I$ commutes with $N$, we have

$$\exp(Jt) = e^{\lambda t} K \tag{3.176}$$

where

$$K = \exp(Nt) = \begin{pmatrix} 1 & t & t^2/2! & t^3/3! & \cdots & t^{n-1}/(n-1)! \\ 0 & 1 & t & t^2/2! & \cdots & t^{n-2}/(n-2)! \\ 0 & 0 & 1 & t & \cdots & t^{n-3}/(n-3)! \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & t \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad . \tag{3.177}$$

In component notation,

$$K_{jk} = \begin{cases} t^{k-j}/(k-j)! & \text{if } k \geq j \\ 0 & \text{if } k < j \end{cases} \quad . \tag{3.178}$$

## 3.8 Appendix II: Poisson Processes and Radioactive Decay

Here we consider two examples where the state labels of the master equation denote a number, corresponding to the discrete population of some group. The master equation is

$$\frac{dP_n}{dt} = \sum_m \big( W_{nm}\, P_m - W_{mn}\, P_n \big) \quad . \tag{3.179}$$

Thus $W_{nm}$ is the transition rate for the process $|\, m \,\rangle \to |\, n \,\rangle$.

### 3.8.1 Poisson process

Consider the *Poisson process*, for which

$$W_{nm} = \begin{cases} \lambda & \text{if } n = m+1 \\ 0 & \text{if } n \neq m+1 \; . \end{cases} \tag{3.180}$$

We then have

$$\frac{dP_n}{dt} = \lambda \big( P_{n-1} - P_n \big) \; . \tag{3.181}$$

The *generating function* $P(z,t) = \sum_{n=0}^{\infty} z^n P_n(t)$ then satisfies

$$\frac{\partial P}{\partial t} = \lambda (z-1)\, P \quad \Rightarrow \quad P(z,t) = e^{(z-1)\lambda t}\, P(z,0) \; . \tag{3.182}$$

If the initial distribution is $P_n(0) = \delta_{n,0}$, then

$$P_n(t) = \frac{(\lambda t)^n}{n!}\, e^{-\lambda t} \; , \tag{3.183}$$

which is known as the *Poisson distribution*. If we define $\alpha \equiv \lambda t$, then from $P_n = \alpha^n\, e^{-\alpha}/n!$ we have

$$\langle n^k \rangle = e^{-\alpha} \left( \alpha \frac{\partial}{\partial \alpha} \right)^k e^{\alpha} \; . \tag{3.184}$$

Thus, $\langle n \rangle = \alpha$, $\langle n^2 \rangle = \alpha^2 + \alpha$, *etc.*

### 3.8.2 Radioactive decay

Consider a group of atoms, some of which are in an excited state which can undergo nuclear decay. Let $P_n(t)$ be the probability that $n$ atoms are excited at some time $t$. We then model the decay dynamics by

$$W_{nm} = \begin{cases} 0 & \text{if } n \geq m \\ m\gamma & \text{if } n = m-1 \\ 0 & \text{if } n < m-1 \quad . \end{cases} \tag{3.185}$$

Here, $\gamma$ is the decay rate of an individual atom, which can be determined from quantum mechanics. The master equation then tells us

$$\frac{dP_n}{dt} = (n+1)\,\gamma\,P_{n+1} - n\,\gamma\,P_n \quad . \tag{3.186}$$

The interpretation here is as follows: let $|\,n\,\rangle$ denote a state in which $n$ atoms are excited. We then have $P_n(t) = \langle\, n\,|\,P(t)\,\rangle$. Then $P_n(t)$ will increase due to spontaneous transitions from $|\,n{+}1\,\rangle$ to $|\,n\,\rangle$, and will decrease due to spontaneous transitions from $|\,n\,\rangle$ to $|\,n{-}1\,\rangle$.

The average number of particles in the system is $N(t) = \sum_{n=0}^\infty n\,P_n(t)$. Note that

$$\begin{aligned}\frac{dN}{dt} &= \sum_{n=0}^\infty n\Big[(n+1)\,\gamma\,P_{n+1} - n\,\gamma\,P_n\Big] \\ &= \gamma\sum_{n=0}^\infty\Big[n(n-1)\,P_n - n^2 P_n\Big] = -\gamma\sum_{n=0}^\infty n\,P_n = -\gamma\,N \quad .\end{aligned} \tag{3.187}$$

Thus, $N(t) = N(0)\,e^{-\gamma t}$. The relaxation time is $\tau = \gamma^{-1}$, and the equilibrium distribution is $P_n^{\rm eq} = \delta_{n,0}$. Note that this satisfies detailed balance.

We can go a bit farther here. Let us define

$$P(z,t) \equiv \sum_{n=0}^\infty z^n\,P_n(t) \quad . \tag{3.188}$$

This is sometimes called a *generating function*. Then

$$\begin{aligned}\frac{\partial P}{\partial t} &= \gamma\sum_{n=0}^\infty z^n\Big[(n+1)\,P_{n+1} - n\,P_n\Big] \\ &= \gamma\,\frac{\partial P}{\partial z} - \gamma z\,\frac{\partial P}{\partial z} \quad .\end{aligned} \tag{3.189}$$

Thus,

$$\frac{1}{\gamma}\,\frac{\partial P}{\partial t} - (1-z)\,\frac{\partial P}{\partial z} = 0 \quad . \tag{3.190}$$

We now see that any function $f(\xi)$ satisfies the above equation, where $\xi = \gamma t - \log(1-z)$. Thus, we can write

$$P(z,t) = f\big(\gamma t - \log(1-z)\big) \quad . \tag{3.191}$$

Setting $t=0$ we have $P(z,0) = f\big(-\log(1-z)\big)$, and inverting this result we obtain $f(u) = P(1-e^{-u},0)$, which entails

$$P(z,t) = P\big(1 + (z-1)\,e^{-\gamma t}\,,\,0\big) \quad . \tag{3.192}$$

The total probability is $P(z{=}1,t) = \sum_{n=0}^\infty P_n$, which clearly is conserved: $P(1,t) = P(1,0)$. The average particle number is

$$N(t) = \sum_{n=0}^\infty n\,P_n(t) = \left.\frac{\partial P}{\partial z}\right|_{z=1} = \left.e^{-\gamma t}\,\frac{\partial P(z,0)}{\partial z}\right|_{z=1} = N(0)\,e^{-\gamma t} \quad . \tag{3.193}$$

## 3.9 Appendix III: The Markovian Mpemba Effect

### 3.9.1 Introduction

Intuition and common experience tell us that if we place two objects of identical composition at temperatures $T_{\rm h}$ and $T_{\rm c}$ into the same bath at temperature $T_{\rm b}$, with $T_{\rm h} > T_{\rm c} > T_{\rm b} > 0$, it will take longer for the hotter object to thermalize to the bath temperature than for the cold object. If you want to make ice, you fill your ice tray with cold water before placing it in the freezer and not with hot water. Yet ca. 350-340 BCE, Aristotle wrote,[28]

*"Hail occurs less often in the summer than in spring and autumn, though more often than in winter, because the air is drier in summer; in spring it is still moist, while in autumn it is already becoming moist. Sometimes, as has been said, hail also occurs in late summer, for the same reason. The fact that the water has previously been heated also contributes to the rapidity of its freezing, for it cools more quickly. Hence many people, when they wish to cool water quickly, first place it in the sun. And the people around the Pontus, when they make camp upon the ice for catching fish – for they fish by breaking through the ice – pour hot water around their reeds so that it will freeze more quickly. "*

This may be the first mention of what is known today as the *Mpemba effect*, where a given system takes a shorter time to cool down from a hotter initial temperature than from a colder one. Newton's phenomenological law of cooling says

$$\frac{dT}{dt} = -k(T - T_{\rm env}) \quad , \tag{3.194}$$

where $T(t)$ is the temperature of an object in contact with a heat bath at temperature $T_{\rm env}$ . The constant $k > 0$ thus determines the cooling rate. We then have

$$T(t) = T_{\rm env} + \big[T(0) - T_{\rm env}\big] e^{-kt} \quad , \tag{3.195}$$

which entails that the Mpemba effect is impossible. The existence of such a phenomenon remains controversial[29].

Here we discuss a demonstrable mathematical phenomenon known as the *Markovian Mpemba effect*[30] (MME) in which an initial thermal probability distribution $p_j(t=0) = \pi_j(T)$ at temperature $T$ is evolved according to the master equation. The equilibrium distribution eventually reached is also thermal, corresponding to a bath temperature $T_{\rm b} < T$. Given an appropriate definition of distance between two distributions, $D[\boldsymbol{p}, \boldsymbol{p}']$, the distance between two thermal distributions $T_{1,2}$ is an increasing function of their temperature difference $|T_1 - T_2|$. Now let $\boldsymbol{p}(t;T)$ be the solution of the master equation subject to the initial conditions $\boldsymbol{p}(t=0,T) = \boldsymbol{\pi}(T)$, the Boltzmann distribution at temperature $T$. Then for some choices of the rates $W_{ij}$, it may be that the plots of $D[\boldsymbol{p}(t,T), \boldsymbol{\pi}(T_{\rm b})]$ *versus* $t$ for $T = T_{\rm h}$ and $T = T_{\rm c}$ cross at some time $t$, and that for asymptotically late times the distribution evolving from the hotter initial state is closer to $\boldsymbol{\pi}(T_{\rm b})$.

[28] Aristotle, *Meteorologica* 1.12, 348b26-349a1.

[29] On the yes side, see, *e.g.*, M. Jeng, *Am. J. Phys.* **74**, 514 (2006). On the no side, H. C. Burridge and P. F. Linden. *Sci Rep.* **6**, 37665 (2016).

[30] Z. Lu and O. Raz, *Proc. Natl. Acad. Sci.* **114**, 5083 (2017).

The Master equation was introduced in §3.2:

$$\frac{dP_i}{dt} = \sum_j \left( W_{ij}\, P_j - W_{ji}\, P_i \right) = -\Gamma_{ij} P_j \quad , \tag{3.196}$$

with

$$W_{ij} = \gamma \exp\left( \frac{E_j - B_{ij}}{k_{\rm B} T_{\rm b}} \right) \quad , \tag{3.197}$$

where $T_{\rm b}$ is the bath temperature and $B_{ij} = B_{ji}$ is the energy barrier that must be overcome in transitioning from $|j\rangle$ to $|i\rangle$. If we take

$$B_{ij} = \begin{cases} E_j & \text{if } i < j \\ E_i & \text{if } i \geq j \end{cases} \quad , \tag{3.198}$$

then $W_{ij}$ is given in eqn. 3.2.3. Recall $\sum_i W_{ij} = 0$ and thus $\langle\, L^1 \,| = (1\ ,\ 1\ ,\ \ldots\ ,\ 1)$ is a left eigenvector of $\Gamma$ with eigenvalue $\lambda_1 = 0$. The equilibrium distribution at temperature $T$ is given by

$$\pi_i(T) = \frac{1}{Z(T)}\, e^{-E_i/k_{\rm B}T} \quad , \tag{3.199}$$

where $Z(T)$ is the canonical partition function, which guarantees $\sum_i \pi_i = 1$. We assume detailed balance, which entails $W_{ij}\,\pi_j(T_{\rm b}) = W_{ji}\,\pi_i(T_{\rm b})$ and is satisfied by eqns. 3.197 and 3.199 for arbitrary symmetric $B_{ij}$.

### 3.9.2 Distance between distributions

Lu and Raz list three desiderata which any distance function $D[\boldsymbol{p}(t), \boldsymbol{\pi}(T_{\rm b})]$ should satisfy:

(i) Under the Master equation dynamics, $D[\boldsymbol{p}(t), \boldsymbol{\pi}(T_{\rm b})]$ should be monotonically decreasing in $t$.

(ii) For all $T_{\rm h} > R_{\rm c} > R_{\rm b}$, we should have $D[\boldsymbol{\pi}(T_{\rm h}), \boldsymbol{\pi}(T_{\rm b})] > D[\boldsymbol{\pi}(T_{\rm c}), \boldsymbol{\pi}(T_{\rm b})] > 0$.

(iii) $D[\boldsymbol{p}, \boldsymbol{\pi}(T_{\rm b})]$ should be a continuous convex function of the components $\{\pi_i(T_{\rm b})\}$.

One possibility is the Kullback-Leibler divergence,

$$D_{\rm KL}[\boldsymbol{p}, \boldsymbol{\pi}^{\rm b}] = \sum_i p_i \log\left( \frac{p_i}{\pi_i^{\rm b}} \right) \quad , \tag{3.200}$$

where $\pi_i^{\rm b} = \pi_i(T_{\rm b})$ is the equilibrium distribution. This is also known as the "relative entropy" of the distributions. Extremizing with respect to $\boldsymbol{p}$ at fixed and normalized $\boldsymbol{\pi}^{\rm b}$ yields $p_i = \pi_i^{\rm b}$. Another possibility is the $L^1$ distance,

$$D_{L^1}[\boldsymbol{p}, \boldsymbol{\pi}^{\rm b}] = \sum_i |p_i - \pi_i^{\rm b}| \quad , \tag{3.201}$$

where in this case we assume $\boldsymbol{p}'$ is normalized. Again the minimum occurs for $\boldsymbol{p} = \boldsymbol{p}'$, where $D = 0$.

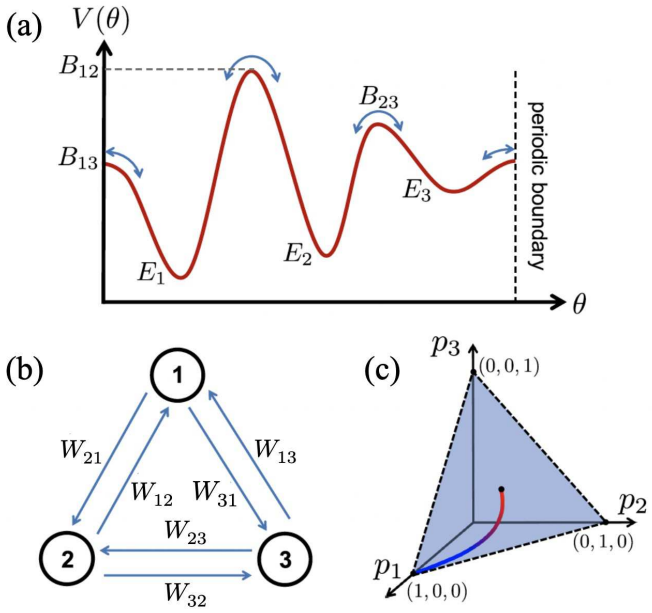



Figure 3.12: Three-state MME from fig. 1 of Lu and Raz (2017). (a) Sketch of potential $V(\theta)$ *versus* coordinate $\theta$ in a periodic geometry. (b) Master equation rates. (c) The probability distribution is specified by $(p_1, p_2, p_3)$ where $p_1 + p_2 + p_3 = 1$. The curved line is the set of Boltzmann distributions $\boldsymbol{\pi}(T)$, with $\boldsymbol{\pi}(T = 0) = (1, 0, 0)$ and $\boldsymbol{\pi}(T = \infty) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

### 3.9.3 Conditions for the Markovian Mpemba effect

The general solution to the Master equation may be written

$$p_i(t) = \pi_i^{\rm b} + \sum_{a=2}^{N} C_a \, R_i^a \, e^{-\lambda_a t} \quad , \tag{3.202}$$

where $N$ is the dimension of the state space and $C_a \equiv \langle\, L^a \,|\, p(0) \,\rangle = \sum_j L_j^a p_j(0)$. Again we assume there are no Jordan blocks in the spectral decomposition of $\Gamma$. If we order the eigenmodes such that

$$0 = \lambda_1 < \mathsf{Re}\,\lambda_2 < \mathsf{Re}\,\lambda_3 < \cdots < \mathsf{Re}\,\lambda_N \quad , \tag{3.203}$$

then, with $\boldsymbol{p}(0) = \boldsymbol{\pi}(T)$ a Boltzmann distribution at a temperature $T > T_{\rm b}$, we find that at long times the solution is

$$|\, p(t) \,\rangle = |\, \pi^{\rm b} \,\rangle + C_2(T)\, e^{-\lambda_2 t} \,|\, R^2 \,\rangle \quad . \tag{3.204}$$

The MME holds in whenever $|C_2(T)|$ is a decreasing function of $T$. If $T < T_{\rm b}$, then there is an *inverse Markovian Mpemba effect* whenever $|C_2(T)|$ is an increasing function of $T$.

### 3.9.4 Numerical simulation

Lu and Raz (2017) investigated numerically a three-state model with energies $E_1 = 0$, $E_2 = 0.1$, and $E_3 = 0.7$ and barriers $B_{12} = 1.5$, $B_{13} = 0.8$, and $B_{23} = 1.2$. A sketch is shown in fig. 3.12. Evolution

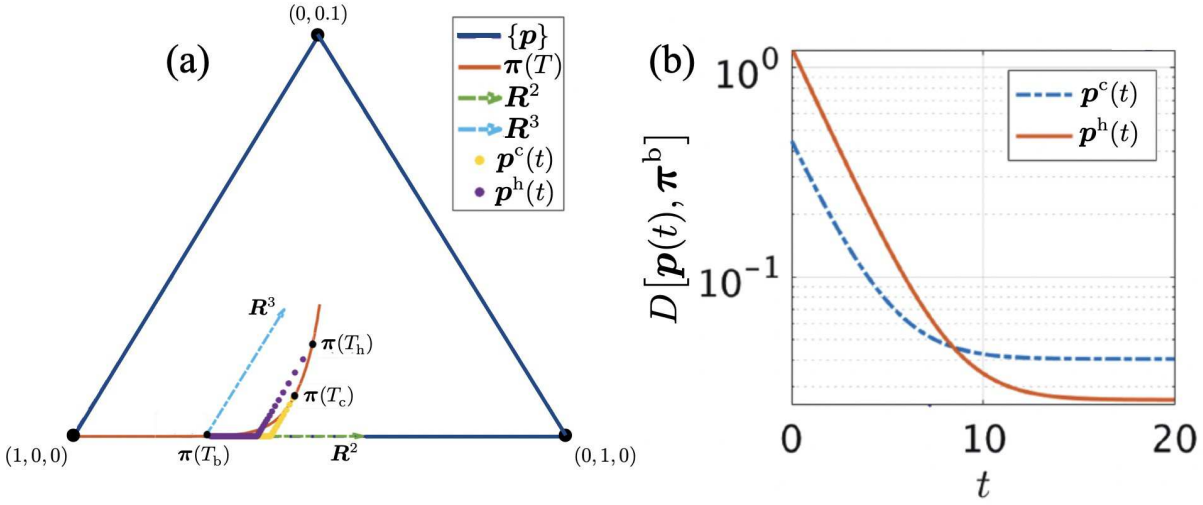



Figure 3.13: Fig. 2 from Lu and Raz (2017). (a) The triangular simplex corresponds to the set of all normalized three-state probability distributions $\boldsymbol{p} = (p_1, p_2, p_3)$. The green and blue dashed lines show the right eigenvectors $\boldsymbol{R}^{2,3}$ of $\Gamma$. The dotted lines show the relaxations starting from thermal distributions at $T_{\rm h} = 1.3$ (purple) and $T_{\rm c} = 0.42$ to the bath at $T_{\rm b} = 0.1$. (b) The distances $D_{\rm KL}[\boldsymbol{p}^{\rm h,c}(t), \boldsymbol{\pi}^{\rm b}]$ *versus time*, showing a MME. At late times, the distribution $\boldsymbol{p}^{\rm h}(t)$ is closer to the bath distribution $\boldsymbol{\pi}^{\rm b}$ than $\boldsymbol{p}^{\rm c}(t)$.

of the distribution $\boldsymbol{p}(t)$ starting from hot and cold initial thermal distributions $\boldsymbol{p}(0) = \boldsymbol{\pi}(T^{\rm h,c})$ and their relaxation to the bath temperature $T^{\rm b}$ is shown in fig. 3.13.

## 3.10 Appendix IV: Transition to Ergodicity in a Simple Model

A ball of mass $m$ executes perfect one-dimensional motion along the symmetry axis of a piston. Above the ball lies a mobile piston head of mass $M$ which slides frictionlessly inside the piston. Both the ball and piston head execute ballistic motion, with two types of collision possible: (i) the ball may bounce off the floor, which is assumed to be infinitely massive and fixed in space, and (ii) the ball and piston head may engage in a one-dimensional elastic collision. The Hamiltonian is

$$H = \frac{P^2}{2M} + \frac{p^2}{2m} + MgX + mgx \quad ,$$

where $X$ is the height of the piston head and $x$ the height of the ball. Another quantity is conserved by the dynamics: $\Theta(X - x)$. *I.e.*, the ball always is below the piston head.

(a) Choose an arbitrary length scale $L$, and then energy scale $E_0 = MgL$, momentum scale $P_0 = M\sqrt{gL}$, and time scale $\tau_0 = \sqrt{L/g}$. Show that the dimensionless Hamiltonian becomes

$$\bar{H} = \tfrac{1}{2}\bar{P}^2 + \bar{X} + \frac{\bar{p}^2}{2r} + r\bar{x} \quad ,$$

with $r = m/M$, and with equations of motion $d\bar{X}/dt = \partial\bar{H}/\partial\bar{P}$, *etc.* (Here the bar indicates dimensionless variables: $\bar{P} = P/P_0$, $\bar{t} = t/\tau_0$, *etc.*) What special dynamical consequences hold for $r = 1$?

(b) Compute the microcanonical average piston height $\langle X \rangle$. The analogous dynamical average is

$$\langle X \rangle_t = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt\, X(t) \quad .$$

When computing microcanonical averages, it is helpful to use the Laplace transform, discussed in §4.2.2 of the notes. (It is possible to compute the microcanonical average by more brute force methods as well.)

(c) Compute the microcanonical average of the rate of collisions between the ball and the floor. Show that this is given by

$$\Big\langle \sum_i \delta(t - t_i) \Big\rangle = \big\langle \Theta(v)\, v\, \delta(x - 0^+) \big\rangle \quad .$$

The analogous dynamical average is

$$\langle \gamma \rangle_t = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt \sum_i \delta(t - t_i) \quad ,$$

where $\{t_i\}$ is the set of times at which the ball hits the floor.

(d) How do your results change if you do not enforce the dynamical constraint $X \geq x$?

(e) Write a computer program to simulate this system. The only input should be the mass ratio $r$ (set $\bar{E} = 10$ to fix the energy). You also may wish to input the initial conditions, or perhaps to choose the initial conditions randomly (all satisfying energy conservation, of course!). Have your program compute the microcanonical as well as dynamical averages in parts (b) and (c). Plot out the Poincaré section of $P$ *vs.* $X$ for those times when the ball hits the floor. Investigate this for several values of $r$. Just to show you that this is interesting, I've plotted some of my own numerical results in fig. 3.14.

Solution:

(a) Once we choose a length scale $L$ (arbitrary), we may define $E_0 = M\,gL$, $P_0 = M\sqrt{gL}$, $V_0 = \sqrt{gL}$, and $\tau_0 = \sqrt{L/g}$ as energy, momentum, velocity, and time scales, respectively, the result follows directly. Rather than write $\bar{P} = P/P_0$ *etc.*, we will drop the bar notation and write

$$H = \tfrac{1}{2}P^2 + X + \frac{p^2}{2r} + rx \quad .$$

(b) What is missing from the Hamiltonian of course is the interaction potential between the ball and the piston head. We assume that both objects are impenetrable, so the potential energy is infinite when the two overlap. We further assume that the ball is a point particle (otherwise reset ground level to minus the diameter of the ball). We can eliminate the interaction potential from $H$ if we enforce that each time $X = x$ the ball and the piston head undergo an elastic collision. From energy and momentum

conservation, it is easy to derive the elastic collision formulae

$$P' = \frac{1-r}{1+r}\,P + \frac{2}{1+r}\,p$$

$$p' = \frac{2r}{1+r}\,P - \frac{1-r}{1+r}\,p \quad .$$

We can now answer the last question from part (a). When $r = 1$, we have that $P' = p$ and $p' = P$, *i.e.* the ball and piston simply exchange momenta. The problem is then equivalent to two identical particles elastically bouncing off the bottom of the piston, and moving through each other as if they were completely transparent. When the trajectories cross, however, the particles exchange identities.

Averages within the microcanonical ensemble are normally performed with respect to the phase space distribution

$$\varrho(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{\mathsf{Tr}\;\delta\big(E - H(\boldsymbol{\varphi})\big)} \quad ,$$

where $\boldsymbol{\varphi} = (P, X, p, x)$, and

$$\mathsf{Tr}\;F(\boldsymbol{\varphi}) = \int\limits_{-\infty}^{\infty} dP \int\limits_{0}^{\infty} dX \int\limits_{-\infty}^{\infty} dp \int\limits_{0}^{\infty} dx\; F(P, X, p, x) \quad .$$

Since $X \geq x$ is a dynamical constraint, we should define an appropriately restricted microcanonical average:

$$\big\langle F(\boldsymbol{\varphi})\big\rangle_{\mu\text{ce}} \equiv \widetilde{\mathsf{Tr}}\,\Big[F(\boldsymbol{\varphi})\,\delta\big(E - H(\boldsymbol{\varphi})\big)\Big] \Big/ \widetilde{\mathsf{Tr}}\,\delta\big(E - H(\boldsymbol{\varphi})\big)$$

where

$$\widetilde{\mathsf{Tr}}\,F(\boldsymbol{\varphi}) \equiv \int\limits_{-\infty}^{\infty} dP \int\limits_{0}^{\infty} dX \int\limits_{-\infty}^{\infty} dp \int\limits_{0}^{X} dx\; F(P, X, p, x)$$

is the modified trace. Note that the integral over $x$ has an upper limit of $X$ rather than $\infty$, since the region of phase space with $x > X$ is *dynamically inaccessible*.

When computing the traces, we shall make use of the following result from the theory of Laplace transforms. The Laplace transform of a function $K(E)$ is

$$\widehat{K}(\beta) = \int\limits_{0}^{\infty} dE\; K(E)\, e^{-\beta E} \quad .$$

The inverse Laplace transform is given by

$$K(E) = \int\limits_{c-i\infty}^{c+i\infty} \frac{d\beta}{2\pi i}\;\widehat{K}(\beta)\, e^{\beta E} \quad ,$$

where the integration contour, which is a line extending from $\beta = c - i\infty$ to $\beta = c + i\infty$, lies to the right of any singularities of $\widehat{K}(\beta)$ in the complex $\beta$-plane. For this problem, all we shall need is the following:

$$K(E) = \frac{E^{t-1}}{\Gamma(t)} \quad \Longleftrightarrow \quad \widehat{K}(\beta) = \beta^{-t} \quad .$$

For a proof, see §4.2.2 of the lecture notes.

We're now ready to compute the microcanonical average of $X$. We have

$$\langle X \rangle = \frac{N(E)}{D(E)} \quad ,$$

where $N(E) = \widetilde{\mathsf{Tr}}\left[X\,\delta(E-H)\right]$ and $D(E) = \widetilde{\mathsf{Tr}}\,\delta(E-H)$. Let's first compute $D(E)$. To do this, we compute the Laplace transform $\widehat{D}(\beta)$:

$$\begin{aligned}
\widehat{D}(\beta) &= \widetilde{\mathsf{Tr}}\, e^{-\beta H} \\
&= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp\, e^{-\beta p^2/2r} \int\limits_{0}^{\infty} dX\, e^{-\beta X} \int\limits_{0}^{X} dx\, e^{-\beta r x} \\
&= \frac{2\pi\sqrt{r}}{\beta} \int\limits_{0}^{\infty} dX\, e^{-\beta X} \left(\frac{1 - e^{-\beta r X}}{\beta r}\right) = \frac{\sqrt{r}}{1+r} \cdot \frac{2\pi}{\beta^3} \quad .
\end{aligned}$$

Similarly for $\widehat{N}(\beta)$ we have

$$\begin{aligned}
\widehat{N}(\beta) &= \widetilde{\mathsf{Tr}}\, X\, e^{-\beta H} \\
&= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp\, e^{-\beta p^2/2r} \int\limits_{0}^{\infty} dX\, X\, e^{-\beta X} \int\limits_{0}^{X} dx\, e^{-\beta r x} \\
&= \frac{2\pi\sqrt{r}}{\beta} \int\limits_{0}^{\infty} dX\, X\, e^{-\beta X} \left(\frac{1 - e^{-\beta r X}}{\beta r}\right) = \frac{(2+r)\, r^{3/2}}{(1+r)^2} \cdot \frac{2\pi}{\beta^4} \quad .
\end{aligned}$$

Taking the inverse Laplace transform, we then have

$$D(E) = \frac{\sqrt{r}}{1+r} \cdot \pi E^2 \qquad , \qquad N(E) = \frac{(2+r)\,\sqrt{r}}{(1+r)^2} \cdot \tfrac{1}{3}\pi E^3 \quad .$$

We then have

$$\langle X \rangle = \frac{N(E)}{D(E)} = \left(\frac{2+r}{1+r}\right) \cdot \tfrac{1}{3} E \quad .$$

The 'brute force' evaluation of the integrals isn't so bad either. We have

$$D(E) = \int\limits_{-\infty}^{\infty} dP \int\limits_{0}^{\infty} dX \int\limits_{-\infty}^{\infty} dp \int\limits_{0}^{X} dx\, \delta\big(\tfrac{1}{2}P^2 + \tfrac{1}{2r}p^2 + X + rx - E\big) \quad .$$

To evaluate, define $P = \sqrt{2}\,u_x$ and $p = \sqrt{2r}\,u_y$. Then we have $dP\,dp = 2\sqrt{r}\,du_x\,du_y$ and $\frac{1}{2}P^2 + \frac{1}{2r}\,p^2 = u_x^2 + u_y^2$. Now convert to 2D polar coordinates with $w \equiv u_x^2 + u_y^2$. Thus,

$$\begin{aligned} D(E) &= 2\pi\sqrt{r}\int_0^\infty dw \int_0^\infty dX \int_0^X dx\;\delta\big(w + X + rx - E\big) \\ &= \frac{2\pi}{\sqrt{r}}\int_0^\infty dw \int_0^\infty dX \int_0^X dx\;\;\Theta(E - w - X)\,\Theta(X + rX - E + w) \\ &= \frac{2\pi}{\sqrt{r}}\int_0^E dw \int_{\frac{E-w}{1+r}}^{E-w} dX = \frac{2\pi\sqrt{r}}{1+r}\int_0^E dq\;q = \frac{\sqrt{r}}{1+r}\cdot\pi E^2 \quad , \end{aligned}$$

with $q = E - w$. Similarly,

$$\begin{aligned} N(E) &= 2\pi\sqrt{r}\int_0^\infty dw \int_0^\infty dX\;X \int_0^X dx\;\delta\big(w + X + rx - E\big) \\ &= \frac{2\pi}{\sqrt{r}}\int_0^\infty dw \int_0^\infty dX\;X \int_0^X dx\;\;\Theta(E - w - X)\,\Theta(X + rX - E + w) \\ &= \frac{2\pi}{\sqrt{r}}\int_0^E dw \int_{\frac{E-w}{1+r}}^{E-w} dX\;X = \frac{2\pi}{\sqrt{r}}\int_0^E dq\left(1 - \frac{1}{(1+r)^2}\right)\cdot\tfrac{1}{2}q^2 = \left(\frac{2+r}{1+r}\right)\cdot\frac{\sqrt{r}}{1+r}\cdot\tfrac{1}{3}\pi E^3 \quad . \end{aligned}$$

(c) Using the general result

$$\delta\big(F(x) - A\big) = \sum_i \frac{\delta(x - x_i)}{\big|F'(x_i)\big|} \quad ,$$

where $F(x_i) = A$, we recover the desired expression. We should be careful not to double count, so to avoid this difficulty we can evaluate $\delta(t - t_i^+)$, where $t_i^+ = t_i + 0^+$ is infinitesimally later than $t_i$. The point here is that when $t = t_i^+$ we have $p = r\,v > 0$ (*i.e.* just after hitting the bottom). Similarly, at times $t = t_i^-$ we have $p < 0$ (*i.e.* just prior to hitting the bottom). Note $v = p/r$. Again we write $\gamma(E) = N(E)/D(E)$, this time with

$$N(E) = \widetilde{\mathsf{Tr}}\left[\Theta(p)\,r^{-1}p\,\delta(x - 0^+)\,\delta(E - H)\right] \quad .$$

The Laplace transform is

$$\begin{aligned} \widehat{N}(\beta) &= \int_{-\infty}^\infty dP\,e^{-\beta P^2/2}\int_0^\infty dp\;r^{-1}\,p\,e^{-\beta p^2/2r}\int_0^\infty dX\,e^{-\beta X} \\ &= \sqrt{\frac{2\pi}{\beta}}\cdot\frac{1}{\beta}\cdot\frac{1}{\beta} = \sqrt{2\pi}\,\beta^{-5/2} \quad . \end{aligned}$$

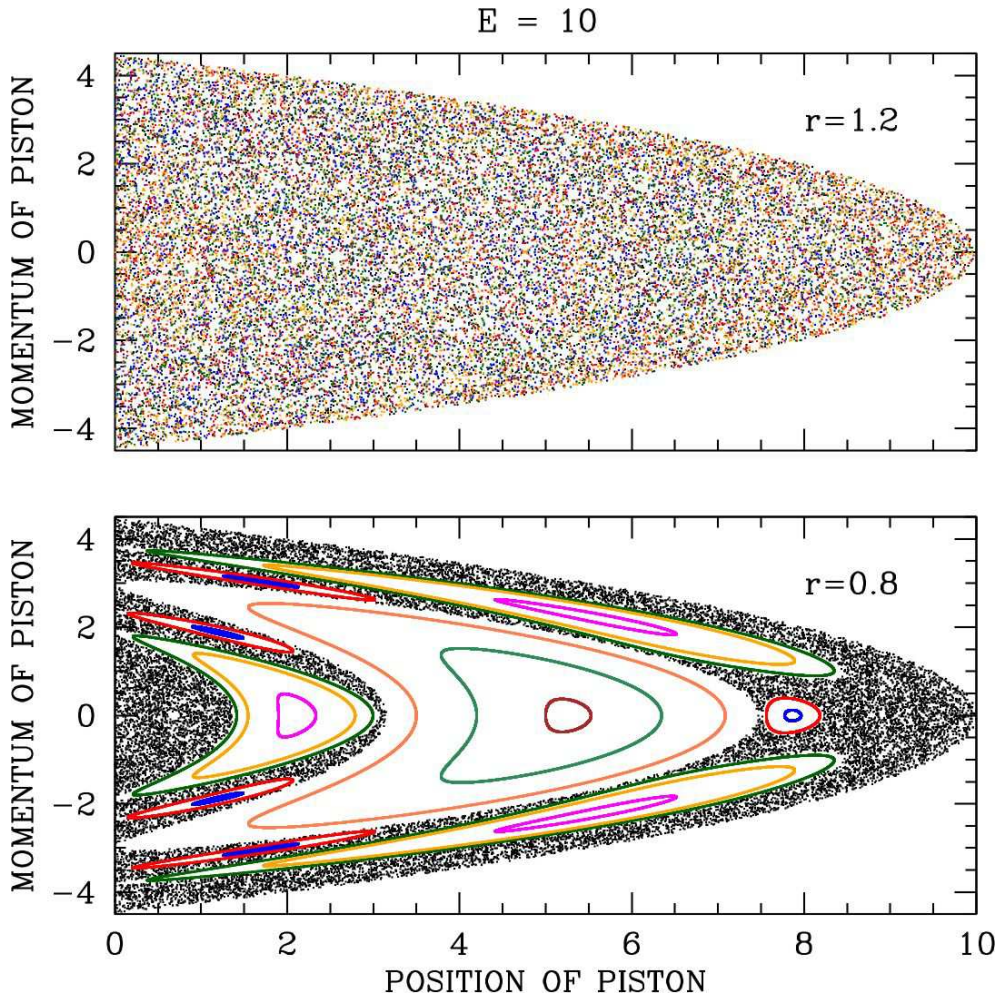



Figure 3.14: Poincaré sections for the ball and piston head problem. Each color corresponds to a different initial condition. When the mass ratio $r = m/M$ exceeds unity, the system apparently becomes ergodic.

Thus, $N(E) = \frac{4\sqrt{2}}{3}\,E^{3/2}$ and

$$\langle\gamma\rangle = \frac{N(E)}{D(E)} = \frac{4\sqrt{2}}{3\pi}\left(\frac{1+r}{\sqrt{r}}\right)E^{-1/2} \quad .$$

(d) When the constraint $X \geq x$ is removed, we integrate over all phase space. We then have

$$\begin{aligned}\widehat{D}(\beta) &= \mathsf{Tr}\; e^{-\beta H} \\ &= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp\, e^{-\beta p^2/2r} \int\limits_{0}^{\infty} dX\, e^{-\beta X} \int\limits_{0}^{\infty} dx\, e^{-\beta r x} = \frac{2\pi\sqrt{r}}{\beta^3} \quad .\end{aligned}$$

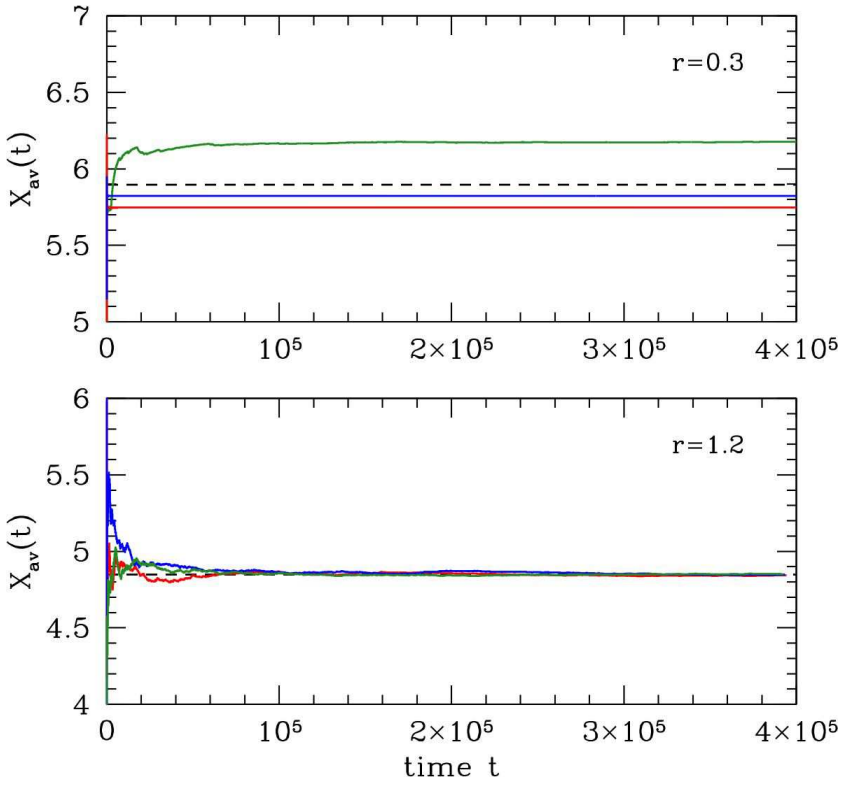



Figure 3.15: Long time running numerical averages $X_{\rm av}(t) \equiv t^{-1}\int_0^t dt' X(t')$ for $r = 0.3$ (top) and $r = 1.2$ (bottom), each for three different initial conditions, with $E = 10$ in all cases. Note how in the $r = 0.3$ case the long time average is dependent on the initial condition, while the $r = 1.2$ case is ergodic and hence independent of initial conditions. The dashed black line shows the restricted microcanonical average, $\langle X\rangle_{\mu{\rm ce}} = \frac{(2+r)}{(1+r)}\cdot\frac{1}{3}E$.

| $r$ | $X(0)$ | $\langle X(t)\rangle$ | $\langle X\rangle_{\mu{\rm ce}}$ | $\langle\gamma(t)\rangle$ | $\langle\gamma\rangle_{\mu{\rm ce}}$ | $r$ | $X(0)$ | $\langle X(t)\rangle$ | $\langle X\rangle_{\mu{\rm ce}}$ | $\langle\gamma(t)\rangle$ | $\langle\gamma\rangle_{\mu{\rm ce}}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.3 | 0.1 | 6.1743 | 5.8974 | 0.5283 | 0.4505 | 1.2 | 0.1 | 4.8509 | 4.8545 | 0.3816 | 0.3812 |
| 0.3 | 1.0 | 5.7303 | 5.8974 | 0.4170 | 0.4505 | 1.2 | 1.0 | 4.8479 | 4.8545 | 0.3811 | 0.3812 |
| 0.3 | 3.0 | 5.7876 | 5.8974 | 0.4217 | 0.4505 | 1.2 | 3.0 | 4.8493 | 4.8545 | 0.3813 | 0.3812 |
| 0.3 | 5.0 | 5.8231 | 5.8974 | 0.4228 | 0.4505 | 1.2 | 5.0 | 4.8482 | 4.8545 | 0.3813 | 0.3812 |
| 0.3 | 7.0 | 5.8227 | 5.8974 | 0.4228 | 0.4505 | 1.2 | 7.0 | 4.8472 | 4.8545 | 0.3808 | 0.3812 |
| 0.3 | 9.0 | 5.8016 | 5.8974 | 0.4234 | 0.4505 | 1.2 | 9.0 | 4.8466 | 4.8545 | 0.3808 | 0.3812 |
| 0.3 | 9.9 | 6.1539 | 5.8974 | 0.5249 | 0.4505 | 1.2 | 9.9 | 4.8444 | 4.8545 | 0.3807 | 0.3812 |

Table 3.2: Comparison of time averages and microcanonical ensemble averages for $r = 0.3$ and $r = 1.2$. Initial conditions are $P(0) = x(0) = 0$, with $X(0)$ given in the table and $E = 10$. Averages were performed over a period extending for $N_{\rm b} = 10^7$ bounces.

For part (b) we would then have

$$\begin{aligned}\widehat{N}(\beta) &= \mathsf{Tr}\ X\, e^{-\beta H} \\ &= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp\, e^{-\beta p^2/2r} \int\limits_{0}^{\infty} dX\, X\, e^{-\beta X} \int\limits_{0}^{\infty} dx\, e^{-\beta r x} = \frac{2\pi\sqrt{r}}{\beta^4} \quad .\end{aligned}$$

The respective inverse Laplace transforms are $D(E) = \pi\sqrt{r}\,E^2$ and $N(E) = \frac{1}{3}\pi\sqrt{r}\,E^3$. The microcanonical average of $X$ would then be $\langle X\rangle = \frac{1}{3}E$. Using the restricted phase space, we obtained a value which is greater than this by a factor of $(2+r)/(1+r)$. That the restricted average gives a larger value makes good sense, since $X$ is not allowed to descend below $x$ in that case. For part (c), we would obtain the same result for $N(E)$ since $x = 0$ in the average. We would then obtain

$$\langle\gamma\rangle = \tfrac{4\sqrt{2}}{3\pi}\, r^{-1/2}\, E^{-1/2} \quad .$$

The restricted microcanonical average yields a rate which is larger by a factor $1 + r$. Again, it makes good sense that the restricted average should yield a higher rate, since the ball is not allowed to attain a height greater than the instantaneous value of $X$.

(e) It is straightforward to simulate the dynamics. So long as $0 < x(t) < X(t)$, we have

$$\dot{X} = P \quad , \quad \dot{P} = -1 \quad , \quad \dot{x} = \frac{p}{r} \quad , \quad \dot{p} = -r \quad .$$

Starting at an arbitrary time $t_0$, these equations are integrated to yield

$$\begin{aligned} X(t) &= X(t_0) + P(t_0)\,(t-t_0) - \tfrac{1}{2}(t-t_0)^2 \\ P(t) &= P(t_0) - (t-t_0) \\ x(t) &= x(t_0) + \frac{p(t_0)}{r}\,(t-t_0) - \tfrac{1}{2}(t-t_0)^2 \\ p(t) &= p(t_0) - r(t-t_0) \quad . \end{aligned}$$

We must stop the evolution when one of two things happens. The first possibility is a bounce at $t = t_{\rm b}$, meaning $x(t_{\rm b}) = 0$. The momentum $p(t)$ changes discontinuously at the bounce, with $p(t_{\rm b}^+) = -p(t_{\rm b}^-)$, and where $p(t_{\rm b}^-) < 0$ necessarily. The second possibility is a collision at $t = t_{\rm c}$, meaning $X(t_{\rm c}) = x(t_{\rm c})$. Integrating across the collision, we must conserve both energy and momentum. This means

$$P(t_{\rm c}^+) = \frac{1-r}{1+r}\,P(t_{\rm c}^-) + \frac{2}{1+r}\,p(t_{\rm c}^-)$$

$$p(t_{\rm c}^+) = \frac{2r}{1+r}\,P(t_{\rm c}^-) - \frac{1-r}{1+r}\,p(t_{\rm c}^-) \quad .$$

In the following tables I report on the results of numerical simulations, comparing dynamical averages with (restricted) phase space averages within the microcanonical ensemble. For $r = 0.3$ the microcanonical averages poorly approximate the dynamical averages, and the dynamical averages are dependent on the initial conditions, indicating that the system is not ergodic. For $r = 1.2$, the agreement between

| $r$ | $X(0)$ | $N_{\rm b}$ | $\langle X(t)\rangle$ | $\langle X\rangle_{\mu{\rm ce}}$ | $\langle\gamma(t)\rangle$ | $\langle\gamma\rangle_{\mu{\rm ce}}$ |
|---|---|---|---|---|---|---|
| 1.2 | 7.0 | $10^4$ | 4.8054892 | 4.8484848 | 0.37560388 | 0.38118510 |
| 1.2 | 7.0 | $10^5$ | 4.8436969 | 4.8484848 | 0.38120356 | 0.38118510 |
| 1.2 | 7.0 | $10^6$ | 4.8479414 | 4.8484848 | 0.38122778 | 0.38118510 |
| 1.2 | 7.0 | $10^7$ | 4.8471686 | 4.8484848 | 0.38083749 | 0.38118510 |
| 1.2 | 7.0 | $10^8$ | 4.8485825 | 4.8484848 | 0.38116282 | 0.38118510 |
| 1.2 | 7.0 | $10^9$ | 4.8486682 | 4.8484848 | 0.38120259 | 0.38118510 |
| 1.2 | 1.0 | $10^9$ | 4.8485381 | 4.8484848 | 0.38118069 | 0.38118510 |
| 1.2 | 9.9 | $10^9$ | 4.8484886 | 4.8484848 | 0.38116295 | 0.38118510 |

Table 3.3: Comparison of time averages and microcanonical ensemble averages for $r = 1.2$, with $N_{\rm b}$ ranging from $10^4$ to $10^9$.

dynamical and microcanonical averages generally improves with averaging time. Indeed, it has been shown by N. I. Chernov, *Physica D* **53**, 233 (1991), building on the work of M. P. Wojtkowski, *Comm. Math. Phys.* **126**, 507 (1990) that this system is ergodic for $r > 1$. Wojtkowski also showed that this system is equivalent to the *wedge billiard*, in which a single point particle of mass $m$ bounces inside a two-dimensional wedge-shaped region $\big\{(\mathsf{x},\mathsf{y})\,\big|\,\mathsf{x}\geq 0\,,\ \mathsf{y}\geq \mathsf{x}\,\mathrm{ctn}\,\phi\big\}$ for some fixed angle $\phi = \tan^{-1}\sqrt{\frac{m}{M}}$. To see this, pass to relative ($\mathcal{X}$) and center-of-mass ($\mathcal{Y}$) coordinates,

$$\mathcal{X} = X - x \qquad\qquad \mathcal{P}_x = \frac{mP - Mp}{M+m}$$
$$\mathcal{Y} = \frac{MX + mx}{M+m} \qquad\qquad \mathcal{P}_y = P + p \quad .$$

Then

$$H = \frac{(M+m)\,\mathcal{P}_x^2}{2Mm} + \frac{\mathcal{P}_y^2}{2(M+m)} + (M+m)\,g\mathcal{Y} \quad .$$

There are two constraints. One requires $X \geq x$, *i.e.* $\mathcal{X} \geq 0$. The second requires $x > 0$, *i.e.*

$$x = \mathcal{Y} - \frac{M}{M+m}\,\mathcal{X} \geq 0 \quad .$$

Now define $\mathsf{x} \equiv \mathcal{X}$, $\mathsf{p}_x \equiv \mathcal{P}_x$, and rescale $\mathsf{y} \equiv \frac{M+m}{\sqrt{Mm}}\,\mathcal{Y}$ and $\mathsf{p}_y \equiv \frac{\sqrt{Mm}}{M+m}\,\mathcal{P}_y$ to obtain

$$H = \frac{1}{2\mu}\big(\mathsf{p}_x^2 + \mathsf{p}_y^2\big) + \mathsf{M}g\mathsf{y}$$

with $\mu = \frac{Mm}{M+m}$ the familiar reduced mass and $\mathsf{M} = \sqrt{Mm}$. The constraints are then $\mathsf{x} \geq 0$ and $\mathsf{y} \geq \sqrt{\frac{M}{m}}\,\mathsf{x}$.

# Chapter 4

# Statistical Ensembles

## 4.1 References

- F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, 1987)
  This has been perhaps the most popular undergraduate text since it first appeared in 1967, and with good reason.
- A. H. Carter, *Classical and Statistical Thermodynamics*
  (Benjamin Cummings, 2000)
  A very relaxed treatment appropriate for undergraduate physics majors.
- D. V. Schroeder, *An Introduction to Thermal Physics* (Addison-Wesley, 2000)
  This is the best undergraduate thermodynamics book I've come across, but only 40% of the book treats statistical mechanics.
- C. Kittel, *Elementary Statistical Physics* (Dover, 2004)
  Remarkably crisp, though dated, this text is organized as a series of brief discussions of key concepts and examples. Published by Dover, so you can't beat the price.
- M. Kardar, *Statistical Physics of Particles* (Cambridge, 2007)
  A superb modern text, with many insightful presentations of key concepts.
- M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* ($3^{\rm rd}$ edition, World Scientific, 2006)
  An excellent graduate level text. Less insightful than Kardar but still a good modern treatment of the subject. Good discussion of mean field theory.
- E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics* (part I, $3^{\rm rd}$ edition, Pergamon, 1980)
  This is volume 5 in the famous Landau and Lifshitz *Course of Theoretical Physics*. Though dated, it still contains a wealth of information and physical insight.

## 4.2 Microcanonical Ensemble ($\mu$CE)

### 4.2.1 The microcanonical distribution function

We have seen how in an *ergodic* dynamical system, time averages can be replaced by phase space averages[1]:

$$\text{ergodicity} \quad \Longleftrightarrow \quad \big\langle f(\boldsymbol{\varphi})\big\rangle_t = \big\langle f(\boldsymbol{\varphi})\big\rangle_E \quad , \tag{4.1}$$

where

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_t = \lim_{T\to\infty} \frac{1}{T}\int\limits_0^T dt\, f\big(\boldsymbol{\varphi}(t)\big) \tag{4.2}$$

and

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_E = \int d\mu\, f(\boldsymbol{\varphi})\,\delta\big(E-\hat{H}(\boldsymbol{\varphi})\big)\Big/\int d\mu\,\delta\big(E-\hat{H}(\boldsymbol{\varphi})\big) \tag{4.3}$$

Here $\hat{H}(\boldsymbol{\varphi}) = \hat{H}(\boldsymbol{q},\boldsymbol{p})$ is the Hamiltonian, and $\delta(x)$ is the Dirac $\delta$-function[2]. The dynamics are presumed to conserve energy, with $\hat{H}(\boldsymbol{\varphi}(t=0)) = E$. Thus, averages are taken over a constant energy hypersurface which is a subset of the entire phase space.

We've also seen how any phase space distribution $\varrho(\Lambda_1,\ldots,\Lambda_k)$ which is a function of conserved quantities $\Lambda_a(\boldsymbol{\varphi})$ is automatically a stationary (time-independent) solution to Liouville's equation. Note that the microcanonical distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \delta\big(E-\hat{H}(\boldsymbol{\varphi})\big)\Big/\int d\mu\,\delta\big(E-\hat{H}(\boldsymbol{\varphi})\big) \quad , \tag{4.4}$$

is of this form, since $\hat{H}(\boldsymbol{\varphi})$ is conserved by the dynamics. Linear and angular momentum conservation generally are broken by elastic scattering off the walls of the sample.

So averages in the microcanonical ensemble are computed by evaluating the ratio

$$\big\langle A\big\rangle = \frac{\mathsf{Tr}\; A\,\delta(E-\hat{H})}{\mathsf{Tr}\;\delta(E-\hat{H})} \quad , \tag{4.5}$$

where $\mathsf{Tr}$ means 'trace', which entails an integration over all phase space:

$$\mathsf{Tr}\; A(q,p) \equiv \frac{1}{N!}\prod_{i=1}^{N}\int \frac{d^dp_i\, d^dq_i}{(2\pi\hbar)^d}\, A(q,p) \quad . \tag{4.6}$$

Here $N$ is the total number of particles and $d$ is the dimension of physical space in which each particle moves. The factor of $1/N!$, which cancels in the ratio between numerator and denominator, is present for *indistinguishable particles*[3]. The normalization factor $(2\pi\hbar)^{-Nd}$ renders the trace dimensionless. Again,

[1] Technically this is the case only for suitable observables and almost every initial condition on the invariant energy surface.

[2] We write the Hamiltonian as $\hat{H}$ (classical or quantum) in order to distinguish it from magnetic field, $H$.

[3] More on this in chapter 5.

this cancels between numerator and denominator. These factors may then seem arbitrary in the definition of the trace, but we'll see how they in fact are required from quantum mechanical considerations. So we now adopt the following metric for classical phase space integration:

$$d\mu = \frac{1}{N!}\prod_{i=1}^{N}\frac{d^dp_i\,d^dq_i}{(2\pi\hbar)^d} \quad . \tag{4.7}$$

### 4.2.2 Density of states

The denominator,

$$D(E) = \mathsf{Tr}\;\delta(E-\hat{H}) \quad , \tag{4.8}$$

is called the *density of states*. It has dimensions of inverse energy, such that

$$\begin{aligned} D(E)\,\Delta E &= \int_E^{E+\Delta E} dE' \int d\mu\,\delta(E'-\hat{H}) = \int\limits_{E<\hat{H}<E+\Delta E} d\mu \\ &= \#\text{ of states with energies between } E \text{ and } E+\Delta E \quad . \end{aligned} \tag{4.9}$$

Let us now compute $D(E)$ for the nonrelativistic ideal gas. The Hamiltonian is

$$\hat{H}(q,p) = \sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m} \quad . \tag{4.10}$$

We assume that the gas is enclosed in a region of volume $V$, and we'll do a purely classical calculation, neglecting discreteness of its quantum spectrum. We must compute

$$D(E) = \frac{1}{N!}\int\prod_{i=1}^{N}\frac{d^dp_i\,d^dq_i}{(2\pi\hbar)^d}\,\delta\Big(E-\sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m}\Big) \quad . \tag{4.11}$$

We shall calculate $D(E)$ in two ways. The first method utilizes the *Laplace transform*, $Z(\beta)$:

$$Z(\beta) = \mathcal{L}\big[D(E)\big] \equiv \int_0^\infty dE\,e^{-\beta E}\,D(E) = \mathsf{Tr}\;e^{-\beta\hat{H}} \quad . \tag{4.12}$$

The inverse Laplace transform is then

$$D(E) = \mathcal{L}^{-1}\big[Z(\beta)\big] \equiv \int_{c-i\infty}^{c+i\infty}\frac{d\beta}{2\pi i}\,e^{\beta E}\,Z(\beta) \quad , \tag{4.13}$$

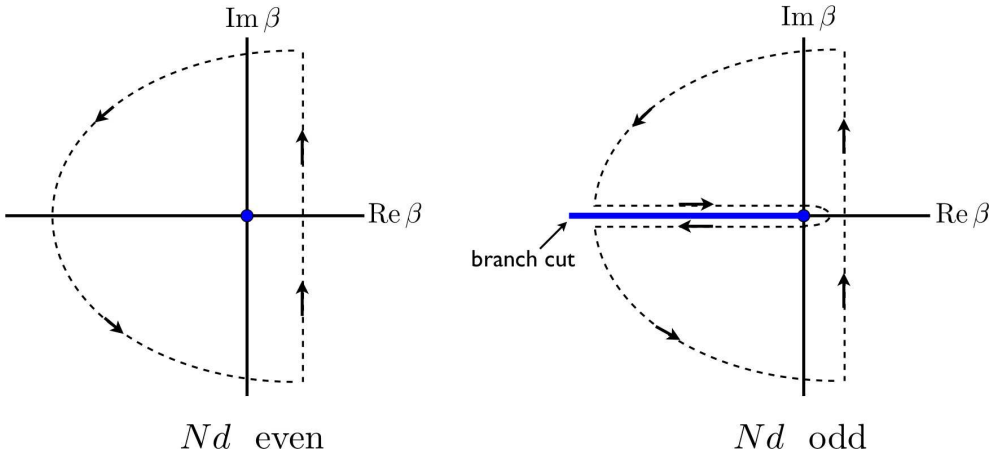



Figure 4.1: Complex integration contours $\mathcal{C}$ for inverse Laplace transform $\mathcal{L}^{-1}\big[Z(\beta)\big] = D(E)$. When the product $dN$ is odd, there is a branch cut along the negative $\mathrm{Re}\,\beta$ axis.

where $c$ is such that the integration contour is to the right of any singularities of $Z(\beta)$ in the complex $\beta$-plane. We then have

$$\begin{aligned} Z(\beta) &= \frac{1}{N!}\prod_{i=1}^{N}\int\frac{d^dx_i\,d^dp_i}{(2\pi\hbar)^d}\,e^{-\beta\boldsymbol{p}_i^2/2m} \\ &= \frac{V^N}{N!}\left(\int_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}\,e^{-\beta p^2/2m}\right)^{Nd} = \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\beta^{-Nd/2} \quad . \end{aligned} \tag{4.14}$$

The inverse Laplace transform is then

$$\begin{aligned} D(E) &= \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\oint_{\mathcal{C}}\frac{d\beta}{2\pi i}\,e^{\beta E}\,\beta^{-Nd/2} \\ &= \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\frac{E^{\frac{1}{2}Nd-1}}{\Gamma(Nd/2)} \quad , \end{aligned} \tag{4.15}$$

exactly as before. The integration contour for the inverse Laplace transform is extended in an infinite semicircle in the left half $\beta$-plane. When $Nd$ is even, the function $\beta^{-Nd/2}$ has a simple pole of order $Nd/2$ at the origin. When $Nd$ is odd, there is a branch cut extending along the negative $\mathrm{Re}\,\beta$ axis, and the integration contour must avoid the cut, as shown in fig. 4.1. One can check that this results in the same expression above, *i.e.* we may analytically continue from even values of $Nd$ to all positive values of $Nd$.

For a general system, the Laplace transform, $Z(\beta) = \mathcal{L}\big[D(E)\big]$ also is called the *partition function*. We shall again meet up with $Z(\beta)$ when we discuss the ordinary canonical ensemble.

Our final result, then, is

$$D(E,V,N) = \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\frac{E^{\frac{1}{2}Nd-1}}{\Gamma(Nd/2)} \quad . \tag{4.16}$$

Here we have emphasized that the density of states is a function of $E$, $V$, and $N$. Using Stirling's approximation,

$$\log N! = N\log N - N + \tfrac{1}{2}\log N + \tfrac{1}{2}\log(2\pi) + \mathcal{O}\big(N^{-1}\big) \quad , \tag{4.17}$$

we may define the *statistical entropy*,

$$S(E,V,N) \equiv k_{\rm B}\log D(E,V,N) = Nk_{\rm B}\,\phi\left(\frac{E}{N}\,,\,\frac{V}{N}\right) + \mathcal{O}(\log N) \quad , \tag{4.18}$$

where

$$\phi\left(\frac{E}{N}\,,\,\frac{V}{N}\right) = \frac{d}{2}\,\log\left(\frac{E}{N}\right) + \log\left(\frac{V}{N}\right) + \frac{d}{2}\,\log\left(\frac{m}{d\pi\hbar^2}\right) + \big(1+\tfrac{1}{2}d\big) \quad . \tag{4.19}$$

Recall $k_{\rm B} = 1.3806503\times10^{-16}\,{\rm erg/K}$ is Boltzmann's constant.

**Second method**

The second method invokes a mathematical trick. First, let's rescale $p_i^\alpha \equiv \sqrt{2mE}\,u_i^\alpha$. We then have

$$D(E) = \frac{V^N}{N!}\left(\frac{\sqrt{2mE}}{h}\right)^{Nd}\frac{1}{E}\int d^Mu\;\delta\big(u_1^2+u_2^2+\ldots+u_M^2-1\big) \quad . \tag{4.20}$$

Here we have written $\boldsymbol{u} = (u_1,u_2,\ldots,u_M)$ with $M = Nd$ as a $M$-dimensional vector. We've also used the rule $\delta(Ex) = E^{-1}\delta(x)$ for $\delta$-functions. We can now write

$$d^Mu = u^{M-1}\,du\,d\Omega_M \quad , \tag{4.21}$$

where $d\Omega_M$ is the $M$-dimensional differential solid angle. We now have our answer:[4]

$$D(E) = \frac{V^N}{N!}\left(\frac{\sqrt{2m}}{h}\right)^{Nd}E^{\frac{1}{2}Nd-1}\cdot\tfrac{1}{2}\,\Omega_{Nd} \quad . \tag{4.22}$$

What remains is for us to compute $\Omega_M$, the total solid angle in $M$ dimensions. We do this by a nifty mathematical trick. Consider the integral

$$\begin{aligned}\mathcal{I}_M &= \int d^Mu\,e^{-u^2} = \Omega_M\int_0^\infty du\,u^{M-1}\,e^{-\boldsymbol{u}^2}\\ &= \tfrac{1}{2}\Omega_M\int_0^\infty ds\,s^{\frac{1}{2}M-1}\,e^{-s} = \tfrac{1}{2}\Omega_M\,\Gamma\big(\tfrac{1}{2}M\big) \quad ,\end{aligned} \tag{4.23}$$

where $s = u^2$, and where

$$\Gamma(z) = \int_0^\infty dt\,t^{z-1}\,e^{-t} \tag{4.24}$$

[4]The factor of $\frac{1}{2}$ preceding $\Omega_M$ in eqn. 4.22 appears because $\delta(u^2-1) = \frac{1}{2}\,\delta(u-1)+\frac{1}{2}\,\delta(u+1)$. Since $u = |\boldsymbol{u}| \geq 0$, the second term can be dropped.

is the Gamma function, which satisfies $z\,\Gamma(z) = \Gamma(z+1)$.[5] On the other hand, we can compute $\mathcal{I}_M$ in Cartesian coordinates, writing

$$\mathcal{I}_M = \left( \int_{-\infty}^{\infty} du_1 \, e^{-u_1^2} \right)^M = \left(\sqrt{\pi}\right)^M \quad . \tag{4.25}$$

Therefore

$$\Omega_M = \frac{2\pi^{M/2}}{\Gamma(M/2)} \quad . \tag{4.26}$$

Thus we obtain $\Omega_2 = 2\pi$, $\Omega_3 = 4\pi$, $\Omega_4 = 2\pi^2$, *etc.*, the first two of which are familiar.

### 4.2.3 Arbitrariness in the definition of $S(E)$

Note that $D(E)$ has dimensions of inverse energy, so one might ask how we are to take the logarithm of a dimensionful quantity in eqn. 4.18. We must introduce an energy scale, such as $\Delta E$ in eqn. 4.9, and define $\tilde{D}(E;\Delta E) = D(E)\,\Delta E$ and $S(E;\Delta E) \equiv k_{\rm B} \log \tilde{D}(E;\Delta E)$. The definition of statistical entropy then involves the arbitrary parameter $\Delta E$, however this only affects $S(E)$ in an additive way. That is,

$$S(E,V,N;\Delta E_1) = S(E,V,N;\Delta E_2) + k_{\rm B} \log\left(\frac{\Delta E_1}{\Delta E_2}\right) \quad . \tag{4.27}$$

Note that the difference between the two definitions of $S$ depends only on the ratio $\Delta E_1/\Delta E_2$, and is independent of $E$, $V$, and $N$.

### 4.2.4 Ultra-relativistic ideal gas

Consider an ultrarelativistic ideal gas, with single particle dispersion $\varepsilon(p) = cp$. We then have

$$Z(\beta) = \frac{V^N}{N!}\frac{\Omega_d^N}{h^N d}\left( \int_0^{\infty} dp\, p^{d-1}\, e^{-\beta cp} \right)^N = \frac{V^N}{N!}\left(\frac{\Gamma(d)\,\Omega_d}{c^d\,h^d\,\beta^d}\right)^N \quad . \tag{4.28}$$

The statistical entropy is $S(E,V,N) = k_{\rm B} \log D(E,V,N) = N k_{\rm B}\,\phi(\frac{E}{N},\frac{V}{N})$, with

$$\phi\left(\frac{E}{N}\,,\,\frac{V}{N}\right) = d\,\log\left(\frac{E}{N}\right) + \log\left(\frac{V}{N}\right) + \log\left(\frac{\Omega_d\,\Gamma(d)}{(dhc)^d}\right) + (d+1) \tag{4.29}$$

### 4.2.5 Discrete systems

For classical systems where the energy levels are discrete, the states of the system $|\,\boldsymbol{\sigma}\,\rangle$ are labeled by a set of discrete quantities $\{\sigma_1, \sigma_2, \ldots\}$, where each variable $\sigma_i$ takes discrete values. The number of ways

[5] Note that for integer argument, $\Gamma(k) = (k-1)!$

of configuring the system at fixed energy $E$ is then

$$\Omega(E,N,\boldsymbol{\lambda}) = \sum_{\boldsymbol{\sigma}} \delta_{\hat{H}(\boldsymbol{\sigma}),E} \quad , \tag{4.30}$$

where the sum is over all possible configurations, and where $\boldsymbol{\lambda}$ is a vector of parameters which enter into $\hat{H}(\boldsymbol{\sigma})$. Here $N$ labels the total number of particles. For example, if we have $N$ spin-$\frac{1}{2}$ particles on a lattice which are placed in a magnetic field $B$,[6] so the individual particle energy is $\varepsilon_i = -\mu_0 B\sigma_i$ , where $\sigma = \pm 1$, then in a configuration in which $N_+$ particles have $\sigma_i = +1$ and $N_- = N - N_+$ particles have $\sigma_i = -1$, the energy is $E = (N_- - N_+)\,\mu_0 B$. The number of configurations at fixed energy $E$ is

$$\Omega(E,N,B) = \binom{N}{N_+} = \frac{N!}{N_+!\,N_-!} \quad , \tag{4.31}$$

We may write $N_\pm = \frac{1}{2}N(1 \pm m)$ where $m = (N_+ - N_-)/N \in [-1,1]$ is the "magnetization." Thus $m = -E/N\mu_0 B$, and it is left as an exercise to the reader to show, using Stirling's formula, that the statistical entropy is

$$S(E,N,B) = k_{\rm B} \log \Omega(E,N,B) = -Nk_{\rm B}\left[\left(\frac{1+m}{2}\right)\log\left(\frac{1+m}{2}\right) + \left(\frac{1-m}{2}\right)\log\left(\frac{1-m}{2}\right)\right] \quad . \tag{4.32}$$

### 4.2.6 Two systems in thermal contact

Consider two systems in thermal contact, as depicted in fig. 4.2. The two subsystems #1 and #2 are free to exchange energy, but their respective volumes and particle numbers remain fixed. We assume the contact is made over a surface, and that the energy associated with that surface is negligible when compared with the bulk energies $E_1$ and $E_2$. Let the total energy be $E = E_1 + E_2$. Then the density of states $D(E)$ for the combined system is

$$D(E) = \int dE_1\, D_1(E_1)\, D_2(E - E_1) \quad . \tag{4.33}$$

The probability density for system #1 to have energy $E_1$ is then

$$P_1(E_1) = \frac{D_1(E_1)\, D_2(E - E_1)}{D(E)} \quad . \tag{4.34}$$

Note that $P_1(E_1)$ is normalized: $\int dE_1\, P_1(E_1) = 1$. We now ask: what is the most probable value of $E_1$? We find out by differentiating $P_1(E_1)$ with respect to $E_1$ and setting the result to zero. This requires

$$\begin{aligned} 0 &= \frac{1}{P_1(E_1)}\frac{dP_1(E_1)}{dE_1} = \frac{\partial}{\partial E_1}\log P_1(E_1) \\ &= \frac{\partial}{\partial E_1}\log D_1(E_1) + \frac{\partial}{\partial E_1}\log D_2(E - E_1) \quad . \end{aligned} \tag{4.35}$$

[6] Properly, we should use $H$ here rather than $B$, but to obviate any confusion between $H$ and the Hamiltonian $\hat{H}$, we use $B$ instead.

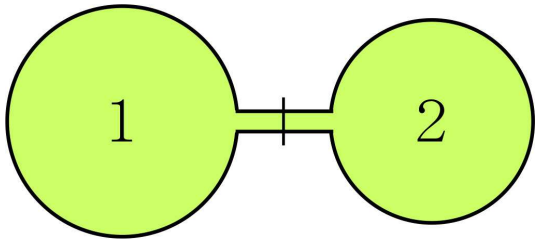



Figure 4.2: Two systems in thermal contact.

We conclude that the maximally likely partition of energy between systems #1 and #2 is realized when

$$\frac{\partial S_1}{\partial E_1} = \frac{\partial S_2}{\partial E_2} \quad . \tag{4.36}$$

This guarantees that

$$S(E, E_1) = S_1(E_1) + S_2(E - E_1) \tag{4.37}$$

is a maximum with respect to the energy $E_1$, at fixed total energy $E$.

The *temperature* $T$ is defined as

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N} \quad , \tag{4.38}$$

a result familiar from thermodynamics. The difference is now we have a more rigorous definition of the entropy. When the total entropy $S$ is extremized we have that $T_1 = T_2$. Together with the usual stability/concavity condition, this entails a maximum for $S$. Once again, two systems in thermal contact can exchange energy and in equilibrium will have equal temperatures.

According to eqns. 4.19 and 4.29, the entropies of nonrelativistic and ultrarelativistic ideal gases in $d$ space dimensions are given by

$$\begin{aligned} S_{\rm NR} &= \tfrac{1}{2} N d\, k_{\rm B} \log\left(\frac{E}{N}\right) + N k_{\rm B} \log\left(\frac{V}{N}\right) + \text{const.} \\ S_{\rm UR} &= N d\, k_{\rm B} \log\left(\frac{E}{N}\right) + N k_{\rm B} \log\left(\frac{V}{N}\right) + \text{const.} \quad . \end{aligned} \tag{4.39}$$

Invoking eqn. 4.38, we then have $E_{\rm NR} = \frac{1}{2} N d\, k_{\rm B} T$ and $E_{\rm UR} = N d\, k_{\rm B} T$.

We saw that the probability distribution $P_1(E_1)$ is maximized when $T_1 = T_2$, but how sharp is the peak in the distribution? Let us write $E_1 = E_1^* + \Delta E_1$, where $E_1^*$ is the solution to eqn. 4.35. We then have

$$\log P_1(E_1^* + \Delta E_1) = \log P_1(E_1^*) + \frac{1}{2k_{\rm B}} \frac{\partial^2 S_1}{\partial E_1^2}\bigg|_{E_1^*} (\Delta E_1)^2 + \frac{1}{2k_{\rm B}} \frac{\partial^2 S_2}{\partial E_2^2}\bigg|_{E_2^*} (\Delta E_1)^2 + \ldots \quad , \tag{4.40}$$

where $E_2^* = E - E_1^*$. We must now evaluate

$$\frac{\partial^2 S}{\partial E^2} = \frac{\partial}{\partial E}\left(\frac{1}{T}\right) = -\frac{1}{T^2}\left(\frac{\partial T}{\partial E}\right)_{V,N} = -\frac{1}{T^2\, C_V} \quad , \tag{4.41}$$

where $C_V = \big(\partial E/\partial T\big)_{V,N}$ is the heat capacity. Thus, in the vicinity of a stable entropy maximum,

$$P_1 = P_1^* \, e^{-(\Delta E_1)^2/2k_{\rm B}T^2\bar{C}_V} \quad , \tag{4.42}$$

where

$$\bar{C}_V = \frac{C_{V,1}\,C_{V,2}}{C_{V,1}+C_{V,2}} \quad . \tag{4.43}$$

The distribution is therefore a Gaussian, and the fluctuations in $\Delta E_1$ can now be computed:

$$\big\langle (\Delta E_1)^2 \big\rangle = k_{\rm B}T^2\,\bar{C}_V \qquad \Longrightarrow \qquad (\Delta E_1)_{\rm RMS} = T\sqrt{k_{\rm B}\bar{C}_V} \quad . \tag{4.44}$$

The individual heat capacities $C_{V,1}$ and $C_{V,2}$ scale with the volumes $V_1$ and $V_2$, respectively. If $V_2 \gg V_1$, then $C_{V,2} \gg C_{V,1}$, in which case $\bar{C}_V \approx C_{V,1}$. Therefore the RMS fluctuations in $\Delta E_1$ are proportional to the *square root* of the system size, whereas $E_1$ itself is extensive. Thus, the ratio $(\Delta E_1)_{\rm RMS}/E_1 \propto V^{-1/2}$ scales as the inverse square root of the volume. The distribution $P_1(E_1)$ is thus *extremely sharp*.

## 4.3 The Quantum Mechanical Trace

Thus far our understanding of ergodicity is rooted in the dynamics of classical mechanics. A Hamiltonian flow which is ergodic is one in which time averages can be replaced by phase space averages using the microcanonical ensemble. What happens, though, if our system is quantum mechanical, as all systems ultimately are?

### 4.3.1 The density matrix

Consider a system S in contact with a world W as depicted in fig. 4.3. Their union $\mathsf{U} = \mathsf{W} \cup \mathsf{S}$ forms the universe. Let us choose orthonormal bases $|\sigma\rangle$ for the system and $|i\rangle$ for the world, with $\sigma \in \{1, \ldots, D_{\mathsf{S}}\}$ and $i \in \{1, \ldots, D_{\mathsf{W}}\}$. We may use the states $|\sigma\rangle \otimes |i\rangle$ as a basis for the universe, for which a general wavefunction may be written

$$|\,\Psi\,\rangle = \sum_{\sigma=1}^{D_{\mathsf{S}}} \sum_{i=1}^{D_{\mathsf{W}}} \Psi_{\sigma i}\,|\,\sigma\,\rangle \otimes |\,i\,\rangle \quad . \tag{4.45}$$

Now consider an operator $\hat{A}$ which only acts within the system's Hilbert space $\mathcal{H}_{\mathsf{S}}$. The expectation of $\hat{A}$ is given by

$$\langle\,\Psi\,|\,\hat{A}\,|\,\Psi\,\rangle = \sum_{\sigma,i}\sum_{\tau,j} \Psi_{\sigma i}\Psi^*_{\tau j}\,\langle\,\tau\,|\,\hat{A}\,|\,\sigma\,\rangle\langle\,j\,|\,i\,\rangle = \sum_{\sigma,\tau} \varrho_{\sigma\tau}\,A_{\tau\sigma} \tag{4.46}$$

where $\varrho_{\sigma\tau} = \sum_j \Psi_{\sigma j}\Psi^*_{\tau j}$ is the *reduced density matrix* of the system. In operator form,

$$\hat{\varrho} = \sum_{\sigma,\tau} \varrho_{\sigma\tau}\,|\,\sigma\,\rangle\langle\,\tau\,| \qquad \Rightarrow \qquad \langle\hat{A}\rangle = \mathsf{Tr}\,(\hat{\varrho}\,\hat{A}) \quad . \tag{4.47}$$

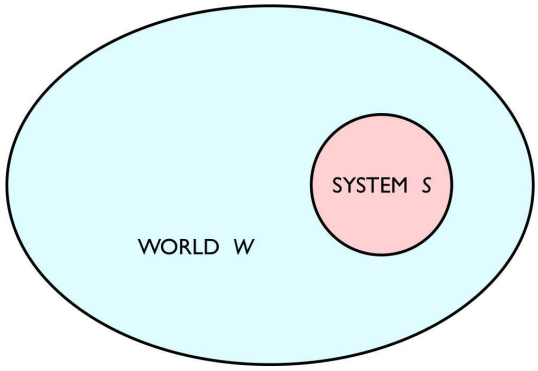



Figure 4.3: A system S in contact with a world W. The union of the two is the universe U = W ∪ S.

Note that $\varrho_{\sigma\tau} = \varrho^*_{\tau\sigma}$ hence $\hat{\varrho} = \hat{\varrho}^\dagger$ is Hermitian. Note also that $\hat{\varrho}$ is normalized:

$$\mathsf{Tr}\ \hat{\varrho} = \sum_\sigma \varrho_{\sigma\sigma} = \sum_{\sigma,j} \Psi_{\sigma j} \Psi^*_{\sigma j} = \langle \Psi | \Psi \rangle = 1 \quad . \tag{4.48}$$

The Hermitian operator $\hat{\varrho}$ has $D_{\mathsf{S}}$ eigenvalues and $D_{\mathsf{S}}$ eigenvectors. In the eigenbasis of $\hat{\varrho}$, we have

$$\hat{\varrho} = \sum_{\alpha=1}^{D_{\mathsf{S}}} \lambda_\alpha \, | \, \alpha \, \rangle \langle \, \alpha \, | \qquad \text{with} \qquad \mathsf{Tr}\ \hat{\varrho} = \sum_{\alpha=1}^{D_{\mathsf{S}}} \lambda_\alpha = 1 \tag{4.49}$$

and the matrix $\varrho_{\sigma\tau} = \lambda_\sigma \, \delta_{\sigma\tau}$ (no sum on $\sigma$) is diagonal. Another constraint is that each $\lambda_\alpha \in [0, 1]$. This is because $\langle \phi | \hat{\varrho} | \phi \rangle$ is the probability that the system is in the quantum state $|\phi\rangle$, and choosing $|\phi\rangle$ to be an eigenstate we have that its corresponding eigenvalue lies on the interval $[0, 1]$. This further entails the (weak) inequality $\mathsf{Tr}\ \hat{\varrho}^2 \leq 1$.

One example of a density matrix is a simple projection operator $\hat{P}_\sigma = |\sigma\rangle\langle\sigma|$. All projectors $\hat{P}$ are *idempotent*, meaning $\hat{P}^2 = \hat{P}$. If $\hat{\varrho} = \hat{P}_\sigma$, then $|\sigma\rangle$ is an eigenvector of $\hat{\varrho}$ with eigenvalue $\lambda = 1$, and all other vectors in the $(D_{\mathsf{S}} - 1)$-dimensional subspace of $\mathcal{H}_{\mathsf{S}}$ orthogonal to $|\sigma\rangle$ are annihilated by $\hat{\varrho}$, *i.e.* they are null eigenvectors. The eigenspectrum of $\hat{\varrho}$ thus consists of $\lambda_1 = 1$ and $\lambda_2 = \cdots = \lambda_{D_{\mathsf{S}}} = 0$. The system is then said to be in a *pure state*. This is a very special state of affairs. In the space of allowed density matrices, the vast majority describe *mixed states*. As an example of a mixed state, let $\hat{\varrho} = x \, |1\rangle\langle 1| + (1 - x) \, |2\rangle\langle 2|$. Here we have $\lambda_1 = x$, $\lambda_2 = 1 - x$, and $\lambda_3 = \cdots = \lambda_{D_S} = 0$, with $x \in (0, 1)$. A crisp distinction between pure and mixed states is that while $\mathsf{Tr}\, \hat{\varrho} = 1$ in all cases, for pure states $\mathsf{Tr}\, \hat{\varrho}^2 = 1$ while for mixed states $\mathsf{Tr}\, \hat{\varrho}^2 < 1$.

How does $\hat{\varrho}(t)$ evolve in time? If the Hamiltonian of the universe is given by $\hat{H}_{\mathsf{U}} = \hat{H}_{\mathsf{S}} + \hat{H}_{\mathsf{W}}$ and involves no coupling between the system and the world, then the analysis is simple. Let's work in eigenbases where $\hat{H}_{\mathsf{S}} \, |\alpha\rangle = \omega_\alpha \, |\alpha\rangle$ and $\hat{H}_{\mathsf{W}} \, |k\rangle = \varepsilon_k \, |k\rangle$. Then

$$| \, \Psi(t) \, \rangle = \sum_{\alpha,k} \Psi_{\alpha k}(t) \, | \, \alpha \, \rangle \otimes | \, k \, \rangle \tag{4.50}$$

where $\Psi_{\alpha k}(t) = e^{-i(\omega_\alpha + \varepsilon_k)t/\hbar} \, \Psi_{\alpha k}(0)$. The reduced density matrix for the system then obeys *Liouville's equation*,

$$i\hbar \, \frac{\partial \hat{\varrho}}{\partial t} = \big[ \hat{H}_{\mathsf{S}}, \hat{\varrho} \big] \quad . \tag{4.51}$$

Note that the density matrix evolves according to a slightly different equation than an operator in the Heisenberg picture, for which

$$\hat{A}(t) = e^{+iHt/\hbar}\, A\, e^{-i\hat{H}t/\hbar} \qquad \Longrightarrow \qquad i\hbar\,\frac{\partial \hat{A}}{\partial t} = \big[\hat{A}, \hat{H}\big] = -\big[\hat{H}, \hat{A}\big] \quad . \tag{4.52}$$

In the eigenbasis of $\hat{H}_{\mathsf{S}}$, Liouville's equation says $\varrho_{\alpha\beta}(t) = e^{i(\omega_\beta - \omega_\alpha)t/\hbar}\varrho_{\alpha\beta}(0)$, hence both $\mathsf{Tr}\ \hat{\varrho}(t) = \mathsf{Tr}\ \hat{\varrho}(0)$ and $\mathsf{Tr}\ \hat{\varrho}^2(t) = \mathsf{Tr}\ \hat{\varrho}^2(0)$. In other words, pure states evolve into pure states and mixed states evolve into mixed states. This is true for any isolated quantum system evolving purely internally, with no external couplings. What happens if there are such couplings between S and W? Consider an initial density matrix

$$\hat{\varrho}_{\mathsf{U}}(0) = \hat{\varrho}_{\mathsf{S}}(0) \otimes \hat{\varrho}_{\mathsf{W}}(0) \tag{4.53}$$

and let the (time-independent) Hamiltonian be $\hat{H}_{\mathsf{U}} = \hat{H}_{\mathsf{S}} + \hat{H}_{\mathsf{W}} + \hat{H}_{\mathsf{SW}}$ where the last term couples $S$ and $W$. The initial density matrices $\hat{\varrho}_{\mathsf{S}}(0)$ and $\hat{\varrho}_{\mathsf{W}}(0)$ are arbitrary – they could be pure or mixed. The evolution operator $\hat{U}_{\mathsf{U}}(t) = \exp(-i\hat{H}_{\mathsf{U}}t/\hbar)$ acts on the full Hilbert space $\mathcal{H}_{\mathsf{U}} = \mathcal{H}_{\mathsf{S}} \otimes \mathcal{H}_{\mathsf{W}}$. The reduced density matrix for S at time $t$ is given by

$$\hat{\varrho}_{\mathsf{S}}(t) = \operatorname*{\mathsf{Tr}}_{\mathsf{W}}\ \hat{U}(t)\,\hat{\varrho}_{\mathsf{U}}(0)\,\hat{U}^\dagger(t) \tag{4.54}$$

Let $|k\rangle$ denote an eigenstate of $\hat{\varrho}_{\mathsf{W}}(0)$ with eigenvalue $\xi_k \in [0,1]$ and $\sum_k \xi_k = 1$. We may then spectrally decompose $\hat{\varrho}_{\mathsf{W}}(0)$ as

$$\hat{\varrho}_{\mathsf{W}}(0) = \sum_k \xi_k\, |\,k\,\rangle\langle\,k\,| \quad . \tag{4.55}$$

Now define the operators

$$\hat{V}_{kl}(t) \equiv \langle\,k\,|\,\hat{U}_{\mathsf{U}}(t)\,|\,l\,\rangle \quad , \tag{4.56}$$

each of which acts as an *operator* within $\mathcal{H}_{\mathsf{S}}$. Note $\hat{V}^\dagger_{kl}(t) = \langle\,l\,|\,\hat{U}^\dagger_{\mathsf{U}}(t)\,|\,k\,\rangle$. We then have

$$\hat{\varrho}_{\mathsf{S}}(t) = \sum_{k,l} \xi_l\,\hat{V}_{kl}(t)\,\hat{\varrho}_{\mathsf{S}}(0)\,\hat{V}^\dagger_{kl}(t) = \sum_r \hat{K}_r(t)\,\hat{\varrho}_{\mathsf{S}}(0)\,\hat{K}^\dagger_r(t) \tag{4.57}$$

where $r \equiv (k,l)$ is a composite index and

$$\hat{K}_r(t) = \xi_l^{1/2}\,\hat{V}_{kl}(t) \qquad , \qquad \hat{K}^\dagger_r(t) = \xi_l^{1/2}\,\hat{V}^\dagger_{kl}(t) \quad , \tag{4.58}$$

which satisfy

$$\sum_r \hat{K}^\dagger_r(t)\,\hat{K}_r(t) = \mathbb{1}_{\mathsf{S}} \quad . \tag{4.59}$$

Fixing and suppressing $t$, eqns. 4.57 and 4.59 define the *Kraus map* on the space of density matrices, *viz.*

$$\hat{\varrho}' = \sum_{r=1}^{M} \hat{K}_r\,\hat{\varrho}\,\hat{K}^\dagger_r \qquad , \qquad \sum_{r=1}^{M} \hat{K}^\dagger_r \hat{K}_r = \mathbb{1} \quad . \tag{4.60}$$

Here $M$ is the total number of Kraus operators; in the above discussion we have $M = D^2_{\mathsf{W}}$ as there is one per $(k,l)$ pair. The Kraus map is linear, trace-preserving, Hermiticity preserving, and satisfies an

additional desideratum, called "complete positivity," which we will not discuss here. When there is only one Kraus operator $\hat{K}$, we have $\hat{K}^\dagger \hat{K} = \mathbb{1}$ and unitary evolution. When there is more than one Kraus operator $\hat{K}_r$, the image $\hat{\varrho}'$ of $\hat{\varrho}$ under the Kraus map is Hermitian and satisfies $\mathsf{Tr}\ \hat{\varrho}' = 1$, but unlike unitary evolution the Kraus map does not separately preserve pure and mixed states. Owing to the environmental coupling, a pure state of S can evolve into a mixed state and *vice versa*. If $\varrho_{\mathsf{S}}(0)$ is pure, its evolution into a mixed state is a signature of *quantum entanglement* between S and W. In general, environmental coupling allows a pure reduced state to evolve into a mixed one, although subsequent recoherence or disentanglement is also possible.

In order to iterate the Kraus map using the same Kraus operators, we adopt the following protocol. We start with $\hat{\varrho}_{\mathsf{U}} \equiv \hat{\varrho}_n \otimes \hat{\varrho}^0_{\mathsf{W}}$. We evolve $\hat{\varrho}_{\mathsf{U}}$ with $\hat{U}_{\mathsf{U}}(\Delta t)$ over a fixed time interval $\Delta t$, then trace out over the W degrees of freedom, leaving us with the next iterate of the reduced density matrix of S, *i.e.* $\hat{\varrho}_{n+1}$. Then we disentangle S and W, reset $\hat{\varrho}_{\mathsf{U}} \equiv \hat{\varrho}_{n+1} \otimes \hat{\varrho}^0_{\mathsf{W}}$, and repeat the process. Physically, we can imagine that the degrees of freedom (*i.e.* qubits) in W that were entangled with S during the previous iteration have been advected away from the interface of S and W, and new unentangled qubits described by the same W density matrix take their place. We then have $\hat{\varrho}_{n+1} = \sum_r \hat{K}_r\, \hat{\varrho}_n\, \hat{K}_r^\dagger$, with the same set of Kraus operators. Lather, rinse, repeat.

For classical Hamiltonian systems, we found that the phase space distribution $\varrho(q, p, t)$ evolved according to the Liouville equation, $i\,\partial\varrho/\partial t = \hat{L}\,\varrho$, where the Liouvillian $\hat{L}$ is the differential operator

$$\hat{L} = -i \sum_{j=1}^{Nd} \left\{ \frac{\partial \hat{H}}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial \hat{H}}{\partial q_j} \frac{\partial}{\partial p_j} \right\} \quad . \tag{4.61}$$

Accordingly, any distribution $\varrho(I_1, \ldots, I_k)$ which is a function of constants of the motion $I_a(q, p)$ is a stationary solution to the Liouville equation: $\partial_t\, \varrho(I_1, \ldots, I_k) = 0$. Similarly, any quantum mechanical density matrix which commutes with the Hamiltonian is a stationary solution to eqn. 4.51. The corresponding microcanonical distribution is $\hat{\varrho}_E = \delta\big(E - \hat{H}\big)/D(E)$ where $D(E)$ is the density of states.

### 4.3.2 Averaging the DOS

If our quantum mechanical system is placed in a finite volume, the energy levels will be discrete, rather than continuous, and the density of states (DOS) will be of the form

$$D(E) = \mathsf{Tr}\ \delta\big(E - \hat{H}\big) = \sum_l \delta(E - E_l) \quad , \tag{4.62}$$

where $\{E_l\}$ are the eigenvalues of the Hamiltonian $\hat{H}$. In the thermodynamic limit, $V \to \infty$, and the discrete spectrum of kinetic energies remains discrete for all finite $V$ but must approach the continuum result. To recover the continuum result, we average the DOS over a window of width $\Delta E$:

$$\overline{D(E)} = \frac{1}{\Delta E} \int_E^{E+\Delta E} dE'\, D(E') \quad . \tag{4.63}$$

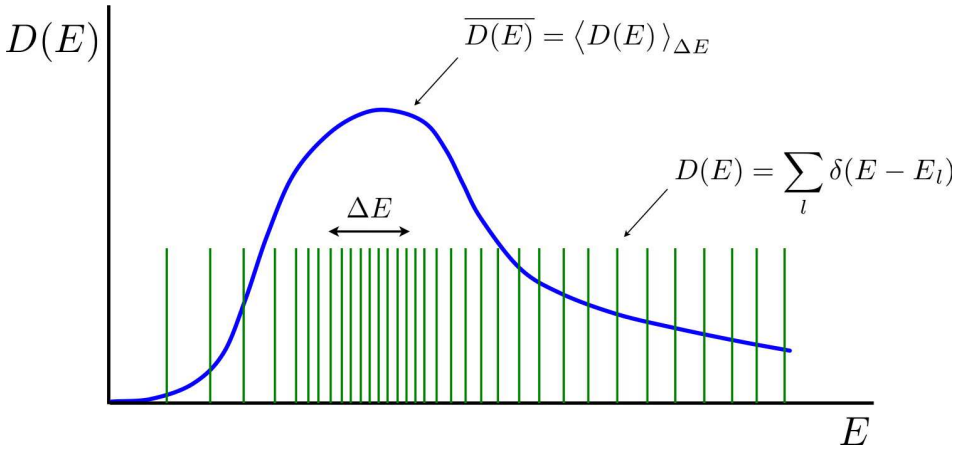



Figure 4.4: Averaging the quantum mechanical discrete density of states yields a continuous curve.

If we take the limit $\Delta E \to 0$ but with $\Delta E \gg \delta E$, where $\delta E$ is the spacing between successive quantized levels, we recover a smooth function, as shown in fig. 4.4. We will in general drop the bar and refer to this function as $D(E)$. Note that $\delta E \sim 1/D(E) = \exp\bigl[-N\phi(\varepsilon, v)\bigr]$ is (typically) *exponentially* small in the size of the system, hence if we took $\Delta E \propto V^{-1}$ which vanishes in the thermodynamic limit, there are still exponentially many energy levels within an interval of width $\Delta E$.

### 4.3.3 Coherent states

The quantum-classical correspondence is elucidated with the use of *coherent states*. Recall that the one-dimensional harmonic oscillator Hamiltonian may be written

$$\hat{H}_0 = \frac{p^2}{2m} + \tfrac{1}{2} m\,\omega_0^2\,q^2 = \hbar\omega_0\left(a^\dagger a + \tfrac{1}{2}\right) \quad , \tag{4.64}$$

where $a$ and $a^\dagger$ are *ladder operators* satisfying $\bigl[a, a^\dagger\bigr] = 1$, which can be taken to be

$$a = \ell\,\frac{\partial}{\partial q} + \frac{q}{2\ell} \qquad , \qquad a^\dagger = -\ell\,\frac{\partial}{\partial q} + \frac{q}{2\ell} \quad , \tag{4.65}$$

with $\ell = \sqrt{\hbar/2m\omega_0}$ . Note that $q = \ell\,(a + a^\dagger)$ and $p = \frac{\hbar}{2i\ell}\,(a - a^\dagger)$.

The ground state satisfies $a\,\psi_0(q) = 0$, which yields

$$\psi_0(q) = (2\pi\ell^2)^{-1/4}\,e^{-q^2/4\ell^2} \quad . \tag{4.66}$$

The normalized *coherent state* $|\,z\,\rangle$ is defined as

$$|\,z\,\rangle = e^{-\frac{1}{2}|z|^2}\,e^{za^\dagger}\,|\,0\,\rangle = e^{-\frac{1}{2}|z|^2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,|\,n\,\rangle \quad . \tag{4.67}$$

The overlap of coherent states is given by

$$\langle\,z_1\,|\,z_2\,\rangle = e^{-\frac{1}{2}|z_1|^2}\,e^{-\frac{1}{2}|z_2|^2}\,e^{\bar{z}_1 z_2} \quad , \tag{4.68}$$

hence different coherent states are not orthogonal. Despite this nonorthogonality, the coherent states allow a simple resolution of the identity,

$$1 = \int \frac{d^2z}{2\pi i} \, |\, z \,\rangle\langle\, z \,| \quad ; \quad \frac{d^2z}{2\pi i} \equiv \frac{d\,\mathrm{Re}z \, d\,\mathrm{Im}z}{\pi} \tag{4.69}$$

which is straightforward to establish.

To gain some physical intuition about the coherent states, define

$$z \equiv \frac{Q}{2\ell} + \frac{i\ell P}{\hbar} \tag{4.70}$$

and write $|\, z \,\rangle \equiv |\, Q, P \,\rangle$. One finds (*exercise!*)

$$\psi_{Q,P}(q) = \langle\, q \,|\, z \,\rangle = (2\pi\ell^2)^{-1/4} \, e^{-iPQ/2\hbar} \, e^{iPq/\hbar} \, e^{-(q-Q)^2/4\ell^2} \quad , \tag{4.71}$$

hence the coherent state $\psi_{Q,P}(q)$ is a wavepacket Gaussianly localized about $q = Q$, but oscillating with average momentum $P$.

For example, we can compute

$$\langle\, Q, P \,|\, q \,|\, Q, P \,\rangle = \langle\, z \,|\, \ell\,(a + a^\dagger) \,|\, z \,\rangle = 2\ell\,\mathrm{Re}\, z = Q \tag{4.72}$$

$$\langle\, Q, P \,|\, p \,|\, Q, P \,\rangle = \langle\, z \,|\, \frac{\hbar}{2i\ell}\,(a - a^\dagger) \,|\, z \,\rangle = \frac{\hbar}{\ell}\,\mathrm{Im}\, z = P \tag{4.73}$$

as well as

$$\langle\, Q, P \,|\, q^2 \,|\, Q, P \,\rangle = \langle\, z \,|\, \ell^2\,(a + a^\dagger)^2 \,|\, z \,\rangle = Q^2 + \ell^2 \tag{4.74}$$

$$\langle\, Q, P \,|\, p^2 \,|\, Q, P \,\rangle = -\langle\, z \,|\, \frac{\hbar^2}{4\ell^2}(a - a^\dagger)^2 \,|\, z \,\rangle = P^2 + \frac{\hbar^2}{4\ell^2} \quad . \tag{4.75}$$

Thus, the root mean square fluctuations in the coherent state $|\, Q, P \,\rangle$ are

$$\Delta q = \ell = \sqrt{\frac{\hbar}{2m\omega_0}} \qquad , \qquad \Delta p = \frac{\hbar}{2\ell} = \sqrt{\frac{m\hbar\omega_0}{2}} \quad , \tag{4.76}$$

and $\Delta q \cdot \Delta p = \frac{1}{2}\,\hbar$. Thus we learn that the coherent state $\psi_{Q,P}(q)$ is localized in phase space, *i.e.* in both position and momentum. If we have a general operator $\hat{A}(q, p)$, we can then write

$$\langle\, Q, P \,|\, \hat{A}(q, p) \,|\, Q, P \,\rangle = A(Q, P) + \mathcal{O}(\hbar) \quad , \tag{4.77}$$

where $A(Q, P)$ is formed from $\hat{A}(q, p)$ by replacing $q \to Q$ and $p \to P$.

Since

$$\frac{d^2z}{2\pi i} \equiv \frac{d\,\mathrm{Re}z \, d\,\mathrm{Im}z}{\pi} = \frac{dQ\, dP}{2\pi\hbar} \quad , \tag{4.78}$$

we can write the trace using coherent states as

$$\mathsf{Tr}\ \hat{A} = \frac{1}{2\pi\hbar} \int\limits_{-\infty}^{\infty} dQ \int\limits_{-\infty}^{\infty} dP \, \langle\, Q, P \,|\, \hat{A} \,|\, Q, P \,\rangle \quad . \tag{4.79}$$

We now can understand the origin of the factor $2\pi\hbar$ in the denominator of each $(q_i, p_i)$ integral over classical phase space in eqn. 4.6.

Note that $\omega_0$ is arbitrary in our discussion. By increasing $\omega_0$, the states become more localized in $q$ and more plane wave like in $p$. However, so long as $\omega_0$ is finite, the width of the coherent state in each direction is proportional to $\hbar^{1/2}$, and thus vanishes in the classical limit.

## 4.4 Ordinary Canonical Ensemble (OCE)

### 4.4.1 Canonical distribution and partition function

Consider a system $S$ in contact with a world $W$, and let their union $U = W \cup S$ be called the 'universe'. The situation is depicted in fig. 4.3. The volume $V_{\rm S}$ and particle number $N_{\rm S}$ of the system are held fixed, but the energy is allowed to fluctuate by exchange with the world $W$. We are interested in the limit $N_{\rm S} \to \infty$, $N_{\rm W} \to \infty$, with $N_{\rm S} \ll N_{\rm W}$, with similar relations holding for the respective volumes and energies. We now ask what is the probability that $S$ is in a state $|\, n \,\rangle$ with energy $E_n$. This is given by the ratio

$$\begin{aligned} P_n &= \lim_{\Delta E \to 0} \frac{D_{\rm W}(E_{\rm U} - E_n)\,\Delta E}{D_{\rm U}(E_{\rm U})\,\Delta E} \\ &= \frac{\text{\# of states accessible to } W \text{ given that } E_{\rm S} = E_n}{\text{total \# of states in } U} \quad . \end{aligned} \tag{4.80}$$

Then

$$\begin{aligned} \log P_n &= \log D_{\rm W}(E_{\rm U} - E_n) - \log D_{\rm U}(E_{\rm U}) \\ &= \log D_{\rm W}(E_{\rm U}) - \log D_{\rm U}(E_{\rm U}) - E_n \left.\frac{\partial \log D_{\rm W}(E)}{\partial E}\right|_{E=E_{\rm U}} + \ldots \quad . \end{aligned} \tag{4.81}$$

The higher order terms are negligible if $\mathsf{vol}(S) \ll \mathsf{vol}(W)$. In this case we have $\log P_n = -\alpha - \beta E_n$, with

$$\beta = \left.\frac{\partial \log D_{\rm W}(E)}{\partial E}\right|_{E=E_{\rm U}} \equiv \frac{1}{k_{\rm B}T} \quad . \tag{4.82}$$

The constant $\alpha$ is fixed by the requirement that $\sum_n P_n = 1$, and thus we obtain

$$P_n = \frac{1}{Z}\, e^{-E_n/k_{\rm B}T} \qquad , \qquad Z(T,V,N) = \mathsf{Tr}\; e^{-\beta\hat{H}} = \sum_n e^{-E_n/k_{\rm B}T} \quad . \tag{4.83}$$

We define the Helmholtz free energy $F(T,V,N)$ as

$$F(T,V,N) = -k_{\rm B}T \log Z(T,V,N) \quad . \tag{4.84}$$

We've already met $Z(\beta)$ in eqn. 4.12 – it is the Laplace transform of the density of states. It is also called the *partition function* of the system $S$. Quantum mechanically, we can write the ordinary canonical density matrix as

$$\hat{\varrho} = \frac{e^{-\beta\hat{H}}}{\mathsf{Tr}\; e^{-\beta\hat{H}}} \quad , \tag{4.85}$$

which is known as the *Gibbs distribution*. Note that $\left[\hat{\varrho}, \hat{H}\right] = 0$, hence the ordinary canonical distribution is a stationary solution to the evolution equation for the density matrix. Note that the OCE is specified by three parameters: $T$, $V$, and $N$.

### 4.4.2 The difference between $P(E_n)$ and $P_n$

Let the total energy of the Universe be fixed at $E_{\rm U}$. The joint probability density $P(E_{\rm S}, E_{\rm W})$ for the system to have energy $E_{\rm S}$ and the world to have energy $E_{\rm W}$ is

$$P(E_{\rm S}, E_{\rm W}) = D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm W})\, \delta(E_{\rm U} - E_{\rm S} - E_{\rm W}) / D_{\rm U}(E_{\rm U}) \quad , \tag{4.86}$$

where

$$D_{\rm U}(E_{\rm U}) = \int\limits_{-\infty}^{\infty} dE_{\rm S}\; D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm U} - E_{\rm S}) \quad , \tag{4.87}$$

which ensures that $\int dE_{\rm S} \int dE_{\rm W}\, P(E_{\rm S}, E_{\rm W}) = 1$. The *probability density* $P(E_{\rm S})$ is defined such that $P(E_{\rm S})\, dE_{\rm S}$ is the (differential) probability for the system to have an energy in the range $[E_{\rm S}, E_{\rm S} + dE_{\rm S}]$. The units of $P(E_{\rm S})$ are $E^{-1}$. To obtain $P(E_{\rm S})$, we simply integrate the joint probability density $P(E_{\rm S}, E_{\rm W})$ over all possible values of $E_{\rm W}$, obtaining

$$P(E_{\rm S}) = \frac{D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm U} - E_{\rm S})}{D_{\rm U}(E_{\rm U})} \quad , \tag{4.88}$$

as we have in eqn. 4.80. Suppose we wish to know the probability $P_n$ that the system is in a *particular* state $|\, n \,\rangle$ with energy $E_n$. Clearly

$$P_n = \lim_{\Delta E \to 0} \frac{\text{probability that } E_{\rm S} \in [E_n, E_n + \Delta E]}{\text{\# of S states with } E_{\rm S} \in [E_n, E_n + \Delta E]} = \frac{P(E_n)\, \Delta E}{D_{\rm S}(E_n)\, \Delta E} = \frac{D_{\rm W}(E_{\rm U} - E_n)}{D_{\rm U}(E_{\rm U})} \quad . \tag{4.89}$$

### 4.4.3 Additional remarks

The formula of eqn. 4.80 is quite general and holds in the case where $N_{\rm S}/N_{\rm W} = \mathcal{O}(1)$, so long as we are in the thermodynamic limit, where the energy associated with the interface between S and W may be neglected. In this case, however, one is not licensed to perform the subsequent Taylor expansion, and the distribution $P_n$ is no longer of the Gibbs form. It is also valid for quantum systems[7], in which case we interpret $P_n = \langle n | \varrho_{\rm S} | n \rangle$ as a diagonal element of the density matrix $\varrho_{\rm S}$. The density of states functions may then be replaced by

$$\begin{aligned} D_{\rm W}(E_{\rm U} - E_n)\, \Delta E \to e^{S_{\rm W}(E_{\rm U} - E_n\,,\, \Delta E)} &\equiv \mathop{\rm Tr}_{\rm W} \int\limits_{E_{\rm U} - E_n}^{E_{\rm U} - E_n + \Delta E} dE\; \delta(E - \hat{H}_{\rm W}) \\ D_{\rm U}(E_{\rm U})\, \Delta E \to e^{S_{\rm U}(E_{\rm U}\,,\, \Delta E)} &\equiv \mathop{\rm Tr}_{\rm U} \int\limits_{E_{\rm U}}^{E_{\rm U} + \Delta E} dE\; \delta(E - \hat{H}_{\rm U}) \quad . \end{aligned} \tag{4.90}$$

[7] See T.-C. Lu and T. Grover, arXiv 1709.08784.

The off-diagonal matrix elements of $\varrho_{\rm S}$ are negligible in the thermodynamic limit.

### 4.4.4 Averages within the OCE

To compute averages within the OCE,

$$\big\langle \hat{A} \big\rangle = \mathsf{Tr}\,\big(\hat{\varrho}\,\hat{A}\big) = \frac{\sum_n \langle n|\hat{A}|n\rangle\, e^{-\beta E_n}}{\sum_j e^{-\beta E_j}} \quad , \tag{4.91}$$

where we have conveniently taken the trace in a basis of energy eigenstates. In the classical limit, we have

$$\varrho(\boldsymbol{\varphi}) = \frac{1}{Z}\, e^{-\beta \hat{H}(\boldsymbol{\varphi})} \quad , \quad Z = \mathsf{Tr}\; e^{-\beta \hat{H}} = \int d\mu\, e^{-\beta \hat{H}(\boldsymbol{\varphi})} \quad , \tag{4.92}$$

with $d\mu = \frac{1}{N!}\prod_{j=1}^{N}(d^d q_j\, d^d p_j/h^d)$ for identical particles ('Maxwell-Boltzmann statistics'). Thus,

$$\langle A \rangle = \mathsf{Tr}\,(\varrho A) = \frac{\int d\mu\; A(\boldsymbol{\varphi})\, e^{-\beta \hat{H}(\boldsymbol{\varphi})}}{\int d\mu\, e^{-\beta \hat{H}(\boldsymbol{\varphi})}} \quad . \tag{4.93}$$

### 4.4.5 Entropy and free energy

The *Boltzmann entropy* is defined by

$$S = -k_{\rm B}\mathsf{Tr}\,(\hat{\varrho}\log\hat{\varrho}) = -k_{\rm B}\sum_n P_n\, \log P_n \quad . \tag{4.94}$$

The Boltzmann entropy and the statistical entropy $S = k_{\rm B}\log D(E)$ are identical in the thermodynamic limit. Since $\log P_n = \beta(F - E_n)$, we have

$$S = -k_{\rm B}\sum_n P_n\,\big(\beta F - \beta E_n\big) = -\frac{F}{T} + \frac{\langle\, \hat{H}\,\rangle}{T} \quad , \tag{4.95}$$

which is to say $F = E - TS$, where

$$E = \sum_n P_n\, E_n = \frac{\mathsf{Tr}\;\hat{H}\, e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} \tag{4.96}$$

is the average energy. We also see that

$$Z = \mathsf{Tr}\; e^{-\beta \hat{H}} = \sum_n e^{-\beta E_n} \quad \Longrightarrow \quad E = \frac{\sum_n E_n\, e^{-\beta E_n}}{\sum_j e^{-\beta E_j}} = -\frac{\partial}{\partial\beta}\,\log Z = \frac{\partial}{\partial\beta}\big(\beta F\big) \quad . \tag{4.97}$$

Thus, $F(T,V,N)$ is a Legendre transform of $E(S,V,N)$, with

$$dF = -S\,dT - p\,dV + \mu\,dN \quad , \tag{4.98}$$

which means

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \quad , \quad p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \quad , \quad \mu = +\left(\frac{\partial F}{\partial N}\right)_{T,V} \quad . \tag{4.99}$$

### 4.4.6 Fluctuations in the OCE

In the OCE, the energy is not fixed. It therefore fluctuates about its average value $E = \langle \hat{H} \rangle$. Note that

$$\begin{aligned} -\frac{\partial E}{\partial \beta} &= k_{\rm B} T^2 \, \frac{\partial E}{\partial T} = \frac{\partial^2 \log Z}{\partial \beta^2} \\ &= \frac{\mathsf{Tr}\ \hat{H}^2 \, e^{-\beta \hat{H}}}{\mathsf{Tr}\ e^{-\beta \hat{H}}} - \left( \frac{\mathsf{Tr}\ \hat{H} \, e^{-\beta \hat{H}}}{\mathsf{Tr}\ e^{-\beta \hat{H}}} \right)^2 \\ &= \big\langle \hat{H}^2 \big\rangle - \big\langle \hat{H} \big\rangle^2 \quad . \end{aligned} \tag{4.100}$$

Thus, the heat capacity is related to the fluctuations in the energy, just as we saw in §4.2.6:

$$C_V = \left( \frac{\partial E}{\partial T} \right)_{V,N} = \frac{1}{k_{\rm B} T^2} \left( \big\langle \hat{H}^2 \big\rangle - \big\langle \hat{H} \big\rangle^2 \right) \qquad \Rightarrow \qquad \frac{(\Delta E)_{\rm RMS}}{E} = \frac{\sqrt{k_{\rm B} T^2 \, C_V}}{E} \quad . \tag{4.101}$$

For the nonrelativistic ideal gas, we found $C_V = \frac{d}{2} N k_{\rm B}$, hence the ratio of RMS fluctuations in the energy to the energy itself is

$$\left( \frac{(\Delta E)_{\rm RMS}}{E} \right)_{\rm NRIG} = \sqrt{\frac{2}{Nd}} \quad , \tag{4.102}$$

which scales as $N^{-1/2}$ and thus vanishes in the thermodynamic limit.

### 4.4.7 Energy distribution within the OCE

The full distribution function for the energy within the ordinary canonical ensemble is

$$P(\mathcal{E}) = \big\langle \delta(\mathcal{E} - \hat{H}) \big\rangle = \frac{\mathsf{Tr} \big[ \delta(\mathcal{E} - \hat{H}) \, e^{-\beta \hat{H}} \big]}{\mathsf{Tr}\ e^{-\beta \hat{H}}} = \frac{1}{Z} \, D(\mathcal{E}) \, e^{-\beta \mathcal{E}} \quad . \tag{4.103}$$

Thus,

$$P(\mathcal{E}) = \frac{e^{-\beta [\mathcal{E} - T S(\mathcal{E})]}}{\int d\mathcal{E}' \, e^{-\beta [\mathcal{E}' - T S(\mathcal{E}')]}} \quad , \tag{4.104}$$

where $S(\mathcal{E}) = k_{\rm B} \log D(\mathcal{E})$ is the statistical entropy. Let's write $\mathcal{E} = E + \delta\mathcal{E}$, where $E$ extremizes the combination $\mathcal{E} - T\,S(\mathcal{E})$, *i.e.* the solution to $T\,S'(E) = 1$, where the energy derivative of $S$ is performed at fixed volume $V$ and particle number $N$. We now expand $S(E + \delta\mathcal{E})$ to second order in $\delta\mathcal{E}$, obtaining

$$S(E + \delta\mathcal{E}) = S(E) + \frac{\delta\mathcal{E}}{T} - \frac{(\delta\mathcal{E})^2}{2T^2 \, C_V} + \ldots \tag{4.105}$$

Recall that $S''(E) = \partial(T^{-1})/\partial E = -1/T^2 C_V$. Thus,

$$\mathcal{E} - T\,S(\mathcal{E}) = E - T\,S(E) + \frac{(\delta\mathcal{E})^2}{2T\,C_V} + \mathcal{O}\big((\delta\mathcal{E})^3\big) \quad . \tag{4.106}$$

Applying this to both numerator and denominator of eqn. 4.104, we obtain[8]

$$P(\mathcal{E}) = \mathcal{N} \exp\left[ - \frac{(\delta\mathcal{E})^2}{2 k_{\rm B} T^2 \, C_V} \right] \quad , \tag{4.107}$$

where $\mathcal{N} = (2\pi k_{\rm B} T^2 C_V)^{-1/2}$ is a normalization constant which guarantees $\int d\mathcal{E}\, P(\mathcal{E}) = 1$. Once again, we see that the distribution is a Gaussian centered at $\langle \mathcal{E} \rangle = E$, and of width $(\Delta\mathcal{E})_{\rm RMS} = \sqrt{k_{\rm B} T^2 \, C_V}$. This is a consequence of the Central Limit Theorem.

### 4.4.8 Thermodynamics revisited

The average energy within the OCE is

$$E = \sum_n E_n P_n \quad , \tag{4.108}$$

and therefore

$$dE = \sum_n E_n \, dP_n + \sum_n P_n \, dE_n = đQ - đW \quad , \tag{4.109}$$

where

$$đQ = \sum_n E_n \, dP_n \qquad , \qquad đW = -\sum_n P_n \, dE_n \quad . \tag{4.110}$$

Finally, from $P_n = Z^{-1} \, e^{-E_n/k_{\rm B}T}$, we can write

$$E_n = -k_{\rm B} T \log Z - k_{\rm B} T \log P_n \quad , \tag{4.111}$$

with which we obtain

$$\begin{aligned} đQ = \sum_n E_n \, dP_n &= -k_{\rm B} T \log Z \sum_n dP_n - k_{\rm B} T \sum_n \log P_n \, dP_n \\ &= T \, d\Big( - k_{\rm B} \sum_n P_n \log P_n \Big) = T \, dS \quad . \end{aligned} \tag{4.112}$$

Note also that

$$\begin{aligned} đW = -\sum_n P_n \, dE_n &= -\sum_n P_n \left( \sum_i \frac{\partial E_n}{\partial X_i} \, dX_i \right) \\ &= -\sum_{n,i} P_n \, \langle \, n \, | \, \frac{\partial \hat{H}}{\partial X_i} \, | \, n \, \rangle \, dX_i \equiv \sum_i F_i \, dX_i \quad , \end{aligned} \tag{4.113}$$

so the generalized force $F_i$ conjugate to the generalized displacement $dX_i$ is

$$F_i = -\sum_n P_n \, \frac{\partial E_n}{\partial X_i} = -\left\langle \frac{\partial \hat{H}}{\partial X_i} \right\rangle \quad . \tag{4.114}$$

[8] In applying eqn. 4.106 to the denominator of eqn. 4.104, we shift $\mathcal{E}'$ by $E$ and integrate over the difference $\delta\mathcal{E}' \equiv \mathcal{E}' - E$, retaining terms up to quadratic order in $\delta\mathcal{E}'$ in the argument of the exponent.

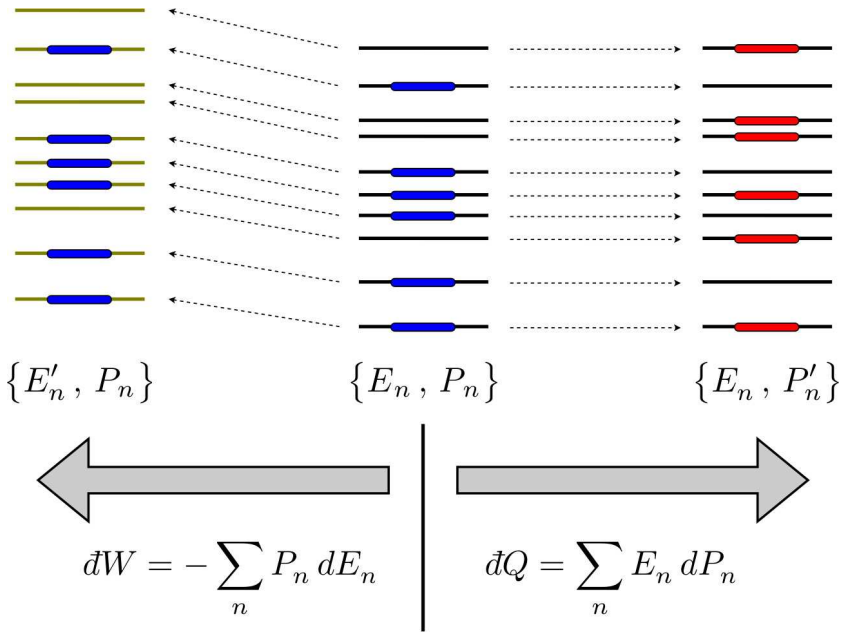



Figure 4.5: Microscopic, statistical interpretation of the First Law of Thermodynamics.

This is the force acting *on* the system[9]. In the chapter on thermodynamics, we defined the generalized force conjugate to $X_i$ as $y_i \equiv -F_i$.

Thus we see from eqn. 4.109 that there are two ways that the average energy can change; these are depicted in the sketch of fig. 4.5. Starting from a set of energy levels $\{E_n\}$ and probabilities $\{P_n\}$, we can shift the energies to $\{E'_n\}$. The resulting change in energy $(\Delta E)_{\rm I} = -W$ is identified with the work done *on* the system. We could also modify the probabilities to $\{P'_n\}$ without changing the energies. The energy change in this case is the heat absorbed *by* the system: $(\Delta E)_{\rm II} = Q$. This provides us with a statistical and microscopic interpretation of the First Law of Thermodynamics.

### 4.4.9 Generalized susceptibilities

Suppose our Hamiltonian is of the form

$$\hat{H} = \hat{H}(\lambda) = \hat{H}_0 - \lambda\,\hat{Q} \quad , \tag{4.115}$$

where $\lambda$ is an intensive parameter, such as magnetic field. Then $Z(\lambda) = \mathrm{Tr}\; e^{-\beta(\hat{H}_0 - \lambda\hat{Q})}$ and

$$\frac{1}{Z}\frac{\partial Z}{\partial \lambda} = \beta \cdot \frac{1}{Z}\,\mathrm{Tr}\left(\hat{Q}\,e^{-\beta\hat{H}(\lambda)}\right) = \beta\,\langle\hat{Q}\rangle \quad . \tag{4.116}$$

But then from $Z = e^{-\beta F}$ we have

$$Q(\lambda, T) = \langle\,\hat{Q}\,\rangle = -\left(\frac{\partial F}{\partial \lambda}\right)_T \quad . \tag{4.117}$$

[9] In deriving eqn. 4.114, we have used the so-called Feynman-Hellman theorem of quantum mechanics: $d\langle n|\hat{H}|n\rangle = \langle n|\,d\hat{H}\,|n\rangle$, if $|n\rangle$ is an energy eigenstate.

Typically we will take $Q$ to be an extensive quantity. We can now define the *susceptibility* $\chi$ as

$$\chi = \frac{1}{V}\frac{\partial Q}{\partial \lambda} = -\frac{1}{V}\frac{\partial^2 F}{\partial \lambda^2} \quad . \tag{4.118}$$

The volume factor in the denominator ensures that $\chi$ is intensive.

It is important to realize that we have assumed here that $\left[\hat{H}_0\,,\,\hat{Q}\right] = 0$, *i.e.* the 'bare' Hamiltonian $\hat{H}_0$ and the operator $\hat{Q}$ commute. If they do not commute, then the response functions must be computed within a proper quantum mechanical formalism, which we shall not discuss here.

Note also that we can imagine an entire family of observables $\left\{\hat{Q}_k\right\}$ satisfying $\left[\hat{Q}_k\,,\,\hat{Q}_{k'}\right] = 0$ and $\left[\hat{H}_0\,,\,\hat{Q}_k\right] = 0$, for all $k$ and $k'$. Then for the Hamiltonian

$$\hat{H}(\boldsymbol{\lambda}) = \hat{H}_0 - \sum_k \lambda_k \hat{Q}_k \quad , \tag{4.119}$$

we have that

$$Q_k(\boldsymbol{\lambda}, T) = \langle\, \hat{Q}_k \,\rangle = -\left(\frac{\partial F}{\partial \lambda_k}\right)_{T,\,\{N_a\},\,\{\lambda_j (\neq k)\}} \tag{4.120}$$

and we may define an entire matrix of susceptibilities,

$$\chi_{kl} \equiv \frac{1}{V}\frac{\partial Q_k}{\partial \lambda_l} = -\frac{1}{V}\frac{\partial^2 F}{\partial \lambda_k\,\partial \lambda_l} \quad . \tag{4.121}$$

## 4.5 Grand Canonical Ensemble (GCE)

### 4.5.1 Grand canonical distribution and partition function

Consider once again the situation depicted in fig. 4.3, where a system $S$ is in contact with a world $W$, their union $U = W \cup S$ being called the 'universe'. We assume that the system's volume $V_{\rm S}$ is fixed, but otherwise it is allowed to exchange energy and particle number with $W$. Hence, the system's energy $E_{\rm S}$ and particle number $N_{\rm S}$ will fluctuate. We ask what is the probability that $S$ is in a state $|\, n \,\rangle$ with energy $E_n$ and particle number $N_n$. This is given by the ratio

$$\begin{aligned} P_n &= \lim_{\Delta E \to 0}\,\lim_{\Delta N \to 0}\,\frac{D_{\rm W}(E_{\rm U} - E_n\,,\,N_{\rm U} - N_n)\,\Delta E\,\Delta N}{D_{\rm U}(E_{\rm U}, N_{\rm U})\,\Delta E\,\Delta N} \\ &= \frac{\text{\# of states accessible to } W \text{ given that } E_{\rm S} = E_n \text{ and } N_{\rm S} = N_n}{\text{total \# of states in } U} \quad . \end{aligned} \tag{4.122}$$

Then

$$\begin{aligned} \log P_n &= \log D_{\rm W}(E_{\rm U} - E_n\,,\,N_{\rm U} - N_n) - \log D_{\rm U}(E_{\rm U}, N_{\rm U}) \\ &= \log D_{\rm W}(E_{\rm U}, N_{\rm U}) - \log D_{\rm U}(E_{\rm U}, N_{\rm U}) \\ &\qquad - E_n \left.\frac{\partial \log D_{\rm W}(E,N)}{\partial E}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} - N_n \left.\frac{\partial \log D_{\rm W}(E,N)}{\partial N}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} + \ldots \end{aligned} \tag{4.123}$$

and thus $\log P_n = -\alpha - \beta E_n + \beta\mu N_n$, with

$$\beta = \left.\frac{\partial \log D_{\rm W}(E,N)}{\partial E}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} = \frac{1}{k_{\rm B}T} \qquad , \qquad \mu = -k_{\rm B}T\left.\frac{\partial \log D_{\rm W}(E,N)}{\partial N}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} \quad . \tag{4.124}$$

The quantity $\mu$ has dimensions of energy and is called the *chemical potential*. *Nota bene:* Some texts define the 'grand canonical Hamiltonian' $\hat{K}$ as $\hat{K} \equiv \hat{H} - \mu\hat{N}$. Thus, $P_n = e^{-\alpha}\, e^{-\beta(E_n - \mu N_n)}$. Once again, the constant $\alpha$ is fixed by the requirement that $\sum_n P_n = 1$:

$$\begin{aligned} P_n &= \frac{1}{\Xi}\, e^{-\beta(E_n-\mu N_n)} \\ \Xi(\beta,V,\mu) &= \sum_n e^{-\beta(E_n-\mu N_n)} = \mathsf{Tr}\; e^{-\beta\hat{K}} \quad . \end{aligned} \tag{4.125}$$

Thus, the quantum mechanical *grand canonical density matrix* is given by

$$\hat{\varrho} = \frac{e^{-\beta\hat{K}}}{\mathsf{Tr}\; e^{-\beta\hat{K}}} \quad , \tag{4.126}$$

with $\hat{K} = \hat{H} - \mu\hat{N}$. Note that $\left[\hat{\varrho}, \hat{K}\right] = 0$. The quantity $\Xi(T,V,\mu)$ is called the *grand partition function*. It stands in relation to a corresponding free energy in the usual way:

$$\Omega(T,V,\mu) = -k_{\rm B}T\log\Xi(T,V,\mu) \qquad \Longleftrightarrow \qquad \Xi = \exp(-\Omega/k_{\rm B}T) \quad , \tag{4.127}$$

where $\Omega(T,V,\mu)$ is the *grand potential*, also known as the *Landau free energy*. The dimensionless quantity $z \equiv \exp(\mu/k_{\rm B}T)$ is called the *fugacity*. Thus, $\log P_n = (\Omega - E_n + \mu N_n)/k_{\rm B}T$.

If $\left[\hat{H}, \hat{N}\right] = 0$, the grand potential may be expressed as a sum over contributions from each $N$ sector, *viz.*

$$\Xi(T,V,\mu) = \sum_N e^{\beta\mu N}\, Z(T,V,N) \quad . \tag{4.128}$$

When there is more than one species, we have several chemical potentials $\{\mu_a\}$, and accordingly we define

$$\hat{K} = \hat{H} - \sum_a \mu_a \hat{N}_a \quad , \tag{4.129}$$

with $\Xi = \mathsf{Tr}\; e^{-\beta\hat{K}}$ as before. To compute averages within the GCE, we use the grand canonical density matrix $\hat{\varrho}$:

$$\left\langle \hat{A} \right\rangle = \mathsf{Tr}\left(\hat{\varrho}\,\hat{A}\right) = \frac{\sum_n \langle n|\hat{A}|n\rangle\, e^{-\beta(E_n-\mu N_n)}}{\sum_j e^{-\beta(E_j-\mu N_j)}} \quad , \tag{4.130}$$

### 4.5.2 Entropy and Gibbs-Duhem relation

In the GCE, the Boltzmann entropy is

$$\begin{aligned} S = -k_{\rm B}\sum_n P_n \log P_n &= -k_{\rm B}\sum_n P_n\left(\beta\Omega - \beta E_n + \beta\mu N_n\right) \\ &= -\frac{\Omega}{T} + \frac{\langle\hat{H}\rangle}{T} - \frac{\mu\,\langle\hat{N}\rangle}{T} \quad , \end{aligned} \tag{4.131}$$

which says

$$\Omega = E - TS - \mu N \quad , \tag{4.132}$$

where

$$E = \sum_n E_n \, P_n = \mathsf{Tr}\big(\hat{\varrho}\, \hat{H}\big) \qquad , \qquad N = \sum_n N_n \, P_n = \mathsf{Tr}\big(\hat{\varrho}\, \hat{N}\big) \quad . \tag{4.133}$$

Therefore, $\Omega(T, V, \mu)$ is a double Legendre transform of $E(S, V, N)$, with

$$d\Omega = -S\, dT - p\, dV - N d\mu \quad , \tag{4.134}$$

which entails

$$S = -\left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} \qquad , \qquad p = -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} \qquad , \qquad N = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} \quad . \tag{4.135}$$

Since $\Omega(T, V, \mu)$ is an extensive quantity, we must be able to write $\Omega = V\omega(T, \mu)$. We identify the function $\omega(T, \mu)$ as the negative of the pressure:

$$\frac{\partial \Omega}{\partial V} = -\frac{k_{\rm B} T}{\Xi}\left(\frac{\partial \Xi}{\partial V}\right)_{T,\mu} = \frac{1}{\Xi}\sum_n \frac{\partial E_n}{\partial V}\, e^{-\beta(E_n - \mu N_n)} = \left(\frac{\partial E}{\partial V}\right)_{T,\mu} = -p(T, \mu) \quad . \tag{4.136}$$

Therefore, $\Omega = -pV$, and $p = p(T, \mu)$ is an *equation of state*. This is consistent with the result from thermodynamics that $G = E - TS + pV = \mu N$. Taking the differential, we recover the Gibbs-Duhem relation,

$$d\Omega = -S\, dT - p\, dV - N d\mu = -p\, dV - V dp \quad \Rightarrow \quad S\, dT - V dp + N d\mu = 0 \quad . \tag{4.137}$$

### 4.5.3 Fluctuations in the GCE

Both energy and particle number fluctuate in the GCE. Let us compute the fluctuations in particle number. We have

$$N = \langle\, \hat{N}\, \rangle = \frac{\mathsf{Tr}\ \hat{N}\, e^{-\beta(\hat{H} - \mu \hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H} - \mu \hat{N})}} = \frac{1}{\beta}\frac{\partial}{\partial \mu}\log \Xi = -\frac{\partial \Omega}{\partial \mu} \quad . \tag{4.138}$$

Therefore,

$$\frac{1}{\beta}\frac{\partial N}{\partial \mu} = -\frac{1}{\beta}\frac{\partial^2 \Omega}{\partial \mu^2} = \frac{\mathsf{Tr}\ \hat{N}^2\, e^{-\beta(\hat{H} - \mu \hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H} - \mu \hat{N})}} - \left(\frac{\mathsf{Tr}\ \hat{N}\, e^{-\beta(\hat{H} - \mu \hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H} - \mu \hat{N})}}\right)^2 = \big\langle \hat{N}^2 \big\rangle - \big\langle \hat{N} \big\rangle^2 \quad , \tag{4.139}$$

and thus

$$\frac{\big\langle \hat{N}^2 \big\rangle - \big\langle \hat{N} \big\rangle^2}{\big\langle \hat{N} \big\rangle^2} = \frac{k_{\rm B} T}{N^2}\left(\frac{\partial N}{\partial \mu}\right)_{T,V} = \frac{k_{\rm B} T}{V}\, \kappa_T \quad , \tag{4.140}$$

where $\kappa_T$ is the isothermal compressibility. Note:

$$\begin{aligned}\left(\frac{\partial N}{\partial \mu}\right)_{T,V} &= \frac{\partial(N,T,V)}{\partial(\mu,T,V)} = -\frac{\partial(N,T,V)}{\partial(V,T,\mu)} \\ &= -\frac{\partial(N,T,V)}{\partial(N,T,p)} \cdot \overbrace{\frac{\partial(N,T,p)}{\partial(V,T,p)}}^{N/V} \cdot \overbrace{\frac{\partial(V,T,p)}{\partial(N,T,\mu)}}^{1} \cdot \overbrace{\frac{\partial(N,T,\mu)}{\partial(V,T,\mu)}}^{N/V} \\ &= -\frac{N^2}{V^2}\left(\frac{\partial V}{\partial p}\right)_{T,N} = \frac{N^2}{V}\,\kappa_T \quad .\end{aligned} \tag{4.141}$$

We thus arrive at the result

$$\frac{(\Delta N)_{\rm RMS}}{N} = \sqrt{\frac{k_{\rm B}T\,\kappa_T}{V}} \quad , \tag{4.142}$$

which again scales as $V^{-1/2}$. For the nonrelativistic ideal gas $\kappa_T = 1/p$ and the ratio is $N^{-1/2}$. Compare with the OCE result in eqn. 4.101.

### 4.5.4 Generalized susceptibilities in the GCE

We can appropriate the results from §4.4.9 and apply them, *mutatis mutandis*, to the GCE. Suppose we have a family of observables $\{\hat{Q}_k\}$ satisfying $\big[\hat{Q}_k\,,\,\hat{Q}_{k'}\big] = 0$ and $\big[\hat{H}_0\,,\,\hat{Q}_k\big] = 0$ and $\big[\hat{N}_a\,,\,\hat{Q}_k\big] = 0$ for all $k$, $k'$, and $a$. We define the grand canonical Hamiltonian,

$$\hat{K} = \hat{H}_0 - \sum_a \mu_a\,\hat{N}_a - \sum_k \lambda_k\,\hat{Q}_k \quad . \tag{4.143}$$

We then have

$$\Omega\big(T,V,\{\mu_a\},\{\lambda_k\}\big) = -k_{\rm B}T\log\mathsf{Tr}\;e^{-\hat{K}/k_{\rm B}T} \quad , \tag{4.144}$$

whence

$$d\Omega = -S\,dT - p\,dV - \sum_a N_a\,d\mu_a - \sum_k Q_k\,d\lambda_k \quad , \tag{4.145}$$

where the un-hatted quantities $\{N_a, Q_k\}$ are statistical averages within the GCE, *viz.*

$$N_a = \langle\hat{N}_a\rangle = -\frac{\partial\Omega}{\partial\mu_a} \qquad , \qquad Q_k = \langle\hat{Q}_k\rangle = -\frac{\partial\Omega}{\partial\lambda_k} \quad . \tag{4.146}$$

This leads to various generalized susceptibilities and cross-susceptibilities,

$$\frac{\partial N_a}{\partial\mu_b} = -\frac{\partial^2\Omega}{\partial\mu_a\partial\mu_b} \qquad , \qquad \frac{\partial Q_k}{\partial\lambda_l} = -\frac{\partial^2\Omega}{\partial\lambda_k\partial\lambda_l} \qquad , \qquad \frac{\partial N_a}{\partial\lambda_k} = \frac{\partial Q_k}{\partial\mu_a} = -\frac{\partial^2\Omega}{\partial\mu_a\partial\lambda_k} \quad . \tag{4.147}$$

Note that the mixed second derivatives are independent of order and yield various Maxwell relations.

### 4.5.5 Gibbs ensemble

Let the system's particle number $N$ be fixed, but let it exchange energy and volume with the world $W$. *Mutatis mutandis*, we have

$$P_n = \lim_{\Delta E\to 0}\lim_{\Delta V\to 0}\frac{D_{\rm W}(E_{\rm U}-E_n\,,\,V_{\rm U}-V_n)\,\Delta E\,\Delta V}{D_{\rm U}(E_{\rm U},V_{\rm U})\,\Delta E\,\Delta V}. \tag{4.148}$$

Then

$$\begin{aligned}\log P_n &= \log D_{\rm W}(E_{\rm U}-E_n\,,\,V_{\rm U}-V_n) - \log D_{\rm U}(E_{\rm U},V_{\rm U}) \\ &= \log D_{\rm W}(E_{\rm U},V_{\rm U}) - \log D_{\rm U}(E_{\rm U},V_{\rm U}) - E_n\,\frac{\partial\log D_{\rm W}(E,V)}{\partial E}\Big|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} - V_n\,\frac{\partial\log D_{\rm W}(E,V)}{\partial V}\Big|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} + \ldots,\end{aligned} \tag{4.149}$$

and thus $\log P_n = -\alpha-\beta E_n-\beta p\,V_n$, where the constants $\beta$ and $p$ are given by

$$\beta = \frac{\partial\log D_{\rm W}(E,V)}{\partial E}\Big|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} = \frac{1}{k_{\rm B}T}\quad,\qquad p = k_{\rm B}T\,\frac{\partial\log D_{\rm W}(E,V)}{\partial V}\Big|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}}\quad. \tag{4.150}$$

The corresponding partition function is

$$Y(T,p,N) = \mathsf{Tr}\;e^{-\beta(\hat H+pV)} = \frac{1}{V_0}\int_0^\infty dV\,e^{-\beta pV}\,Z(T,V,N) \equiv e^{-\beta G(T,p,N)}\quad, \tag{4.151}$$

where $V_0$ is a constant which has dimensions of volume. The factor $V_0^{-1}$ in front of the integral renders $Y$ dimensionless. Note that $G(V_0') = G(V_0) + k_{\rm B}T\log(V_0'/V_0)$, so the difference is not extensive and can be neglected in the thermodynamic limit. In other words, it doesn't matter what constant we choose for $V_0$ since it contributes subextensively to $G$. Moreover, in computing averages, the constant $V_0$ divides out in the ratio of numerator and denominator. Like the Helmholtz free energy, the Gibbs free energy $G(T,p,N)$ is also a double Legendre transform of the energy $E(S,V,N)$, *viz.*

$$\begin{aligned}G &= E - TS + pV \\ dG &= -S\,dT + V\,dp + \mu\,dN\quad,\end{aligned} \tag{4.152}$$

which entails

$$S = -\left(\frac{\partial G}{\partial T}\right)_{p,N}\quad,\qquad V = +\left(\frac{\partial G}{\partial p}\right)_{T,N}\quad,\qquad \mu = +\left(\frac{\partial G}{\partial N}\right)_{T,p}\quad. \tag{4.153}$$

## 4.6 Statistical Ensembles from Maximum Entropy

The basic principle: maximize the Boltzmann entropy,

$$S = -k_{\rm B}\sum_n P_n\log P_n\quad, \tag{4.154}$$

subject to a set of *constraints*. Constrained extremization using Lagrange's method of undetermined multipliers is reviewed in the appendix, §4.10.

### 4.6.1 Microcanonical ensemble

We maximize $S$ subject to the single constraint

$$C = \sum_n P_n - 1 = 0 \quad . \tag{4.155}$$

We implement the constraint $C = 0$ with a Lagrange multiplier, $\bar{\lambda} \equiv k_{\rm B}\,\lambda$, writing

$$S^* = S - k_{\rm B}\,\lambda\,C \quad , \tag{4.156}$$

and freely extremizing over the distribution $\{P_n\}$ and the Lagrange multiplier $\lambda$. Thus,

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B}\lambda\,\delta C - k_{\rm B}\,C\,\delta\lambda \\ &= -k_{\rm B}\sum_n \Big[\log P_n + 1 + \lambda\Big]\delta P_n - k_{\rm B}\,C\,\delta\lambda \equiv 0 \quad . \end{aligned} \tag{4.157}$$

We conclude that $C = 0$ and that

$$\log P_n = -\big(1 + \lambda\big) \quad , \tag{4.158}$$

and we fix $\lambda$ by the normalization condition $\sum_n P_n = 1$. This gives $P_n = 1/\Omega$, with

$$\Omega = \sum_n \Theta(E + \Delta E - E_n)\,\Theta(E_n - E) \quad , \tag{4.159}$$

*i.e.* the total number of energy states lying in the interval $[E, E + \Delta E]$.

### 4.6.2 Ordinary canonical ensemble

We maximize $S$ subject to the two constraints

$$C_1 = \sum_n P_n - 1 = 0 \qquad , \qquad C_2 = \sum_n E_n\,P_n - E = 0 \quad . \tag{4.160}$$

We now have two Lagrange multipliers. We write

$$S^* = S - k_{\rm B}\sum_{j=1}^{2}\lambda_j\,C_j \quad , \tag{4.161}$$

and we freely extremize over $\{P_n\}$ and $\{\lambda_j\}$. We therefore have

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B}\sum_n \big(\lambda_1 + \lambda_2\,E_n\big)\,\delta P_n - k_{\rm B}\sum_{j=1}^{2} C_j\,\delta\lambda_j \\ &= -k_{\rm B}\sum_n \Big[\log P_n + 1 + \lambda_1 + \lambda_2\,E_n\Big]\delta P_n - k_{\rm B}\sum_{j=1}^{2} C_j\,\delta\lambda_j \equiv 0 \quad . \end{aligned} \tag{4.162}$$

Thus, $C_1 = C_2 = 0$ and

$$\log P_n = -\big(1 + \lambda_1 + \lambda_2\,E_n\big) \quad . \tag{4.163}$$

We define $\lambda_2 \equiv \beta$ and we fix $\lambda_1$ by normalization. This yields

$$P_n = \frac{1}{Z}\,e^{-\beta E_n} \qquad , \qquad Z = \sum_n e^{-\beta E_n} = \mathsf{Tr}\;e^{-\beta\hat{H}} \quad . \tag{4.164}$$

### 4.6.3 Grand canonical ensemble

We maximize $S$ subject to the three constraints

$$C_1 = \sum_n P_n - 1 = 0 \quad , \quad C_2 = \sum_n E_n\,P_n - E = 0 \quad , \quad C_3 = \sum_n N_n\,P_n - N = 0 \quad . \tag{4.165}$$

We now have three Lagrange multipliers. We write

$$S^* = S - k_{\rm B} \sum_{j=1}^{3} \lambda_j\,C_j \quad , \tag{4.166}$$

and hence

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B} \sum_n \big(\lambda_1 + \lambda_2\,E_n + \lambda_3\,N_n\big)\,\delta P_n - k_{\rm B} \sum_{j=1}^{3} C_j\,\delta\lambda_j \\ &= -k_{\rm B} \sum_n \Big[\log P_n + 1 + \lambda_1 + \lambda_2\,E_n + \lambda_3\,N_n\Big]\delta P_n - k_{\rm B} \sum_{j=1}^{3} C_j\,\delta\lambda_j \equiv 0 \quad . \end{aligned} \tag{4.167}$$

Thus, $C_1 = C_2 = C_3 = 0$ and

$$\log P_n = -\big(1 + \lambda_1 + \lambda_2\,E_n + \lambda_3\,N_n\big) \quad . \tag{4.168}$$

We define $\lambda_2 \equiv \beta$ and $\lambda_3 \equiv -\beta\mu$, and we fix $\lambda_1$ by normalization. This yields

$$P_n = \frac{1}{\Xi}\,e^{-\beta(E_n - \mu N_n)} \qquad , \qquad \Xi = \sum_n e^{-\beta(E_n - \mu N_n)} = \mathsf{Tr}\; e^{-\beta(\hat{H} - \mu\hat{N})} \quad . \tag{4.169}$$

### 4.6.4 Generalized Gibbs ensembles

Suppose we have $\mathsf{p}$ constraints of the form $\langle\hat{Q}_k\rangle = Q_k$ with $k \in \{1, \ldots, \mathsf{p}\}$ and $\mathsf{c}$ constraints of the form $\langle\hat{N}_a\rangle = N_a$ with $a \in \{1, \ldots, \mathsf{c}\}$, in addition to the constraint of overall normalization of the probability distribution $P_n$. Then construct the extended function $S^*$, with

$$\begin{aligned} \frac{1}{k_{\rm B}}\,S^*(P_1, \ldots, P_\Omega\,,\,\lambda\,,\,\eta_1\,,\,\ldots,\eta_{\mathsf{c}}\,,\,\theta_1, \ldots, \theta_{\mathsf{p}}) = &-\sum_n P_n \log P_n - \lambda\Big(\sum_n P_n - 1\Big) \\ &- \sum_{k=1}^{\mathsf{p}} \theta_k \Big(\sum_n P_n\,Q_{k,n} - Q_k\Big) - \sum_{a=1}^{\mathsf{c}} \eta_a \Big(\sum_n P_n\,N_{a,n} - N_a\Big) \quad , \end{aligned} \tag{4.170}$$

where $\lambda$, $\{\theta_k\}$ and $\{\eta_a\}$ comprise $1+\mathsf{p}+\mathsf{c}$ undetermined Lagrange multipliers. The total number of states is taken to be $\Omega$, *i.e.* $n \in \{1, \ldots, \Omega\}$. Setting the variation $\delta S^* = 0$, we obtain the following $\Omega + 1 + \mathsf{p} + \mathsf{c}$

equations:

$$\begin{aligned}
0 &= \frac{1}{k_{\rm B}}\frac{\partial S^*}{\partial P_n} = -\log P_n - (1+\lambda) - \sum_{k=1}^{\mathsf{p}} \theta_k\, Q_{k,n} - \sum_{a=1}^{\mathsf{c}} \eta_a\, N_{a,n} \\
0 &= \frac{1}{k_{\rm B}}\frac{\partial S^*}{\partial \lambda} = 1 - \sum_n P_n \\
0 &= \frac{1}{k_{\rm B}}\frac{\partial S^*}{\partial \theta_k} = Q_k - \sum_n P_n\, Q_{k,n} \\
0 &= \frac{1}{k_{\rm B}}\frac{\partial S^*}{\partial \eta_a} = N_a - \sum_n P_n\, N_{a,n} \quad .
\end{aligned} \tag{4.171}$$

Thus, the probability distribution is given by

$$P_n(\boldsymbol{\theta},\boldsymbol{\eta}) = \frac{1}{\mathcal{Z}(\boldsymbol{\theta},\boldsymbol{\eta})}\,\exp\left\{-\sum_{k=1}^{\mathsf{p}}\theta_k\,Q_{k,n} - \sum_{a=1}^{\mathsf{c}}\eta_a\,N_{a,n}\right\} \tag{4.172}$$

with

$$\mathcal{Z}(\boldsymbol{\theta},\boldsymbol{\eta}) = \sum_n \exp\left\{-\sum_{k=1}^{\mathsf{p}}\theta_k\,Q_{k,n} - \sum_{a=1}^{\mathsf{c}}\eta_a\,N_{a,n}\right\} \tag{4.173}$$

as well as

$$Q_k = \sum_n P_n(\boldsymbol{\theta},\boldsymbol{\eta})\,Q_{k,n} \qquad , \qquad N_a = \sum_n P_n(\boldsymbol{\theta},\boldsymbol{\eta})\,N_{a,n} \quad , \tag{4.174}$$

where $Q_{k,n} = \langle\, n \,|\, \hat{Q}_k \,|\, n \,\rangle$ and $N_{a,n} = \langle\, n \,|\, \hat{N}_a \,|\, n \,\rangle$.

If for $k = 1$ we have $\hat{Q}_{k=1} = \hat{H}$, then $Q_{1,n} = E_n$ and we may define $\theta_1 \equiv \beta = 1/k_{\rm B}T$ and $\theta_k \equiv -\beta\lambda_k$ for $k \geq 2$. Similarly we may define $\eta_a \equiv -\beta\mu_a$. Further defining the generalized grand canonical Hamiltonian as

$$\hat{K}(\boldsymbol{\lambda},\boldsymbol{\mu}) \equiv \hat{H} - \sum_{k=2}^{\mathsf{p}}\lambda_k\,\hat{Q}_k - \sum_{a=1}^{\mathsf{c}}\mu_a\,\hat{N}_a \quad , \tag{4.175}$$

we have

$$\mathcal{Z}(T,\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathsf{Tr}\; e^{-\beta\hat{K}} = \sum_n e^{-K_n/k_{\rm B}T} \tag{4.176}$$

and

$$P_n(T,\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathcal{Z}^{-1}\,e^{-K_n/k_{\rm B}T} \qquad , \qquad K_n = \langle\, n \,|\, \hat{K} \,|\, n \,\rangle \quad . \tag{4.177}$$

Note the correspondence of these results with those of §4.5.4.

## 4.7 Ideal Gas Statistical Mechanics

The ordinary canonical partition function for the ideal gas was computed in eqn. 4.14. We found

$$\begin{aligned} Z(T,V,N) &= \frac{1}{N!}\prod_{i=1}^{N}\int\frac{d^dx_i\,d^dp_i}{(2\pi\hbar)^d}\,e^{-\beta\boldsymbol{p}_i^2/2m} \\ &= \frac{V^N}{N!}\left(\int_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}\,e^{-\beta p^2/2m}\right)^{Nd} = \frac{1}{N!}\left(\frac{V}{\lambda_T^d}\right)^N \quad , \end{aligned} \tag{4.178}$$

where $\lambda_T$ is the *thermal wavelength*,

$$\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T} \quad . \tag{4.179}$$

The physical interpretation of $\lambda_T$ is that it is proportional to the de Broglie wavelength for a particle of mass $m$ which has a kinetic energy of $k_{\rm B}T$.

In the GCE, we have

$$\begin{aligned} \Xi(T,V,\mu) &= \sum_{N=0}^{\infty} e^{\beta\mu N}\,Z(T,V,N) \\ &= \sum_{N=0}^{\infty}\frac{1}{N!}\left(\frac{Ve^{\mu/k_{\rm B}T}}{\lambda_T^d}\right)^N = \exp\left(\frac{Ve^{\mu/k_{\rm B}T}}{\lambda_T^d}\right) \quad . \end{aligned} \tag{4.180}$$

From $\Xi = e^{-\Omega/k_{\rm B}T}$, we have the grand potential is

$$\Omega(T,V,\mu) = -Vk_{\rm B}T\,e^{\mu/k_{\rm B}T}/\lambda_T^d \quad . \tag{4.181}$$

Since $\Omega = -pV$ (see §4.5.2), we have

$$p(T,\mu) = k_{\rm B}T\,\lambda_T^{-d}\,e^{\mu/k_{\rm B}T} \quad . \tag{4.182}$$

The number density can also be calculated:

$$n = \frac{N}{V} = -\frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = \lambda_T^{-d}\,e^{\mu/k_{\rm B}T} \quad . \tag{4.183}$$

Combined, the last two equations recapitulate the ideal gas law, $pV = Nk_{\rm B}T$.

### 4.7.1 Maxwell velocity distribution

The distribution function for momenta is given by

$$g(\boldsymbol{p}) = \left\langle\frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{p}_i - \boldsymbol{p})\right\rangle \quad . \tag{4.184}$$

Note that $g(\boldsymbol{p}) = \big\langle \delta(\boldsymbol{p}_i - \boldsymbol{p}) \big\rangle$ is the same for every particle, independent of its label $i$. We compute the average $\langle A \rangle = \mathrm{Tr}\,\big(A e^{-\beta\hat{H}}\big) / \mathrm{Tr}\; e^{-\beta\hat{H}}$. Setting $i = 1$, all the integrals other than that over $\boldsymbol{p}_1$ divide out between numerator and denominator. We then have

$$g(\boldsymbol{p}) = \frac{\int d^3p_1\; \delta(\boldsymbol{p}_1 - \boldsymbol{p})\, e^{-\beta \boldsymbol{p}_1^2/2m}}{\int d^3p_1\; e^{-\beta \boldsymbol{p}_1^2/2m}} = (2\pi m k_{\rm B} T)^{-3/2}\, e^{-\beta \boldsymbol{p}^2/2m} \quad . \tag{4.185}$$

Textbooks commonly refer to the *velocity distribution* $f(\boldsymbol{v})$, which is related to $g(\boldsymbol{p})$ by

$$f(\boldsymbol{v})\, d^3v = g(\boldsymbol{p})\, d^3p \quad . \tag{4.186}$$

Hence,

$$f(\boldsymbol{v}) = \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3/2} e^{-m\boldsymbol{v}^2/2k_{\rm B}T} \quad . \tag{4.187}$$

This is known as the *Maxwell velocity distribution*. Note that the distributions are normalized, *viz.*

$$\int d^3p\; g(\boldsymbol{p}) = \int d^3v\; f(\boldsymbol{v}) = 1 \quad . \tag{4.188}$$

If we are only interested in averaging functions of $v = |\boldsymbol{v}|$ which are isotropic, then we can define the *Maxwell speed distribution*, $\tilde{f}(v)$, as

$$\tilde{f}(v) = 4\pi\, v^2 f(\boldsymbol{v}) = 4\pi \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3/2} v^2\, e^{-mv^2/2k_{\rm B}T} \quad . \tag{4.189}$$

Note that $\tilde{f}(v)$ is normalized according to $\int\limits_0^\infty dv\, \tilde{f}(v) = 1$. It is convenient to represent $v$ in units of $v_0 = \sqrt{k_{\rm B}T/m}$, in which case

$$\tilde{f}(v) = \frac{1}{v_0}\, \varphi(v/v_0) \qquad , \qquad \varphi(s) = \sqrt{\tfrac{2}{\pi}}\, s^2\, e^{-s^2/2} \quad . \tag{4.190}$$

The distribution $\varphi(s)$ is shown in fig. 4.6. Computing averages, we have

$$C_k \equiv \langle s^k \rangle = \int\limits_0^\infty ds\; s^k\, \varphi(s) = \frac{2}{\sqrt{\pi}} \cdot 2^{k/2}\, \Gamma\left(\tfrac{3}{2} + \tfrac{k}{2}\right) \quad . \tag{4.191}$$

Thus, $C_0 = 1$, $C_1 = \sqrt{8/\pi}$, $C_2 = 3$, *etc.* The speed averages are

$$\big\langle v^k \big\rangle = C_k \left(\frac{k_{\rm B}T}{m}\right)^{k/2} \quad . \tag{4.192}$$

Note that the average *velocity* is $\langle \boldsymbol{v} \rangle = 0$, but the average *speed* is $\langle v \rangle = \sqrt{8k_{\rm B}T/\pi m}$. The speed distribution is plotted in fig. 4.6.

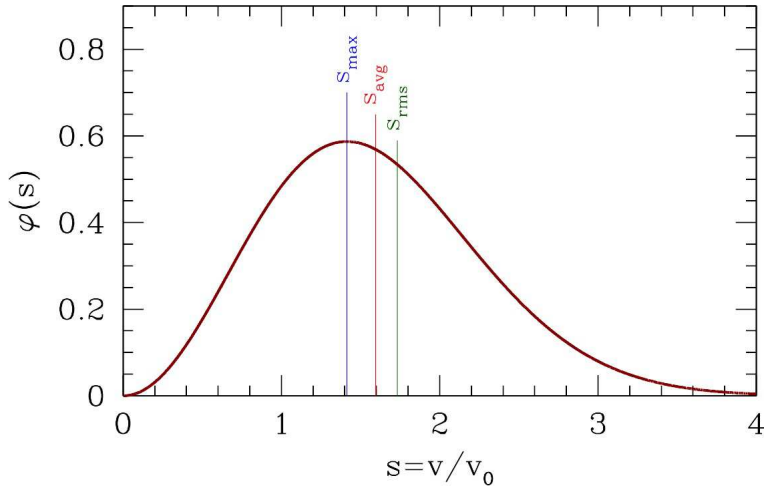



Figure 4.6: Maxwell distribution of speeds $\varphi(v/v_0)$. The most probable speed is $v_{\rm max} = \sqrt{2}\,v_0$. The average speed is $v_{\rm avg} = \sqrt{8/\pi}\,v_0$. The RMS speed is $v_{\rm rms} = \sqrt{3}\,v_0$.

### 4.7.2 Equipartition

The Hamiltonian for ballistic (*i.e.* massive nonrelativistic) particles is quadratic in the individual components of each momentum $\boldsymbol{p}_i$. There are other cases in which a classical degree of freedom appears quadratically in $\hat{H}$ as well. For example, an individual normal mode $\xi$ of a system of coupled oscillators has the Lagrangian $L = \frac{1}{2}\dot{\xi}^2 - \frac{1}{2}\,\omega_0^2\,\xi^2$, where the dimensions of $\xi$ are $[\xi] = M^{1/2}L$ by convention. The Hamiltonian for this normal mode is then $\hat{H} = \frac{1}{2}p^2 + \frac{1}{2}\,\omega_0^2\,\xi^2$, from which we see that both the kinetic as well as potential energy terms enter quadratically into the Hamiltonian. The classical rotational kinetic energy is also quadratic in the angular momentum components.

Let us compute the contribution of a single quadratic degree of freedom in $\hat{H}$ to the partition function. We'll call this degree of freedom $u$ – it may be a position or momentum or an angle or an angular momentum or some other generalized coordinate or conjugate momentum – and we'll write its contribution to $\hat{H}$ as $\hat{H}_u = \frac{1}{2}Ku^2$, where $K$ is some constant. Integrating over $u$ yields the following factor in the partition function:

$$\zeta(\beta) \equiv \int\limits_{-\infty}^{\infty} du\, e^{-\beta K u^2/2} = \left(\frac{2\pi}{K\beta}\right)^{1/2} \quad , \tag{4.193}$$

where $\beta = 1/k_{\rm B}T$. The contribution to the Helmholtz free energy is then

$$\Delta F_u = -k_{\rm B}T\log\zeta = \tfrac{1}{2}k_{\rm B}T\log\left(\frac{K}{2\pi k_{\rm B}T}\right) \quad , \tag{4.194}$$

and therefore the contribution to the internal energy $E$ is

$$\Delta E_u = -\frac{\partial\log\zeta}{\partial\beta} = \tfrac{1}{2}k_{\rm B}T \quad . \tag{4.195}$$

We have thus derived what is commonly called the *equipartition theorem* of classical statistical mechanics:

*To each degree of freedom which enters the Hamiltonian quadratically is associated a contribution $\frac{1}{2}k_{\rm B}T$ to the internal energy of the system. This results in a concomitant contribution of $\frac{1}{2}k_{\rm B}$ to the heat capacity.*

We now see why the internal energy of a classical ideal gas with $f$ degrees of freedom per molecule is $E = \frac{1}{2}fNk_{\rm B}T$, and $C_V = \frac{1}{2}fNk_{\rm B}$. This result also has applications in the theory of solids. The atoms in a solid possess kinetic energy due to their motion, and potential energy due to the spring-like interatomic potentials which tend to keep the atoms in their preferred crystalline positions. Thus, for a three-dimensional crystal, there are *six* quadratic degrees of freedom (three positions and three momenta) per atom, and the classical energy should be $E = 3Nk_{\rm B}T$, and the heat capacity $C_V = 3Nk_{\rm B}$. As we shall see, quantum mechanics modifies this result considerably at temperatures below the highest normal mode (*i.e.* phonon) frequency, but the high temperature limit is given by the classical value $C_V = 3\nu R$ (where $\nu = N/N_{\rm A}$ is the number of moles) derived here, known as the *Dulong-Petit limit*.

For a degree of freedom which enters the Hamiltonian as a power, such as $\hat{H}_u = \frac{1}{2}K|u|^\sigma$ we have

$$\zeta(\beta) = 2\int_0^\infty du\, e^{-\beta K u^\sigma/2} = \frac{2^{1+\sigma^{-1}}\,\Gamma(1/\sigma)}{\sigma K^{1/\sigma}}\,\beta^{-1/\sigma} \tag{4.196}$$

after substituting $u(t) = (2t/\beta K)^{1/\sigma}$ and integrating over $t$. We then obtain $\Delta E_u = k_{\rm B}T/\sigma$.

## 4.8 Selected Examples

### 4.8.1 Spins in an external magnetic field

Consider a system of $N_{\rm s}$ spins, each of which can be either up ($\sigma = +1$) or down ($\sigma = -1$). The Hamiltonian for this system is

$$\hat{H} = -\mu_0 H \sum_{j=1}^{N_{\rm s}} \sigma_j \quad , \tag{4.197}$$

where now we write $\hat{H}$ for the Hamiltonian, to distinguish it from the external magnetic field $H$, and $\mu_0$ is the magnetic moment per particle. We treat this system within the ordinary canonical ensemble. The partition function is

$$Z = \sum_{\sigma_1} \cdots \sum_{\sigma_{N_{\rm s}}} e^{-\beta\hat{H}} = \zeta^{N_{\rm s}} \quad , \tag{4.198}$$

where $\zeta$ is the single particle partition function:

$$\zeta = \sum_{\sigma=\pm1} e^{\mu_0 H\sigma/k_{\rm B}T} = 2\cosh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.199}$$

The Helmholtz free energy is then

$$F(T,H,N_{\rm s}) = -k_{\rm B}T\log Z = -N_{\rm s}k_{\rm B}T\log\left[2\cosh\left(\frac{\mu_0 H}{k_{\rm B}T}\right)\right] \quad . \tag{4.200}$$

The magnetization is

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T,\,N_{\rm s}} = N_{\rm s}\mu_0 \tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.201}$$

The energy is

$$E = \frac{\partial}{\partial\beta}\left(\beta F\right) = -N_{\rm s}\mu_0 H \tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.202}$$

Hence, $E = -HM$, which we already knew, from the form of $\hat{H}$ itself.

Each spin here is independent. The probability that a given spin has polarization $\sigma$ is

$$P_\sigma = \frac{e^{\beta\mu_0 H\sigma}}{e^{\beta\mu_0 H} + e^{-\beta\mu_0 H}} \quad . \tag{4.203}$$

The total probability is unity, *i.e.* $P_+ + P_- = 1$, and the average polarization is a weighted average of $\sigma = +1$ and $\sigma = -1$ contributions:

$$\langle\sigma\rangle = P_+ - P_- = \tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.204}$$

At low temperatures $T \ll \mu_0 H/k_{\rm B}$, we have $P_\uparrow \approx 1 - e^{-2\mu_0 H/k_{\rm B}T}$. At high temperatures $T > \mu_0 H/k_{\rm B}$, the two polarizations are equally likely, and $P_\sigma \approx \frac{1}{2}\left(1 + \frac{\sigma\mu_0 H}{k_{\rm B}T}\right)$.

The *isothermal magnetic susceptibility* is defined as

$$\chi_T = \frac{1}{N}\left(\frac{\partial M}{\partial H}\right)_T = \frac{\mu_0^2}{k_{\rm B}T}\,\mathrm{sech}^2\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.205}$$

(Typically this is computed per unit volume rather than per particle.) At $H = 0$, we have $\chi_T = \mu_0^2/k_{\rm B}T$, which is known as the *Curie law*.

**Aside**

The energy $E = -HM$ here is not the same quantity we discussed in our study of thermodynamics. In fact, the thermodynamic energy for this problem vanishes! Here is why. To avoid confusion, we'll need to invoke a new symbol for the thermodynamic energy, $\mathcal{E}$. Recall that the thermodynamic energy $\mathcal{E}$ is a function of extensive quantities, meaning $\mathcal{E} = \mathcal{E}(S, M, N_{\rm s})$. It is obtained from the free energy $F(T, H, N_{\rm s})$ by a double Legendre transform:

$$\mathcal{E}(S, M, N_{\rm s}) = F(T, H, N_{\rm s}) + TS + HM \quad . \tag{4.206}$$

Now from eqn. 4.200 we derive the entropy

$$S = -\frac{\partial F}{\partial T} = N_{\rm s}k_{\rm B}\log\left[2\cosh\left(\frac{\mu_0 H}{k_{\rm B}T}\right)\right] - \frac{N_{\rm s}\mu_0 H}{T}\,\tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.207}$$

Thus, using eqns. 4.200 and 4.201, we obtain $\mathcal{E}(S, M, N_{\rm s}) = 0$.

The potential confusion here arises from our use of the expression $F(T, H, N_{\rm s})$. In thermodynamics, it is the Gibbs free energy $G(T, p, N_{\rm s})$ which is a double Legendre transform of the energy: $G = \mathcal{E} - TS + pV$. By analogy, with magnetic systems we should perhaps write $G = \mathcal{E} - TS - HM$, but in keeping with many textbooks we shall use the symbol $F$ and refer to it as the Helmholtz free energy. The quantity we've called $E$ in eqn. 4.202 is in fact $E = \mathcal{E} - HM$, which means $\mathcal{E} = 0$. The energy $\mathcal{E}(S, M, N)$ vanishes here because the spins are noninteracting.

### 4.8.2 Negative temperature (!)

Consider again a system of $N_{\rm s}$ spins, each of which can be either up $(+)$ or down $(-)$. Let $N_\sigma$ be the number of sites with spin $\sigma$, where $\sigma = \pm 1$. Clearly $N_+ + N_- = N_{\rm s}$. We now treat this system within the microcanonical ensemble.

The energy of the system is $E = -HM$, where $H$ is an external magnetic field, and $M = (N_+ - N_-)\,\mu_0$ is the total magnetization. We now compute $S(E)$ using the microcanonical ensemble, as described in §4.2.5. The number of ways of arranging the system with $N_+$ up spins is

$$\Omega = \binom{N_{\rm s}}{N_+} = \frac{N_{\rm s}!}{N_+!\,N_-!} \quad . \tag{4.208}$$

Using Stirling's expression $\log K! = K \log K - K + \mathcal{O}(\log K)$ for large $K$, we have

$$\begin{aligned} S = k_{\rm B} \log \Omega &= k_{\rm B}(N_{\rm s} \log N_{\rm s} - N_{\rm s}) - k_{\rm B}(N_+ \log N_+ - N_+) - k_{\rm B}(N_- \log N_- - N_-) \\ &= -N_{\rm s} k_{\rm B} \Big\{ x \log x + (1-x) \log(1-x) \Big\} \end{aligned} \tag{4.209}$$

in the thermodynamic limit: $N_{\rm s} \to \infty$, $N_+ \to \infty$, $x = N_+/N_{\rm s}$ constant. Now the magnetization is $M = (N_+ - N_-)\mu_0 = (2N_+ - N_{\rm s})\mu_0$ , hence if we define the maximum energy $E_0 \equiv N_{\rm s}\mu_0 H$, then

$$\frac{E}{E_0} = -\frac{M}{N_{\rm s}\mu_0} = 1 - 2x \qquad \Longrightarrow \qquad x = \frac{E_0 - E}{2E_0} \quad . \tag{4.210}$$

We therefore have

$$S(E, N_{\rm s}) = -N_{\rm s} k_{\rm B} \left[ \left( \frac{E_0 - E}{2E_0} \right) \log\left( \frac{E_0 - E}{2E_0} \right) + \left( \frac{E_0 + E}{2E_0} \right) \log\left( \frac{E_0 + E}{2E_0} \right) \right] \quad . \tag{4.211}$$

We now have

$$\frac{1}{T} = \left( \frac{\partial S}{\partial E} \right)_{N_{\rm s}} = \frac{\partial S}{\partial x}\,\frac{\partial x}{\partial E} = \frac{N_{\rm s} k_{\rm B}}{2E_0} \log\left( \frac{E_0 - E}{E_0 + E} \right) \quad . \tag{4.212}$$

We see that the temperature is positive for $-E_0 \le E < 0$ and is *negative* for $0 < E \le E_0$.

What has gone wrong? The answer is that *nothing* has gone wrong – all our calculations are perfectly correct. This system *does* exhibit the possibility of negative temperature. It is, however, unphysical in that we have neglected kinetic degrees of freedom, which result in an entropy function $S(E, N_{\rm s})$ which is an increasing function of energy. In this system, $S(E, N_{\rm s})$ achieves a maximum of $S_{\rm max} = N_{\rm s} k_{\rm B} \log 2$ at $E = 0$ (*i.e.* $x = \frac{1}{2}$), and then turns over and starts decreasing. In fact, our results are completely consistent with eqn. 4.202 : the energy $E$ is an odd function of temperature. Positive energy requires negative temperature! Another example of this peculiarity is provided in the appendix in §4.11.3.

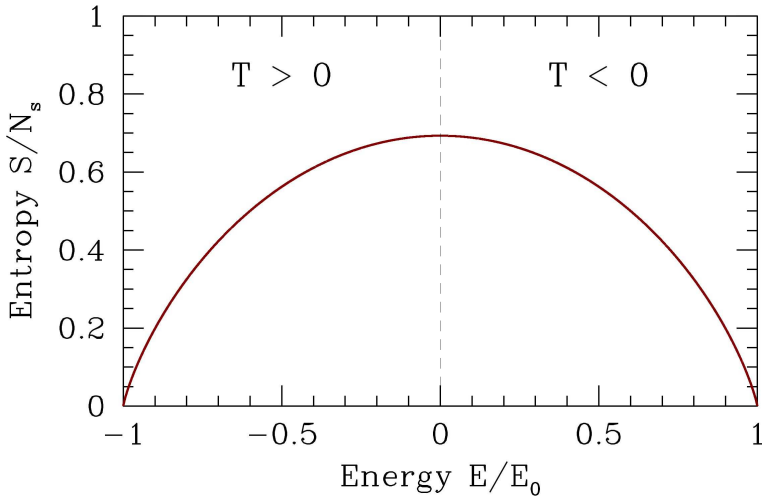



Figure 4.7: When entropy decreases with increasing energy, the temperature is negative. Typically, kinetic degrees of freedom prevent this peculiarity from manifesting in physical systems.

### 4.8.3 Adsorption

PROBLEM: A surface containing $N_{\rm s}$ adsorption sites is in equilibrium with a monatomic ideal gas. Atoms adsorbed on the surface have an energy $-\Delta$ and no kinetic energy. Each adsorption site can accommodate at most one atom. Calculate the fraction $f$ of occupied adsorption sites as a function of the gas density $n$, the temperature $T$, the binding energy $\Delta$, and physical constants.

SOLUTION: The grand partition function for the surface is

$$\Xi_{\rm surface} = e^{-\Omega_{\rm surface}/k_{\rm B}T} = \sum_{j=0}^{N_{\rm s}} \binom{N_{\rm s}}{j} e^{j(\mu+\Delta)/k_{\rm B}T} = \left(1 + e^{\mu/k_{\rm B}T}\, e^{\Delta/k_{\rm B}T}\right)^{N_{\rm s}} \quad . \tag{4.213}$$

The fraction of occupied sites is

$$f = \frac{\langle \hat{N}_{\rm surface} \rangle}{N_{\rm s}} = -\frac{1}{N_{\rm s}} \frac{\partial \Omega_{\rm surface}}{\partial \mu} = \frac{e^{\mu/k_{\rm B}T}}{e^{\mu/k_{\rm B}T} + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.214}$$

Since the surface is in equilibrium with the gas, its fugacity $z = \exp(\mu/k_{\rm B}T)$ and temperature $T$ are the same as in the gas.

For a monatomic ideal gas, the single particle partition function is $\zeta = V\lambda_T^{-3}$, where $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T}$ is the thermal wavelength. Thus, the grand partition function, for indistinguishable particles, is

$$\Xi_{\rm gas} = \exp\left(V\lambda_T^{-3}\, e^{\mu/k_{\rm B}T}\right) \quad . \tag{4.215}$$

The gas density is

$$n = \frac{\langle \hat{N}_{\rm gas} \rangle}{V} = -\frac{1}{V} \frac{\partial \Omega_{\rm gas}}{\partial \mu} = \lambda_T^{-3}\, e^{\mu/k_{\rm B}T} \quad . \tag{4.216}$$

We can now solve for the fugacity: $z = e^{\mu/k_{\rm B}T} = n\lambda_T^3$. Thus, the fraction of occupied adsorption sites is

$$f = \frac{n\lambda_T^3}{n\lambda_T^3 + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.217}$$

Interestingly, the solution for $f$ involves the constant $\hbar$.

It is always advisable to check that the solution makes sense in various limits. First of all, if the gas density tends to zero at fixed $T$ and $\Delta$, we have $f \to 0$. On the other hand, if $n \to \infty$ we have $f \to 1$, which also makes sense. At fixed $n$ and $T$, if the adsorption energy is $-\Delta \to -\infty$, then once again $f = 1$ since every adsorption site wants to be occupied. Conversely, taking $-\Delta \to +\infty$ results in $f \to 0$, since the energetic cost of adsorption is infinitely high.

### 4.8.4 Dipole gas

Consider a gas of polar diatomic molecules, each endowed with a dipole moment $\boldsymbol{d}_j = \mu_0 \hat{\boldsymbol{n}}_j$. Here $\hat{\boldsymbol{n}}_j = (\sin\theta_j \cos\phi_j, \sin\theta_j \sin\phi_j, \cos\theta_j)$ is the unit orientation vector for the symmetry axis of molecule $j$, which points in the direction of the molecule's dipole moment. The single molecule Hamiltonian is then

$$\hat{h} = \frac{\boldsymbol{p}^2}{2m} + \frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I\sin^2\theta} - \mu_0 \boldsymbol{E}\cdot\hat{\boldsymbol{n}} \quad . \tag{4.218}$$

Without loss of generality we may choose $\boldsymbol{E} = E\hat{\boldsymbol{z}}$, hence the dipole energy function is $u(\theta) = -\mu_0 E\cos\theta$. Here $\boldsymbol{p} = (p_x, p_y, p_z)$ is the vector of linear momenta and $(p_\phi, p_\theta, p_\psi)$ are the *angular* momenta conjugate to the Euler angles $\phi$, $\theta$, and $\psi$. For a symmetric top, the Euler angles $(\theta, \phi)$ are, respectively, the polar and azimuthal angles of the symmetry axis, and $\psi$ is the angle of rotation about the symmetry axis. The inertia tensor in the center-of-mass frame is then $\mathsf{diag}(I_1, I_1, I_3)$, and for diatomic molecules $I_3$ is on the order of $ma_{\rm nuc}^2$, where $a_{\rm nuc}$ is the nuclear diameter. Thus $I_3$ is so small that the lowest nonzero quantized rotation energies $\propto \hbar^2/2I_3$ are much larger than $k_{\rm B}T$ in all conceivable applications, and we may forget about the $\{\psi, p_\psi\}$ degrees of freedom and set $I_1 = I_2 \equiv I$. This results in the rotational kinetic energy term in the above molecular Hamiltonian and a phase space metric

$$d\mu = \frac{dx\,dp_x}{h}\cdot\frac{dy\,dp_y}{h}\cdot\frac{dz\,dp_z}{h}\cdot\frac{d\theta\,dp_\theta}{h}\cdot\frac{d\phi\,dp_\phi}{h} \tag{4.219}$$

and the single molecule partition function is

$$\begin{aligned}\zeta(T,V,E) &= \int d\mu\, e^{-\hat{h}/k_{\rm B}T} = \frac{V}{\lambda_T^3}\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \int_{-\infty}^{\infty}\frac{dp_\theta}{h}\int_{-\infty}^{\infty}\frac{dp_\phi}{h}\, e^{-\beta p_\theta^2/2I}\, e^{-\beta p_\phi^2/2I\sin^2\theta}\, e^{-\beta u(\theta)} \\ &= \frac{V}{\lambda_T^3}\cdot\frac{Ik_{\rm B}T}{\hbar^2}\int_0^{\pi} d\theta\,\sin\theta\, e^{-\beta u(\theta)} = \underbrace{\frac{V}{\lambda_T^3}}_{\zeta_{\rm trans}}\cdot\underbrace{\frac{Ik_{\rm B}T}{\hbar^2}}_{\zeta_{\rm rot}}\cdot\underbrace{\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}}_{\zeta_{\rm pol}} \quad ,\end{aligned} \tag{4.220}$$

where

$$\zeta_{\rm trans}(T,V) = \frac{V}{\lambda_T^3} \qquad , \qquad \zeta_{\rm rot}(T) = \frac{Ik_{\rm B}T}{\hbar^2} \qquad , \qquad \zeta_{\rm pol}(T,E) = \frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T} \tag{4.221}$$

are the contributions to the single molecule partition function from translational, rotational, and polarization energies, respectively. The $N$-molecule partition function is $Z(T,V,N,E) = \zeta(T,V,E)^N/N!$ and thus in the thermodynamic limit, using Stirling's rule,

$$\begin{aligned} F(T,V,N,E) &= -Nk_{\rm B}T\log\big[\zeta(T,V,E)/N\big] - Nk_{\rm B}T \\ &= -Nk_{\rm B}T\left\{\log\left(\frac{V}{N\lambda_T^3}\right) + \log\left(\frac{2Ik_{\rm B}T}{\hbar^2}\right) + \log\left(\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}\right) + 1\right\} \quad . \end{aligned} \tag{4.222}$$

The average dipole moment per molecule is given by

$$\langle \boldsymbol{d}\rangle = \mu_0\,\langle\hat{\boldsymbol{n}}\rangle = -\frac{1}{N}\left(\frac{\partial F}{\partial \boldsymbol{E}}\right)_{T,V,N} = \mu_0\left\{\mathrm{ctnh}\left(\frac{\mu_0 E}{k_{\rm B}T}\right) - \frac{k_{\rm B}T}{\mu_0 E}\right\}\widehat{\boldsymbol{E}} \quad , \tag{4.223}$$

where $\widehat{\boldsymbol{E}} = \hat{\boldsymbol{z}}$.

Let's work out the physics in the GCE. The grand partition function is given by

$$\begin{aligned} \Xi(T,V,\mu,E) &= \sum_{N=0}^{\infty} z^N Z(T,V,N,E) = \exp\big(z\,\zeta_{\rm trans}\,\zeta_{\rm rot}\,\zeta_{\rm pol}\big) \\ &= \exp\left\{\left(\frac{z\,V}{\lambda_T^3}\right)\left(\frac{2Ik_{\rm B}T}{\hbar^2}\right)\left(\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}\right)\right\} \quad , \end{aligned} \tag{4.224}$$

where $z = \exp(\mu/k_{\rm B}T)$ is the fugacity. Thus, the grand potential is

$$\begin{aligned} \Omega(T,V,\mu,E) &= -k_{\rm B}T\log\Xi(T,V,\mu,E) = -k_{\rm B}T\,z\,\zeta_{\rm trans}\,\zeta_{\rm rot}\,\zeta_{\rm pol} \\ &= -k_{\rm B}T\left(\frac{z\,V}{\lambda_T^3}\right)\left(\frac{2Ik_{\rm B}T}{\hbar^2}\right)\left(\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}\right) \quad . \end{aligned} \tag{4.225}$$

As a sanity check, we compute $\langle\boldsymbol{d}\rangle$ within the GCE and make sure we get the same result in the thermodynamic limit. We have the average *total* dipole moment is

$$\boldsymbol{D} = \Big\langle\sum_{j=1}^{N}\boldsymbol{d}_j\Big\rangle = -\left(\frac{\partial\Omega}{\partial\boldsymbol{E}}\right)_{T,V,\mu} = k_{\rm B}T\,z\,\zeta_{\rm trans}\,\zeta_{\rm rot}\left(\frac{\partial\zeta_{\rm pol}}{\partial\boldsymbol{E}}\right)_{T,V,\mu} \quad . \tag{4.226}$$

The average number of particles is

$$N = -\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V,\boldsymbol{E}} = z\,\zeta_{\rm trans}\,\zeta_{\rm rot}\,\zeta_{\rm pol} \tag{4.227}$$

and therefore

$$\boldsymbol{d} = \frac{\boldsymbol{D}}{N} = k_{\rm B}T\left(\frac{\partial\log\zeta_{\rm pol}}{\partial\boldsymbol{E}}\right)_{T,V,\mu} = \mu_0\left\{\mathrm{ctnh}\left(\frac{\mu_0 E}{k_{\rm B}T}\right) - \frac{k_{\rm B}T}{\mu_0 E}\right\}\widehat{\boldsymbol{E}} \quad . \tag{4.228}$$

### 4.8.5 Elasticity of wool

Wool consists of interlocking protein molecules which can stretch into an elongated configuration, but reversibly so. This feature gives wool its very useful elasticity. Let us model a chain of these proteins by assuming they can exist in one of two states, which we will call A and B, with energies $\varepsilon_{\rm A}$ and $\varepsilon_{\rm B}$ and lengths $\ell_{\rm A}$ and $\ell_{\rm B}$. The situation is depicted in fig. 4.8. We model these conformational degrees of freedom by a spin variable $\sigma = \pm 1$ for each molecule, where $\sigma = +1$ in the A state and $\sigma = -1$ in the B state. Suppose a chain consisting of $N_{\rm m}$ monomers is placed under a tension $\tau$. We then have

$$\hat{H} = \sum_{j=1}^{N_{\rm m}} \Big[ \varepsilon_{\rm A}\, \delta_{\sigma_j,+1} + \varepsilon_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad . \tag{4.229}$$

Similarly, the length is

$$\hat{L} = \sum_{j=1}^{N_{\rm m}} \Big[ \ell_{\rm A}\, \delta_{\sigma_j,+1} + \ell_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad . \tag{4.230}$$

The Gibbs partition function is $Y = {\rm Tr}\; e^{-\hat{K}/k_{\rm B}T}$, with $\hat{K} = \hat{H} - \tau \hat{L}$ :

$$\hat{K} = \sum_{j=1}^{N_{\rm m}} \Big[ \tilde{\varepsilon}_{\rm A}\, \delta_{\sigma_j,+1} + \tilde{\varepsilon}_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad , \tag{4.231}$$

where $\tilde{\varepsilon}_{\rm A} \equiv \varepsilon_{\rm A} - \tau \ell_{\rm A}$ and $\tilde{\varepsilon}_{\rm B} \equiv \varepsilon_{\rm B} - \tau \ell_{\rm B}$. At $\tau = 0$ the A state is preferred for each monomer, but when $\tau$ exceeds $\tau^*$, defined by the relation $\tilde{\varepsilon}_{\rm A} = \tilde{\varepsilon}_{\rm B}$, the B state is preferred. One finds

$$\tau^* = \frac{\varepsilon_{\rm B} - \varepsilon_{\rm A}}{\ell_{\rm B} - \ell_{\rm A}} \quad . \tag{4.232}$$

Once again, we have a set of $N_{\rm m}$ noninteracting spins. The partition function is $Y = \zeta^{N_{\rm m}}$, where $\zeta$ is the single monomer partition function, $\zeta = {\rm Tr}\; \exp(-\beta \hat{h})$, where

$$\hat{h} = \tilde{\varepsilon}_{\rm A}\, \delta_{\sigma,1} + \tilde{\varepsilon}_{\rm B}\, \delta_{\sigma,-1} \tag{4.233}$$

is the single "spin" Hamiltonian. Thus,

$$\zeta = {\rm Tr}\; e^{-\beta \hat{h}} = e^{-\beta \tilde{\varepsilon}_{\rm A}} + e^{-\beta \tilde{\varepsilon}_{\rm B}} \quad , \tag{4.234}$$

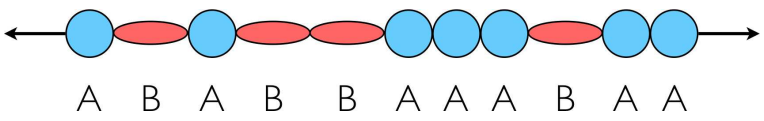



Figure 4.8: The monomers in wool are modeled as existing in one of two states. The low energy undeformed state is A, and the higher energy deformed state is B. Applying tension induces more monomers to enter the B state.

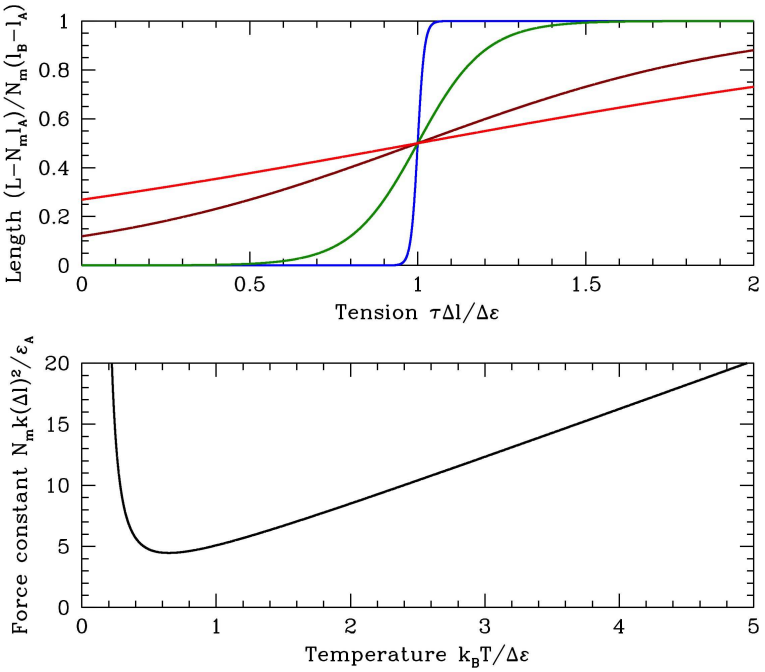



Figure 4.9: Upper panel: length $L(\tau, T)$ for $k_{\rm B}T/\tilde{\varepsilon} = 0.01$ (blue), $0.1$ (green), $0.5$ (dark red), and $1.0$ (red). Bottom panel: dimensionless force constant $N_{\rm m}(\Delta\ell)^2 k/\varepsilon_{\rm A}$ *versus* dimensionless temperature $k_{\rm B}T/\Delta\varepsilon$.

It is convenient to define the differences

$$\Delta\varepsilon = \varepsilon_{\rm B} - \varepsilon_{\rm A} \quad , \quad \Delta\ell = \ell_{\rm B} - \ell_{\rm A} \quad , \quad \Delta\tilde{\varepsilon} = \tilde{\varepsilon}_{\rm B} - \tilde{\varepsilon}_{\rm A} \tag{4.235}$$

in which case the partition function $Y$ and Gibbs free energy $G$ are

$$\begin{aligned} Y(T,\tau,N_{\rm m}) &= e^{-N_{\rm m}\beta\,\tilde{\varepsilon}_{\rm A}}\Big[1 + e^{-\beta\Delta\tilde{\varepsilon}}\Big]^{N_{\rm m}} \\ G(T,\tau,N_{\rm m}) &= N_{\rm m}\,\tilde{\varepsilon}_{\rm A} - N_{\rm m}\,k_{\rm B}T\log\Big[1 + e^{-\Delta\tilde{\varepsilon}/k_{\rm B}T}\Big] \end{aligned} \tag{4.236}$$

The average length is

$$L = \langle\hat{L}\rangle = -\left(\frac{\partial G}{\partial\tau}\right)_{T,N_{\rm m}} = N_{\rm m}\,\ell_{\rm A} + \frac{N_{\rm m}\,\Delta\ell}{e^{(\Delta\varepsilon-\tau\Delta\ell)/k_{\rm B}T}+1} \quad . \tag{4.237}$$

The polymer behaves as a spring, and for small $\tau$ the spring constant is

$$k = \left.\frac{\partial\tau}{\partial L}\right|_{\tau=0} = \frac{4k_{\rm B}T}{N_{\rm m}(\Delta\ell)^2}\,\cosh^2\!\left(\frac{\Delta\varepsilon}{2k_{\rm B}T}\right) \quad . \tag{4.238}$$

The results are shown in fig. 4.9. Note that length increases with temperature for $\tau < \tau^*$ and decreases with temperature for $\tau > \tau^*$. Note also that $k$ diverges at both low and high temperatures. At low $T$, the energy gap $\Delta\varepsilon$ dominates and all monomers are in the A state, hence $L = N_{\rm m}\,\ell_{\rm A}$. At high temperatures $k_{\rm B}T$ dominates and both A and B configurations are equally likely. The mean length is then $L = \frac{1}{2}N_{\rm m}(\ell_{\rm A} + \ell_{\rm B})$.

## 4.9 Statistical Mechanics of Molecular Gases

### 4.9.1 Separation of translational and internal degrees of freedom

The states of a noninteracting atom or molecule are labeled by its total momentum $\boldsymbol{p}$ and its internal quantum numbers, which we will simply write with a collective index $\alpha$, specifying rotational, vibrational, and electronic degrees of freedom. The single particle Hamiltonian is then

$$\hat{h} = \frac{\boldsymbol{p}^2}{2m} + \hat{h}_{\rm int} \quad , \tag{4.239}$$

with

$$\hat{h}\,|\,\boldsymbol{k}\,,\,\alpha\,\rangle = \left(\frac{\hbar^2\boldsymbol{k}^2}{2m} + \varepsilon_\alpha\right)|\,\boldsymbol{k}\,,\,\alpha\,\rangle \quad . \tag{4.240}$$

The partition function is

$$\zeta = {\rm Tr}\; e^{-\beta\hat{h}} = \sum_{\boldsymbol{p}} e^{-\beta\boldsymbol{p}^2/2m} \sum_j g_j\, e^{-\beta\varepsilon_j} \quad . \tag{4.241}$$

Here we have replaced the internal label $\alpha$ with a label $j$ of energy eigenvalues, with $g_j$ being the degeneracy of the internal state with energy $\varepsilon_j$. To do the $\boldsymbol{p}$ sum, we quantize in a box of dimensions $L_1 \times L_2 \times \cdots \times L_d$, using periodic boundary conditions. Then

$$\boldsymbol{p} = \left(\frac{2\pi\hbar n_1}{L_1}\,,\,\frac{2\pi\hbar n_2}{L_2}\,,\,\ldots\,,\,\frac{2\pi\hbar n_d}{L_d}\right) \quad , \tag{4.242}$$

where each $n_i$ is an integer. Since the differences between neighboring quantized $\boldsymbol{p}$ vectors are very tiny, we can replace the sum over $\boldsymbol{p}$ by an integral:

$$\sum_{\boldsymbol{p}} \longrightarrow \int \frac{d^dp}{\Delta p_1 \cdots \Delta p_d} \tag{4.243}$$

where the volume in momentum space of an elementary rectangle is

$$\Delta p_1 \cdots \Delta p_d = \frac{(2\pi\hbar)^d}{L_1 \cdots L_d} = \frac{(2\pi\hbar)^d}{V} \quad . \tag{4.244}$$

Thus,

$$\zeta = V\!\int \frac{d^dp}{(2\pi\hbar)^d}\, e^{-\boldsymbol{p}^2/2mk_{\rm B}T} \sum_j g_j\, e^{-\varepsilon_j/k_{\rm B}T} = V\lambda_T^{-d}\,\xi(T) \tag{4.245}$$

where

$$\xi(T) = \sum_j g_j\, e^{-\varepsilon_j/k_{\rm B}T} \tag{4.246}$$

is the *internal coordinate partition function*. The $N$-particle ordinary canonical partition function is then

$$Z_N = \frac{1}{N!}\left(\frac{V}{\lambda_T^d}\right)^N \xi^N(T) \quad . \tag{4.247}$$

Using Stirling's approximation, we find the Helmholtz free energy $F = -k_{\rm B}T \log Z$ is

$$F(T,V,N) = -Nk_{\rm B}T\left[\log\left(\frac{V}{N\lambda_T^d}\right) + 1 + \log\xi(T)\right] \quad , \tag{4.248}$$

which entails the ideal gas law $pV = Nk_{\rm B}T$ since $dF = -S\,dT - p\,dV + \mu\,dN$. Making the Legendre transform to $G(T,p,N) = F(T,V,N) + pV$, we have

$$G(T,p,N) = Nk_{\rm B}T\log\left(\frac{p\,\lambda_T^d}{k_{\rm B}T}\right) + N\varphi(T) \quad , \tag{4.249}$$

where

$$\varphi(T) = -k_{\rm B}T\log\xi(T) \tag{4.250}$$

is the internal coordinate contribution to the single particle free energy. Recall that $G = \mu N$, which is to say $\mu(T,p) = G(T,p,N)/N$.

The entropy is

$$S = -\left(\frac{\partial G}{\partial T}\right)_{p,N} = Nk_{\rm B}\left[\log\left(\frac{k_{\rm B}T}{p\lambda_T^d}\right) + 1 + \tfrac{1}{2}d\right] - N\varphi'(T) \quad , \tag{4.251}$$

and therefore the heat capacities are

$$\begin{aligned} C_{p,N} &= T\left(\frac{\partial S}{\partial T}\right)_{p,N} = \left(\tfrac{1}{2}d+1\right)Nk_{\rm B} - NT\varphi''(T) \\ C_{V,N} &= T\left(\frac{\partial S}{\partial T}\right)_{V,N} = \tfrac{1}{2}dNk_{\rm B} - NT\varphi''(T) \quad , \end{aligned} \tag{4.252}$$

and we have $C_{p,N} - C_{V,N} = Nk_{\rm B}$. Furthermore, any temperature variation in the specific heats must be due to the internal degrees of freedom.

### 4.9.2 The internal coordinate partition function

At energy scales of interest we can separate the internal degrees of freedom into distinct classes, writing

$$\hat{h}_{\rm int} = \hat{h}_{\rm rot} + \hat{h}_{\rm vib} + \hat{h}_{\rm elec} \tag{4.253}$$

as a sum over internal Hamiltonians governing rotational, vibrational, and electronic degrees of freedom. Then

$$\xi_{\rm int}(T) = \xi_{\rm rot}(T)\cdot\xi_{\rm vib}(T)\cdot\xi_{\rm elec}(T) \quad . \tag{4.254}$$

Associated with each class of excitation is a characteristic temperature $\Theta$. Rotational and vibrational temperatures of a few common molecules are listed in tab. 4.1.

| molecule | $\Theta_{\rm rot}$(K) | $\Theta_{\rm vib}$(K) |
|---|---|---|
| $H_2$ | 85.4 | 6100 |
| $N_2$ | 2.86 | 3340 |
| $H_2O$ | 13.7 , 21.0 , 39.4 | 2290 , 5180 , 5400 |

Table 4.1: Some rotational and vibrational temperatures of common molecules.

### 4.9.3 Quantum theory of rotations

Consider a class of molecules which can be approximated as axisymmetric tops. Let $\boldsymbol{L}$ be the molecular angular momentum vector with components $L^{x,y,z}$ along *space-fixed* axes. We write $\vec{\mathfrak{L}} = (\mathfrak{L}_{\rm a}, \mathfrak{L}_{\rm b}, \mathfrak{L}_{\rm c})$ as the angular momentum vector about instantaneous *body-fixed* principal axes. Thus $\mathfrak{L}_{\rm a} = \hat{\boldsymbol{n}}_{\rm a} \cdot \boldsymbol{L}$, $\mathfrak{L}_{\rm b} = \hat{\boldsymbol{n}}_{\rm b} \cdot \boldsymbol{L}$, and $\mathfrak{L}_{\rm c} = \hat{\boldsymbol{n}}_{\rm c} \cdot \boldsymbol{L}$, where $\hat{\boldsymbol{n}}_{\rm a,b,c}(t)$ are the instantaneous principal axes. We take $\hat{\boldsymbol{n}}_{\rm c}$ to be the symmetry axis of the molecule. The rotational Hamiltonian is then

$$\begin{aligned}\hat{h}_{\rm rot} &= \frac{\mathfrak{L}_{\rm a}^2 + \mathfrak{L}_{\rm b}^2}{2I_1} + \frac{\mathfrak{L}_{\rm c}^2}{2I_3} \\ &= \frac{\hbar^2 L(L+1)}{2I_1} + \left(\frac{1}{2I_3} - \frac{1}{2I_1}\right)\mathfrak{L}_c^2 \quad .\end{aligned} \tag{4.255}$$

Note that

$$\left[L^\mu\,,\,\mathfrak{L}_c\right] = n_{\rm c}^\nu\left[L^\mu\,,\,L^\nu\right] + \left[L^\mu\,,\,n_{\rm c}^\nu\right]L^\nu = i\epsilon_{\mu\nu\lambda}\,\hbar n_{\rm c}^\nu\,L^\lambda + i\epsilon_{\mu\nu\lambda}\,\hbar n_{\rm c}^\lambda\,L^\nu = 0 \quad , \tag{4.256}$$

which is equivalent to the statement that $\mathfrak{L}_{\rm c} = \hat{\boldsymbol{n}}_{\rm c} \cdot \boldsymbol{L}$ is a rotational scalar. We can therefore simultaneously specify the eigenvalues of $\{\boldsymbol{L}^2, L^z, \mathfrak{L}_{\rm c}\}$, which form a complete set of commuting observables (CSCO)[10]. The eigenvalues of $L^z$ are $m\hbar$ with $m \in \{-L, \ldots, L\}$, while those of $\mathfrak{L}_{\rm c}$ are $k\hbar$ with $k \in \{-L, \ldots, L\}$. There is a $(2L+1)$-fold degeneracy associated with the $L^z$ quantum number.

We assume the molecule is prolate, so that $I_3 < I_1$. We can the define two temperature scales,

$$\Theta = \frac{\hbar^2}{2I_1 k_{\rm B}} \qquad , \qquad \widetilde{\Theta} = \frac{\hbar^2}{2I_3 k_{\rm B}} \quad . \tag{4.257}$$

Prolateness then means $\widetilde{\Theta} > \Theta$. We conclude that the rotational partition function for an axisymmetric molecule is given by

$$\xi_{\rm rot}(T) = \sum_{L=0}^{\infty}(2L+1)\,e^{-L(L+1)\,\Theta/T}\sum_{k=-L}^{L} e^{-k^2\,(\widetilde{\Theta}-\Theta)/T} \tag{4.258}$$

[10]Note that while we cannot simultaneously specify the eigenvalues of two components of $\boldsymbol{L}$ along axes fixed in space, we can simultaneously specify the components of $\boldsymbol{L}$ along one axis fixed in space and one axis rotating with a body. See Landau and Lifshitz, *Quantum Mechanics*, §103.

In diatomic molecules, $I_3$ is extremely small, and $\widetilde{\Theta} \gg k_{\rm B}T$ at all relevant temperatures. Only the $k = 0$ term contributes to the partition sum, and we have

$$\xi_{\rm rot}(T) = \sum_{L=0}^{\infty} (2L+1)\, e^{-L(L+1)\,\Theta/T} \quad . \tag{4.259}$$

When $T \ll \Theta$, only the first few terms contribute, and

$$\xi_{\rm rot}(T) = 1 + 3\, e^{-2\Theta/T} + 5\, e^{-6\Theta/T} + \ldots \tag{4.260}$$

In the high temperature limit, we have a slowly varying summand. The *Euler-MacLaurin summation formula* may be used to evaluate such a series:

$$\sum_{k=0}^{n} F_k = \int_0^n dk\, F(k) + \tfrac{1}{2}\big[F(0) + F(n)\big] + \sum_{j=1}^{\infty} \frac{\mathsf{B}_{2j}}{(2j)!}\Big[F^{(2j-1)}(n) - F^{(2j-1)}(0)\Big] \tag{4.261}$$

where $\mathsf{B}_j$ is the $j^{\rm th}$ *Bernoulli number*:

$$\mathsf{B}_0 = 1 \quad , \quad \mathsf{B}_1 = -\tfrac{1}{2} \quad , \quad \mathsf{B}_2 = \tfrac{1}{6} \quad , \quad \mathsf{B}_4 = -\tfrac{1}{30} \quad , \quad \mathsf{B}_6 = \tfrac{1}{42} \quad . \tag{4.262}$$

Thus,

$$\sum_{k=0}^{\infty} F_k = \int_0^{\infty} dx\, F(x) + \tfrac{1}{2}F(0) - \tfrac{1}{12}F'(0) - \frac{1}{720}\,F'''(0) + \ldots \quad . \tag{4.263}$$

We have $F(x) = (2x+1)\, e^{-x(x+1)\Theta/T}$, for which $\int_0^{\infty} dx\, F(x) = T/\Theta$, hence

$$\xi_{\rm rot} = \frac{T}{\Theta} + \frac{1}{3} + \frac{1}{15}\,\frac{\Theta}{T} + \frac{4}{315}\left(\frac{\Theta}{T}\right)^2 + \ldots \quad . \tag{4.264}$$

Recall that $\varphi(T) = -k_{\rm B}T \log \xi(T)$. We conclude that $\varphi_{\rm rot}(T) \approx -3k_{\rm B}T\, e^{-2\Theta/T}$ for $T \ll \Theta$ and $\varphi_{\rm rot}(T) \approx -k_{\rm B}T \log(T/\Theta)$ for $T \gg \Theta$. We have seen that the internal coordinate contribution to the heat capacity is $\Delta C_V = -NT\varphi''(T)$. For diatomic molecules, then, this contribution is exponentially suppressed for $T \ll \Theta$, while for high temperatures we have $\Delta C_V = Nk_{\rm B}$. One says that the rotational excitations are 'frozen out' at temperatures much below $\Theta$. Including the first few terms, we have

$$\begin{aligned} \Delta C_V(T \ll \Theta) &= 12\, Nk_{\rm B}\left(\frac{\Theta}{T}\right)^2 e^{-2\Theta/T} + \ldots \\ \Delta C_V(T \gg \Theta) &= Nk_{\rm B}\left\{1 + \frac{1}{45}\left(\frac{\Theta}{T}\right)^2 + \frac{16}{945}\left(\frac{\Theta}{T}\right)^3 + \ldots\right\} \quad . \end{aligned} \tag{4.265}$$

Note that $C_V$ overshoots its limiting value of $Nk_{\rm B}$ and asymptotically approaches it from above.

Special care must be taken in the case of homonuclear diatomic molecules, for then only even or odd $L$ states are allowed, depending on the total nuclear spin. This is discussed below in §4.9.6.

For polyatomic molecules, the moments of inertia generally are large enough that the molecule's rotations can be considered classically. We then have the rotational kinetic energy

$$K(\mathfrak{L}_{\rm a},\mathfrak{L}_{\rm b},\mathfrak{L}_{\rm c}) = \frac{\mathfrak{L}_{\rm a}^2}{2I_1} + \frac{\mathfrak{L}_{\rm b}^2}{2I_2} + \frac{\mathfrak{L}_{\rm c}^2}{2I_3} \quad . \tag{4.266}$$

We then have

$$\xi_{\rm rot}(T) = \frac{1}{g_{\rm rot}} \int \frac{d\mathfrak{L}_{\rm a}\, d\mathfrak{L}_{\rm b}\, d\mathfrak{L}_{\rm c}\, d\phi\, d\theta\, d\psi}{(2\pi\hbar)^3}\, e^{-K(\mathfrak{L}_{\rm a}\,\mathfrak{L}_{\rm b}\,\mathfrak{L}_{\rm c})/k_{\rm B}T} \quad , \tag{4.267}$$

where $(\phi,\theta,\psi)$ are the Euler angles. Recall $\phi \in [0,2\pi]$, $\theta \in [0,\pi]$, and $\psi \in [0,2\pi]$. The factor $g_{\rm rot}$ accounts for physically indistinguishable orientations of the molecule brought about by rotations, which can occur when more than one of the nuclei is the same. We then have

$$\xi_{\rm rot}(T) = \frac{1}{g_{\rm rot}} \left(\frac{2k_{\rm B}T}{\hbar^2}\right)^{3/2} \sqrt{\pi I_1 I_2 I_3} \quad . \tag{4.268}$$

This leads to $\Delta C_V = \frac{3}{2}Nk_{\rm B}$.

Quantum mechanically, the angular momentum operators with respect to principal axes may be expressed as differential operators in the Euler angles, *viz.*

$$\begin{aligned}
\mathfrak{L}_1 &= -i\hbar \left\{ \cos\psi \,\mathrm{ctn}\,\theta\, \frac{\partial}{\partial\psi} + \sin\psi\, \frac{\partial}{\partial\theta} - \cos\psi \csc\theta\, \frac{\partial}{\partial\phi} \right\} \\
\mathfrak{L}_2 &= -i\hbar \left\{ -\sin\psi \,\mathrm{ctn}\,\theta\, \frac{\partial}{\partial\psi} + \cos\psi\, \frac{\partial}{\partial\theta} + \sin\psi \csc\theta\, \frac{\partial}{\partial\phi} \right\} \\
\mathfrak{L}_3 &= -i\hbar\, \frac{\partial}{\partial\psi} \quad ,
\end{aligned} \tag{4.269}$$

for which $[\mathfrak{L}_i,\mathfrak{L}_j] = i\hbar\,\epsilon_{ijk}\,\mathfrak{L}_k$. The rigid body Hamiltonian is then

$$\hat{h}_{\rm rot} = \sum_{i=1}^{3} \frac{\mathfrak{L}_i^2}{2I_i} + u(\phi,\theta,\psi) \quad , \tag{4.270}$$

where, for the sake of generality, we include a potential energy $u(\phi,\theta,\psi)$.

### 4.9.4 Vibrations

Vibrational frequencies are often given in units of inverse wavelength, such as $\mathrm{cm}^{-1}$, called a *wavenumber*. To convert to a temperature scale $T^*$, we write $k_{\rm B}T^* = h\nu = hc/\lambda$, hence $T^* = (hc/k_{\rm B})\,\lambda^{-1}$, and we multiply by

$$\frac{hc}{k_{\rm B}} = 1.436\,\mathrm{K}\cdot\mathrm{cm} \quad . \tag{4.271}$$

For example, infrared absorption ($\sim 50$ cm$^{-1}$ to $10^4$ cm$^{-1}$) reveals that the 'asymmetric stretch' mode of the $H_2O$ molecule has a vibrational frequency of $\nu = 3756\,\mathrm{cm}^{-1}$. The corresponding temperature scale is $T^* = 5394\,\mathrm{K}$.

Vibrations are normal modes of oscillations. A single normal mode Hamiltonian is of the form

$$\hat{h} = \frac{p^2}{2m} + \tfrac{1}{2}m\omega^2 q^2 = \hbar\omega\big(a^\dagger a + \tfrac{1}{2}\big) \quad . \tag{4.272}$$

In general there are many vibrational modes, hence many normal mode frequencies $\omega_\alpha$. We then must sum over all of them, resulting in

$$\xi_{\rm vib} = \prod_\alpha \xi_{\rm vib}^{(\alpha)} \quad . \tag{4.273}$$

For each such normal mode, the contribution is

$$\begin{aligned}\xi &= \sum_{n=0}^\infty e^{-(n+\frac{1}{2})\hbar\omega/k_{\rm B}T} = e^{-\hbar\omega/2k_{\rm B}T}\sum_{n=0}^\infty \Big(e^{-\hbar\omega/k_{\rm B}T}\Big)^n \\ &= \frac{e^{-\hbar\omega/2k_{\rm B}T}}{1-e^{-\hbar\omega/k_{\rm B}T}} = \frac{1}{2\sinh(\Theta/2T)} \quad ,\end{aligned} \tag{4.274}$$

where $\Theta = \hbar\omega/k_{\rm B}$. Then

$$\varphi = k_{\rm B}T\log\Big(2\sinh(\Theta/2T)\Big) = \tfrac{1}{2}k_{\rm B}\Theta + k_{\rm B}T\log\big(1-e^{-\Theta/T}\big) \quad . \tag{4.275}$$

The contribution to the heat capacity is

$$\Delta C_V = Nk_{\rm B}\left(\frac{\Theta}{T}\right)^2\frac{e^{\Theta/T}}{(e^{\Theta/T}-1)^2} = \begin{cases} Nk_{\rm B}\,(\Theta/T)^2\,\exp(-\Theta/T) & (T\to 0) \\ Nk_{\rm B} & (T\to\infty) \end{cases} \quad . \tag{4.276}$$

### 4.9.5 Two-level systems : Schottky anomaly

Consider now a two-level system, with energies $\varepsilon_0$ and $\varepsilon_1$. We define $\Delta \equiv \varepsilon_1 - \varepsilon_0$ and assume without loss of generality that $\Delta > 0$. The partition function is

$$\zeta = e^{-\beta\varepsilon_0} + e^{-\beta\varepsilon_1} = e^{-\beta\varepsilon_0}\big(1+e^{-\beta\Delta}\big) \quad . \tag{4.277}$$

The free energy is

$$f = -k_{\rm B}T\log\zeta = \varepsilon_0 - k_{\rm B}T\log\big(1+e^{-\Delta/k_{\rm B}T}\big) \quad . \tag{4.278}$$

The entropy for a given two level system is then

$$s = -\frac{\partial f}{\partial T} = k_{\rm B}\log\big(1+e^{-\Delta/k_{\rm B}T}\big) + \frac{\Delta}{T}\cdot\frac{1}{e^{\Delta/k_{\rm B}T}+1} \tag{4.279}$$

and the heat capacity is $= T\,(\partial s/\partial T)$, *i.e.*

$$c(T) = \frac{\Delta^2}{k_{\rm B}T^2}\cdot\frac{e^{\Delta/k_{\rm B}T}}{\big(e^{\Delta/k_{\rm B}T}+1\big)^2} \quad . \tag{4.280}$$

Thus,

$$c\,(k_{\rm B}T \ll \Delta) = \frac{\Delta^2}{k_{\rm B}T^2}\,e^{-\Delta/k_{\rm B}T} \qquad , \qquad c\,(k_{\rm B}T \gg \Delta) = \frac{\Delta^2}{4k_{\rm B}T^2} \quad . \tag{4.281}$$

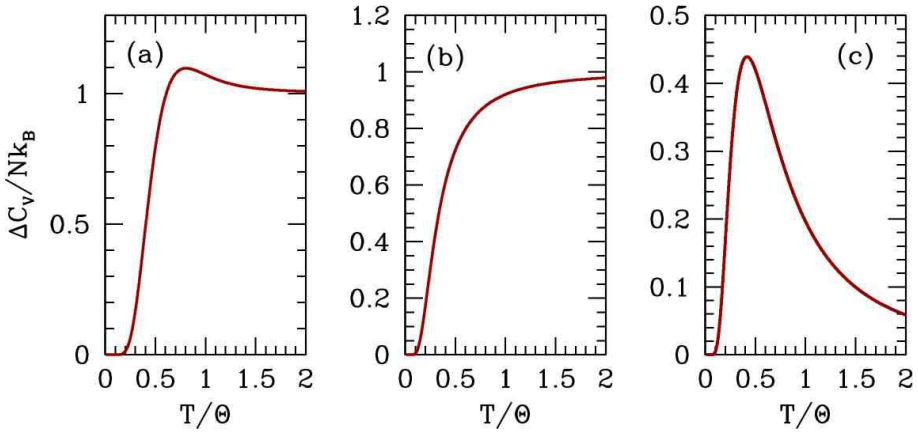



Figure 4.10: Heat capacity per molecule as a function of temperature for (a) heteronuclear diatomic gases, (b) a single vibrational mode, and (c) a single two-level system.

We find that $c(T)$ has a characteristic peak at $T^* \approx 0.42\,\Delta/k_{\rm B}$. The heat capacity vanishes in both the low temperature and high temperature limits. At low temperatures, the gap to the excited state is much greater than $k_{\rm B}T$, and it is not possible to populate it and store energy. At high temperatures, both ground state and excited state are equally populated, and once again there is no way to store energy.

If we have a distribution of independent two-level systems, the heat capacity of such a system is a sum over the individual Schottky functions:

$$C(T) = \sum_i \widetilde{c}\,(\Delta_i/k_{\rm B}T) = N\int\limits_0^\infty d\Delta\, P(\Delta)\,\widetilde{c}(\Delta/k_{\rm B}T) \quad , \tag{4.282}$$

where $N$ is the number of two level systems, $\widetilde{c}(x) = k_{\rm B}\,x^2\,e^x/(e^x+1)^2$, and where $P(\Delta)$ is the normalized distribution function, which satisfies the normalization condition

$$\int\limits_0^\infty d\Delta\, P(\Delta) = 1 \quad . \tag{4.283}$$

$N$ is the total number of two level systems. If $P(\Delta) \propto \Delta^r$ for $\Delta \to 0$, then the low temperature heat capacity behaves as $C(T) \propto T^{1+r}$. Many amorphous or glassy systems contain such a distribution of two level systems, with $r \approx 0$ for glasses, leading to a linear low-temperature heat capacity. The origin of these two-level systems is not always so clear but is generally believed to be associated with local atomic configurations for which there are two low-lying states which are close in energy. The paradigmatic example is the mixed crystalline solid $({\rm KBr})_{1-x}({\rm KCN})_x$ which over the range $0.1 \lesssim x \lesssim 0.6$ forms an 'orientational glass' at low temperatures. The two level systems are associated with different orientations of the cyanide (CN) dipoles.

### 4.9.6 Electronic and nuclear excitations

For a monatomic gas, the internal coordinate partition function arises due to electronic and nuclear degrees of freedom. Let's first consider the electronic degrees of freedom. We assume that $k_{\rm B}T$ is small compared with energy differences between successive electronic shells. The atomic ground state is then computed by filling up the hydrogenic orbitals until all the electrons are used up. If the atomic number is a 'magic number' ($Z = 2$ (He), 10 (Ne), 18 (Ar), 36 (Kr), 54 (Xe), *etc.*) then the atom has all shells filled and $L = 0$ and $S = 0$. Otherwise the last shell is partially filled and one or both of $L$ and $S$ will be nonzero. The atomic ground state configuration ${}^{2S+1}L_J$ is then determined by *Hund's rules*:

1. The $LS$ multiplet with the *largest* $S$ has the lowest energy.
2. If the largest value of $S$ is associated with several multiplets, the multiplet with the *largest* $L$ has the lowest energy.
3. If an incomplete shell is not more than half-filled, then the lowest energy state has $J = |L - S|$. If the shell is more than half-filled, then $J = L + S$.

The last of Hund's rules distinguishes between the $(2S+1)(2L+1)$ states which result upon fixing $S$ and $L$ as per rules #1 and #2. It arises due to the atomic spin-orbit coupling, whose effective Hamiltonian may be written $\hat{H} = \Lambda \boldsymbol{L} \cdot \boldsymbol{S}$, where $\Lambda$ is the Russell-Saunders coupling. If the last shell is less than or equal to half-filled, then $\Lambda > 0$ and the ground state has $J = |L - S|$. If the last shell is more than half-filled, the coupling is *inverted*, *i.e.* $\Lambda < 0$, and the ground state has $J = L + S$.[11]

The electronic contribution to $\xi$ is then

$$\xi_{\rm elec} = \sum_{J=|L-S|}^{L+S} (2J+1)\, e^{-\Delta\varepsilon(L,S,J)/k_{\rm B}T} \tag{4.284}$$

where

$$\Delta\varepsilon(L,S,J) = \tfrac{1}{2}\Lambda\Big[J(J+1) - L(L+1) - S(S+1)\Big] \quad . \tag{4.285}$$

At high temperatures, $k_{\rm B}T$ is larger than the energy difference between the different $J$ multiplets, and we have $\xi_{\rm elec} \sim (2L+1)(2S+1)\, e^{-\beta\varepsilon_0}$, where $\varepsilon_0$ is the ground state energy. At low temperatures, a particular value of $J$ is selected – that determined by Hund's third rule – and we have $\xi_{\rm elec} \sim (2J+1)\, e^{-\beta\varepsilon_0}$. If, in addition, there is a nonzero nuclear spin $I$, then we also must include a factor $\xi_{\rm nuc} = (2I+1)$, neglecting the small hyperfine splittings due to the coupling of nuclear and electronic angular momenta.

For *heteronuclear* diatomic molecules, *i.e.* molecules composed from two different atomic nuclei, the internal partition function simply receives a factor of $\xi_{\rm elec} \cdot \xi^{(1)}_{\rm nuc} \cdot \xi^{(2)}_{\rm nuc}$, where the first term is a sum over molecular electronic states, and the second two terms arise from the spin degeneracies of the two nuclei. For *homonuclear* diatomic molecules, the exchange of nuclear centers is a symmetry operation, and does not represent a distinct quantum state. To correctly count the electronic states, we first assume that the total electronic spin is $S = 0$. This is generally a very safe assumption. Exchange symmetry now puts restrictions on the possible values of the molecular angular momentum $L$, depending on the total nuclear

[11] See *e.g.* §72 of Landau and Lifshitz, *Quantum Mechanics*.

| $2I$ | $g_g$ | $g_u$ |
|---|---|---|
| odd | $I(2I+1)$ | $(I+1)(2I+1)$ |
| even | $(I+1)(2I+1)$ | $I(2I+1)$ |

Table 4.2: Number of even ($g_g$) and odd ($g_u$) total nuclear angular momentum states for a homonuclear diatomic molecule. $I$ is the ground state nuclear spin.

angular momentum $I_{\rm tot}$. If $I_{\rm tot}$ is even, then the molecular angular momentum $L$ must also be even. If the total nuclear angular momentum is odd, then $L$ must be odd. This is so because the molecular ground state configuration is ${}^1\Sigma_g^+$.[12]

The total number of nuclear states for the molecule is $(2I+1)^2$, of which some are even under nuclear exchange, and some are odd. The number of even states, corresponding to even total nuclear angular momentum is written as $g_g$, where the subscript conventionally stands for the (mercifully short) German word *gerade*, meaning 'even'. The number of odd (Ger. *ungerade*) states is written $g_u$. Table 4.2 gives the values of $g_{g,u}$ corresponding to half-odd-integer $I$ and integer $I$.

The final answer for the rotational component of the internal molecular partition function is then

$$\xi_{\rm rot}(T) = g_g\,\zeta_g(T) + g_u\,\zeta_u(T) \quad , \tag{4.286}$$

where

$$\zeta_g(T) = \sum_{L\ \rm even} (2L+1)\,e^{-L(L+1)\,\Theta_{\rm rot}/T} \quad , \qquad \zeta_u(T) = \sum_{L\ \rm odd} (2L+1)\,e^{-L(L+1)\,\Theta_{\rm rot}/T} \quad . \tag{4.287}$$

For hydrogen, the molecules with the larger nuclear statistical weight are called *orthohydrogen* and those with the smaller statistical weight are called *parahydrogen*. For $H_2$, we have $I = \frac{1}{2}$ hence the ortho state has $g_u = 3$ and the para state has $g_g = 1$. In $D_2$, we have $I = 1$ and the ortho state has $g_g = 6$ while the para state has $g_u = 3$. In equilibrium, the ratio of ortho to para states is then

$$\frac{N^{\rm ortho}_{\rm H_2}}{N^{\rm para}_{\rm H_2}} = \frac{g_u\,\zeta_u(T)}{g_g\,\zeta_g(T)} = \frac{3\,\zeta_u(T)}{\zeta_g(T)} \quad , \qquad \frac{N^{\rm ortho}_{\rm D_2}}{N^{\rm para}_{\rm D_2}} = \frac{g_g\,\zeta_g(T)}{g_u\,\zeta_u(T)} = \frac{2\,\zeta_g(T)}{\zeta_u(T)} \quad . \tag{4.288}$$

Incidentally, how do we derive the results in Tab. 4.2? The total nuclear angular momentum $I_{\rm tot}$ is the quantum mechanical sum of the two individual nuclear angular momenta, each of which are of magnitude $I$. From elementary addition of angular momenta, we have

$$I \otimes I = 0 \oplus 1 \oplus 2 \oplus \cdots \oplus 2I \quad . \tag{4.289}$$

The right hand side of the above equation lists all the possible multiplets. Thus, $I_{\rm tot} \in \{0, 1, \ldots, 2I\}$. Now let us count the total number of states with *even* $I_{\rm tot}$. If $2I$ is even, which is to say if $I$ is an integer, we have

$$g_g^{\rm even} = \sum_{n=0}^{I} \Big\{ 2\cdot(2n) + 1 \Big\} = (I+1)(2I+1) \quad , \tag{4.290}$$

[12] *Ibid.* §86.

because the degeneracy of each multiplet is $2I_{\rm tot}+1$. It follows that

$$g_u^{\sf even} = (2I+1)^2 - g_g^{\sf even} = I(2I+1) \quad . \tag{4.291}$$

On the other hand, if $2I$ is odd, which is to say $I$ is a half odd integer, then

$$g_g^{\sf odd} = \sum_{n=0}^{I-\frac{1}{2}} \Big\{ 2\cdot(2n)+1 \Big\} = I(2I+1) \quad . \tag{4.292}$$

It follows that

$$g_u^{\sf odd} = (2I+1)^2 - g_g^{\sf odd} = (I+1)(2I+1) \quad . \tag{4.293}$$

## 4.10 Appendix I: Constrained Extremization of Functions

Given $F(x_1,\ldots,x_n)$ to be extremized subject to $k$ constraints of the form $G_j(x_1,\ldots,x_n)=0$ where $j=1,\ldots,k$, construct

$$F^*\big(x_1,\ldots,x_n;\lambda_1,\ldots,\lambda_k\big) \equiv F(x_1,\ldots,x_n) + \sum_{j=1}^k \lambda_j\, G_j(x_1,\ldots,x_n) \tag{4.294}$$

which is a function of the $(n+k)$ variables $\big\{x_1,\ldots,x_n;\lambda_1,\ldots,\lambda_k\big\}$, where the quantities $\{\lambda_1,\ldots,\lambda_k\}$ are *Lagrange undetermined multipliers*. We now *freely* extremize the extended function $F^*$:

$$\begin{aligned} dF^* &= \sum_{\sigma=1}^n \frac{\partial F^*}{\partial x_\sigma}\,dx_\sigma + \sum_{j=1}^k \frac{\partial F^*}{\partial \lambda_j}\,d\lambda_j \\ &= \sum_{\sigma=1}^n \left( \frac{\partial F}{\partial x_\sigma} + \sum_{j=1}^k \lambda_j \frac{\partial G_j}{\partial x_\sigma} \right) dx_\sigma + \sum_{j=1}^k G_j\,d\lambda_j = 0 \end{aligned} \tag{4.295}$$

This results in the $(n+k)$ equations

$$\begin{aligned} \frac{\partial F}{\partial x_\sigma} + \sum_{j=1}^k \lambda_j \frac{\partial G_j}{\partial x_\sigma} &= 0 \qquad (\sigma=1,\ldots,n) \\ G_j &= 0 \qquad (j=1,\ldots,k) \quad . \end{aligned} \tag{4.296}$$

The interpretation of all this is as follows. The first $n$ equations in 4.296 can be written in vector form as

$$\boldsymbol{\nabla} F + \sum_{j=1}^k \lambda_j\, \boldsymbol{\nabla} G_j = 0 \quad . \tag{4.297}$$

This says that the ($n$-component) vector $\boldsymbol{\nabla} F$ is linearly dependent upon the $k$ vectors $\boldsymbol{\nabla} G_j$. Thus, any movement in the direction of $\boldsymbol{\nabla} F$ must necessarily entail movement along one or more of the directions

$\boldsymbol{\nabla} G_j$. This would require violating the constraints, since movement along $\boldsymbol{\nabla} G_j$ takes us off the level set $G_j = 0$. Were $\boldsymbol{\nabla} F$ linearly *independent* of the set $\{\boldsymbol{\nabla} G_j\}$, this would mean that we could find a differential displacement $d\boldsymbol{x}$ which has finite overlap with $\boldsymbol{\nabla} F$ but zero overlap with each $\boldsymbol{\nabla} G_j$. Thus $\boldsymbol{x} + d\boldsymbol{x}$ would still satisfy $G_j(\boldsymbol{x} + d\boldsymbol{x}) = 0$, but $F$ would change by the finite amount $dF = \boldsymbol{\nabla} F(\boldsymbol{x}) \cdot d\boldsymbol{x}$.

Put another way, when we extremize $F(\boldsymbol{x})$ without constraints, we identify points $\boldsymbol{x} \in \mathbb{R}^n$ where the gradient $\boldsymbol{\nabla} F$ vanishes. However, when we have $k$ constraints of the form $G_j(\boldsymbol{x}) = 0$, the subset

$$\Sigma = \big\{ \boldsymbol{x} \in \mathbb{R}^n \,|\, G_j(\boldsymbol{x}) = 0 \;\forall\; j \in \{1, \ldots, k\} \big\} \tag{4.298}$$

is a hypersurface of dimension $n - k$. Generically we should not expect any of the solutions to $\boldsymbol{\nabla} F = 0$ to lie within the subspace $\Sigma$. Extremizing $F(\boldsymbol{x})$ subject to the $k$ constraints $G_j(\boldsymbol{x}) = 0$ means that we must find the extrema of $F(\boldsymbol{x})$ for $\boldsymbol{x} \in \Sigma \subset \mathbb{R}^n$. All such extrema satisfy that $\boldsymbol{\nabla} F(\boldsymbol{x})$ is *perpendicular* to the hypersurface $\Sigma$, *i.e.* $\boldsymbol{\nabla} F(\boldsymbol{x})$ must lie in the $k$-dimensional subspace spanned by the vectors $\boldsymbol{\nabla} G_j(\boldsymbol{x})$.

### 4.10.1 Example : volume of a cylinder

To see how this formalism works in practice, let's extremize the volume $V = \pi a^2 h$ of a cylinder of radius $a$ and height $h$, subject to the constraint

$$G(a, h) = 2\pi a + \frac{h^2}{b} - \ell = 0 \quad . \tag{4.299}$$

Here, $b$ and $\ell$ are constant parameters, each of which has dimensions of length.

Following Lagrange's method, we define the extended function

$$V^*(a, h, \lambda) \equiv V(a, h) + \lambda\, G(a, h) \quad , \tag{4.300}$$

and set

$$\begin{aligned} \frac{\partial V^*}{\partial a} &= 2\pi a h + 2\pi\lambda = 0 \\ \frac{\partial V^*}{\partial h} &= \pi a^2 + 2\lambda\,\frac{h}{b} = 0 \\ \frac{\partial V^*}{\partial \lambda} &= 2\pi a + \frac{h^2}{b} - \ell = 0 \quad . \end{aligned} \tag{4.301}$$

Solving these three equations simultaneously gives

$$a = \frac{2\ell}{5\pi} \quad , \quad h = \sqrt{\frac{b\ell}{5}} \quad , \quad \lambda = -\frac{2}{5^{3/2}\pi}\, b^{1/2}\, \ell^{3/2} \quad , \quad V^* = \frac{4}{5^{5/2}\,\pi}\, \ell^{5/2}\, b^{1/2} \quad . \tag{4.302}$$

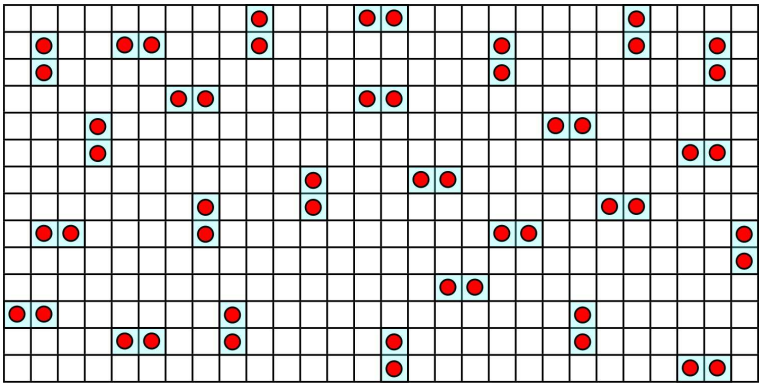


Figure 4.11: A model of noninteracting spin dimers on a lattice. Each red dot represents a classical spin for which $\sigma_j = \pm 1$.

## 4.11 Appendix II: Additional Examples of Statistical Ensembles

### 4.11.1 Noninteracting spin dimers

Consider a system of noninteracting spin dimers as depicted in fig. 4.11. Each dimer contains two spins, and is described by the Hamiltonian

$$\hat{h} = -J\,\sigma_1\sigma_2 - \mu_0 H\,(\sigma_1 + \sigma_2) \quad . \tag{4.303}$$

Here, $J$ is an *interaction energy* between the spins which comprise the dimer. If $J > 0$ the interaction is *ferromagnetic*, which prefers that the dimer spins are aligned. That is, the lowest energy states are $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$. If $J < 0$ the interaction is *antiferromagnetic*, which prefers that the dimer spins are anti-aligned, *i.e.* in configurations $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$[13].

Suppose there are $N_{\rm d}$ dimers. Then the OCE partition function is $Z = \zeta^{N_{\rm d}}$, where $\zeta(T,H)$ is the single dimer partition function. To obtain $\zeta(T,H)$, we sum over the four possible states of the two spins, obtaining

$$\zeta = {\rm Tr}\, e^{-\hat{h}/k_{\rm B}T} = 2\,e^{-J/k_{\rm B}T} + 2\,e^{J/k_{\rm B}T}\cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.304}$$

Thus, the free energy is

$$F(T,H,N_{\rm d}) = -N_{\rm d}\,k_{\rm B}T\log 2 - N_{\rm d}\,k_{\rm B}T\log\left[e^{-J/k_{\rm B}T} + e^{J/k_{\rm B}T}\cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)\right] \quad . \tag{4.305}$$

The magnetization is

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T,N_{\rm d}} = 2N_{\rm d}\,\mu_0\cdot\frac{e^{J/k_{\rm B}T}\sinh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)}{e^{-J/k_{\rm B}T} + e^{J/k_{\rm B}T}\cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)} \tag{4.306}$$

[13] *Nota bene* we are concerned with classical spin configurations only – there is no superposition of states allowed in this model.

It is instructive to consider the zero field isothermal susceptibility per spin,

$$\chi_T = \frac{1}{2N_{\rm d}} \frac{\partial M}{\partial H}\bigg|_{H=0} = \frac{\mu_0^2}{k_{\rm B}T} \cdot \frac{2\,e^{J/k_{\rm B}T}}{e^{J/k_{\rm B}T} + e^{-J/k_{\rm B}T}} \quad . \tag{4.307}$$

The quantity $\mu_0^2/k_{\rm B}T$ is simply the Curie susceptibility for noninteracting classical spins. Note that we correctly recover the Curie result when $J = 0$, since then the individual spins comprising each dimer are in fact noninteracting. For the ferromagnetic case, if $J \gg k_{\rm B}T$, then we obtain

$$\chi_T(J \gg k_{\rm B}T) \approx \frac{2\mu_0^2}{k_{\rm B}T} \quad . \tag{4.308}$$

This has the following simple interpretation. When $J \gg k_{\rm B}T$, the spins of each dimer are effectively locked in parallel. Thus, each dimer has an effective magnetic moment $\mu_{\rm eff} = 2\mu_0$. On the other hand, there are only half as many dimers as there are spins, so the resulting Curie susceptibility per spin is $\frac{1}{2} \times (2\mu_0)^2/k_{\rm B}T$.

When $-J \gg k_{\rm B}T$, the spins of each dimer are effectively locked in one of the two antiparallel configurations. We then have

$$\chi_T(-J \gg k_{\rm B}T) \approx \frac{2\mu_0^2}{k_{\rm B}T}\, e^{-2|J|/k_{\rm B}T} \quad . \tag{4.309}$$

In this case, the individual dimers have essentially zero magnetic moment.

### 4.11.2 Three-state system

Consider a spin-1 particle where $\sigma = -1, 0, +1$. We model this with the single particle Hamiltonian

$$\hat{h} = -\mu_0 H\,\sigma + \Delta(1 - \sigma^2) \quad . \tag{4.310}$$

We can also interpret this as describing a spin if $\sigma = \pm 1$ and a vacancy if $\sigma = 0$. The parameter $\Delta$ then represents the vacancy formation energy. The single particle partition function is

$$\zeta = \mathsf{Tr}\; e^{-\beta\hat{h}} = e^{-\beta\Delta} + 2\cosh(\beta\mu_0 H) \quad . \tag{4.311}$$

With $N$ distinguishable noninteracting spins (*e.g.* at different sites in a crystalline lattice), we have $Z = \zeta^N$ and

$$F \equiv Nf = -k_{\rm B}T \log Z = -Nk_{\rm B}T \log\Big[e^{-\beta\Delta} + 2\cosh(\beta\mu_0 H)\Big] \quad , \tag{4.312}$$

where $f = -k_{\rm B}T \log\zeta$ is the free energy of a single particle. Note that

$$\hat{n}_{\rm V} = 1 - \sigma^2 = \frac{\partial\hat{h}}{\partial\Delta} \qquad , \qquad \hat{m} = \mu_0\,\sigma = -\frac{\partial\hat{h}}{\partial H} \quad , \tag{4.313}$$

are the vacancy number and magnetization, respectively. Thus,

$$n_{\rm V} = \big\langle\,\hat{n}_{\rm V}\,\big\rangle = \frac{\partial f}{\partial\Delta} = \frac{e^{-\Delta/k_{\rm B}T}}{e^{-\Delta/k_{\rm B}T} + 2\cosh(\mu_0 H/k_{\rm B}T)} \tag{4.314}$$

and

$$m = \langle\, \hat{m} \,\rangle = -\frac{\partial f}{\partial H} = \frac{2\mu_0 \, \sinh(\mu_0 H/k_{\rm B}T)}{e^{-\Delta/k_{\rm B}T} + 2\cosh(\mu_0 H/k_{\rm B}T)} \quad . \tag{4.315}$$

At weak fields we can compute

$$\chi_T = \left.\frac{\partial m}{\partial H}\right|_{H=0} = \frac{\mu_0^2}{k_{\rm B}T} \cdot \frac{2}{2 + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.316}$$

We thus obtain a modified Curie law. At temperatures $T \ll \Delta/k_{\rm B}$, the vacancies are frozen out and we recover the usual Curie behavior. At high temperatures, where $T \gg \Delta/k_{\rm B}$, the low temperature result is reduced by a factor of $\frac{2}{3}$, which accounts for the fact that one third of the time the particle is in a nonmagnetic state with $\sigma = 0$.

### 4.11.3 Spins and vacancies on a surface

PROBLEM: A collection of spin-$\frac{1}{2}$ particles is confined to a surface with $N$ sites. For each site, let $\sigma = 0$ if there is a vacancy, $\sigma = +1$ if there is a particle present with spin up, and $\sigma = -1$ if there is a particle present with spin down. The particles are non-interacting, and the energy for each site is given by $\varepsilon = -W\sigma^2$, where $-W < 0$ is the binding energy.

(a) Let $Q = N_\uparrow + N_\downarrow$ be the number of spins, and $N_0$ be the number of vacancies. The surface magnetization is $M = N_\uparrow - N_\downarrow$. Compute, in the microcanonical ensemble, the statistical entropy $S(Q, M)$.

(b) Let $q = Q/N$ and $m = M/N$ be the dimensionless particle density and magnetization density, respectively. Assuming that we are in the thermodynamic limit, where $N$, $Q$, and $M$ all tend to infinity, but with $q$ and $m$ finite, find the temperature $T(q, m)$. Recall Stirling's formula

$$\log(N!) = N \log N - N + \mathcal{O}(\log N) \quad .$$

(c) Show explicitly that $T$ can be negative for this system. What does negative $T$ mean? What physical degrees of freedom have been left out that would avoid this strange property?

SOLUTION: There is a constraint on $N_\uparrow$, $N_0$, and $N_\downarrow$:

$$N_\uparrow + N_0 + N_\downarrow = Q + N_0 = N \quad . \tag{4.317}$$

The total energy of the system is $E = -WQ$.

(a) The number of states available to the system is

$$\Omega = \frac{N!}{N_\uparrow!\, N_0!\, N_\downarrow!} \quad . \tag{4.318}$$

Fixing $Q$ and $M$, along with the above constraint, is enough to completely determine $\{N_\uparrow, N_0, N_\downarrow\}$:

$$N_\uparrow = \tfrac{1}{2}\,(Q+M) \qquad , \qquad N_0 = N - Q \qquad , \qquad N_\downarrow = \tfrac{1}{2}\,(Q-M) \quad , \tag{4.319}$$

whence

$$\Omega(Q,M) = \frac{N!}{\left[\frac{1}{2}(Q+M)\right]!\left[\frac{1}{2}(Q-M)\right]!\,(N-Q)!} \quad . \tag{4.320}$$

The statistical entropy is $S = k_{\rm B}\log\Omega$:

$$S(Q,M) = k_{\rm B}\log(N!) - k_{\rm B}\log\left[\tfrac{1}{2}(Q+M)!\right] - k_{\rm B}\log\left[\tfrac{1}{2}(Q-M)!\right] - k_{\rm B}\log\left[(N-Q)!\right] \quad . \tag{4.321}$$

(b) Now we invoke Stirling's rule,

$$\log(N!) = N\log N - N + \mathcal{O}(\log N) \quad , \tag{4.322}$$

to obtain

$$\log\Omega(Q,M) = N\log N - N - \tfrac{1}{2}(Q+M)\log\left[\tfrac{1}{2}(Q+M)\right] + \tfrac{1}{2}(Q+M) \tag{4.323}$$

$$- \tfrac{1}{2}(Q-M)\log\left[\tfrac{1}{2}(Q-M)\right] + \tfrac{1}{2}(Q-M) - (N-Q)\log(N-Q) + (N-Q)$$

$$= N\log N - \tfrac{1}{2}Q\log\left[\tfrac{1}{4}(Q^2-M^2)\right] - \tfrac{1}{2}M\log\left(\frac{Q+M}{Q-M}\right) \tag{4.324}$$

Combining terms,

$$\log\Omega(Q,M) = -Nq\log\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right] - \tfrac{1}{2}Nm\log\left(\frac{q+m}{q-m}\right) - N(1-q)\log(1-q) \quad , \tag{4.325}$$

where $Q = Nq$ and $M = Nm$. Note that the entropy $S = k_{\rm B}\log\Omega$ is extensive. The statistical entropy per site is thus

$$s(q,m) = -k_{\rm B}\,q\log\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right] - \tfrac{1}{2}k_{\rm B}\,m\log\left(\frac{q+m}{q-m}\right) - k_{\rm B}\,(1-q)\log(1-q) \quad . \tag{4.326}$$

The temperature is obtained from the relation

$$\begin{aligned}\frac{1}{T} &= \left(\frac{\partial S}{\partial E}\right)_M = -\frac{1}{W}\left(\frac{\partial s}{\partial q}\right)_m \\ &= -\frac{k_{\rm B}}{W}\,\log(1-q) + \frac{k_{\rm B}}{W}\,\log\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right] \quad .\end{aligned} \tag{4.327}$$

Thus,

$$T = -\frac{W/k_{\rm B}}{\log\left[2(1-q)/\sqrt{q^2-m^2}\right]} \quad . \tag{4.328}$$

(c) We have $0 \le q \le 1$ and $-q \le m \le q$, so $T$ is real (thank heavens!). But it is easy to choose $\{q, m\}$ such that $T < 0$. For example, when $m = 0$ we have $T = -W/k_{\rm B} \log(2q^{-1} - 2)$ and $T < 0$ for all $q \in [0, \frac{2}{3})$. The reason for this strange state of affairs is that the entropy $S$ is bounded, and is not a monotonically increasing function of the energy $E$ (or the dimensionless quantity $Q$). The entropy is maximized for $N_\uparrow = N_0 = N_\downarrow = \frac{1}{3}N$, which says $m = 0$ and $q = \frac{2}{3}$. Increasing $q$ beyond this point (with $m = 0$ fixed) starts to reduce the entropy, and hence $(\partial S/\partial E) < 0$ in this range, which immediately gives $T < 0$. What we've left out are kinetic degrees of freedom, such as vibrations and rotations, whose energies are unbounded, and which result in an increasing $S(E)$ function.

### 4.11.4 Fluctuating interface

Consider an interface between two dissimilar fluids. In equilibrium, in a uniform gravitational field, the denser fluid is on the bottom. Let $z = z(x, y)$ be the height of the interface between the fluids, relative to equilibrium. The potential energy is a sum of gravitational and surface tension terms, with

$$U_{\rm grav} = \int d^2x \int_0^z dz' \, \Delta\rho \, g \, z' \qquad , \qquad U_{\rm surf} = \tfrac{1}{2}\sigma \int d^2x \, (\boldsymbol{\nabla} z)^2 \quad . \tag{4.329}$$

We won't need the kinetic energy in our calculations, but we can include it just for completeness. It isn't so clear how to model it *a priori* so we will assume a rather general form

$$T = \int d^2x \int d^2x' \, \tfrac{1}{2}\mu(\boldsymbol{x}, \boldsymbol{x}') \, \frac{\partial z(\boldsymbol{x}, t)}{\partial t} \, \frac{\partial z(\boldsymbol{x}', t)}{\partial t} \quad . \tag{4.330}$$

We assume that the $(x, y)$ plane is a rectangle of dimensions $L_x \times L_y$. We also assume $\mu(\boldsymbol{x}, \boldsymbol{x}') = \mu\big(|\boldsymbol{x} - \boldsymbol{x}'|\big)$. We can then Fourier transform

$$z(\boldsymbol{x}) = \big(L_x \, L_y\big)^{-1/2} \sum_{\boldsymbol{k}} z_{\boldsymbol{k}} \, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \quad , \tag{4.331}$$

where the wavevectors $\boldsymbol{k}$ are quantized according to

$$\boldsymbol{k} = \frac{2\pi n_x}{L_x} \, \hat{\boldsymbol{x}} + \frac{2\pi n_y}{L_y} \, \hat{\boldsymbol{y}} \quad , \tag{4.332}$$

with integer $n_x$ and $n_y$, if we impose periodic boundary conditions (for calculational convenience). The Lagrangian is then

$$L = \frac{1}{2} \sum_{\boldsymbol{k}} \Big[ \mu_{\boldsymbol{k}} \, \big|\dot{z}_{\boldsymbol{k}}\big|^2 - \big(g \, \Delta\rho + \sigma \boldsymbol{k}^2\big) \, \big|z_{\boldsymbol{k}}\big|^2 \Big] \quad , \tag{4.333}$$

where

$$\mu_{\boldsymbol{k}} = \int d^2x \, \mu\big(|\boldsymbol{x}|\big) \, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \quad . \tag{4.334}$$

Since $z(\boldsymbol{x}, t)$ is real, we have the relation $z_{-\boldsymbol{k}} = z_{\boldsymbol{k}}^*$, therefore the Fourier coefficients at $\boldsymbol{k}$ and $-\boldsymbol{k}$ are not independent. The canonical momenta are given by

$$p_{\boldsymbol{k}} = \frac{\partial L}{\partial \dot{z}_{\boldsymbol{k}}^*} = \mu_{\boldsymbol{k}} \, \dot{z}_{\boldsymbol{k}} \qquad , \qquad p_{\boldsymbol{k}}^* = \frac{\partial L}{\partial \dot{z}_{\boldsymbol{k}}} = \mu_{\boldsymbol{k}} \, \dot{z}_{\boldsymbol{k}}^* \tag{4.335}$$

The Hamiltonian is then

$$\begin{aligned}\hat{H} &= \sum_{\boldsymbol{k}}{}' \Big[ p_{\boldsymbol{k}}\,\dot{z}^*_{\boldsymbol{k}} + p^*_{\boldsymbol{k}}\,\dot{z}_{\boldsymbol{k}} \Big] - L \\ &= \sum_{\boldsymbol{k}}{}' \left[ \frac{|p_{\boldsymbol{k}}|^2}{\mu_{\boldsymbol{k}}} + \big(g\,\Delta\rho + \sigma \boldsymbol{k}^2\big)\,|z_{\boldsymbol{k}}|^2 \right] \quad ,\end{aligned} \tag{4.336}$$

where the prime on the $\boldsymbol{k}$ sum indicates that only one of the pair $\{\boldsymbol{k}, -\boldsymbol{k}\}$ is to be included, for each $\boldsymbol{k}$.

We may now compute the ordinary canonical partition function:

$$\begin{aligned}Z &= \prod_{\boldsymbol{k}}{}' \int \frac{d^2p_{\boldsymbol{k}}\,d^2z_{\boldsymbol{k}}}{(2\pi\hbar)^2}\, e^{-|p_{\boldsymbol{k}}|^2/\mu_{\boldsymbol{k}}k_{\rm B}T}\, e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \\ &= \prod_{\boldsymbol{k}}{}' \left(\frac{k_{\rm B}T}{2\hbar}\right)^2 \left(\frac{\mu_{\boldsymbol{k}}}{g\,\Delta\rho + \sigma\boldsymbol{k}^2}\right) \quad .\end{aligned} \tag{4.337}$$

Thus,

$$F = -k_{\rm B}T \sum_{\boldsymbol{k}} \log\left(\frac{k_{\rm B}T}{2\hbar\Omega_{\boldsymbol{k}}}\right) \quad , \tag{4.338}$$

where[14]

$$\Omega_{\boldsymbol{k}} = \left(\frac{g\,\Delta\rho + \sigma\boldsymbol{k}^2}{\mu_{\boldsymbol{k}}}\right)^{1/2} \quad . \tag{4.339}$$

is the normal mode frequency for surface oscillations at wavevector $\boldsymbol{k}$. For deep water waves, it is appropriate to take $\mu_{\boldsymbol{k}} = \Delta\rho/|\boldsymbol{k}|$, where $\Delta\rho = \rho_{\rm L} - \rho_{\rm G} \approx \rho_{\rm L}$ is the difference between the densities of water and air.

It is now easy to compute the thermal average

$$\begin{aligned}\big\langle\, |z_{\boldsymbol{k}}|^2 \,\big\rangle &= \int d^2z_{\boldsymbol{k}}\, |z_{\boldsymbol{k}}|^2\, e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \Bigg/ \int d^2z_{\boldsymbol{k}}\, e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \\ &= \frac{k_{\rm B}T}{g\,\Delta\rho + \sigma\boldsymbol{k}^2} \quad .\end{aligned} \tag{4.340}$$

Note that this result does not depend on $\mu_{\boldsymbol{k}}$, *i.e.* on our choice of kinetic energy. One defines the *correlation function*

$$\begin{aligned}C(\boldsymbol{x}) \equiv \big\langle z(\boldsymbol{x})\,z(0)\big\rangle &= \frac{1}{L_xL_y}\sum_{\boldsymbol{k}} \big\langle\, |z_{\boldsymbol{k}}|^2 \,\big\rangle\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = \int \frac{d^2k}{(2\pi)^2} \left(\frac{k_{\rm B}T}{g\,\Delta\rho + \sigma\boldsymbol{k}^2}\right) e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \\ &= \frac{k_{\rm B}T}{2\pi\sigma} \int\limits_0^\infty dk\, k\, \frac{J_0\big(k\,|\boldsymbol{x}|\big)}{k^2 + \xi^{-2}} = \frac{k_{\rm B}T}{2\pi\sigma}\, K_0\big(|\boldsymbol{x}|/\xi\big) \quad ,\end{aligned} \tag{4.341}$$

[14]Note that there is no prime on the $\boldsymbol{k}$ sum for $F$, as we have divided the logarithm of $Z$ by two and replaced the half sum by the whole sum.

where $\xi = \sqrt{\sigma/g\,\Delta\rho}$ is the correlation length, and where $K_0(z)$ is the modified Bessel function of the second kind. The asymptotic behavior of $K_0(z)$ for small $z$ is $K_0(z) \sim \log(2/z)$, whereas for large $z$ one has $K_0(z) \sim (\pi/2z)^{1/2}\,e^{-z}$. We see that on large length scales the correlations decay exponentially, but on small length scales they diverge. This divergence is due to the improper energetics we have assigned to short wavelength fluctuations of the interface. Roughly, it can be cured by imposing a cutoff on the integral, or by insisting that the shortest distance scale is a molecular diameter.

### 4.11.5 Dissociation of atomic hydrogen

Suppose we fill a volume $V$ with a number density $n$ of hydrogen atoms. Now let the reaction

$$\mathrm{H} \rightleftharpoons \mathrm{p}^+ + \mathrm{e}^- \tag{4.342}$$

run until thermal equilibrium is attained. Under equilibrium conditions, some fraction $x$ of the original $nV$ hydrogen atoms will have dissociated, and we will have

$$n_{\mathrm{H}} = (1-x)\,n \qquad , \qquad n_{\mathrm{p}} = xn \qquad , \qquad n_{\mathrm{e}} = xn \quad . \tag{4.343}$$

What is the equilibrium relationship between the temperature $T$, the dissociation fraction $x$, and the original number density $n$? To find this, we must invoke the machinery of statistical mechanics, invoking the key relation (see eqn. 2.379)

$$\mu_{\mathrm{H}} = \mu_{\mathrm{p}} + \mu_{\mathrm{e}} \quad . \tag{4.344}$$

The single particle partition function for each species is

$$\zeta = \left(\frac{gV}{\lambda_T^3}\right) e^{-\varepsilon_{\mathrm{int}}/k_{\mathrm{B}}T} \quad , \tag{4.345}$$

where $g$ is the degeneracy and $\varepsilon_{\mathrm{int}}$ the internal energy for a given species. We have $\varepsilon_{\mathrm{int}} = 0$ for p and e, and $\varepsilon_{\mathrm{int}} = -\Delta$ for H, where $\Delta = e^2/2a_{\mathrm{B}} = 13.6\,\mathrm{eV}$, the binding energy of hydrogen. Neglecting hyperfine splittings[15], we have $g_{\mathrm{H}} = 4$, while $g_{\mathrm{e}} = g_{\mathrm{p}} = 2$ because each has spin $S = \frac{1}{2}$. Thus, the associated grand potentials are

$$\begin{aligned}
\Omega_{\mathrm{H}}(T,V,\mu_{\mathrm{H}}) &= -g_{\mathrm{H}}\,Vk_{\mathrm{B}}T\,\lambda_{T,\mathrm{H}}^{-3}\,e^{(\mu_{\mathrm{H}}+\Delta)/k_{\mathrm{B}}T} \\
\Omega_{\mathrm{p}}(T,V,\mu_{\mathrm{p}}) &= -g_{\mathrm{p}}\,Vk_{\mathrm{B}}T\,\lambda_{T,\mathrm{p}}^{-3}\,e^{\mu_{\mathrm{p}}/k_{\mathrm{B}}T} \\
\Omega_{\mathrm{e}}(T,V,\mu_{\mathrm{e}}) &= -g_{\mathrm{e}}\,Vk_{\mathrm{B}}T\,\lambda_{T,\mathrm{e}}^{-3}\,e^{\mu_{\mathrm{e}}/k_{\mathrm{B}}T} \quad ,
\end{aligned} \tag{4.346}$$

where

$$\lambda_{T,a} = \sqrt{\frac{2\pi\hbar^2}{m_a k_{\mathrm{B}} T}} \tag{4.347}$$

[15]The hyperfine splitting in hydrogen is on the order of $(m_{\mathrm{e}}/m_{\mathrm{p}})\,\alpha^4\,m_{\mathrm{e}}c^2 \sim 10^{-6}\,\mathrm{eV}$, which is on the order of $0.01\,\mathrm{K}$. Here $\alpha = e^2/\hbar c$ is the fine structure constant.

for species $a$. The corresponding number densities are

$$n = \frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = g\,\lambda_T^{-3}\,e^{(\mu-\varepsilon_{\rm int})/k_{\rm B}T} \quad , \tag{4.348}$$

and the fugacity $z = e^{\mu/k_{\rm B}T}$ of a given species is thus given by

$$z = g^{-1} n\,\lambda_T^3\,e^{\varepsilon_{\rm int}/k_{\rm B}T} \quad . \tag{4.349}$$

We now invoke $\mu_{\rm H} = \mu_{\rm p} + \mu_{\rm e}$, which says $z_{\rm H} = z_{\rm p}\,z_{\rm e}$, *i.e.*

$$g_{\rm H}^{-1} n_{\rm H}\,\lambda_{T,{\rm H}}^3\,e^{-\Delta/k_{\rm B}T} = \left(g_{\rm p}^{-1} n_{\rm p}\,\lambda_{T,{\rm p}}^3\right)\left(g_{\rm e}^{-1} n_{\rm e}\,\lambda_{T,{\rm e}}^3\right) \quad , \tag{4.350}$$

which yields

$$\left(\frac{x^2}{1-x}\right) n\tilde{\lambda}_T^3 = e^{-\Delta/k_{\rm B}T} \quad , \tag{4.351}$$

where $\tilde{\lambda}_T = \sqrt{2\pi\hbar^2/m^* k_{\rm B}T}$, with $m^* = m_{\rm p}m_{\rm e}/m_{\rm H} \approx m_{\rm e}$. Note that

$$\tilde{\lambda}_T = a_{\rm B}\sqrt{\frac{4\pi m_{\rm H}}{m_{\rm p}}}\sqrt{\frac{\Delta}{k_{\rm B}T}} \quad , \tag{4.352}$$

where $a_{\rm B} = 0.529\,\text{Å}$ is the Bohr radius. Thus, we have

$$\left(\frac{x^2}{1-x}\right)\cdot(4\pi)^{3/2}\,\nu = \left(\frac{T}{T_0}\right)^{3/2} e^{-T_0/T} \quad , \tag{4.353}$$

where $T_0 = \Delta/k_{\rm B} = 1.578\times10^5\,{\rm K}$ and $\nu = n a_{\rm B}^3$. Consider for example a temperature $T = 3000\,{\rm K}$, for which $T_0/T = 52.6$, and assume that $x = \frac{1}{2}$. We then find $\nu = 1.69\times10^{-27}$, corresponding to a density of $n = 1.14\times10^{-2}\,{\rm cm}^{-3}$. At this temperature, the fraction of hydrogen atoms in their first excited (2s) state is $x' \sim e^{-T_0/2T} = 3.8\times10^{-12}$. This is quite striking: half the hydrogen atoms are completely dissociated, which requires an energy of $\Delta$, yet the number in their first excited state, requiring energy $\frac{1}{2}\Delta$, is twelve orders of magnitude smaller. The student should reflect on why this can be the case.

# Chapter 5

# Noninteracting Quantum Systems

## 5.1 References

- F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, 1987)
  This has been perhaps the most popular undergraduate text since it first appeared in 1967, and with good reason.
- A. H. Carter, *Classical and Statistical Thermodynamics*
  (Benjamin Cummings, 2000)
  A very relaxed treatment appropriate for undergraduate physics majors.
- D. V. Schroeder, *An Introduction to Thermal Physics* (Addison-Wesley, 2000)
  This is the best undergraduate thermodynamics book I've come across, but only 40% of the book treats statistical mechanics.
- C. Kittel, *Elementary Statistical Physics* (Dover, 2004)
  Remarkably crisp, though dated, this text is organized as a series of brief discussions of key concepts and examples. Published by Dover, so you can't beat the price.
- R. K. Pathria, *Statistical Mechanics* ($2^{\rm nd}$ edition, Butterworth-Heinemann, 1996)
  This popular graduate level text contains many detailed derivations which are helpful for the student.
- M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* ($3^{\rm rd}$ edition, World Scientific, 2006)
  An excellent graduate level text. Less insightful than Kardar but still a good modern treatment of the subject. Good discussion of mean field theory.
- E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics* (part I, $3^{\rm rd}$ edition, Pergamon, 1980)
  This is volume 5 in the famous Landau and Lifshitz *Course of Theoretical Physics*. Though dated, it still contains a wealth of information and physical insight.

## 5.2 Statistical Mechanics of Noninteracting Quantum Systems

### 5.2.1 Bose and Fermi systems in the grand canonical ensemble

A noninteracting many-particle quantum Hamiltonian may be written as[1]

$$\hat{H} = \sum_\alpha \varepsilon_\alpha \, \hat{n}_\alpha \quad , \tag{5.1}$$

where $\hat{n}_\alpha$ is the number of particles in the quantum state $\alpha$ with energy $\varepsilon_\alpha$. This form is called the *second quantized representation* of the Hamiltonian. The number eigenbasis is therefore also an energy eigenbasis. Any eigenstate of $\hat{H}$ may be labeled by the integer eigenvalues of the $\hat{n}_\alpha$ number operators, and written as $|\, n_1 \,,\, n_2 \,,\, \ldots \,\rangle$. We then have

$$\hat{n}_\alpha \, |\, \boldsymbol{n} \,\rangle = n_\alpha \, |\, \boldsymbol{n} \,\rangle \tag{5.2}$$

and

$$\hat{H} \, |\, \boldsymbol{n} \,\rangle = \sum_\alpha n_\alpha \, \varepsilon_\alpha \, |\, \boldsymbol{n} \,\rangle \quad . \tag{5.3}$$

The eigenvalues $n_\alpha$ take on different possible values depending on whether the constituent particles are *bosons* or *fermions*, *viz.*

$$\begin{aligned} \text{bosons} &: n_\alpha \in \big\{ 0 \,,\, 1 \,,\, 2 \,,\, 3 \,,\, \ldots \big\} \\ \text{fermions} &: n_\alpha \in \big\{ 0 \,,\, 1 \big\} \quad . \end{aligned} \tag{5.4}$$

In other words, for bosons, the occupation numbers are nonnegative integers. For fermions, the occupation numbers are either 0 or 1 due to the *Pauli principle*, which says that at most one fermion can occupy any single particle quantum state. There is no Pauli principle for bosons.

The $N$-particle partition function $Z_N$ is then

$$Z_N = \sum_{\{n_\alpha\}} e^{-\beta \sum_\alpha n_\alpha \varepsilon_\alpha} \, \delta_{N \,, \sum_\alpha n_\alpha} \quad , \tag{5.5}$$

where the sum is over all allowed values of the set $\{n_\alpha\}$, which depends on the *statistics* of the particles. Bosons satisfy *Bose-Einstein* (BE) statistics, in which $n_\alpha \in \{0 \,,\, 1 \,,\, 2 \,,\, \ldots\}$. Fermions satisfy *Fermi-Dirac* (FD) statistics, in which $n_\alpha \in \{0 \,, 1\}$.

The OCE partition sum is difficult to perform, owing to the constraint $\sum_\alpha n_\alpha = N$ on the total number of particles. This constraint is relaxed in the GCE, where

$$\begin{aligned} \Xi &= \sum_N e^{\beta \mu N} \, Z_N \\ &= \sum_{\{n_\alpha\}} e^{-\beta \sum_\alpha n_\alpha \varepsilon_\alpha} \, e^{\beta \mu \sum_\alpha n_\alpha} = \prod_\alpha \left( \sum_{n_\alpha} e^{-\beta (\varepsilon_\alpha - \mu) \, n_\alpha} \right) \quad . \end{aligned} \tag{5.6}$$

[1] For a review of the formalism of second quantization, see the appendix in §5.8.

Note that the grand partition function $\Xi = \exp(-\Omega/k_{\rm B}T)$ takes the form of a product over contributions from the individual single particle states.

We now perform the single particle sums:

$$\begin{aligned}\sum_{n=0}^{\infty} e^{-\beta(\varepsilon-\mu)\,n} &= \frac{1}{1-e^{-\beta(\varepsilon-\mu)}} \qquad \text{(bosons)} \\ \sum_{n=0}^{1} e^{-\beta(\varepsilon-\mu)\,n} &= 1 + e^{-\beta(\varepsilon-\mu)} \qquad \text{(fermions)} \quad .\end{aligned} \tag{5.7}$$

Therefore we have[2]

$$\begin{aligned}\Xi^{\sf BE} &= \prod_\alpha \frac{1}{1-e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T}} \\ \Omega^{\sf BE} &= -k_{\rm B}T\log\Xi^{\sf BE} = k_{\rm B}T\sum_\alpha \log\Big(1-e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T}\Big)\end{aligned} \tag{5.8}$$

for Bose-Einstein statistics and

$$\begin{aligned}\Xi^{\sf FD} &= \prod_\alpha \Big(1+e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T}\Big) \\ \Omega^{\sf FD} &= -k_{\rm B}T\log\Xi^{\sf FD} = -k_{\rm B}T\sum_\alpha \log\Big(1+e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T}\Big)\end{aligned} \tag{5.9}$$

for Fermi-Dirac statistics. We can combine these expressions into one, writing

$$\Omega^{\sf B/F}(T,V,\mu) = \pm k_{\rm B}T\sum_\alpha \log\Big(1\mp e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T}\Big) \quad , \tag{5.10}$$

where we take the upper sign for Bose-Einstein statistics and the lower sign for Fermi-Dirac statistics. Note that the average occupancy of single particle state $\alpha$ is

$$\langle\hat n_\alpha\rangle = \frac{\partial\Omega^{\sf B/F}}{\partial\varepsilon_\alpha} = \frac{1}{e^{(\varepsilon_\alpha-\mu)/k_{\rm B}T}\mp 1} \qquad , \qquad n_\varepsilon^{\sf B/F}(T,\mu) = \frac{1}{e^{(\varepsilon-\mu)/k_{\rm B}T}\mp 1} \quad . \tag{5.11}$$

Thus $n_\varepsilon^{\sf B/F}(T,\mu)$ is the average occupation of single particle states of energy $\varepsilon$ for bosons (B) and fermions (F) at temperature $T$ and chemical potential $\mu$. The total particle number is then

$$N^{\sf B/F}(T,V,\mu) = \sum_\alpha \frac{1}{e^{(\varepsilon_\alpha-\mu)/k_{\rm B}T}\mp 1} \quad . \tag{5.12}$$

We will henceforth write $n_\alpha(\mu,T) = \langle\hat n_\alpha\rangle$ for the thermodynamic average of this occupancy.

[2]Note that convergence of the partition sum for bosons requires $\exp((\mu-\varepsilon_\alpha)/k_{\rm B}T) < 1$, which is to say that $\mu < \min_\alpha \varepsilon_\alpha$ for all $\alpha$.

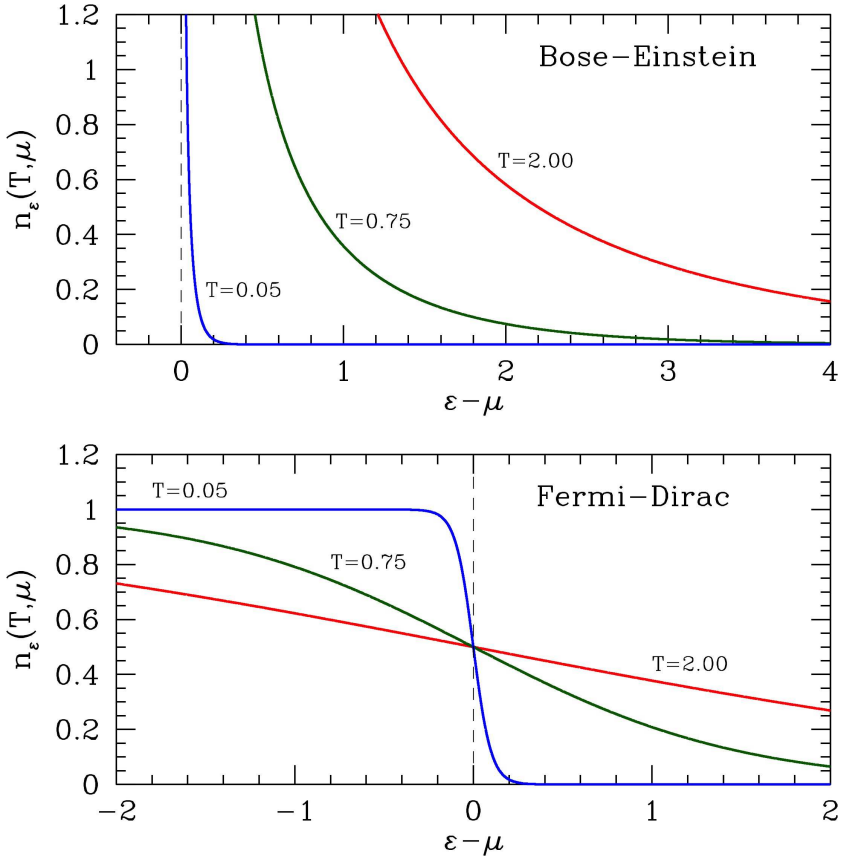



Figure 5.1: Occupation number *versus* $\varepsilon - \mu$ for Bose and Fermi distributions.

### 5.2.2 Quantum statistics and the Maxwell-Boltzmann limit

Consider a system composed of $N$ noninteracting particles. The Hamiltonian is

$$\hat{H} = \sum_{j=1}^{N} \hat{h}_j \quad . \tag{5.13}$$

The single particle Hamiltonian $\hat{h}$ has eigenstates $|\alpha\rangle$ with corresponding energy eigenvalues $\varepsilon_\alpha$. What is the partition function? Is it

$$Z \stackrel{?}{=} \sum_{\alpha_1} \cdots \sum_{\alpha_N} e^{-\beta(\varepsilon_{\alpha_1} + \varepsilon_{\alpha_2} + \ldots + \varepsilon_{\alpha_N})} = \zeta^N \quad , \tag{5.14}$$

where $\zeta$ is the single particle partition function, $\zeta = \sum_\alpha e^{-\beta\varepsilon_\alpha}$ . For systems where the individual particles are *distinguishable*, such as spins on a lattice which have fixed positions, this is indeed correct. But for particles free to move in a gas, this equation is *wrong*. The reason is that for *indistinguishable particles* the many particle quantum mechanical states are specified by a collection of *occupation numbers* $n_\alpha$, which tell us how many particles are in the single-particle state $|\,\alpha\,\rangle$. The energy is $E = \sum_\alpha n_\alpha \varepsilon_\alpha$ and

the total number of particles is $N = \sum_\alpha n_\alpha$ . That is, each collection of occupation numbers $\{n_\alpha\}$ labels a unique many particle state $|\,\{n_\alpha\}\,\rangle$. In the product $\zeta^N$, the collection $\{n_\alpha\}$ occurs many times. We have therefore *overcounted* the contribution to $Z_N$ due to this state. By what factor have we overcounted? It is easy to see that the overcounting factor is

$$\text{degree of overcounting} \quad = \quad \frac{N!}{\prod_\alpha n_\alpha!} \qquad ,$$

which is the number of ways we can rearrange the labels $\alpha_j$ to arrive at the same collection $\{n_\alpha\}$. This follows from the multinomial theorem,

$$\left(\sum_{\alpha=1}^{K} x_\alpha\right)^N = \sum_{n_1}\sum_{n_2}\cdots\sum_{n_K} \frac{N!}{n_1!\,n_2!\cdots n_K!}\, x_1^{n_1}\, x_2^{n_2}\cdots x_K^{n_K}\, \delta_{N,n_1+\ldots+n_K} \quad . \tag{5.15}$$

Thus, the correct expression for $Z_N$ is

$$\begin{aligned} Z_N &= \sum_{\{n_\alpha\}} e^{-\beta\sum_\alpha n_\alpha\varepsilon_\alpha}\,\delta_{N,\sum_\alpha n_\alpha} \\ &= \sum_{\alpha_1}\sum_{\alpha_2}\cdots\sum_{\alpha_N}\left(\frac{\prod_\alpha n_\alpha!}{N!}\right)\, e^{-\beta(\varepsilon_{\alpha_1}+\varepsilon_{\alpha_2}+\,\ldots\,+\varepsilon_{\alpha_N})} \quad . \end{aligned} \tag{5.16}$$

In the high temperature limit, almost all the $n_\alpha$ are either 0 or 1, hence $Z_N \approx \zeta^N/N!$ . This is the classical *Maxwell-Boltzmann limit* of quantum statistical mechanics. We now see the origin of the $1/N!$ term which is so important in the thermodynamics of entropy of mixing.

Finally, starting with the expressions for the grand partition function for Bose-Einstein or Fermi-Dirac particles, and working in the low density limit where $n_\alpha(\mu,T) \ll 1$ , we have $\varepsilon_\alpha - \mu \gg k_{\rm B}T$, which is to say $\mu \ll \varepsilon_0 - k_{\rm B}T$ where $\varepsilon_0$ is the minimum single particle energy. Consequently,

$$\begin{aligned} \Omega^{\rm B/F} &= \pm k_{\rm B}T\sum_\alpha \log\Big(1 \mp e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T}\Big) \\ &\approx -k_{\rm B}T\sum_\alpha e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T} \equiv \Omega^{\rm MB} \quad . \end{aligned} \tag{5.17}$$

This is the Maxwell-Boltzmann limit of quantum statistical mechanics. The occupation number average in the Maxwell-Boltzmann limit is then

$$\langle\hat{n}_\alpha\rangle = e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T} \quad . \tag{5.18}$$

### 5.2.3 Entropy and counting states

Suppose we are to partition $N$ particles among $J$ possible distinct single particle states. How many ways $\Omega$ are there of accomplishing this task? The answer depends on the statistics of the particles. If the particles are fermions, the answer is easy: $\Omega_{\rm FD} = \binom{J}{N}$. For bosons, the number of possible partitions can be evaluated via the following argument. Imagine that we line up all the $N$ particles in a row, and we

place $J-1$ barriers among the particles, as shown below in fig. 5.2. The number of partitions is then the total number of ways of placing the $N$ particles among these $N+J-1$ objects (particles plus barriers), hence we have $\Omega_{\rm BE} = \binom{N+J-1}{N}$. For Maxwell-Boltzmann statistics, we take $\Omega_{\rm MB} = J^N/N!$ Note that $\Omega_{\rm MB}$ is not necessarily an integer, so Maxwell-Boltzmann statistics does not represent any actual state counting. Rather, it manifests itself as a common limit of the Bose and Fermi distributions, as we have seen and shall see again shortly.

The entropy in each case is simply $S = k_{\rm B} \log \Omega$. We assume $N \gg 1$ and $J \gg 1$, with $n \equiv N/J$ finite. Then using Stirling's approximation, $\log(K!) = K \log K - K + \mathcal{O}(\log K)$, we have

$$\begin{aligned} S^{\rm MB} &= -Jk_{\rm B}\big[n \log n - n\big] \\ S^{\rm BE} &= -Jk_{\rm B}\big[n \log n - (1+n)\log(1+n)\big] \\ S^{\rm F} &= -Jk_{\rm B}\big[n \log n + (1-n)\log(1-n)\big] \quad . \end{aligned} \tag{5.19}$$

In the Maxwell-Boltzmann limit, $n \ll 1$, and all three expressions agree. Note that

$$\left(\frac{\partial S^{\rm MB}}{\partial N}\right)_J = -k_{\rm B} \log n \quad , \quad \left(\frac{\partial S^{\rm BE}}{\partial N}\right)_J = k_{\rm B} \log\big(n^{-1}+1\big) \quad , \quad \left(\frac{\partial S^{\rm F}}{\partial N}\right)_J = k_{\rm B} \log\big(n^{-1}-1\big) \quad . \tag{5.20}$$

Now let's imagine grouping the single particle spectrum into intervals of $J$ consecutive energy states. If $J$ is finite and the spectrum is continuous and we are in the thermodynamic limit, then these states will all be approximately degenerate over a sufficiently narrow energy interval. Therefore, using $\alpha$ as a label for the energies, we have that the grand potential $\Omega = E - TS - \mu N$ is given in each case by

$$\begin{aligned} \Omega^{\rm MB} &= J \sum_\alpha \Big[(\varepsilon_\alpha - \mu)\, n_\alpha + k_{\rm B}T(n_\alpha \log n_\alpha - n_\alpha)\Big] \\ \Omega^{\rm BE} &= J \sum_\alpha \Big[(\varepsilon_\alpha - \mu)\, n_\alpha + k_{\rm B}T\, n_\alpha \log n_\alpha - k_{\rm B}T\,(1+n_\alpha)\log(1+n_\alpha)\Big] \\ \Omega^{\rm FD} &= J \sum_\alpha \Big[(\varepsilon_\alpha - \mu)\, n_\alpha + k_{\rm B}T\, n_\alpha \log n_\alpha + k_{\rm B}T\,(1-n_\alpha)\log(1-n_\alpha)\Big] \quad . \end{aligned} \tag{5.21}$$

Now - *lo and behold!* - treating $\Omega$ as a function of the distribution $\{n_\alpha\}$ and extremizing in each case, one obtains the Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac distributions, respectively:

$$\frac{\delta\Omega}{\delta n_\alpha} = 0 \quad \Rightarrow \quad \begin{cases} n_\alpha^{\rm MB} = e^{(\mu-\varepsilon_\alpha)/k_{\rm B}T} \\ n_\alpha^{\rm B} = \big[e^{(\varepsilon_\alpha-\mu)/k_{\rm B}T} - 1\big]^{-1} \\ n_\alpha^{\rm F} = \big[e^{(\varepsilon_\alpha-\mu)/k_{\rm B}T} + 1\big]^{-1} \quad . \end{cases} \tag{5.22}$$

As long as $J$ is finite, so the states in each block remain approximately degenerate, the results are independent of $J$.



Figure 5.2: Partitioning $N$ bosons into $J$ possible states ($N = 14$ and $J = 5$ shown). The $N$ black dots represent bosons, while the $J-1$ white dots represent markers separating the different single particle populations. Here $n_1 = 3$, $n_2 = 1$, $n_3 = 4$, $n_4 = 2$, and $n_5 = 4$.

### 5.2.4 Single particle density of states

The *single particle density of states per unit volume* $g(\varepsilon)$ is defined as

$$g(\varepsilon) = \frac{1}{V}\sum_\alpha \delta(\varepsilon - \varepsilon_\alpha) \quad . \tag{5.23}$$

The concept of density of states is an important one and the student should develop some facility with it. Note that the dimensions of $g(\varepsilon)$ are $\big[g(\varepsilon)\big] = E^{-1}L^{-d}$, where $E$ stands for energy and $L$ for length. We may now write

$$\Omega^{\text{B/F}}(T,V,\mu) = \pm V k_{\text{B}} T \int\limits_{-\infty}^{\infty} d\varepsilon\; g(\varepsilon)\; \log\Big(1 \mp e^{-(\varepsilon-\mu)/k_{\text{B}}T}\Big) \quad . \tag{5.24}$$

For particles with a dispersion $\varepsilon(\boldsymbol{k})$, with $\boldsymbol{p} = \hbar\boldsymbol{k}$, we have

$$g(\varepsilon) = \mathsf{g}\int \frac{d^dk}{(2\pi)^d}\,\delta\big(\varepsilon - \varepsilon(\boldsymbol{k})\big) = \frac{\mathsf{g}\,\Omega_d}{(2\pi)^d}\,\frac{k^{d-1}}{d\varepsilon/dk} \quad , \tag{5.25}$$

where $\mathsf{g} = 2S+1$ is the spin degeneracy, and where we assume that $\varepsilon(\boldsymbol{k})$ is both isotropic and a monotonically increasing function of $k$. Thus, we have

$$g_{d=1}(\varepsilon) = \frac{\mathsf{g}}{\pi}\,\frac{dk}{d\varepsilon} \qquad , \qquad g_{d=2}(\varepsilon) = \frac{\mathsf{g}}{2\pi}\,k\,\frac{dk}{d\varepsilon} \qquad , \qquad g_{d=3}(\varepsilon) = \frac{\mathsf{g}}{2\pi^2}\,k^2\,\frac{dk}{d\varepsilon} \quad . \tag{5.26}$$

In order to obtain $g(\varepsilon)$ as a function of the energy $\varepsilon$ one must invert the dispersion relation $\varepsilon = \varepsilon(k)$ to obtain $k = k(\varepsilon)$. A quick way to derive the result in eqn. 5.25 is to write

$$g(\varepsilon)\,d\varepsilon = \mathsf{g}\,\frac{d^dk}{(2\pi)^d} = \frac{\mathsf{g}\,\Omega_d}{(2\pi)^d}\,k^{d-1}\,dk \quad . \tag{5.27}$$

For a general isotropic power law dispersion $\varepsilon(\boldsymbol{k}) = A\,|\boldsymbol{k}|^\sigma$, this yields

$$\varepsilon(\boldsymbol{k}) = A\,k^\sigma \qquad \Rightarrow \qquad g(\varepsilon) = \frac{\mathsf{g}\,\Omega_d}{(2\pi)^d\sigma}\,A^{-d/\sigma}\,\varepsilon^{\frac{d}{\sigma}-1}\,\Theta(\varepsilon) \quad . \tag{5.28}$$

For a ballistic dispersion $\varepsilon(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m$, we have[3]

$$g(\varepsilon) = \frac{\mathsf{g}}{\Gamma(d/2)}\left(\frac{m}{2\pi\hbar^2}\right)^{d/2}\varepsilon^{(d-2)/2}\,\Theta(\varepsilon) \quad , \tag{5.29}$$

where $\Theta(\varepsilon)$ is the step function, which takes the value 0 for $\varepsilon < 0$ and 1 for $\varepsilon \geq 0$. The appearance of $\Theta(\varepsilon)$ simply says that all the single particle energy eigenvalues are nonnegative.

[3]Throughout, we assume a box of volume $V$, but we assume that the separation $\delta\varepsilon$ between successive quantized single particle energy eigenvalues is negligible so that $g(\varepsilon)$ can be replaced by its energy average over a small window $\Delta\varepsilon \gg \delta\varepsilon$.

Note that

$$n_\varepsilon^{\text{B/F}}(T,\mu) = \frac{1}{e^{(\varepsilon-\mu)/k_{\text{B}}T} \mp 1} \quad . \tag{5.30}$$

This result is valid independent of the form of $g(\varepsilon)$. The average total number of particles is then

$$N^{\text{B/F}}(T,V,\mu) = V \int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, \overbrace{\frac{1}{e^{(\varepsilon-\mu)/k_{\text{B}}T} \mp 1}}^{n_\varepsilon^{\text{B/F}}(T,\mu)} \quad , \tag{5.31}$$

which does depend on $g(\varepsilon)$.

## 5.3 Quantum Ideal Gases : Low Density Expansions

### 5.3.1 Expansion in powers of the fugacity

From eqn. 5.31, we have that the number density $n = N/V$ is

$$n(T,z) = \int\limits_{-\infty}^{\infty} d\varepsilon\, \frac{g(\varepsilon)}{z^{-1}\, e^{\varepsilon/k_{\text{B}}T} \mp 1} = \sum_{j=1}^{\infty} C_j(T)\, z^j \quad , \tag{5.32}$$

where $z = \exp(\mu/k_{\text{B}}T)$ is the *fugacity* and

$$C_j(T) = (\pm 1)^{j-1} \int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, e^{-j\varepsilon/k_{\text{B}}T} \quad . \tag{5.33}$$

Note that $[C_j] = V^{-1}$ for all $j$. From $\Omega = -pV$ and our expression above for $\Omega(T,V,\mu)$, we have

$$p(T,z) = \mp k_{\text{B}}T \int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, \log\Big(1 \mp z\, e^{-\varepsilon/k_{\text{B}}T}\Big) = \int\limits_{-\infty}^{\infty} d\varepsilon\, \frac{H(\varepsilon)}{z^{-1}\, e^{\varepsilon/k_{\text{B}}T} \mp 1} \quad , \tag{5.34}$$

where in the second expression we have integrated by parts, defining

$$H(\varepsilon) = \int\limits_{-\infty}^{\varepsilon} d\varepsilon'\, g(\varepsilon') \quad , \tag{5.35}$$

which is the integral of the density of states, *i.e.* $g(\varepsilon) = H'(\varepsilon)$. We then obtain

$$p(T,z) = k_{\text{B}}T \sum_{j=1}^{\infty} j^{-1} C_j(T)\, z^j \quad . \tag{5.36}$$

### 5.3.2 Virial expansion of the equation of state

Eqns. 5.32 and 5.36 express $n(T,z)$ and $p(T,z)$ as power series in the fugacity $z$, with $T$-dependent coefficients. In principle, we can eliminate $z$ using eqn. 5.32, writing $z = z(T,n)$ as a power series in the number density $n$, and substitute this into eqn. 5.36 to obtain an equation of state $p = p(T,n)$ of the form

$$p(T,n) = n\,k_{\rm B}T\left(1 + B_2(T)\,n + B_3(T)\,n^2 + \ldots\right) \quad . \tag{5.37}$$

Note that the low density limit $n \to 0$ yields the ideal gas law independent of the density of states $g(\varepsilon)$. This follows from expanding $n(T,z)$ and $p(T,z)$ to lowest order in $z$, yielding $n = C_1\,z + \mathcal{O}(z^2)$ and $p = k_{\rm B}T\,C_1\,z + \mathcal{O}(z^2)$. Dividing the second of these equations by the first yields $p = n\,k_{\rm B}T + \mathcal{O}(n^2)$, which is the ideal gas law. Note that $z = n/C_1 + \mathcal{O}(n^2)$ can formally be written as a power series in $n$.

There is a wonderful `Mathematica` function called `InverseSeries` which tells us the following. If

$$n(z) = C_1 z + C_2 z^2 + C_3 z^3 + C_4 z^4 + \mathcal{O}(z^5) \quad , \tag{5.38}$$

then

$$z(n) = \frac{n}{C_1} - \frac{C_2\,n^2}{C_1^3} + \frac{(2\,C_2^2 - C_1C_3)\,n^3}{C_1^5} - \frac{(5\,C_2^3 - 5\,C_1C_2C_3 + C_1^2C_4)\,n^4}{C_1^7} + \mathcal{O}(n^5) \quad , \tag{5.39}$$

and

$$\begin{aligned} y(z) &= D_1 z + D_2 z^2 + D_3 z^3 + D_4 z^4 + \mathcal{O}(z^5) \\ &= \frac{D_1\,n}{C_1} + \frac{(C_1D_2 - C_2D_1)\,n^2}{C_1^3} + \frac{(C_1^2D_3 - 2\,C_1C_2D_2 - C_1C_3D_1 + 2\,C_2^2D_1)\,n^3}{C_1^5} + \mathcal{O}(n^4) \quad . \end{aligned} \tag{5.40}$$

The order $n^4$ term is already very cumbersome and is omitted in the last line above.

Unfortunately, there is no general analytic expression for the virial coefficients $B_j(T)$ in terms of the expansion coefficients $n_j(T)$. However our work is made somewhat easier by appealing to a method of Lagrange. We regard the series

$$\pi(T,n) \equiv \frac{p(T,n)}{k_{\rm B}T} = \sum_{k=1}^{\infty} B_k\,n^k \tag{5.41}$$

as a power series in a *complex* variable $n$. We then have

$$B_k = \oint \frac{dn}{2\pi i}\,\frac{\pi(n)}{n^{k+1}} = \oint \frac{dz}{2\pi i}\,\frac{n'(z)\,\pi(z)}{\big[n(z)\big]^{k+1}} = -\frac{1}{k}\oint \frac{dz}{2\pi i}\,\pi(z)\,\frac{d}{dz}\big[n(z)\big]^{-k} \quad , \tag{5.42}$$

where the contour encloses the origin in the complex plane. Integrating by parts, and using the relation $z\,\pi'(z) = n(z)$, we obtain[4]

$$\begin{aligned} B_k &= \frac{1}{k}\oint \frac{dz}{2\pi i}\,\pi'(z)\,\big[n(z)\big]^{-k} = \frac{1}{k}\oint \frac{dz}{2\pi i}\,\frac{1}{z}\,\big[n(z)\big]^{-(k-1)} \\ &= \frac{1}{k\,C_1^{k-1}}\oint \frac{dz}{2\pi i}\,\frac{1}{z^k}\left(1 + \frac{C_2}{C_1}\,z + \frac{C_3}{C_1}\,z^2 + \ldots\right)^{1-k} \quad . \end{aligned} \tag{5.43}$$

[4]Since there is no term proportional to $\log z$ in the Laurent expansion of $\pi(z)\big[n(z)\big]^{-k}$, there is no residue arising from integrating its derivative around the unit circle.

Expanding by hand to order $z^2$ isn't so difficult, and we obtain

$$\left(1+\frac{C_2}{C_1}\,z+\frac{C_3}{C_1}\,z^2+\ldots\right)^{1-k}=1-(k-1)\,\frac{C_2}{C_1}\,z+(k-1)\left(\frac{k\,C_2^2}{2\,C_1^2}-\frac{C_3}{C_1}\right)z^2+\mathcal{O}(z^3)\quad . \tag{5.44}$$

Plugging this into eqn. 5.43, we may read off

$$B_1=1\qquad,\qquad B_2=-\frac{C_2}{2\,C_1^2}\qquad,\qquad B_3=\frac{C_2^2}{C_1^4}-\frac{2\,C_3}{3\,C_1^3}\quad . \tag{5.45}$$

It is easy to see that, in general, $B_j^{\rm F}=(-1)^{j-1}B_j^{\rm B}$, where the superscripts denote Fermi (F) or Bose (B) statistics.

We remark that the equation of state for classical (and quantum) *interacting* systems also can be expanded in terms of virial coefficients. Consider, for example, the van der Waals equation of state,

$$\left(p+\frac{aN^2}{V^2}\right)(V-Nb)=Nk_{\rm B}T\quad . \tag{5.46}$$

This may be recast as

$$\begin{aligned}p&=\frac{nk_{\rm B}T}{1-bn}-an^2\\&=nk_{\rm B}T+\left(b\,k_{\rm B}T-a\right)n^2+k_{\rm B}T\,b^2n^3+k_{\rm B}T\,b^3n^4+\ldots\quad ,\end{aligned} \tag{5.47}$$

where $n=N/V$. Thus, for the van der Waals system, we have $B_2=b-\frac{a}{k_{\rm B}T}$ and $B_k=b^{k-1}$ for all $k\geq 3$.

### 5.3.3 Ballistic dispersion

For the ballistic dispersion $\varepsilon(\boldsymbol{p})=\boldsymbol{p}^2/2m$ we computed the density of states in eqn. 5.29. One finds

$$C_j(T)=(\pm1)^{j-1}\frac{\mathsf{g}\,\lambda_T^{-d}}{\Gamma(d/2)}\int\limits_0^\infty dt\,t^{\frac{d}{2}-1}\,e^{-jt}=(\pm1)^{j-1}j^{-d/2}\,\mathsf{g}\,\lambda_T^{-d}\quad , \tag{5.48}$$

where $\lambda_T=\sqrt{2\pi\hbar^2/mk_{\rm B}T}$ is the thermal wavelength. We then have

$$\begin{aligned}B_2(T)&=\mp\,2^{-\left(\frac{d}{2}+1\right)}\cdot\mathsf{g}^{-1}\,\lambda_T^d\\B_3(T)&=\left(2^{-(d+1)}-3^{-\left(\frac{d}{2}+1\right)}\right)\cdot 2\,\mathsf{g}^{-2}\,\lambda_T^{2d}\quad .\end{aligned} \tag{5.49}$$

Note that $B_2(T)$ is negative for bosons and positive for fermions. This is because bosons have a tendency to bunch and under certain circumstances may exhibit a phenomenon known as *Bose-Einstein condensation* (BEC). Fermions, on the other hand, obey the Pauli principle, which results in an extra positive correction to the pressure in the low density limit.

We may also write

$$n(T,z)=\pm\mathsf{g}\,\lambda_T^{-d}\;\mathsf{Li}_{\frac{d}{2}}(\pm z) \tag{5.50}$$

and

$$p(T,z) = \pm \mathsf{g}\, k_{\rm B} T\, \lambda_T^{-d}\ \mathsf{Li}_{\frac{d}{2}+1}(\pm z) \quad , \tag{5.51}$$

where

$$\mathsf{Li}_s(z) \equiv \sum_{n=1}^{\infty} \frac{z^n}{n^s} \tag{5.52}$$

is the *polylogarithm function*[5]. Note that $\mathsf{Li}_s(z)$ obeys a recursion relation in its index, *viz.*

$$z\,\frac{\partial}{\partial z}\,\mathsf{Li}_s(z) = \mathsf{Li}_{s-1}(z) \quad , \tag{5.53}$$

and that

$$\zeta(s) = \mathsf{Li}_s(1) = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{5.54}$$

is the Riemann zeta function.

To evaluate $\mathsf{Li}_s(z)$ for $|z| \ll 1$, we use the series expansion in eqn. 5.52. For $|z| \gg 1$, use[6]

$$\mathsf{Li}_s(z) = \sum_{j=0}^{\infty} (-1)^j \big(1 - 2^{1-2j}\big) \frac{(2\pi)^{2j}\,\mathsf{B}_{2j}}{(2j)!}\,\frac{\big[\log(-z)\big]^{s-2j}}{\Gamma(1+s-2j)} \quad , \tag{5.55}$$

where $\mathsf{B}_{2j}$ is a Bernoulli number, with $\mathsf{B}_0 = 1$, $\mathsf{B}_2 = \frac{1}{6}$, $\mathsf{B}_4 = -\frac{1}{30}$, $\mathsf{B}_6 = \frac{1}{42}$, *etc.* For intermediate values of $z$, where $|\log z| < 2\pi$, one has[7]

$$\mathsf{Li}_s(z) = \Gamma(1-s)\,(-\log z)^{s-1} + \sum_{k=0}^{\infty} \frac{\zeta(s-k)}{k!}\,(\log z)^k \quad . \tag{5.56}$$

where $s \notin \{1, 2, 3, \ldots\}$.

## 5.4 Photon Statistics

### 5.4.1 Thermodynamics of the photon gas

There exists a certain class of particles, including photons and certain elementary excitations in solids such as phonons (*i.e.* lattice vibrations) and magnons (*i.e.* spin waves) which obey bosonic statistics but with zero chemical potential. This is because their overall number is not conserved (under typical conditions) – photons can be emitted and absorbed by the atoms in the wall of a container, phonon and

[5]Several texts, such as those by Pathria and by Reichl, write $g_s(z)$ for $\mathsf{Li}_s(z)$. I adopt the latter notation since we are already using the symbol $g$ for the density of states function $g(\varepsilon)$ and for the internal degeneracy g.

[6]See the Wikipedia entry on "Polylogarithm."

[7]See *Digital Library of Mathematical Functions* §25.12.12.

magnon number is also not conserved due to various processes, *etc.* In such cases, the free energy attains its minimum value with respect to particle number when

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} = 0 \quad . \tag{5.57}$$

The number distribution, from eqn. 5.11, is then

$$n_\varepsilon(T) = \frac{1}{e^{\varepsilon/k_{\rm B}T} - 1} \quad . \tag{5.58}$$

The grand partition function for a system of particles with $\mu = 0$ is

$$\Omega(T,V) = V k_{\rm B} T \int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, \log\left(1 - e^{-\varepsilon/k_{\rm B}T}\right) \quad , \tag{5.59}$$

where $g(\varepsilon)$ is the density of states per unit volume.

Suppose a particle with $\mu = 0$ exhibits a power law dispersion $\varepsilon(\boldsymbol{k}) = A\,|\boldsymbol{k}|^\sigma$. We can compute the density of states $g(\varepsilon)$, *viz.*

$$g(\varepsilon) = \frac{{\rm g}\Omega_d}{(2\pi)^d}\,\frac{k^{d-1}}{d\varepsilon/dk} = \left(\frac{{\rm g}\,\Omega_d}{(2\pi)^d\,\sigma A^{d/\sigma}}\right)\varepsilon^{\frac{d}{\sigma}-1}\,\Theta(\varepsilon) \quad . \tag{5.60}$$

Recall that $\Omega_d = 2\pi^{d/2}/\Gamma(d/2)$ is the total solid angle in $d$ dimensions. The step function $\Theta(\varepsilon)$ enforces that the energy spectrum is bounded from below by $\varepsilon = 0$, *i.e.* there are no negative energy states. For a general power law density of states $g(\varepsilon) = C\,\varepsilon^{r-1}\,\Theta(\varepsilon)$ and $\mu = 0$, we have

$$n(T) = C \int\limits_0^\infty d\varepsilon\,\frac{\varepsilon^{r-1}}{e^{\varepsilon/k_{\rm B}T} - 1} = C\,\zeta(r)\,\Gamma(r)\,(k_{\rm B}T)^r \tag{5.61}$$

and

$$p(T) = -C \int\limits_0^\infty d\varepsilon\,\varepsilon^{r-1}\log\left(1 - e^{-\varepsilon/k_{\rm B}T}\right) = C\,\zeta(r+1)\,\Gamma(r)\,(k_{\rm B}T)^{r+1} \quad . \tag{5.62}$$

Note that we require $r > 1$ in order for the particle density to be finite. To work out the above integrals, first change variables to $t = \varepsilon/k_{\rm B}T$, then expand in powers of $\exp(-t)$, integrate over $t$, and express the sums in terms of Riemann zeta (see eqn. 5.54). Dividing these two equations, we obtain the equation of state

$$p = \frac{\zeta(r+1)}{\zeta(r)}\,nk_{\rm B}T \quad . \tag{5.63}$$

To find the entropy, we use Gibbs-Duhem:

$$d\mu = 0 = -s\,dT + v\,dp \quad \Longrightarrow \quad s = v\,\frac{dp}{dT} \quad , \tag{5.64}$$

where $s$ is the entropy per particle and $v = n^{-1}$ is the volume per particle. In $d$ space dimensions,

$$s(T) = (r+1)\,\frac{\zeta(r+1)}{\zeta(r)}\,k_{\rm B} \quad . \tag{5.65}$$

The entropy per particle is constant. The internal energy is

$$E = -\frac{\partial \log \Xi}{\partial \beta} = -\frac{\partial}{\partial \beta}(\beta p V) = rpV \quad , \tag{5.66}$$

and hence the energy per particle is

$$\varepsilon = \frac{E}{N} = rpv = \frac{r\,\zeta(r+1)}{\zeta(r)}\,k_{\rm B}T \quad . \tag{5.67}$$

For the photon gas, we have $\varepsilon(\boldsymbol{k}) = \hbar c k$, whence $\sigma = 1$, $r = d$, $A = \hbar c$, and

$$g(\varepsilon) = \left(\frac{2{\rm g}\,\pi^{d/2}}{\Gamma(d/2)(hc)^d}\right)\varepsilon^{d-1}\,\Theta(\varepsilon) \quad . \tag{5.68}$$

In $d = 3$ dimensions the degeneracy is ${\rm g} = 2$, which is the number of independent polarization states. Thus, for the photon gas in $d = 3$ dimensions, we have ${\rm g} = 2$ and $C = 1/\pi^2\hbar^3c^3$, and

$$n(T) = \frac{2\,\zeta(3)}{\pi^2}\,\frac{(k_{\rm B}T)^3}{(\hbar c)^3} \qquad , \qquad p(T) = \frac{2\,\zeta(4)}{\pi^2}\,\frac{(k_{\rm B}T)^4}{(\hbar c)^3} \quad . \tag{5.69}$$

It turns out that $\zeta(3) = 1.20206$ and $\zeta(4) = \pi^4/90$. We also have $\hbar c/k_{\rm B} = 0.2290\,{\rm cm}\cdot{\rm K}$, and thus $k_{\rm B}T/\hbar c = 4.3755\,T[{\rm K}]\,{\rm cm}^{-1}$. We then obtain

$$p = 0.90039\,nk_{\rm B}T \qquad , \qquad n(T) = 20.286\times T^3[{\rm K}^3]\,{\rm cm}^{-3} \quad . \tag{5.70}$$

### 5.4.2 Classical arguments for the photon gas

A number of thermodynamic properties of the photon gas can be determined from purely classical arguments. Here we recapitulate a few important ones.

1. Suppose our photon gas is confined to a rectangular box of dimensions $L_x \times L_y \times L_z$. Suppose further that the dimensions are all expanded by a factor $\lambda^{1/3}$, *i.e.* the volume is isotropically expanded by a factor of $\lambda$. The cavity modes of the electromagnetic radiation have quantized wavevectors, even within classical electromagnetic theory, given by

$$\boldsymbol{k} = \left(\frac{2\pi n_x}{L_x}\,,\,\frac{2\pi n_y}{L_y}\,,\,\frac{2\pi n_z}{L_z}\right) \quad . \tag{5.71}$$

Since the energy for a given mode is $\varepsilon(\boldsymbol{k}) = \hbar c|\boldsymbol{k}|$, we see that the energy changes by a factor $\lambda^{-1/3}$ under an adiabatic volume expansion $V \to \lambda V$, where the distribution of different electromagnetic mode occupancies remains fixed. Thus,

$$V\left(\frac{\partial E}{\partial V}\right)_S = \lambda\left(\frac{\partial E}{\partial \lambda}\right)_S = -\tfrac{1}{3}E \quad . \tag{5.72}$$

Thus,

$$p = -\left(\frac{\partial E}{\partial V}\right)_S = \frac{E}{3V} \quad , \tag{5.73}$$

as we found in eqn. 5.66. Since $E = E(T, V)$ is extensive, we must have $p = p(T)$ alone.

2. Since $p = p(T)$ alone, we have

$$\begin{aligned}\left(\frac{\partial E}{\partial V}\right)_T &= \left(\frac{\partial E}{\partial V}\right)_p = 3p \\ &= T\left(\frac{\partial p}{\partial T}\right)_V - p \quad ,\end{aligned} \tag{5.74}$$

where the second line follows the Maxwell relation $\left(\frac{\partial S}{\partial V}\right)_p = \left(\frac{\partial p}{\partial T}\right)_V$, after invoking the First Law $dE = T dS - p\, dV$. Thus,

$$T\,\frac{dp}{dT} = 4p \quad \Longrightarrow \quad p(T) = A\,T^4 \quad , \tag{5.75}$$

where $A$ is a constant. Thus, we recover the temperature dependence found microscopically in eqn. 5.62.

3. Given an energy density $E/V$, the differential energy flux emitted in a direction $\theta$ relative to a surface normal is

$$dj_\varepsilon = c \cdot \frac{E}{V} \cdot \cos\theta \cdot \frac{d\Omega}{4\pi} \quad , \tag{5.76}$$

where $d\Omega$ is the differential solid angle. Thus, the power emitted per unit area is

$$\frac{dP}{dA} = \frac{cE}{4\pi V}\int_0^{\pi/2} d\theta \int_0^{2\pi} d\phi\, \sin\theta \cdot \cos\theta = \frac{cE}{4V} = \tfrac{3}{4}\,c\,p(T) \equiv \sigma\,T^4 \quad , \tag{5.77}$$

where $\sigma = \frac{3}{4}cA$, with $p(T) = A\,T^4$ as we found above. From quantum statistical mechanical considerations, we have

$$\sigma = \frac{\pi^2 k_{\rm B}^4}{60\,c^2\,\hbar^3} = 5.67 \times 10^{-8}\,\frac{\rm W}{\rm m^2\,K^4} \tag{5.78}$$

is *Stefan's constant*.

### 5.4.3 Surface temperature of the earth

We derived the result $P = \sigma T^4 \cdot A$ where $\sigma = 5.67 \times 10^{-8}\,{\rm W/m^2\,K^4}$ for the power emitted by an electromagnetic 'black body'. Let's apply this result to the earth-sun system. We'll need three lengths: the radius of the sun $R_\odot = 6.96 \times 10^8\,{\rm m}$, the radius of the earth $R_{\rm e} = 6.38 \times 10^6\,{\rm m}$, and the radius of the earth's orbit $a_{\rm e} = 1.50 \times 10^{11}\,{\rm m}$. Let's assume that the earth has achieved a steady state temperature of $T_{\rm e}$. We balance the total power incident upon the earth with the power radiated by the earth. The power incident upon the earth is

$$P_{\rm incident} = \frac{\pi R_{\rm e}^2}{4\pi a_{\rm e}^2} \cdot \sigma T_\odot^4 \cdot 4\pi R_\odot^2 = \frac{R_{\rm e}^2\,R_\odot^2}{a_{\rm e}^2} \cdot \pi\sigma T_\odot^4 \quad . \tag{5.79}$$

The power radiated by the earth is

$$P_{\rm radiated} = \sigma T_{\rm e}^4 \cdot 4\pi R_{\rm e}^2 \quad . \tag{5.80}$$

Setting $P_{\rm incident} = P_{\rm radiated}$, we obtain

$$T_{\rm e} = \left(\frac{R_\odot}{2\,a_{\rm e}}\right)^{1/2} T_\odot \quad . \tag{5.81}$$

Thus, we find $T_{\rm e} = 0.04817\,T_\odot$, and with $T_\odot = 5780\,{\rm K}$, we obtain $T_{\rm e} = 278.4\,{\rm K}$. The mean surface temperature of the earth is $\bar{T}_{\rm e} = 287\,{\rm K}$, which is only about $10\,{\rm K}$ higher. The difference is due to the fact that the earth is not a perfect blackbody, *i.e.* an object which absorbs all incident radiation upon it and emits radiation according to Stefan's law. As you know, the earth's atmosphere retraps a fraction of the emitted radiation – a phenomenon known as the *greenhouse effect*. The earth also has a nonzero *albedo*,

$$A = \frac{\text{reflected power}}{\text{incident power}} \quad , \tag{5.82}$$

due to ice, snow, clouds, atmospheric particles, and other factors. Fresh snow has a very high albedo of about $0.8$ to $0.9$. Earth's overall albedo is $A_{\rm e} \approx 0.3$. As a consequence of earth's nonzero albedo, its effective radiating temperature is about $255\,{\rm K}$. The greenhouse effect then raises the mean surface temperature to approximately $288\,{\rm K}$.

### 5.4.4 Distribution of blackbody radiation

Recall that the frequency of an electromagnetic wave of wavevector $\boldsymbol{k}$ is $\nu = c/\lambda = ck/2\pi$. Therefore the number of photons $\mathcal{N}(\nu, T)$ per unit frequency in thermodynamic equilibrium is (recall there are two polarization states)

$$\mathcal{N}(\nu,T)\,d\nu = \frac{2\,V}{8\pi^3}\cdot\frac{d^3k}{e^{\hbar ck/k_{\rm B}T}-1} = \frac{V}{\pi^2}\cdot\frac{k^2\,dk}{e^{\hbar ck/k_{\rm B}T}-1} \quad . \tag{5.83}$$

We therefore have

$$\mathcal{N}(\nu,T) = \frac{8\pi V}{c^3}\cdot\frac{\nu^2}{e^{h\nu/k_{\rm B}T}-1} \quad . \tag{5.84}$$

Since a photon of frequency $\nu$ carries energy $h\nu$, the energy per unit frequency $\mathcal{E}(\nu)$ is

$$\mathcal{E}(\nu,T) = \frac{8\pi h V}{c^3}\cdot\frac{\nu^3}{e^{h\nu/k_{\rm B}T}-1} \quad . \tag{5.85}$$

Note what happens if Planck's constant $h$ vanishes, as it does in the classical (Maxwell-Boltzmann) limit. The denominator can then be written

$$e^{h\nu/k_{\rm B}T} - 1 = \frac{h\nu}{k_{\rm B}T} + \mathcal{O}(h^2) \tag{5.86}$$

and

$$\mathcal{E}_{\rm MB}(\nu,T) = \lim_{h\to 0}\mathcal{E}(\nu) = V\cdot\frac{8\pi k_{\rm B}T}{c^3}\,\nu^2 \quad . \tag{5.87}$$

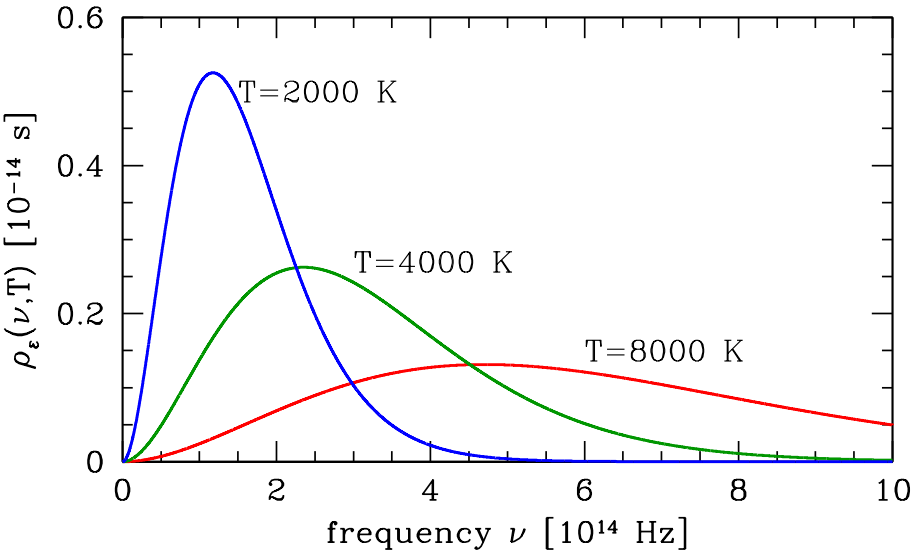



Figure 5.3: Spectral density $\rho_\varepsilon(\nu, T)$ for blackbody radiation at three temperatures.

In classical electromagnetic theory, then, the total energy integrated over all frequencies *diverges*. This is known as the *ultraviolet catastrophe*, since the divergence comes from the large $\nu$ part of the integral, which in the optical spectrum is the ultraviolet portion. With quantization, the Bose-Einstein factor imposes an effective ultraviolet cutoff $k_{\rm B}T/h$ on the frequency integral, and the total energy, as we found above, is finite:

$$E(T) = \int_0^\infty d\nu\, \mathcal{E}(\nu) = 3pV = V \cdot \frac{\pi^2}{15}\,\frac{(k_{\rm B}T)^4}{(\hbar c)^3} \quad . \tag{5.88}$$

We can define the spectral density $\rho_\varepsilon(\nu)$ of the radiation as

$$\rho_\varepsilon(\nu, T) \equiv \frac{\mathcal{E}(\nu, T)}{E(T)} = \frac{15}{\pi^4}\,\frac{h}{k_{\rm B}T}\,\frac{(h\nu/k_{\rm B}T)^3}{e^{h\nu/k_{\rm B}T} - 1} \tag{5.89}$$

so that $\rho_\varepsilon(\nu, T)\, d\nu$ is the fraction of the electromagnetic energy, under equilibrium conditions, between frequencies $\nu$ and $\nu + d\nu$, *i.e.* $\int_0^\infty d\nu\, \rho_\varepsilon(\nu, T) = 1$. In fig. 5.3 we plot this for three different temperatures. The maximum occurs at a frequency $\nu^* = s\, k_{\rm B}T/h$, where

$$\frac{d}{ds}\left(\frac{s^3}{e^s - 1}\right) = 0 \qquad \Longrightarrow \qquad \frac{s}{1 - e^{-s}} = 3 \qquad \Longrightarrow \qquad s = 2.82144 \quad . \tag{5.90}$$

Thus $\nu^* = 58.8\,\text{GHz} \cdot T[\text{K}]$.

### 5.4.5 What if the sun emitted ferromagnetic spin waves?

We saw in eqn. 5.76 that the power emitted per unit surface area by a blackbody is $\sigma T^4$. The power law here follows from the ultrarelativistic dispersion $\varepsilon = \hbar c k$ of the photons. Suppose that we replace this

dispersion with the general form $\varepsilon = \varepsilon(\boldsymbol{k})$. Now consider a large box in equilibrium at temperature $T$. The *energy current* incident on a differential area $dA$ of surface normal to $\hat{\boldsymbol{z}}$ is

$$dP = dA \cdot \int \frac{d^3k}{(2\pi)^3}\,\Theta(\cos\theta)\cdot\varepsilon(\boldsymbol{k})\cdot\frac{1}{\hbar}\frac{\partial\varepsilon(\boldsymbol{k})}{\partial k_z}\cdot\frac{1}{e^{\varepsilon(\boldsymbol{k})/k_{\rm B}T}-1} \quad . \tag{5.91}$$

Let us assume an isotropic power law dispersion of the form $\varepsilon(\boldsymbol{k}) = Ck^\alpha$. Then after a straightforward calculation we obtain

$$\frac{dP}{dA} = \widetilde{\sigma}\,T^{2+2\alpha^{-1}} \quad , \tag{5.92}$$

where

$$\widetilde{\sigma} = \zeta(2+2\alpha^{-1})\,\Gamma(2+2\alpha^{-1})\cdot\frac{\mathsf{g}\,k_{\rm B}^2}{8\pi^2\hbar}\left(\frac{k_{\rm B}}{C}\right)^{2\alpha^{-1}} \quad . \tag{5.93}$$

One can check that for $\mathsf{g} = 2$, $C = \hbar c$, and $\alpha = 1$ that this result reduces to that of eqn. 5.78. For the case of ferromagnetic spin waves, $\alpha = 2$, in which case $P = \widetilde{\sigma}T^3A$. What would be the surface temperature of the earth if the photon dispersion were $\varepsilon = Ck^\alpha$? Generalizing the results from §5.4.3, we find

$$T_{\rm e} = \left(\frac{R_\odot}{2\,a_{\rm e}}\right)^{\alpha/(1+\alpha)} T_\odot \quad . \tag{5.94}$$

With $R_\odot/2a_{\rm e} = 2.32\times10^{-3}$, assuming the same value for $T_\odot = 5780\,{\rm K}$, and with $\alpha = 2$, we obtain $T_{\rm e} = 101\,{\rm K}$.

## 5.5 Lattice Vibrations : Einstein and Debye Models

Crystalline solids support propagating waves called *phonons*, which are quantized vibrations of the lattice. Recall that the quantum mechanical harmonic oscillator Hamiltonian, $\hat{H} = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2q^2$, may be written as $\hat{H} = \hbar\omega_0(a^\dagger a + \frac{1}{2})$, where $a$ and $a^\dagger$ are 'ladder operators' satisfying commutation relations $\left[a\,,a^\dagger\right] = 1$.

### 5.5.1 One-dimensional chain

Consider the linear chain of masses and springs depicted in fig. 5.4. We assume that our system consists of $N$ mass points on a large ring of circumference $L$. In equilibrium, the masses are spaced evenly by a distance $b = L/N$. That is, $x_n^0 = nb$ is the equilibrium position of particle $n$. We define $u_n = x_n - x_n^0$ to be the difference between the position of mass $n$ and its equilibrium position. The Hamiltonian is then

$$\begin{aligned}\hat{H} &= \sum_n\left[\frac{p_n^2}{2m} + \tfrac{1}{2}\kappa\,(x_{n+1}-x_n-a)^2\right] \\ &= \sum_n\left[\frac{p_n^2}{2m} + \tfrac{1}{2}\kappa\,(u_{n+1}-u_n)^2\right] + \tfrac{1}{2}N\kappa(b-a)^2 \quad ,\end{aligned} \tag{5.95}$$

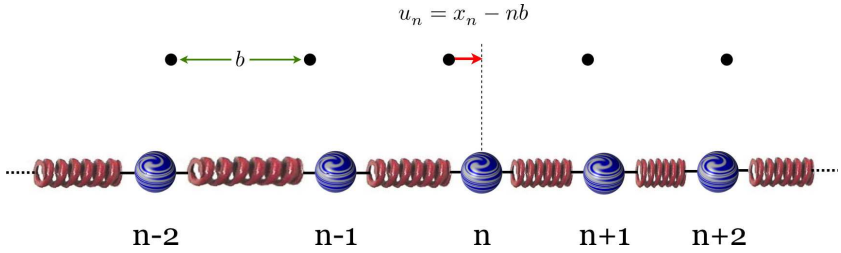



Figure 5.4: A linear chain of masses and springs. The black circles represent the equilibrium positions of the masses. The displacement of mass $n$ relative to its equilibrium value is $u_n$.

where $a$ is the unstretched length of each spring, $m$ is the mass of each mass point, $\kappa$ is the force constant of each spring, and $N$ is the total number of mass points. If $b \neq a$ the springs are under tension in equilibrium, but as we see this only leads to an additive constant in the Hamiltonian, and hence does not enter the equations of motion.

The classical equations of motion are

$$\begin{aligned} \dot{u}_n &= \frac{\partial \hat{H}}{\partial p_n} = \frac{p_n}{m} \\ \dot{p}_n &= -\frac{\partial \hat{H}}{\partial u_n} = \kappa \left( u_{n+1} + u_{n-1} - 2u_n \right) \quad . \end{aligned} \tag{5.96}$$

Taking the time derivative of the first equation and substituting into the second yields

$$\ddot{u}_n = \frac{\kappa}{m} \left( u_{n+1} + u_{n-1} - 2u_n \right) \quad . \tag{5.97}$$

We now write

$$u_n = \frac{1}{\sqrt{N}} \sum_k \hat{u}_k \, e^{iknb} \quad , \tag{5.98}$$

where periodicity $u_{N+n} = u_n$ requires that the $k$ values are quantized so that $e^{ikNb} = 1$, *i.e.* $k = 2\pi j / Nb$ where $j \in \{0, 1, \ldots, N-1\}$. The inverse of this discrete Fourier transform is

$$\hat{u}_k = \frac{1}{\sqrt{N}} \sum_n u_n \, e^{-iknb} \quad . \tag{5.99}$$

Note that $\hat{u}_k$ is in general complex, but that $\hat{u}_k^* = \hat{u}_{-k}$. In terms of the $\hat{u}_k$, the equations of motion take the form

$$\ddot{\hat{u}}_k = -\frac{2\kappa}{m} \left( 1 - \cos(kb) \right) \hat{u}_k \equiv -\omega_k^2 \, \hat{u}_k \quad . \tag{5.100}$$

Thus, each $\hat{u}_k$ is a normal mode, and the normal mode frequencies are

$$\Omega_k = 2 \sqrt{\frac{\kappa}{m}} \left| \sin\left( \tfrac{1}{2} kb \right) \right| \quad . \tag{5.101}$$

The density of states for this band of phonon excitations is

$$g(\varepsilon) = \int\limits_{-\pi/b}^{\pi/b} \frac{dk}{2\pi}\,\delta(\varepsilon - \hbar\Omega_k) = \frac{2}{\pi b}\big(J^2 - \varepsilon^2\big)^{-1/2}\,\Theta(\varepsilon)\,\Theta(J-\varepsilon) \quad , \tag{5.102}$$

where $J = 2\hbar\sqrt{\kappa/m}$ is the phonon bandwidth. The step functions require $0 \le \varepsilon \le J$; outside this range there are no phonon energy levels and the density of states accordingly vanishes.

The entire theory can be quantized, taking $\big[p_n\,,u_{n'}\big] = -i\hbar\,\delta_{nn'}$. We then define

$$p_n = \frac{1}{\sqrt{N}}\sum_k \hat{p}_k\,e^{iknb} \qquad , \qquad \hat{p}_k = \frac{1}{\sqrt{N}}\sum_n p_n\,e^{-iknb} \quad , \tag{5.103}$$

in which case $\big[\hat{p}_k\,,\hat{u}_{k'}\big] = -i\hbar\,\delta_{k+k',0}$. Note that $\hat{u}_k^\dagger = \hat{u}_{-k}$ and $\hat{p}_k^\dagger = \hat{p}_{-k}$. We then define the ladder operator

$$a_k = \left(\frac{1}{2m\hbar\Omega_k}\right)^{1/2}\hat{p}_k - i\left(\frac{m\Omega_k}{2\hbar}\right)^{1/2}\hat{u}_k \tag{5.104}$$

and its Hermitean conjugate $a_k^\dagger$, in terms of which the Hamiltonian is

$$\hat{H} = \sum_k \hbar\Omega_k\,\big(a_k^\dagger a_k + \tfrac{1}{2}\big) \quad , \tag{5.105}$$

which is a sum over independent harmonic oscillator modes. Note that the sum over $k$ is restricted to $k \in \big[-\frac{\pi}{b},\frac{\pi}{b}\big]$, an interval of width $\frac{2\pi}{b}$, known as the *first Brillouin zone* for the one-dimensional lattice. The state at wavevector $k + \frac{2\pi}{b}$ is equivalent to that at $k$, as we see from eqn. 5.99.

### 5.5.2 General theory of lattice vibrations

The most general model of a harmonic solid is described by a Hamiltonian of the form

$$\hat{H} = \sum_{\boldsymbol{R},i}\frac{\boldsymbol{p}_i^2(\boldsymbol{R})}{2M_i} + \frac{1}{2}\sum_{i,j}\sum_{\alpha,\beta}\sum_{\boldsymbol{R},\boldsymbol{R}'} u_i^\alpha(\boldsymbol{R})\,\Phi_{ij}^{\alpha\beta}(\boldsymbol{R}-\boldsymbol{R}')\,u_j^\beta(\boldsymbol{R}') \quad , \tag{5.106}$$

where the *dynamical matrix* is

$$\Phi_{ij}^{\alpha\beta}(\boldsymbol{R}-\boldsymbol{R}') = \frac{\partial^2 U}{\partial u_i^\alpha(\boldsymbol{R})\,\partial u_j^\beta(\boldsymbol{R}')} \quad , \tag{5.107}$$

where $U$ is the potential energy of interaction among all the atoms. Here we have simply expanded the potential to second order in the local displacements $u_i^\alpha(\boldsymbol{R})$. The lattice sites $\boldsymbol{R}$ are elements of a *Bravais lattice*. The indices $i$ and $j$ specify *basis elements* with respect to this lattice, and the indices $\alpha$ and $\beta$ range over $\{1,\ldots,d\}$, the number of possible directions in space. The subject of crystallography is beyond the scope of these notes, but, very briefly, a Bravais lattice in $d$ dimensions is specified by a set of $d$ linearly independent *primitive direct lattice vectors* $\boldsymbol{a}_l$, such that any point in the Bravais lattice may be written as a sum over the primitive vectors with integer coefficients: $\boldsymbol{R} = \sum_{l=1}^d n_l\,\boldsymbol{a}_l$. The set of all such vectors

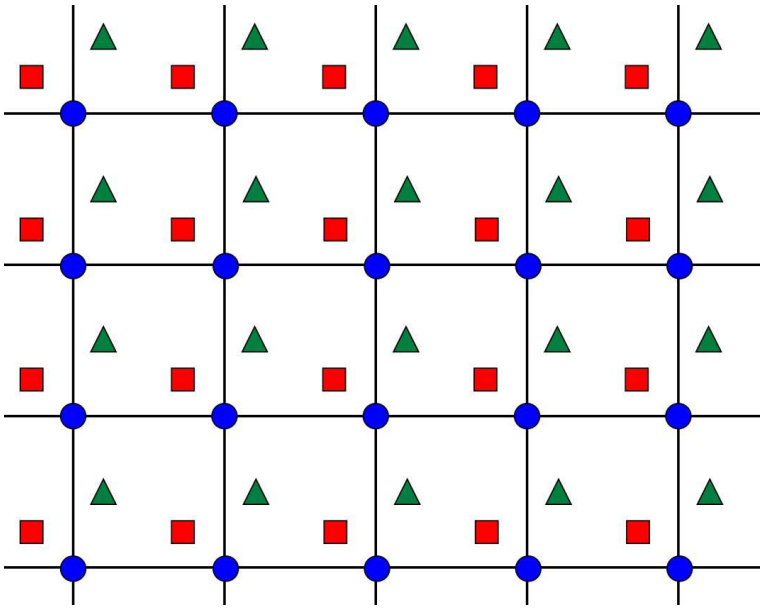


Figure 5.5: A crystal structure with an underlying square Bravais lattice and a three element basis.

$\{\boldsymbol{R}\}$ is called the *direct lattice*. The direct lattice is closed under the operation of vector addition: if $\boldsymbol{R}$ and $\boldsymbol{R}'$ are points in a Bravais lattice, then so is $\boldsymbol{R} + \boldsymbol{R}'$.

A crystal is a periodic arrangement of lattice sites. The fundamental repeating unit is called the *unit cell*. Not every crystal is a Bravais lattice, however. Indeed, Bravais lattices are special crystals in which there is only one atom per unit cell. Consider, for example, the structure in fig. 5.5. The blue dots form a square Bravais lattice with primitive direct lattice vectors $\boldsymbol{a}_1 = a\,\hat{\boldsymbol{x}}$ and $\boldsymbol{a}_2 = a\,\hat{\boldsymbol{y}}$, where $a$ is the *lattice constant*, which is the distance between any neighboring pair of blue dots. The red squares and green triangles, along with the blue dots, form a *basis* for the crystal structure which label each *sublattice*. Our crystal in fig. 5.5 is formally classified as a *square Bravais lattice with a three element basis*. To specify an arbitrary site in the crystal, we must specify both a direct lattice vector $\boldsymbol{R}$ as well as a basis index $j \in \{1, \ldots, r\}$, so that the location is $\boldsymbol{R} + \boldsymbol{\eta}_j$. The vectors $\{\boldsymbol{\eta}_j\}$ are the *basis vectors* for our crystal structure. We see that a general crystal structure consists of a repeating unit, known as a *unit cell*. The centers (or corners, if one prefers) of the unit cells form a Bravais lattice. Within a given unit cell, the individual sublattice sites are located at positions $\boldsymbol{\eta}_j$ with respect to the unit cell position $\boldsymbol{R}$.

Upon diagonalization, the Hamiltonian of eqn. 5.106 takes the form

$$\hat{H} = \sum_{\boldsymbol{k},a} \hbar\Omega_a(\boldsymbol{k}) \left( A_a^\dagger(\boldsymbol{k})\, A_a(\boldsymbol{k}) + \tfrac{1}{2} \right) \quad , \tag{5.108}$$

where

$$\left[ A_a(\boldsymbol{k})\,,\, A_b^\dagger(\boldsymbol{k}') \right] = \delta_{ab}\,\delta_{\boldsymbol{k}\boldsymbol{k}'} \quad . \tag{5.109}$$

The eigenfrequencies are solutions to the eigenvalue equation

$$\sum_{j,\beta} \tilde{\Phi}_{ij}^{\alpha\beta}(\boldsymbol{k})\, \mathsf{e}_{j\beta}^{(a)}(\boldsymbol{k}) = m_i\, \omega_a^2(\boldsymbol{k})\, \mathsf{e}_{i\alpha}^{(a)}(\boldsymbol{k}) \quad , \tag{5.110}$$

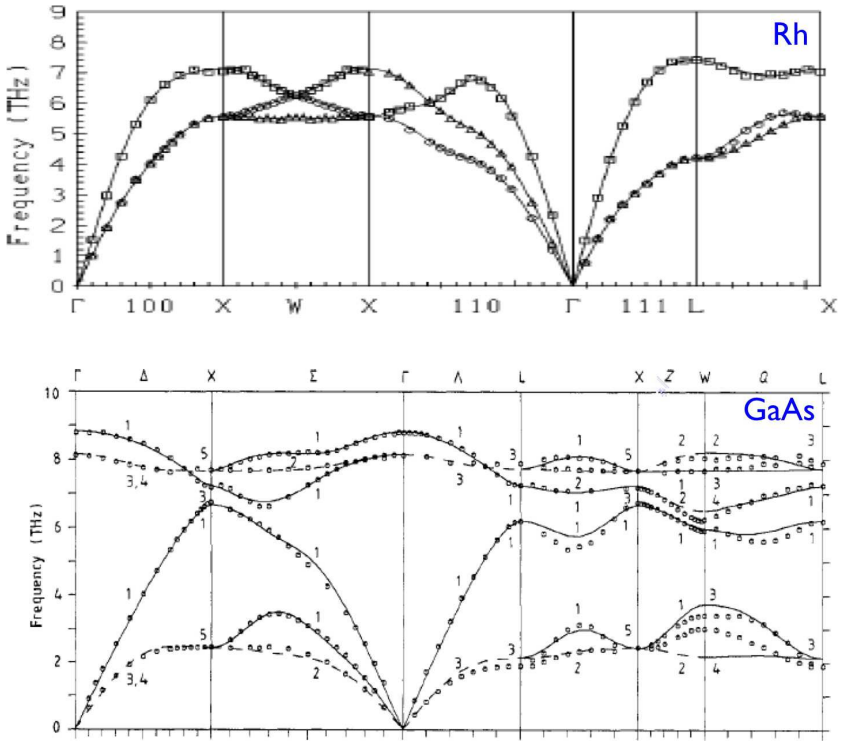



Figure 5.6: Upper panel: phonon spectrum in elemental rhodium (Rh) at $T = 297\,\mathrm{K}$ measured by high precision inelastic neutron scattering (INS) by A. Eichler *et al.*, *Phys. Rev. B* **57**, 324 (1998). Note the three acoustic branches and no optical branches, corresponding to $d = 3$ and $r = 1$. Lower panel: phonon spectrum in gallium arsenide (GaAs) at $T = 12\,\mathrm{K}$, comparing theoretical lattice-dynamical calculations with INS results of D. Strauch and B. Dorner, *J. Phys.: Condens. Matter* **2**, 1457 (1990). Note the three acoustic branches and three optical branches, corresponding to $d = 3$ and $r = 2$. The Greek letters along the $x$-axis indicate points of high symmetry in the Brillouin zone.

where $m_i$ is the mass of the $i^{\rm th}$ basis ion and

$$\tilde\Phi_{ij}^{\alpha\beta}(\boldsymbol{k}) = \sum_{\boldsymbol{R}} \Phi_{ij}^{\alpha\beta}(\boldsymbol{R})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \tag{5.111}$$

is the (discrete) Fourier transform of the dynamical matrix. Here, $\boldsymbol{k}$ lies within the first *Brillouin zone*, which is the unit cell of the *reciprocal lattice* of points $\boldsymbol{G}$ satisfying $e^{i\boldsymbol{G}\cdot\boldsymbol{R}} = 1$ for all $\boldsymbol{G}$ and $\boldsymbol{R}$. The reciprocal lattice is also a Bravais lattice, with primitive reciprocal lattice vectors $\boldsymbol{b}_l$, such that any point on the reciprocal lattice may be written $\boldsymbol{G} = \sum_{l=1}^{d} p_l\,\boldsymbol{b}_l$ with $p_l \in \mathbb{Z}$. One also has that $\boldsymbol{a}_l \cdot \boldsymbol{b}_{l'} = 2\pi\delta_{ll'}$. The index $a$ ranges from 1 to $d \cdot r$ and labels the *mode* of oscillation at wavevector $\boldsymbol{k}$. The vector $\mathrm{e}_{i\alpha}^{(a)}(\boldsymbol{k})$ is the *polarization vector* for the $a^{\rm th}$ phonon branch. In solids, along directions of high symmetry, phonon modes can be classified as longitudinal or transverse excitations.

For a crystalline lattice with an $r$-element basis, there are then $d \cdot r$ phonon modes for each wavevector $\boldsymbol{k}$ lying in the first Brillouin zone. If we impose periodic boundary conditions, then the $\boldsymbol{k}$ points within the first Brillouin zone are themselves quantized, as in the $d = 1$ case where we found $k = 2\pi n/N$. There are $N$ distinct $\boldsymbol{k}$ points in the first Brillouin zone – one for every direct lattice site. The total number

of modes is than $d \cdot r \cdot N$, which is the total number of translational degrees of freedom in our system: $rN$ total atoms ($N$ unit cells each with an $r$ atom basis) each free to vibrate in $d$ dimensions. Of the $d \cdot r$ branches of phonon excitations, $d$ of them will be *acoustic modes* whose frequency vanishes as $\boldsymbol{k} \to 0$. The remaining $d(r-1)$ branches are *optical modes* and oscillate at finite frequencies. Basically, in an acoustic mode, for $\boldsymbol{k}$ close to the (Brillouin) zone center $\boldsymbol{k} = 0$, all the atoms in each unit cell move together in the same direction at any moment of time. In an optical mode, the different basis atoms move in different directions.

There is no number conservation law for phonons – they may be freely created or destroyed in anharmonic processes, where two photons with wavevectors $\boldsymbol{k}$ and $\boldsymbol{q}$ can combine into a single phonon with wavevector $\boldsymbol{k} + \boldsymbol{q}$, and *vice versa*. Therefore the chemical potential for phonons is $\mu = 0$. We define the density of states $g_a(\omega)$ per unit cell for the $a^{\rm th}$ phonon mode to be

$$g_a(\omega) = \frac{1}{N} \sum_{\boldsymbol{k}} \delta\big(\omega - \Omega_a(\boldsymbol{k})\big) = \mathcal{V}_0 \int\limits_{\rm BZ} \frac{d^d k}{(2\pi)^d}\, \delta\big(\omega - \Omega_a(\boldsymbol{k})\big) \quad , \tag{5.112}$$

where $N$ is the number of unit cells, $\mathcal{V}_0$ is the unit cell volume of the direct lattice, and the $\boldsymbol{k}$ sum and integral are over the first Brillouin zone only. Note that $\omega$ here has dimensions of frequency. The functions $g_a(\omega)$ is normalized to unity: $\int\limits_0^\infty d\omega\, g_a(\omega) = 1$. The total phonon density of states per unit cell is given by[8] $g(\omega) = \sum_{a=1}^{dr} g_a(\omega)$.

The grand potential for the phonon gas is

$$\begin{aligned} \Omega(T,V) &= -k_{\rm B}T \log \prod_{\boldsymbol{k},a} \sum_{n_a(\boldsymbol{k})=0}^{\infty} e^{-\beta\hbar\omega_a(\boldsymbol{k})\left(n_a(\boldsymbol{k})+\frac{1}{2}\right)} \\ &= k_{\rm B}T \sum_{\boldsymbol{k},a} \log\left[2\sinh\left(\frac{\hbar\Omega_a(\boldsymbol{k})}{2k_{\rm B}T}\right)\right] = Nk_{\rm B}T \int\limits_0^\infty d\omega\, g(\omega)\, \log\left[2\sinh\left(\frac{\hbar\omega}{2k_{\rm B}T}\right)\right] \quad . \end{aligned} \tag{5.113}$$

Note that $V = N\mathcal{V}_0$ since there are $N$ unit cells, each of volume $\mathcal{V}_0$. The entropy is given by $S = -\left(\frac{\partial\Omega}{\partial T}\right)_V$ and thus the heat capacity is

$$C_V = -T\,\frac{\partial^2\Omega}{\partial T^2} = Nk_{\rm B} \int\limits_0^\infty d\omega\, g(\omega) \left(\frac{\hbar\omega}{2k_{\rm B}T}\right)^2 {\rm csch}^2\left(\frac{\hbar\omega}{2k_{\rm B}T}\right) \tag{5.114}$$

Note that as $T \to \infty$ we have ${\rm csch}(\hbar\omega/2k_{\rm B}T) \to 2k_{\rm B}T/\hbar\omega$, and therefore

$$\lim_{T\to\infty} C_V(T) = Nk_{\rm B} \int\limits_0^\infty d\omega\, g(\omega) = rdNk_{\rm B} \quad . \tag{5.115}$$

This is the classical Dulong-Petit limit of $\frac{1}{2}k_{\rm B}$ per quadratic degree of freedom; there are $rN$ atoms moving in $d$ dimensions, hence $d \cdot rN$ positions and an equal number of momenta, resulting in a high temperature limit of $C_V = rdNk_{\rm B}$.

[8] Note the dimensions of $g(\omega)$ are $(\text{frequency})^{-1}$. By contrast, the dimensions of $g(\varepsilon)$ in eqn. 5.29 are $(\text{energy})^{-1} \cdot (\text{volume})^{-1}$. The difference lies in the factor $\mathcal{V}_0 \cdot \hbar$, where $\mathcal{V}_0$ is the unit cell volume.

### 5.5.3 Einstein and Debye models

Historically, two models of lattice vibrations have received wide attention. First is the so-called *Einstein model*, in which there is no dispersion to the individual phonon modes. We take $g_a(\omega) \approx \delta(\omega - \Omega_a)$, in which case

$$C_V(T) = N k_{\rm B} \sum_a \left(\frac{\hbar\Omega_a}{2k_{\rm B}T}\right)^2 {\rm csch}^2\left(\frac{\hbar\Omega_a}{2k_{\rm B}T}\right) \quad . \tag{5.116}$$

At low temperatures, the contribution from each branch vanishes exponentially, since ${\rm csch}^2(\hbar\Omega_a/2k_{\rm B}T) \simeq 4\exp(-\hbar\Omega_a/k_{\rm B}T) \to 0$. Real solids don't behave this way.

A more realistic model, due to Debye, accounts for the low-lying acoustic phonon branches. Since the acoustic phonon dispersion vanishes linearly with $|\boldsymbol{k}|$ as $\boldsymbol{k} \to 0$, there is no temperature at which the acoustic phonons 'freeze out' exponentially, as in the case of Einstein phonons. Indeed, the Einstein model is appropriate in describing the $d\,(r-1)$ optical phonon branches, though it fails miserably for the acoustic branches.

In the vicinity of the zone center $\boldsymbol{k} = 0$ (also called $\Gamma$ in crystallographic notation) the $d$ acoustic modes obey a linear dispersion, with $\Omega_a(\boldsymbol{k}) = c_a(\hat{\boldsymbol{k}})\,k$. This results in an acoustic phonon density of states in $d = 3$ dimensions of

$$\tilde{g}(\omega) = \frac{\mathcal{V}_0\,\omega^2}{2\pi^2} \sum_a \int \frac{d\hat{\boldsymbol{k}}}{4\pi}\, \frac{1}{c_a^3(\boldsymbol{k})}\, \Theta(\Omega_{\rm D} - \omega) = \frac{3\mathcal{V}_0}{2\pi^2 \bar{c}^3}\,\omega^2\,\Theta(\Omega_{\rm D} - \omega) \quad , \tag{5.117}$$

where $\bar{c}$ is an average acoustic phonon velocity (*i.e.* speed of sound) defined by

$$\frac{3}{\bar{c}^3} = \sum_a \int \frac{d\hat{\boldsymbol{k}}}{4\pi}\, \frac{1}{c_a^3(\boldsymbol{k})} \tag{5.118}$$

and $\Omega_{\rm D}$ is a cutoff known as the *Debye frequency*. The cutoff is necessary because the phonon branch does not extend forever, but only to the boundaries of the Brillouin zone. Thus, $\Omega_{\rm D}$ should roughly be equal to the energy of a zone boundary phonon. Alternatively, we can define $\Omega_{\rm D}$ by the normalization condition

$$\int_0^\infty d\omega\; \tilde{g}(\omega) = 3 \quad \Longrightarrow \quad \Omega_{\rm D} = (6\pi^2/\mathcal{V}_0)^{1/3}\,\bar{c} \quad . \tag{5.119}$$

This allows us to write $\tilde{g}(\omega) = \left(9\omega^2/\omega_{\rm D}^3\right)\Theta(\Omega_{\rm D} - \omega)$.

The specific heat due to the acoustic phonons is then

$$\begin{aligned} C_V(T) &= \frac{9Nk_{\rm B}}{\omega_{\rm D}^3} \int_0^{\Omega_{\rm D}} d\omega\;\omega^2 \left(\frac{\hbar\omega}{2k_{\rm B}T}\right)^2 {\rm csch}^2\left(\frac{\hbar\omega}{2k_{\rm B}T}\right) \\ &= 9Nk_{\rm B}\left(\frac{2T}{\Theta_{\rm D}}\right)^3 \phi\big(\Theta_{\rm D}/2T\big) \quad , \end{aligned} \tag{5.120}$$

where $\Theta_{\rm D} = \hbar\Omega_{\rm D}/k_{\rm B}$ is the *Debye temperature* and

$$\phi(x) = \int_0^x dt\, t^4\, \mathrm{csch}^2 t = \begin{cases} x^3/3 & x \to 0 \\ \pi^4/30 & x \to \infty \end{cases} \quad . \tag{5.121}$$

Therefore,

$$C_V(T) = \begin{cases} \frac{12}{5}\pi^4 N k_{\rm B}\,(T/\Theta_{\rm D})^3 & T \ll \Theta_{\rm D} \\ 3Nk_{\rm B} & T \gg \Theta_{\rm D} \end{cases} \quad . \tag{5.122}$$

Thus, the heat capacity due to acoustic phonons obeys the Dulong-Petit rule in that $C_V(T \to \infty) = 3Nk_{\rm B}$, corresponding to the three acoustic degrees of freedom per unit cell. The remaining contribution of $3(r-1)Nk_{\rm B}$ to the high temperature heat capacity comes from the optical modes not considered in the Debye model. The low temperature $T^3$ behavior of the heat capacity of crystalline solids is a generic feature, and its detailed description is a triumph of the Debye model.

### 5.5.4 Melting and the Lindemann criterion

#### Atomic fluctuations in a crystal

For the one-dimensional chain, eqn. 5.104 gives

$$\hat{u}_k = i\left(\frac{\hbar}{2m\Omega_k}\right)^{1/2}\left(a_k - a_{-k}^\dagger\right) \quad . \tag{5.123}$$

Therefore the RMS fluctuations at each site are given by

$$\langle u_n^2\rangle = \frac{1}{N}\sum_k \langle \hat{u}_k\, \hat{u}_{-k}\rangle = \frac{1}{N}\sum_k \frac{\hbar}{m\Omega_k}\left(n(k) + \tfrac{1}{2}\right) \quad , \tag{5.124}$$

where $n(k,T) = \left[\exp(\hbar\Omega_k/k_{\rm B}T) - 1\right]^{-1}$ is the Bose occupancy function.

Let us now generalize this expression to the case of a $d$-dimensional solid. The appropriate expression for the RMS position fluctuations of the $i^{\rm th}$ basis atom in each unit cell is

$$\langle \boldsymbol{u}_i^2(\boldsymbol{R})\rangle = \frac{1}{N}\sum_{\boldsymbol{k}}\sum_{a=1}^{dr} \frac{\hbar}{M_{ia}(\boldsymbol{k})\,\Omega_a(\boldsymbol{k})}\left(n_a(\boldsymbol{k}) + \tfrac{1}{2}\right) \quad . \tag{5.125}$$

Here we sum over all wavevectors $\boldsymbol{k}$ in the first Brilliouin zone, and over all normal modes $a$. There are $dr$ normal modes per unit cell *i.e.* $dr$ branches of the phonon dispersion $\Omega_a(\boldsymbol{k})$. (For the one-dimensional chain with $d = 1$ and $r = 1$ there was only one such branch to consider). Note also the quantity $M_{ia}(\boldsymbol{k})$, which has units of mass and is defined in terms of the polarization vectors $\mathsf{e}_{i\alpha}^{(a)}(\boldsymbol{k})$ as

$$\frac{1}{M_{ia}(\boldsymbol{k})} = \sum_{\mu=1}^{d}\left|\mathsf{e}_{i\mu}^{(a)}(\boldsymbol{k})\right|^2 \quad . \tag{5.126}$$

The dimensions of the polarization vector are $[\mathrm{mass}]^{-1/2}$, since the generalized orthonormality condition on the normal modes is

$$\sum_{i,\mu} m_i \, \mathsf{e}_{i\mu}^{(a)\,*}(\boldsymbol{k}) \, \mathsf{e}_{i\mu}^{(b)}(\boldsymbol{k}) = \delta^{ab} \quad , \tag{5.127}$$

where $m_i$ is the mass of the ion of species $i$ within the unit cell ($i \in \{1, \ldots, r\}$). For our purposes we cheat a little and replace $M_{ia}(\boldsymbol{k})$ by an appropriately averaged quantity which we call $\bar{m}_i$ ; this 'effective mass' is then independent of the mode index $a$ as well as the wavevector $\boldsymbol{k}$, greatly simplifying the calculations. We may then write

$$\langle \, \boldsymbol{u}_i^2 \, \rangle \approx \int\limits_0^\infty d\omega \; g(\omega) \, \frac{\hbar}{\bar{m}_i \, \omega} \cdot \left\{ \frac{1}{e^{\hbar\omega/k_{\rm B}T} - 1} + \frac{1}{2} \right\} \quad , \tag{5.128}$$

where we have dropped the site label $\boldsymbol{R}$ since translational invariance guarantees that the fluctuations are the same from one unit cell to the next. Note that the fluctuations $\langle \, \boldsymbol{u}_i^2 \, \rangle$ can be divided into a temperature-dependent part $\langle \, \boldsymbol{u}_i^2 \, \rangle_{\rm th}$ and a temperature-independent quantum contribution $\langle \, \boldsymbol{u}_i^2 \, \rangle_{\rm qu}$, where

$$\begin{aligned} \langle \, \boldsymbol{u}_i^2 \, \rangle_{\rm th} &= \frac{\hbar}{\bar{m}_i} \int\limits_0^\infty d\omega \; \frac{g(\omega)}{\omega} \cdot \frac{1}{e^{\hbar\omega/k_{\rm B}T} - 1} \\ \langle \, \boldsymbol{u}_i^2 \, \rangle_{\rm qu} &= \frac{\hbar}{2\bar{m}_i} \int\limits_0^\infty d\omega \; \frac{g(\omega)}{\omega} \quad . \end{aligned} \tag{5.129}$$

Let's evaluate these contributions within the Debye model, where we replace $g(\omega)$ by

$$\bar{g}(\omega) = \frac{d^2 \, \omega^{d-1}}{\omega_{\rm D}^d} \, \Theta(\Omega_{\rm D} - \omega) \quad . \tag{5.130}$$

We then find

$$\begin{aligned} \langle \, \boldsymbol{u}_i^2 \, \rangle_{\rm th} &= \frac{d^2\hbar}{\bar{m}_i \, \Omega_{\rm D}} \left( \frac{k_{\rm B}T}{\hbar\Omega_{\rm D}} \right)^{d-1} F_d(\hbar\Omega_{\rm D}/k_{\rm B}T) \\ \langle \, \boldsymbol{u}_i^2 \, \rangle_{\rm qu} &= \frac{d^2}{d-1} \cdot \frac{\hbar}{2\bar{m}_i \, \Omega_{\rm D}} \quad , \end{aligned} \tag{5.131}$$

where

$$F_d(x) = \int\limits_0^x ds \; \frac{s^{d-2}}{e^s - 1} = \begin{cases} x^{d-2}/(d-2) & x \to 0 \\ \zeta(d-1) & x \to \infty \end{cases} \quad . \tag{5.132}$$

We can now extract from these expressions several important conclusions:

1) The $T = 0$ contribution to the fluctuations, $\langle \, \boldsymbol{u}_i^2 \, \rangle_{\rm qu}$, diverges in $d = 1$ dimensions. *Therefore there are no one-dimensional quantum solids.*

| Element | Ag | Al | Au | C | Cd | Cr | Cu | Fe | Mn |
|---|---|---|---|---|---|---|---|---|---|
| $\Theta_{\rm D}$ (K) | 227 | 433 | 162 | 2250 | 210 | 606 | 347 | 477 | 409 |
| $T_{\rm melt}$ (K) | 1235 | 933.5 | 1337 | 3823 | 594.2 | 2180 | 1358 | 1811 | 1519 |
| Element | Ni | Pb | Pt | Si | Sn | Ta | Ti | W | Zn |
| $\Theta_{\rm D}$ (K) | 477 | 105 | 237 | 645 | 199 | 246 | 420 | 383 | 329 |
| $T_{\rm melt}$ (K) | 1728 | 600.2 | 2042 | 1687 | 505.1 | 3290 | 1941 | 3695 | 692.7 |

Table 5.1: Debye temperatures (at $T = 0$) and melting points for some common elements (carbon is assumed to be diamond and not graphite). (*Source: the internet!*)

2) The thermal contribution to the fluctuations, $\langle \boldsymbol{u}_i^2 \rangle_{\rm th}$, diverges for any $T > 0$ whenever $d \leq 2$. This is because the integrand of $F_d(x)$ goes as $s^{d-3}$ as $s \to 0$. *Therefore, there are no two-dimensional classical solids*[9].

3) Both the above conclusions are valid in the thermodynamic limit. Finite size imposes a cutoff on the frequency integrals, because there is a smallest wavevector $k_{\rm min} \sim 2\pi/L$, where $L$ is the (finite) linear dimension of the system. This leads to a low frequency cutoff $\Omega_{\rm min} = 2\pi\bar{c}/L$, where $\bar{c}$ is the appropriately averaged acoustic phonon velocity from eqn. 5.118, which mitigates any divergences.

**Lindemann melting criterion**

An old phenomenological theory of melting due to Lindemann says that a crystalline solid melts when the RMS fluctuations in the atomic positions exceeds a certain fraction $x^*$ of the lattice constant $a$. We therefore define the ratios

$$\begin{aligned} x_{i,\rm th}^2 &\equiv \frac{\langle \boldsymbol{u}_i^2 \rangle_{\rm th}}{a^2} = d^2 \cdot \left( \frac{\hbar^2}{\bar{m}_i \, a^2 \, k_{\rm B}} \right) \cdot \frac{T^{d-1}}{\Theta_{\rm D}^d} \cdot F(\Theta_{\rm D}/T) \\ x_{i,\rm qu}^2 &\equiv \frac{\langle \boldsymbol{u}_i^2 \rangle_{\rm qu}}{a^2} = \frac{d^2}{2(d-1)} \cdot \left( \frac{\hbar^2}{\bar{m}_i \, a^2 \, k_{\rm B}} \right) \cdot \frac{1}{\Theta_{\rm D}} \quad , \end{aligned} \tag{5.133}$$

with $x_i = \sqrt{x_{i,\rm th}^2 + x_{i,\rm qu}^2} = \sqrt{\langle \boldsymbol{u}_i^2 \rangle}\Big/a$.

Let's now work through an example of a three-dimensional solid. We'll assume a single element basis ($r = 1$). We have that $\hbar^2/k_{\rm B} = 48.4\,{\rm K} \cdot {\rm amu} \cdot {\rm \AA}^2$. According to table 5.1, the melting temperature always exceeds the Debye temperature, and often by a great amount. We therefore assume $T \gg \Theta_{\rm D}$, which puts us in the small $x$ limit of $F_d(x)$ by invoking the small-$x$ limit of eqn. 5.131. We obtain

$$x_{\rm qu}^2 = \frac{\Theta^\star}{\Theta_{\rm D}} \quad , \quad x_{\rm th}^2 = \frac{\Theta^\star}{\Theta_{\rm D}} \cdot \frac{4T}{\Theta_{\rm D}} \quad \Rightarrow \quad x = \sqrt{\left(1 + \frac{4T}{\Theta_{\rm D}}\right) \frac{\Theta^\star}{\Theta_{\rm D}}} \quad . \tag{5.134}$$

[9]There are two-dimensional crystals which possess "quasi-long-range order" at finite temperature.

where

$$\Theta^* = \frac{109\,\mathrm{K}}{m[\mathsf{amu}] \cdot a^2[\mathrm{\AA}^2]} \quad . \tag{5.135}$$

The total position fluctuation is of course the sum $x^2 = x_{i,\rm th}^2 + x_{i,\rm qu}^2$. Consider for example the case of copper, with $m = 63.5\,\mathsf{amu}$ and $a = 2.87\,\mathrm{\AA}$. The Debye temperature is $\Theta_{\rm D} = 347\,\mathrm{K}$. From this we find $x_{\rm qu} = 0.0245$, which says that at $T = 0$ the RMS fluctuations of the atomic positions are not quite three percent of the lattice spacing (*i.e.* the distance between neighboring copper atoms). At room temperature, $T = 293\,\mathrm{K}$, one finds $x_{\rm th} = 0.0450$, which is about twice as large as the quantum contribution. How big are the atomic position fluctuations at the melting point? According to our table, $T_{\rm melt} = 1358\,\mathrm{K}$ for copper, and from our formulae we obtain $x_{\rm melt} = 0.10$. The *Lindemann criterion* says that solids melt when $x(T) \approx 0.1$, which, remarkably, is the value we computed for Cu.

We were very lucky to hit the magic number $x_{\rm melt} = 0.1$ with copper. Let's try another example. Lead has $M = 208\,\mathsf{amu}$ and $a = 4.95\,\mathrm{\AA}$. The Debye temperature is $\Theta_{\rm D} = 105\,\mathrm{K}$ ('soft phonons'), and the melting point is $T_{\rm melt} = 600.2\,\mathrm{K}$. From these data we obtain $x(T = 0) = 0.0143$, $x(293\,\mathrm{K}) = 0.0477$ and $x(T = 600.2\,\mathrm{K}) = 0.0683$. Not as lucky this time, but at least we're in the right ballpark.

We can turn the analysis around and predict a melting temperature based on the Lindemann criterion $x(T_{\rm melt}) = x^* \approx 0.1$. We obtain

$$T_{\rm L} = \left( \frac{\Theta_{\rm D}}{\Theta^\star}\, x^{*2} - 1 \right) \cdot \frac{\Theta_{\rm D}}{4} \quad . \tag{5.136}$$

We call $T_{\rm L}$ the *Lindemann temperature*. Most treatments of the Lindemann criterion ignore the quantum correction, which gives the $-1$ contribution inside the above parentheses. But if we are more careful and include it, we see that it may be possible to have $T_{\rm L} < 0$. This occurs for any crystal where $\Theta_{\rm D} < \Theta^\star / x^{*2}$. In this case we might expect the crystalline solid to be unstable to a liquid phase even at $T = 0$ owing to quantum fluctuations.

This is indeed the case for $^4$He, which at atmospheric pressure condenses into a liquid at $T_{\rm c} = 4.2\,\mathrm{K}$ and remains in the liquid state down to absolute zero. At $p = 1\,\mathrm{atm}$, it never solidifies! Why? The number density of liquid $^4$He at $p = 1\,\mathrm{atm}$ and $T = 0\,\mathrm{K}$ is $2.2 \times 10^{22}\,\mathrm{cm}^{-3}$. Let's say the helium atoms want to form a crystalline lattice. We don't know *a priori* what the lattice structure will be, so let's for the sake of simplicity assume a simple cubic lattice. From the number density we obtain a lattice spacing of $a = 3.57\,\mathrm{\AA}$. OK now what do we take for the Debye temperature? Theoretically this should depend on the microscopic force constants which enter the small oscillations problem (*i.e.* the spring constants between pairs of helium atoms in equilibrium). We'll use the expression we derived for the Debye frequency, $\Omega_{\rm D} = (6\pi^2/\mathcal{V}_0)^{1/3}\bar{c}$, where $\mathcal{V}_0$ is the unit cell volume. We'll take $\bar{c} = 238\,\mathrm{m/s}$, which is the speed of sound in liquid helium at $T = 0$. This gives $\Theta_{\rm D} = 19.8\,\mathrm{K}$. We find $\Theta^\star = 2.13\,\mathrm{K}$, and if we take $x^* = 0.1$ this gives $\Theta^\star / x^{*2} = 213\,\mathrm{K}$, which significantly exceeds $\Theta_{\rm D}$. Thus, the solid should melt because the RMS fluctuations in the atomic positions at absolute zero are huge: $x_{\rm qu} = (\Theta^\star/\Theta_{\rm D})^{1/2} = 0.33$. By applying pressure, one can get $^4$He to crystallize above $p_{\rm c} = 25\,\mathrm{atm}$ at $T = 0$. Under pressure, the unit cell volume $\mathcal{V}_0$ decreases and the phonon velocity $\bar{c}$ increases, so the Debye temperature increases.

It is important to recognize that the Lindemann criterion does not provide us with a theory of melting *per se*. Rather it provides us with a heuristic allowing us to predict roughly when a solid should melt.

### 5.5.5 Goldstone bosons

The vanishing of the acoustic phonon dispersion at $\boldsymbol{k} = 0$ is a consequence of *Goldstone's theorem* which says that associated with every *broken generator* of a *continuous symmetry* there is an associated bosonic gapless excitation (*i.e.* one whose frequency $\omega$ vanishes in the long wavelength limit). In the case of phonons, the 'broken generators' are the symmetries under spatial translation in the $x$, $y$, and $z$ directions. The crystal selects a particular location for its center-of-mass, which breaks this symmetry. There are, accordingly, three gapless acoustic phonons.

Magnetic materials support another branch of elementary excitations known as spin waves, or *magnons*. In *isotropic* magnets, there is a global symmetry associated with rotations in internal spin space, described by the group $\mathrm{SU}(2)$. If the system spontaneously magnetizes, meaning there is long-ranged ferromagnetic order ($\uparrow\uparrow\uparrow\cdots$), or long-ranged antiferromagnetic order ($\uparrow\downarrow\uparrow\downarrow\cdots$), then global spin rotation symmetry is broken. Typically a particular direction is chosen for the magnetic moment (or staggered moment, in the case of an antiferromagnet). Symmetry under rotations about this axis is then preserved, but rotations which do not preserve the selected axis are 'broken'. In the most straightforward case, that of the antiferromagnet, there are two such rotations for $\mathrm{SU}(2)$, and concomitantly two gapless magnon branches, with linearly vanishing dispersions $\Omega_a(\boldsymbol{k})$. The situation is more subtle in the case of ferromagnets, because the total magnetization is conserved by the dynamics (unlike the total staggered magnetization in the case of antiferromagnets). Another wrinkle arises if there are long-ranged interactions present.

For our purposes, we can safely ignore the deep physical reasons underlying the gaplessness of Goldstone bosons and simply posit a gapless dispersion relation of the form $\omega(\boldsymbol{k}) = A\,|\boldsymbol{k}|^\sigma$. The density of states for this excitation branch is then $g(\omega) = \mathcal{C}\,\omega^{r-1}\,\Theta(\Omega_{\rm c} - \omega)$, where $r = d/\sigma$ and where $\mathcal{C}$ is a constant and $\Omega_{\rm c}$ is the cutoff, which is the bandwidth for this excitation branch.[10] Normalizing the density of states for this branch results in the identification $\Omega_{\rm c} = (r/\mathcal{C})^{1/r}$.

The heat capacity is then found to be

$$C_V = N k_{\rm B}\,\mathcal{C}\int\limits_0^{\Omega_{\rm c}} d\omega\,\omega^{r-1}\left(\frac{\hbar\omega}{k_{\rm B}T}\right)^2 \mathrm{csch}^2\!\left(\frac{\hbar\omega}{2k_{\rm B}T}\right) = \frac{d}{\sigma}\,N k_{\rm B}\left(\frac{2T}{\Theta}\right)^r \phi\big(\Theta/2T\big) \quad , \tag{5.137}$$

where $\Theta = \hbar\Omega_{\rm c}/k_{\rm B}$ and

$$\phi(x) = \int\limits_0^x dt\,t^{r+1}\,\mathrm{csch}^2 t = \begin{cases} r^{-1}\,x^r & x \to 0 \\ 2^{-r}\,\Gamma\big(2+r\big)\,\zeta\big(2+r\big) & x \to \infty \end{cases} \quad , \tag{5.138}$$

which is a generalization of our earlier results. Once again, we recover Dulong-Petit for $k_{\rm B}T \gg \hbar\Omega_{\rm c}$, with $C_V(T \gg \hbar\Omega_{\rm c}/k_{\rm B}) = N k_{\rm B}$.

In an isotropic ferromagnet, *i.e.*a ferromagnetic material where there is full SU(2) symmetry in internal 'spin' space, the magnons have a $k^2$ dispersion. Thus, a bulk three-dimensional isotropic ferromagnet will exhibit a heat capacity due to spin waves which behaves as $T^{3/2}$ at low temperatures. For sufficiently low temperatures this will overwhelm the phonon contribution, which behaves as $T^3$.

[10] If $\omega(\boldsymbol{k}) = Ak^\sigma$, then $\mathcal{C} = 2^{1-d}\,\pi^{-\frac{d}{2}}\,\sigma^{-1}\,A^{-\frac{d}{\sigma}}\,\mathrm{g}\,/\,\Gamma(d/2)$.

## 5.6 The Ideal Bose Gas

### 5.6.1 General formulation for noninteracting systems

Recall that the grand partition function for noninteracting bosons is given by

$$\Xi(T,V,\mu) = \prod_\alpha \left( \sum_{n_\alpha=0}^{\infty} e^{\beta(\mu-\varepsilon_\alpha)n_\alpha} \right) = \prod_\alpha \left(1 - e^{\beta(\mu-\varepsilon_\alpha)}\right)^{-1} \quad , \tag{5.139}$$

In order for the sum to converge to the RHS above, we must have $\mu < \varepsilon_\alpha$ for all single-particle states $|\alpha\rangle$. The density of particles is then

$$n(T,\mu) = -\frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = \frac{1}{V}\sum_\alpha \frac{1}{e^{\beta(\varepsilon_\alpha-\mu)}-1} = \int_{\varepsilon_0}^{\infty} d\varepsilon\, g(\varepsilon) \overbrace{\frac{1}{e^{(\varepsilon-\mu)/k_{\rm B}T}-1}}^{n_\varepsilon^{\rm B}(T,\mu)} \quad , \tag{5.140}$$

where $g(\varepsilon) = V^{-1}\sum_\alpha \delta(\varepsilon-\varepsilon_\alpha)$ is the density of single particle states per unit volume. We must have that $g(\varepsilon) = 0$ for $\varepsilon < \varepsilon_0$ since $\varepsilon_0$ is the ground state energy of the single particle system. Typically we will have $\varepsilon_0 = 0$, as is the case for any dispersion of the form $\varepsilon(\boldsymbol{k}) = A\,|\boldsymbol{k}|^\theta$, where $\theta$ is an exponent ($\theta = 2$ for ballistic particles), for example. However, in the presence of a magnetic field, we could have $\varepsilon(\boldsymbol{k},\sigma) = A\,|\boldsymbol{k}|^\theta - g\mu_0 H\sigma$, in which case $\varepsilon_0 = -g\mu_0|H|$. The equation $n = n(T,\mu)$ is an equation of state relating the three intensive quantities $n$, $T$, and $\mu$.

For fixed $T$ and $\mu < 0$, it is easy to see that $n(T,\mu)$ is a monotonically increasing function of $\mu$,[11] and the maximum possible value for $n(T,\mu)$, called the *critical density* $n_{\rm c}(T)$, is achieved for $\mu = \varepsilon_0$, *i.e.*

$$n_{\rm c}(T) = \int_{\varepsilon_0}^{\infty} d\varepsilon\, \frac{g(\varepsilon)}{e^{(\varepsilon-\varepsilon_0)/k_{\rm B}T}-1} \quad . \tag{5.141}$$

If $g(\varepsilon_0) > 0$, then setting $\varepsilon = \varepsilon_0 + \delta\varepsilon$ and analyzing the integrand in the vicinity of $\delta\varepsilon = 0$, we conclude that $n_{\rm c}(T)$ diverges. In this case, we can always invert eqn. 5.140 to obtain $\mu = \mu(T,n)$. This is the case for power law dispersions $\varepsilon(\boldsymbol{k}) = Ak^\theta$ with $\theta \geq d$, since, as we've seen, $g(\varepsilon) = C\varepsilon^{r-1}$ with $r = d/\theta$. But if $g(\varepsilon_0) = 0$, assuming $g(\varepsilon)$ is continuous (*e.g.*, whenever $\theta < d$), $n_{\rm c}(T)$ is finite[12]. In this case, for $n > n_{\rm c}(T)$, there is no allowed value of the chemical potential for which eqn. 5.140 can be satisfied. What happens then?

When $n_{\rm c}(T)$ is finite, we can equivalently restate the problem in terms of a *critical temperature* $T_{\rm c}(n)$, defined by the equation $n_{\rm c}(T_{\rm c}) = n$. For $T < T_{\rm c}$, we apparently can no longer invert to obtain $\mu(T,n)$, so clearly something has gone wrong. The remedy is to recognize that the single particle energy levels

[11] At fixed $\mu < 0$, $n(T,\mu)$ is a monotonically increasing function of $T$.

[12] Well, that isn't quite true. For example, if $g(\varepsilon) \sim 1/\log(1/\varepsilon)$ as $\varepsilon \to 0^+$, then $\lim_{\varepsilon\to 0^+} g(\varepsilon) = 0$ yet the integral has a very weak $\log\log(1/\eta)$ divergence, where $\eta$ is the lower cutoff. But for any power law density of states where $g(\varepsilon) \propto \varepsilon^{r-1}$ with $r > 1$ as $\varepsilon \to 0^+$, the integral converges.

are discrete, and separate out the contribution from the lowest energy state $\varepsilon_0$. The number density of particles in the one-body ground state, *i.e.* the single particle state with energy $\varepsilon_0$, is

$$n_0 = \frac{1}{V}\,\frac{\mathrm{g}_0}{e^{-\beta\,\delta\mu}-1} \tag{5.142}$$

To make up for the missing density, we must have that $n_0$ is finite. This entails that $\delta\mu$ is infinitesimal in the thermodynamic limit, and solving for $\delta\mu$ we obtain

$$\delta\mu = \mu - \varepsilon_0 = -k_{\rm B}T\log\left(1+\frac{\mathrm{g}_0}{n_0 V}\right) \approx -\frac{\mathrm{g}_0 k_{\rm B}T}{N_0} \quad , \tag{5.143}$$

where $\mathrm{g}_0$ is the degeneracy of the single particle state with energy $\varepsilon_0$. Thus, for $T < T_{\rm c}(n)$, the chemical potential $\mu$ lies below $\varepsilon_0$ by an infinitesimal amount. The total density is then a sum of contributions:

$$n = n_0 + n_{\rm c}(T) \quad . \tag{5.144}$$

What has happened is this: At fixed density $n$, for $T < T_{\rm c}(n)$, the number of bosons $N_0 = n_0 V$ in the one-body ground state $|\psi_0\rangle$ with energy $\varepsilon_0$ is *extensive*, and we say that these particles have *condensed*. A macroscopic number of particles occupy a single quantum state, which is allowed according to the rules of Bose-Einstein statistics. The quantity $n_0 = N_0/V$ is the *condensate number density*. The remaining particles, with number density $n_{\rm c}(T)$, are said to comprise the *overcondensate*. The relation $n = n_0 + n_{\rm c}(T)$ may be regarded as an equation of state relating the three intensive quantities $T$, $n$, and $n_0$.

To reiterate, in the thermodynamic limit, when $T < T_{\rm c}(n)$ we have $\mu = \varepsilon_0$ and $n_0 \in [0, n]$, with

$$n(T, n_0) = n_0 + \overbrace{\int_{\varepsilon_0}^{\infty} d\varepsilon\,\frac{g(\varepsilon)}{e^{(\varepsilon-\varepsilon_0)/k_{\rm B}T}-1}}^{n_{\rm c}(T)} \quad , \tag{5.145}$$

an equation of state relating $T$, $n$, and $n_0$. For $T > T_{\rm c}(n)$, we have $\mu < \varepsilon_0$ and $n_0 = 0$, with

$$n(T,\mu) = \int_{\varepsilon_0}^{\infty} d\varepsilon\,\frac{g(\varepsilon)}{e^{(\varepsilon-\mu)/k_{\rm B}T}-1} \quad , \tag{5.146}$$

an equation of state relating $T$, $n$, and $\mu$. The critical temperature $T_{\rm c}(n)$ is determined by simultaneously setting $T = T_{\rm c}$ and $n_0 = 0$ in eqn. 5.145, or, equivalently, by setting $T = T_{\rm c}$ and $\mu = \varepsilon_0$ in eqn. 5.146, yielding the condition $n = n_{\rm c}(T_{\rm c})$. In the condensed phase, the *condensate fraction* is defined to be

$$f_0(T,n) \equiv \frac{n_0(T,n)}{n} = 1 - \frac{n_{\rm c}(T)}{n} \quad . \tag{5.147}$$

Note that $f_0(T,n)$ at fixed $n$ is a monotonically decreasing function of $T$ with $f_0(T=0,n) = 1$ and $f_0(T = T_{\rm c}(n), n) = 0$.

Incidentally, why are we allowed to ignore $\delta\mu$ in the expression for $n_{\rm c}(T)$ in eqn. 5.144? Suppose our particles have a dispersion $\varepsilon(\boldsymbol{k}) \propto k^\theta$ as $k \to 0$, *i.e.* at low energies. With box quantization, the first excited state energy will be $\varepsilon_1 \sim L^{-\theta}$ since its wavevector is $k_1 \sim L^{-1}$. Since $\delta\mu \sim V^{-1} \sim L^{-d}$, we see that $\delta\mu$ is negligible compared with $\varepsilon_1$ provided $\theta < d$, which, as we have seen, is precisely the condition for the existence of a condensate.

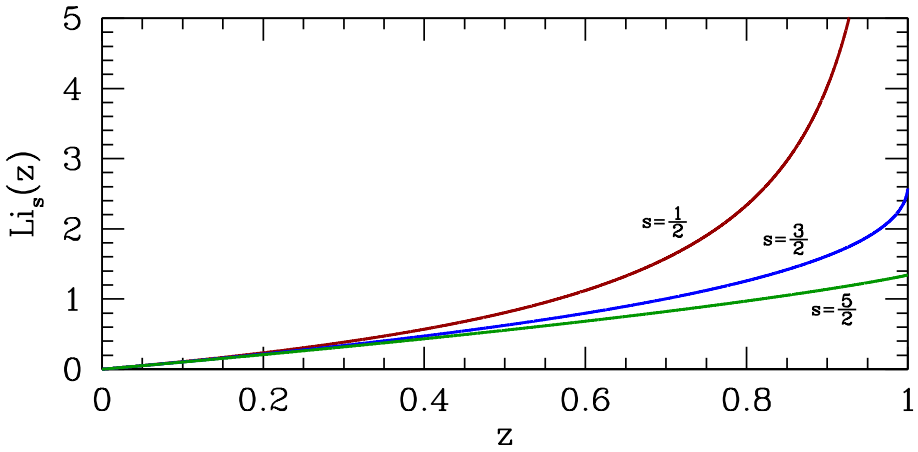



Figure 5.7: The polylogarithm function $\mathrm{Li}_s(z)$ *versus* $z$ for $s = \frac{1}{2}$, $s = \frac{3}{2}$, and $s = \frac{5}{2}$. Note that $\mathrm{Li}_s(1) = \zeta(s)$ diverges for $s \leq 1$.

### 5.6.2 Ballistic dispersion

We already derived, in §5.3.3, expressions for $n(T,z)$ and $p(T,z)$ for the ideal Bose gas (IBG) with ballistic dispersion $\varepsilon(\boldsymbol{p}) = \boldsymbol{p}^2/2m$. We found

$$\begin{aligned} n(T,z) &= \mathsf{g}\,\lambda_T^{-d}\ \mathrm{Li}_{\frac{d}{2}}(z) \\ p(T,z) &= \mathsf{g}\,k_{\mathrm{B}}T\,\lambda_T^{-d}\ \mathrm{Li}_{\frac{d}{2}+1}(z), \end{aligned} \tag{5.148}$$

where $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\mathrm{B}}T}$ is the thermal wavelength, and where $\mathsf{g}$ is the internal (*e.g.* spin) degeneracy of each single particle energy level. Here $z = \exp(\mu/k_{\mathrm{B}}T)$ is the fugacity and $\mathrm{Li}_s(z) = \sum_{m=1}^{\infty} z^m/m^s$ is the polylogarithm function. For bosons with a spectrum bounded below by $\varepsilon_0 = 0$, the fugacity takes values on the interval $z \in [0,1]$[13]. Note that $\mathrm{Li}_s(z=1) = \zeta(s)$, which is Riemann's zeta function; $\zeta(s)$ is finite for $s > 1$.

Clearly $n(T,z) = \mathsf{g}\,\lambda_T^{-d}\ \mathrm{Li}_{d/2}(z)$ is an increasing function of $z$ for fixed $T$. In fig. 5.7 we plot the function $\mathrm{Li}_s(z)$ *versus* $z$ for three different values of $s$. We note that the maximum value $\mathrm{Li}_s(z=1)$ is *finite* if $s > 1$. Thus, for $d > 2$, there is a *maximum density* $n_{\max}(T) = \mathsf{g}\,\zeta(d/2)\,\lambda_T^{-d}$ which is an increasing function of temperature $T$. Put another way, if we fix the density $n$, then there is a *critical temperature* $T_{\mathrm{c}}$ below which there is no solution to the equation $n = n(T,z)$. The critical temperature $T_{\mathrm{c}}(n)$ is then determined by the relation

$$n = \mathsf{g}\,\zeta\big(\tfrac{d}{2}\big)\lambda_{T_{\mathrm{c}}}^{-d/2} = \mathsf{g}\,\zeta\big(\tfrac{d}{2}\big)\left(\frac{mk_{\mathrm{B}}T_{\mathrm{c}}}{2\pi\hbar^2}\right)^{d/2} \qquad \Longrightarrow \qquad k_{\mathrm{B}}T_{\mathrm{c}} = \frac{2\pi\hbar^2}{m}\left(\frac{n}{\mathsf{g}\,\zeta\big(\frac{d}{2}\big)}\right)^{2/d} . \tag{5.149}$$

What happens for $T < T_{\mathrm{c}}$ ?

[13] It is easy to see that the chemical potential for noninteracting bosons can never exceed the minimum value $\varepsilon_0$ of the single particle dispersion.

As shown above in §5.6.1, we must separate out the contribution from the lowest energy single particle mode, which for ballistic dispersion lies at $\varepsilon_0 = 0$. Thus writing

$$n = \frac{1}{V}\,\frac{1}{z^{-1}-1} + \frac{1}{V}\sum_{\substack{\alpha\\(\varepsilon_\alpha>0)}}\frac{1}{z^{-1}\,e^{\varepsilon_\alpha/k_{\rm B}T}-1} \quad , \tag{5.150}$$

where we have taken $\mathsf{g} = 1$. Now $V^{-1}$ is of course very small, since $V$ is thermodynamically large, but if $\mu \to 0$ then $z^{-1}-1$ is also very small and their ratio can be finite, as we have seen. Indeed, if the density of $\boldsymbol{k} = 0$ bosons $n_0$ is *finite*, then their total number $N_0$ satisfies

$$N_0 = V n_0 = \frac{1}{z^{-1}-1} \qquad \Longrightarrow \qquad z = \frac{1}{1+N_0^{-1}} \quad . \tag{5.151}$$

The chemical potential is given by

$$\mu = k_{\rm B}T\log z = -k_{\rm B}T\log\left(1+N_0^{-1}\right) \approx -\frac{k_{\rm B}T}{N_0} \to 0^- \quad . \tag{5.152}$$

In other words, the chemical potential is infinitesimally negative, because $N_0$ is assumed to be thermodynamically large. According to eqn. 5.10, the contribution to the pressure from the $\boldsymbol{k} = 0$ states is

$$p_0 = -\frac{k_{\rm B}T}{V}\,\log(1-z) = \frac{k_{\rm B}T}{V}\log(1+N_0) = 0^+ \quad . \tag{5.153}$$

So the $\boldsymbol{k} = 0$ bosons, which we identify as the *condensate*, contribute nothing to the pressure.

Having separated out the $\boldsymbol{k} = 0$ mode, we can now replace the remaining sum over $\alpha$ by the usual integral over $\boldsymbol{k}$. We then have

$$T < T_{\rm c} \quad : \quad n = n_0 + \mathsf{g}\,\zeta\left(\tfrac{d}{2}\right)\lambda_T^{-d} \qquad , \qquad p = \mathsf{g}\,\zeta\left(\tfrac{d}{2}+1\right)k_{\rm B}T\,\lambda_T^{-d} \tag{5.154}$$

and

$$T > T_{\rm c} \quad : \quad n = \mathsf{g}\,\mathsf{Li}_{\frac{d}{2}}(z)\,\lambda_T^{-d} \qquad , \qquad p = \mathsf{g}\,\mathsf{Li}_{\frac{d}{2}+1}(z)\,k_{\rm B}T\,\lambda_T^{-d} \quad . \tag{5.155}$$

The *condensate fraction* $n_0/n$ is unity at $T = 0$, when all particles are in the condensate with $\boldsymbol{k} = 0$, and decreases with increasing $T$ until $T = T_{\rm c}$, at which point it vanishes. Explicitly, we have

$$\frac{n_0(T,n)}{n} = 1 - \frac{\mathsf{g}\,\zeta\left(\frac{d}{2}\right)}{n\,\lambda_T^d} = 1 - \left(\frac{T}{T_{\rm c}(n)}\right)^{d/2} \quad . \tag{5.156}$$

### 5.6.3 Energy and entropy, heat capacity, and compressibility

#### Energy and entropy

Let us compute the internal energy $E$ for the ideal Bose gas. We have

$$\frac{\partial}{\partial\beta}\left(\beta\Omega\right) = \Omega + \beta\,\frac{\partial\Omega}{\partial\beta} = \Omega - T\,\frac{\partial\Omega}{\partial T} = \Omega + TS \tag{5.157}$$

and therefore

$$\begin{aligned} E = \Omega + TS + \mu N &= \mu N + \frac{\partial}{\partial\beta}\,(\beta\Omega) \\ &= V\Big(\mu n - \frac{\partial}{\partial\beta}\,(\beta p)\Big) = \tfrac{1}{2} d\,\mathsf{g} V k_{\rm B} T\,\lambda_T^{-d}\,\mathsf{Li}_{\frac{d}{2}+1}(z) \quad . \end{aligned} \tag{5.158}$$

This expression is valid at all temperatures, both above and below $T_{\rm c}$. Note that the condensate particles do not contribute to $E$, because the $\boldsymbol{k} = 0$ condensate particles carry no energy.

Working along similar lines, it is left as an exercise to the reader to obtain the following formulae for the entropy per particle of the ideal Bose gas in $d$ dimensions:

$$\begin{aligned} \frac{S(T<T_{\rm c},V,N)}{Nk_{\rm B}} &= \big(\tfrac{d}{2}+1\big)\,\zeta\big(\tfrac{d}{2}+1\big)\cdot\frac{\mathsf{g}}{n\lambda_T^d} \\ \frac{S(T>T_{\rm c},V,N)}{Nk_{\rm B}} &= \big(\tfrac{d}{2}+1\big)\,\frac{\mathsf{Li}_{\frac{d}{2}+1}(z)}{\mathsf{Li}_{\frac{d}{2}}(z)} - \log z \quad , \end{aligned} \tag{5.159}$$

with $n = N/V$. In the second equation above, one must invert $n\lambda_T^d = \mathsf{g}\,\mathsf{Li}_{\frac{d}{2}}(z)$ to obtain $z(T,n)$. Note that the above equations agree at $T = T_{\rm c}$, where $z = 1$.

Let us contrast these results with those for the classical nonrelativistic ideal gas, for which

$$Z_{\rm CIG}(T,V,N) = \frac{1}{N!}\left(\frac{\mathsf{g}V}{\lambda_T^d}\right)^N \qquad \Rightarrow \qquad \frac{S_{\rm CIG}(T,V,N)}{Nk_{\rm B}} = -\log(n\lambda_T^d/\mathsf{g}) + \tfrac{d}{2} + 1 \quad . \tag{5.160}$$

For the ideal Bose gas in the Maxwell-Boltzmann limit, where $z \to 0$ and $\mathsf{Li}_s(z) = z + 2^{-s}z^2 + \ldots$, we have $z = n\lambda_T^d/\mathsf{g}$, and the classical and quantum entropies agree to lowest nontrivial order in the dimensionless quantity $n\lambda_T^3$. But as $T \to 0^+$, the classical entropy per particle $s(T,n)$ diverges logarithmically toward negative infinity. On the other hand, in the low temperature (condensed) phase of the ideal Bose gas, we have $z = 1$ and $s(T,n) \propto n^{-1}T^{d/2}$, which vanishes as $T \to 0$.

**Heat capacity**

We now investigate the heat capacity $C_{V,N} = \big(\frac{\partial E}{\partial T}\big)_{V,N}$. Since we have been working in the GCE, it is very important to note that $N$ is held constant when computing $C_{V,N}$. We'll also restrict our attention to the case $d = 3$ since the ideal Bose gas does not condense at finite $T$ for $d \le 2$ and $d > 3$ is unphysical. While we're at it, we'll also set $\mathsf{g} = 1$.

The number of particles is

$$N = \begin{cases} N_0 + \zeta(3/2)\,V\lambda_T^{-3} & (T < T_{\rm c}) \\ V\lambda_T^{-3}\,\mathsf{Li}_{3/2}(z) & (T > T_{\rm c}) \end{cases} , \tag{5.161}$$

and the energy is

$$E = \tfrac{3}{2}\,k_{\rm B}T\,\frac{V}{\lambda_T^3}\,\mathsf{Li}_{5/2}(z) = \tfrac{3}{2}\,pV \quad . \tag{5.162}$$

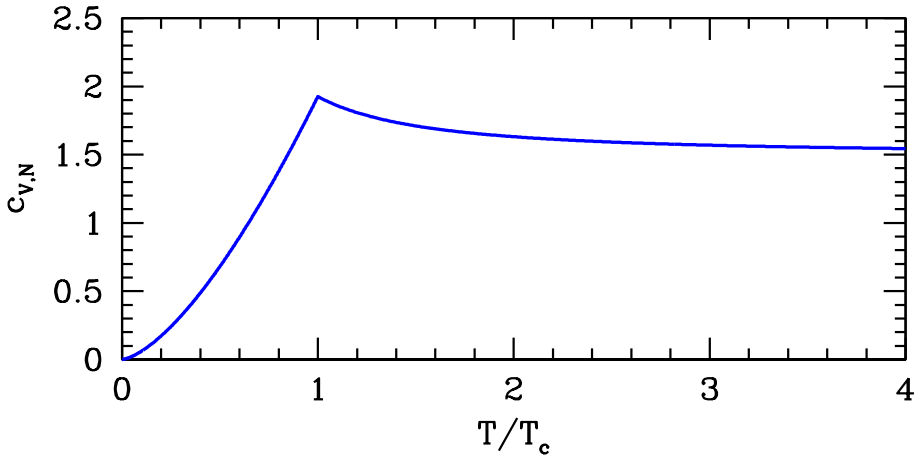



Figure 5.8: Molar heat capacity of the ideal Bose gas (units of $R$). Note the cusp at $T = T_{\rm c}$.

For $T < T_{\rm c}$, we have $z = 1$ and

$$C_{V,N} = \left(\frac{\partial E}{\partial T}\right)_{V,N} = \tfrac{15}{4}\,\zeta(5/2)\,\frac{V}{\lambda_T^3}\,k_{\rm B} \quad . \tag{5.163}$$

The molar heat capacity is therefore

$$c_{V,N}(T,n) = N_{\rm A}\cdot\frac{C_{V,N}}{N} = \tfrac{15}{4}\,\zeta(5/2)(n\,\lambda_T^3)^{-1}R \quad . \tag{5.164}$$

For $T > T_{\rm c}$, we have

$$dE\big|_V = \tfrac{15}{4}\,k_{\rm B}T\,{\rm Li}_{5/2}(z)\,\frac{V}{\lambda_T^3}\cdot\frac{dT}{T} + \tfrac{3}{2}\,k_{\rm B}T\,{\rm Li}_{3/2}(z)\,\frac{V}{\lambda_T^3}\cdot\frac{dz}{z} \quad , \tag{5.165}$$

where we have invoked eqn. 5.53. Taking the differential of $N$, we have

$$dN\big|_V = \tfrac{3}{2}\,{\rm Li}_{3/2}(z)\,\frac{V}{\lambda_T^3}\cdot\frac{dT}{T} + {\rm Li}_{1/2}(z)\,\frac{V}{\lambda_T^3}\cdot\frac{dz}{z} \quad . \tag{5.166}$$

We set $dN = 0$, which fixes $dz$ in terms of $dT$, resulting in

$$c_{V,N}(T,z) = \tfrac{3}{2}R\cdot\left[\frac{\frac{5}{2}\,{\rm Li}_{5/2}(z)}{{\rm Li}_{3/2}(z)} - \frac{\frac{3}{2}\,{\rm Li}_{3/2}(z)}{{\rm Li}_{1/2}(z)}\right] \quad . \tag{5.167}$$

To obtain $c_{V,N}(T,n)$, we must then invert the relation

$$n(T,z) = \lambda_T^{-3}\,{\rm Li}_{3/2}(z) \tag{5.168}$$

in order to obtain $z(T,n)$, and then insert this into eqn. 5.167. The results are shown in fig. 5.8. There are several noteworthy features of this plot. First of all, by dimensional analysis the function $c_{V,N}(T,n)$

is $R$ times a function of the dimensionless ratio $T/T_{\rm c}(n) \propto T\,n^{-2/3}$. Second, the high temperature limit is $\frac{3}{2}R$, which is the classical value. Finally, there is a *cusp* at $T = T_{\rm c}(n)$. At the cusp, we have

$$R^{-1}\,c_{V,N}(T_{\rm c}(n),n) = \frac{15\,\zeta(5/2)}{4\,\zeta(3/2)} = 1.92567\ldots \quad . \tag{5.169}$$

For a calculation of $c_{p,N}(T,p)$, see §5.11.

### Isothermal compressibility

The isothermal compressibility is given by

$$\kappa_T = -\frac{1}{v}\left(\frac{\partial v}{\partial p}\right)_T = -\frac{1}{v}\cdot\frac{\partial(v,T)}{\partial(p,T)}\cdot\frac{\partial(\mu,T)}{\partial(\mu,T)} = -\frac{1}{v}\left(\frac{\partial v}{\partial \mu}\right)_T\left(\frac{\partial \mu}{\partial p}\right)_T = n^{-2}\left(\frac{\partial n}{\partial \mu}\right)_T \quad , \tag{5.170}$$

where $v = V/N$ is the volume per particle[14], using $(\partial\mu/\partial p)_T = v$. In the condensed phase, when $T < T_{\rm c}(n)$, we have that $\mu = 0^-$ and thus $\kappa_T = \infty$ – the condensed phase is *infinitely compressible* (but see §5.6.5 below!). This is an unphysical consequence of the ideal (*i.e.* noninteracting) limit. In the superfluid phase of $^4$He, the individual helium atoms behave as hard spheres with a weak attraction as long distances. There is then a limit to the number of helium atoms one can force into a box of a given volume $V$, hence a limit to the number density $n$. Rather than being infinitely compressible, superfluid helium is almost *incompressible*, like generic liquids.

Gases are highly compressible, and for ideal gases $\kappa_T = 1/p$ To compute the compressibility of the ideal Bose gas for $T > T_{\rm c}(n)$, we shall need to invert the relation $n(T,z) = \lambda_T^{-3}\,\mathsf{Li}_{3/2}(z)$ to obtain $\mu(T,n)$ for $T$ just above the critical temperature, where $\mu$ is just below zero but still finite. To do this we must appeal to the relation in eqn. 5.56, which says

$$\mathsf{Li}_{\frac{3}{2}}(z) = \Gamma\big(-\tfrac{1}{2}\big)\big(-\log z\big)^{1/2} + \zeta\big(\tfrac{3}{2}\big) + \mathcal{O}(\log z) \quad , \tag{5.171}$$

with $\Gamma\big(-\frac{1}{2}\big) = -2\sqrt{\pi}$. The first term is proportional to $(-\mu)^{1/2}$. This vanishes as $\mu \to 0^-$ but is nonanalytic. The second term is the leading contribution. What remains is $\mathcal{O}(\mu)$. From the equation for $n(T,z)$, we then have

$$n\lambda_T^3 = \zeta\big(\tfrac{3}{2}\big) - \sqrt{-\frac{4\pi\mu}{k_{\rm B}T}} + \mathcal{O}(\mu) \quad , \tag{5.172}$$

which says

$$n_{\rm c}(T) - n = 2\sqrt{\pi}\,\lambda_T^{-3}\big(-\log z\big)^{1/2} + \mathcal{O}(\log z) \quad , \tag{5.173}$$

with $n_{\rm c}(T) = \zeta\big(\frac{3}{2}\big)\lambda_T^{-3}$. The isothermal compressibility is

$$\kappa_T = n^{-2}\left(\frac{\partial n}{\partial \mu}\right)_T = \frac{\lambda_T^3}{k_{\rm B}T}\,\frac{\mathsf{Li}_{1/2}(z)}{\big[\mathsf{Li}_{3/2}(z)\big]^2} \quad , \tag{5.174}$$

[14]Note that in the thermodynamics chapter we used $v$ to denote the molar volume, $N_{\rm A}V/N$.

where we make use of $\partial_\mu = (k_{\rm B}T)^{-1} z\,\partial_z$ and $z\,\partial_z\,\mathrm{Li}_s(z) = \mathrm{Li}_{s-1}(z)$. As we approach the transition temperature, $z \to 1^-$ and $\mathrm{Li}_{1/2}(z) = \Gamma(\frac{1}{2})(-\log z)^{-1/2} + \mathcal{O}\big((\log z)^0\big)$. We then have

$$\kappa_T(T,n) = \frac{2\pi}{\big[\zeta(3/2)\big]^2 k_{\rm B} T} \cdot \frac{1}{n_{\rm c}(T) - n} + \text{finite} \quad . \tag{5.175}$$

Thus, the isothermal compressibility diverges as we approach the condensed phase.

### 5.6.4 Isotherms for the ideal Bose gas

Let $a$ be some length scale and define

$$v_a = a^3 \qquad , \qquad p_a = \frac{2\pi\hbar^2}{m a^5} \qquad , \qquad T_a = \frac{2\pi\hbar^2}{m a^2 k_{\rm B}} \tag{5.176}$$

Then we have

$$\frac{v_a}{v} = \left(\frac{T}{T_a}\right)^{3/2} \mathrm{Li}_{3/2}(z) + v_a\, n_0 \qquad , \qquad \frac{p}{p_a} = \left(\frac{T}{T_a}\right)^{5/2} \mathrm{Li}_{5/2}(z) \quad , \tag{5.177}$$

where $v = V/N$ is the volume per particle and $n_0$ is the condensate number density; $n_0$ vanishes for $T \geq T_{\rm c}$, where $z = 1$. One identifies a critical volume $v_{\rm c}(T)$ by setting $z = 1$ and $n_0 = 0$, leading to $v_{\rm c}(T) = v_a\,(T/T_a)^{3/2}$. For $v < v_{\rm c}(T)$, we set $z = 1$ in eqn. 5.177 to find a relation between $v$, $T$, and $n_0$. For $v > v_{\rm c}(T)$, we set $n_0 = 0$ in eqn. 5.177 to relate $v$, $T$, and $z$. Note that the pressure is independent of volume for $T < T_{\rm c}$. The isotherms in the $(p, v)$ plane are then flat for $v < v_{\rm c}$. This resembles the coexistence region familiar from our study of the thermodynamics of the liquid-gas transition. The situation is depicted in fig. 5.9. In the $(T, p)$ plane, we identify $p_{\rm c}(T) = p_a(T/T_a)^{5/2}$ as the critical temperature at which condensation starts to occur.

Recall the Gibbs-Duhem equation, $d\mu = -s\,dT + v\,dp$. Along a coexistence curve, we have the Clausius-Clapeyron relation,

$$\left(\frac{dp}{dT}\right)_{\rm coex} = \frac{s_2 - s_1}{v_2 - v_1} = \frac{\ell}{T\,\Delta v} \quad , \tag{5.178}$$

where $\ell = T\,(s_2 - s_1)$ is the latent heat per mole, and $\Delta v = v_2 - v_1$. For ideal gas Bose condensation, the coexistence curve resembles the red curve in the right hand panel of fig. 5.9. There is no meaning to the shaded region where $p > p_{\rm c}(T)$. Nevertheless, it is tempting to associate the curve $p = p_{\rm c}(T)$ with the coexistence of the $\boldsymbol{k} = 0$ condensate and the remaining uncondensed ($\boldsymbol{k} \neq 0$) bosons.

The entropy in the coexistence region is given by

$$s = -\frac{1}{N}\left(\frac{\partial\Omega}{\partial T}\right)_V = \tfrac{5}{2}\,\zeta(5/2)\,k_{\rm B}\,v\,\lambda_T^{-3} = \frac{\frac{5}{2}\,\zeta(5/2)}{\zeta(3/2)}\,k_{\rm B}\left(1 - \frac{n_0}{n}\right) \quad . \tag{5.179}$$

All the entropy is thus carried by the uncondensed bosons, and the condensate carries zero entropy. The Clausius-Clapeyron relation can then be interpreted as describing a phase equilibrium between the condensate, for which $s_0 = v_0 = 0$, and the uncondensed bosons, for which $s' = s(T)$ and $v' = v_{\rm c}(T)$. So this identification forces us to conclude that the specific volume of the condensate is zero. This is certainly false in an interacting Bose gas!

While one can identify, by analogy, a 'latent heat' $\ell = T\,\Delta s = Ts$ in the Clapeyron equation, it is important to understand that there is no distinct thermodynamic phase associated with the region in which $p > p_{\rm c}(T)$. Ideal Bose gas condensation is a second order transition, and not a first order transition.

### 5.6.5 Mean field theory of the interacting Bose gas

For a simplified model of an interacting Bose system, consider the Hamiltonian

$$\hat{H} = \sum_{\boldsymbol{k}} \varepsilon(\boldsymbol{k})\,\hat{n}_{\boldsymbol{k}} + \frac{U}{2V}\,\hat{N}^2 \quad , \tag{5.180}$$

where $\hat{n}_{\boldsymbol{k}}$ is the number operator for particles with wavevector $\boldsymbol{k}$, $\hat{N} = \sum_{\boldsymbol{k}} \hat{n}_{\boldsymbol{k}}$ is the total number of particles, $V$ is the volume of the system, and $U$ is an interaction parameter with dimensions of (energy)$\times$(volume). We assume no internal state degeneracies (*i.e.* spin $S = 0$), and that the single particle dispersion $\varepsilon(\boldsymbol{k})$ is minimized at $\boldsymbol{k} = 0$ with $\varepsilon_0 = 0$. In this model, the interactions are very weak, proportional to $V^{-1}$, but infinite in range.

We treat the interaction term within mean field theory. That is, we define $\hat{N} = \langle \hat{N} \rangle + \delta\hat{N}$, where $\langle \hat{N} \rangle$ is the thermodynamic average of $\hat{N}$, and derive the mean field self-consistency equation for the number density $n = \langle \hat{N} \rangle / V$ by neglecting terms quadratic in the fluctuations $\delta\hat{N}$. Thus, we have

$$\begin{aligned} \hat{N}^2 &= \big(\langle \hat{N} \rangle + \delta\hat{N}\big)^2 \\ &= \langle \hat{N} \rangle^2 + 2\langle \hat{N} \rangle\,\delta\hat{N} + (\delta\hat{N})^2 \\ &= -\langle \hat{N} \rangle^2 + 2\langle \hat{N} \rangle\,\hat{N} + (\delta\hat{N})^2 \quad . \end{aligned} \tag{5.181}$$

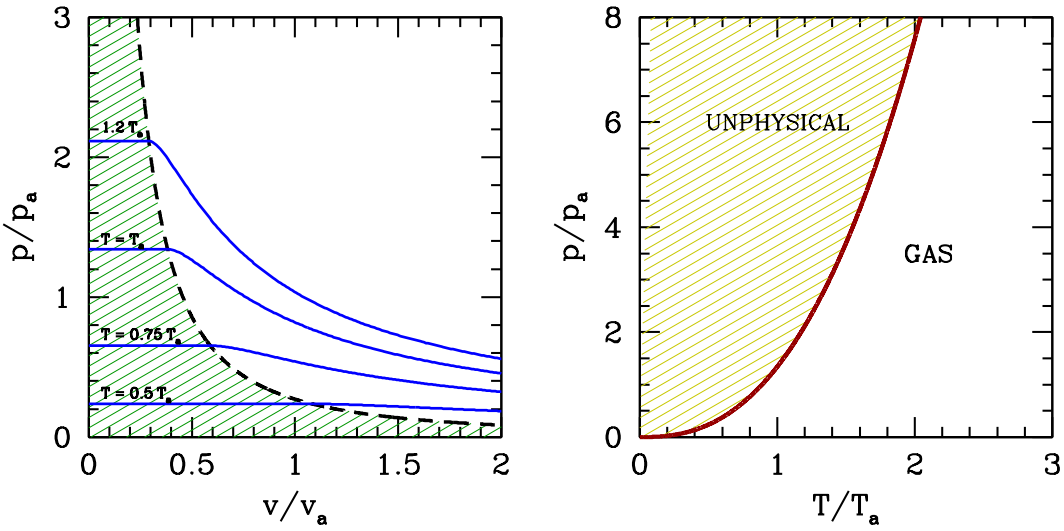



Figure 5.9: Phase diagrams for the ideal Bose gas. Left panel: $(p, v)$ plane. The solid blue curves are isotherms, and the green hatched region denotes $v < v_{\rm c}(T)$, where the system is partially condensed. Right panel: $(p, T)$ plane. The solid red curve is the coexistence curve $p_{\rm c}(T)$, along which Bose condensation occurs. No distinct thermodynamic phase exists in the yellow hatched region above $p = p_{\rm c}(T)$.

We drop the last term, $(\delta\hat{N})^2$, because it is quadratic in the fluctuations. This is the mean field assumption. The Hamiltonian now becomes

$$\hat{H}_{\rm MF} = -\tfrac{1}{2}VUn^2 + \sum_{\boldsymbol{k}} \Big(\varepsilon(\boldsymbol{k}) + Un\Big)\,\hat{n}_{\boldsymbol{k}} \quad , \tag{5.182}$$

where $n = \langle\hat{N}\rangle/V$ is the mean field number density. The dispersion is thus effectively changed from $\varepsilon(\boldsymbol{k})$ to $\tilde{\varepsilon}(\boldsymbol{k})$, where

$$\tilde{\varepsilon}(\boldsymbol{k}) = \frac{\hbar^2\boldsymbol{k}^2}{2m} + Un \quad , \tag{5.183}$$

and the minimum energy is also shifted, to $\tilde{\varepsilon}_0 = Un$. The average number of particles in state $|\boldsymbol{k}\rangle$ is given by the Bose function,

$$n_{\boldsymbol{k}}(T,\mu) = \langle\hat{n}_{\boldsymbol{k}}\rangle = \frac{1}{\exp\big[\big(\tilde{\varepsilon}(\boldsymbol{k}) - \mu\big)/k_{\rm B}T\big] - 1} \quad . \tag{5.184}$$

The grand potential is

$$\Omega = -\tfrac{1}{2}VUn^2 + Vk_{\rm B}T\int\limits_0^\infty d\varepsilon\; g(\varepsilon)\,\ln\big(1 - e^{(\mu-\varepsilon-Un)/k_{\rm B}T}\big) \quad , \tag{5.185}$$

where $g(\varepsilon)$ is the density of states for the original dispersion $\varepsilon(\boldsymbol{k})$. Extremizing $\Omega$ with respect to the mean field value $n$ yields

$$\frac{\partial\Omega}{\partial n} = -VUn + VU\int\limits_0^\infty d\varepsilon\;\frac{g(\varepsilon)}{e^{(\varepsilon+Un-\mu)/k_{\rm B}T}-1} \equiv 0 \qquad \Rightarrow \qquad n = \int\limits_0^\infty d\varepsilon\;\frac{g(\varepsilon)}{e^{(\varepsilon+Un-\mu)/k_{\rm B}T}-1} \quad . \tag{5.186}$$

At fixed $T$ and $\mu$, this is a self-consistent equation for the number density $n(T,\mu)$.

At criticality, we must have $\mu = \min_{\boldsymbol{k}}\tilde{\varepsilon}(\boldsymbol{k}) = Un$, and the critical number density $n_{\rm c}(T)$ is given by

$$n_{\rm c}(T) = \int\limits_0^\infty d\varepsilon\;\frac{g(\varepsilon)}{e^{\varepsilon/k_{\rm B}T}-1} \quad , \tag{5.187}$$

which may be inverted to yield $T_{\rm c}(n)$. Note that $n_{\rm c}(T)$ and hence $T_{\rm c}(n)$ are independent of the interaction parameter $U$ within mean field theory. We then have

$$\begin{aligned} T < T_{\rm c}(n) \;\;:\;\;\; & n(T,n_0) = n_0 + n_{\rm c}(T) \qquad (n_0 > 0, \mu = Un) \\ T > T_{\rm c}(n) \;\;:\;\;\; & n(T,\mu) = \int\limits_0^\infty d\varepsilon\;\frac{g(\varepsilon)}{e^{(\varepsilon+Un-\mu)/k_{\rm B}T}-1} \quad (n_0 = 0, \mu < Un) \quad . \end{aligned} \tag{5.188}$$

For ballistic particles we define $\nu \equiv n\lambda_T^3$, where $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T}$ is the thermal wavelength. Thus $T_{\rm c}(n) = (2\pi\hbar^2/m)\big(n/\zeta(3/2)\big)^{2/3}$ and, for $T > T_{\rm c}(n)$,

$$n\lambda_T^3 = {\rm Li}_{3/2}\big(z\,e^{-Un/k_{\rm B}T}\big) \qquad \Rightarrow \qquad \nu = {\rm Li}_{3/2}(z\,e^{-\theta\nu}) \quad , \tag{5.189}$$

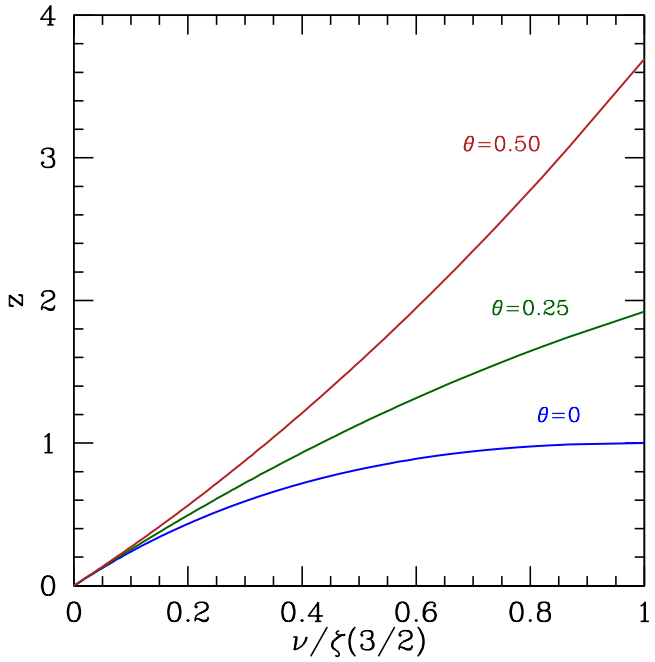



Figure 5.10: Fugacity $z$ *versus* dimensionless density $\nu = n\lambda_T^3$ for the mean field interacting Bose gas.

where $\theta \equiv U/k_{\rm B}T\lambda_T^3 \propto UT^{1/2}$ a dimensionless parameter formed from $U$ and $T$. This may be considered an equation of state relating the three intensive quantities $(\theta, z, \nu)$, or, equivalently, $(T, \mu, n)$. In Fig. 5.10 we show $z(\nu)$ for three different values of $\theta$. Note that $U = 0$ entails $\theta = 0$, whence we have the familiar relation $n\lambda_T^3 = {\rm Li}_{3/2}(z)$.

In the condensed phase, $\mu$ and $n$ are locked, with $\mu = Un$. Thus, the isothermal compressibility for $T < T_{\rm c}(n)$ is then simply obtained as

$$\kappa_T(T,n) = n^{-2}\left(\frac{\partial n}{\partial \mu}\right)_T = \frac{1}{Un^2} \quad . \tag{5.190}$$

In the noninteracting limit $U \to 0$, we recover the diverging compressibility found in §5.6.3. In the opposite limit $U \to \infty$ we find that the system is incompressible.

### 5.6.6 Properties of liquid $^4$He

#### The $\lambda$-transition

Helium has two stable isotopes. $^4$He is a boson, consisting of two protons, two neutrons, and two electrons (hence an even number of fermions). $^3$He is a fermion, with one less neutron than $^4$He. Each $^4$He atom can be regarded as a tiny hard sphere of mass $m = 6.65 \times 10^{-24}\,$g and diameter $a = 2.65\,$Å. The gas-liquid transition is first order, as usual. However, as one continues to cool, a second transition sets in at $T = T_\lambda = 2.17\,$K (at $p = 1\,$atm). The $\lambda$-transition, so named for the $\lambda$-shaped anomaly in the

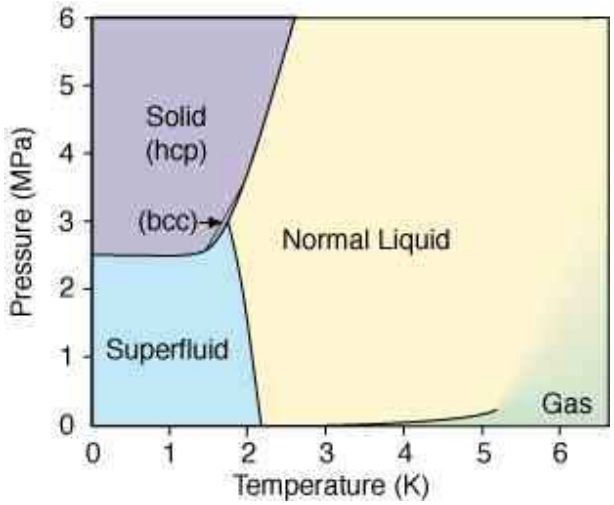



Figure 5.11: Phase diagram of $^4$He. All phase boundaries are first order transition lines, with the exception of the normal liquid-superfluid transition, which is second order. (*Source: University of Helsinki*)

specific heat in the vicinity of the transition, as shown in fig. 5.12, is continuous (*i.e.* second order). A sketch of the low temperature phase diagram is shown in fig. 5.11. At atmospheric pressure, helium liquefies at $T_{\rm l} = 4.2\,{\rm K}$.

If we pretend that $^4$He is a noninteracting Bose gas, then from a density of $n = 2.2 \times 10^{22}\,{\rm cm}^{-3}$, we obtain a Bose-Einstein condensation temperature $T_{\rm c}(n) = \frac{2\pi\hbar^2}{m}\big(n/\zeta(3/2)\big)^{2/3} = 3.16\,{\rm K}$, which is in the right ballpark. The specific heat $C_p(T)$ is found to be singular at $T = T_\lambda$, with

$$C_p(T) = A_\pm \big|T - T_\lambda(p)\big|^{-\alpha} + \text{nonsingular} \quad . \tag{5.191}$$

$\alpha$ is an example of a *critical exponent*[15]. We shall study the physics of critical phenomena later on in this course. For the ideal nonrelativistic Bose gas, the cusp singularity found in fig. 5.8 for $C_{V,N}(T)$ corresponds to $\alpha = -1$. Perhaps surprisingly, for $C_{p,N}(T)$ one finds instead $\alpha = \frac{1}{2}$ (see the calculation in §5.11). The observed behavior of $C_p(T)$ in $^4$He is very nearly logarithmic in $|T - T_\lambda|$. In fact, both theory (renormalization group on the O(2) model) and experiment concur that $\alpha$ is almost zero but in fact slightly negative, with $\alpha = -0.0127 \pm 0.0003$ in the best experiments (Lipa *et al.*, 2003). The $\lambda$ transition is most definitely *not* an ideal Bose gas condensation. Theoretically, in the parlance of critical phenomena, IBG condensation and the $\lambda$-transition in $^4$He lie in different *universality classes*[16]. Unlike the IBG, the condensed phase in $^4$He is a distinct thermodynamic phase, known as a *superfluid*. Note that $C_p(T < T_{\rm c})$ for the IBG is not even defined, since for $T < T_{\rm c}$ we have $p = p(T)$ and therefore $dp = 0$ requires $dT = 0$.

[15] $A_\pm$ is called a *critical amplitude* and is separately defined on either side of the transition, *i.e.* $A_+$ for $T \to T_{\rm c}$ from above and $A_0$ for $T \to T_{\rm c}$ from below.

[16] IBG condensation is in the universality class of the spherical model. The $\lambda$-transition is in the universality class of the $XY$ model.

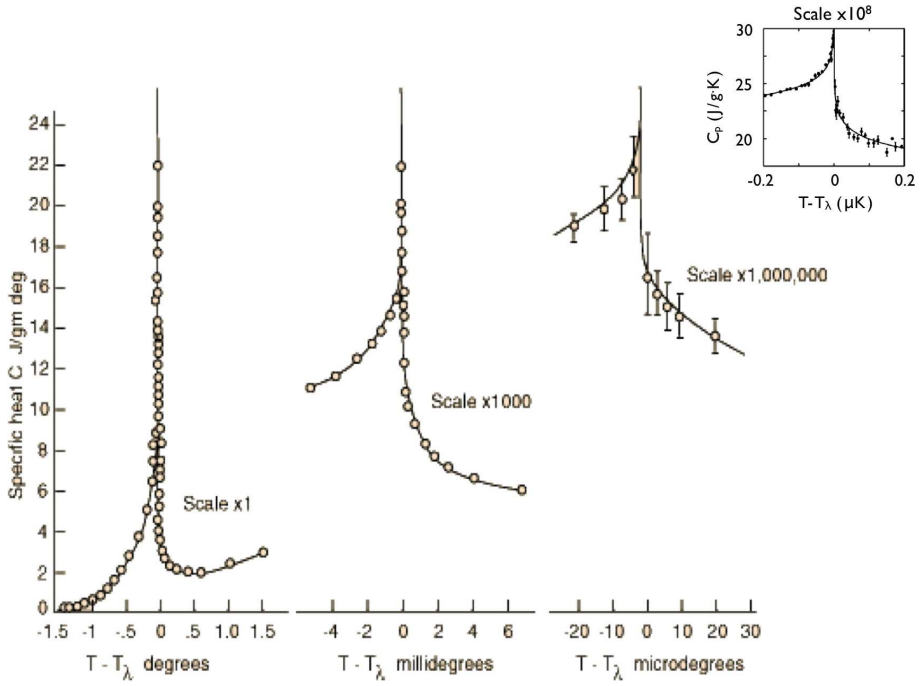



Figure 5.12: Specific heat of liquid $^4$He in the vicinity of the $\lambda$-transition. Data from M. J. Buckingham and W. M. Fairbank, in *Progress in Low Temperature Physics*, C. J. Gortner, ed. (North-Holland, 1961). Inset at upper right: more recent data of J. A. Lipa *et al.*, *Phys. Rev. B* **68**, 174518 (2003) performed in zero gravity earth orbit, to within $\Delta T = 2\,\mathrm{nK}$ of the transition.

**Fountain effect in superfluid $^4$He**

At temperatures $T < T_\lambda$, liquid $^4$He has a superfluid component which is a type of Bose condensate. In fact, there is an important difference between condensate fraction $N_{\boldsymbol{k}=0}/N$ and superfluid density, which is denoted by the symbol $\rho_{\rm s}$. In $^4$He, for example, at $T = 0$ the condensate fraction is only about $8\%$, while the superfluid fraction $\rho_{\rm s}/\rho = 1$. The distinction between $N_0$ and $\rho_s$ is very interesting but lies beyond the scope of this course.

One aspect of the superfluid phase is its complete absence of viscosity of its superfluid component. For this reason, superfluids can flow through tiny cracks called *microleaks* that will not pass normal fluid. Consider then a porous plug which permits the passage of superfluid but not of normal fluid. The key feature of the superfluid component is that it carries no entropy. For the ideal Bose gas, it also has zero energy density, but that is no longer true for interacting systems like $^4$He. In the latter case, we must invoke thermomechanical equilibrium: $d\mu = -s\,dT + v\,dp = 0$, where $s$ and $v$ are the entropy and volume per particle, respectively. The reason that the chemical potential difference $d\mu$ across the plug vanishes is that helium atoms in the zero viscosity superfluid component can shuttle from one side of the plug to the other. The two sides can then maintain different temperatures and pressures while reaching equilibrium via superfluid particle transfer. From hydrostatics we have

$$\frac{dh}{dT} = \frac{1}{\rho g}\frac{dp}{dT} = \frac{ns}{\rho g} \tag{5.192}$$

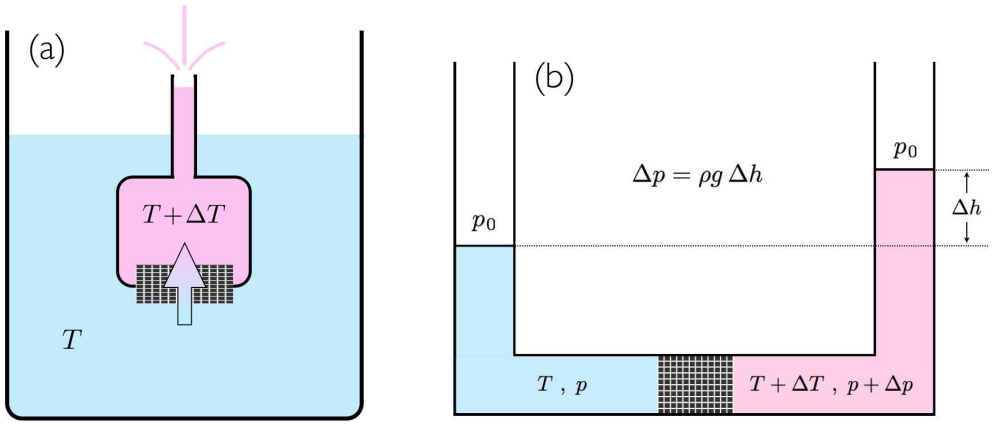



Figure 5.13: The fountain effect. In each case, a temperature gradient is maintained across a porous plug through which only superfluid can flow. This results in a pressure gradient which can result in a fountain or an elevated column in a U-tube.

where $n = v^{-1}$ is the number density and $ns = S/V$ is the entropy per unit volume, which is entirely due to the non-superfluid ("normal") component.

The elementary excitations in the superfluid state are sound waves called *phonons*. They are compressional waves, just like longitudinal phonons in a solid, but here in a liquid. Their dispersion is acoustic, given by $\omega(k) = ck$ where $c = 238\,\mathrm{m/s}$.[17] They have no internal degrees of freedom, hence $\mathsf{g} = 1$. Like phonons in a solid, the phonons in liquid helium are not conserved. Hence their chemical potential vanishes and these excitations are described by photon statistics. Appealing to the quantum thermodynamics of phonons, we may compute the height difference $\Delta h$ in the U-tube.

Clearly $\Delta h = \Delta p/\rho g$. so we must find $p(T)$ for the helium. In the grand canonical ensemble, we have

$$\begin{aligned} p = -\Omega/V &= -k_{\rm B}T \int \frac{d^3k}{(2\pi)^3}\, \log\left(1 - e^{-\hbar ck/k_{\rm B}T}\right) \\ &= -\frac{(k_{\rm B}T)^4}{(\hbar c)^3}\,\frac{4\pi}{8\pi^3}\int\limits_0^\infty du\, u^2\, \log(1-e^{-u}) = \frac{\pi^2}{90}\,\frac{(k_{\rm B}T)^4}{(\hbar c)^3} \quad . \end{aligned} \tag{5.193}$$

Let's assume $T = 1\,\mathrm{K}$. We'll need the density of liquid helium, $\rho = 148\,\mathrm{kg/m^3}$. We thus obtain

$$\begin{aligned} \frac{dh}{dT} &= \frac{2\pi^2}{45}\left(\frac{k_{\rm B}T}{\hbar c}\right)^3 \frac{k_{\rm B}}{\rho g} \\ &= \frac{2\pi^2}{45}\left(\frac{(1.38\times10^{-23}\,\mathrm{J/K})(1\,\mathrm{K})}{(1.055\times10^{-34}\,\mathrm{J\cdot s})(238\,\mathrm{m/s})}\right)^3 \times \frac{(1.38\times10^{-23}\,\mathrm{J/K})}{(148\,\mathrm{kg/m^3})(9.8\,\mathrm{m/s^2})} \simeq 32\,\mathrm{cm/K} \quad , \end{aligned} \tag{5.194}$$

a very noticeable effect!

[17] The phonon velocity $c$ is slightly temperature dependent.

**Andronikashvili experiment, superfluid density, and condensate fraction**

At temperatures below $T_\lambda$, the superfluid component of liquid $^4$He has zero viscosity. This entails several remarkable properties. For example, the superfluid can climb the walls of an open vessel in which it is contained and leak out, seemingly defying gravity, as depicted in panels (a), (b), and (c) of fig. 5.14. Panel (d) depicts the famous Andronikashvili experiment (1946) in which a set of metallic plates is immersed in a vat of liquid $^4$He and used as a torsional oscillator. The restoring torque of a torsional fiber twisted by an angle $\theta$ is $\tau = -K\theta$, where $K$ is the torsion elastic modulus. If the fiber is connected to an element of inertial moment $I$, then $I\ddot{\theta} = -K\theta$ and the frequency of oscillations (neglecting damping) is $\omega = \sqrt{K/I}$. In Andronikashvili's experiment, $I = I_{\rm app} + I_{\rm He}$, where $I_{\rm app}$ is the moment of inertia of the apparatus and $I_{\rm He}$ that of the $^4$He which is dragged along due to frictional forces exerted on the fluid by the plates. In viscous liquids there is a length scale $\xi(\omega) \propto \sqrt{\nu/\omega}$, where $\nu$ is the viscosity such that if the distance between the plates is less than $\xi(\omega)$, then all the normal fluid between the plates is dragged along with the apparatus as it oscillates. Any superfluid, however, remains decoupled from the oscillation of the plates. Thus,

$$I_{\rm He}(T) \approx \tfrac{1}{2}\pi R^4 h_{\rm eff}\rho_n(T) \quad , \tag{5.195}$$

where $\rho_n(T)$ is the mass density of the normal component (with $\rho_n(T > T_\lambda) = \rho$, the total helium mass density), and $R$ and $h_{\rm eff}$ are the radius and height of the helium liquid between the plates. The experimental results are sketched in panel (e) of fig. 5.14. As the temperature $T$ decreases below $T_\lambda$, the moment of inertia $I_{\rm He}(T)$ decreases in proportion to the normal fluid fraction. At $T = 0$, 100% of the helium is in the superfluid state and there is no contribution from the fluid to the moment of inertia $I$.

**The many-body wavefunction**

It is important to apprehend that there is a difference between *superfluid fraction* $\rho_{\rm s}/\rho$, which is measured in the Andronikashvili experiment, and *condensate fraction* $f_0 = N_{\boldsymbol{k}=0}/N$, which can be measured in scattering experiments. Theoretically, $f_0$ is a property of the many-body ground state wavefunction.

The ground state wavefunction of an $N$-particle system of point particles with no internal degrees of freedom is written as $\Psi_0(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$. For bosons, $\Psi_0$ is completely symmetric under any permutation of the labels $\boldsymbol{x}_i \to \boldsymbol{x}_{\sigma(i)}$ for all $i$, where $\sigma \in S_N$ is an element of the $N$ object permutation group. The wavefunction is also assumed to be normalized, *i.e.*

$$\int \prod_{i=1}^{N} d^3x_i \, \big|\Psi_0(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)\big|^2 = 1 \quad . \tag{5.196}$$

The condensate fraction $f_0$ can be expressed in terms of $\Psi_0$ as the following integral:

$$f_0 = \frac{N_0}{N} = \lim_{|\boldsymbol{x}-\boldsymbol{x}'|\to\infty} \frac{n_1(\boldsymbol{x}, \boldsymbol{x}')}{n} \quad , \tag{5.197}$$

where $n_1(\boldsymbol{x}, \boldsymbol{x}')$ is the off-diagonal one-body density matrix in the ground state[18] and $n$ is the total $^4$He

[18] See S. Moroni and M. Boninsegni, *J. Low Temp. Phys.* **136**, 129 (2004).

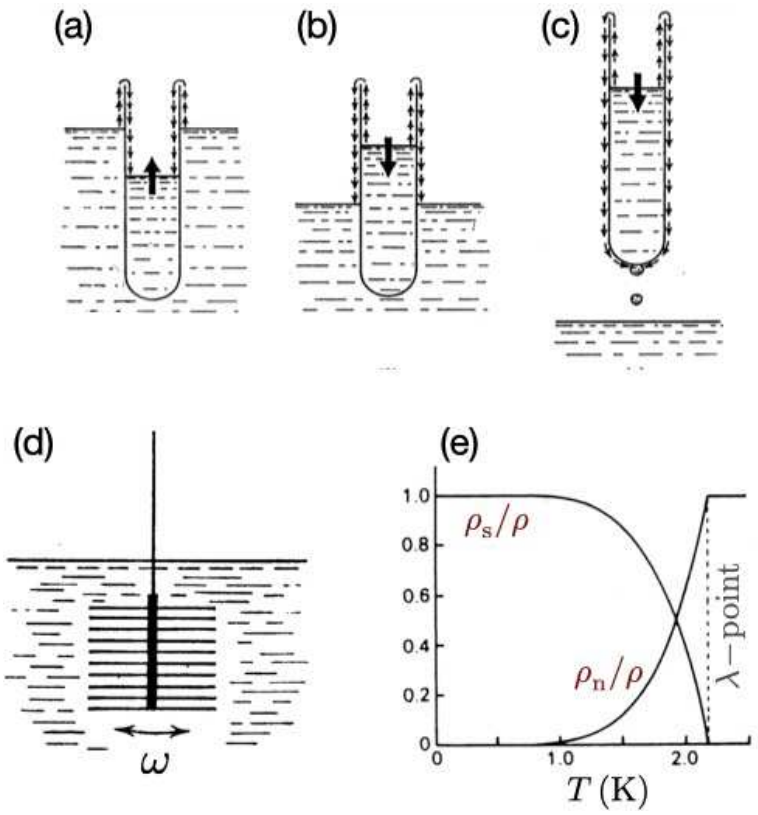



Figure 5.14: Strange properties of superfluid $^4$He: (a)-(c) Driven by van der Waals forces, superfluid $^4$He can climb walls and a container of superfluid can empty itself. (d) Andronikashvili's experiment (1946), in which a set of metallic disks immersed in a fluid is used as a torsional oscillator. Below $T_\lambda$, the period of the oscillations depends on temperature. (e) Normal and superfluid density fractions as a function of temperature.

number density

$$n_1(\boldsymbol{x},\boldsymbol{x}') = N\int\prod_{i=2}^{N} d^3x_i\;\Psi_0^*(\boldsymbol{x},\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N)\,\Psi_0(\boldsymbol{x}',\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N) \quad . \tag{5.198}$$

Theoretically, an estimate of $f_0$ requires an accurately approximated ground state wavefunction $\Psi_0$. The simplest trial wavefunctions for $^4$He are of the Jastrow pair product form,

$$\Psi_0^{\rm var}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = \prod_{j<k}^{N} f\big(|\boldsymbol{x}_j-\boldsymbol{x}_k|\big) \quad , \tag{5.199}$$

where the pair function $f(r)$ is roughly constant for $r > d$, and rapidly decays to zero for $r < d$, with $d \approx 4.2$ Å the hard core length scale, or Jastrow parameter, for helium. (Note that $d$ is greater than the effective diameter for He of $a = 2.65$ Å.) Its precise functional form is then determined by minimizing the energy $E = \langle\hat{H}\rangle$. Accurate numerical work requires much more sophisticated techniques, such as Diffusion Monte Carlo or, at finite temperature, Path Integral Monte Carlo. Moroni and Boninsegni

(2004) obtained $f_0 = 0.069 \pm 0.005$ and an equilibrium density of $\rho_0 = 0.02186\,\text{Å}^{-3}$. In dilute atomic gases, discussed in the next section, the interactions are weak, $f_0 \approx 1$, and the condensate may be imaged experimentally. In $^4$He $f_0$ can be ascertained experimentally from deep inelastic neutron scattering, and at the lowest temperatures under saturated vapor pressure measurements values from $f_0 \approx 0.06$ to $f_0 \approx 0.14$ have been reported.

### 5.6.7 Bose condensation in optical traps

The 2001 Nobel Prize in Physics was awarded to Wieman, Cornell, and Ketterle for the experimental observation of Bose condensation in dilute atomic gases. The experimental techniques required to trap and cool such systems are a true *tour de force*, and we shall not enter into a discussion of the details here[19].

The optical trapping of neutral bosonic atoms, such as $^{87}$Rb, results in a confining potential $V(\boldsymbol{r})$ which is quadratic in the atomic positions. Thus, the single particle Hamiltonian for a given atom is written

$$\hat{H} = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + \tfrac{1}{2}m\big(\omega_1^2\,x^2 + \omega_2^2\,y^2 + \omega_3^2\,z^2\big) \quad , \tag{5.200}$$

where $\omega_{1,2,3}$ are the angular frequencies of the trap. This is an anisotropic three-dimensional harmonic oscillator, the solution of which is separable into a product of one-dimensional harmonic oscillator wavefunctions. The eigenspectrum is then given by a sum of one-dimensional spectra, *viz.*

$$E_{n_1,n_2,n_3} = \big(n_1 + \tfrac{1}{2}\big)\,\hbar\omega_1 + \big(n_2 + \tfrac{1}{2}\big)\,\hbar\omega_2 + \big(n_3 + \tfrac{1}{2}\big)\,\hbar\omega_3 \quad . \tag{5.201}$$

According to eqn. 5.12, the number of particles in the system is

$$\begin{aligned} N &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\sum_{n_3=0}^{\infty}\Big[y^{-1}\,e^{n_1\hbar\omega_1/k_{\rm B}T}\,e^{n_2\hbar\omega_2/k_{\rm B}T}\,e^{n_3\hbar\omega_3/k_{\rm B}T} - 1\Big]^{-1} \\ &= \sum_{k=1}^{\infty} y^k \left(\frac{1}{1-e^{-k\hbar\omega_1/k_{\rm B}T}}\right)\left(\frac{1}{1-e^{-k\hbar\omega_2/k_{\rm B}T}}\right)\left(\frac{1}{1-e^{-k\hbar\omega_3/k_{\rm B}T}}\right) \quad , \end{aligned} \tag{5.202}$$

where we've defined $y \equiv e^{\mu/k_{\rm B}T}\,e^{-\hbar\omega_1/2k_{\rm B}T}\,e^{-\hbar\omega_2/2k_{\rm B}T}\,e^{-\hbar\omega_3/2k_{\rm B}T} \in [0,1]$.

Let's assume that the trap is modestly anisotropic, which entails that the frequency ratios $\omega_1/\omega_2$ *etc.* are all numbers on the order of one. Let us further assume that $k_{\rm B}T \gg \hbar\omega_{1,2,3}$. Then

$$\frac{1}{1-e^{-k\hbar\omega_j/k_{\rm B}T}} \;\approx\; \begin{cases} k_{\rm B}T/k\hbar\omega_j & k \lesssim k^*(T) \\ 1 & k > k^*(T) \end{cases} \tag{5.203}$$

where $k^*(T) = k_{\rm B}T/\hbar\bar{\omega} \gg 1$, with $\bar{\omega} = \big(\omega_1\,\omega_2\,\omega_3\big)^{1/3}$. We then have

$$N(T,y) \approx \frac{y^{k^*+1}}{1-y} + \left(\frac{k_{\rm B}T}{\hbar\bar{\omega}}\right)^3 \sum_{k=1}^{k^*}\frac{y^k}{k^3} \quad , \tag{5.204}$$

[19] Many reliable descriptions may be found on the web. Check Wikipedia, for example.

where the first term on the RHS is due to $k > k^*$ and the second term from $k \le k^*$ in the previous sum. Since $k^* \gg 1$ and since the sum of inverse cubes is convergent, we may safely extend the limit on the above sum to infinity. To help make more sense of the first term, write $N_0 = \left(y^{-1} - 1\right)^{-1}$ for the number of particles in the $(n_1, n_2, n_3) = (0, 0, 0)$ state. Then $y = N_0/(N_0 + 1)$, which is true always. The issue *vis-a-vis* Bose-Einstein condensation is whether $N_0 \gg 1$. At any rate, we now see that we can write

$$N \approx N_0 \left(1 + N_0^{-1}\right)^{-k^*} + \left(\frac{k_{\rm B} T}{\hbar\bar{\omega}}\right)^3 {\rm Li}_3(y) \quad . \tag{5.205}$$

As for the first term, we have

$$N_0 \left(1 + N_0^{-1}\right)^{-k^*} = \begin{cases} 0 & N_0 \ll k^* \\ N_0 & N_0 \gg k^* \end{cases} \tag{5.206}$$

Thus, as in the case of IBG condensation of ballistic particles, we identify the critical temperature by the condition $y = N_0/(N_0 + 1) \approx 1$, and we have

$$T_{\rm c} = \frac{\hbar\bar{\omega}}{k_{\rm B}} \left(\frac{N}{\zeta(3)}\right)^{1/3} = 4.5 \left(\frac{\bar{\nu}}{100\,{\rm Hz}}\right) N^{1/3}\,[\,{\rm nK}\,] \quad , \tag{5.207}$$

where $\bar{\nu} = \bar{\omega}/2\pi$. We see that $k_{\rm B} T_{\rm c} \gg \hbar\bar{\omega}$ if the number of particles in the trap is large: $N \gg 1$. In this regime, we have

$$\begin{aligned} T < T_{\rm c} \;:\quad & N = N_0 + \zeta(3) \left(\frac{k_{\rm B} T}{\hbar\bar{\omega}}\right)^3 \\ T > T_{\rm c} \;:\quad & N = \left(\frac{k_{\rm B} T}{\hbar\bar{\omega}}\right)^3 {\rm Li}_3(y) \quad . \end{aligned} \tag{5.208}$$

It is interesting to note that BEC can also occur in two-dimensional traps, which is to say traps which are very anisotropic, with oblate equipotential surfaces $V(\boldsymbol{r}) = V_0$. This happens when $\hbar\omega_3 \gg k_{\rm B} T \gg \hbar\omega_{1,2}$. We then have

$$T_{\rm c}^{(d=2)} = \frac{\hbar\bar{\omega}}{k_{\rm B}} \cdot \left(\frac{6N}{\pi^2}\right)^{1/2} \tag{5.209}$$

with $\bar{\omega} = \left(\omega_1\,\omega_2\right)^{1/2}$. The particle number then obeys a set of equations like those in eqns. 5.208, *mutatis mutandis*[20].

For extremely prolate traps, with $\omega_3 \ll \omega_{1,2}$, the situation is different because ${\rm Li}_1(y)$ diverges for $y = 1$. We then have

$$N = N_0 + \frac{k_{\rm B} T}{\hbar\omega_3} \log\left(1 + N_0\right) \quad . \tag{5.210}$$

Here we have simply replaced $y$ by the equivalent expression $N_0/(N_0 + 1)$. If our criterion for condensation is that $N_0 = \alpha N$, where $\alpha$ is some fractional value, then we have

$$T_{\rm c}(\alpha) = (1 - \alpha)\,\frac{\hbar\omega_3}{k_{\rm B}} \cdot \frac{N}{\log N} \quad . \tag{5.211}$$

[20]Explicitly, one replaces $\zeta(3)$ with $\zeta(2) = \frac{\pi^2}{6}$, ${\rm Li}_3(y)$ with ${\rm Li}_2(y)$, and $\left(k_{\rm B} T/\hbar\bar{\omega}\right)^3$ with $\left(k_{\rm B} T/\hbar\bar{\omega}\right)^2$.

## 5.7 The Ideal Fermi Gas

### 5.7.1 Grand potential and particle number

The grand potential of the ideal Fermi gas is, per eqn. 5.10,

$$\begin{aligned}\Omega(T,V,\mu) &= -k_{\rm B}T\sum_\alpha \log\left(1+e^{\mu/k_{\rm B}T}\,e^{-\varepsilon_\alpha/k_{\rm B}T}\right) \\ &= -Vk_{\rm B}T\int\limits_{-\infty}^{\infty}\!\!d\varepsilon\;g(\varepsilon)\;\log\left(1+e^{(\mu-\varepsilon)/k_{\rm B}T}\right) \quad .\end{aligned} \tag{5.212}$$

The average number of particles in a state with energy $\varepsilon$ is

$$n^{\rm F}_\varepsilon(T,\mu) = \frac{1}{e^{(\varepsilon-\mu)/k_{\rm B}T}+1} \quad , \tag{5.213}$$

hence the total number of particles is

$$N(T,V,\mu) = V\int\limits_{-\infty}^{\infty}\!\!d\varepsilon\;g(\varepsilon)\,n^{\rm F}_\varepsilon(T,\mu) = V\int\limits_{-\infty}^{\infty}\!\!d\varepsilon\;\frac{g(\varepsilon)}{e^{(\varepsilon-\mu)/k_{\rm B}T}+1} \quad . \tag{5.214}$$

### 5.7.2 The Fermi distribution

We define the function

$$f_T(\omega) \equiv \frac{1}{e^{\omega/k_{\rm B}T}+1} \quad , \tag{5.215}$$

known as the *Fermi distribution* (with $T$ implicit). In the $T\to\infty$ limit, $f_T(\omega)\to\frac{1}{2}$ for all finite values of $\omega$. As $T\to 0$, $f_T(\omega)$ approaches a step function $\Theta(-\omega)$. The average number of particles in a state of energy $\varepsilon$ in a system at temperature $T$ and chemical potential $\mu$ is $n^{\rm F}_\varepsilon(T,\mu) = f_T(\varepsilon-\mu)$. In fig. 5.1 we plot $f_T(\varepsilon-\mu)$ *versus* $\varepsilon$ for three representative temperatures. Oftentimes we will drop the subscript and write the Fermi distribution simply as $f(\varepsilon-\mu)$, with an implicit $T$-dependence.

### 5.7.3 $T=0$ and the Fermi surface

At $T=0$, we have $n_\varepsilon(T=0,\mu)=\Theta(\mu-\varepsilon)$, which says that all single particle energy states up to $\varepsilon=\mu$ are filled, and all energy states above $\varepsilon=\mu$ are empty. We call $\mu(T=0)$ the *Fermi energy*: $\varepsilon_{\rm F}=\mu(T=0)$, and we define $T_{\rm F}=\varepsilon_{\rm F}/k_{\rm B}$ to be the *Fermi temperature*. If the single particle dispersion $\varepsilon(\boldsymbol{k})$ depends only on the wavevector $\boldsymbol{k}$, then the locus of points in $\boldsymbol{k}$-space for which $\varepsilon(\boldsymbol{k})=\varepsilon_{\rm F}$, is called the *Fermi surface*. Thus, a Fermi surface is a $(d-1)$-dimensional *level set* of the function $\varepsilon(\boldsymbol{k})$, labeled by the value $\varepsilon_{\rm F}$. For isotropic systems, $\varepsilon(\boldsymbol{k})=\varepsilon(k)$ is a function only of the wavevector magnitude $k=|\boldsymbol{k}|$, and the Fermi surface is a sphere in $d=3$ or a circle in $d=2$. The radius of this circle is the *Fermi wavevector*, $k_{\rm F}$.

When there is internal (*e.g.* spin) degree of freedom, there is a Fermi surface and Fermi wavevector (for isotropic systems) for each polarization state of the internal degree of freedom.

Let's compute the Fermi wavevector $k_{\rm F}$ and Fermi energy $\varepsilon_{\rm F}$ for the IFG with a ballistic dispersion $\varepsilon(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m$. The number density is

$$n = \mathsf{g}\int \frac{d^dk}{(2\pi)^d}\,\Theta(k_{\rm F}-k) = \frac{\mathsf{g}\,\omega_d}{(2\pi)^d}\cdot\frac{k_{\rm F}^d}{d} \tag{5.216}$$

and thus

$$n^{(d=1)} = \mathsf{g}\,k_{\rm F}/\pi \quad , \quad n^{(d=2)} = \mathsf{g}\,k_{\rm F}^2/4\pi \quad , \quad n^{(d=3)} = \mathsf{g}\,k_{\rm F}^3/6\pi^2 \quad , \tag{5.217}$$

where $\omega_d = 2\pi^{d/2}/\Gamma(d/2)$ is the area of the unit sphere in $d$ space dimensions. Note that the form of $n(k_{\rm F})$ is independent of the dispersion relation, so long as it remains isotropic. Inverting the above expressions, we obtain $k_{\rm F}(n) = 2\pi(dn/\mathsf{g}\,\Omega_d)^{1/d}$:

$$k_{\rm F}^{(d=1)} = \pi n/\mathsf{g} \quad , \quad k_{\rm F}^{(d=2)} = (4\pi n/\mathsf{g})^{1/2} \quad , \quad k_{\rm F}^{(d=3)} = (6\pi^2 n/\mathsf{g})^{1/3} \quad . \tag{5.218}$$

The Fermi energy in each case, for ballistic dispersion, is given by

$$\varepsilon_{\rm F} = \frac{\hbar^2k_{\rm F}^2}{2m} = \frac{2\pi^2\hbar^2}{m}\left(\frac{d\,n}{\mathsf{g}\,\Omega_d}\right)^{2/d} \quad , \tag{5.219}$$

and so

$$\varepsilon_{\rm F}^{(d=1)} = \frac{\pi^2\hbar^2n^2}{2\mathsf{g}^2m} \quad , \quad \varepsilon_{\rm F}^{(d=2)} = \frac{2\pi\hbar^2\,n}{\mathsf{g}\,m} \quad , \quad \varepsilon_{\rm F}^{(d=3)} = \frac{\hbar^2}{2m}\left(\frac{6\pi^2n}{\mathsf{g}}\right)^{2/3} \quad . \tag{5.220}$$

Another useful result for the ballistic dispersion, which follows from the above, is that the density of states at the Fermi level is given by

$$g(\varepsilon_{\rm F}) = \frac{\mathsf{g}\,\Omega_d}{(2\pi)^d}\cdot\frac{mk_{\rm F}^{d-2}}{\hbar^2} = \frac{d}{2}\cdot\frac{n}{\varepsilon_{\rm F}} \quad . \tag{5.221}$$

That $g(\varepsilon_{\rm F})$ must be a numerical factor multiplied by $n/\varepsilon_{\rm F}$ is obvious on dimensional grounds.

For the electron gas, we have $\mathsf{g} = 2$. In a metal, one typically has $k_{\rm F} \sim 0.5\,\text{Å}^{-1}$ to $2\,\text{Å}^{-1}$, and $\varepsilon_{\rm F} \sim 1\,\text{eV} - 10\,\text{eV}$. Due to the effects of the crystalline lattice, electrons in a solid behave as if they have an *effective mass* $m^*$ which is typically on the order of the electron mass but very often about an order of magnitude smaller, particularly in semiconductors.

Nonisotropic dispersions $\varepsilon(\boldsymbol{k})$ are more interesting in that they give rise to non-spherical Fermi surfaces. The simplest example is that of a two-dimensional 'tight-binding' model of electrons hopping on a square lattice, as may be appropriate in certain layered materials. The dispersion relation is then

$$\varepsilon(k_x,k_y) = -2t\,\cos(k_xa) - 2t\,\cos(k_ya) \quad , \tag{5.222}$$

where $k_x$ and $k_y$ are confined to the interval $\left[-\frac{\pi}{a},\frac{\pi}{a}\right]$. The quantity $t$ has dimensions of energy and is known as the *hopping integral*. The Fermi surface is the set of points $(k_x,k_y)$ which satisfies $\varepsilon(k_x,k_y) =$

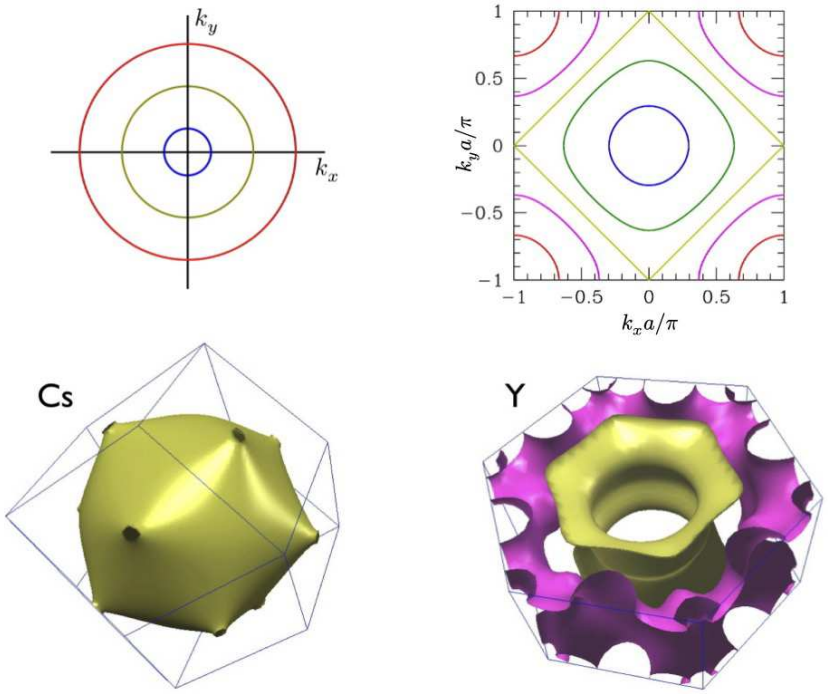



Figure 5.15: Fermi surfaces for two and three-dimensional structures. The wavevector $\boldsymbol{k}$ ranges over the first Brillouin zone in each case. Upper left: free particles in two dimensions. Upper right: 'tight binding' electrons on a square lattice. Lower left: Fermi surface for cesium, which is predominantly composed of electrons in the $6s$ orbital shell. Lower right: the Fermi surface of yttrium has two parts. One part (yellow) is predominantly due to $5s$ electrons, while the other (pink) is due to $4d$ electrons. (*Source: www.phys.ufl.edu/fermisurface/*)

$\varepsilon_{\rm F}$. When $\varepsilon_{\rm F}$ achieves its minimum value of $\varepsilon_{\rm F}^{\rm min} = -4t$, the Fermi surface collapses to a point at $(k_x, k_y) = (0, 0)$. For energies just above this minimum value, we can expand the dispersion in a power series, writing

$$\varepsilon(k_x, k_y) = -4t + ta^2\left(k_x^2 + k_y^2\right) - \tfrac{1}{12}\,ta^4\left(k_x^4 + k_y^4\right) + \ldots \quad . \tag{5.223}$$

If we only work to quadratic order in $k_x$ and $k_y$, the dispersion is isotropic, and the Fermi surface is a circle, with $k_{\rm F}^2 = (\varepsilon_{\rm F} + 4t)/ta^2$. As the energy increases further, the continuous $\mathrm{O}(2)$ rotational invariance is broken down to the discrete group of rotations of the square, $C_{4v}$. The Fermi surfaces distort and eventually, at $\varepsilon_{\rm F} = 0$, the Fermi surface is itself a square. As $\varepsilon_{\rm F}$ increases further, the square turns back into a circle, but centered about the point $\left(\frac{\pi}{a}, \frac{\pi}{a}\right)$. Note that everything is periodic in $k_x$ and $k_y$ modulo $\frac{2\pi}{a}$. The Fermi surfaces for this model are depicted in the upper right panel of fig. 5.15.

Fermi surfaces in three dimensions can be very interesting indeed, and of great importance in understanding the electronic properties of solids. Two examples are shown in the bottom panels of fig. 5.15. The electronic configuration of cesium (Cs) is $[\mathrm{Xe}]\,6s^1$. The $6s$ electrons 'hop' from site to site on a body centered cubic (BCC) lattice, a generalization of the simple two-dimensional square lattice hopping model discussed above. The elementary unit cell in $\boldsymbol{k}$ space, known as the *first Brillouin zone*, turns out to be a dodecahedron. In yttrium, the electronic structure is $[\mathrm{Kr}]\,5s^2\,4d^1$, and there are two electronic energy bands at the Fermi level, meaning two Fermi surfaces. Yttrium forms a hexagonal close packed

(HCP) crystal structure, and its first Brillouin zone is shaped like a hexagonal pillbox.

#### Spin-split Fermi surfaces

Consider an electron gas in an external magnetic field $H$. The single particle Hamiltonian is then[21]

$$\hat{H} = \frac{\boldsymbol{p}^2}{2m} + \mu_{\rm B} H \sigma \quad , \tag{5.224}$$

where $\mu_{\rm B}$ is the *Bohr magneton*, $\mu_{\rm B} = e\hbar/2mc = 5.788 \times 10^{-9}\,{\rm eV/G}$. It is convenient to keep in mind the ratio $\mu_{\rm B}/k_{\rm B} = 6.717 \times 10^{-5}\,{\rm K/G}$ when $m$ is the electron mass. What happens at $T = 0$ to a noninteracting electron gas in a magnetic field?

Electrons of each spin polarization form their own Fermi surfaces. That is, there is an up spin Fermi surface, with Fermi wavevector $k_{\rm F\uparrow}$, and a down spin Fermi surface, with Fermi wavevector $k_{\rm F\downarrow}$. The individual Fermi energies, on the other hand, must be equal, hence

$$\frac{\hbar^2 k_{\rm F\uparrow}^2}{2m} + \mu_{\rm B} H = \frac{\hbar^2 k_{\rm F\downarrow}^2}{2m} - \mu_{\rm B} H \quad , \tag{5.225}$$

which says

$$k_{\rm F\downarrow}^2 - k_{\rm F\uparrow}^2 = \frac{2eH}{\hbar c} \quad . \tag{5.226}$$

The total density is

$$n = \frac{k_{\rm F\uparrow}^3}{6\pi^2} + \frac{k_{\rm F\downarrow}^3}{6\pi^2} \qquad \Longrightarrow \qquad k_{\rm F\uparrow}^3 + k_{\rm F\downarrow}^3 = 6\pi^2 n \quad . \tag{5.227}$$

Clearly the down spin Fermi surface grows and the up spin Fermi surface shrinks with increasing $H$. Eventually, the minority spin Fermi surface vanishes altogether. This happens for the up spins when $k_{\rm F\uparrow} = 0$. Solving for the critical field, we obtain

$$H_{\rm c} = \frac{\hbar c}{2e} \cdot \left(6\pi^2 n\right)^{2/3} \quad . \tag{5.228}$$

In real magnetic solids, like cobalt and nickel, the spin-split Fermi surfaces are not spheres, just like the case of the (spin degenerate) Fermi surfaces for Cs and Y shown in fig. 5.15.

### 5.7.4 The Sommerfeld expansion

In dealing with the ideal Fermi gas, we will repeatedly encounter integrals of the form

$$\mathcal{I}(T,\mu) \equiv \int\limits_{-\infty}^{\infty} d\varepsilon\; f(\varepsilon - \mu)\, \phi(\varepsilon) \quad . \tag{5.229}$$

[21] Sorry about the notational resemblance between the Hamiltonian $\hat{H}$ and the magnetic field $H$. The Hamiltonian is adorned with a stylish hat to obviate any potential confusion.

The Sommerfeld expansion provides a systematic way of expanding these expressions in powers of $T$ and is an important analytical tool in analyzing the low temperature properties of the ideal Fermi gas (IFG).

We start by defining

$$\Phi(\varepsilon) \equiv \int\limits_{-\infty}^{\varepsilon} d\varepsilon' \, \phi(\varepsilon') \tag{5.230}$$

so that $\phi(\varepsilon) = \Phi'(\varepsilon)$. We then have

$$\mathcal{I}(T,\mu) = \int\limits_{-\infty}^{\infty} d\varepsilon \, f(\varepsilon-\mu) \, \frac{d\Phi}{d\varepsilon} = -\int\limits_{-\infty}^{\infty} d\varepsilon \, f_T'(\varepsilon) \, \Phi(\mu+\varepsilon) \quad , \tag{5.231}$$

where we assume $\Phi(-\infty) = 0$. Next, we invoke Taylor's theorem, to write

$$\Phi(\mu+\varepsilon) = \sum_{n=0}^{\infty} \frac{\varepsilon^n}{n\,!} \, \frac{d^n\Phi}{d\mu^n} = \exp\left(\varepsilon \, \frac{d}{d\mu}\right) \Phi(\mu) \quad . \tag{5.232}$$

This last expression involving the exponential of a differential operator may appear overly formal but it proves extremely useful. Since

$$f_T'(\varepsilon) = -\frac{1}{k_{\rm B}T} \, \frac{e^{\varepsilon/k_{\rm B}T}}{\left(e^{\varepsilon/k_{\rm B}T}+1\right)^2} \quad , \tag{5.233}$$

we define

$$\widetilde{\mathcal{I}}(T,\mu) = \int\limits_{-\infty}^{\infty} dv \, \frac{e^{vD}}{(e^v+1)(e^{-v}+1)} \, \Phi(\mu) \quad , \tag{5.234}$$

with $v = \varepsilon/k_{\rm B}T$, where $D = k_{\rm B}T \, \frac{d}{d\mu}$ is a dimensionless differential operator. The integral can now be done using the methods of complex integration:[22]

$$\begin{aligned} \int\limits_{-\infty}^{\infty} dv \, \frac{e^{vD}}{(e^v+1)(e^{-v}+1)} &= 2\pi i \sum_{n=1}^{\infty} \mathrm{Res}\left[\frac{e^{vD}}{(e^v+1)(e^{-v}+1)}\right]_{v=(2n+1)i\pi} \\ &= -2\pi i \sum_{n=0}^{\infty} D \, e^{(2n+1)i\pi D} = -\frac{2\pi i D \, e^{i\pi D}}{1-e^{2\pi i D}} = \pi D \, \csc \pi D \quad . \end{aligned} \tag{5.235}$$

Thus,

$$\widetilde{\mathcal{I}}(T,\mu) = \pi D \csc(\pi D) \, \Phi(\mu) \quad , \tag{5.236}$$

which is to be understood as the differential operator $\pi D \csc(\pi D)$ acting on the function $\Phi(\mu)$. Appealing once more to Taylor's theorem, we have

$$\pi D \csc(\pi D) = 1 + \frac{\pi^2}{6} \, (k_{\rm B}T)^2 \, \frac{d^2}{d\mu^2} + \frac{7\pi^4}{360} \, (k_{\rm B}T)^4 \, \frac{d^4}{d\mu^4} + \ldots \quad . \tag{5.237}$$

[22] Note that writing $v = (2n+1)\,i\pi + \varepsilon$ we have $e^{\pm v} = -1 \mp \varepsilon - \frac{1}{2}\varepsilon^2 + \ldots$, so $(e^v+1)(e^{-v}+1) = -\varepsilon^2 + \ldots$ We then expand $e^{vD} = e^{(2n+1)i\pi D}\big(1 + \varepsilon D + \ldots\big)$ to find the residue: $\mathrm{Res} = -D\,e^{(2n+1)i\pi D}$.

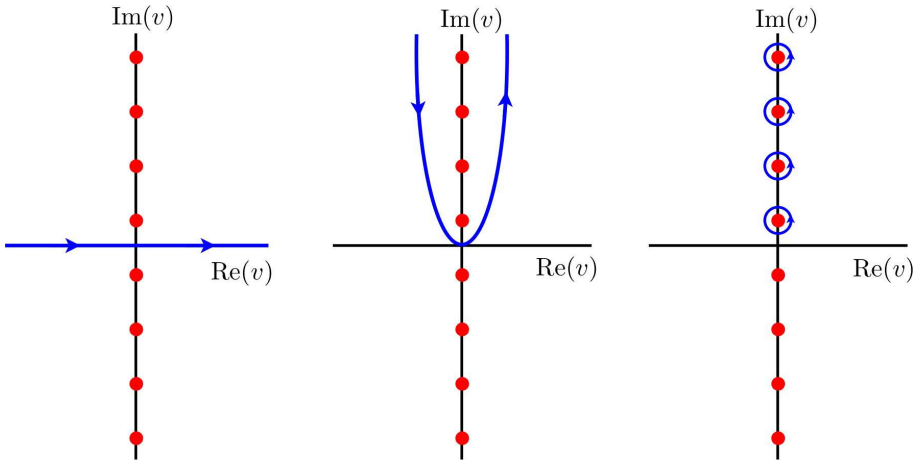



Figure 5.16: Deformation of the complex integration contour in eqn. 5.235.

Thus,

$$\widetilde{\mathcal{I}}(T,\mu) = \int\limits_{-\infty}^{\infty} d\varepsilon\, f(\varepsilon-\mu)\,\phi(\varepsilon) = \int\limits_{-\infty}^{\mu} d\varepsilon\,\phi(\varepsilon) + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,\phi'(\mu) + \frac{7\pi^4}{360}\,(k_{\rm B}T)^4\,\phi'''(\mu) + \ldots \quad . \tag{5.238}$$

If $\phi(\varepsilon)$ is a polynomial function of its argument, then each derivative effectively reduces the order of the polynomial by one degree, and the dimensionless parameter of the expansion is $(k_{\rm B}T/\mu)^2$. This procedure is known as the *Sommerfeld expansion*. We introduce the notation $\widetilde{\mathcal{I}}(T,\mu)$ because the function $\mathcal{I}(T,\mu)$ may contain nonanalytic terms which are invisible in the Taylor series expansion $\widetilde{\mathcal{I}}(T,\mu)$, as we will see below.

### Chemical potential shift

As our first application of the Sommerfeld expansion formalism, let us compute $\mu(n,T)$ for the ideal Fermi gas. The number density $n(T,\mu)$ is

$$n = \int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, f(\varepsilon-\mu) = \int\limits_{-\infty}^{\mu} d\varepsilon\, g(\varepsilon) + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\, g'(\mu) + \ldots \quad . \tag{5.239}$$

Let us write $\mu = \varepsilon_{\rm F} + \delta\mu$, where $\varepsilon_{\rm F} = \mu(T=0,n)$ is the Fermi energy, which is the chemical potential at $T=0$. We then have

$$\begin{aligned} n &= \int\limits_{-\infty}^{\varepsilon_{\rm F}+\delta\mu} d\varepsilon\, g(\varepsilon) + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\, g'(\varepsilon_{\rm F}+\delta\mu) + \ldots \\ &= \int\limits_{-\infty}^{\varepsilon_{\rm F}} d\varepsilon\, g(\varepsilon) + g(\varepsilon_{\rm F})\,\delta\mu + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\, g'(\varepsilon_{\rm F}) + \ldots \quad , \end{aligned} \tag{5.240}$$

from which we derive

$$\delta\mu = -\frac{\pi^2}{6}\,(k_{\rm B}T)^2\,\frac{g'(\varepsilon_{\rm F})}{g(\varepsilon_{\rm F})} + \mathcal{O}(T^4) \quad . \tag{5.241}$$

Note that $g'/g = (\log g)'$. For a ballistic dispersion, assuming $\mathsf{g} = 2$,

$$g(\varepsilon) = 2\int \frac{d^3k}{(2\pi)^3}\,\delta\!\left(\varepsilon - \frac{\hbar^2k^2}{2m}\right) = \frac{m\,k(\varepsilon)}{\pi^2\hbar^2}\bigg|_{k(\varepsilon)=\frac{1}{\hbar}\sqrt{2m\varepsilon}} \tag{5.242}$$

Thus, $g(\varepsilon) \propto \varepsilon^{1/2}$ and $(\log g)' = \frac{1}{2}\,\varepsilon^{-1}$, so

$$\mu(n,T) = \varepsilon_{\rm F} - \frac{\pi^2}{12}\,\frac{(k_{\rm B}T)^2}{\varepsilon_{\rm F}} + \ldots \quad , \tag{5.243}$$

where $\varepsilon_{\rm F}(n) = \frac{\hbar^2}{2m}(3\pi^2 n)^{2/3}$.

**Specific heat**

The energy of the electron gas is

$$\begin{aligned}\frac{E}{V} &= \int\limits_{-\infty}^{\infty} d\varepsilon\; g(\varepsilon)\,\varepsilon\,f(\varepsilon-\mu) = \int\limits_{-\infty}^{\mu} d\varepsilon\; g(\varepsilon)\,\varepsilon + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,\frac{d}{d\mu}\Big(\mu\,g(\mu)\Big) + \ldots \\ &= \int\limits_{-\infty}^{\varepsilon_{\rm F}} d\varepsilon\; g(\varepsilon)\,\varepsilon + g(\varepsilon_{\rm F})\,\varepsilon_{\rm F}\,\delta\mu + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,\varepsilon_{\rm F}\,g'(\varepsilon_{\rm F}) + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,g(\varepsilon_{\rm F}) + \ldots \\ &= \varepsilon_0 + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,g(\varepsilon_{\rm F}) + \ldots \quad ,\end{aligned} \tag{5.244}$$

where $\varepsilon_0 = \int\limits_{-\infty}^{\varepsilon_{\rm F}} d\varepsilon\; g(\varepsilon)\,\varepsilon$ is the ground state energy density (*i.e.* ground state energy per unit volume). Thus, to order $T^2$,

$$C_{V,N} = \left(\frac{\partial E}{\partial T}\right)_{V,N} = \frac{\pi^2}{3}\,V k_{\rm B}^2\,T\,g(\varepsilon_{\rm F}) \equiv V\gamma\,T \quad , \tag{5.245}$$

where $\gamma(n) = \frac{\pi^2}{3}\,k_{\rm B}^2\,g\big(\varepsilon_{\rm F}(n)\big)$ . Note that the molar heat capacity is

$$c_V = \frac{N_{\rm A}}{N}\cdot C_V = \frac{\pi^2}{3}R\cdot\frac{k_{\rm B}T\,g(\varepsilon_{\rm F})}{n} = \frac{\pi^2}{2}\left(\frac{k_{\rm B}T}{\varepsilon_{\rm F}}\right)R \quad , \tag{5.246}$$

where in the last expression on the RHS we have assumed a ballistic dispersion, for which

$$\frac{g(\varepsilon_{\rm F})}{n} = \frac{\mathsf{g}\,mk_{\rm F}}{2\pi^2\hbar^2}\cdot\frac{6\pi^2}{\mathsf{g}\,k_{\rm F}^3} = \frac{3}{2\,\varepsilon_{\rm F}} \quad . \tag{5.247}$$

The molar heat capacity in eqn. 5.246 is to be compared with the classical ideal gas value of $\frac{3}{2}R$. Relative to the classical ideal gas, the IFG value is reduced by a fraction of $(\pi^2/3)\times(k_{\rm B}T/\varepsilon_{\rm F})$, which in most metals

is very small and even at room temperature is only on the order of $10^{-2}$. Most of the heat capacity of metals at room temperature is due to the energy stored in lattice vibrations.

A niftier way to derive the heat capacity[23]: Starting with eqn. 5.241 for $\mu(T) - \varepsilon_{\rm F} \equiv \delta\mu(T)$, note that $g(\varepsilon_{\rm F}) = dn/d\varepsilon_{\rm F}$, so we may write $\delta\mu = -\frac{\pi^2}{6}(k_{\rm B}T)^2(dg/dn) + \mathcal{O}(T^4)$. Next, use the Maxwell relation $(\partial S/\partial N)_{T,V} = -(\partial\mu/\partial T)_{N,V}$ to arrive at

$$\left(\frac{\partial s}{\partial n}\right)_T = \frac{\pi^2}{3}\,k_{\rm B}^2 T\,\frac{\partial g(\varepsilon_{\rm F})}{\partial n} + \mathcal{O}(T^3) \quad , \tag{5.248}$$

where $s = S/V$ is the entropy per unit volume. Now use $S(T=0) = 0$ and integrate with respect to the density $n$ to arrive at $S(T,V,N) = V\gamma T$, where $\gamma(n)$ is defined above.

### Nonanalytic terms

As we've seen, the Sommerfeld expansion is an expansion in powers of $T$. Consider the case where $\phi(\varepsilon) = \Theta(\varepsilon)$. We then have

$$\mathcal{I}(T,\mu) = \int_0^\infty d\varepsilon\, f(\varepsilon-\mu) = \mu + k_{\rm B}T\log\bigl(1+e^{-\mu/k_{\rm B}T}\bigr) \quad . \tag{5.249}$$

By contrast, the Sommerfeld expansion, assuming $\mu \neq 0$, yields $\widetilde{I}(T,\mu) = \mu$, and is missing the second term above. This is because $\exp(-\mu/k_{\rm B}T)$ is nonanalytic in $T$ and cannot appear in any order of a Taylor expansion about $T=0$. As a second example, consider the case $\phi(\varepsilon) = \varepsilon\,\Theta(\varepsilon)$. The Sommerfeld expansion yields

$$\widetilde{\mathcal{I}}(T,\mu) = \tfrac{1}{2}\mu^2 + \frac{\pi^2}{6}(k_{\rm B}T)^2 \quad , \tag{5.250}$$

while the exact result is

$$\mathcal{I}(T,\mu) = \tfrac{1}{2}\mu^2 + \frac{\pi^2}{6}(k_{\rm B}T)^2 + (k_{\rm B}T)^2\sum_{j=1}^{\infty}\frac{(-1)^j}{j^2}\,e^{-j\mu/k_{\rm B}T} \quad , \tag{5.251}$$

which follows from the polylogarithm identity

$$\mathsf{Li}_2(z) + \mathsf{Li}_2(1/z) = -\tfrac{1}{2}\bigl[\log(-z)\bigr]^2 - \frac{\pi^2}{6} \quad . \tag{5.252}$$

Again we see that the Sommerfeld expansion terminates at a finite order in $T$, and is missing nonanalytic terms in the $T \to 0$ limit[24]. This is a generic state of affairs for the case where $\phi(\varepsilon)$ is a finite order polynomial in $\varepsilon$.

[23] I thank my colleague Tarun Grover for this observation.

[24] Once again I thank my colleague Tarun Grover for pointing this out to me.

### 5.7.5 Magnetic susceptibility

#### Pauli paramagnetism

Magnetism has two origins: (i) orbital currents of charged particles, and (ii) intrinsic magnetic moment. The intrinsic magnetic moment $\boldsymbol{m}$ of a particle is related to its quantum mechanical *spin* according to $\boldsymbol{m} = g\mu_0 \boldsymbol{S}/\hbar$, where $\mu_0 = q\hbar/2mc$ is the magneton. Here $g$ is the particle's $g$-factor, $\mu_0$ its magnetic moment, and $\boldsymbol{S}$ is the vector of quantum mechanical spin operators satisfying $\left[S^\alpha\,,\,S^\beta\right] = i\hbar\epsilon_{\alpha\beta\gamma}\,S^\gamma$, *i.e.* SU(2) commutation relations. The Hamiltonian for a single particle is then

$$\hat{H} = \frac{1}{2m^*}\Big(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\Big)^2 - \boldsymbol{H}\cdot\boldsymbol{m} = \frac{1}{2m^*}\Big(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}\Big)^2 + \frac{g}{2}\,\mu_{\rm B} H\,\sigma \quad , \tag{5.253}$$

where in the last line we've restricted our attention to the electron, for which $q = -e$. The $g$-factor for an electron is $g = 2$ at tree level, and when radiative corrections are accounted for using quantum electrodynamics (QED) one finds $g = 2.0023193043617(15)$. For our purposes we can take $g = 2$, although we can always absorb the small difference into the definition of $\mu_{\rm B}$, writing $\mu_{\rm B} \to \tilde{\mu}_{\rm B} = ge\hbar/4mc$. We've chosen the $\hat{\boldsymbol{z}}$-axis in spin space to point in the direction of the magnetic field, and we wrote the eigenvalues of $S^z$ as $\frac{1}{2}\hbar\sigma$, where $\sigma = \pm 1$. The quantity $m^*$ is the *effective mass* of the electron, which we mentioned earlier. An important distinction is that it is $m^*$ which enters into the kinetic energy term $\boldsymbol{p}^2/2m^*$, but it is the electron mass $m$ itself ($m = 511\,\text{keV}$) which enters into the definition of the Bohr magneton. We shall discuss the consequences of this further below.

In the absence of orbital magnetic coupling, the single particle dispersion is

$$\varepsilon_\sigma(\boldsymbol{k}) = \frac{\hbar^2\boldsymbol{k}^2}{2m^*} + \tilde{\mu}_{\rm B} H\,\sigma \quad . \tag{5.254}$$

At $T = 0$, we have the results of §5.7.3. At finite $T$, we once again use the Sommerfeld expansion. We then have

$$\begin{aligned} n &= \int\limits_{-\infty}^{\infty} d\varepsilon\; g_\uparrow(\varepsilon)\, f(\varepsilon-\mu) \;+\; \int\limits_{-\infty}^{\infty} d\varepsilon\; g_\downarrow(\varepsilon)\, f(\varepsilon-\mu) \\ &= \tfrac{1}{2}\int\limits_{-\infty}^{\infty} d\varepsilon\,\Big\{g(\varepsilon - \tilde{\mu}_{\rm B}H) + g(\varepsilon + \tilde{\mu}_{\rm B}H)\Big\}\, f(\varepsilon-\mu) \\ &= \int\limits_{-\infty}^{\infty} d\varepsilon\,\Big\{g(\varepsilon) + \tfrac{1}{2}(\tilde{\mu}_{\rm B}H)^2\, g''(\varepsilon) + \ldots\Big\}\, f(\varepsilon-\mu) \quad . \end{aligned} \tag{5.255}$$

We now invoke the Sommerfeld expansion to find the temperature dependence:

$$\begin{aligned} n &= \int\limits_{-\infty}^{\mu} d\varepsilon\; g(\varepsilon) + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\, g'(\mu) + \tfrac{1}{2}(\tilde{\mu}_{\rm B}H)^2\, g'(\mu) + \ldots \\ &= \int\limits_{-\infty}^{\varepsilon_{\rm F}} d\varepsilon\; g(\varepsilon) + g(\varepsilon_{\rm F})\,\delta\mu + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\, g'(\varepsilon_{\rm F}) + \tfrac{1}{2}(\tilde{\mu}_{\rm B}H)^2\, g'(\varepsilon_{\rm F}) + \ldots \quad , \end{aligned} \tag{5.256}$$

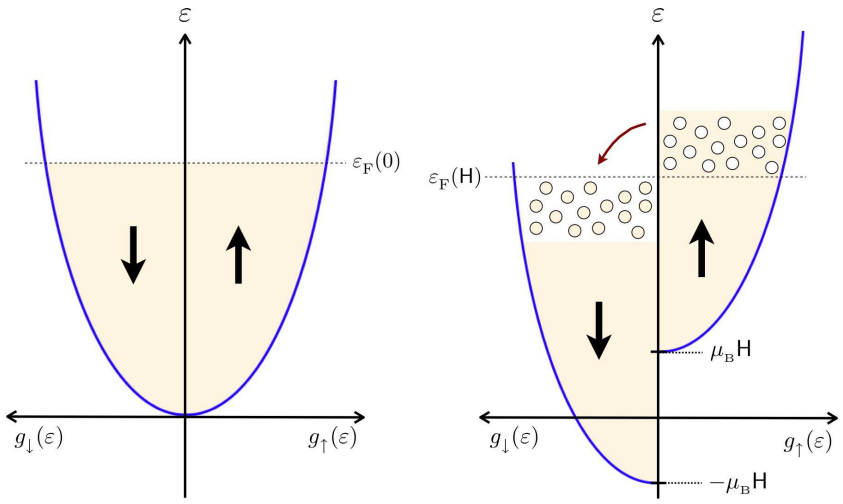



Figure 5.17: Fermi distributions in the presence of an external Zeeman-coupled magnetic field.

where $\varepsilon_{\rm F} = \mu(T=0, n, H=0)$. The density of states for spin species $\sigma$ is $g_\sigma(\varepsilon) = \frac{1}{2}\, g(\varepsilon - \tilde{\mu}_{\rm B} H\sigma)$, where $g(\varepsilon)$ is the total density of states per unit volume, for both spin species, in the absence of a magnetic field. We conclude that the chemical potential shift in an external field is

$$\delta\mu(T, n, H) = -\left\{ \frac{\pi^2}{6}\,(k_{\rm B}T)^2 + \tfrac{1}{2}(\tilde{\mu}_{\rm B}H)^2 \right\} \frac{g'(\varepsilon_{\rm F})}{g(\varepsilon_{\rm F})} + \ldots \quad . \tag{5.257}$$

We next compute the difference $n_\uparrow - n_\downarrow$ in the densities of up and down spin electrons:

$$\begin{aligned} n_\uparrow - n_\downarrow &= \int\limits_{-\infty}^{\infty} d\varepsilon \left\{ g_\uparrow(\varepsilon) - g_\downarrow(\varepsilon) \right\} f(\varepsilon - \mu) \\ &= \tfrac{1}{2} \int\limits_{-\infty}^{\infty} d\varepsilon \left\{ g(\varepsilon - \tilde{\mu}_{\rm B}H) - g(\varepsilon + \tilde{\mu}_{\rm B}H) \right\} f(\varepsilon - \mu) \\ &= -\tilde{\mu}_{\rm B} H \cdot \pi D \csc(\pi D)\, g(\mu) + \mathcal{O}(H^3) \quad . \end{aligned} \tag{5.258}$$

We needn't go beyond the trivial lowest order term in the Sommerfeld expansion, because $H$ is already assumed to be small. Thus, the magnetization density is

$$M = -\tilde{\mu}_{\rm B}(n_\uparrow - n_\downarrow) = \tilde{\mu}_{\rm B}^2\, g(\varepsilon_{\rm F})\, H \quad . \tag{5.259}$$

in which the magnetic susceptibility is

$$\chi = \left( \frac{\partial M}{\partial H} \right)_{T,N} = \tilde{\mu}_{\rm B}^2\, g(\varepsilon_{\rm F}) \quad . \tag{5.260}$$

This is called the *Pauli paramagnetic susceptibility*.

**Landau diamagnetism**

When orbital effects are included, the single particle energy levels are given by

$$\varepsilon(n,k_z,\sigma) = (n+\tfrac{1}{2})\hbar\omega_{\rm c} + \frac{\hbar^2 k_z^2}{2m^*} + \sigma\tilde{\mu}_{\rm B}H \quad . \tag{5.261}$$

Here $n$ is a Landau level index, and $\omega_{\rm c} = eH/m^*c$ is the *cyclotron frequency*. Note that

$$\frac{\tilde{\mu}_{\rm B}H}{\hbar\omega_{\rm c}} = \frac{ge\hbar H}{4mc}\cdot\frac{m^*c}{\hbar eH} = \frac{g}{4}\cdot\frac{m^*}{m} \quad . \tag{5.262}$$

Accordingly, we define the ratio $r \equiv (g/2)\times(m^*/m)$. We can then write

$$\varepsilon(n,k_z,\sigma) = \big(n+\tfrac{1}{2}+\tfrac{1}{2}r\sigma\big)\hbar\omega_{\rm c} + \frac{\hbar^2k_z^2}{2m^*} \quad . \tag{5.263}$$

The grand potential is then given by

$$\Omega = -\frac{HA}{\phi_0}\cdot L_z\cdot k_{\rm B}T\int\limits_{-\infty}^{\infty}\frac{dk_z}{2\pi}\sum_{n=0}^{\infty}\sum_{\sigma=\pm1}\log\Big[1+e^{\mu/k_{\rm B}T}\,e^{-(n+\frac{1}{2}+\frac{1}{2}r\sigma)\hbar\omega_{\rm c}/k_{\rm B}T}\,e^{-\hbar^2k_z^2/2m^*k_{\rm B}T}\Big] \quad . \tag{5.264}$$

A few words are in order here regarding the prefactor. In the presence of a uniform magnetic field, the energy levels of a two-dimensional ballistic charged particle collapse into Landau levels. The number of states per Landau level scales with the area of the system, and is equal to the number of flux quanta through the system: $N_\phi = HA/\phi_0$, where $\phi_0 = hc/e$ is the Dirac flux quantum. Note that

$$\frac{HA}{\phi_0}\cdot L_z\cdot k_{\rm B}T = \hbar\omega_{\rm c}\cdot\frac{V}{\lambda_T^2} \quad , \tag{5.265}$$

hence we can write

$$\Omega(T,V,\mu,H) = \hbar\omega_{\rm c}\sum_{n=0}^{\infty}\sum_{\sigma=\pm1}Q\big((n+\tfrac{1}{2}+\tfrac{1}{2}r\sigma)\hbar\omega_{\rm c}-\mu\big) \quad , \tag{5.266}$$

where

$$Q(\varepsilon) = -\frac{V}{\lambda_T^2}\int\limits_{-\infty}^{\infty}\frac{dk_z}{2\pi}\,\log\Big[1+e^{-\varepsilon/k_{\rm B}T}e^{-\hbar^2k_z^2/2m^*k_{\rm B}T}\Big] = \frac{V}{\lambda_T^3}\,{\rm Li}_{3/2}\big(e^{-\varepsilon/k_{\rm B}T}\big) \quad . \tag{5.267}$$

We now invoke the Euler-MacLaurin formula,

$$\sum_{n=0}^{\infty}F(n) = \int\limits_0^{\infty}dx\,F(x) + \tfrac{1}{2}\,F(0) - \tfrac{1}{12}\,F'(0) + \ldots \quad , \tag{5.268}$$

resulting in

$$\Omega = \sum_\sigma\int\limits_{-\infty}^{\infty}d\varepsilon\,Q(\varepsilon-\mu)\,\Theta\big(\varepsilon-\tfrac{1}{2}(1+\sigma r)\hbar\omega_{\rm c}\big) \tag{5.269}$$

$$+\sum_\sigma\Big[\tfrac{1}{2}\,\hbar\omega_{\rm c}\,Q\big(\tfrac{1}{2}(1+\sigma r)\hbar\omega_{\rm c}-\mu\big) - \tfrac{1}{12}\,(\hbar\omega_{\rm c})^2\,Q'\big(\tfrac{1}{2}(1+\sigma r)\hbar\omega_{\rm c}-\mu\big)+\ldots\Big] \quad .$$

We next expand in powers of the magnetic field $H$ to obtain

$$\Omega(T,V,\mu,H) = 2\int_0^\infty d\varepsilon\, Q(\varepsilon-\mu) + \left(\tfrac{1}{4}r^2 - \tfrac{1}{12}\right)(\hbar\omega_{\rm c})^2\, Q'(-\mu) + \dots \quad . \tag{5.270}$$

Thus, the magnetic susceptibility is

$$\begin{aligned}\chi &= -\frac{1}{V}\frac{\partial^2\Omega}{\partial H^2} = \left(r^2 - \tfrac{1}{3}\right)\cdot\tilde\mu_{\rm B}^2\cdot\left(m/m^*\right)^2\cdot(-2V^{-1})\,Q'(-\mu) \\ &= \left(\frac{g^2}{4} - \frac{m^2}{3m^{*2}}\right)\cdot\tilde\mu_{\rm B}^2\cdot n^2\kappa_T \quad ,\end{aligned} \tag{5.271}$$

where $\kappa_T$ is the isothermal compressibility[25]. For temperatures $T \ll T_{\rm F}$, we have to lowest order $(\partial n/\partial\mu)_T = g(\varepsilon_{\rm F})$, the density of states at the Fermi energy. In most metals we have $m^* \approx m$ and the term in brackets is positive (recall $g \approx 2$). In semiconductors, however, we can have $m^* \ll m$; for example in GaAs we have $m^* = 0.067\,m$. Thus, semiconductors can have a *diamagnetic* response. If we take $g = 2$ and $m^* = m$, we see that the orbital currents give rise to a diamagnetic contribution to the magnetic susceptibility which is exactly $-\frac{1}{3}$ times as large as the contribution arising from Zeeman coupling. The net result is then paramagnetic ($\chi > 0$) and $\frac{2}{3}$ as large as the Pauli susceptibility. The orbital currents can be understood within the context of *Lenz's law*.

### 5.7.6 Moment formation in interacting itinerant electron systems

#### The Hubbard model

A noninteracting electron gas exhibits paramagnetism or diamagnetism, depending on the sign of $\chi$, but never develops a spontaneous magnetic moment: $\boldsymbol{M}(\boldsymbol{H} = 0) = 0$. What gives rise to magnetism in solids? Overwhelmingly, the answer is that Coulomb repulsion between electrons is responsible for magnetism, in those instances in which magnetism arises. At first thought this might seem odd, since the Coulomb interaction is spin-independent. How then can it lead to a spontaneous magnetic moment?

To understand how Coulomb repulsion leads to magnetism, it is useful to consider a model interacting system, described by the Hamiltonian

$$\hat H = -t\sum_{ij,\sigma}\left(c^\dagger_{i\sigma}c_{j\sigma} + c^\dagger_{j\sigma}c_{i\sigma}\right) + U\sum_i n_{i\uparrow}\,n_{i\downarrow} + \mu_{\rm B}\boldsymbol{H}\cdot\sum_{i,\alpha,\beta} c^\dagger_{i\alpha}\,\boldsymbol{\sigma}_{\alpha\beta}\,c_{i\beta} \quad . \tag{5.272}$$

This is none other than the famous *Hubbard model*, which has served as a kind of Rosetta stone for interacting electron systems. The first term describes hopping of electrons along the links of some regular lattice (the symbol $ij$ denotes a link between sites $i$ and $j$). The second term describes the local (on-site) repulsion of electrons. This is a single orbital model, so the repulsion exists when one tries to put two electrons in the orbital, with opposite spin polarization. Typically the Hubbard $U$ parameter is on the

[25] We've used $-\frac{2}{V}Q'(\mu) = -\frac{1}{V}\frac{\partial^2\Omega}{\partial\mu^2} = n^2\kappa_T$.

order of electron volts. The last term is the Zeeman interaction of the electron spins with an external magnetic field. Orbital effects can be modeled by associating a phase $\exp(iA_{ij})$ to the hopping matrix element $t$ between sites $i$ and $j$, where the directed sum of $A_{ij}$ around a plaquette yields the total magnetic flux through the plaquette in units of $\phi_0 = hc/e$. We will ignore orbital effects here. Note that the interaction term is short-ranged, whereas the Coulomb interaction falls off as $1/|\boldsymbol{R}_i - \boldsymbol{R}_j|$. The Hubbard model is thus unrealistic, although screening effects in metals do effectively render the interaction to be short-ranged.

Within the Hubbard model, the interaction term is local and written as $Un_\uparrow n_\downarrow$ on any given site. This term favors a local moment. This is because the chemical potential will fix the mean value of the total occupancy $n_\uparrow + n_\downarrow$, in which case it always pays to maximize the difference $|n_\uparrow - n_\downarrow|$.

**Stoner mean field theory**

There are no general methods available to solve for even the ground state of an interacting many-body Hamiltonian. We'll solve this problem using a *mean field* theory due to Stoner. The idea is to write the occupancy $n_{i\sigma}$ as a sum of average and fluctuating terms:

$$n_{i\sigma} = \langle n_{i\sigma} \rangle + \delta n_{i\sigma} \quad . \tag{5.273}$$

Here, $\langle n_{i\sigma} \rangle$ is the thermodynamic average; the above equation may then be taken as a definition of the fluctuating piece, $\delta n_{i\sigma}$. We assume that the average is site-independent. This is a significant assumption, for while we understand why each site should favor developing a moment, it is not clear that all these local moments should want to line up parallel to each other. Indeed, on a bipartite lattice, it is possible that the individual local moments on neighboring sites will be antiparallel, corresponding to an *antiferromagnetic* order of the pins. Our mean field theory will be one for *ferromagnetic* states.

We now write the interaction term as

$$\begin{aligned} n_{i\uparrow} n_{i\downarrow} &= \langle n_\uparrow \rangle \langle n_\downarrow \rangle + \langle n_\uparrow \rangle \, \delta n_{i\downarrow} + \langle n_\downarrow \rangle \, \delta n_{i\uparrow} + \overbrace{\delta n_{i\uparrow} \, \delta n_{i\downarrow}}^{(\text{fluct}^{\text{s}})^2} \\ &= -\langle n_\uparrow \rangle \langle n_\downarrow \rangle + \langle n_\uparrow \rangle \, n_{i\downarrow} + \langle n_\downarrow \rangle \, n_{i\uparrow} + \mathcal{O}\big((\delta n)^2\big) \\ &= \tfrac{1}{4}(m^2 - n^2) + \tfrac{1}{2} n \, (n_{i\uparrow} + n_{i\downarrow}) + \tfrac{1}{2} m \, (n_{i\uparrow} - n_{i\downarrow}) + \mathcal{O}\big((\delta n)^2\big) \quad , \end{aligned} \tag{5.274}$$

where $n$ and $m$ are the average occupancy per spin and average spin polarization, each per unit cell:

$$\begin{aligned} n &= \langle n_\downarrow \rangle + \langle n_\uparrow \rangle \\ m &= \langle n_\downarrow \rangle - \langle n_\uparrow \rangle \quad , \end{aligned} \tag{5.275}$$

*i.e.* $\langle n_\sigma \rangle = \frac{1}{2}(n - \sigma m)$. The mean field grand canonical Hamiltonian $\mathcal{K} = \hat{H} - \mu \mathcal{N}$, may then be written as

$$\begin{aligned} \mathcal{K}^{\text{MF}} = &-\tfrac{1}{2} \sum_{i,j,\sigma} t_{ij} \Big( c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma} \Big) - \big( \mu - \tfrac{1}{2} U n \big) \sum_{i\sigma} c^\dagger_{i\sigma} c_{i\sigma} \\ &+ \big( \mu_{\text{B}} H + \tfrac{1}{2} U m \big) \sum_{i\sigma} \sigma \, c^\dagger_{i\sigma} c_{i\sigma} + \tfrac{1}{4} N_{\text{sites}} \, U (m^2 - n^2) \quad , \end{aligned} \tag{5.276}$$

where we've quantized spins along the direction of $\boldsymbol{H}$, defined as $\hat{\boldsymbol{z}}$. You should take note of two things here. First, the chemical potential is shifted *downward* (or the electron energies shifted *upward*) by an amount $\frac{1}{2}Un$, corresponding to the average energy of repulsion with the background. Second, the effective magnetic field has been shifted by an amount $\frac{1}{2}Um/\mu_{\rm B}$, so the effective field is

$$H_{\rm eff} = H + \frac{Um}{2\mu_{\rm B}} \quad . \tag{5.277}$$

The *bare* single particle dispersions are given by $\varepsilon_\sigma(\boldsymbol{k}) = -\hat{t}(\boldsymbol{k}) + \sigma\mu_{\rm B}H$, where

$$\hat{t}(\boldsymbol{k}) = \sum_{\boldsymbol{R}} t(\boldsymbol{R})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \quad , \tag{5.278}$$

and $t_{ij} = t(\boldsymbol{R}_i - \boldsymbol{R}_j)$. On a $d$-dimensional cubic lattice, and with nearest-neighbor (NN) hopping, one obtains $\hat{t}(\boldsymbol{k}) = -t\sum_{\mu=1}^{d}\cos(k_\mu a)$, where $a$ is the lattice constant. Including the mean field effects, the *effective* single particle dispersions become

$$\widetilde{\varepsilon}_\sigma(\boldsymbol{k}) = -\hat{t}(\boldsymbol{k}) - \tfrac{1}{2}Un + \big(\mu_{\rm B}H + \tfrac{1}{2}Um\big)\,\sigma \quad . \tag{5.279}$$

We now solve the mean field theory, by obtaining the free energy per site, $\varphi(n, T, H)$. First, note that $\varphi = \omega + \mu n$, where $\omega = \Omega/N_{\rm sites}$ is the Landau, or grand canonical, free energy per site. This follows from the general relation $\Omega = F - \mu N$; note that the total electron number is $N = nN_{\rm sites}$, since $n$ is the electron number per unit cell (including both spin species). If $g(\varepsilon)$ is the density of states per unit cell (rather than per unit volume), then we have[26]

$$\varphi = \tfrac{1}{4}U(m^2+n^2) + \bar{\mu}n - \tfrac{1}{2}k_{\rm B}T\int_{-\infty}^{\infty}\!\!d\varepsilon\, g(\varepsilon)\left\{\log\left(1+e^{(\bar{\mu}-\varepsilon-\Delta)/k_{\rm B}T}\right) + \log\left(1+e^{(\bar{\mu}-\varepsilon+\Delta)/k_{\rm B}T}\right)\right\} \tag{5.280}$$

where $\bar{\mu} \equiv \mu - \frac{1}{2}Un$ and $\Delta \equiv \mu_{\rm B}H + \frac{1}{2}Um$. From this free energy we derive two self-consistent equations for $\mu$ and $m$. The first comes from demanding that $\varphi$ be a function of $n$ and not of $\mu$, *i.e.* $\partial\varphi/\partial\mu = 0$, which leads to

$$n = \tfrac{1}{2}\int_{-\infty}^{\infty}\!\!d\varepsilon\, g(\varepsilon)\Big\{f(\varepsilon-\Delta-\bar{\mu}) + f(\varepsilon+\Delta-\bar{\mu})\Big\} \quad , \tag{5.281}$$

where $f(y) = \big[\exp(y/k_{\rm B}T)+1\big]^{-1}$ is the Fermi function. The second equation comes from minimizing $f$ with respect to average moment $m$:

$$m = \tfrac{1}{2}\int_{-\infty}^{\infty}\!\!d\varepsilon\, g(\varepsilon)\Big\{f(\varepsilon-\Delta-\bar{\mu}) - f(\varepsilon+\Delta-\bar{\mu})\Big\} \quad . \tag{5.282}$$

Here, we will solve the first equation, eq. 5.281, and use the results to generate a Landau expansion of the free energy $\varphi$ in powers of $m^2$. We assume that $\Delta$ is small, in which case we may write

$$n = \int_{-\infty}^{\infty}\!\!d\varepsilon\, g(\varepsilon)\Big\{f(\varepsilon-\bar{\mu}) + \tfrac{1}{2}\Delta^2 f''(\varepsilon-\bar{\mu}) + \tfrac{1}{24}\Delta^4 f''''(\varepsilon-\bar{\mu}) + \ldots\Big\} \quad . \tag{5.283}$$

[26] Note that we have written $\mu n = \bar{\mu}n + \frac{1}{2}Un^2$, which explains the sign of the coefficient of $n^2$.

We write $\bar\mu(\Delta) = \bar\mu_0 + \delta\bar\mu$ and expand in $\delta\bar\mu$. Since $n$ is fixed in our (canonical) ensemble, we have

$$n = \int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, f(\varepsilon - \bar\mu_0) \quad , \tag{5.284}$$

which defines $\bar\mu_0(n,T)$.[27] The remaining terms in the $\delta\bar\mu$ expansion of eqn. 5.283 must sum to zero. This yields

$$D(\bar\mu_0)\,\delta\bar\mu + \tfrac{1}{2}\Delta^2\, D'(\bar\mu_0) + \tfrac{1}{2}(\delta\bar\mu)^2\, D'(\bar\mu_0) + \tfrac{1}{2}D''(\bar\mu_0)\,\Delta^2\,\delta\bar\mu + \tfrac{1}{24}\,D'''(\bar\mu_0)\,\Delta^4 + \mathcal{O}(\Delta^6) = 0 \quad , \tag{5.285}$$

where[28]

$$D(\mu) = -\int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, f_T'(\varepsilon - \mu) \tag{5.286}$$

is the thermally averaged *bare* density of states at energy $\mu$. Note that the $k^{\rm th}$ derivative is

$$D^{(k)}(\mu) = -\int\limits_{-\infty}^{\infty} d\varepsilon\, g^{(k)}(\varepsilon)\, f_T'(\varepsilon - \mu) \quad . \tag{5.287}$$

Solving for $\delta\bar\mu$, we obtain

$$\delta\bar\mu = -\tfrac{1}{2}a_1\Delta^2 - \tfrac{1}{24}\big(3a_1^3 - 6a_1a_2 + a_3\big)\,\Delta^4 + \mathcal{O}(\Delta^6) \quad , \tag{5.288}$$

where $a_k \equiv D^{(k)}(\bar\mu_0)/D(\bar\mu_0)$.

After integrating by parts and inserting this result for $\delta\bar\mu$ into our expression for the free energy $f$, we obtain the expansion

$$\varphi(n,T,m) = \varphi_0(n,T) + \tfrac{1}{4}Um^2 - \tfrac{1}{2}D(\bar\mu_0)\,\Delta^2 + \frac{1}{8}\left(\frac{\big[D'(\bar\mu_0)\big]^2}{D(\bar\mu_0)} - \tfrac{1}{3}D''(\bar\mu_0)\right)\Delta^4 + \ldots \quad , \tag{5.289}$$

where prime denotes differentiation with respect to argument, at $m = 0$, and

$$\varphi_0(n,T) = \tfrac{1}{4}Un^2 + n\bar\mu_0 - \int\limits_{-\infty}^{\infty} d\varepsilon\, \mathcal{N}(\varepsilon)\, f(\varepsilon - \bar\mu_0) \quad , \tag{5.290}$$

where $g(\varepsilon) = \mathcal{N}'(\varepsilon)$, so $\mathcal{N}(\varepsilon)$ is the integrated bare density of states per unit cell in the absence of any magnetic field (including both spin species).

We assume that $H$ and $m$ are small, in which case

$$\varphi = \varphi_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 - \tfrac{1}{2}\chi_0\,H^2 - \frac{U\chi_0}{2\mu_{\rm B}}\,Hm + \ldots \quad , \tag{5.291}$$

[27] The Gibbs-Duhem relation guarantees that such an equation of state exists, relating any three intensive thermodynamic quantities.

[28] $D(\mu)$ is also a function of $T$ through $f(\varepsilon - \mu)$. We write $D(\mu)$ rather than $D(T,\mu)$ for notational simplicity.



Figure 5.18: A graduate student experiences the Stoner enhancement

where $\chi_0 = \mu_{\rm B}^2\, D(\bar\mu_0)$ is the Pauli susceptibility, and

$$a = \tfrac{1}{2}U\big(1 - \tfrac{1}{2}UD\big) \quad , \quad b = \frac{1}{32}\left(\frac{(D')^2}{D} - \tfrac{1}{3}D''\right)U^4 \quad , \tag{5.292}$$

where the argument of each $D^{(k)}$ above is $\bar\mu_0(n,T)$. The magnetization density (per unit cell, rather than per unit volume) is given by

$$M = -\frac{\partial\varphi}{\partial H} = \chi_0 H + \frac{U\chi_0}{2\mu_{\rm B}}\, m \quad . \tag{5.293}$$

Minimizing with respect to $m$ yields

$$am + bm^3 - \frac{U\chi_0}{2\mu_{\rm B}}\, H = 0 \quad , \tag{5.294}$$

which gives, for small $m$,

$$m = \frac{\chi_0}{\mu_{\rm B}}\,\frac{H}{1 - \frac{1}{2}UD} \quad . \tag{5.295}$$

We therefore obtain $M = \chi\, H$ with

$$\chi = \frac{\chi_0}{1 - \frac{U}{U_{\rm c}}} \quad , \tag{5.296}$$

where $U_{\rm c} = 2/D(\bar\mu_0)$ is the critical value of $U$. The denominator of $\chi$ increases the susceptibility above the bare Pauli value $\chi_0$, and is referred to as – I kid you not – the *Stoner enhancement* (see fig. 5.18).

It is worth emphasizing that the magnetization per unit cell is given by

$$M = -\frac{1}{N_{\rm sites}}\,\frac{\delta\hat H}{\delta H} = \mu_{\rm B}\hat m \quad . \tag{5.297}$$

This is an operator identity and is valid for any value of $\hat M$, and not only small $m$.

When $H = 0$ we can still get a magnetic moment, provided $U > U_{\rm c}$. This is a consequence of the simple Landau theory we have derived. Solving for $m$ when $H = 0$ gives $m = 0$ when $U < U_{\rm c}$ and

$$m(U) = \pm\left(\frac{U}{2b\,U_{\rm c}}\right)^{1/2}\sqrt{U - U_{\rm c}} \quad , \tag{5.298}$$

when $U > U_{\rm c}$, and assuming $b > 0$. Thus we have the usual mean field order parameter exponent of $\beta = \frac{1}{2}$.

**Antiferromagnetic solution**

In addition to ferromagnetism, there may be other ordered states which solve the mean field theory. One such example is antiferromagnetism. On a bipartite lattice, the antiferromagnetic mean field theory is obtained from

$$\langle n_{i\sigma}\rangle = \tfrac{1}{2}n + \tfrac{1}{2}\sigma\, e^{i\boldsymbol{Q}\cdot\boldsymbol{R}_i}\, m \quad , \tag{5.299}$$

where $\boldsymbol{Q} = (\pi/a, \pi/a, \ldots, \pi/a)$ is the antiferromagnetic ordering wavevector. The grand canonical Hamiltonian is then

$$\begin{aligned}\mathcal{K}^{\rm MF} &= -\tfrac{1}{2}\sum_{i,j,\sigma} t_{ij}\left(c^\dagger_{i\sigma}c_{j\sigma} + c^\dagger_{j\sigma}c_{i\sigma}\right) - (\mu - \tfrac{1}{2}Un)\sum_{i\sigma} c^\dagger_{i\sigma}c_{i\sigma} \\ &\qquad + \tfrac{1}{2}Um\sum_{i\sigma} e^{i\boldsymbol{Q}\cdot\boldsymbol{R}_i}\,\sigma\, c^\dagger_{i\sigma}c_{i\sigma} + \tfrac{1}{4}N_{\rm sites}\,U(m^2 - n^2) \\ &= \tfrac{1}{2}\sum_{\boldsymbol{k}\sigma}\begin{pmatrix} c^\dagger_{\boldsymbol{k},\sigma} & c^\dagger_{\boldsymbol{k}+\boldsymbol{Q},\sigma}\end{pmatrix}\begin{pmatrix}\varepsilon(\boldsymbol{k}) - \mu + \frac{1}{2}Un & \frac{1}{2}\sigma\,Um \\ \frac{1}{2}\sigma\,Um & \varepsilon(\boldsymbol{k}+\boldsymbol{Q}) - \mu + \frac{1}{2}Un\end{pmatrix}\begin{pmatrix} c_{\boldsymbol{k},\sigma} \\ c_{\boldsymbol{k}+\boldsymbol{Q},\sigma}\end{pmatrix} + \tfrac{1}{4}N_{\rm sites}\,U(m^2 - n^2) \quad ,\end{aligned} \tag{5.300}$$

where $\varepsilon(\boldsymbol{k}) = -\hat{t}(\boldsymbol{k})$, as before. On a bipartite lattice, with nearest neighbor hopping only, we have the relation $\varepsilon(\boldsymbol{k}+\boldsymbol{Q}) = -\varepsilon(\boldsymbol{k})$. The above matrix is diagonalized by a unitary transformation, yielding the eigenvalues

$$\lambda_\pm = \pm\sqrt{\varepsilon^2(\boldsymbol{k}) + \Delta^2} - \bar{\mu} \tag{5.301}$$

with $\Delta = \frac{1}{2}Um$ and $\bar{\mu} = \mu - \frac{1}{2}Un$ as before. The free energy per unit cell is then

$$\varphi = \tfrac{1}{4}U(m^2 + n^2) + \bar{\mu}n - \tfrac{1}{2}k_{\rm B}T\int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\left\{\log\left(1 + e^{(\bar{\mu} - \sqrt{\varepsilon^2+\Delta^2})/k_{\rm B}T}\right) + \log\left(1 + e^{(\bar{\mu} + \sqrt{\varepsilon^2+\Delta^2})/k_{\rm B}T}\right)\right\} \quad . \tag{5.302}$$

The mean field equations are then

$$\begin{aligned} n &= \tfrac{1}{2}\int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\left\{f\left(-\sqrt{\varepsilon^2+\Delta^2} - \bar{\mu}\right) + f\left(\sqrt{\varepsilon^2+\Delta^2} - \bar{\mu}\right)\right\} \\ \frac{1}{U} &= \tfrac{1}{2}\int\limits_{-\infty}^{\infty} d\varepsilon\, \frac{g(\varepsilon)}{\sqrt{\varepsilon^2+\Delta^2}}\left\{f\left(-\sqrt{\varepsilon^2+\Delta^2} - \bar{\mu}\right) - f\left(\sqrt{\varepsilon^2+\Delta^2} - \bar{\mu}\right)\right\}.\end{aligned} \tag{5.303}$$

As in the case of the ferromagnet, a paramagnetic solution with $m = 0$ always exists, in which case the second of the above equations is no longer valid.

**Mean field phase diagram of the Hubbard model**

Let us compare the mean field theories for the ferromagnetic and antiferromagnetic states at $T = 0$ and $H = 0$. We then have $f(\omega) = \Theta(-\omega)$. Due to particle-hole symmetry, we may assume $0 \le n \le 1$. (The solutions repeat themselves under $n \to 2 - n$.) For the paramagnet, we have

$$\varphi = \tfrac{1}{4}Un^2 + \int\limits_{-\infty}^{\bar\mu} d\varepsilon\, g(\varepsilon)\,\varepsilon \qquad , \qquad n = \int\limits_{-\infty}^{\bar\mu} d\varepsilon\, g(\varepsilon) \quad , \tag{5.304}$$

with $\bar\mu = \mu - \frac{1}{2}Un$ is the 'renormalized' Fermi energy and $g(\varepsilon)$ is the density of states per unit cell in the absence of any explicit ($H$) or implicit ($m$) symmetry breaking, including both spin polarizations.

For the ferromagnet,

$$\begin{aligned}
\varphi &= \tfrac{1}{4}Un^2 \;-\; \frac{\Delta^2}{U} \;+ \int\limits_{-\infty}^{\bar\mu-\Delta} d\varepsilon\, g(\varepsilon)\,\varepsilon \;+ \int\limits_{-\infty}^{\bar\mu+\Delta} d\varepsilon\, g(\varepsilon)\,\varepsilon \\
n &= \tfrac{1}{2}\int\limits_{-\infty}^{\bar\mu-\Delta} d\varepsilon\, g(\varepsilon) + \tfrac{1}{2}\int\limits_{-\infty}^{\bar\mu+\Delta} d\varepsilon\, g(\varepsilon) \\
\frac{4\Delta}{U} &= \int\limits_{\bar\mu-\Delta}^{\bar\mu+\Delta} d\varepsilon\, g(\varepsilon) \quad .
\end{aligned} \tag{5.305}$$

Here, $\Delta = \frac{1}{2}Um$ is nonzero in the ordered phase.

Finally, the antiferromagnetic mean field equations are

$$\begin{aligned}
\varphi &= \tfrac{1}{4}Un^2 + \frac{\Delta^2}{U} - \int\limits_{\varepsilon_0}^{\infty} d\varepsilon\, g(\varepsilon)\,\sqrt{\varepsilon^2+\Delta^2} \\
n_{\bar\mu<0} &= \int\limits_{\varepsilon_0}^{\infty} d\varepsilon\, g(\varepsilon) \qquad , \qquad n_{\bar\mu>0} = 2 - \int\limits_{\varepsilon_0}^{\infty} d\varepsilon\, g(\varepsilon) \\
\frac{2}{U} &= \int\limits_{\varepsilon_0}^{\infty} d\varepsilon\, \frac{g(\varepsilon)}{\sqrt{\varepsilon^2+\Delta^2}} \quad ,
\end{aligned} \tag{5.306}$$

where $\varepsilon_0 = \sqrt{\bar\mu^2 - \Delta^2}$ and $\Delta = \frac{1}{2}Um$ as before. Note that $|\bar\mu| \ge \Delta$ for these solutions. Exactly at half-filling, we have $n = 1$ and $\bar\mu = 0$. We then set $\varepsilon_0 = 0$.

The paramagnet to ferromagnet transition may be first or second order, depending on the details of $g(\varepsilon)$. If second order, it occurs at $U_{\rm c}^{\rm F} = 1/g(\bar\mu_{\rm P})$, where $\bar\mu_{\rm P}(n)$ is the paramagnetic solution for $\bar\mu$. The paramagnet to antiferromagnet transition is always second order in this mean field theory, since the

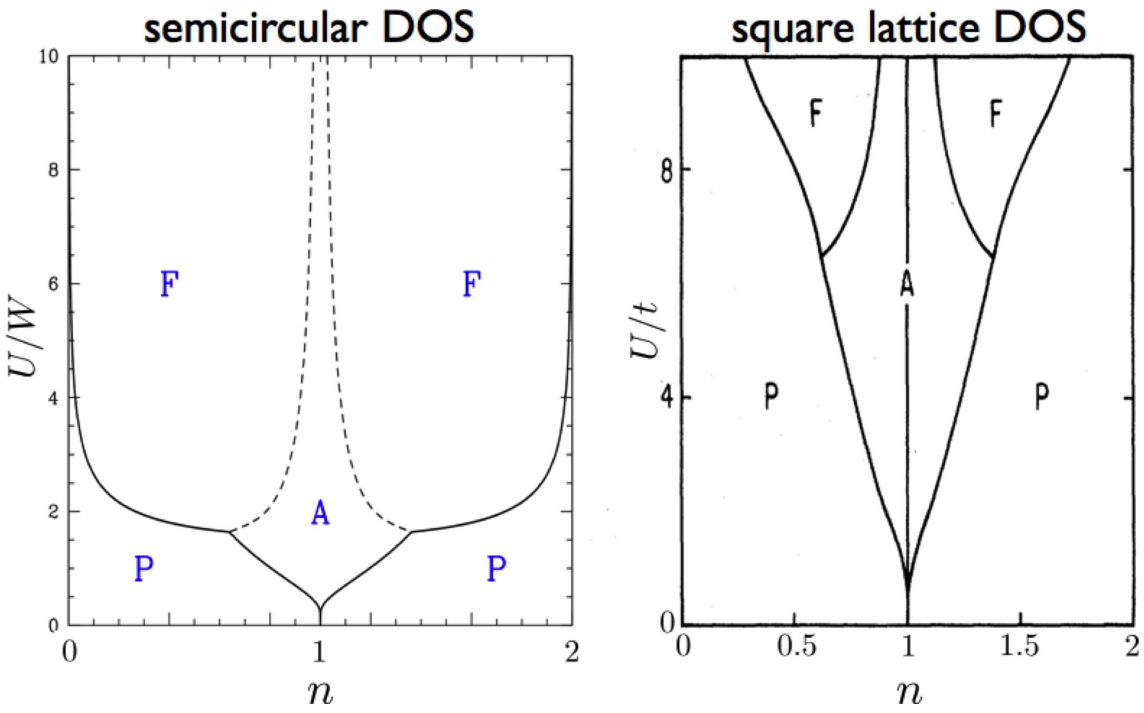



Figure 5.19: Mean field phase diagram of the Hubbard model, including paramagnetic (P), ferromagnetic (F), and antiferromagnetic (A) phases. Left panel: results using a semicircular density of states function of half-bandwidth $W$. Right panel: results using a two-dimensional square lattice density of states with nearest neighbor hopping $t$, from J. E. Hirsch, *Phys. Rev. B* **31**, 4403 (1985). The phase boundary between F and A phases is first order.

RHS of eqn. (5.306) is a monotonic function of $\Delta$. This transition occurs at $U_{\rm c}^{\rm A} = 2 \Big/ \int\limits_{\bar{\mu}_{\rm P}}^{\infty} d\varepsilon \, g(\varepsilon) \, \varepsilon^{-1}$. Note that $U_{\rm c}^{\rm A} \to 0$ logarithmically for $n \to 1$, since $\bar{\mu}_{\rm P} = 0$ at half-filling.

For large $U$, the ferromagnetic solution always has the lowest energy, and therefore if $U_{\rm c}^{\rm A} < U_{\rm c}^{\rm F}$, there will be a first-order antiferromagnet to ferromagnet transition at some value $U^* > U_{\rm c}^{\rm F}$. In fig. 5.19, I plot the phase diagram obtained by solving the mean field equations assuming a semicircular density of states $g(\varepsilon) = \frac{2}{\pi} W^{-2} \sqrt{W^2 - \varepsilon^2}$. Also shown is the phase diagram for the $d = 2$ square lattice Hubbard model obtained by J. Hirsch (1985).

How well does Stoner theory describe the physics of the Hubbard model? Quantum Monte Carlo calculations by J. Hirsch (1985) found that the actual phase diagram of the $d = 2$ square lattice Hubbard Model exhibits no ferromagnetism for any $n$ up to $U = 10$. Furthermore, he found the antiferromagnetic phase to be entirely confined to the vertical line $n = 1$. For $n \neq 1$ and $0 \leq U \leq 10$, the system is a paramagnet[29]. These results were state-of-the art at the time, but both computing power as well as numerical algorithms for interacting quantum systems have advanced considerably since 1985. Yet as of 2018, we *still* don't have a clear understanding of the $d = 2$ Hubbard model's $T = 0$ phase diagram! There is an

[29] A theorem due to Nagaoka establishes that the ground state is ferromagnetic for the case of a single hole in the $U = \infty$ system on bipartite lattices.

emerging body of numerical evidence[30] that in the underdoped ($n < 1$) regime, there are portions of the phase diagram which exhibit a *stripe* ordering, in which antiferromagnetic order is interrupted by a parallel array of line defects containing excess holes (*i.e.* the absence of an electron)[31]. This problem has turned out to be unexpectedly rich, complex, and numerically difficult to resolve due to the presence of *competing ordered states*, such as $d$-wave superconductivity and spiral magnetic phases, which lie nearby in energy with respect to the putative stripe ground state.

In order to achieve a ferromagnetic solution, it appears necessary to introduce geometric frustration, either by including a next-nearest-neighbor hopping amplitude $t'$ or by defining the model on non-bipartite lattices. Numerical work by M. Ulmke (1997) showed the existence of a ferromagnetic phase at $T = 0$ on the FCC lattice Hubbard model for $U = 6$ and $n \in [0.15, 0.87]$ (approximately).

### 5.7.7 White dwarf stars

The term *white dwarf*, coined by W. J. Luyten in 1922, refers to an end stage of stellar evolution in cases where the star's mass is too small to transition to a neutron star or collapse into a black hole. White dwarfs are typically earth-sized objects which have masses $M \sim M_\odot$ on the order of that of the sun, hence they are extremely compact. They are also very faint, and occupy a region in the lower left of the Hertzsprung-Russell diagram[32]. In so-called main sequence stars, heat and radiation is generated by hydrogen-to-helium fusion, which roughly occurs in three steps: (i) $p + p \to d + e^+ + \nu_e$, (ii) $d + p \to {}^3\mathrm{He}$, and (iii) ${}^3\mathrm{He} + {}^3\mathrm{He} \to {}^4\mathrm{He} + 2p$, where ${}^4\mathrm{He}$ is a helium-4 nucleus, *i.e.* an $\alpha$-particle. For low mass stars ($M < M_\odot$), the radius typically scales as $R(M) \propto M^{0.8}$.

In a main-sequence star, thermal pressure generated by hydrogen fusion helps support the star against gravitational contraction. After the hydrogen in the core is exhausted, a star of sufficiently large mass undergoes further stages of evolution, including helium burning through the triple-alpha process, whose net reaction is $3\alpha \to {}^{12}\mathrm{C} + \gamma$. Additional reactions, particularly ${}^{12}\mathrm{C} + \alpha \to {}^{16}\mathrm{O} + \gamma$, produce oxygen. For a low- or intermediate-mass star, nuclear fusion eventually ceases and the outer layers are expelled, leaving behind a white dwarf. The white dwarf ordinarily consists principally of carbon and oxygen, although its composition depends on the mass and evolutionary history of the progenitor: some white dwarfs have helium cores, while especially massive ones may contain substantial oxygen, neon, and magnesium. With no continuing nuclear heat source, the remnant is supported against gravitational collapse primarily by electron degeneracy pressure and gradually cools by radiating its residual thermal energy.

Here we follow the nice discussion of this material in R. K. Pathria, *Statistical Mechanics*. As a model, consider a mass $M \sim 10^{33}\,\mathrm{g}$ of helium at white dwarf densities of $\rho \sim 10^7\,\mathrm{g/cm^3}$ and temperature $T \sim 10^7\,\mathrm{K}$. (We neglect the effects of helium burning.) This temperature is much larger than the ionization energy of ${}^4\mathrm{He}$, hence we may safely assume that all helium atoms are ionized. If there are $N$ electrons, then the number of $\alpha$ particles (*i.e.* ${}^4\mathrm{He}$ nuclei) must be $\frac{1}{2}N$. The mass of the $\alpha$ particle is $m_\alpha \approx 4m_\mathrm{p}$. The total stellar mass $M$ is almost completely due to $\alpha$ particle cores.

[30] See J. P. F. LeBlanc *et al.*, *Phys. Rev. X* **5**, 041041 (2015) and B. Zheng *et al.*, *Science* **358**, 1155 (2017).
[31] The best case for stripe order has been made at $T = 0$, $U/t = 8$, and hole doping $x = \frac{1}{8}$ (*i.e.* $n = \frac{7}{8}$).
[32] The H-R diagram plots luminosity *versus* surface temperature (*i.e.* color or spectral class).

The electron density is then

$$n = \frac{N}{V} = \frac{2 \cdot M/4m_{\rm p}}{V} = \frac{\rho}{2m_{\rm p}} \approx 10^{30}\,{\rm cm}^{-3} \quad , \tag{5.307}$$

since $M = N \cdot m_{\rm e} + \frac{1}{2}N \cdot 4m_{\rm p}$. From the number density $n$, including both spin polarizations, we find for the electrons

$$\begin{aligned} k_{\rm F} = (3\pi^2 n)^{1/3} &= 2.14 \times 10^{10}\,{\rm cm}^{-1} \\ p_{\rm F} = \hbar k_{\rm F} &= 2.26 \times 10^{-17}\,{\rm g\,cm/s} \\ mc = (9.1 \times 10^{-28}\,{\rm g})(3 \times 10^{10}\,{\rm cm/s}) &= 2.7 \times 10^{-17}\,{\rm g\,cm/s} \quad . \end{aligned} \tag{5.308}$$

Since $p_{\rm F} \sim mc$, we conclude that the electrons are relativistic. The Fermi temperature will then be $T_{\rm F} \sim mc^2 \sim 10^6\,{\rm eV} \sim 10^{10}\,{\rm K}$. Thus, $T \ll T_{\rm F}$ which says that the electron gas is degenerate and may be considered to be at $T \sim 0$. So we need to understand the ground state properties of the relativistic electron gas.

The kinetic energy is given by

$$\varepsilon(\boldsymbol{p}) = \sqrt{\boldsymbol{p}^2c^2 + m^2c^4} - mc^2 \quad . \tag{5.309}$$

The velocity is

$$\boldsymbol{v} = \frac{\partial\varepsilon}{\partial\boldsymbol{p}} = \frac{\boldsymbol{p}c^2}{\sqrt{p^2c^2 + m^2c^4}} \quad . \tag{5.310}$$

The pressure in the ground state is

$$\begin{aligned} p_0 = \tfrac{1}{3}n\langle \boldsymbol{p}\cdot\boldsymbol{v}\rangle &= \frac{1}{3\pi^2\hbar^3}\int\limits_0^{p_{\rm F}} dp\,p^2 \cdot \frac{p^2c^2}{\sqrt{p^2c^2 + m^2c^4}} \\ &= \frac{m^4c^5}{3\pi^2\hbar^3}\int\limits_0^{\theta_{\rm F}} d\theta\ \sinh^4\!\theta = \frac{m^4c^5}{96\pi^2\hbar^3}\left(\sinh(4\theta_{\rm F}) - 8\sinh(2\theta_{\rm F}) + 12\,\theta_{\rm F}\right) \quad , \end{aligned} \tag{5.311}$$

where we use the substitution

$$p = mc\sinh\theta \quad , \quad v = c\tanh\theta \quad \Longrightarrow \quad \theta = \tfrac{1}{2}\log\left(\frac{c+v}{c-v}\right) \quad . \tag{5.312}$$

Note that $p_{\rm F} = \hbar k_{\rm F} = \hbar(3\pi^2 n)^{1/3}$, and that

$$n = \frac{M}{2m_{\rm p}V} \qquad \Longrightarrow \qquad 3\pi^2 n = \frac{9\pi}{8}\,\frac{M}{R^3\,m_{\rm p}} \quad . \tag{5.313}$$

Now in equilibrium the pressure $p$ is balanced by gravitational pressure. Let $R$ be the radius of the star. Then we have

$$dE_0 = -p_0\,dV = -p_0(R)\cdot 4\pi R^2\,dR \quad . \tag{5.314}$$

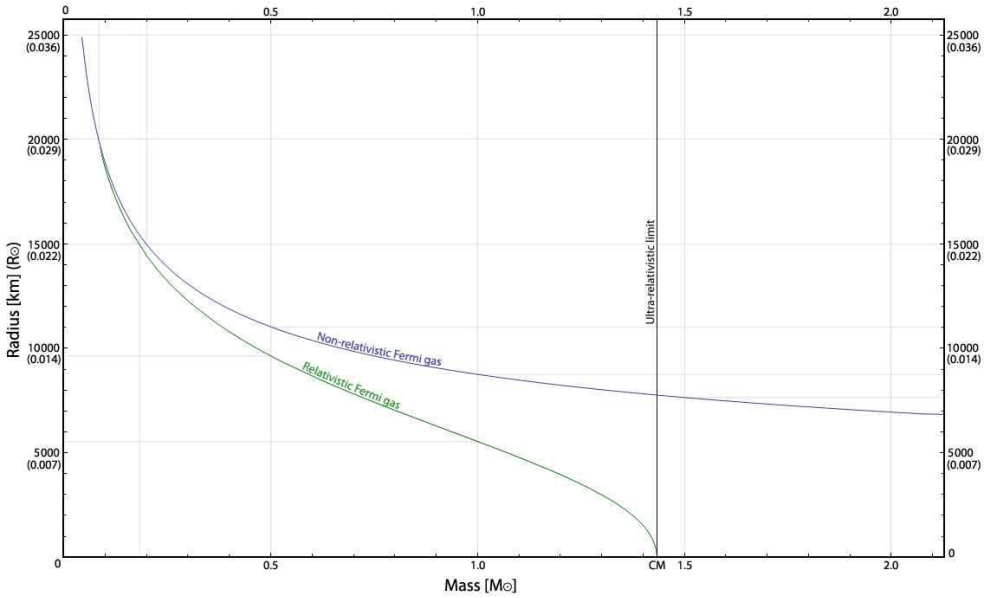



Figure 5.20: Mass-radius relationship for white dwarf stars. (*Source: Wikipedia*).

This must be balanced by gravity:

$$dE_{\rm g} = \alpha \cdot \frac{GM^2}{R^2}\, dR \quad , \tag{5.315}$$

where the phenomenological constant $\alpha$ depends on the radial mass distribution. Equilibrium then implies

$$p_0(R) = \frac{\alpha}{4\pi}\,\frac{GM^2}{R^4} \quad . \tag{5.316}$$

To find the relation $R = R(M)$, we must solve

$$\frac{\alpha}{4\pi}\,\frac{GM^2}{R^4} = \frac{m^4c^5}{96\pi^2\hbar^3}\left(\sinh(4\theta_{\rm F}) - 8\sinh(2\theta_{\rm F}) + 12\,\theta_{\rm F}\right) \quad . \tag{5.317}$$

Note that

$$\sinh(4\theta_{\rm F}) - 8\sinh(2\theta_{\rm F}) + 12\theta_{\rm F} = \begin{cases} \frac{96}{15}\,\theta_{\rm F}^5 & \theta_{\rm F} \to 0 \\ \\ \frac{1}{2}\,e^{4\theta_{\rm F}} & \theta_{\rm F} \to \infty \quad . \end{cases} \tag{5.318}$$

Thus, we may write

$$p_0(R) = \frac{\alpha}{4\pi}\,\frac{gM^2}{R^4} = \begin{cases} \frac{\hbar^2}{15\pi^2 m}\left(\frac{9\pi}{8}\,\frac{M}{R^3\,m_{\rm p}}\right)^{5/3} & \theta_{\rm F} \to 0 \\ \\ \frac{\hbar c}{12\pi^2}\left(\frac{9\pi}{8}\,\frac{M}{R^3\,m_{\rm p}}\right)^{4/3} & \theta_{\rm F} \to \infty \quad . \end{cases} \tag{5.319}$$

In the limit $\theta_{\rm F} \to 0$, we solve for $R(M)$ and find

$$R(M) = \frac{3\,(9\pi)^{2/3}}{40\,\alpha}\,\frac{\hbar^2}{G\,m_{\rm p}^{5/3}\,m\,M^{1/3}} \propto M^{-1/3} \quad . \tag{5.320}$$

In the opposite limit $\theta_{\rm F} \to \infty$, the $R$ factors on either side of the equality in eqn. 5.319 cancel, and we obtain $M = M_0$ with

$$M_0 = \frac{9}{64}\left(\frac{3\pi}{\alpha^3}\right)^{1/2}\left(\frac{\hbar c}{G}\right)^{3/2}\frac{1}{m_{\rm p}^2} \approx 0.8\,\alpha^{-3/2} M_\odot \quad . \tag{5.321}$$

To find the $R$ dependence in this limit, we must go beyond the lowest order expansion of eqn. 5.318, in which case we find

$$R(M) = \left(\frac{9\pi}{8}\right)^{1/3}\left(\frac{\hbar}{mc}\right)\left(\frac{M}{m_{\rm p}}\right)^{1/3}\left[1 - \left(\frac{M}{M_0}\right)^{2/3}\right]^{1/2} \quad . \tag{5.322}$$

Since we must have $R(M) > 0$, we conclude that the value $M_0$ is the maximum mass for a white dwarf. Beyond this mass, the degeneracy pressure cannot resist the force of gravity. More correctly stated, there is no possibility for a stable equilibrium between pressure and gravitational forces. This is called the *Chandrasekhar limit*. The actual value is $M_0 \simeq 1.44\,M_\odot$, which requires a fudge factor $\alpha = 0.677$. Remarkably, the mass-radius relation is *inverted* for white dwarfs, *i.e.* radius *decreases* with increasing stellar mass.

### The Lane-Emden equation

We've assumed a uniform mass density $\rho$ in our derivation, but this is a significant simplification of the physics. If we assume hydrostatic equilibrium throughout, then for all $r \in [0, R]$ we must have

$$\frac{dp(r)}{dr} = -\frac{GM(r)\,\rho(r)}{r^2} \qquad , \qquad \frac{dM(r)}{dr} = 4\pi r^2 \rho(r) \quad . \tag{5.323}$$

Here $M(r)$ is the total mass up to a radius $r$. To close this system we need the adiabatic equation of state relating pressure to density, which takes the form $p = K\rho^\gamma$ where the polytropic/adiabatic exponent is $\gamma = C_p/C_V$ for ideal gases. This leads to the dimensionless *Lane-Emden equation*,

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2 \frac{d\theta}{d\xi}\right) = -\theta^{1/(\gamma-1)} \tag{5.324}$$

where

$$\theta = \left(\frac{\rho}{\rho_{\rm c}}\right)^{\gamma-1} \qquad , \qquad \xi = \frac{r}{b} \qquad , \qquad b = \sqrt{\frac{\gamma K \rho_{\rm c}^{\gamma-2}}{4\pi(\gamma-1)G}} \quad , \tag{5.325}$$

and $\rho_{\rm c} = \rho(r=0)$ the central (maximum) density. One conventionally defines $n \equiv 1/(\gamma-1)$ as the *polytropic index*. For nonrelativistic particles $\gamma = \frac{5}{3}$ ($n = \frac{3}{2}$) while for ultrarelativistic $\gamma = \frac{4}{3}$ ($n = 3$). The radius of the star is then given by $R = b\xi_0$, where $\theta(\xi_0) = 0$. By adjusting the constant $\alpha$ in the simplified derivation one can better approximate the exact results.

## 5.8 Appendix I: Second Quantization

### 5.8.1 Basis states and creation/annihilation operators

Second quantization is a convenient scheme to label basis states of a many particle quantum system. We are ultimately interested in solutions of the many-body Schrödinger equation,

$$\hat{H}\Psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = E\,\Psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) \tag{5.326}$$

where the Hamiltonian is

$$\begin{aligned}\hat{H} &= \sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}_i^2 + v_{\text{ext}}(\boldsymbol{x}_i)\right) + \sum_{j<k}^{N} u(\boldsymbol{x}_j - \boldsymbol{x}_k) \\ &\equiv \hat{T} + \hat{V} + \hat{U} \quad ,\end{aligned} \tag{5.327}$$

where $\hat{T}$ is the kinetic energy, $\hat{V}$ the one-body potential energy, and $\hat{U}$ the two-body potential energy. To the coordinate labels $\{\boldsymbol{x}_1,\ldots\boldsymbol{x}_N\}$ we may also append labels for internal degrees of freedom, such as spin polarization, denoted $\{\zeta_1,\ldots,\zeta_N\}$. Since $\left[\hat{H},\sigma\right] = 0$ for all permutations $\sigma \in S_N$, the many-body wavefunctions may be chosen to transform according to irreducible representations of the symmetric group $S_N$. Thus, for any $\sigma \in S_N$,

$$\Psi\big(\boldsymbol{x}_{\sigma(1)},\ldots,\boldsymbol{x}_{\sigma(N)}\big) = \left\{\begin{array}{c} 1 \\ \text{sgn}(\sigma)\end{array}\right\}\Psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) \quad , \tag{5.328}$$

where the upper choice is for Bose-Einstein statistics and the lower sign for Fermi-Dirac statistics. Here $\boldsymbol{x}_j$ may include not only the spatial coordinates of particle $j$, but its internal quantum number(s) as well, such as $\zeta_j$.

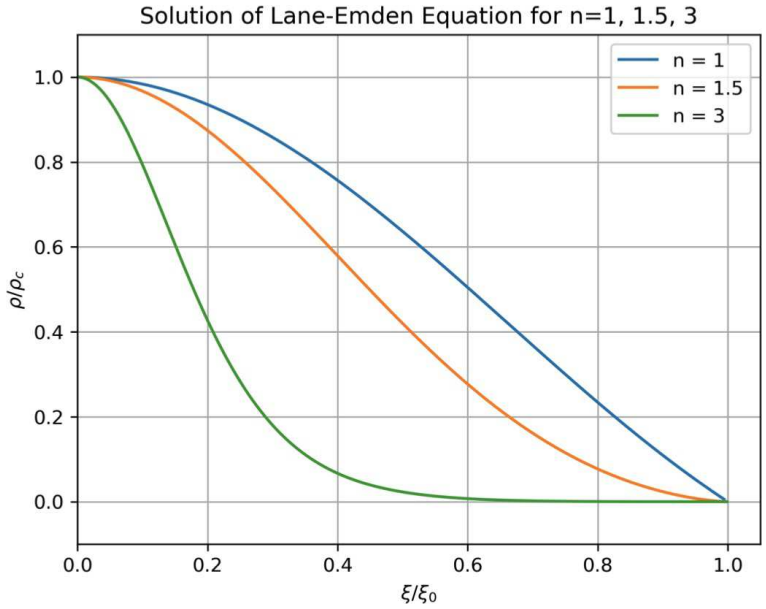



Figure 5.21: Solutions to the Lane-Emden equation for $n = 1$, $n = \frac{3}{2}$, and $n = 3$.

A convenient basis for the many body states is obtained from the single-particle eigenstates $\{|\alpha\rangle\}$ of some single-particle Hamiltonian $\hat{H}_0$, with $\langle\, \boldsymbol{x} \,|\, \alpha \,\rangle = \varphi_\alpha(\boldsymbol{x})$ and $\hat{H}_0\,|\alpha\rangle = \varepsilon_\alpha\,|\alpha\rangle$. The basis may be taken as orthonormal, *i.e.* $\langle\, \alpha \,|\, \alpha' \,\rangle = \delta_{\alpha\alpha'}$. Now define

$$\Psi_{\alpha_1,\ldots,\alpha_N}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = \frac{1}{\sqrt{N!\prod_\alpha n_\alpha!}} \sum_{\sigma\in S_N} \left\{ \begin{matrix} 1 \\ \mathsf{sgn}(\sigma) \end{matrix} \right\} \varphi_{\alpha_{\sigma(1)}}(\boldsymbol{x}_1)\cdots\varphi_{\alpha_{\sigma(N)}}(\boldsymbol{x}_N) \quad . \tag{5.329}$$

Here $n_\alpha$ is the number of times the index $\alpha$ appears among the set $\{\alpha_1,\ldots,\alpha_N\}$. For BE statistics, $n_\alpha \in \{0,1,2,\ldots\}$, whereas for FD statistics, $n_\alpha \in \{0,1\}$. Note that the above states are normalized[33]:

$$\begin{aligned}\int d^dx_1\cdots\int d^dx_N\,\left|\Psi_{\alpha_1\cdots\alpha_N}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\right|^2 &= \frac{1}{N!\prod_\alpha n_\alpha!}\sum_{\sigma,\mu\in S_N}\left\{\begin{matrix}1\\ \mathsf{sgn}(\sigma\mu)\end{matrix}\right\}\prod_{j=1}^N\int d^dx_j\;\varphi^*_{\alpha_{\sigma(j)}}(\boldsymbol{x}_j)\,\varphi_{\alpha_{\mu(j)}}(\boldsymbol{x}_j) \\ &= \frac{1}{\prod_\alpha n_\alpha!}\sum_{\sigma\in S_N}\prod_{j=1}^N\delta_{\alpha_j,\alpha_{\sigma(j)}} = 1 \quad .\end{aligned} \tag{5.330}$$

Note that

$$\begin{aligned}\sum_{\sigma\in S_N}\varphi_{\alpha_{\sigma(1)}}(\boldsymbol{x}_1)\cdots\varphi_{\alpha_{\sigma(N)}}(\boldsymbol{x}_N) &\equiv \mathsf{per}\big\{\varphi_{\alpha_i}(\boldsymbol{x}_j)\big\} \\ \sum_{\sigma\in S_N}\mathsf{sgn}(\sigma)\,\varphi_{\alpha_{\sigma(1)}}(\boldsymbol{x}_1)\cdots\varphi_{\alpha_{\sigma(N)}}(\boldsymbol{x}_N) &\equiv \mathsf{det}\big\{\varphi_{\alpha_i}(\boldsymbol{x}_j)\big\} \quad ,\end{aligned} \tag{5.331}$$

which stand for *permanent* and *determinant*, respectively. We may now write

$$\Psi_{\alpha_1\cdots\alpha_N}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = \big\langle\, \boldsymbol{x}_1,\cdots,\boldsymbol{x}_N \,\big|\, \alpha_1\cdots\alpha_N \,\big\rangle \quad , \tag{5.332}$$

where

$$\big|\, \alpha_1\cdots\alpha_N \,\big\rangle = \frac{1}{\sqrt{N!\prod_\alpha n_\alpha!}}\sum_{\sigma\in S_N}\left\{\begin{matrix}1\\ \mathsf{sgn}(\sigma)\end{matrix}\right\}\big|\,\alpha_{\sigma(1)}\,\big\rangle\otimes\big|\,\alpha_{\sigma(2)}\,\big\rangle\otimes\cdots\otimes\big|\,\alpha_{\sigma(N)}\,\big\rangle \quad . \tag{5.333}$$

Note that $|\,\alpha_{\sigma(1)}\cdots\alpha_{\sigma(N)}\,\rangle = (\pm1)^\sigma\,|\,\alpha_1\cdots\alpha_N\,\rangle$, where by $(\pm1)^\sigma$ we mean 1 in the case of BE statistics and $\mathsf{sgn}(\sigma)$ in the case of FD statistics.

We may express $|\,\alpha_1\cdots\alpha_N\,\rangle$ as a product of creation operators acting on a vacuum $|\,0\,\rangle$ in *Fock space*. For bosons,

$$\big|\,\alpha_1\cdots\alpha_N\,\big\rangle = \prod_\alpha\frac{(b^\dagger_\alpha)^{n_\alpha}}{\sqrt{n_\alpha!}}\,\big|\,0\,\big\rangle \equiv \big|\,\{n_\alpha\}\,\big\rangle \quad , \tag{5.334}$$

with

$$\big[b_\alpha\,,\,b_\beta\big] = 0 \qquad , \qquad \big[b^\dagger_\alpha\,,\,b^\dagger_\beta\big] = 0 \qquad , \qquad \big[b_\alpha\,,\,b^\dagger_\beta\big] = \delta_{\alpha\beta} \quad , \tag{5.335}$$

where $[\,\bullet\,,\bullet\,]$ is the commutator. For fermions,

$$\big|\,\alpha_1\cdots\alpha_N\,\big\rangle = c^\dagger_{\alpha_1}c^\dagger_{\alpha_2}\cdots c^\dagger_{\alpha_N}\,\big|\,0\,\big\rangle \equiv \big|\,\{n_\alpha\}\,\big\rangle \quad , \tag{5.336}$$

[33] In the normalization integrals, each $\int d^dx$ implicitly includes a sum $\sum_\zeta$ over any internal indices that may be present.

with

$$\{c_\alpha\,,\,c_\beta\}=0\qquad,\qquad\{c_\alpha^\dagger\,,\,c_\beta^\dagger\}=0\qquad,\qquad\{c_\alpha\,,\,c_\beta^\dagger\}=\delta_{\alpha\beta}\quad, \tag{5.337}$$

where $\{\bullet\,,\bullet\}$ is the anticommutator.

### 5.8.2 Second quantized operators

Now consider the action of permutation-symmetric first quantized operators such as the kinetic energy $\hat{T}=-\frac{\hbar^2}{2m}\sum_{i=1}^N\boldsymbol{\nabla}_i^2=\sum_{i=1}^N\hat{t}_i$ and the potential energy $\hat{U}=\sum_{i<j}^N u(\boldsymbol{x}_i-\boldsymbol{x}_j)=\sum_{i<j}^N\hat{u}_{ij}$. For a one-body operator such as $\hat{T}$, we have

$$\begin{aligned}\langle\,\alpha_1\cdots\alpha_N\,|\,\hat{T}\,|\,\alpha_1'\cdots\alpha_N'\,\rangle &= \int d^dx_1\cdots\int d^dx_N\,\Big(\prod_\alpha n_\alpha!\Big)^{-1/2}\Big(\prod_\alpha n_\alpha'!\Big)^{-1/2} \\ &\qquad\times\sum_{\pi\in\mathcal{S}_N}\mathsf{sgn}(\pi)\,\varphi^*_{\alpha_{\pi(1)}}(\boldsymbol{x}_1)\cdots\varphi^*_{\alpha_{\pi(N)}}(\boldsymbol{x}_N)\,\sum_{k=1}^N\hat{t}_k\,\varphi_{\alpha'_{\pi(1)}}(\boldsymbol{x}_1)\cdots\varphi_{\alpha'_{\pi(N)}}(\boldsymbol{x}_N) \\ &=\sum_{\pi\in\mathcal{S}_N}\mathsf{sgn}(\pi)\,\Big(\prod_\alpha n_\alpha!\,n_\alpha'!\Big)^{-1/2}\sum_{i=1}^N\prod_{\substack{j\\(j\neq i)}}\delta_{\alpha_j,\alpha'_{\pi(j)}}\int d^dx_1\,\varphi^*_{\alpha_i}(\boldsymbol{x}_1)\,\hat{t}_1\,\varphi_{\alpha'_{\pi(i)}}(\boldsymbol{x}_1)\quad.\end{aligned} \tag{5.338}$$

One may verify that any permutation-symmetric one-body operator such as $\hat{T}$ is faithfully represented by the second quantized expression,

$$\hat{T}=\sum_{\alpha,\beta}\langle\,\alpha\,|\,\hat{t}\,|\,\beta\,\rangle\,\psi_\alpha^\dagger\,\psi_\beta\quad, \tag{5.339}$$

where $\psi_\alpha^\dagger$ is $b_\alpha^\dagger$ or $c_\alpha^\dagger$ as the application determines, and

$$\langle\,\alpha\,|\,\hat{t}\,|\,\beta\,\rangle=\int d^dx\,\varphi^*_\alpha(\boldsymbol{x})\,\hat{t}(\boldsymbol{\nabla})\,\varphi_\beta(\boldsymbol{x})\equiv t_{\alpha\beta}\quad. \tag{5.340}$$

Similarly,

$$\hat{V}=\sum_{\alpha,\beta}\langle\,\alpha\,|\,\hat{v}_{\mathsf{ext}}\,|\,\beta\,\rangle\,\psi_\alpha^\dagger\,\psi_\beta\quad, \tag{5.341}$$

where

$$\langle\,\alpha\,|\,\hat{v}_{\mathsf{ext}}\,|\,\beta\,\rangle=\int d^dx\,\varphi^*_\alpha(\boldsymbol{x})\,\hat{v}_{\mathsf{ext}}(\boldsymbol{x})\,\varphi_\beta(\boldsymbol{x})\equiv v^{\mathsf{ext}}_{\alpha\beta}\quad. \tag{5.342}$$

Two-body operators such as $\hat{U}$ are represented as

$$\hat{U}=\tfrac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}\langle\,\alpha\beta\,|\,\hat{u}\,|\,\gamma\delta\,\rangle\,\psi_\alpha^\dagger\,\psi_\beta^\dagger\,\psi_\delta\,\psi_\gamma\quad, \tag{5.343}$$

where

$$\langle\,\alpha\beta\,|\,\hat{u}\,|\,\gamma\delta\,\rangle=\int d^dx\int d^dx'\,\varphi^*_\alpha(\boldsymbol{x})\,\varphi^*_\beta(\boldsymbol{x}')\,\hat{u}(\boldsymbol{x},\boldsymbol{x}')\,\varphi_\delta(\boldsymbol{x}')\,\varphi_\gamma(\boldsymbol{x})\equiv u_{\alpha\beta\gamma\delta}\quad. \tag{5.344}$$

The general form for an $n$-body operator is then

$$\hat{R} = \frac{1}{n!} \sum_{\substack{\alpha_1 \cdots \alpha_n \\ \beta_1 \cdots \beta_n}} \langle\, \alpha_1 \cdots \alpha_n \,|\, \hat{r} \,|\, \beta_1 \cdots \beta_n \,\rangle\, \psi^\dagger_{\alpha_n} \cdots \psi^\dagger_{\alpha_n}\, \psi_{\beta_n} \cdots \psi_{\beta_1} \tag{5.345}$$

where

$$\langle\, \alpha_1 \cdots \alpha_n \,|\, \hat{r} \,|\, \beta_1 \cdots \beta_n \,\rangle = \int d^d x_1 \cdots \int d^d x_n\, \varphi^*_{\alpha_1}(\boldsymbol{x}_1) \cdots \varphi^*_{\alpha_n}(\boldsymbol{x}_n)\, \hat{r}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\, \varphi_{\beta_n}(\boldsymbol{x}_n) \cdots \varphi_{\beta_1}(\boldsymbol{x}_1) \quad . \tag{5.346}$$

If the particles have no internal degrees of freedom, then the operators $\hat{r}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ are just functions of the spatial coordinates $\{\boldsymbol{x}_i\}$. If there are (discrete) internal degrees of freedom, then $\hat{r}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ also has operator content in the internal Hilbert space as well.

Finally, if the Hamiltonian $\hat{H}_0 = \sum_{i=1}^N \hat{h}(\boldsymbol{x}_i, \boldsymbol{p}_i, \boldsymbol{\mu}_i)$ is noninteracting (with $\boldsymbol{\mu}$ a set of operators acting on internal degrees of freedom, such as spin), consisting solely of one-body operators, then

$$\hat{H}_0 = \sum_\alpha \varepsilon_\alpha\, \psi^\dagger_\alpha\, \psi_\alpha \quad , \tag{5.347}$$

where $\{\varepsilon_\alpha\}$ is the spectrum of the single particle Hamiltonian $\hat{h}$.

## 5.9 Appendix II: Ideal Bose Gas Condensation

We begin with the grand canonical Hamiltonian $K = H - \mu N$ for the ideal Bose gas,

$$K = \sum_{\boldsymbol{k}} (\varepsilon_{\boldsymbol{k}} - \mu)\, b^\dagger_{\boldsymbol{k}} b_{\boldsymbol{k}} - \sqrt{N} \sum_{\boldsymbol{k}} \left( \nu_{\boldsymbol{k}}\, b^\dagger_{\boldsymbol{k}} + \bar{\nu}_{\boldsymbol{k}}\, b_{\boldsymbol{k}} \right) \quad . \tag{5.348}$$

Here $b^\dagger_{\boldsymbol{k}}$ is the creation operator for a boson in a state of wavevector $\boldsymbol{k}$, hence $\left[ b_{\boldsymbol{k}}\,,\, b^\dagger_{\boldsymbol{k}'} \right] = \delta_{\boldsymbol{k}\boldsymbol{k}'}$. The dispersion relation is given by the function $\varepsilon_{\boldsymbol{k}}$, which is the energy of a particle with wavevector $\boldsymbol{k}$. We must have $\varepsilon_{\boldsymbol{k}} - \mu \geq 0$ for all $\boldsymbol{k}$, lest the spectrum of $K$ be unbounded from below. The fields $\{\nu_{\boldsymbol{k}}, \bar{\nu}_{\boldsymbol{k}}\}$ break a global $\mathsf{O}(2)$ symmetry.

Students who have not taken a course in solid state physics can skip the following paragraph, and be aware that $N = V/v_0$ is the total volume of the system in units of a fundamental "unit cell" volume. The thermodynamic limit is then $N \to \infty$. Note that $N$ is not the boson particle number, which we'll call $N_b$.

*Solid state physics boilerplate* : We presume a setting in which the real space Hamiltonian is defined by some boson hopping model on a Bravais lattice. The wavevectors $\boldsymbol{k}$ are then restricted to the first Brillouin zone, $\hat{\Omega}$, and assuming periodic boundary conditions are quantized according to the condition $\exp\big(i N_l\, \boldsymbol{k} \cdot \boldsymbol{a}_l\big) = 1$ for all $l \in \{1, \ldots, d\}$, where $\boldsymbol{a}_l$ is the $l^{\rm th}$ fundamental direct lattice vector and $N_l$ is the size of the system in the $\boldsymbol{a}_l$ direction; $d$ is the dimension of space. The total number of unit cells is $N \equiv \prod_l N_l$. Thus, quantization entails $\boldsymbol{k} = \sum_l (2\pi n_l / N_l)\, \boldsymbol{b}_l$, where $\boldsymbol{b}_l$ is the $l^{\rm th}$ elementary reciprocal lattice vector ($\boldsymbol{a}_l \cdot \boldsymbol{b}_{l'} = 2\pi \delta_{ll'}$) and $n_l$ ranges over $N_l$ distinct integers such that the allowed $\boldsymbol{k}$ points form a discrete approximation to $\hat{\Omega}$.

To solve, we first shift the boson creation and annihilation operators, writing

$$K = \sum_{\boldsymbol{k}} (\varepsilon_{\boldsymbol{k}} - \mu)\, \beta^\dagger_{\boldsymbol{k}} \beta_{\boldsymbol{k}} - N \sum_{\boldsymbol{k}} \frac{|\nu_{\boldsymbol{k}}|^2}{\varepsilon_{\boldsymbol{k}} - \mu} \quad , \tag{5.349}$$

where

$$\beta_{\boldsymbol{k}} = b_{\boldsymbol{k}} - \frac{\sqrt{N}\,\nu_{\boldsymbol{k}}}{\varepsilon_{\boldsymbol{k}} - \mu} \qquad , \qquad \beta^\dagger_{\boldsymbol{k}} = b^\dagger_{\boldsymbol{k}} - \frac{\sqrt{N}\,\bar\nu_{\boldsymbol{k}}}{\varepsilon_{\boldsymbol{k}} - \mu} \quad . \tag{5.350}$$

Note that $\left[\beta_{\boldsymbol{k}}\,,\,\beta^\dagger_{\boldsymbol{k}'}\right] = \delta_{\boldsymbol{k}\boldsymbol{k}'}$ so the above transformation is canonical. The Landau free energy $\Omega = -k_{\rm B}T \log \Xi$, where $\Xi = \mathsf{Tr}\ e^{-K/k_{\rm B}T}$, is given by

$$\Omega = N k_{\rm B} T \int\limits_{-\infty}^{\infty} d\varepsilon\; g(\varepsilon)\; \log\left(1 - e^{(\mu-\varepsilon)/k_{\rm B}T}\right) - N \sum_{\boldsymbol{k}} \frac{|\nu_{\boldsymbol{k}}|^2}{\varepsilon_{\boldsymbol{k}} - \mu} \quad , \tag{5.351}$$

where $g(\varepsilon)$ is the density of energy states per unit cell,

$$g(\varepsilon) = \frac{1}{N} \sum_{\boldsymbol{k}} \delta\big(\varepsilon - \varepsilon_{\boldsymbol{k}}\big) \xrightarrow[N\to\infty]{} v_0 \int\limits_{\hat\Omega} \frac{d^d k}{(2\pi)^d}\, \delta\big(\varepsilon - \varepsilon_{\boldsymbol{k}}\big) \quad , \tag{5.352}$$

where $v_0$ is the volume of the unit cell in real space. Note that

$$\psi_{\boldsymbol{k}} \equiv \frac{1}{\sqrt{N}} \left\langle b_{\boldsymbol{k}} \right\rangle = -\frac{1}{N} \frac{\partial \Omega}{\partial \bar\nu_{\boldsymbol{k}}} = \frac{\nu_{\boldsymbol{k}}}{\varepsilon_{\boldsymbol{k}} - \mu} \quad . \tag{5.353}$$

In the condensed phase, $\psi_{\boldsymbol{k}}$ is nonzero.

The Landau free energy (grand potential) is a function $\Omega(T, N, \mu, \nu, \bar\nu)$. We now make a Legendre transformation,

$$Y(T, N, \mu, \psi, \bar\psi) = \Omega(T, N, \mu, \nu, \bar\nu) + N \sum_{\boldsymbol{k}} \left(\nu_{\boldsymbol{k}} \bar\psi_{\boldsymbol{k}} + \bar\nu_{\boldsymbol{k}} \psi_{\boldsymbol{k}}\right) \quad . \tag{5.354}$$

Note that

$$\frac{\partial Y}{\partial \bar\nu_{\boldsymbol{k}}} = \frac{\partial \Omega}{\partial \bar\nu_{\boldsymbol{k}}} + N \psi_{\boldsymbol{k}} = 0 \quad , \tag{5.355}$$

by the definition of $\psi_{\boldsymbol{k}}$. Similarly, $\partial Y / \partial \nu_{\boldsymbol{k}} = 0$. We now have

$$Y(T, N, \mu, \psi, \bar\psi) = N k_{\rm B} T \int\limits_{-\infty}^{\infty} d\varepsilon\; g(\varepsilon)\; \log\left(1 - e^{(\mu-\varepsilon)/k_{\rm B}T}\right) + N \sum_{\boldsymbol{k}} (\varepsilon_{\boldsymbol{k}} - \mu)\, |\psi_{\boldsymbol{k}}|^2 \quad . \tag{5.356}$$

Therefore, the boson particle number per unit cell is given by the *dimensionless density*,

$$n = \frac{N_b}{N} = -\frac{1}{N} \frac{\partial Y}{\partial \mu} = \sum_{\boldsymbol{k}} |\psi_{\boldsymbol{k}}|^2 + \int\limits_{-\infty}^{\infty} d\varepsilon\; \frac{g(\varepsilon)}{e^{(\varepsilon-\mu)/k_{\rm B}T} - 1} \quad , \tag{5.357}$$

and the condensate amplitude at wavevector $\boldsymbol{k}$ is

$$\nu_{\boldsymbol{k}} = \frac{1}{N}\frac{\partial Y}{\partial \bar{\psi}_{\boldsymbol{k}}} = (\varepsilon_{\boldsymbol{k}} - \mu)\,\psi_{\boldsymbol{k}} \quad . \tag{5.358}$$

Recall that $\nu_{\boldsymbol{k}}$ acts as an external field. Let the dispersion $\varepsilon_{\boldsymbol{k}}$ be minimized at $\boldsymbol{k} = \boldsymbol{K}$. Without loss of generality, we may assume this minimum value is $\varepsilon_{\boldsymbol{K}} = 0$. We see that if $\nu_{\boldsymbol{k}} = 0$ then one of two must be true:

(i) $\psi_{\boldsymbol{k}} = 0$ for all $\boldsymbol{k}$

(ii) $\mu = \varepsilon_{\boldsymbol{K}}$, in which case $\psi_{\boldsymbol{K}}$ can be nonzero.

Thus, for $\nu = \bar{\nu} = 0$ and $\mu < 0$, we have the usual equation of state,

$$n(T,\mu) = \int_0^\infty d\varepsilon\, \frac{g(\varepsilon)}{e^{(\varepsilon-\mu)/k_{\rm B}T} - 1} \quad , \tag{5.359}$$

which relates the intensive variables $n$, $T$, and $\mu$. When $\mu = 0$, the equation of state becomes

$$n(T,\mu=0) = \overbrace{\sum_{\boldsymbol{K}} |\psi_{\boldsymbol{K}}|^2}^{n_0} + \overbrace{\int_0^\infty d\varepsilon\, \frac{g(\varepsilon)}{e^{\varepsilon/k_{\rm B}T} - 1}}^{n'(T)} \quad , \tag{5.360}$$

where now the sum is over only those $\boldsymbol{K}$ for which $\varepsilon_{\boldsymbol{K}} = 0$. Typically this set has only one member, $\boldsymbol{K} = 0$, but it is quite possible, due to symmetry reasons, that there are more such $\boldsymbol{K}$ values. This last equation of state is one which relates the intensive variables $n$, $T$, and $n_0$, where $n_0 = \sum_{\boldsymbol{K}} |\psi_{\boldsymbol{K}}|^2$ is the dimensionless condensate density. If the integral $n'(T)$ in eqn. 5.360 is finite, then for $n > n'(T)$ we must have $n_0 > 0$. Note that, for any $T$, $n'(T)$ diverges logarithmically whenever $g(0)$ is finite. This means that eqn. 5.359 can always be inverted to yield a finite $\mu(n,T)$, no matter how large the value of $n$, in which case there is no condensation and $n_0 = 0$. If $g(\varepsilon) \propto \varepsilon^\alpha$ with $\alpha > 0$, the integral converges and $n'(T)$ is finite and monotonically increasing for all $T$. Thus, for fixed dimensionless number $n$, there will be a *critical temperature* $T_{\rm c}$ for which $n = n'(T_{\rm c})$. For $T < T_{\rm c}(n)$, eqn. 5.359 has no solution for any $\mu$ and we must appeal to eqn. 5.360. The condensate density, given by $n_0(n,T) = n - n'(T)$, is then finite for $T < T_{\rm c}(n)$, and vanishes for $T \geq T_{\rm c}(n)$.

In the condensed phase, the phase of the order parameter $\psi$ inherits its phase from the external field $\nu$, which is taken to zero, in the same way the magnetization in the symmetry-broken phase of an Ising ferromagnet inherits its direction from an applied field $h$ which is taken to zero. The important feature is that in both cases the applied field is taken to zero *after* the approach to the thermodynamic limit.

## 5.10 Appendix III: Example Bose Condensation Problem

PROBLEM: A three-dimensional gas of noninteracting bosonic particles obeys the dispersion relation $\varepsilon(\boldsymbol{k}) = A\,|\boldsymbol{k}|^{1/2}$.

(a) Obtain an expression for the density $n(T,z)$ where $z = \exp(\mu/k_{\rm B}T)$ is the fugacity. Simplify your expression as best you can, adimensionalizing any integral or infinite sum which may appear. You may find it convenient to define

$$\mathsf{Li}_\nu(z) \equiv \frac{1}{\Gamma(\nu)} \int\limits_0^\infty dt\, \frac{t^{\nu-1}}{z^{-1}\,e^t - 1} = \sum_{k=1}^\infty \frac{z^k}{k^\nu} \quad . \tag{5.361}$$

Note $\mathsf{Li}_\nu(1) = \zeta(\nu)$, the Riemann zeta function.

(b) Find the critical temperature for Bose condensation, $T_{\rm c}(n)$. Your expression should only include the density $n$, the constant $A$, physical constants, and numerical factors (which may be expressed in terms of integrals or infinite sums).

(c) What is the condensate density $n_0$ when $T = \frac{1}{2}\,T_{\rm c}$?

(d) Do you expect the second virial coefficient to be positive or negative? Explain your reasoning. (You don't have to do any calculation.)

SOLUTION: We work in the grand canonical ensemble, using Bose-Einstein statistics.

(a) The density for Bose-Einstein particles is given by

$$\begin{aligned} n(T,z) &= \int \frac{d^3k}{(2\pi)^3}\, \frac{1}{z^{-1}\,\exp(Ak^{1/2}/k_{\rm B}T) - 1} \\ &= \frac{1}{\pi^2}\left(\frac{k_{\rm B}T}{A}\right)^6 \int\limits_0^\infty ds\, \frac{s^5}{z^{-1}\,e^s - 1} = \frac{120}{\pi^2}\left(\frac{k_{\rm B}T}{A}\right)^6 \mathsf{Li}_6(z) \quad , \end{aligned} \tag{5.362}$$

where we have changed integration variables from $k$ to $s = Ak^{1/2}/k_{\rm B}T$, and we have defined the functions $\mathsf{Li}_\nu(z)$ as above, in eqn. 5.361. Note $\mathsf{Li}_\nu(1) = \zeta(\nu)$, the Riemann zeta function.

(b) Bose condensation sets in for $z = 1$, *i.e.* $\mu = 0$. Thus, the critical temperature $T_{\rm c}$ and the density $n$ are related by

$$n = \frac{120\,\zeta(6)}{\pi^2}\left(\frac{k_{\rm B}T_{\rm c}}{A}\right)^6, \tag{5.363}$$

or

$$T_{\rm c}(n) = \frac{A}{k_{\rm B}}\left(\frac{\pi^2\,n}{120\,\zeta(6)}\right)^{1/6} \quad . \tag{5.364}$$

(c) For $T < T_{\rm c}$, we have

$$n = n_0 + \frac{120\,\zeta(6)}{\pi^2}\left(\frac{k_{\rm B}T}{A}\right)^6 = n_0 + \left(\frac{T}{T_{\rm c}}\right)^6 n \quad , \tag{5.365}$$

where $n_0$ is the condensate density. Thus, at $T = \frac{1}{2}\,T_{\rm c}$,

$$n_0\big(T = \tfrac{1}{2}T_{\rm c}\big) = \tfrac{63}{64}\,n. \tag{5.366}$$

(d) The virial expansion of the equation of state is

$$p = nk_{\rm B}T\Big(1 + B_2(T)\,n + B_3(T)\,n^2 + \dots\Big) \quad . \tag{5.367}$$

We expect $B_2(T) < 0$ for noninteracting bosons, reflecting the tendency of the bosons to condense. (Correspondingly, for noninteracting fermions we expect $B_2(T) > 0$.)

For the curious, we compute $B_2(T)$ by eliminating the fugacity $z$ from the equations for $n(T,z)$ and $p(T,z)$. First, we find $p(T,z)$ :

$$\begin{aligned} p(T,z) &= -k_{\rm B}T\int\frac{d^3k}{(2\pi)^3}\,\log\Big(1 - z\,\exp(-Ak^{1/2}/k_{\rm B}T)\Big) \\ &= -\frac{k_{\rm B}T}{\pi^2}\left(\frac{k_{\rm B}T}{A}\right)^6\int_0^\infty ds\,s^5\,\log\left(1 - z\,e^{-s}\right) = \frac{120\,k_{\rm B}T}{\pi^2}\left(\frac{k_{\rm B}T}{A}\right)^6{\rm Li}_7(z). \end{aligned} \tag{5.368}$$

Expanding in powers of the fugacity, we have

$$\begin{aligned} n &= \frac{120}{\pi^2}\left(\frac{k_{\rm B}T}{A}\right)^6\left\{z + \frac{z^2}{2^6} + \frac{z^3}{3^6} + \dots\right\} \\ \frac{p}{k_{\rm B}T} &= \frac{120}{\pi^2}\left(\frac{k_{\rm B}T}{A}\right)^6\left\{z + \frac{z^2}{2^7} + \frac{z^3}{3^7} + \dots\right\} \quad . \end{aligned} \tag{5.369}$$

Solving for $z(n)$ using the first equation, we obtain, to order $n^2$,

$$z = \left(\frac{\pi^2A^6\,n}{120\,(k_{\rm B}T)^6}\right) - \frac{1}{2^6}\left(\frac{\pi^2A^6\,n}{120\,(k_{\rm B}T)^6}\right)^2 + \mathcal{O}(n^3) \quad . \tag{5.370}$$

Plugging this into the equation for $p(T,z)$, we obtain the first nontrivial term in the virial expansion, with

$$B_2(T) = -\frac{\pi^2}{15360}\left(\frac{A}{k_{\rm B}T}\right)^6 \quad , \tag{5.371}$$

which is negative, as expected. Note that the ideal gas law is recovered for $T \to \infty$, for fixed $n$.

## 5.11 Appendix IV: $C_{p,N}$ for the Ideal Bose Gas

The phase diagram for the ideal Bose gas in the $(T,p)$ plane was considered in §5.6.4 and in fig. 5.9. Let's compute the behavior of $C_{p,N}(T,p,N)$ and explore how it behaves as one approaches the critical curve $p = p_{\rm c}(T) = \zeta(5/2)\,k_{\rm B}T/\lambda_T^3$ . We found that when the fugacity $z = \exp(\mu/k_{\rm B}T) \in (0,1)$, then the density and pressure are given by

$$n(z,T) = {\rm Li}_{3/2}(z)\,\lambda_T^{-3} \qquad , \qquad p(z,T) = {\rm Li}_{5/2}(z)\,k_{\rm B}T\,\lambda_T^{-3} \quad . \tag{5.372}$$

The energy is $E = \frac{3}{2}pV$, as we obtained in eqn. 5.162.

To obtain $C_{p,N}$, we first set $dp = 0$, which is of course equivalent to setting $d\log p = 0$:

$$d\log p = 0 = \frac{\mathsf{Li}_{3/2}(z)}{\mathsf{Li}_{5/2}(z)}\,d\log z + \tfrac{5}{2}\,d\log T \quad \Rightarrow \quad \left(\frac{\partial \log z}{\partial \log T}\right)_p = -\frac{5\,\mathsf{Li}_{5/2}(z)}{3\,\mathsf{Li}_{3/2}(z)} \quad . \tag{5.373}$$

We wish to evaluate

$$C_{p,N} = \left(\frac{\partial E}{\partial T}\right)_{p,N} + p\left(\frac{\partial V}{\partial T}\right)_{p,N} = \frac{5}{2}\,p\left(\frac{\partial V}{\partial T}\right)_{p,N} \quad . \tag{5.374}$$

Thus, we need

$$\begin{aligned}\left(\frac{\partial V}{\partial T}\right)_{p,N} &= \left(\frac{\partial N/n}{\partial T}\right)_{p,N} = -\frac{N}{n}\left(\frac{\partial \log n}{\partial T}\right)_{p,N} \\ &= -\frac{N}{nT}\left\{\frac{\mathsf{Li}_{1/2}(z)}{\mathsf{Li}_{3/2}(z)}\left(\frac{\partial \log z}{\partial \log T}\right)_p + \frac{3}{2}\right\} = \frac{N}{nT}\left\{\frac{5\,\mathsf{Li}_{5/2}(z)\,\mathsf{Li}_{1/2}(z)}{2\,\mathsf{Li}^2_{3/2}(z)} - \frac{3}{2}\right\} \quad .\end{aligned} \tag{5.375}$$

We therefore have

$$c_{p,N} \equiv \frac{N_{\rm A}\,C_{p,N}}{N} = \frac{25R}{4}\,\frac{\mathsf{Li}_{5/2}(z)}{\mathsf{Li}_{3/2}(z)}\left\{\frac{\mathsf{Li}_{5/2}(z)\,\mathsf{Li}_{1/2}(z)}{\mathsf{Li}^2_{3/2}(z)} - \frac{3}{5}\right\} \quad , \tag{5.376}$$

where $R = N_{\rm A}k_{\rm B}$ is the gas constant.

As we approach the critical line $p = p_{\rm c}(T)$, the fugacity approaches unity: $z \to 1$. In this limit we have $\mathsf{Li}_{3/2}(z \to 1) = \zeta(3/2)$ and $\mathsf{Li}_{5/2}(z \to 1) = \zeta(5/2)$, but $\mathsf{Li}_{1/2}(z \to 1)$ is divergent. We write $z \equiv \exp(-\epsilon)$, with

$$\frac{p}{p_{\rm c}(T)} = \frac{\mathsf{Li}_{5/2}(z)}{\zeta(5/2)} = 1 - \frac{\zeta(3/2)}{\zeta(5/2)}\,\epsilon + \mathcal{O}(\epsilon^{3/2}) \tag{5.377}$$

allows us to write

$$\epsilon = \frac{\zeta(5/2)}{\zeta(3/2)}\left(1 - \frac{p}{p_{\rm c}(T)}\right) \quad . \tag{5.378}$$

Here we have appealed to the expansion in eqn. 5.56, which also gives

$$\mathsf{Li}_{1/2}(e^{-\epsilon}) = \sqrt{\pi}\left[\frac{\zeta(5/2)}{\zeta(3/2)}\cdot\frac{p_{\rm c}(T) - p}{p_{\rm c}(T)}\right]^{-1/2} + \ldots \quad . \tag{5.379}$$

Thus, from eqn. 5.375, we have that $c_{p,N}$ diverges as we approach $p = p_{\rm c}(T)$ from below as

$$c_{p,N}(T,p) = \frac{25\sqrt{\pi}}{4}\,\frac{\zeta^{3/2}(5/2)}{\zeta^{5/2}(3/2)}\left(\frac{p_{\rm c}(T) - p}{p_{\rm c}(T)}\right)^{-1/2} + \ldots \quad . \tag{5.380}$$

Equivalently, we can consider approaching the curve $T = T_{\rm c}(p)$ from the right. In both cases we have

$$c_{p,N}(T,p) \propto \big|p_{\rm c}(T) - p\big|^{-1/2} \quad , \quad c_{p,N}(T,p) \propto \big|T - T_{\rm c}(p)\big|^{-1/2} \quad . \tag{5.381}$$

In other words, the critical exponent is $\alpha = \frac{1}{2}$, and unlike $c_{V,N}(T,n)$ which has a cusp at $T = T_{\rm c}(n)$ yet remains finite, the specific heat at constant pressure diverges[34].

[34] I thank Andre Vieira for prompting me to clarify the differences between $c_{V,N}(T,n)$ and $c_{p,N}(T,p)$ in this context.

# Chapter 6

# Classical Interacting Systems

## 6.1 References

- M. Kardar, *Statistical Physics of Particles* (Cambridge, 2007)
  A superb modern text, with many insightful presentations of key concepts.
- L. E. Reichl, *A Modern Course in Statistical Physics* (2nd edition, Wiley, 1998)
  A comprehensive graduate level text with an emphasis on nonequilibrium phenomena.
- M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* ($3^{\rm rd}$ edition, World Scientific, 2006)
  An excellent graduate level text. Less insightful than Kardar but still a good modern treatment of the subject. Good discussion of mean field theory.
- E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics* (part I, $3^{\rm rd}$ edition, Pergamon, 1980)
  This is volume 5 in the famous Landau and Lifshitz *Course of Theoretical Physics*. Though dated, it still contains a wealth of information and physical insight.
- J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, 1990)
  An advanced, detailed discussion of liquid state physics.

## 6.2 The Ising Model

The simplest model of an interacting system consists of a lattice $\mathcal{L}$ of $N$ sites, each of which hosts a spin $\sigma_i$ which may be either up ($\sigma_i = +1$) or down ($\sigma_i = -1$). The Hamiltonian is

$$\hat{H} = -J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j - \mu_0 H \sum_i \sigma_i \quad . \tag{6.1}$$

The energy $J$ accounts for an interaction between sites $i$ and $j$, here taken to be nearest neighbors on $\mathcal{L}$. When $J = 0$, the model describes $N$ noninteracting spins. The geometry is then irrelevant, and as we saw in §4.8.1, for each spin we have $\langle \sigma_i \rangle =$
$tanh(\mu_0 H / k_{\rm B} T)$.

In zero external field (*i.e.* $H = 0$), energetic considerations dictate that the interaction term on each link $\langle ij \rangle$ prefers neighboring spins to be *ferromagnetically* aligned in a $|\uparrow\uparrow\rangle$ or $|\downarrow\downarrow\rangle$ configuration (*i.e.* $\sigma_i \, \sigma_j = +1$) when $J > 0$, and *antiferromagnetically* aligned in a $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$ configuration (*i.e.* $\sigma_i \, \sigma_j = -1$) when $J < 0$.

This model is not exactly solvable in general. In one dimension, the solution is quite straightforward. In two dimensions, Onsager's solution of the model (with $H = 0$) is among the most celebrated results in statistical physics. In higher dimensions the system has been studied by numerical simulations (the Monte Carlo method), by series expansions, and by field theoretic calculations (renormalization group), but no exact solutions exist.

One important aspect of interacting systems is the emergence of *correlations*. We define the *correlation matrix* of the Ising model $C_{ij}(T, H) \equiv \langle \sigma_i \, \sigma_j \rangle$, which is the average of the product $\sigma_i \, \sigma_j$[1]. When $J = 0$ the system is noninteracting, and the average of the product is the product of the averages for all $i \neq j$, *i.e.* $C_{ij}(T, H) = m_i \, m_j$, where $m_i \equiv \langle \sigma_i \rangle$. Thus the *connected correlator*,

$$\widetilde{C}_{ij}(T, H) \equiv \langle \sigma_i \, \sigma_j \rangle - \langle \sigma_i \rangle \, \langle \sigma_j \rangle \quad , \tag{6.2}$$

which is the average of the product *minus* the product of the averages, vanishes identically for $J = 0$ whenever $i \neq j$.

The general expression for the *thermal density matrix* at temperature $T$ is

$$\varrho(T, N, H) = \frac{\sum_{\boldsymbol{\sigma}} e^{-E(\boldsymbol{\sigma})/k_{\rm B}T} \, |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}|}{\sum_{\boldsymbol{\sigma}'} e^{-E(\boldsymbol{\sigma}')/k_{\rm B}T}} = \frac{1}{Z(T, N, H)} \sum_{\boldsymbol{\sigma}} e^{-E(\boldsymbol{\sigma})/k_{\rm B}T} \, |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}| \quad . \tag{6.3}$$

where

$$Z(T, N, H) = \mathsf{Tr} \; e^{-\hat{H}/k_{\rm B}T} = \sum_{\boldsymbol{\sigma}} e^{-E(\boldsymbol{\sigma})/k_{\rm B}T} \tag{6.4}$$

is the partition function. Here $N$ is the number of sites, which is the volume $|\mathcal{L}|$ of the lattice (*i.e.* the total number of sites). Thermodynamic averages of any quantity $\mathcal{O}(\boldsymbol{\sigma})$ are computed via

$$\langle \mathcal{O} \rangle = \mathsf{Tr} \, (\mathcal{O} \varrho) = \frac{1}{Z(T, N, H)} \sum_{\boldsymbol{\sigma}} \mathcal{O}(\boldsymbol{\sigma}) \, e^{-E(\boldsymbol{\sigma})/k_{\rm B}T} \quad . \tag{6.5}$$

[1]The correlation matrix $C_{ij}(T, H)$ is defined for arbitrary sites $i$ and $j$. In the thermodynamic limit, it is independent of the system size $N$.

When $J \neq 0$, there are nontrivial correlations between different sites. Consider, for example, the case $H = 0$. In this case the model has what we call a $\mathbb{Z}_2$ symmetry, which means that for any spin configuration $\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_N\}$, the energy $E(\boldsymbol{\sigma})$ is invariant under reversing the direction of all the spins[2] . *I.e.* if $\varepsilon\boldsymbol{\sigma} \equiv \{-\sigma_1, \ldots, -\sigma_N\}$, we have $E(\varepsilon\boldsymbol{\sigma}) = E(\boldsymbol{\sigma})$ for each of the $2^N$ possible spin configurations $\boldsymbol{\sigma}$. This means that $\langle\sigma_i\rangle = 0$ for all $i$ . At $T = 0$ there are two ground states, $|\Uparrow\rangle \equiv |\uparrow\uparrow\uparrow\cdots\rangle$ and $|\Downarrow\rangle \equiv |\downarrow\downarrow\downarrow\cdots\rangle$ , each with energy $E_0 = -N_{\rm nn}J$, where $N_{\rm nn}$ is the number of nearest neighbor links[3] on $\mathcal{L}$. The thermal density matrix as $T \to 0$ is given by $\varrho(T = 0, H = 0) = \frac{1}{2}|\Uparrow\rangle\langle\Uparrow| + \frac{1}{2}|\Downarrow\rangle\langle\Downarrow|$ , from which we obtain $\langle\sigma_i\rangle = 0$ and $\widetilde{C}_{ij}(T = 0, H = 0) = 1$. If we apply an external field $H > 0$, then the two ground states are no longer degenerate in energy, and $E_\Uparrow = -N_{\rm nn}J - N\mu_0 H$ while $E_\Downarrow = -N_{\rm nn}J + N\mu_0 H$. Thus, as $T \to 0$, even for very small $H > 0$, in the thermodynamic limit the $|\Uparrow\rangle$ state is selected. and we obtain a very different density matrix: $\varrho(T = 0, H > 0) = |\Uparrow\rangle\langle\Uparrow|$. Using this density matrix, we obtain $\langle\sigma_i\rangle = +1$ and $\langle\sigma_i\sigma_j\rangle = +1$ $(i \neq j)$, hence $\widetilde{C}_{ij}(T = 0, H > 0) = 0$.

### 6.2.1 Spontaneous symmetry breaking

In quantum mechanics, the eigenstates of a Hamiltonian $\hat{H}_0$ which commutes with all the *generators* of a symmetry group G may be classified according to the *representations* of that group. Typically this entails the appearance of degeneracies in the eigenspectrum, with degenerate states transforming into each other under the group operations, as we saw above. Adding a perturbation $V$ to the Hamiltonian which breaks G down to a subgroup H will accordingly split these degeneracies, and the new multiplets of $\hat{H} = \hat{H}_0 + \hat{V}$ are characterized by representations of the lower symmetry group H.

In quantum field theory, or in the thermodynamic limit of a classical system, as a consequence of the infinite number of degrees of freedom, symmetries may be *spontaneously broken*. This means that even if the Hamiltonian $H$ (or action $S$) for the field theory is invariant under a group G of symmetry transformations, the ground state or thermodynamic density matrix may not be invariant under the full symmetry group G. The presence or absence of spontaneous symmetry breaking (SSB), and its detailed manifestations, will in general depend on the couplings, or the temperature in the case of quantum statistical mechanics. SSB is usually associated with the presence of a local *order parameter* which transforms nontrivially under some group operations, and whose quantum statistical average vanishes in a fully symmetric phase, but takes nonzero values in symmetry-broken phase[4]. The parade example is the Ising model, $\hat{H} = -\sum_{i<j} J_{ij}\,\sigma_i\,\sigma_j$ , where each $\sigma_i = \pm 1$ , the subscript $i$ indexes a physical location in space, such as a site $\boldsymbol{R}_i$ on a particular lattice. The model is explicitly $\mathbb{Z}_2$ symmetric under $\sigma_i \to \varepsilon\sigma_i$ for all $i$, where $\varepsilon \in \{+1, -1\}$, yet if the interaction matrix $J_{ij} = J(\boldsymbol{R}_i - \boldsymbol{R}_j)$ is short-ranged and the space dimension $d$ is greater than one, there is a *critical temperature* $T_{\rm c}$ below which SSB sets in, and the system develops a spontaneous magnetization $m = \langle\sigma_i\rangle$. You know how in quantum mechanics, the eigenstates of a particle moving in one-dimensional double-well potential $V(x) = V(-x)$ can be classified by their parity eigenvalues P, and the lowest two energy states are respectively symmetric (P $= +1$) and

[2]The group $\mathbb{Z}_2$ contains two elements, $\{1, \varepsilon\}$, and the multiplication table is $1 \times 1 = 1$, $1 \times \varepsilon = \varepsilon \times 1 = \varepsilon$, and $\varepsilon \times \varepsilon = 1$.

[3]On a lattice with coordination number $z$ and with periodic boundary conditions, $N_{\rm nn} = \frac{1}{2}zN$.

[4]While SSB is generally associated with the existence of a phase transition, not all phase transitions involve SSB. Exceptions include the Kosterlitz-Thouless transition, and also those topological phases which have no local order parameter.

antisymmetric ($\mathsf{P} = -1$), and are delocalized among both wells. For a quantum field theory, however, with (Euclidean) Lagrangian density $\mathcal{L}_{\rm E} = \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi)$, for $d > 1$ and $T < T_{\rm c}$, the system actually picks the left or the right well, so that $\langle\phi(\boldsymbol{r})\rangle \neq 0$. Another example is the spontaneously broken $\mathsf{O}(2)$ invariance of superfluids, where the boson annihilation operator $\psi(\boldsymbol{r})$ develops a spontaneous average $\big\langle\psi(\boldsymbol{r})\big\rangle = \sqrt{n_0}\, e^{i\theta}$, where $n_0$ is the condensate density and $\theta$ the condensate phase.

But this is an obvious swindle, because, as we have seen, $\langle\sigma_i\rangle = 0$ in zero external field, due to the $\mathbb{Z}_2$ symmetry of the model. Rather, we may understand the phenomenon of spontaneous symmetry breaking in either of the following ways:

- First, rather than defining the order parameter of the Ising model, for example, to be the expected value $m = \langle\sigma_i\rangle$ of the local spin[5], consider instead the behavior of the correlator $C_{ij} = \langle\sigma_i\,\sigma_j\rangle$ in the limit $d_{ij} = |\boldsymbol{R}_i - \boldsymbol{R}_j| \to \infty$. In a disordered phase, there is no correlation between infinitely far separated spins, hence $\lim_{d_{ij}\to\infty} C_{ij} = 0$. In the ordered phase, this is no longer true, and we define the *spontaneous magnetization* $m$ from the long distance correlator: $m^2 \equiv \lim_{d_{ij}\to\infty}\langle\sigma_i\,\sigma_j\rangle$. In this formulation, SSB is associated with the emergence of *long-ranged order* in the correlators of operators which transform nontrivially under the symmetry group.
- Second, we could impose an external field which *explicitly* breaks the symmetry, such as a Zeeman term $\hat{V} = -\mu_0 H\sum_i \sigma_i$ in the Ising model. We compute the magnetization per site $m(T,N,H) = \langle\sigma_i\rangle$ as a function of temperature $T$, the external field $H$, *and* the system volume (*i.e.* number of sites) $N$. The order parameter $m(T)$ in zero field is then defined as
$$m(T) = \lim_{H\to 0}\lim_{N\to\infty} m(T,N,H) \quad . \tag{6.6}$$
The order of limits here is crucially important. The thermodynamic limit $N \to \infty$ is taken first, which means that the energy difference between $|\Uparrow\rangle$ and $|\Downarrow\rangle$ diverges, being proportional to $N$, thus infinitely suppressing the $|\Downarrow\rangle$ state if $H > 0$ (and the $|\Uparrow\rangle$ state if $H < 0$). The magnitude of the order parameter will be independent of the way in which we take $H \to 0$, but its sign will depend on whether $H \to 0^+$ or $H \to 0^-$, with $\mathsf{sgn}(m) = \mathsf{sgn}(H)$. Physically, the direction in which a system orders can be decided by the presence of small stray fields or impurities.

Note that in both formulations, SSB is necessarily associated with the existence of a local operator $\mathcal{O}_i$ which is identified as the order parameter field. In the first scheme, the correlations $\big\langle\mathcal{O}_i\,\mathcal{O}_j\big\rangle$ exhibit long-ranged order in the symmetry-broken phase. In the second scheme, $\mathcal{O}_i$ is the operator to which the local external field $H_i$ couples. We will discuss the mean field theory of phase transitions and spontaneous symmetry breaking in chapter 7.

### 6.2.2 Solution in $d = 1$ dimension

Consider a one-dimensional ring of $L$ sites. The ordinary canonical partition function is then
$$Z_{\mathsf{PBC}}(T,L,H) = \mathsf{Tr}\; e^{-\beta\hat{H}} = \sum_{\{\sigma_n\}}\prod_{n=1}^{L} e^{\beta J\sigma_n\sigma_{n+1}}\, e^{\beta\mu_0 H\sigma_n} \quad , \tag{6.7}$$

[5]We assume translational invariance, which means $\langle\sigma_i\rangle$ is independent of the site index $i$.

where $\sigma_{L+1} \equiv \sigma_1$ owing to periodic (ring) boundary conditions (PBC). We can replace the factor $e^{\beta\mu_0 H\sigma_n}$ in the above expression with $e^{\beta\mu_0 H(\sigma_n+\sigma_{n+1})/2}$, since the product over $n$ yields the same result. We then obtain $Z = \mathsf{Tr}\,(R^L)$, where $R$ is a $2\times 2$ matrix with entries

$$R_{\sigma\sigma'} = e^{\beta J\sigma\sigma'}\, e^{\beta\mu_0 H(\sigma+\sigma')/2} = \begin{pmatrix} e^{\beta J}\, e^{\beta\mu_0 H} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J}\, e^{-\beta\mu_0 H} \end{pmatrix}_{\sigma\sigma'} \quad , \tag{6.8}$$

called the *transfer matrix*. Expressed in terms of the Pauli matrices, we have[6]

$$R = e^{\beta J}\cosh(\beta\mu_0 H)\, I + e^{-\beta J} X + e^{\beta J}\sinh(\beta\mu_0 H)\, Z \quad , \tag{6.9}$$

where $I$, $X$, and $Z$ are the $2\times 2$ identity, Pauli $X$, and Pauli $Z$ matrices.

Since the trace of a matrix is invariant under a similarity transformation, we have

$$Z(T, L, H) = \lambda_+^L + \lambda_-^L \quad , \tag{6.10}$$

where $\lambda_\pm$ are the eigenvalues of $R$, *viz.*

$$\lambda_\pm(T, H) = e^{\beta J}\cosh(\beta\mu_0 H) \pm \sqrt{e^{2\beta J}\sinh^2(\beta\mu_0 H) + e^{-2\beta J}} \quad . \tag{6.11}$$

When $H = 0$, we have $\lambda_+ = 2\cosh(\beta J)$ and $\lambda_- = 2\sinh(\beta J)$. In the thermodynamic limit, $L\to\infty$, and the larger $\lambda_+^L$ term dominates exponentially. We then have

$$F(T, L, H) = -Lk_{\rm B}T\log\lambda_+(T, H) \quad . \tag{6.12}$$

From the free energy, we can compute the magnetization,

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T,L} = \frac{L\mu_0\,\sinh(\beta\mu_0 H)}{\sqrt{\sinh^2(\beta\mu_0 H) + e^{-4\beta J}}} \tag{6.13}$$

and the zero field isothermal susceptibility,

$$\chi(T) = \frac{1}{L}\,\frac{\partial M}{\partial H}\bigg|_{H=0} = \frac{\mu_0^2}{k_{\rm B}T}\, e^{2J/k_{\rm B}T} \quad . \tag{6.14}$$

Note that in the noninteracting limit $J\to 0$ we recover the familiar result for a free spin. The effect of the interactions at low temperature is to vastly increase the susceptibility. Rather than a set of independent single spins, the system effectively behaves as if it were composed of large blocks of spins, where the block size $\xi$ is the *correlation length*, to be derived below.

The physical properties of the system are often elucidated by evaluation of various correlation functions. Accordingly, we define $C(n) \equiv \langle\sigma_1\,\sigma_{n+1}\rangle$, where

$$\big\langle\sigma_1\,\sigma_{n+1}\big\rangle = \frac{\mathsf{Tr}\,\big(\sigma_1\, R_{\sigma_1\sigma_2}\cdots R_{\sigma_n\sigma_{n+1}}\,\sigma_{n+1}\, R_{\sigma_{n+1}\sigma_{n+2}}\cdots R_{\sigma_L\sigma_1}\big)}{\mathsf{Tr}\,\left(R^L\right)} = \frac{\mathsf{Tr}\,\left(Z\, R^n\, Z\, R^{L-n}\right)}{\mathsf{Tr}\,\left(R^L\right)} \quad , \tag{6.15}$$

[6]Take care not to confuse Pauli $Z$ with the partition function!

with $0 < n < L$, and where $Z$ is the Pauli matrix. To compute this ratio, we decompose $R$ in terms of its eigenvectors, writing $R = \lambda_+ \, |+\rangle\langle+| + \lambda_- \, |-\rangle\langle-|$. Then

$$C(n) = \frac{\lambda_+^L \, Z_{++}Z_{++} + \lambda_-^L \, Z_{--}Z_{--} + \left(\lambda_+^{L-n} \, \lambda_-^n + \lambda_+^n \, \lambda_-^{L-n}\right) Z_{+-}Z_{-+}}{\lambda_+^L + \lambda_-^L} \quad , \tag{6.16}$$

with $Z_{\mu\mu'} = \langle \, \mu \, | \, Z \, | \, \mu' \, \rangle$ being the matrix elements of $Z$ in the eigenbasis of $R$.

### Zero external field

Consider $H = 0$, where $R = e^{\beta J} + e^{-\beta J} \, X$. Then $|\pm\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\rangle \pm |\downarrow\rangle\big)$, *i.e.* the eigenvectors of $R$ are

$$\psi_\pm = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \quad , \tag{6.17}$$

which entails $Z_{++} = Z_{--} = 0$, while $Z_{+-} = Z_{-+} = 1$. The corresponding eigenvalues are given by $\lambda_+ = 2\cosh(\beta J)$ and $\lambda_- = 2\sinh(\beta J)$. The correlation function is then found to be

$$\begin{aligned} C(n) \equiv \big\langle \sigma_1 \, \sigma_{n+1} \big\rangle &= \frac{\lambda_+^{L-|n|} \, \lambda_-^{|n|} + \lambda_+^{|n|} \, \lambda_-^{L-|n|}}{\lambda_+^L + \lambda_-^L} \\ &= \frac{\tanh^{|n|}(\beta J) + \tanh^{L-|n|}(\beta J)}{1 + \tanh^L(\beta J)} \approx \tanh^{|n|}(\beta J) \quad \text{for } L \to \infty \quad . \end{aligned} \tag{6.18}$$

This result is also valid for $n < 0$, provided $|n| \le L$. We see that we may write $C(n) = e^{-|n|/\xi(T)}$, where the *correlation length* is

$$\xi(T) = \frac{1}{\log \operatorname{ctnh}(J/k_{\rm B}T)} \quad . \tag{6.19}$$

Note that $\xi(T)$ grows as $T \to 0$ as $\xi \approx \frac{1}{2} \, e^{2J/k_{\rm B}T}$.

### Chain with free ends

When the chain has free ends, *i.e.* open boundary conditions (OBC), there are $(L-1)$ links and the partition function is

$$Z_{\rm OBC}(T, L, H) = \sum_{\sigma,\sigma'} \big(R^{L-1}\big)_{\sigma\sigma'} = \sum_{\sigma,\sigma'} \Big\{ \lambda_+^{L-1} \, \psi_+(\sigma) \, \psi_+(\sigma') + \lambda_-^{L-1} \, \psi_-(\sigma) \, \psi_-(\sigma') \Big\} \quad , \tag{6.20}$$

where $\psi_\pm(\sigma) = \langle \, \sigma \, | \, \pm \, \rangle$. When $H = 0$, we make use of eqn. 6.17 to obtain

$$R^{L-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \big(2\cosh \beta J\big)^{L-1} + \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \big(2\sinh \beta J\big)^{L-1} \quad , \tag{6.21}$$

and therefore $Z_{\text{OBC}} = 2^L \cosh^{L-1}(\beta J)$. There's a simple trick to obtain the zero field partition function which amounts to a change of variables. We define $\nu_n \equiv \sigma_n \sigma_{n+1}$ for $1 \le n < L$. Thus, $\nu_1 = \sigma_1\sigma_2$, $\nu_2 = \sigma_2\sigma_3$, *etc.* Note that each $\nu_j$ takes the values $\pm 1$. The Hamiltonian for the chain is

$$H_{\text{OBC}} = -J \sum_{n=1}^{L-1} \sigma_n \, \sigma_{n+1} = -J \sum_{n=1}^{L-1} \nu_n \quad . \tag{6.22}$$

The state of the system is defined by the $L$ Ising variables $\{\sigma_1 \, , \, \nu_1 \, , \, \ldots \, , \, \nu_{L-1}\}$. Note that $\sigma_1$ doesn't appear in the Hamiltonian. Thus, the interacting model is recast as $L-1$ noninteracting Ising spins, and the partition function is

$$\begin{aligned} Z_{\text{OBC}}(T, L, H) = \operatorname{Tr} e^{-\beta H_{\text{OBC}}} &= \sum_{\sigma_1} \sum_{\nu_1} \cdots \sum_{\nu_{L-1}} e^{\beta J \nu_1} e^{\beta J \nu_2} \cdots e^{\beta J \nu_{L-1}} \\ &= \sum_{\sigma_1} \left( \sum_{\nu} e^{\beta J \nu} \right)^{L-1} = 2^L \cosh^{L-1}(\beta J) \quad . \end{aligned} \tag{6.23}$$

### 6.2.3 Domain walls

We have just seen how in one dimension, the Ising model never achieves long-ranged spin order. That is, the spin-spin correlation function decays asymptotically as an exponential function of the distance with a correlation length $\xi(T)$ which is finite for all $T > 0$. Only for $T = 0$ does the correlation length diverge. At $T = 0$, there are two ground states, $|\Uparrow\rangle$ and $|\Downarrow\rangle$. To choose between these ground states, we can specify a boundary condition at the ends of our one-dimensional chain, where we demand that the spins are up. Equivalently, we can apply a magnetic field $H$ of order $1/L$, which vanishes in the thermodynamic limit, but which at zero temperature will select the 'all up' ground state. At finite temperature, there is always a finite probability for any consecutive pair of sites $(n, n+1)$ to be in a high energy state, *i.e.* either $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$. Such a configuration is called a *domain wall*, and in one-dimensional systems domain walls live on individual links. Relative to the configurations $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$, a domain wall costs energy $2J$. For a system with $M = xL$ domain walls, the free energy in the thermodynamic limit is

$$F = E_0 + 2MJ - k_{\text{B}} T \log \binom{L}{M} = L \cdot \Big\{ (2x-1)J + k_{\text{B}} T \Big[ x \log x + (1-x) \log(1-x) \Big] \Big\} \quad , \tag{6.24}$$

where $E_0 = -LJ$ is the ground state energy. Minimizing the free energy with respect to $x$, one finds $x = 1/\big(e^{2J/k_{\text{B}}T} + 1\big)$, so the equilibrium concentration of domain walls is finite, meaning there can be no long-ranged spin order. In one dimension, entropy wins and there is always a thermodynamically large number of domain walls in equilibrium. And since the correlation length for $T > 0$ is finite, any boundary conditions imposed at spatial infinity will have no thermodynamic consequences since they will only be 'felt' over a finite range.

As we shall discuss in the following chapter, this consideration is true for any system with sufficiently short-ranged interactions and a discrete global symmetry. Another example is the $q$-state Potts model,

$$H = -J \sum_{\langle ij \rangle} \delta_{\sigma_i, \sigma_j} - h \sum_i \delta_{\sigma_i, 1} \quad . \tag{6.25}$$

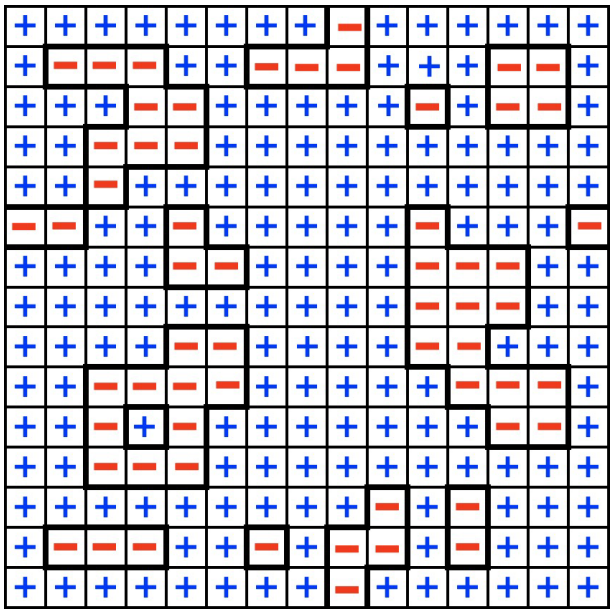


Figure 6.1: Domain walls (thick black lines) in the $d = 2$ Ising model.

Here, the spin variables $\sigma_i$ take values in the set $\{1, 2, \ldots, q\}$ on each site. The equivalent of an external magnetic field in the Ising case is a field $h$ which prefers a particular value of $\sigma$ ($\sigma = 1$ in the above Hamiltonian). See the appendix in §6.8 for a transfer matrix solution of the one-dimensional Potts model.

In dimensions $d > 1$ the Ising model exhibits spontaneous symmetry breaking below a critical temperature $T_c$. Consider the two-dimensional Ising model on a square lattice. We'll impose periodic boundary conditions on a $L \times L$ region and later imagine taking the thermodynamic limit. Each spin configuration $|\,\boldsymbol{\sigma}\,\rangle$ can be represented as an $L^2$-digit binary number $\boldsymbol{n}$, with $n_{\boldsymbol{r}} = 0$ if $\sigma_{\boldsymbol{r}} = -1$ and $n_{\boldsymbol{r}} = 1$ if $\sigma_{\boldsymbol{r}} = +1$. Thus $n_{\boldsymbol{r}} = \frac{1}{2}(1 + \sigma_{\boldsymbol{r}})$ where $\boldsymbol{r}$ labels a plaquette. We can also label each spin configuration by a configuration of *closed loops* in the following way. Let $\Gamma$ represent a set of closed loops, represented by the thick black lines in fig. 6.1. Start with each $\sigma_{\boldsymbol{r}} = +1$ and in each plaquette $\boldsymbol{r}$ flip the spin to $\sigma_{\boldsymbol{r}} = -1$ if $\boldsymbol{r}$ is enclosed by an odd number of loops. The energy of this configuration $|\,\Gamma\,\rangle$ is given by $E_\Gamma = 2JL_\Gamma$, where $L_\Gamma$ is the total perimeter of all the individual loops comprising $\Gamma$. This is because spins lying on opposite sides of a loop contribute an energy $-J\sigma_{\boldsymbol{r}}\sigma_{\boldsymbol{r}'} = +J$ instead of $-J$, which is their energy in the fiducial configuration with all spins up[7]. Thus the partition sum may be written as

$$Z = \mathrm{Tr}\; e^{-\beta\hat{H}} = 2\sum_{\Gamma} e^{-2\beta JL_\Gamma} = 2\sum_{P=0}^{\infty} N_P\, e^{-2\beta JP} \quad . \tag{6.26}$$

Here we have subtracted off the ground state energy from the start, setting $E_0 \equiv 0$. The second sum is over the total perimeter $P$, with $N_P$ the number of loop configurations with total perimeter $P$. But what is $N_P$? For a very crude approximation, we might take $N_P \sim (z-1)^P$, where $z = 4$ is the coordination number of the square lattice[8]. This is because at each step of an individual loop, if we rule out retracing

[7] For each $\Gamma$ there are in fact *two* corresponding spin configurations, the second obtained by flipping all the spins of the first.

[8] A better approximation for the asymptotic behavior of $N_P$ is $N_P = CP^{-theta}(z-1)^P$, *i.e.* including a power

our previous step, the number of possible directions in which the domain wall can be continued is $z-1$. If we approximate $N_P \sim (z-1)^P$, then we have

$$N_P\, e^{-2\beta JP} \sim \exp\Big(\big[\log(z-1) - 2\beta J\big]P\Big) \tag{6.27}$$

and the argument of the exponential starts to increase with $P$ when $T > T_{\rm c}$, where $k_{\rm B}T_{\rm c} = 2J/\log(z-1)$. Thus, below $T_{\rm c}$ the loops are suppressed, while above $T_{\rm c}$ the loops proliferate. This provides us a very rough picture of the phase transition and an estimate of where it occurs. For the square lattice, where $z = 4$, we obtain an estimate of $k_{\rm B}T_{\rm c}/J \approx 2/\log 3 = 1.8205$. The exact location of the phase transition, as we will derive below, is given by $k_{\rm B}T_{\rm c}/J = 2/\log(1+\sqrt{2}) = 2.2692$.

### 6.2.4 Ising model in two dimensions : Peierls' argument

A more careful version of our previous argument, due to Rudolf Peierls, shows rigorously that there will be a finite temperature phase transition for the Ising model on the square lattice[9]. Consider the Ising model, in zero magnetic field, on a $N = L_x \times L_y$ square lattice, with $L_{x,y} \to \infty$ in the thermodynamic limit. Along the perimeter of the system we impose the boundary condition $\sigma_i = +1$. Any configuration of the spins may then be represented uniquely in the following manner. Start with a configuration in which all spins are up. Next, draw a set of closed loops on the lattice, as we did previously. By definition, the loops cannot share any links along their boundaries, *i.e.* each link on the lattice is associated with at most one such loop. Now flip all the spins inside each loop from up to down. Identify each such loop configuration with a label $\Gamma$. The partition function is

$$Z = \mathsf{Tr}\; e^{-\beta\hat{H}} = \sum_{\Gamma} e^{-2\beta J L_\Gamma} \quad , \tag{6.28}$$

where $L_\Gamma$ is the total perimeter of the loop configuration $\Gamma$. The domain walls are now loops, rather than individual links, but as in the one-dimensional case, each link of each domain wall contributes an energy $+2J$ relative to the ground state.

Now we wish to compute the average magnetization of the central site (assume $L_{x,y}$ are both odd, so there is a unique central site). This is given by the difference $P_+(0) - P_-(0)$, where $P_\mu(0) = \big\langle \delta_{\sigma_0,\mu} \big\rangle$ is the probability that the central spin has spin polarization $\mu$. If $P_+(0) > P_-(0)$, then the magnetization per site $m = P_+(0) - P_-(0)$ is finite in the thermodynamic limit, and the system is ordered. Clearly

$$P_+(0) = \frac{1}{Z} \sum_{\Gamma\in\Sigma_+} e^{-2\beta J L_\Gamma} \quad , \tag{6.29}$$

where the restriction on the sum indicates that only those configurations where the central spin is up ($\sigma_0 = +1$) are to be included. (see fig. 6.2a). Similarly,

$$P_-(0) = \frac{1}{Z} \sum_{\tilde{\Gamma}\in\Sigma_-} e^{-2\beta J L_{\tilde{\Gamma}}} \quad , \tag{6.30}$$

law prefactor in addition to the exponential.

[9] Here we modify slightly the discussion in chapter 5 of the book by L. Peliti.

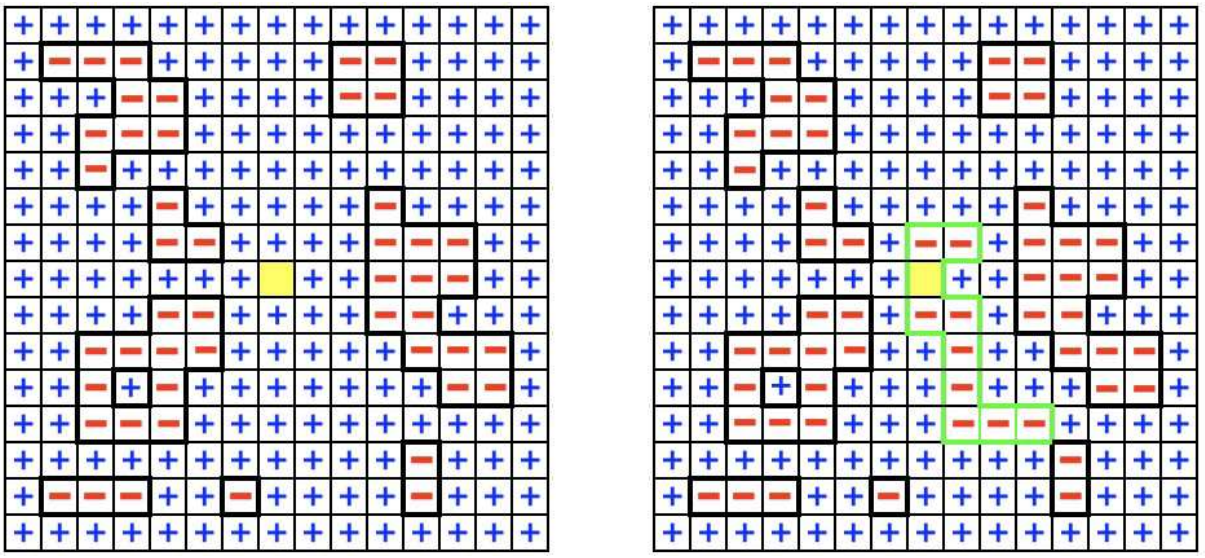


Figure 6.2: Clusters and boundaries for the square lattice Ising model. Left panel: a configuration $\Gamma$ where the central spin is up. Right panel: a configuration $C_\gamma \circ \Gamma$ where the interior spins of a new loop $\gamma$ containing the central spin have been flipped.

where only configurations in which $\sigma_0 = -1$ are included in the sum, where $\Sigma_\pm \equiv \Big\{ \Gamma \;\Big|\; \sigma_0 = \pm \Big\}$. That is, $\Sigma_+(\Sigma_-)$ is the set of configurations $\Gamma$ in which the central spin is always up (down). Consider now the construction in fig. 6.2b. Any loop configuration $\widetilde{\Gamma} \in \Sigma_-$ may be associated with a unique loop configuration $\Gamma \in \Sigma_+$ by reversing all the spins within the loop of $\widetilde{\Gamma}$ which contains the origin. Note that the map from $\widetilde{\Gamma}$ to $\Gamma$ is many-to-one. That is, we can write $\widetilde{\Gamma} = C_\gamma \circ \Gamma$, where $C_\gamma$ overturns the spins within the loop $\gamma$, with the conditions that (i) $\gamma$ contains the origin, and (ii) none of the links in the perimeter of $\gamma$ coincide with any of the links from the constituent loops of $\Gamma$. Let us denote this set of loops as $\Upsilon_\Gamma$:

$$\Upsilon_\Gamma = \Big\{ \gamma \;:\; 0 \in \mathrm{int}(\gamma) \text{ and } \gamma \cap \Gamma = \emptyset \Big\} \quad . \tag{6.31}$$

Then

$$m = P_+(0) - P_-(0) = \frac{1}{Z} \sum_{\Gamma \in \Sigma_+} e^{-2\beta J L_\Gamma} \left( 1 - \sum_{\gamma \in \Upsilon_\Gamma} e^{-2\beta J L_\gamma} \right) \quad . \tag{6.32}$$

If we can prove that $\sum_{\gamma \in \Upsilon_\Gamma} e^{-2\beta J L_\gamma} < 1$, then we will have established that $m > 0$. Let us ask: how many loops $\gamma$ are there in $\Upsilon_\Gamma$ with perimeter $L$? We cannot answer this question exactly, but we can derive a rigorous upper bound for this number, which, following Peliti, we call $g(L)$. We claim that

$$g(L) < \frac{2}{3L} \cdot 3^L \cdot \left( \frac{L}{4} \right)^2 = \frac{L}{24} \cdot 3^L \quad . \tag{6.33}$$

To establish this bound, consider any site on such a loop $\gamma$. Initially we have $4$ possible directions to proceed to the next site, but thereafter there are only $3$ possibilities for each subsequent step, since the loop cannot run into itself. This gives $4 \cdot 3^{L-1}$ possibilities. But we are clearly overcounting, since any

point on the loop could have been chosen as the initial point, and moreover we could have started by proceeding either clockwise or counterclockwise. So we are justified in dividing this by $2L$. We are still overcounting, because we have not accounted for the constraint that $\gamma$ is a closed loop, nor that $\gamma \cap \Gamma = \emptyset$. We won't bother trying to improve our estimate to account for these constraints. However, we are clearly undercounting due to the fact that a given loop can be translated in space so long as the origin remains within it. To account for this, we multiply by the area of a square of side length $L/4$, which is the maximum area that can be enclosed by a loop of perimeter $L$. We therefore arrive at eqn. 6.33. Finally, we note that the smallest possible value of $L$ is $L = 4$, corresponding to a square enclosing the central site alone. Therefore

$$\sum_{\gamma \in \Upsilon_\Gamma} e^{-2\beta J L_\gamma} < \frac{1}{12} \sum_{k=2}^{\infty} k \cdot \left(3\, e^{-2\beta J}\right)^{2k} = \frac{x^4\,(2-x^2)}{12\,(1-x^2)^2} \equiv r \quad , \tag{6.34}$$

where $x = 3\, e^{-2\beta J}$. Note that we have accounted for the fact that the perimeter $L$ of each loop $\gamma$ must be an even integer. The sum is smaller than unity provided $x < x_0 = 0.869756\ldots$, hence the system is ordered provided

$$\frac{k_{\rm B} T}{J} < \frac{2}{\log(3/x_0)} = 1.61531 \quad , \tag{6.35}$$

which establishes a rigorous lower bound for the critical temperature. The exact result, which we will establish in §6.2.6 below, is $k_{\rm B} T_{\rm c}/J = 2/\sinh^{-1}(1) \simeq 2.26918$. Peierls' argument has been generalized to higher dimensional lattices as well[10].

With a little more work we can derive a bound for the magnetization. We have shown that

$$P_-(0) = \frac{1}{Z} \sum_{\Gamma \in \Sigma_+} e^{-2\beta J L_\Gamma} \sum_{\gamma \in \Upsilon_\Gamma} e^{-2\beta J L_\gamma} < r \cdot \frac{1}{Z} \sum_{\Gamma \in \Sigma_+} e^{-2\beta J L_\Gamma} = r\, P_+(0) \quad . \tag{6.36}$$

Thus,

$$1 = P_+(0) + P_-(0) < (1+r)\, P_+(0) \tag{6.37}$$

and therefore

$$m = P_+(0) - P_-(0) > (1-r)\, P_+(0) > \frac{1-r}{1+r} \quad , \tag{6.38}$$

where $r(T)$ is given in eqn. 6.34.

### Importance of interaction range

We showed that the one-dimensional Ising model has no finite temperature phase transition, and is disordered at any finite temperature $T$, but in two dimensions on the square lattice there is a finite critical temperature $T_{\rm c}$ below which there is long-ranged order. Consider now the construction depicted in fig. 6.3, where the sites of a two-dimensional square lattice are mapped onto those of a linear chain[11]. Clearly we can elicit a one-to-one mapping between the sites of a two-dimensional square lattice and

[10] See. *e.g.* J. L. Lebowitz and A. E. Mazel, *J. Stat. Phys.* **90**, 1051 (1998).

[11] A corresponding mapping can be found between a cubic lattice and the linear chain as well.

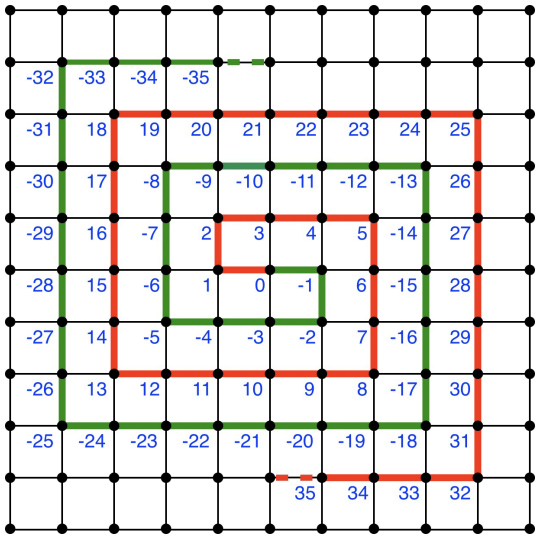



Figure 6.3: A two-dimensional square lattice mapped onto a one-dimensional chain.

those of a one-dimensional chain. That is, the two-dimensional square lattice Ising model may be written as a one-dimensional Ising model, *i.e.*

$$\hat{H} = -J \sum_{\langle ij \rangle}^{\substack{\text{square} \\ \text{lattice}}} \sigma_i \, \sigma_j = -\sum_{n,n'}^{\substack{\text{linear} \\ \text{chain}}} J_{nn'} \, \sigma_n \, \sigma_{n'} \quad . \tag{6.39}$$

How can this be consistent with the results we have just proven?

The fly in the ointment here is that the interaction along the chain $J_{n,n'}$ is long-ranged. This is apparent from inspecting the site labels in fig. 6.3. Note that site $n = 15$ is linked to sites $n' = 14$ and $n' = 16$, but also to sites $n' = -6$ and $n' = -28$. With each turn of the concentric spirals in the figure, the range of the interaction increases. To complicate matters further, the interactions are no longer translationally invariant, *i.e.* $J_{nn'} \neq J(n - n')$. But it is the long-ranged nature of the interactions on our contrived one-dimensional chain which spoils our previous energy-entropy argument, because now the domain walls themselves interact via a long-ranged potential. Consider the linear chain with $J_{n,n'} = J \, |n - n'|^{-\alpha}$, where $\alpha > 0$. Let us compute the energy of a domain wall configuration where $\sigma_n = +1$ if $n > 0$ and $\sigma_n = -1$ if $n \leq 0$. The domain wall energy is then

$$\Delta = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \frac{2J}{|m+n|^\alpha} \quad . \tag{6.40}$$

Here we have written one of the sums in terms of $m = -n'$. For asymptotically large $m$ and $n$, we can write $\boldsymbol{R} = (m, n)$ and we obtain an integral over the upper right quadrant of the plane:

$$\int_1^\infty dR \, R \int_0^{\pi/2} d\phi \, \frac{2J}{R^\alpha \, (\cos\phi + \sin\phi)^\alpha} = 2^{-\alpha/2} \int_{-\pi/4}^{\pi/4} \frac{d\phi}{\cos^\alpha \phi} \int_1^\infty \frac{dR}{R^{\alpha - 1}} \quad . \tag{6.41}$$

The $\phi$ integral is convergent, but the $R$ integral diverges for $\alpha \leq 2$. For a finite system, the upper bound on the $R$ integral becomes the system size $L$. For $\alpha > 2$ the domain wall energy is finite in the thermodynamic limit $L \to \infty$. In this case, entropy again wins. *I.e.* the entropy associated with a single domain wall is $k_{\rm B} \log L$, and therefore $F = E - TS$ is always lowered by having a finite density of domain walls. For $\alpha < 2$, the energy of a single domain wall scales as $L^{2-\alpha}$. It was first proven by F. J. Dyson in 1969 that this model has a finite temperature phase transition provided $1 < \alpha < 2$. There is no transition for $\alpha < 1$ or $\alpha > 2$. The case $\alpha = 2$ is special, and is discussed in the beautiful renormalization group analysis by J. M. Kosterlitz in *Phys. Rev. Lett.* **37**, 1577 (1976).

### 6.2.5 High temperature expansion

Consider once again the ferromagnetic Ising model in zero field ($H = 0$), but on an arbitrary lattice. The partition function is

$$Z = \mathsf{Tr}\, e^{\beta J \sum_{\langle ij \rangle} \sigma_i \sigma_j} = \big(\cosh \beta J\big)^{N_{\rm nn}} \, \mathsf{Tr} \left\{ \prod_{\langle ij \rangle} \big(1 + x\, \sigma_i \sigma_j\big) \right\} \quad , \tag{6.42}$$

where $x = \tanh \beta J$ and $N_{\rm nn}$ is the number of links (*i.e.* bonds). For regular lattices, $N_{\rm nn} = \frac{1}{2} z N$, where $N$ is the number of lattice sites and $z$ is the lattice coordination number, *i.e.* the number of nearest neighbors for each site. We have used

$$e^{\beta J \sigma \sigma'} = \cosh \beta J \cdot \Big\{ 1 + \sigma \sigma' \tanh \beta J \Big\} = \begin{cases} e^{+\beta J} & \text{if } \sigma \sigma' = +1 \\ e^{-\beta J} & \text{if } \sigma \sigma' = -1 \end{cases} \quad . \tag{6.43}$$

We expand eqn. 6.42 in powers of $x$, resulting in a sum of $2^{N_{\rm nn}}$ terms, each of which can be represented graphically in terms of so-called *lattice animals*. A lattice animal is a distinct (including reflections and rotations) arrangement of adjacent plaquettes on a lattice. In order that the trace not vanish, only such configurations and their compositions are permitted. This is because each $\sigma_i$ for every given site $i$ must occur an even number of times in order for a given term in the sum not to vanish. For all such terms, the trace is $2^N$. Let $\{\Gamma\}$ represent a collection of lattice animals, and $g_\Gamma$ the multiplicity of $\Gamma$. Then

$$Z = 2^N \big(\cosh \beta J\big)^{N_{\rm nn}} \sum_\Gamma g_\Gamma \, \big(\tanh \beta J\big)^{L_\Gamma} \quad , \tag{6.44}$$

where $L_\Gamma$ is the total number of links in the diagram $\Gamma$, and $g_\Gamma$ is the multiplicity of $\Gamma$. Since $x$ vanishes as $T \to \infty$, this procedure is known as the *high temperature expansion* (HTE).

For the square lattice, the enumeration of all lattice animals with up to order eight is given in fig. 6.4. For the diagram represented as a single elementary plaquette, there are $N$ possible locations for the lower left vertex. For the $2 \times 1$ plaquette animal, one has $g = 2N$, because there are two inequivalent orientations as well as $N$ translations. For two disjoint elementary squares, one has $g = \frac{1}{2} N(N-5)$, which arises from subtracting $5N$ 'illegal' configurations involving double lines (remember each link in the partition sum appears only once!), shown in the figure, and finally dividing by two because the individual squares are identical. Note that $N(N-5)$ is always even for any integer value of $N$. Thus, to lowest interesting order on the square lattice,

$$Z = 2^N \big(\cosh \beta J\big)^{2N} \Big\{ 1 + N x^4 + 2N x^6 + \big(7 - \tfrac{5}{2}\big) N x^8 + \tfrac{1}{2} N^2 x^8 + \mathcal{O}(x^{10}) \Big\} \quad . \tag{6.45}$$

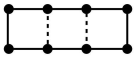
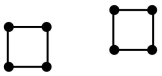
| diagram $\Gamma$ | $L_\Gamma$ | $g_\Gamma$ | remarks |
|---|---|---|---|
|  | 0 | 1 | empty lattice |
|  | 4 | $N$ | $N$ translations |
|  | 6 | $2N$ | 2 rotations, $N$ translations |
|  | 8 | $2N$ | 2 rotations, $N$ translations |
|  | 8 | $N$ | $N$ translations |
|  | 8 | $4N$ | 4 rotations, $N$ translations |
|  | 8 | $\frac{1}{2}(N^2-5N)$ | exclude five invalid diagrams; halve for equivalent squares  (1) (4) |

Figure 6.4: HTE diagrams on the square lattice and their multiplicities.

The free energy is therefore

$$\begin{aligned} F &= -k_{\rm B}T\log 2 + Nk_{\rm B}T\log(1-x^2) - Nk_{\rm B}T\Big[x^4 + 2\,x^6 + \tfrac{9}{2}\,x^8 + \mathcal{O}(x^{10})\Big] \\ &= Nk_{\rm B}T\log 2 - Nk_{\rm B}T\Big\{x^2 + \tfrac{3}{2}\,x^4 + \tfrac{7}{3}\,x^6 + \tfrac{19}{4}\,x^8 + \mathcal{O}(x^{10})\Big\}, \end{aligned} \tag{6.46}$$

again with $x = \tanh\beta J$. Here we have substituted $\cosh^2\beta J = 1/(1-x^2)$ to write the final result as a power series in $x$. Notice that the $\mathcal{O}(N^2)$ factor in $Z$ has cancelled upon taking the logarithm, so the free energy is properly extensive.

For the one-dimensional chain or ring, the high temperature expansion yields

$$Z_{\rm OBC}(T,N) = 2^N\cosh^{N-1}\beta J \qquad , \qquad Z_{\rm PBC}(T,N) = 2^N\big(\cosh^N\beta J + \sinh^N\beta J\big) \quad , \tag{6.47}$$

in agreement with the transfer matrix calculations. Notice how for OBC there is only one lattice animal which contributes, *i.e.* the empty lattice. For PBC there are two contributing animals, the empty lattice and the entire ring. In higher dimensions, where there is a finite temperature phase transition, one typically computes the specific heat $c(T)$ and tries to extract its singular behavior in the vicinity of $T_{\rm c}$, where $c(T) \sim A\,(T-T_{\rm c})^{-\alpha}$. Since $x(T) = \tanh(J/k_{\rm B}T)$ is analytic in $T$, we have $c(x) \sim A'\,(x-x_{\rm c})^{-\alpha}$, where $x_{\rm c} = x(T_{\rm c})$. One assumes $x_{\rm c}$ is the singularity closest to the origin and corresponds to the radius of convergence of the high temperature expansion. If we write

$$c(x) = \sum_{n=0}^{\infty} a_n\,x^n \sim A''\left(1-\frac{x}{x_{\rm c}}\right)^{-\alpha} \quad , \tag{6.48}$$

then according to the binomial theorem we should expect

$$\frac{a_n}{a_{n-1}} = \frac{1}{x_{\rm c}}\left[1 - \frac{1-\alpha}{n}\right] \quad . \tag{6.49}$$

Thus, by plotting $a_n/a_{n-1}$ *versus* $1/n$, one extracts $1/x_{\rm c}$ as the intercept, and $(\alpha - 1)/x_{\rm c}$ as the slope.

### High temperature expansion for correlation functions

Can we also derive a high temperature expansion for the spin-spin correlation function $C_{kl} = \langle \sigma_k \, \sigma_l \rangle$ ? Yes we can. We have

$$C_{kl} = \frac{{\rm Tr}\left[\sigma_k \, \sigma_l \, e^{\beta J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j}\right]}{{\rm Tr}\left[e^{\beta J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j}\right]} \equiv \frac{Y_{kl}}{Z} \quad . \tag{6.50}$$

Recall our analysis of the partition function $Z$. We concluded that in order for the trace not to vanish, the spin variable $\sigma_i$ on each site $i$ must occur an even number of times in the expansion of the product. Similar considerations hold for $Y_{kl}$, except now due to the presence of $\sigma_k$ and $\sigma_l$, those variables now must occur an *odd* number of times when expanding the product. It is clear that the only nonvanishing diagrams will be those in which there is a finite string connecting sites $k$ and $l$, in addition to the usual closed HTE loops. See fig. 6.5 for an instructive sketch. One then expands both $Y_{kl}$ as well as $Z$ in powers of $x = \tanh \beta J$, taking the ratio to obtain the correlator $C_{kl}$. At high temperatures ($x \to 0$), both numerator and denominator are dominated by the configurations $\Gamma$ with the shortest possible total perimeter. For $Z$, this means the trivial path $\Gamma = \{\emptyset\}$, while for $Y_{kl}$ this means finding the shortest length path from $k$ to $l$. (If there is no straight line path from $k$ to $l$, there will in general be several such minimizing paths.) Note, however, that the presence of the string between sites $k$ and $l$ complicates the analysis of $g_\Gamma$ for the closed loops, since none of the links of $\Gamma$ can intersect the string. It is worth stressing that this does not mean that the string and the closed loops cannot intersect at isolated *sites*, but only that they share no common *links*; see once again fig. 6.5.

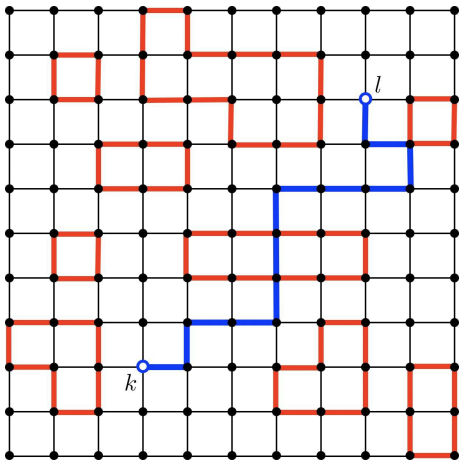



Figure 6.5: HTE diagrams for the numerator $Y_{kl}$ of the correlation function $C_{kl}$. The blue path connecting sites $k$ and $l$ is the string. The remaining red paths are all closed loops.

### 6.2.6 Kramers-Wannier duality

Consider an Ising model on a square lattice of size $L_x \times L_y$, with periodic boundary conditions, and let us set $H = 0$, so there is no explicit symmetry breaking. We label each plaquette on the square lattice by a vector $\boldsymbol{R}$ which may be taken to be the position of its center. We write the Ising model Hamiltonian as

$$\hat{H}(J_x, J_y) = -J_x \sum_{\boldsymbol{R}} \sigma_{\boldsymbol{R}} \sigma_{\boldsymbol{R}+\hat{\boldsymbol{x}}} - J_y \sum_{\boldsymbol{R}} \sigma_{\boldsymbol{R}} \sigma_{\boldsymbol{R}+\hat{\boldsymbol{y}}} \quad , \tag{6.51}$$

taking $J_{x,y} > 0$.[12] The ground state $|\Uparrow\rangle$ satisfies $\hat{H}\,|\Uparrow\rangle = E_0\,|\Uparrow\rangle$, where $E_0 = -N(J_x + J_y)$. Note that $|\Downarrow\rangle$ is also a ground state, obtained from $|\Uparrow\rangle$ by a $\mathbb{Z}_2$ transformation which flips all the plaquette spins and leaves the Hamiltonian invariant. Starting from $|\Uparrow\rangle$, let the state $|\,\Gamma\,\rangle$ represent the configuration where all spins within each of the individual loops comprising $\Gamma$ have been flipped, as discussed in §6.2.4 and depicted in fig. 6.2. Let $n^\Gamma_{x,y}$ denote the total number of $x$- and $y$-directed links in $\Gamma$, respectively. The partition function $Z(T, J_x, J_y)$ may be written as

$$Z(T, J_x, J_y) = 2 \sum_\Gamma e^{(N-2n^\Gamma_x)\beta J_y}\, e^{(N-2n^\Gamma_y)\beta J_x} \quad , \tag{6.52}$$

since the spins on the plaquettes located at $\boldsymbol{R}$ and $\boldsymbol{R} + \hat{\boldsymbol{x}}$ on opposite sides of each $\hat{\boldsymbol{y}}$-directed link in $\Gamma$ are anti-aligned, according to the construction of $|\,\Gamma\,\rangle$, thereby raising the energy by $2J_x$ relative to $E_0$. Similarly, the anti-aligned plaquette spins at $\boldsymbol{R}$ and $\boldsymbol{R} + \hat{\boldsymbol{y}}$ on opposite sides of an $\hat{\boldsymbol{x}}$-directed link raise the energy by $2J_y$. The overall factor of 2 is because the set of states $\Gamma$ is only $2^{N-1}$ in number. For example, the state $|\Downarrow\rangle$ cannot be constructed according to the rules. For each state $|\,\Gamma\,\rangle$, we can obtain a unique symmetry-related state $|\,\Gamma'\,\rangle$ by applying a $\mathbb{Z}_2$ global spin flip, which completes the full set of $2^N$ spin configurations.

Now define the *dual* couplings $K_{x,y}$ as follows[13]:

$$e^{-2\beta J_x} = \tanh(\beta K_y) \quad , \quad e^{-2\beta J_y} = \tanh(\beta K_x) \quad \Rightarrow \quad \beta K_{x,y} = \tfrac{1}{2} \log\, \mathrm{ctnh}\,(\beta J_{y,x}) \quad . \tag{6.53}$$

In the isotropic case $J_x = J_y = J$ we have $\exp(-2\beta J) = \tanh(\beta K)$ and thus for $J \to 0$ we have $K \to \infty$ while for $J \to \infty$ we have $K \to 0$. Thus, in the phase diagram of the dual couplings, *the temperature is inverted*. We now have

$$Z(T, J_x, J_y) = \big[\tanh(\beta K_y)\tanh(\beta K_x)\big]^{-N/2} \cdot 2 \sum_\Gamma \big[\tanh(\beta K_y)\big]^{n^\Gamma_x} \big[\tanh(\beta K_x)\big]^{n^\Gamma_y} \quad . \tag{6.54}$$

We next invoke the identity

$$2 \sum_\Gamma \big[\tanh(\beta K_y)\big]^{n^\Gamma_x} \big[\tanh(\beta K_x)\big]^{n^\Gamma_y} = 2^{-N}\, \mathrm{Tr} \prod_{\boldsymbol{r}} \big[1 + \tanh(\beta K_x)\, \mu_{\boldsymbol{r}} \mu_{\boldsymbol{r}+\hat{\boldsymbol{x}}}\big] \big[1 + \tanh(\beta K_y)\, \mu_{\boldsymbol{r}} \mu_{\boldsymbol{r}+\hat{\boldsymbol{y}}}\big] \quad , \tag{6.55}$$

[12] Thus we assume the model is ferromagnetic. If $J_{x,y} < 0$ the model is antiferromagnetic, which we may rewrite as a ferromagnet defining $\sigma'_{\boldsymbol{R}} = -\sigma_{\boldsymbol{R}}$ on one sublattice (*i.e.* the white squares on a chessboard) and retaining $\sigma'_{\boldsymbol{R}} = \sigma_{\boldsymbol{R}}$ on the other, effectively reversing the signs of $J_{x,y}$.

[13] Note that if $\exp(-2y) = \tanh x$ then $\exp(-2x) = \tanh y$.

where $\mu_{\boldsymbol{r}} \in \{+1, -1\}$ is an Ising spin located at the *site* $\boldsymbol{r}$. Note that $\boldsymbol{R}$ denotes a location of a *plaquette* while $\boldsymbol{r}$ denotes the location of a *site*. Each site can be uniquely associated with a plaquette (*e.g.*, the site in the lower left corner), so there are $N$ sites and $2^N$ configurations $\{\mu_{\boldsymbol{r}}\}$. But from where do we get eqn. 6.55? Observe that the product on the RHS can be expressed as giant sum over $2^{2N}$ terms. Each such term has a graphical interpretation as a collection of nearest neighbor links $\langle \boldsymbol{r}\boldsymbol{r}' \rangle$, and we label each such collection by $\Theta$. Unless each site $\boldsymbol{r}$ appears an even number of times (0, 2, or 4) in $\Theta$, when we take the trace over the spin configurations the term will vanish. The surviving $\Theta$ configurations then correspond to the set of loop configurations $\{\Gamma\}$! We then have from eqn. 6.42 that

$$\mathrm{Tr} \prod_{\boldsymbol{r}} \big[1+\tanh(\beta K_x)\,\sigma_{\boldsymbol{r}}\sigma_{\boldsymbol{r}+\hat{\boldsymbol{x}}}\big]\big[1+\tanh(\beta K_y)\,\sigma_{\boldsymbol{r}}\sigma_{\boldsymbol{r}+\hat{\boldsymbol{y}}}\big] = \big[\cosh(\beta K_x)\big]^{-N}\big[\cosh(\beta K_y)\big]^{-N}\,\mathrm{Tr}\; e^{-\beta \hat{H}^{\mathrm{D}}(K_x,K_y)} \quad , \tag{6.56}$$

where

$$\hat{H}^{\mathrm{D}}(K_x, K_y) = -K_x \sum_{\boldsymbol{r}} \sigma_{\boldsymbol{r}}\,\sigma_{\boldsymbol{r}+\hat{\boldsymbol{x}}} - K_y \sum_{\boldsymbol{r}} \sigma_{\boldsymbol{r}}\,\sigma_{\boldsymbol{r}+\hat{\boldsymbol{y}}} \quad . \tag{6.57}$$

is the *dual Hamiltonian* defined on the *dual lattice* of sites. The plaquette lattice is *dual* to the site lattice (and vice versa). Therefore we have

$$Z(T, J_x, J_y) = \big[\sinh(2\beta K_x)\sinh(2\beta K_y)\big]^{-N/2}\, Z(T, K_x, K_y) \quad . \tag{6.58}$$

This establishes a *duality relation* between the original and inverted temperature regions. Assuming that the ferromagnetic square lattice model has a unique transition separating its high- and low-temperature phases, duality requires that transition to lie on the *self-dual line*. Thus, at the critical point the system must satisfy $J_{x,y} = K_{x,y}$, and we have the relation

$$\sinh(2\beta_{\mathrm{c}} J_x)\sinh(2\beta_{\mathrm{c}} J_y) = 1 \quad . \tag{6.59}$$

This equation defines a critical surface $T_{\mathrm{c}}(J_x, J_y)$. If $J_x = J_y = J$, we have

$$k_{\mathrm{B}} T_{\mathrm{c}} = \frac{2J}{\sinh^{-1}(1)} = \frac{2J}{\log\big(1+\sqrt{2}\big)} \simeq 2.26918\, J \quad . \tag{6.60}$$

Thus we have derived the exact critical temperature for the two-dimensional square lattice Ising model.

## 6.3 Nonideal Classical Gases

Let's switch gears now and return to the study of continuous classical systems described by a Hamiltonian $\hat{H}\big(\{\boldsymbol{x}_i\}, \{\boldsymbol{p}_i\}\big)$. In the next chapter, we will see how the critical properties of classical fluids can in fact be modeled by an appropriate *lattice gas* Ising model, and we'll derive methods for describing the liquid-gas phase transition in such a model.

### 6.3.1 The configuration integral

Consider the ordinary canonical partition function for a nonideal system of identical point particles interacting via a central two-body potential $u(r)$. We work in the ordinary canonical ensemble. The

$N$-particle partition function is

$$\begin{aligned} Z(T,V,N) &= \frac{1}{N!}\int\prod_{i=1}^{N}\frac{d^dp_i\,d^dx_i}{h^d}\,e^{-\hat{H}/k_{\rm B}T} \\ &= \frac{\lambda_T^{-Nd}}{N!}\int\prod_{i=1}^{N}d^dx_i\,\exp\left(-\frac{1}{k_{\rm B}T}\sum_{i<j}u\big(|\boldsymbol{x}_i-\boldsymbol{x}_j|\big)\right) \quad . \end{aligned} \tag{6.61}$$

Here, we have assumed a many body Hamiltonian of the form

$$\hat{H} = \sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m} + \sum_{i<j}u\big(|\boldsymbol{x}_i-\boldsymbol{x}_j|\big) \quad , \tag{6.62}$$

in which massive nonrelativistic particles interact via a two-body central potential. As before, the thermal wavelength is $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T}$. We can now write

$$Z(T,V,N) = \lambda_T^{-Nd}\,Q_N(T,V) \quad , \tag{6.63}$$

where the *configuration integral* $Q_N(T,V)$ is given by

$$Q_N(T,V) = \frac{1}{N!}\int d^dx_1\cdots\int d^dx_N\,\prod_{i<j}e^{-\beta u(r_{ij})} \quad . \tag{6.64}$$

There are no general methods for evaluating the configurational integral exactly.

### 6.3.2 One-dimensional Tonks gas

The Tonks gas is a one-dimensional generalization of the hard sphere gas. Consider a one-dimensional gas of indistinguishable particles of mass $m$ interacting via the potential

$$u(x-x') = \begin{cases}\infty & \text{if } |x-x'|<a \\ 0 & \text{if } |x-x'|\geq a\end{cases} \quad . \tag{6.65}$$

Thus, the Tonks gas may be considered to be a gas of *hard rods*. The above potential guarantees that the portion of configuration space in which any rods overlap is *forbidden* in this model. Let the gas be placed in a finite volume $L$. The hard sphere nature of the particles means that no particle can get within a distance $\frac{1}{2}a$ of the ends at $x=0$ and $x=L$. That is, there is a one-body potential $v(x)$ acting as well, where

$$v(x) = \begin{cases}\infty & \text{if } x<\frac{1}{2}a \\ 0 & \text{if } \frac{1}{2}a\leq x\leq L-\frac{1}{2}a \\ \infty & \text{if } x>L-\frac{1}{2}a\end{cases} \quad . \tag{6.66}$$

The configuration integral of the 1D Tonks gas is given by

$$Q_N(T,L) = \frac{1}{N!}\int_0^L dx_1\cdots\int_0^L dx_N\,\chi(x_1,\ldots,x_N) \quad , \tag{6.67}$$

where $\chi = e^{-U/k_{\rm B}T}$ is zero if any two 'rods' (of length $a$) overlap, or if any rod overlaps with either boundary at $x = 0$ and $x = L$, and $\chi = 1$ otherwise. Note that $\chi$ does not depend on the temperature. Due to permutation symmetry, we may integrate over the subspace where $x_1 < x_2 < \cdots < x_N$ and then multiply the result by $N!$. Clearly $x_j$ must lie to the right of $x_{j-1} + a$ as well as to the left of $Y_j \equiv L - Na + (j - \frac{1}{2})a$. Note that since $Y_j - a = Y_{j-1}$. Thus, the configurational integral is

$$Q_N(T,L) = \int_{a/2}^{Y_1} dx_1 \int_{x_1+a}^{Y_2} dx_2 \cdots \int_{x_{N-1}+a}^{Y_N} dx_N = \int_{a/2}^{Y_1} dx_1 \int_{x_1+a}^{Y_2} dx_2 \cdots \int_{x_{N-2}+a}^{Y_{N-1}} dx_{N-1} \left(Y_{N-1} - x_{N-1}\right) \tag{6.68}$$

$$= \frac{1}{2} \int_{a/2}^{Y_1} dx_1 \int_{x_1+a}^{Y_2} dx_2 \cdots \int_{x_{N-3}+a}^{Y_{N-2}} dx_{N-2} \left(Y_{N-2} - x_{N-2}\right)^2 = \cdots = \frac{1}{k!} \int_{a/2}^{Y_1} dx_1 \int_{x_1+a}^{Y_2} dx_2 \cdots \int_{x_{N-k-1}+a}^{Y_{N-k}} dx_{N-k} \left(Y_{N-k} - x_{N-k}\right)^k$$

$$= \frac{1}{N!} \left(Y_1 - \tfrac{1}{2}a\right)^N = \frac{1}{N!} \left(L - Na\right)^N \quad .$$

The partition function is $Z(T,L,N) = \lambda_T^{-N} \, Q_N(T,L)$, and so the free energy is

$$F = -k_{\rm B}T \log Z = -Nk_{\rm B}T \left\{ -\log \lambda_T + 1 + \log\left(\frac{L}{N} - a\right) \right\} \quad , \tag{6.69}$$

where we have used Stirling's rule to write $\log N! \approx N \log N - N$. The pressure is

$$p = -\frac{\partial F}{\partial L} = \frac{nk_{\rm B}T}{1 - na} \quad , \tag{6.70}$$

where $n = N/L$ is the one-dimensional density. Note that the pressure diverges as $n$ approaches $1/a$. The usual one-dimensional ideal gas law, $pL = Nk_{\rm B}T$, is replaced by $pL_{\rm eff} = Nk_{\rm B}T$, where $L_{\rm eff} = L - Na$ is the 'free' volume obtained by subtracting the total "excluded volume" $Na$ from the original volume $L$. Note the similarity here to the van der Waals equation of state, $(p + av^{-2})(v - b) = RT$, where $v = N_{\rm A}V/N$ is the molar volume. Defining $\tilde{a} \equiv a/N_{\rm A}^2$ and $\tilde{b} \equiv b/N_{\rm A}$, we have

$$p + \tilde{a}n^2 = \frac{nk_{\rm B}T}{1 - \tilde{b}n} \quad , \tag{6.71}$$

where $n = N_{\rm A}/v$ is the number density. The term involving the constant $\tilde{a}$ is due to the long-ranged attraction of atoms due to their mutual polarizability. The term involving $\tilde{b}$ is an excluded volume effect. The Tonks gas models only the latter.

### 6.3.3 Mayer cluster expansion

Let us return to the general problem of computing the configuration integral. Consider the function $e^{-\beta u_{ij}}$, where $u_{ij} \equiv u(|\boldsymbol{x}_i - \boldsymbol{x}_j|)$. We assume that at very short distances there is a strong repulsion between particles, *i.e.* $u_{ij} \to \infty$ as $r_{ij} = |\boldsymbol{x}_i - \boldsymbol{x}_j| \to 0$, and that $u_{ij} \to 0$ as $r_{ij} \to \infty$. Thus, $e^{-\beta u_{ij}}$ vanishes as $r_{ij} \to 0$ and approaches unity as $r_{ij} \to \infty$. For our purposes, it will prove useful to define the function

$$f(r) = e^{-\beta u(r)} - 1 \quad , \tag{6.72}$$

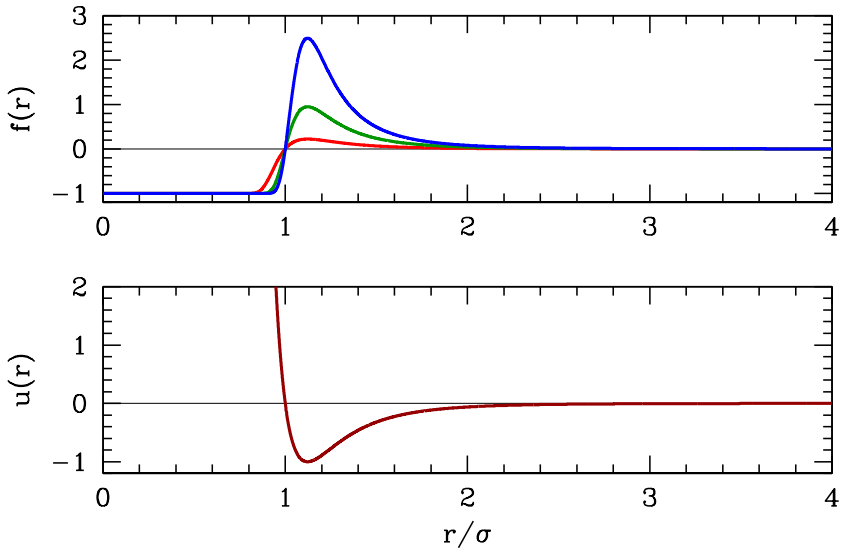



Figure 6.6: Bottom panel: Lennard-Jones potential $u(r) = 4\epsilon\left(x^{-12} - x^{-6}\right)$, with $x = r/\sigma$ and $\epsilon = 1$. Note the weak attractive tail and the strong repulsive core. Top panel: Mayer function $f(r,T) = e^{-u(r)/k_{\rm B}T} - 1$ for $k_{\rm B}T = 0.8\,\epsilon$ (blue), $k_{\rm B}T = 1.5\,\epsilon$ (green), and $k_{\rm B}T = 5\,\epsilon$ (red).

called the *Mayer function* after Josef Mayer. We may now write

$$Q_N(T,V) = \frac{1}{N!}\int d^dx_1 \cdots \int d^dx_N \prod_{i<j}\left(1 + f_{ij}\right) \quad . \tag{6.73}$$

A typical potential we might consider is the semi-phenomenological Lennard-Jones potential,

$$u(r) = 4\,\epsilon\left\{\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right\} \quad . \tag{6.74}$$

This accounts for a long-distance attraction due to mutually induced electric dipole fluctuations, and a strong short-ranged repulsion, phenomenologically modelled with a $r^{-12}$ potential, which mimics a hard core due to overlap of the atomic electron distributions. Setting $u'(r) = 0$ we obtain $r^* = 2^{1/6}\,\sigma \approx 1.12246\,\sigma$ at the minimum, where $u(r^*) = -\epsilon$. In contrast to the Boltzmann weight $e^{-\beta u(r)}$, the Mayer function $f(r)$ vanishes as $r \to \infty$, behaving as $f(r) \sim -\beta u(r)$. The Mayer function also depends on temperature. Sketches of $u(r)$ and $f(r)$ for the Lennard-Jones model are shown in fig. 6.6.

The Lennard-Jones potential[14] is realistic for certain simple fluids, but it leads to a configuration integral which is in general impossible to evaluate. Indeed, even a potential as simple as that of the hard sphere gas is intractable in more than one space dimension. We can however make progress by deriving a series expansion for the equation of state in powers of the particle density. This is known as the *virial*

[14]Disambiguation footnote: Take care not to confuse Philipp Lenard (Hungarian-German, cathode ray tubes, Nazi), Alfred-Marie Liénard (French, Liénard-Wiechert potentials, not a Nazi), John Lennard-Jones (British, molecular structure, also not a Nazi), and Lynyrd Skynyrd (American, "Free Bird"). I thank my colleague Oleg Shpyrko for setting me straight on this.

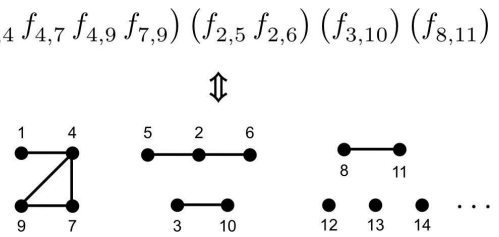



Figure 6.7: Diagrammatic interpretation of a term involving a product of eight Mayer functions.

*expansion*. As was the case when we investigated noninteracting quantum statistics, it is convenient to work in the grand canonical ensemble and to derive series expansions for the density $n(T,z)$ and the pressure $p(T,z)$ in terms of the fugacity $z$, then solve for $z(T,n)$ to obtain $p(T,n)$. These expansions in terms of fugacity have a nifty diagrammatic interpretation, due to Mayer.

We begin by expanding the product in eqn. 6.73 as

$$\prod_{i<j}\big(1+f_{ij}\big) = 1 + \sum_{i<j} f_{ij} + \sum_{\substack{i<j\ ,\ k<l \\ (ij)\neq(kl)}} f_{ij}\, f_{kl} + \ldots \quad . \tag{6.75}$$

As there are $\frac{1}{2}N(N-1)$ possible pairings, there are $2^{N(N-1)/2}$ terms in the expansion of the above product. Each such term may be represented by a graph, as shown in fig. 6.7. For each such term, we draw a connection between dots representing different particles $i$ and $j$ if the factor $f_{ij}$ appears in the term under consideration. The contribution for any given graph may be written as a product over contributions from each of its disconnected component clusters. For example, in the case of the term in fig. 6.7, the contribution to the configurational integral would be

$$\begin{aligned}\Delta Q = \frac{V^{N-11}}{N!} \int d^dx_1\, d^dx_4\, d^dx_7\, d^dx_9\; f_{1,4}\, f_{4,7}\, f_{4,9}\, f_{7,9} \\ \times \int d^dx_2\, d^dx_5\, d^dx_6\; f_{2,5}\, f_{2,6} \times \int d^dx_3\, d^dx_{10}\; f_{3,10} \times \int d^dx_8\, d^dx_{11}\; f_{8,11} \quad .\end{aligned} \tag{6.76}$$

We will refer to a given product of Mayer functions which arises from this expansion as a *term*.

The particular labels we assign to each vertex of a given graph don't affect the overall value of the graph. Now a given unlabeled graph consists of a certain number of connected subgraphs. For a system with $N$ particles, we may then write

$$N = \sum_\gamma m_\gamma\, n_\gamma \quad , \tag{6.77}$$

where $\gamma$ ranges over all possible connected subgraphs, and

$$\begin{aligned} m_\gamma &= \text{number of connected subgraphs of type } \gamma \text{ in the unlabeled graph} \\ n_\gamma &= \text{number of vertices in the connected subgraph } \gamma \quad . \end{aligned}$$

Note that the single vertex $\bullet$ counts as a connected subgraph, with $n_\bullet = 1$. We now ask: how many ways are there of assigning the $N$ labels to the $N$ vertices of a given unlabeled graph? One might first think the answer is simply $N!$, however this is too big, because different assignments of the labels to the vertices

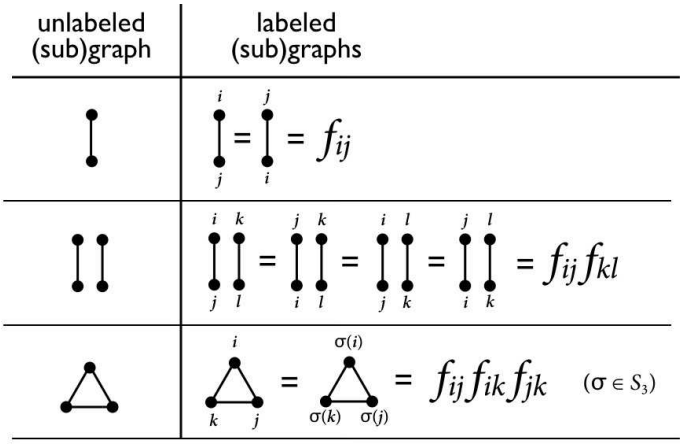



Figure 6.8: Different assignments of labels to vertices may not result in a distinct term in the expansion of the configuration integral.

may not result in a distinct graph. To see this, consider the examples in fig. 6.8. In the first example, an unlabeled graph with four vertices consists of two identical connected subgraphs. Given any assignment of labels to the vertices, then, we can simply exchange the two subgraphs and get the same term. So we should divide $N!$ by the product $\prod_\gamma m_\gamma!$. But even this is not enough, because within each connected subgraph $\gamma$ there may be permutations which leave the integrand unchanged, as shown in the second and third examples in fig. 6.8. We define the *symmetry factor* $s_\gamma$ as the number of permutations of the labels which leaves a given connected subgraphs $\gamma$ invariant. Examples of symmetry factors are shown in fig. 6.9. Consider, for example, the third subgraph in the top row. Clearly one can rotate the figure about its horizontal symmetry axis to obtain a new labeling which represents the same term. This twofold axis is the only symmetry the diagram possesses, hence $s_\gamma = 2$. For the first diagram in the second row, one can rotate *either* of the triangles about the horizontal symmetry axis. One can also rotate the figure in the plane by $180^\circ$ so as to exchange the two triangles. Thus, there are $2 \times 2 \times 2 = 8$ symmetry operations which result in the same term, and $s_\gamma = 8$. Finally, the last subgraph in the second row consists of five vertices each of which is connected to the other four. Therefore any permutation of the labels results in the same term, and $s_\gamma = 5! = 120$. In addition to dividing by the product $\prod_\gamma m_\gamma!$, we must then also divide by $\prod_\gamma s_\gamma^{m_\gamma}$.

We can now write the partition function as

$$\begin{aligned} Z &= \frac{\lambda_T^{-Nd}}{N!} \sum_{\{m_\gamma\}} \frac{N!}{\prod m_\gamma!\, s_\gamma^{m_\gamma}} \cdot \prod_\gamma \left( \int d^d x_1 \cdots d^d x_{n_\gamma} \prod_{i<j}^{\gamma} f_{ij} \right)^{m_\gamma} \delta_{N\,,\sum m_\gamma n_\gamma} \\ &= \sum_{\{m_\gamma\}} \prod_\gamma \frac{1}{m_\gamma!} \left( \frac{V b_\gamma(T)}{\lambda_T^d} \right)^{m_\gamma} \delta_{N\,,\sum m_\gamma n_\gamma} \end{aligned} \tag{6.78}$$

where the product $\prod_{i<j}^{\gamma} f_{ij}$ is over all links in the subgraph $\gamma$. The final Kronecker delta enforces the

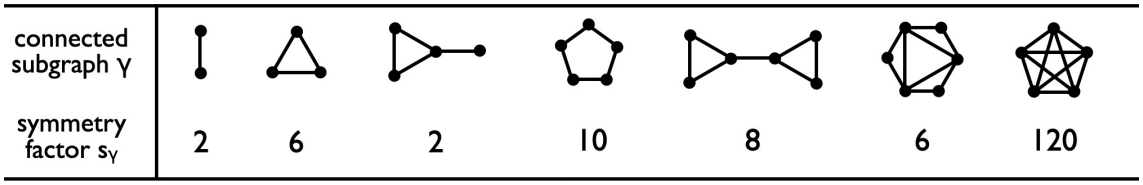



Figure 6.9: The symmetry factor $s_\gamma$ for a connected subgraph $\gamma$ is the number of permutations of its indices which leaves the term $\prod_{(ij)\in\gamma} f_{ij}$ invariant.

constraint $N = \sum_\gamma m_\gamma\, n_\gamma$. We have defined the *dimensionless cluster integrals* $b_\gamma$ as

$$b_\gamma(T) \equiv \frac{1}{s_\gamma}\int\frac{d^dx_1}{\lambda_T^d}\cdots\int\frac{d^dx_{n_\gamma-1}}{\lambda_T^d}\prod_{i<j}^{\gamma} f_{ij} \quad , \tag{6.79}$$

where we assume the limit $V\to\infty$. Since $f_{ij} = f\big(|\boldsymbol{x}_i - \boldsymbol{x}_j|\big)$, the product $\prod_{i<j}^{\gamma} f_{ij}$ is invariant under simultaneous translation of all the coordinate vectors by any constant vector, and hence the integral over the $n_\gamma$ position variables contains exactly one factor of the volume, which yields factor of $V$ within the round brackets in the second line of eqn. 6.78. Thus, each cluster integral is *intensive*[15], scaling as $V^0$.

If we compute the grand partition function, then the fixed $N$ constraint is relaxed, and we can do the sums:

$$\begin{aligned}\Xi(T,V,\mu) &= \sum_{\{m_\gamma\}}\big(e^{\beta\mu}\big)^{\sum m_\gamma n_\gamma}\prod_\gamma\frac{1}{m_\gamma!}\left(\frac{Vb_\gamma(T)}{\lambda_T^d}\right)^{m_\gamma}\\ &= \prod_\gamma\sum_{m_\gamma=0}^{\infty}\frac{1}{m_\gamma!}\left(\frac{Vz^{n_\gamma}\,b_\gamma(T)}{\lambda_T^d}\right)^{m_\gamma} = \exp\left(V\lambda_T^{-d}\sum_\gamma z^{n_\gamma}\,b_\gamma\right) \quad ,\end{aligned} \tag{6.80}$$

where $z = \exp(\beta\mu)$ is the fugacity. Thus, since $\Omega = -k_{\rm B}T\log\Xi$,

$$\Omega(T,V,\mu) = -\frac{Vk_{\rm B}T}{\lambda_T^d}\sum_\gamma z^{n_\gamma}\,b_\gamma(T) \quad , \tag{6.81}$$

and we can write

$$\begin{aligned}p &= k_{\rm B}T\,\lambda_T^{-d}\sum_\gamma z^{n_\gamma}\,b_\gamma(T)\\ n &= \lambda_T^{-d}\sum_\gamma n_\gamma\,z^{n_\gamma}\,b_\gamma(T) \quad ,\end{aligned} \tag{6.82}$$

where $b_\bullet \equiv 1$. As in the case of ideal quantum gas statistical mechanics, we can systematically invert the relation $n = n(z,T)$ to obtain $z = z(n,T)$, and then insert this into the equation for $p(z,T)$ to obtain the equation of state $p = p(n,T)$. This yields the *virial expansion* of the equation of state,

$$p = nk_{\rm B}T\Big\{1 + B_2(T)\,n + B_3(T)\,n^2 + \ldots\Big\} \quad . \tag{6.83}$$

[15]We assume that the long-ranged behavior of $f(r)\approx -\beta u(r)$ is integrable.

It is useful to define the dimensionless quantities $\nu \equiv n\lambda_T^d$ and $\pi = p\lambda_T^d/k_{\rm B}T$, as well as the dimensionless cluster integral sums

$$\mathsf{b}_k \equiv \sum_\gamma b_\gamma\,\delta_{k,n_\gamma} \quad , \tag{6.84}$$

which is the sum of all cluster integrals $b_\gamma$ with $n_\gamma = k$ vertices. Then

$$\nu(z) = \sum_{k=1}^\infty k\,\mathsf{b}_k\,z^k \qquad , \qquad \pi(z) = \sum_{k=1}^\infty \mathsf{b}_k\,z^k \quad . \tag{6.85}$$

The virial expansion of the dimensionless equation of state is then

$$\pi(\nu) = \sum_{k=1}^\infty \mathsf{B}_k\,\nu^k \tag{6.86}$$

We may again apply the Lagrange method introduced in §5.3.2 for the quantum virial coefficients, writing

$$\mathsf{B}_k = \oint \frac{d\nu}{2\pi i}\,\frac{\pi(\nu)}{\nu^{k+1}} = \oint \frac{dz}{2\pi i}\,\frac{\nu'(z)\,\pi(z)}{\big[\nu(z)\big]^{k+1}} = -\frac{1}{k}\oint \frac{dz}{2\pi i}\,\pi(z)\,\frac{d}{dz}\big[\nu(z)\big]^{-k} \quad , \tag{6.87}$$

where the contour encloses the origin in the complex plane. Integrating by parts, and using the relation $\pi'(z) = \nu(z)/z$, we obtain[16]

$$\begin{aligned}\mathsf{B}_k &= \frac{1}{k}\oint \frac{dz}{2\pi i}\,\pi'(z)\,\big[\nu(z)\big]^{-k} = \frac{1}{k}\oint \frac{dz}{2\pi i}\,\frac{1}{z}\,\big[\nu(z)\big]^{1-k} \\ &= \frac{1}{k}\oint \frac{dz}{2\pi i}\,\frac{1}{z^k}\,\Big(1 + 2\mathsf{b}_2 z + 3\mathsf{b}_3 z^2 + 4\mathsf{b}_4 z^4 + \ldots\Big)^{1-k} \quad ,\end{aligned} \tag{6.88}$$

where the contour is a small circle enclosing the origin. Working out the first two virial coefficients, we find

$$\mathsf{B}_2 = -\mathsf{b}_2 \qquad , \qquad \mathsf{B}_3 = 4\mathsf{b}_2^2 - 2\mathsf{b}_3 \quad . \tag{6.89}$$

The dimensionful virial coefficients in eqn. 6.83 are then given by $B_k = \mathsf{B}_k\,\lambda_T^{(k-1)d}$.

**Lowest order expansion**

We have

$$b_-(T) = \tfrac{1}{2}\int \frac{d^dx_1}{\lambda_T^d}\,f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\big) = \tfrac{1}{2}\int \frac{d^dr}{\lambda_T^d}\,f(r) \tag{6.90}$$

and

$$\begin{aligned}b_\wedge(T) &= \tfrac{1}{2}\int \frac{d^dx_1}{\lambda_T^d}\int \frac{d^dx_2}{\lambda_T^d}\,f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\big)\,f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_3|\big) \\ &= \tfrac{1}{2}\int \frac{d^dr}{\lambda_T^d}\int \frac{d^dr'}{\lambda_T^d}\,f(r)\,f(r') = 2\big(b_-\big)^2\end{aligned} \tag{6.91}$$

[16]Since there is no term proportional to $\log w$ in the Laurent expansion of $\pi(w)\big[n(w)\big]^{-k}$, there is no residue arising from integrating its derivative around the unit circle.

and

$$\begin{aligned} b_\triangle(T) &= \tfrac{1}{6}\int\frac{d^dx_1}{\lambda_T^d}\int\frac{d^dx_2}{\lambda_T^d}\,f\big(|\boldsymbol{x}_1-\boldsymbol{x}_2|\big)\,f\big(|\boldsymbol{x}_1-\boldsymbol{x}_3|\big)\,f\big(|\boldsymbol{x}_2-\boldsymbol{x}_3|\big) \\ &= \tfrac{1}{6}\int\frac{d^dr}{\lambda_T^d}\int\frac{d^dr'}{\lambda_T^d}\,f(r)\,f(r')\,f\big(|\boldsymbol{r}-\boldsymbol{r}'|\big) \quad . \end{aligned} \tag{6.92}$$

Thus we have $\mathsf{b}_2 = b_-$ and $\mathsf{b}_3 = b_\wedge + b_\triangle = 2b_-^2 + b_\triangle$. From eqn. 6.89 we now have

$$\begin{aligned} \mathsf{B}_2(T) &= -\mathsf{b}_2(T) = -b_-(T) \\ \mathsf{B}_3(T) &= \big[4\mathsf{b}_2^2(T) - 2\mathsf{b}_3(T)\big] = -2b_\triangle(T) \quad . \end{aligned} \tag{6.93}$$

Note that $b_\wedge$ does not contribute to $\mathsf{B}_3$, even though the graph $\wedge$ has three vertices, and only $b_\triangle$ appears. This is because the virial coefficients $\mathsf{B}_j$ involve only cluster integrals $b_\gamma$ for *one-particle irreducible* clusters, *i.e.* those clusters which remain connected and don't fall into multiple pieces if any of its vertices is removed, as depicted in fig. 6.10.

### Cookbook recipe

Just follow these simple steps:

- The pressure and number density are written as sums over unlabeled connected clusters $\gamma$, *viz.*

$$\begin{aligned} p &= k_{\rm B}T\,\lambda_T^{-d}\sum_\gamma z^{n_\gamma}\,b_\gamma(T) \\ n &= \lambda_T^{-d}\sum_\gamma n_\gamma\,z^{n_\gamma}\,b_\gamma(T) \quad , \end{aligned} \tag{6.94}$$

  where $z = \exp(\beta\mu)$ is the fugacity.

- To compute the dimensionless cluster integral $b_\gamma(T)$, first draw the connected cluster $\gamma$ with *unlabeled* vertices.

- Next, assign labels $1, 2, \ldots, n_\gamma$ to the vertices, where $n_\gamma$ is the total number of vertices in the cluster $\gamma$. It doesn't matter how you assign the labels.

- Write down the product $\prod_{i<j}^\gamma f_{ij}$. The factor $f_{ij}$ appears in the product if there is a link in your (now labeled) cluster between sites $i$ and $j$.

- The symmetry factor $s_\gamma$ is the number of elements of the symmetric group $S_{n_\gamma}$ which leave the product $\prod_{i<j}^\gamma f_{ij}$ invariant. The identity permutation leaves the product invariant, so $s_\gamma \geq 1$.

- The dimensionless cluster integral $b_\gamma(T)$ is given by

$$b_\gamma(T) \equiv \frac{1}{s_\gamma}\int\frac{d^dx_1}{\lambda_T^d}\cdots\int\frac{d^dx_{n_\gamma-1}}{\lambda_T^d}\prod_{i<j}^\gamma f_{ij} \quad , \tag{6.95}$$

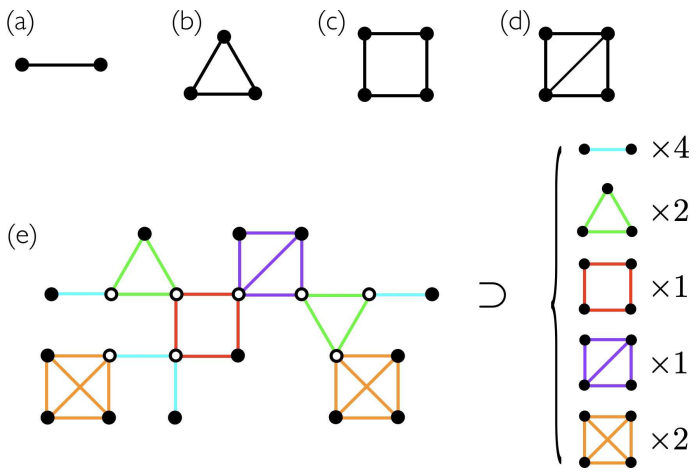



Figure 6.10: Connected *versus* irreducible clusters. Clusters (a) through (d) are *irreducible* in that they remain connected if any component site and its connecting links are removed. Cluster (e) is connected, but is *reducible*. Its integral $b_\gamma$ is proportional to a product over its irreducible components, each shown in a unique color, and occurring with various multiplicities. The open circles denote *articulation points*. Removal of an articulation point and all the links connected to it results in a disconnected diagram. Removal of any of the closed circles and its associated links does not result in a disconnected diagram.

Due to translation invariance, $b_\gamma(T) \propto V^0$. One can therefore set $\boldsymbol{x}_{n_\gamma} \equiv 0$, eliminate the volume factor from the denominator, and perform the integral over the remaining $n_\gamma - 1$ coordinates.

- This procedure generates expansions for $p(T,z)$ and $n(T,z)$ in powers of the fugacity $z = \exp(\beta\mu)$. To obtain something useful like $p(T,n)$, we must invert the equation $n = n(T,z)$ to find $z = z(T,n)$, and then substitute into the equation $p = p(T,z)$ to obtain $p = p\big(T, z(T,n)\big) = p(T,n)$. The result is the virial expansion,

$$p = nk_{\rm B}T\left\{1 + B_2(T)\,n + B_3(T)\,n^2 + \ldots\right\} \quad , \tag{6.96}$$

where

$$B_k(T) = -(k-1)\,\lambda_T^{(k-1)d} \sum_{\gamma\in\Gamma_k} b_\gamma(T) \quad , \tag{6.97}$$

with $\Gamma_k$ the set of all one-particle irreducible (1PI) $k$-site clusters. A 1PI cluster remains connected if any of its sites and all that site's connecting links are removed.

### 6.3.4 Examples

#### Hard sphere gas in three dimensions

The hard sphere potential is given by

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ 0 & \text{if } r > a \end{cases} \quad . \tag{6.98}$$

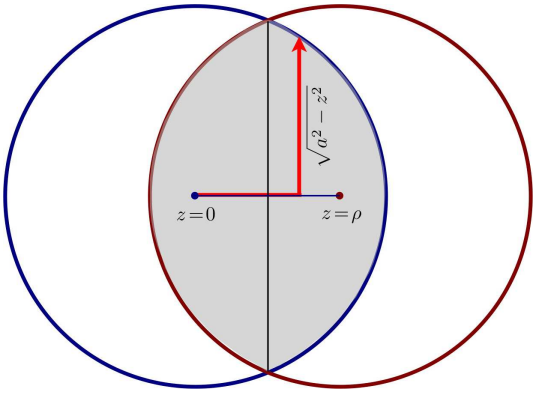



Figure 6.11: The overlap of hard sphere Mayer functions. The shaded volume is $\mathcal{V}$.

Here $a$ is the *diameter* of the spheres. The corresponding Mayer function is then temperature independent, and given by

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ 0 & \text{if } r > a \end{cases} \quad . \tag{6.99}$$

We can change variables to obtain

$$\mathsf{b}_2(T) = \tfrac{1}{2}\int \frac{d^3r}{\lambda_T^3}\, f(r) = -\tfrac{2}{3}\pi a^3 \lambda_T^{-3} \quad . \tag{6.100}$$

The calculation of $\mathsf{b}_3$ is more challenging. We have

$$\mathsf{b}_3 = \tfrac{1}{6}\int \frac{d^3\rho}{\lambda_T^3}\int \frac{d^3r}{\lambda_T^3}\, f(\rho)\, f(r)\, f\big(|\boldsymbol{r}-\boldsymbol{\rho}|\big) \quad . \tag{6.101}$$

We must first compute the volume of overlap for spheres of *radius* $a$ (recall $a$ is the *diameter* of the constituent hard sphere particles) centered at 0 and at $\boldsymbol{\rho}$:

$$\begin{aligned} \mathcal{V} &= \int d^3r\, f(r)\, f\big(|\boldsymbol{r}-\boldsymbol{\rho}|\big) \\ &= 2\int_{\rho/2}^{a} dz\, \pi(a^2-z^2) = \tfrac{4\pi}{3}a^3 - \pi a^2\rho + \tfrac{\pi}{12}\,\rho^3 \quad . \end{aligned} \tag{6.102}$$

We then integrate over region $|\boldsymbol{\rho}| < a$, to obtain

$$\mathsf{b}_3 = -\tfrac{1}{6}\cdot 4\pi\lambda_T^{-6}\int_0^a d\rho\,\rho^2\cdot\Big\{\tfrac{4\pi}{3}a^3 - \pi a^2\rho + \tfrac{\pi}{12}\,\rho^3\Big\} = -\tfrac{5\pi^2}{36}\,a^6\lambda_T^{-6} \quad . \tag{6.103}$$

Thus, we have

$$B_2 = -\lambda_T^3\,\mathsf{b}_2(T) = \tfrac{2\pi}{3}a^3 = 4\,v_0 \qquad , \qquad B_3 = -2\lambda_T^6\,\mathsf{b}_3(T) = \tfrac{5\pi^2}{18}a^6 = 10\,v_0^2 \quad , \tag{6.104}$$

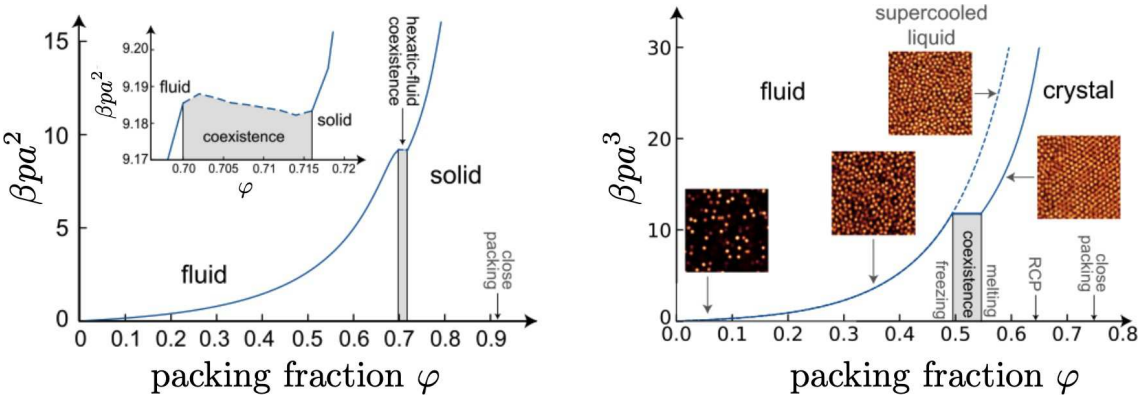



Figure 6.12: Phase diagram of the hard disk (left) and hard sphere gas (right), from G. P. Royall *et al.*, *Rev. Mod. Phys.* **96**, 045003 (2024).

where $v_0 = \pi a^3/6$ is the hard sphere volume. The equation of state is then

$$p = nk_{\rm B}T\left\{1 + B_2\, n + B_3\, n^2 + \mathcal{O}(n^3)\right\} \quad . \tag{6.105}$$

$B_4$ for the three-dimensional hard sphere gas was calculated by van Laar[17] and is given by

$$B_4 = \left(\tfrac{2707}{70} + \tfrac{219\sqrt{2}}{35\pi} - \tfrac{4131}{70\pi}\cos^{-1}(1/3)\right)v_0^3 \quad . \tag{6.106}$$

The coefficients beyond $B_4$ require numerical evaluation. The *packing fraction* of hard spheres is given by $\varphi = nv_0$. There is a first order, entropy-driven phase transition as a function of pressure at $\beta p_{\rm c} a^3 = 11.6$, *i.e.* $p_{\rm c}(T) = 11.6\, k_{\rm B}T a^{-3}$. The coexistence region (see fig. 6.12) lies between the packing fraction for freezing, $\varphi_{\rm f} = 0.492$, and the packing fraction for melting, $\varphi_{\rm m} = 0.543$. The critical volume fraction for dense random packing is $\varphi_{\rm drp} \approx 0.64$, and the maximum possible packing occurs for crystalline close packing (face-centered cubic or hexagonal close-packed structures) at $\varphi_{\rm ccp} = \frac{\pi}{3\sqrt{2}} \approx 0.740$.

For the hard disk gas in $d = 2$, one has $v_0 = \frac{1}{4}\pi a^2$ where $a$ again is the diameter, and

$$B_2 = 2\, v_0 \qquad , \qquad B_3 = \left(\tfrac{16}{3} - \tfrac{4\sqrt{3}}{\pi}\right)v_0^2 \qquad , \qquad B_4 = \left(16 - \tfrac{36\sqrt{3}}{\pi} + \tfrac{80}{\pi^2}\right)v_0^3 \quad . \tag{6.107}$$

**Weakly attractive tail**

Suppose

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ -u_0(r) & \text{if } r > a \end{cases} \quad . \tag{6.108}$$

[17] J. J. Van Laar, Koninklijke Nederlandse Akademie van Wetenschappen Amsterdam *Prov. Sec. Sci.* **1**, 273 (1899).

Then the corresponding Mayer function is

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ e^{\beta u_0(r)} - 1 & \text{if } r > a \quad . \end{cases} \tag{6.109}$$

Thus,

$$\mathsf{b}_2(T) = \tfrac{1}{2}\int \frac{d^3r}{\lambda_T^3}\, f(r) = -\tfrac{2\pi}{3}a^3\lambda_T^{-3} + 2\pi\lambda_T^{-3}\int\limits_a^\infty dr\, r^2 \left[e^{\beta u_0(r)} - 1\right] \quad . \tag{6.110}$$

Thus, the second virial coefficient is

$$B_2(T) = -\lambda_T^3\, \mathsf{b}_2(T) \approx \tfrac{2\pi}{3}a^3 - \frac{2\pi}{k_{\rm B}T}\int\limits_a^\infty dr\, r^2\, u_0(r) \quad , \tag{6.111}$$

where we have assumed $|u_0(r)| \ll k_{\rm B}T$. We see that the second virial coefficient *changes sign* at some temperature $T_0$, from a negative low temperature value to a positive high temperature value.

**Spherical potential well**

Consider an attractive spherical well potential with an infinitely repulsive core,

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ -\epsilon & \text{if } a < r < R \\ 0 & \text{if } r > R \quad . \end{cases} \tag{6.112}$$

Then the corresponding Mayer function is

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ e^{\beta\epsilon} - 1 & \text{if } a < r < R \\ 0 & \text{if } r > R \quad . \end{cases} \tag{6.113}$$

Writing $s \equiv R/a$, we have

$$\begin{aligned} B_2(T) = -\lambda_T^3\, \mathsf{b}_2(T) &= -\tfrac{1}{2}\int d^3r\, f(r) \\ &= -\frac{1}{2}\left\{(-1)\cdot\tfrac{4\pi}{3}a^3 + \big(e^{\beta\epsilon}-1\big)\cdot\tfrac{4\pi}{3}a^3(s^3-1)\right\} \\ &= \tfrac{2\pi}{3}a^3\left\{1 - (s^3-1)\big(e^{\beta\epsilon}-1\big)\right\} \quad . \end{aligned} \tag{6.114}$$

To find the temperature $T_0$ where $B_2(T)$ changes sign, we set $B_2(T_0) = 0$ and obtain

$$k_{\rm B}T_0 = \epsilon \Big/ \log\left(\frac{s^3}{s^3-1}\right) \quad . \tag{6.115}$$

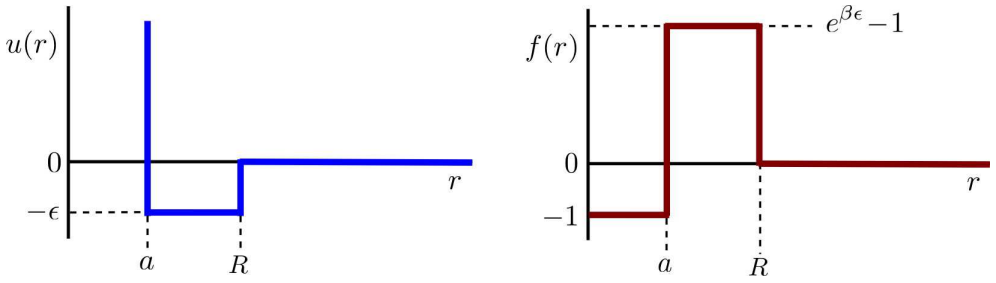



Figure 6.13: An attractive spherical well with a repulsive core $u(r)$ and its associated Mayer function $f(r)$.

Recall in our study of the thermodynamics of the Joule-Thomson effect in §2.10.7 that the throttling process is *isenthalpic*. The temperature change, when a gas is pushed (or escapes) through a porous plug from a high pressure region to a low pressure one is

$$\Delta T = \int_{p_1}^{p_2} dp \left(\frac{\partial T}{\partial p}\right)_H \quad , \tag{6.116}$$

where

$$\left(\frac{\partial T}{\partial p}\right)_H = \frac{1}{C_p}\left[T\left(\frac{\partial V}{\partial T}\right)_p - V\right] \quad . \tag{6.117}$$

Appealing to the virial expansion, and working to lowest order in corrections to the ideal gas law, we have

$$p = \frac{N}{V}\,k_{\rm B}T + \frac{N^2}{V^2}\,k_{\rm B}T\,B_2(T) + \ldots \tag{6.118}$$

and we compute $\left(\frac{\partial V}{\partial T}\right)_p$ by setting

$$0 = dp = -\frac{Nk_{\rm B}T}{V^2}\,dV + \frac{Nk_{\rm B}}{V}\,dT - \frac{2N^2}{V^3}\,k_{\rm B}T\,B_2(T)\,dV + \frac{N^2}{V^2}\,d\big(k_{\rm B}T\,B_2(T)\big) + \ldots \quad . \tag{6.119}$$

Dividing by $dT$, we find

$$T\left(\frac{\partial V}{\partial T}\right)_p - V = N\left[T\,\frac{\partial B_2}{\partial T} - B_2\right] \quad . \tag{6.120}$$

The temperature where $\left(\frac{\partial T}{\partial p}\right)_H$ changes sign is called the *inversion temperature* $T^*$. To find the inversion point, we set $T^*\,B_2'(T^*) = B_2(T^*)$, *i.e.*

$$\left.\frac{d\log B_2}{d\log T}\right|_{T^*} = 1 \quad . \tag{6.121}$$

If we approximate $B_2(T) \approx A - \frac{B}{T}$, then the inversion temperature follows simply:

$$\frac{B}{T^*} = A - \frac{B}{T^*} \qquad \Longrightarrow \qquad T^* = \frac{2B}{A} \quad . \tag{6.122}$$

**Hard spheres with a hard wall**

Consider a hard sphere gas in three dimensions in the presence of a hard wall at $z = 0$. The gas is confined to the region $z > 0$. The total potential energy is now

$$W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) = \sum_i v(\boldsymbol{x}_i) + \sum_{i<j} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \quad , \tag{6.123}$$

where

$$v(\boldsymbol{r}) = v(z) = \begin{cases} \infty & \text{if} \quad z \le \frac{1}{2}a \\ 0 & \text{if} \quad z > \frac{1}{2}a \end{cases} \quad , \tag{6.124}$$

and $u(\boldsymbol{r})$ is given in eqn. 6.98. The grand potential is written as a series in the total particle number $N$, and is given by

$$\Xi = e^{-\beta\Omega} = 1 + \xi \int d^3r\, e^{-\beta v(z)} + \tfrac{1}{2}\xi^2 \int d^3r \int d^3r'\, e^{-\beta v(z)}\, e^{-\beta v(z')}\, e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')} + \ldots \quad , \tag{6.125}$$

where $\xi = z\,\lambda_T^{-3}$, with $z = e^{\mu/k_{\rm B}T}$ the fugacity. Taking the logarithm, and invoking the Taylor series $\log(1+\delta) = \delta - \frac{1}{2}\delta^2 + \frac{1}{3}\delta^3 - \ldots$, we obtain

$$-\beta\Omega = \xi \int\limits_{z>\frac{a}{2}} d^3r + \tfrac{1}{2}\xi^2 \int\limits_{z>\frac{a}{2}} d^3r \int\limits_{z'>\frac{a}{2}} d^3r' \left[ e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')} - 1 \right] + \ldots \tag{6.126}$$

The volume is $V = \int\limits_{z>0} d^3r$. Dividing by $V$, we have, in the thermodynamic limit,

$$\begin{aligned} -\frac{\beta\Omega}{V} = \beta p &= \xi + \tfrac{1}{2}\xi^2 \frac{1}{V} \int\limits_{z>\frac{a}{2}} d^3r \int\limits_{z'>\frac{a}{2}} d^3r' \left[ e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')} - 1 \right] + \ldots \\ &= \xi - \tfrac{2}{3}\pi a^3\,\xi^2 + \mathcal{O}(\xi^3) \quad . \end{aligned} \tag{6.127}$$

The number density is

$$n = \xi\,\frac{\partial}{\partial\xi}\,(\beta p) = \xi - \tfrac{4}{3}\pi a^3\,\xi^2 + \mathcal{O}(\xi^3) \quad , \tag{6.128}$$

and inverting to obtain $\xi(n)$ and then substituting into the pressure equation, we obtain the lowest order virial expansion for the equation of state,

$$p = k_{\rm B}T \left\{ n + \tfrac{2}{3}\pi a^3\, n^2 + \ldots \right\} \quad . \tag{6.129}$$

As expected, the presence of the wall does not affect a bulk property such as the equation of state.

Next, let us compute the number density $n(z)$, given by

$$n(z) = \Big\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i) \Big\rangle \quad . \tag{6.130}$$

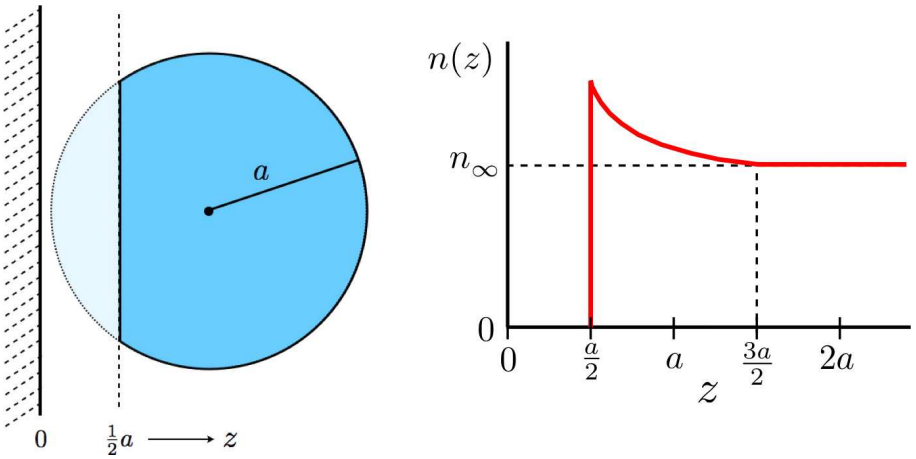



Figure 6.14: In the presence of a hard wall, the Mayer sphere is cut off on the side closest to the wall. The resulting density $n(z)$ vanishes for $z < \frac{1}{2}a$ since the center of each sphere must be at least one radius ($\frac{1}{2}a$) away from the wall. Between $z = \frac{1}{2}a$ and $z = \frac{3}{2}a$ there is a density *enhancement*. If the calculation were carried out to higher order, $n(z)$ would exhibit damped spatial oscillations with wavelength $\lambda \sim a$.

Due to translational invariance in the $(x, y)$ plane, we know that the density must be a function of $z$ alone. The presence of the wall at $z = 0$ breaks translational symmetry in the $z$ direction. The number density is

$$\begin{aligned} n(z) &= \mathrm{Tr}\left[e^{\beta(\mu\hat{N}-\hat{H})}\sum_{i=1}^{N}\delta(\boldsymbol{r}-\boldsymbol{r}_i)\right]\Big/\mathrm{Tr}\,e^{\beta(\mu\hat{N}-\hat{H})} \\ &= \Xi^{-1}\left\{\xi\,e^{-\beta v(z)} + \xi^2\,e^{-\beta v(z)}\int d^3r'\,e^{-\beta v(z')}\,e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')} + \ldots\right\} \\ &= \xi\,e^{-\beta v(z)} + \xi^2\,e^{-\beta v(z)}\int d^3r'\,e^{-\beta v(z')}\left[e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')}-1\right] + \ldots \quad . \end{aligned} \tag{6.131}$$

Note that the term in square brackets in the last line is the Mayer function $f(\boldsymbol{r}-\boldsymbol{r}') = e^{-\beta u(\boldsymbol{r}-\boldsymbol{r}')} - 1$. Consider the function

$$e^{-\beta v(z)}\,e^{-\beta v(z')}\,f(\boldsymbol{r}-\boldsymbol{r}') = \begin{cases} 0 & \text{if } z < \frac{1}{2}a \text{ or } z' < \frac{1}{2}a \\ 0 & \text{if } |\boldsymbol{r}-\boldsymbol{r}'| > a \\ -1 & \text{if } z > \frac{1}{2}a \text{ and } z' > \frac{1}{2}a \text{ and } |\boldsymbol{r}-\boldsymbol{r}'| < a \quad . \end{cases} \tag{6.132}$$

Now consider the integral of the above function with respect to $\boldsymbol{r}'$. Clearly the result depends on the value of $z$. If $z > \frac{3}{2}a$, then there is no excluded region in $\boldsymbol{r}'$ and the integral is $(-1)$ times the full Mayer sphere volume, *i.e.* $-\frac{4}{3}\pi a^3$. If $z < \frac{1}{2}a$ the integral vanishes due to the $e^{-\beta v(z)}$ factor. For $z$ infinitesimally larger than $\frac{1}{2}a$, the integral is $(-1)$ times half the Mayer sphere volume, *i.e.* $-\frac{2}{3}\pi a^3$. For $z \in \left[\frac{a}{2}\,,\,\frac{3a}{2}\right]$ the integral interpolates between $-\frac{2}{3}\pi a^3$ and $-\frac{4}{3}\pi a^3$. Explicitly, one finds by elementary integration,

$$\int d^3r'\,e^{-\beta v(z)}\,e^{-\beta v(z')}\,f(\boldsymbol{r}-\boldsymbol{r}') = \begin{cases} 0 & \text{if } z < \frac{1}{2}a \\ \left[-1-\frac{3}{2}\left(\frac{z}{a}-\frac{1}{2}\right)+\frac{1}{2}\left(\frac{z}{a}-\frac{1}{2}\right)^3\right]\cdot\frac{2}{3}\pi a^3 & \text{if } \frac{1}{2}a < z < \frac{3}{2}a \\ -\frac{4}{3}\pi a^3 & \text{if } z > \frac{3}{2}a \quad . \end{cases} \tag{6.133}$$

After substituting $\xi = n + \frac{4}{3}\pi a^3 n^2 + \mathcal{O}(n^3)$ to relate $\xi$ to the bulk density $n = n_\infty$, we obtain the desired result:

$$n(z) = \begin{cases} 0 & \text{if } z < \frac{1}{2}a \\ n + \left[1 - \frac{3}{2}\left(\frac{z}{a} - \frac{1}{2}\right) + \frac{1}{2}\left(\frac{z}{a} - \frac{1}{2}\right)^3\right] \cdot \frac{2}{3}\pi a^3\, n^2 & \text{if } \frac{1}{2}a < z < \frac{3}{2}a \\ n & \text{if } z > \frac{3}{2}a \quad . \end{cases} \tag{6.134}$$

A sketch is provided in the right hand panel of fig. 6.14. Note that the density $n(z)$ vanishes identically for $z < \frac{1}{2}a$ due to the exclusion of the hard spheres by the wall. For $z$ between $\frac{1}{2}a$ and $\frac{3}{2}a$, there is a density *enhancement*, the origin of which has a simple physical interpretation. Since the wall excludes particles from the region $z < \frac{1}{2}a$, there is an empty slab of thickness $\frac{1}{2}a$ coating the interior of the wall. There are then no particles in this region to exclude neighbors to their right, hence the density builds up just on the other side of this slab. The effect vanishes to the order of the calculation past $z = \frac{3}{2}a$, where $n(z) = n$ returns to its bulk value. Had we calculated to higher order, we'd have found damped oscillations with spatial period $\lambda \sim a$.

## 6.4 Liquid State Physics

### 6.4.1 The many-particle distribution function

The virial expansion is typically applied to low-density systems. When the density is high, *i.e.* when $na^3 \sim 1$, where $a$ is a typical molecular or atomic length scale, the virial expansion is impractical. There are too many terms to compute, and to make progress one must use sophisticated resummation techniques to investigate the high density regime.

To elucidate the physics of liquids, it is useful to consider the properties of various *correlation functions*. These objects are derived from the general $N$-body Boltzmann distribution for identical particles,

$$\varrho_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N) = \frac{1}{Z(T,V,N)}\,\frac{1}{N!}\,e^{-\beta \hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} \quad , \tag{6.135}$$

here given in the ordinary canonical ensemble (OCE), where

$$Z(T,V,N) = \frac{1}{N!}\,\mathsf{Tr}\; e^{-\beta \hat{H}_N} = \frac{1}{N!}\int \prod_{j=1}^{N} \frac{d^d x_j\, d^d p_j}{h^d}\, e^{-\beta \hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} \tag{6.136}$$

We assume a Hamiltonian of the form

$$\hat{H}_N = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} + W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) \quad . \tag{6.137}$$

The quantity

$$\varrho_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N)\,\frac{d^d x_1\, d^d p_1}{h^d} \cdots \frac{d^d x_N\, d^d p_N}{h^d} \tag{6.138}$$

is the probability of finding particle #1 lying within $d^3x_1$ of $\boldsymbol{x}_1$ and having momentum within $d^dp_1$ of $\boldsymbol{p}_1$, *etc.* Note $\mathsf{Tr}\ \varrho_N = 1$. Note that the definition of the trace ($\mathsf{Tr}$) includes a factor $1/N!$ in order to account for particle indistinguishability, and that $\varrho_N$ is normalized according to

$$\int \prod_{j=1}^{N} \frac{d^dx_j\, d^dp_j}{h^d}\, \varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N) = 1 \quad . \tag{6.139}$$

Thus, the canonical average of any function $\Phi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_n)$ is given by

$$\langle\,\Phi\,\rangle_{T,V,N} = \int \prod_{j=1}^{N} \frac{d^dx_j\, d^dp_j}{h^d}\, \varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N)\,\Phi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N) \quad . \tag{6.140}$$

In the grand canonical ensemble (GCE), the probability that there are $N$ particles present is

$$\mathsf{P}_N(T,V,\mu) = \Xi^{-1}\,\mathsf{Tr}\,\left(e^{-\beta\hat{K}}\,\delta_{\hat{N},N}\right) = \frac{Z(T,V,N)\,e^{N\mu/k_{\rm B}T}}{\Xi(T,V,\mu)} \quad , \tag{6.141}$$

with $\hat{K} = \hat{H} - \mu\hat{N}$ and

$$\Xi(T,V,\mu) = \mathsf{Tr}\ e^{-\beta\hat{K}} = \sum_{N=0}^{\infty} Z(T,V,N)\,e^{\beta\mu N} = \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N!} \int \prod_{j=1}^{N} \frac{d^dx_j\, d^dp_j}{h^d}\, e^{-\beta\hat{H}_N\left(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\}\right)} \quad . \tag{6.142}$$

The grand canonical average of $\Phi$ is given by

$$\langle\,\Phi\,\rangle_{T,V,\mu} = \sum_{N=0}^{\infty} \mathsf{P}_N(T,V,\mu)\,\langle\,\Phi\,\rangle_{T,V,N} \quad . \tag{6.143}$$

If we compute averages of quantities which only depend on the positions $\{\boldsymbol{x}_j\}$ and not on the momenta $\{\boldsymbol{p}_j\}$, then we may integrate out the momenta to obtain, in the OCE,

$$\begin{aligned} P(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) &= \int \prod_{j=1}^{N} \frac{d^dp_j}{h^d}\, \varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N) \\ &\equiv Q_N^{-1}\cdot\frac{1}{N!}\,e^{-\beta W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)} \quad , \end{aligned} \tag{6.144}$$

where $W$ is the total potential energy,

$$W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = \sum_i v(\boldsymbol{x}_i) + \sum_{i<j} u(\boldsymbol{x}_i-\boldsymbol{x}_j) + \sum_{i<j<k} w(\boldsymbol{x}_i-\boldsymbol{x}_j\,,\,\boldsymbol{x}_j-\boldsymbol{x}_k) + \ldots \quad , \tag{6.145}$$

and $Q_N$ is the configuration integral,

$$Q_N(T,V) = \frac{1}{N!}\int d^dx_1 \cdots \int d^dx_N\, e^{-\beta W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)} \quad . \tag{6.146}$$

We will, for the most part, consider only two-body central potentials as contributing to $W$, which is to say we will only retain the middle term on the RHS. Note that $P(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)$ is invariant under any permutation of the particle labels, and is normalized according to $\int \prod_{j=1}^{N} d^dx_j\, P(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = 1$.

### 6.4.2 Averages over the distribution

To compute an average, one integrates over the distribution:

$$\big\langle F(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\big\rangle = \int d^dx_1\cdots\int d^dx_N\, P(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)\, F(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) \quad . \tag{6.147}$$

The overall $N$-particle probability density is normalized, with $\int\prod_{j=1}^N d^dx_j\, P(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = 1$.

The average local density is

$$n_1(\boldsymbol{r}) = \big\langle\sum_i \delta(\boldsymbol{r}-\boldsymbol{x}_i)\big\rangle = N\int d^dx_2\cdots\int d^dx_N\, P(\boldsymbol{r},\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N) \quad . \tag{6.148}$$

Note that the local density obeys the sum rule $\int d^dr\; n_1(\boldsymbol{r}) = N$. In a translationally invariant system, $n_1 = n = \frac{N}{V}$ is a constant independent of position. The boundaries of a system will in general break translational invariance, so in order to maintain the notion of a translationally invariant system of finite total volume, one must impose periodic boundary conditions.

The two-particle density matrix $n_2(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is defined by

$$\begin{aligned} n_2(\boldsymbol{r}_1,\boldsymbol{r}_2) &= \big\langle\sum_{i\neq j}\delta(\boldsymbol{r}_1-\boldsymbol{x}_i)\,\delta(\boldsymbol{r}_2-\boldsymbol{x}_j)\big\rangle \\ &= N(N-1)\int d^dx_3\cdots\int d^dx_N\, P(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{x}_3,\ldots,\boldsymbol{x}_N) \quad . \end{aligned} \tag{6.149}$$

As in the case of the one-particle density matrix, *i.e.* the local density $n_1(\boldsymbol{r})$, the two-particle density matrix also satisfies a sum rule:

$$\int d^dr_1\int d^dr_2\, n_2(\boldsymbol{r}_1,\boldsymbol{r}_2) = N(N-1) \quad . \tag{6.150}$$

Generalizing further, one defines the *k-particle density matrix* as

$$\begin{aligned} n_k(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k) &= \big\langle{\sum_{i_1\cdots i_k}}'\delta(\boldsymbol{r}_1-\boldsymbol{x}_{i_1})\cdots\delta(\boldsymbol{r}_k-\boldsymbol{x}_{i_k})\,\big\rangle \\ &= \frac{N!}{(N-k)!}\int d^dx_{k+1}\cdots\int d^dx_N\, P(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k,\boldsymbol{x}_{k+1},\ldots,\boldsymbol{x}_N) \quad , \end{aligned} \tag{6.151}$$

where the prime on the sum indicates that all the indices $i_1,\ldots,i_k$ are distinct. The corresponding sum rule is then

$$\int d^dr_1\cdots\int d^dr_k\, n_k(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k) = \frac{N!}{(N-k)!} \quad . \tag{6.152}$$

The average potential energy can be expressed in terms of the distribution functions. Assuming only

two-body interactions, we have

$$\begin{aligned}\langle W\rangle &= \Big\langle \sum_{i<j} u(\boldsymbol{x}_i - \boldsymbol{x}_j)\Big\rangle \\ &= \tfrac{1}{2}\int d^d r_1 \int d^d r_2\, u(\boldsymbol{r}_1 - \boldsymbol{r}_2)\,\Big\langle \sum_{i\neq j} \delta(\boldsymbol{r}_1 - \boldsymbol{x}_i)\,\delta(\boldsymbol{r}_2 - \boldsymbol{x}_j)\Big\rangle \\ &= \tfrac{1}{2}\int d^d r_1 \int d^d r_2\, u(\boldsymbol{r}_1 - \boldsymbol{r}_2)\, n_2(\boldsymbol{r}_1, \boldsymbol{r}_2) \quad .\end{aligned} \tag{6.153}$$

As the separations $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$ get large, we expect the correlations to vanish, in which case

$$\begin{aligned}n_k(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k) &= \Big\langle {\sum_{i_1\cdots i_k}}' \delta(\boldsymbol{r}_1 - \boldsymbol{x}_{i_1})\cdots\delta(\boldsymbol{r}_k - \boldsymbol{x}_{i_k})\Big\rangle \xrightarrow[r_{ij}\to\infty]{} {\sum_{i_1\cdots i_k}}' \big\langle \delta(\boldsymbol{r}_1 - \boldsymbol{x}_{i_1})\big\rangle\cdots\big\langle\delta(\boldsymbol{r}_k - \boldsymbol{x}_{i_k})\big\rangle \\ &= \frac{N!}{(N-k)!}\cdot\frac{1}{N^k}\, n_1(\boldsymbol{r}_1)\cdots n_1(\boldsymbol{r}_k) \\ &= \left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{k-1}{N}\right) n_1(\boldsymbol{r}_1)\cdots n_1(\boldsymbol{r}_k) \quad .\end{aligned} \tag{6.154}$$

The *k-particle distribution function* is defined as the ratio

$$g_k(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k) \equiv \frac{n_k(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k)}{n_1(\boldsymbol{r}_1)\cdots n_1(\boldsymbol{r}_k)} \quad . \tag{6.155}$$

For large separations, then,

$$g_k(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k) \xrightarrow[r_{ij}\to\infty]{} \prod_{j=1}^{k-1}\left(1-\frac{j}{N}\right) \quad . \tag{6.156}$$

For isotropic systems, the two-particle distribution function $g_2(\boldsymbol{r}_1,\boldsymbol{r}_2)$ depends only on the magnitude $|\boldsymbol{r}_1 - \boldsymbol{r}_2|$. As a function of this scalar separation, the function is known as the *radial distribution function*:

$$g(r) \equiv g_2(\boldsymbol{r}) = \frac{1}{n^2}\,\Big\langle \sum_{i\neq j}\delta(\boldsymbol{r} - \boldsymbol{x}_i)\,\delta(\boldsymbol{x}_j)\Big\rangle = \frac{1}{Vn^2}\,\Big\langle \sum_{i\neq j}\delta(\boldsymbol{r} - \boldsymbol{x}_i + \boldsymbol{x}_j)\Big\rangle \quad . \tag{6.157}$$

The radial distribution function is of great importance in the physics of liquids because

- thermodynamic properties of the system can be related to $g(r)$
- $g(r)$ is directly measurable by scattering experiments

For example, in an isotropic system the average potential energy is given by

$$\begin{aligned}\langle W\rangle &= \tfrac{1}{2}\int d^d r_1 \int d^d r_2\, u(\boldsymbol{r}_1 - \boldsymbol{r}_2)\, n_2(\boldsymbol{r}_1,\boldsymbol{r}_2) \\ &= \tfrac{1}{2}n^2 \int d^d r_1 \int d^d r_2\, u(\boldsymbol{r}_1 - \boldsymbol{r}_2)\, g\big(|\boldsymbol{r}_1 - \boldsymbol{r}_2|\big) = \frac{N^2}{2V}\int d^d r\, u(r)\, g(r) \quad .\end{aligned} \tag{6.158}$$

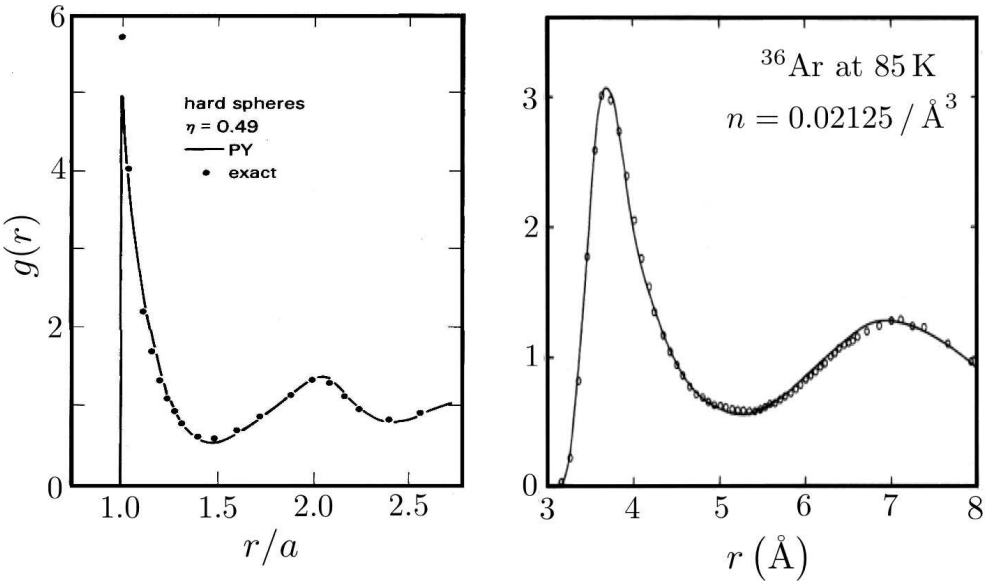



Figure 6.15: Pair distribution functions for hard spheres of diameter $a$ at filling fraction $\varphi = \frac{\pi}{6}a^3 n = 0.49$ (left) and for liquid Argon at $T = 85\,\mathrm{K}$ (right). Molecular dynamics data for hard spheres (points) is compared with the result of the Percus-Yevick approximation (see below in §6.4.8). Reproduced (without permission) from J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, fig 5.5. Experimental data on liquid argon are from the neutron scattering work of J. L. Yarnell *et al.*, *Phys. Rev. A* **7**, 2130 (1973). The data (points) are compared with molecular dynamics calculations by Verlet (1967) for a Lennard-Jones fluid.

For a three-dimensional system, the average internal (*i.e.* potential) energy per particle is

$$\frac{\langle W\rangle}{N} = 2\pi n \int_0^\infty dr\, r^2\, g(r)\, u(r) \quad . \tag{6.159}$$

Intuitively, $f(r)\,dr \equiv 4\pi r^2\, n\, g(r)\,dr$ is the average number of particles lying at a radial distance between $r$ and $r + dr$ from a given reference particle. The total potential energy of interaction with the reference particle is then $f(r)\,u(r)\,dr$. Now integrate over all $r$ and divide by two to avoid double-counting. This recovers eqn. 6.159.

In the OCE, $g(r)$ obeys the sum rule

$$\int d^d r\, g(r) = \frac{V}{N^2}\cdot N(N-1) = V - \frac{V}{N} \quad , \tag{6.160}$$

hence

$$n\int d^d r\, \big[g(r) - 1\big] = -1 \qquad \text{(OCE)} \quad . \tag{6.161}$$

The function $h(r) \equiv g(r) - 1$ is called the *pair correlation function*.

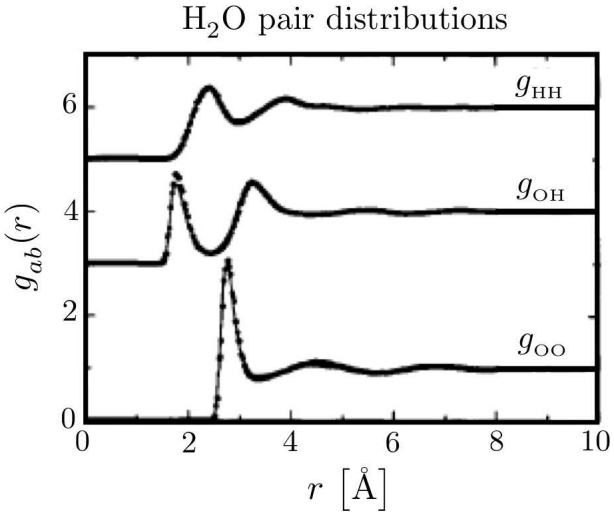



Figure 6.16: Monte Carlo pair distribution functions for liquid water. From A. K. Soper, *Chem Phys.* **202**, 295 (1996).

In the grand canonical formulation, we have

$$\begin{aligned} n\int d^3r\, h(r) &= \frac{\langle N\rangle}{V}\cdot\left[\frac{\langle N(N-1)\rangle}{\langle N\rangle^2}\,V - V\right] \\ &= \frac{\langle N^2\rangle - \langle N\rangle^2}{\langle N\rangle} - 1 = nk_{\rm B}T\kappa_T - 1 \qquad \text{(GCE)} \quad , \end{aligned} \tag{6.162}$$

where $\kappa_T$ is the isothermal compressibility. Note that in an ideal gas we have $h(r) = 0$ and $\kappa_T = \kappa_T^0 \equiv 1/nk_{\rm B}T$. Self-condensed systems, such as liquids and solids far from criticality, are nearly incompressible, hence $0 < nk_{\rm B}T\,\kappa_T \ll 1$, and therefore $n\int d^3r\, h(r) \approx -1$. For incompressible systems, where $\kappa_T = 0$, this becomes an equality.

As we shall see below in §6.4.4, the function $h(r)$, or rather its Fourier transform $\hat{h}(\boldsymbol{k})$, is directly measured in a scattering experiment. The question then arises as to which result applies: the OCE result from eqn. 6.161 or the GCE result from eqn. 6.162. The answer is that under almost all experimental conditions it is the GCE result which applies. The reason for this is that the scattering experiment typically illuminates only a subset of the entire system. This subsystem is in particle equilibrium with the remainder of the system, hence it is appropriate to use the grand canonical ensemble. The OCE results would only apply if the scattering experiment were to measure the entire system.

### 6.4.3 Virial equation of state

The *virial* of a mechanical system is defined to be

$$G = \sum_i \boldsymbol{x}_i \cdot \boldsymbol{F}_i \quad , \tag{6.163}$$

where $\boldsymbol{F}_i$ is the total force acting on particle $i$. If we average $G$ over time, we obtain

$$\begin{aligned}\langle G\rangle &= \lim_{T\to\infty}\frac{1}{T}\int_0^T dt\,\sum_i \boldsymbol{x}_i\cdot\boldsymbol{F}_i \\ &= -\lim_{T\to\infty}\frac{1}{T}\int_0^T dt\,\sum_i m\,\dot{\boldsymbol{x}}_i^2 = -3Nk_{\rm B}T \quad .\end{aligned} \tag{6.164}$$

Here, we have made use of

$$\boldsymbol{x}_i\cdot\boldsymbol{F}_i = m\,\boldsymbol{x}_i\cdot\ddot{\boldsymbol{x}}_i = -\,m\,\dot{\boldsymbol{x}}_i^2 + \frac{d}{dt}\Big(m\,\boldsymbol{x}_i\cdot\dot{\boldsymbol{x}}_i\Big) \quad , \tag{6.165}$$

as well as ergodicity and equipartition of kinetic energy. We have also assumed three space dimensions. In a bounded system, there are two contributions to the force $\boldsymbol{F}_i$. One contribution is from the surfaces which enclose the system. This is given by[18]

$$\langle G\rangle_{\rm surfaces} = \Big\langle\sum_i \boldsymbol{x}_i\cdot\boldsymbol{F}_i^{(\rm surf)}\Big\rangle = -3pV \quad . \tag{6.166}$$

The remaining contribution is due to the interparticle forces. Thus,

$$\frac{p}{k_{\rm B}T} = \frac{N}{V} - \frac{1}{3Vk_{\rm B}T}\Big\langle\sum_i \boldsymbol{x}_i\cdot\boldsymbol{\nabla}_i W\Big\rangle \quad . \tag{6.167}$$

Invoking the definition of $g(r)$, we have

$$p = nk_{\rm B}T\left\{1 - \frac{2\pi n}{3k_{\rm B}T}\int_0^\infty dr\,r^3\,g(r)\,u'(r)\right\} \quad . \tag{6.168}$$

As an alternate derivation, consider the First Law of Thermodynamics,

$$d\Omega = -S\,dT - p\,dV - N\,d\mu \quad , \tag{6.169}$$

from which we derive

$$p = -\left(\frac{\partial\Omega}{\partial V}\right)_{T,\mu} = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \quad . \tag{6.170}$$

Now let $V\to\ell^3 V$, where $\ell$ is a scale parameter. Then

$$p = -\frac{\partial\Omega}{\partial V} = -\frac{1}{3V}\,\frac{\partial}{\partial\ell}\bigg|_{\ell=1}\Omega(T,\ell^3V,\mu) \quad . \tag{6.171}$$

[18]To derive this expression, note that $\boldsymbol{F}^{(\rm surf)}$ is directed inward and vanishes away from the surface. Each Cartesian direction $\alpha = (x,y,z)$ then contributes $-F_\alpha^{(\rm surf)}L_\alpha$, where $L_\alpha$ is the corresponding linear dimension. But $F_\alpha^{(\rm surf)} = pA_\alpha$, where $A_\alpha$ is the area of the corresponding face and $p$ is the pressure. Summing over the three possibilities for $\alpha$, one obtains eqn. 6.166.

Now

$$\begin{aligned}\Xi(T,\ell^3V,\mu) &= \sum_{N=0}^{\infty}\frac{1}{N!}\,e^{\beta\mu N}\,\lambda_T^{-3N}\int\limits_{\ell^3V}\!d^3x_1\cdots\int\limits_{\ell^3V}\!d^3x_N\,e^{-\beta W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)}\\ &= \sum_{N=0}^{\infty}\frac{1}{N!}\left(e^{\beta\mu}\,\lambda_T^{-3}\right)^N\ell^{3N}\int\limits_{V}\!d^3x_1\cdots\int\limits_{V}\!d^3x_N\,e^{-\beta W(\ell\boldsymbol{x}_1,\ldots,\ell\boldsymbol{x}_N)}\end{aligned}\tag{6.172}$$

Thus,

$$\begin{aligned}p &= -\frac{1}{3V}\,\frac{\partial\Omega(\ell^3V)}{\partial\ell}\bigg|_{\ell=1} = \frac{k_{\rm B}T}{3V}\,\frac{1}{\Xi}\,\frac{\partial\Xi(\ell^3V)}{\partial\ell}\bigg|_{\ell=1}\\ &= \frac{k_{\rm B}T}{3V}\,\frac{1}{\Xi}\,\sum_{N=0}^{\infty}\frac{1}{N!}\big(z\lambda_T^{-3}\big)^N\left\{\int\limits_V\!d^3x_1\cdots\int\limits_V\!d^3x_N\,e^{-\beta W(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)}\left[3N-\beta\sum_i\boldsymbol{x}_i\cdot\frac{\partial W}{\partial\boldsymbol{x}_i}\right]\right\}\\ &= nk_{\rm B}T-\frac{1}{3V}\left\langle\frac{\partial W}{\partial\ell}\right\rangle_{\ell=1}\quad.\end{aligned}\tag{6.173}$$

Finally, from $W=\sum_{i<j}u(\ell\boldsymbol{x}_{ij})$ we have

$$\left\langle\frac{\partial W}{\partial\ell}\right\rangle_{\ell=1}=\left\langle\sum_{i<j}\boldsymbol{x}_{ij}\cdot\boldsymbol{\nabla}u(\boldsymbol{x}_{ij})\right\rangle=\frac{2\pi N^2}{V}\int\limits_0^\infty\!dr\,r^3g(r)\,u'(r)\quad,\tag{6.174}$$

and hence

$$p=nk_{\rm B}T-\tfrac{2}{3}\pi n^2\int\limits_0^\infty\!dr\,r^3\,g(r)\,u'(r)\quad.\tag{6.175}$$

Note that the density $n$ enters the equation of state explicitly on the RHS of the above equation, but also implicitly through the pair distribution function $g(r)$, which has implicit dependence on both $n$ and $T$.

### 6.4.4 Correlations and scattering

Consider the scattering of a light or particle beam (*i.e.* photons or neutrons) from a liquid. We label the states of the beam particles by their wavevector $\boldsymbol{k}$ and we assume a general dispersion $\varepsilon_{\boldsymbol{k}}$. For photons, $\varepsilon_{\boldsymbol{k}}=\hbar c|\boldsymbol{k}|$, while for neutrons $\varepsilon_{\boldsymbol{k}}=\hbar^2\boldsymbol{k}^2/2m_{\rm n}$. We assume a single scattering process with the liquid, during which the total momentum and energy of the liquid plus beam are conserved. We write

$$\boldsymbol{k}'=\boldsymbol{k}+\boldsymbol{q}\qquad,\qquad\varepsilon_{\boldsymbol{k}'}=\varepsilon_{\boldsymbol{k}}+\hbar\omega\quad,\tag{6.176}$$

where $\boldsymbol{k}'$ is the final state of the scattered beam particle. Thus, the fluid transfers momentum $\Delta\boldsymbol{p}=\hbar\boldsymbol{q}$ and energy $\hbar\omega$ to the beam.

Now consider the scattering process between an initial state $|\,i,\boldsymbol{k}\,\rangle$ and a final state $|\,j,\boldsymbol{k}'\,\rangle$, where these states describe both the beam and the liquid. According to Fermi's Golden Rule, the scattering rate is

$$\Gamma_{i\boldsymbol{k}\to j\boldsymbol{k}'}=\frac{2\pi}{\hbar}\left|\langle\,j,\boldsymbol{k}'\,|\,\mathcal{V}\,|\,i,\boldsymbol{k}\,\rangle\right|^2\delta(E_j-E_i+\hbar\omega)\quad,\tag{6.177}$$

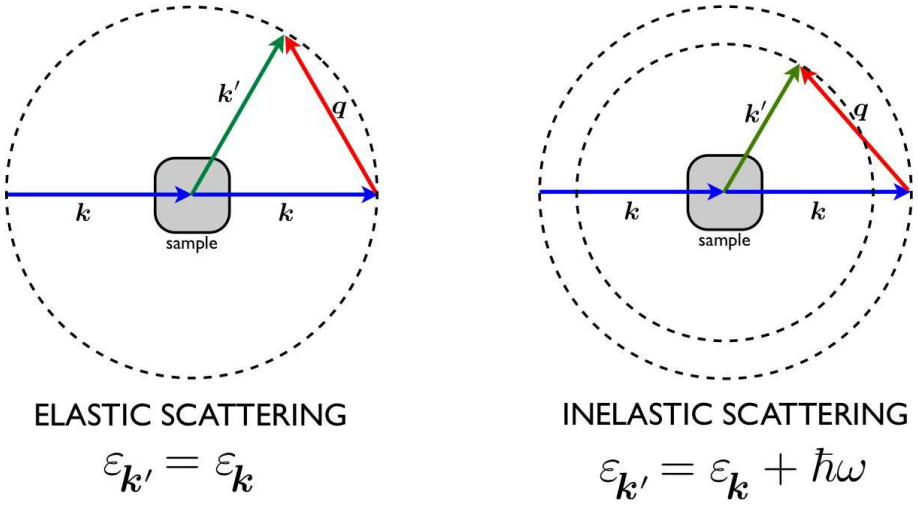



Figure 6.17: In a scattering experiment, a beam of particles interacts with a sample and the beam particles scatter off the sample particles. A momentum $\hbar\boldsymbol{q}$ and energy $\hbar\omega$ are transferred to the beam particle during such a collision. If $\omega = 0$, the scattering is said to be *elastic*. For $\omega \neq 0$, the scattering is *inelastic*.

where $\mathcal{V}$ is the scattering potential and $E_i$ is the initial internal energy of the liquid. Note that overall energy conservation requires $E_j + \varepsilon_{\boldsymbol{k}'} = E_i + \varepsilon_{\boldsymbol{k}}$ and therefore $E_j = E_i - \hbar\omega$. If $\boldsymbol{r}$ is the position of the beam particle and $\{\boldsymbol{x}_l\}$ are the positions of the liquid particles, then

$$\mathcal{V}(\boldsymbol{r}) = \sum_{l=1}^{N} v(\boldsymbol{r} - \boldsymbol{x}_l) \quad . \tag{6.178}$$

For elastic scattering, the *differential scattering cross section* $d\sigma$ is defined to be

$$d\sigma(\hat{\Omega}) = \frac{\text{rate at which particles scattered into solid angle } d\Omega}{\text{incident flux}} \quad . \tag{6.179}$$

For inelastic scattering,

$$d^2\sigma(\hat{\Omega},\omega) = \frac{\text{rate at which particles scattered into solid angle } d\Omega \text{ and energy change within } \hbar\, d\omega}{\text{incident flux}} \quad . \tag{6.180}$$

The inelastic differential scattering cross section (per unit frequency per unit solid angle) is

$$\frac{\partial^2\sigma}{\partial\Omega\,\partial\omega} = \frac{\hbar}{4\pi}\,\frac{g(\varepsilon_{\boldsymbol{k}'})}{|\boldsymbol{v}_{\boldsymbol{k}}|}\sum_{i,j} P_i\,\Gamma_{i\boldsymbol{k}\to j\boldsymbol{k}'} \quad , \tag{6.181}$$

where $g(\varepsilon) = \int \frac{d^dk}{(2\pi)^d}\,\delta(\varepsilon - \varepsilon_{\boldsymbol{k}})$ is the density of states for the beam particles, and $P_i = Z^{-1}\,e^{-\beta E_i}$ is the Boltzmann weight.

Consider now the matrix element

$$\begin{aligned}\left\langle\, j, \boldsymbol{k}' \,\middle|\, \mathcal{V} \,\middle|\, i, \boldsymbol{k} \,\right\rangle &= \left\langle\, j \,\middle|\, \frac{1}{V} \sum_{l=1}^{N} \int d^d r\; e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot \boldsymbol{r}}\, v(\boldsymbol{r}-\boldsymbol{x}_l) \,\middle|\, i \,\right\rangle \\ &= \frac{1}{V}\, \hat{v}(\boldsymbol{q}) \left\langle\, j \,\middle|\, \sum_{l=1}^{N} e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_l} \,\middle|\, i \,\right\rangle \quad ,\end{aligned} \tag{6.182}$$

where we have assumed that the incident and scattered beams are plane waves. We then have

$$\begin{aligned}\frac{\partial^2\sigma}{\partial\Omega\,\partial\omega} &= \frac{\hbar}{2}\, \frac{g(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}})}{|\boldsymbol{\nabla}_{\boldsymbol{k}}\varepsilon_{\boldsymbol{k}}|}\, \frac{|\hat{v}(\boldsymbol{q})|^2}{V^2} \sum_i P_i \sum_j \left|\left\langle\, j \,\middle|\, \sum_{l=1}^{N} e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_l} \,\middle|\, i \,\right\rangle\right|^2 \delta(E_j - E_i + \hbar\omega) \\ &= \frac{g(\varepsilon_{\boldsymbol{k}+\boldsymbol{q}})}{4\pi\,|\boldsymbol{\nabla}_{\boldsymbol{k}}\varepsilon_{\boldsymbol{k}}|}\, \frac{N}{V^2}\, |\hat{v}(\boldsymbol{q})|^2\, S(\boldsymbol{q},\omega) \quad ,\end{aligned} \tag{6.183}$$

where $S(\boldsymbol{q},\omega)$ is the *dynamic structure factor*,

$$S(\boldsymbol{q},\omega) = \frac{2\pi\hbar}{N} \sum_i P_i \sum_j \left|\left\langle\, j \,\middle|\, \sum_{l=1}^{N} e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_l} \,\middle|\, i \,\right\rangle\right|^2 \delta(E_j - E_i + \hbar\omega) \tag{6.184}$$

Note that for an arbitrary operator $A$,

$$\begin{aligned}\sum_j \left|\left\langle\, j \,\middle|\, A \,\middle|\, i \,\right\rangle\right|^2 \delta(E_j - E_i + \hbar\omega) &= \frac{1}{2\pi\hbar} \sum_j \int_{-\infty}^{\infty} dt\; e^{i(E_j - E_i + \hbar\omega)\,t/\hbar} \left\langle\, i \,\middle|\, A^\dagger \,\middle|\, j \,\right\rangle \left\langle\, j \,\middle|\, A \,\middle|\, i \,\right\rangle \\ &= \frac{1}{2\pi\hbar} \sum_j \int_{-\infty}^{\infty} dt\; e^{i\omega t} \left\langle\, i \,\middle|\, A^\dagger \,\middle|\, j \,\right\rangle \left\langle\, j \,\middle|\, e^{i\hat{H}t/\hbar}\, A\, e^{-i\hat{H}t/\hbar} \,\middle|\, i \,\right\rangle \\ &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dt\; e^{i\omega t} \left\langle\, i \,\middle|\, A^\dagger(0)\, A(t) \,\middle|\, i \,\right\rangle \quad .\end{aligned} \tag{6.185}$$

Thus,

$$\begin{aligned}S(\boldsymbol{q},\omega) &= \frac{1}{N} \int_{-\infty}^{\infty} dt\; e^{i\omega t} \sum_i P_i \left\langle\, i \,\middle|\, \sum_{l,l'} e^{i\boldsymbol{q}\cdot\boldsymbol{x}_l(0)}\, e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_{l'}(t)} \,\middle|\, i \,\right\rangle \\ &= \frac{1}{N} \int_{-\infty}^{\infty} dt\; e^{i\omega t} \left\langle \sum_{l,l'} e^{i\boldsymbol{q}\cdot\boldsymbol{x}_l(0)}\, e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_{l'}(t)} \right\rangle \quad ,\end{aligned} \tag{6.186}$$

where the angular brackets in the last line denote a thermal expectation value of a quantum mechanical operator. If we integrate over all frequencies, we obtain the equal time correlator,

$$\begin{aligned}S(\boldsymbol{q}) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, S(\boldsymbol{q},\omega) = \frac{1}{N} \sum_{l,l'} \left\langle e^{i\boldsymbol{q}\cdot(\boldsymbol{x}_l - \boldsymbol{x}_{l'})} \right\rangle \\ &= N\,\delta_{\boldsymbol{q},0} + 1 + n \int d^d r\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} \left[g(r) - 1\right] \quad .\end{aligned} \tag{6.187}$$

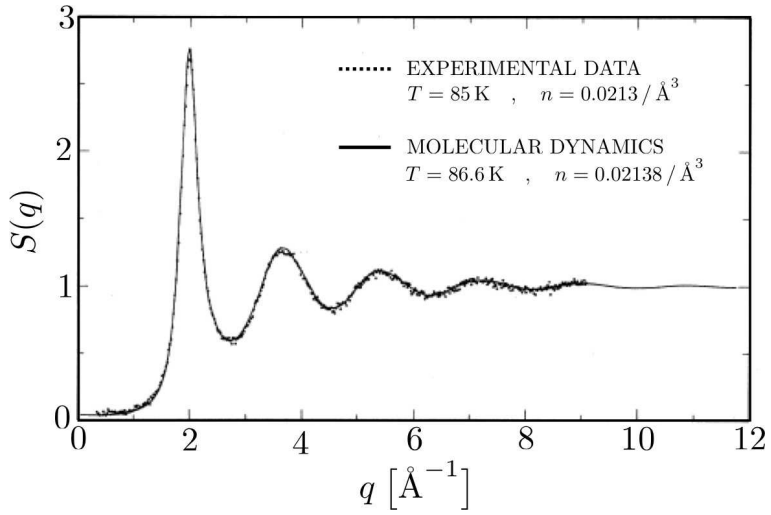



Figure 6.18: Comparison of the static structure factor of Argon as determined by neutron scattering work of J. L. Yarnell *et al.*, *Phys. Rev. A* **7**, 2130 (1973) with molecular dynamics calculations by Verlet (1967) for a Lennard-Jones fluid.

known as the *static structure factor*[19]. Note that $S(\boldsymbol{q} = 0) = N$, since all the phases $e^{i\boldsymbol{q}\cdot(\boldsymbol{x}_i - \boldsymbol{x}_j)}$ are then unity. As $q \to \infty$, the phases oscillate rapidly with changes in the distances $|\boldsymbol{x}_i - \boldsymbol{x}_j|$, and average out to zero. However, the 'diagonal' terms in the sum, *i.e.* those with $i = j$, always contribute a total of 1 to $S(\boldsymbol{q})$. Therefore in the $q \to \infty$ limit we have $S(q \to \infty) = 1$.

In general, the detectors used in a scattering experiment are sensitive to the energy of the scattered beam particles, although there is always a finite experimental resolution, both in $\boldsymbol{q}$ and $\omega$. This means that what is measured is actually something like

$$S_{\rm meas}(\boldsymbol{q},\omega) = \int d^d q' \int d\omega'\, F(\boldsymbol{q} - \boldsymbol{q}')\, G(\omega - \omega')\, S(\boldsymbol{q}', \omega') \quad , \tag{6.188}$$

where $F$ and $G$ are essentially Gaussian functions of their argument, with width given by the experimental resolution. If one integrates over all frequencies $\omega$, *i.e.* if one simply counts scattered particles as a function of $\boldsymbol{q}$ but without any discrimination of their energies, then one measures the static structure factor $S(\boldsymbol{q})$. Elastic scattering is determined by $S(\boldsymbol{q}, \omega = 0)$, *i.e.* at no energy transfer.

### 6.4.5 Correlation and response

Suppose an external potential $v(\boldsymbol{x})$ is also present. Then

$$P(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) = \frac{1}{Q_N[v]} \cdot \frac{1}{N!}\, e^{-\beta W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)}\, e^{-\beta \sum_i v(\boldsymbol{x}_i)} \quad , \tag{6.189}$$

where

$$Q_N[v] = \frac{1}{N!} \int d^d x_1 \cdots \int d^d x_N\, e^{-\beta W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)}\, e^{-\beta \sum_i v(\boldsymbol{x}_i)} \quad . \tag{6.190}$$

[19] We may write $\delta_{\boldsymbol{q},0} = \frac{1}{V}\,(2\pi)^d\,\delta(\boldsymbol{q})$.

The Helmholtz free energy is then

$$F = -\frac{1}{\beta}\,\log\left(\lambda_T^{-dN}\,Q_N[v]\right) \quad . \tag{6.191}$$

Now consider the functional derivative

$$\frac{\delta F}{\delta v(\boldsymbol{r})} = -\frac{1}{\beta}\cdot\frac{1}{Q_N}\cdot\frac{\delta Q_N}{\delta v(\boldsymbol{r})} \quad . \tag{6.192}$$

Using

$$\sum_i v(\boldsymbol{x}_i) = \int d^d r\, v(\boldsymbol{r})\,\sum_i \delta(\boldsymbol{r}-\boldsymbol{x}_i) \quad , \tag{6.193}$$

hence

$$\frac{\delta F}{\delta v(\boldsymbol{r})} = \int d^d x_1 \cdots \int d^d x_N\, P(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)\,\sum_i \delta(\boldsymbol{r}-\boldsymbol{x}_i) = n_1(\boldsymbol{r}) \quad , \tag{6.194}$$

which is the local density at $\boldsymbol{r}$.

Next, consider the *response function*,

$$\begin{aligned}\chi(\boldsymbol{r},\boldsymbol{r}') &\equiv \frac{\delta n_1(\boldsymbol{r})}{\delta v(\boldsymbol{r}')} = \frac{\delta^2 F[v]}{\delta v(\boldsymbol{r})\,\delta v(\boldsymbol{r}')} \\ &= \frac{1}{\beta}\cdot\frac{1}{Q_N^2}\,\frac{\delta Q_N}{\delta v(\boldsymbol{r})}\,\frac{\delta Q_N}{\delta v(\boldsymbol{r}')} - \frac{1}{\beta}\cdot\frac{1}{Q_N}\,\frac{\delta^2 Q_N}{\delta v(\boldsymbol{r})\,\delta v(\boldsymbol{r}')} \\ &= \beta\, n_1(\boldsymbol{r})\, n_1(\boldsymbol{r}') - \beta\, n_1(\boldsymbol{r})\,\delta(\boldsymbol{r}-\boldsymbol{r}') - \beta\, n_2(\boldsymbol{r},\boldsymbol{r}') \quad .\end{aligned} \tag{6.195}$$

In an isotropic system, $\chi(\boldsymbol{r},\boldsymbol{r}') = \chi(\boldsymbol{r}-\boldsymbol{r}')$ is a function of the coordinate separation, and

$$\begin{aligned}-k_{\rm B}T\,\chi(\boldsymbol{r}-\boldsymbol{r}') &= -n^2 + n\,\delta(\boldsymbol{r}-\boldsymbol{r}') + n^2 g\big(|\boldsymbol{r}-\boldsymbol{r}'|\big) \\ &= n^2\,h\big(|\boldsymbol{r}-\boldsymbol{r}'|\big) + n\,\delta(\boldsymbol{r}-\boldsymbol{r}') \quad .\end{aligned} \tag{6.196}$$

Taking the Fourier transform,

$$-k_{\rm B}T\,\hat\chi(\boldsymbol{q}) = n + n^2\,\hat h(\boldsymbol{q}) = n\,S(\boldsymbol{q}) \quad . \tag{6.197}$$

We may also write

$$-\frac{k_{\rm B}T}{n}\,\hat\chi(0) = 1 + n\,\hat h(0) = \frac{\kappa_T}{\kappa_T^0} \quad , \tag{6.198}$$

*i.e.* $n^2\kappa_T = -\hat\chi(0)$, since $\kappa_T^0 = 1/nk_{\rm B}T$.

What does this all mean? Suppose we have an isotropic system which is subjected to a weak, spatially inhomogeneous potential $v(\boldsymbol{r})$. We expect that the density $n(\boldsymbol{r})$ in the presence of the inhomogeneous

potential itself to be inhomogeneous. The first corrections to the $v = 0$ value $n = n_0$ are linear in $v$, and given by

$$\begin{aligned}\delta n(\boldsymbol{r}) &= \int d^d r'\, \chi(\boldsymbol{r}, \boldsymbol{r}')\, v(\boldsymbol{r}') \\ &= -\beta n_0\, v(\boldsymbol{r}) - \beta n_0^2 \int d^d r'\, h(\boldsymbol{r} - \boldsymbol{r})\, v(\boldsymbol{r}') \quad .\end{aligned} \tag{6.199}$$

Note that if $v(\boldsymbol{r}) > 0$ it becomes energetically more costly for a particle to be at $\boldsymbol{r}$. Accordingly, the density response is negative, and proportional to the ratio $v(\boldsymbol{r})/k_{\rm B}T$ – this is the first term in the above equation. If there were no correlations between the particles, then $h = 0$ and this would be the entire story. However, the particles in general *are* correlated. Consider, for example, the case of hard spheres of diameter $a$, and let there be a repulsive potential at $\boldsymbol{r} = 0$. This means that it is less likely for a particle to be centered anywhere within a distance $a$ of the origin. But then it will be *more* likely to find a particle in the next 'shell' of radial thickness $a$.

### 6.4.6 BBGKY hierarchy

The distribution functions satisfy a hierarchy of integro-differential equations known as the *BBGKY hierarchy*[20]. In homogeneous systems, we have

$$g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) = \frac{N!}{(N-k)!}\,\frac{1}{n^k}\int d^d x_{k+1}\cdots\int d^d x_N\, P(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k\,,\,\boldsymbol{x}_{k+1}\,,\,\ldots\,,\,\boldsymbol{x}_N) \quad , \tag{6.200}$$

where

$$P(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) = \frac{1}{Q_N}\cdot\frac{1}{N!}\,e^{-\beta W(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N)} \quad . \tag{6.201}$$

Taking the gradient with respect to $\boldsymbol{r}_1$, we have

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{r}_1} g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) = \frac{1}{Q_N}\cdot\frac{n^{-k}}{(N-k)!}\int d^d x_{k+1}\cdots\int d^d x_N\, e^{-\beta\sum_{k<i<j} u(\boldsymbol{x}_{ij})} \\ \times\,\frac{\partial}{\partial \boldsymbol{r}_1}\left[\, e^{-\beta\sum_{i<j\le k} u(\boldsymbol{r}_{ij})}\cdot e^{-\beta\sum_{i\le k<j} u(\boldsymbol{r}_i - \boldsymbol{x}_j)}\right] \quad ,\end{aligned} \tag{6.202}$$

where $\sum_{k<i<j}$ means to sum on indices $i$ and $j$ such that $i < j$ and $k < i$, *i.e.*

$$\begin{aligned}\sum_{k<i<j} u(\boldsymbol{x}_{ij}) &\equiv \sum_{i=k+1}^{N-1}\sum_{j=i+1}^{N} u\big(\boldsymbol{x}_i - \boldsymbol{x}_j\big) \\ \sum_{i<j\le k} u(\boldsymbol{r}_{ij}) &\equiv \sum_{i=1}^{k-1}\sum_{j=i+1}^{k} u\big(\boldsymbol{r}_i - \boldsymbol{r}_j\big) \\ \sum_{i\le k<j} u(\boldsymbol{r}_i - \boldsymbol{x}_j) &= \sum_{i=1}^{k}\sum_{j=k+1}^{N} u(\boldsymbol{r}_i - \boldsymbol{x}_j) \quad .\end{aligned} \tag{6.203}$$

[20] So named after Bogoliubov, Born, Green, Kirkwood, and Yvon.

Now

$$\frac{\partial}{\partial \boldsymbol{r}_1}\left[e^{-\beta\sum_{i<j\le k}u(r_{ij})}\cdot e^{-\beta\sum_{i\le k<j}u(\boldsymbol{r}_i-\boldsymbol{x}_j)}\right] = \tag{6.204}$$
$$-\beta\left\{\sum_{j=1}^{k}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\partial \boldsymbol{r}_1}+\sum_{j=k+1}^{N}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{x}_j)}{\partial \boldsymbol{r}_1}\right\}\cdot\left[e^{-\beta\sum_{i<j\le k}u(r_{ij})}\cdot e^{-\beta\sum_{i\le k<j}u(\boldsymbol{r}_i-\boldsymbol{x}_j)}\right] \quad ,$$

hence

$$\frac{\partial}{\partial \boldsymbol{r}_1}g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) = -\beta\sum_{j=2}^{k}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\partial \boldsymbol{r}_1}\,g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) \tag{6.205}$$
$$-\beta(N-k)\int d^dx_{k+1}\,\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{x}_{k+1})}{\partial \boldsymbol{r}_1}\,P(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k\,,\,\boldsymbol{x}_{k+1}\,,\,\ldots\,,\,\boldsymbol{x}_N)$$
$$= -\beta\sum_{j=2}^{k}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\partial \boldsymbol{r}_1}\,g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k)$$
$$-\beta n\int d^dx_{k+1}\,\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{x}_{k+1})}{\partial \boldsymbol{r}_1}\,g_{k+1}(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k\,,\,x_{k+1}) \quad .$$

Thus, we obtain the BBGKY hierarchy:

$$-k_{\rm B}T\,\frac{\partial}{\partial \boldsymbol{r}_1}g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) = \sum_{j=2}^{k}\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\partial \boldsymbol{r}_1}\,g_k(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k) \tag{6.206}$$
$$+\,n\int d^dr'\,\frac{\partial u(\boldsymbol{r}_1-\boldsymbol{r}')}{\partial \boldsymbol{r}_1}\,g_{k+1}(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_k\,,\,\boldsymbol{r}') \quad .$$

The BBGKY hierarchy is an infinite tower of coupled integro-differential equations, relating $g_k$ to $g_{k+1}$ for all $k$. If we approximate $g_k$ at some level $k$ in terms of equal or lower order distributions, then we obtain a closed set of equations which in principle can be solved, at least numerically. For example, the *Kirkwood approximation* closes the hierarchy at order $k=2$ by imposing the condition

$$g_3(\boldsymbol{r}_1\,,\,\boldsymbol{r}_2\,,\,\boldsymbol{r}_3) \equiv g(\boldsymbol{r}_1-\boldsymbol{r}_2)\,g(\boldsymbol{r}_1-\boldsymbol{r}_3)\,g(\boldsymbol{r}_2-\boldsymbol{r}_3) \quad . \tag{6.207}$$

This results in the single integro-differential equation

$$-k_{\rm B}T\,\boldsymbol{\nabla}g(\boldsymbol{r}) = g(\boldsymbol{r})\,\boldsymbol{\nabla}u + n\int d^dr'\,g(\boldsymbol{r})\,g(\boldsymbol{r}')\,g(\boldsymbol{r}-\boldsymbol{r}')\,\boldsymbol{\nabla}u(\boldsymbol{r}-\boldsymbol{r}') \quad . \tag{6.208}$$

This is known as the Born-Green-Yvon (BGY) equation. In practice, the BGY equation, which is solved numerically, gives adequate results only at low densities.

### 6.4.7 Ornstein-Zernike theory

The *direct correlation function* $c(\boldsymbol{r})$ is defined by the equation

$$h(\boldsymbol{r}) = c(\boldsymbol{r}) + n\int d^3r'\,h(\boldsymbol{r}-\boldsymbol{r}')\,c(\boldsymbol{r}') \quad , \tag{6.209}$$

where $h(r) = g(r) - 1$ and we assume an isotropic system. This is called the *Ornstein-Zernike equation*. The first term, $c(\boldsymbol{r})$, accounts for local correlations, which are then propagated in the second term to account for long-ranged correlations.

The OZ equation is an integral equation, but it becomes a simple algebraic one upon Fourier transforming:

$$\hat{h}(\boldsymbol{q}) = \hat{c}(\boldsymbol{q}) + n\,\hat{h}(\boldsymbol{q})\,\hat{c}(\boldsymbol{q}) \quad , \tag{6.210}$$

the solution of which is

$$\hat{h}(\boldsymbol{q}) = \frac{\hat{c}(\boldsymbol{q})}{1 - n\,\hat{c}(\boldsymbol{q})} \quad . \tag{6.211}$$

The static structure factor is then

$$S(\boldsymbol{q}) = 1 + n\,\hat{h}(\boldsymbol{q}) = \frac{1}{1 - n\,\hat{c}(\boldsymbol{q})} \quad . \tag{6.212}$$

In the grand canonical ensemble, we can write

$$\kappa_T = \frac{1 + n\,\hat{h}(0)}{nk_{\rm B}T} = \frac{1}{nk_{\rm B}T} \cdot \frac{1}{1 - n\,\hat{c}(0)} \qquad \Longrightarrow \qquad n\,\hat{c}(0) = 1 - \frac{\kappa_T^0}{\kappa_T} \quad , \tag{6.213}$$

where $\kappa_T^0 = 1/nk_{\rm B}T$ is the ideal gas isothermal compressibility.

**Ornstein-Zernike approximation at long wavelengths**

Let's expand the direct correlation function $\hat{c}(\boldsymbol{q})$ in powers of the wavevector $\boldsymbol{q}$, *viz.*

$$\hat{c}(\boldsymbol{q}) = \hat{c}(0) + c_2\,q^2 + c_4\,q^4 + \ldots \quad . \tag{6.214}$$

Here we have assumed spatial isotropy. Then

$$\begin{aligned} 1 - n\,\hat{c}(\boldsymbol{q}) = \frac{1}{S(\boldsymbol{q})} &= 1 - n\,\hat{c}(0) - n\,c_2\,q^2 + \ldots \\ &\equiv \xi^{-2}R^2 + q^2R^2 + \mathcal{O}(q^4) \quad , \end{aligned} \tag{6.215}$$

where

$$R^2 = -n\,c_2 = \frac{2\pi n}{3}\int\limits_0^\infty dr\; r^4\,c(r) \tag{6.216}$$

and

$$\xi^{-2} = \frac{1 - n\,\hat{c}(0)}{R^2} = \frac{3}{2\pi n} \cdot \frac{1 - 4\pi n\int_0^\infty dr\,r^2\,c(r)}{\int_0^\infty dr\,r^4\,c(r)} \quad . \tag{6.217}$$

The quantity $R(T)$ tells us something about the effective range of the interactions, while $\xi(T)$ is the *correlation length*. As we approach a critical point, the correlation length diverges as a power law:

$$\xi(T) \sim A\,|T - T_{\rm c}|^{-\nu} \quad . \tag{6.218}$$

From eqns. 6.199 and 6.212, the susceptibility is given by

$$\hat{\chi}(\boldsymbol{q}) = -n\beta\, S(\boldsymbol{q}) = -\frac{n\beta R^{-2}}{\xi^{-2} + q^2 + \mathcal{O}(q^4)} \tag{6.219}$$

Note that $\hat{\chi}(\boldsymbol{q}) < 0$ because the susceptibility is the functional derivative of the density with respect to the added potential, *i.e.* $\chi(\boldsymbol{r}, \boldsymbol{r}') = \delta n(\boldsymbol{r})/\delta v(\boldsymbol{r}')$. A local positive potential is repulsive, causing the density to decrease in its vicinity. In the *Ornstein-Zernike approximation*, one drops the $\mathcal{O}(q^4)$ terms in the denominator and retains only the long-wavelength behavior in the direct correlation function. Thus,

$$\hat{\chi}^{\rm OZ}(\boldsymbol{q}) = -\frac{n\beta R^{-2}}{\xi^{-2} + q^2} \quad . \tag{6.220}$$

We now apply the inverse Fourier transform back to real space to obtain $\chi^{\rm OZ}(\boldsymbol{r})$. In $d = 1$ dimension the result can be obtained exactly:

$$\chi^{\rm OZ}_{d=1}(x) = -\frac{n}{k_{\rm B}TR^2}\int\limits_{-\infty}^{\infty}\frac{dq}{2\pi}\,\frac{e^{iqx}}{\xi^{-2}+q^2} = -\frac{n\xi}{2k_{\rm B}TR^2}\,e^{-|x|/\xi} \quad . \tag{6.221}$$

In $d = 2$ dimensions,

$$\begin{aligned}\chi^{\rm OZ}_{d=2}(\boldsymbol{x}) &= -\frac{n}{k_{\rm B}TR^2}\int\frac{d^2q}{(2\pi)^2}\,\frac{e^{i\boldsymbol{q}\cdot\boldsymbol{x}}}{\xi^{-2}+\boldsymbol{q}^2} \\ &= -\frac{n}{2\pi k_{\rm B}TR^2}\int\limits_0^\infty dq\,\frac{q\,J_0(qr)}{\xi^{-2}+q^2} = -\frac{n}{2\pi k_{\rm B}TR^2}\,K_0(r/\xi) \quad ,\end{aligned} \tag{6.222}$$

where $K_0(z)$ is the modified Bessel function of the second kind. As $z \to 0$ we have[21]

$$K_0(z) = -\log z + \log 2 - \gamma + \mathcal{O}(z^2 \log z) \quad , \tag{6.223}$$

where $\gamma = 0.57721566\ldots$ is Euler's constant.

In higher dimensions $d > 2$ we can obtain the result asymptotically in two limits:

- Take $r \to \infty$ with $\xi$ fixed. Then

$$\chi^{\rm OZ}_d(\boldsymbol{r}) \simeq -C_d\, n \cdot \frac{\xi^{(3-d)/2}}{k_{\rm B}T\,R^2}\cdot\frac{e^{-r/\xi}}{r^{(d-1)/2}}\cdot\left\{1+\mathcal{O}\left(\frac{d-3}{r/\xi}\right)\right\} \quad , \tag{6.224}$$

  where the $C_d$ are dimensionless constants.

- Take $\xi \to \infty$ with $r$ fixed; this is the limit $T \to T_{\rm c}$ at fixed $r$. In dimensions $d > 2$ we obtain

$$\chi^{\rm OZ}_d(\boldsymbol{r}) \simeq -\frac{C'_d\, n}{k_{\rm B}TR^2}\cdot\frac{e^{-r/\xi}}{r^{d-2}}\cdot\left\{1+\mathcal{O}\left(\frac{d-3}{r/\xi}\right)\right\} \quad . \tag{6.225}$$

[21] For a systematic small $z$ expansion, see formula 10.31.2 in the NIST Handbook of Mathematical Functions.

At criticality, $\xi \to \infty$, and clearly our results in $d = 1$ and $d = 2$ dimensions are nonsensical, as they are divergent. To correct this behavior, M. E. Fisher in 1963 suggested that the OZ correlation functions in the $r \ll \xi$ limit be replaced by

$$\chi(\boldsymbol{r}) \simeq -\frac{C_d'' \, n}{k_{\rm B} T R^2} \cdot \frac{e^{-r/\xi}}{r^{d-2+\eta}} \quad , \tag{6.226}$$

a result known as *anomalous scaling*. Here, $\eta$ is the *anomalous scaling exponent*.

Recall that the isothermal compressibility is given by $\kappa_T = -\hat{\chi}(0)$. Near criticality, the integral in $\hat{\chi}(0)$ is dominated by the $r \ll \xi$ part, since $\xi \to \infty$. Thus, using Fisher's anomalous scaling,

$$\begin{aligned} \kappa_T = -\hat{\chi}(0) &= -\int d^d r \, \chi(\boldsymbol{r}) \\ &\sim A \int d^d r \, \frac{e^{-r/\xi}}{r^{d-2+\eta}} \sim B \, \xi^{2-\eta} \sim C \left| T - T_{\rm c} \right|^{-(2-\eta)\nu} \quad , \end{aligned} \tag{6.227}$$

where $A$, $B$, and $C$ are temperature-dependent constants which are nonsingular at $T = T_{\rm c}$. Thus, since $\kappa_T \propto |T - T_{\rm c}|^{-\gamma}$, we conclude

$$\gamma = (2 - \eta)\, \nu \quad , \tag{6.228}$$

a result known as *Fisher's scaling relation*.

In the above derivation of OZ theory, we have merely substituted one unknown function, $h(\boldsymbol{r})$, for another, namely $c(\boldsymbol{r})$. To close the system of equations, we need to relate $c(\boldsymbol{r})$ to $h(\boldsymbol{r})$ again in some way. There are various approximation schemes which do just this.

### 6.4.8 Percus-Yevick and hypernetted chains approximations

In the Percus-Yevick approximation, we take

$$c(\boldsymbol{r}) = \left[1 - e^{\beta u(\boldsymbol{r})}\right] \cdot g(\boldsymbol{r}) \quad . \tag{6.229}$$

Note that $c(\boldsymbol{r})$ vanishes whenever the potential $u(\boldsymbol{r})$ itself vanishes. This results in the following integro-differential equation for the pair distribution function $g(\boldsymbol{r})$:

$$g(\boldsymbol{r}) = e^{-\beta u(\boldsymbol{r})} + n \, e^{-\beta u(\boldsymbol{r})} \int d^3 r' \, \left[g(\boldsymbol{r} - \boldsymbol{r}') - 1\right] \cdot \left[1 - e^{\beta u(\boldsymbol{r}')}\right] g(\boldsymbol{r}') \quad . \tag{6.230}$$

This is the *Percus-Yevick equation*. Remarkably, the Percus-Yevick (PY) equation can be solved analytically for the case of hard spheres, where $u(r) = \infty$ for $r \le a$ and $u(r) = 0$ for $r > a$, where $a$ is the hard sphere diameter. Define the function $y(\boldsymbol{r}) = e^{\beta u(\boldsymbol{r})} g(\boldsymbol{r})$, in which case

$$c(\boldsymbol{r}) = y(\boldsymbol{r})\, f(\boldsymbol{r}) = \begin{cases} -y(\boldsymbol{r}) & , \quad r \le a \\ 0 & , \quad r > a \quad . \end{cases} \tag{6.231}$$

Here, $f(\boldsymbol{r}) = e^{-\beta u(\boldsymbol{r})} - 1$ is the Mayer function. We remark that the definition of $y(\boldsymbol{r})$ may cause some concern for the hard sphere system, because of the $e^{\beta u(\boldsymbol{r})}$ term, which diverges severely for $r \le a$.

However, $g(r)$ vanishes in this limit, and their product $y(\boldsymbol{r})$ is in fact finite! The PY equation may then be written for the function $y(\boldsymbol{r})$ as

$$y(\boldsymbol{r}) = 1 + n\int\limits_{r'<a} d^3r'\, y(\boldsymbol{r}') - n\int\limits_{\substack{r'<a \\ |\boldsymbol{r}-\boldsymbol{r}'|>a}} d^3r'\, y(\boldsymbol{r}')\, y(\boldsymbol{r}-\boldsymbol{r}') \quad . \tag{6.232}$$

This has been solved using Laplace transform methods by M. S. Wertheim, *J. Math. Phys.* **5**, 643 (1964). The final result for $c(\boldsymbol{r})$ is

$$c(\boldsymbol{r}) = -\left\{\lambda_1 + 6\varphi\,\lambda_2\left(\frac{r}{a}\right) + \tfrac{1}{2}\varphi\,\lambda_1\left(\frac{r}{a}\right)^3\right\}\Theta(a-r) \quad , \tag{6.233}$$

where $\varphi = \frac{1}{6}\pi a^3 n$ is the packing fraction[22] and

$$\lambda_1 = \frac{(1+2\varphi)^2}{(1-\varphi)^4} \qquad , \qquad \lambda_2 = -\frac{(1+\frac{1}{2}\varphi)^2}{(1-\varphi)^4} \quad . \tag{6.234}$$

This leads to the equation of state

$$p = nk_{\rm B}T\,\frac{1+\varphi+\varphi^2}{(1-\varphi)^3} \quad . \tag{6.235}$$

This gets $B_2$ and $B_3$ exactly right. The accuracy of the PY approximation for higher order virial coefficients is shown in table 6.1.

To obtain the equation of state from eqn. 6.233, we invoke the compressibility equation,

$$nk_{\rm B}T\,\kappa_T = \left(\frac{\partial n}{\partial p}\right)_T = \frac{1}{1-n\,\hat{c}(0)} \quad . \tag{6.236}$$

We therefore need

$$\begin{aligned}\hat{c}(0) = \int d^3r\, c(\boldsymbol{r}) &= -4\pi a^3\int\limits_0^1 dx\, x^2\,\left[\lambda_1 + 6\,\varphi\,\lambda_2\,x + \tfrac{1}{2}\,\varphi\,\lambda_1\,x^3\right] \\ &= -4\pi a^3\left[\tfrac{1}{3}\,\lambda_1 + \tfrac{3}{2}\,\varphi\,\lambda_2 + \tfrac{1}{12}\,\varphi\,\lambda_1\right] \quad .\end{aligned} \tag{6.237}$$

With $\varphi = \frac{1}{6}\pi a^3 n$ and using the definitions of $\lambda_{1,2}$ in eqn. 6.234, one finds

$$1 - n\,\hat{c}(0) = \frac{1+4\varphi+4\varphi^2}{(1-\varphi)^4} \quad . \tag{6.238}$$

We then have, from the compressibility equation,

$$\frac{6k_{\rm B}T}{\pi a^3}\,\frac{\partial p}{\partial \varphi} = \frac{1+4\varphi+4\varphi^2}{(1-\varphi)^4} \quad . \tag{6.239}$$

Integrating, we obtain $p(\varphi)$ up to a constant. The constant is set so that $p = 0$ when $n = 0$. The result is eqn. 6.235.

[22] Thus $\varphi = \frac{4}{3}\pi s^3 n$ is the mean number of particles within a hard sphere volume, with the hard sphere radius given by $s = \frac{1}{2}a$.

| quantity | exact | PY | HNC |
|---|---|---|---|
| $B_4/B_2^3$ | 0.28695 | 0.2969 | 0.2092 |
| $B_5/B_2^4$ | 0.1103 | 0.1211 | 0.0493 |
| $B_6/B_2^5$ | 0.0386 | 0.0281 | 0.0449 |
| $B_7/B_2^6$ | 0.0138 | 0.0156 | – |

Table 6.1: Comparison of exact (Monte Carlo) results to those of the Percus-Yevick (PY) and hypernetted chains approximation (HNC) for hard spheres in three dimensions. *Sources*: Hansen and McDonald (1990) and Reichl (1998)

Another commonly used scheme is the *hypernetted chains* (HNC) approximation, for which

$$c(\boldsymbol{r}) = -\beta u(\boldsymbol{r}) + h(\boldsymbol{r}) - \log\Big(1 + h(\boldsymbol{r})\Big) \quad . \tag{6.240}$$

The rationale behind the HNC and other such approximation schemes is rooted in diagrammatic approaches, which are extensions of the Mayer cluster expansion to the computation of correlation functions. For details and references to their application in the literature, see Hansen and McDonald (1990) and Reichl (1998).

## 6.5 Coulomb Systems : Plasmas and the Electron Gas

### 6.5.1 Electrostatic potential

Coulomb systems are particularly interesting in statistical mechanics because of their long-ranged forces, which result in the phenomenon of *screening*. Long-ranged forces wreak havoc with the Mayer cluster expansion, since the Mayer function is no longer integrable. Thus, the virial expansion fails, and new techniques need to be applied to reveal the physics of plasmas.

The potential energy of a Coulomb system is

$$U = \tfrac{1}{2}\int d^d r \int d^d r'\, \rho(\boldsymbol{r})\, u(\boldsymbol{r} - \boldsymbol{r}')\, \rho(\boldsymbol{r}') \quad , \tag{6.241}$$

where $\rho(\boldsymbol{r})$ is the charge density and $u(\boldsymbol{r})$, which has the dimensions of (energy)/(charge)$^2$, satisfies

$$\nabla^2 u(\boldsymbol{r} - \boldsymbol{r}') = -4\pi\, \delta(\boldsymbol{r} - \boldsymbol{r}') \quad . \tag{6.242}$$

Thus,

$$u(\boldsymbol{r}) = \begin{cases} -2\pi\, |x - x'| & , \quad d = 1 \\ -2\, \log |\boldsymbol{r} - \boldsymbol{r}'| & , \quad d = 2 \\ |\boldsymbol{r} - \boldsymbol{r}'|^{-1} & , \quad d = 3 \quad . \end{cases} \tag{6.243}$$

For discrete particles, the charge density $\rho(\boldsymbol{r})$ is given by

$$\rho(\boldsymbol{r}) = \sum_i q_i \, \delta(\boldsymbol{r} - \boldsymbol{x}_i) \quad , \tag{6.244}$$

where $q_i$ is the charge of the $i^{\rm th}$ particle. We will assume two types of charges: $q = \pm e$, with $e > 0$. Such a system is called a *plasma*. The electric potential is

$$\phi(\boldsymbol{r}) = \int d^d r' \, u(\boldsymbol{r} - \boldsymbol{r}') \, \rho(\boldsymbol{r}') = \sum_i q_i \, u(\boldsymbol{r} - \boldsymbol{x}_i) \quad . \tag{6.245}$$

This satisfies the Poisson equation, $\nabla^2 \phi(\boldsymbol{r}) = -4\pi\rho(\boldsymbol{r})$ . The total potential energy can be written as

$$U = \tfrac{1}{2} \int d^d r \, \phi(\boldsymbol{r}) \, \rho(\boldsymbol{r}) = \tfrac{1}{2} \sum_i q_i \, \phi(\boldsymbol{x}_i) \quad , \tag{6.246}$$

where it is understood that we omit self-interaction terms.

### 6.5.2 Debye-Hückel theory

#### A simple 'derivation'

For a gas of noninteracting classical particles with dispersion $\varepsilon(\boldsymbol{k})$, the number density in the GCE is given by

$$n(T, z) = z \int \frac{d^d k}{(2\pi)^d} \, e^{-\varepsilon(\boldsymbol{k})/k_{\rm B}T} \equiv z \lambda_T^{-d} \quad . \tag{6.247}$$

For the ballistic dispersion $\varepsilon(\boldsymbol{k}) = \hbar^2 \boldsymbol{k}^2 / 2m$ one has $\lambda_T = \sqrt{2\pi\hbar^2 / m k_{\rm B} T}$ . Suppose the particles each have charge $q$, and are in the presence of a constant external potential $\phi$. Then the internal energy of each particle is $q\phi$ and $n(T, z) = z\lambda_T^{-d} \exp(-q\phi / k_{\rm B} T)$, which is obvious after replacing $\varepsilon(\boldsymbol{k}) \to \varepsilon(\boldsymbol{k}) + q\phi$. Now let us suppose $\phi(\boldsymbol{r})$ varies smoothly in space. We then have that the *local* number density is

$$n(\boldsymbol{r}) = z \lambda_T^{-d} \exp(-q \, \phi(\boldsymbol{r}) / k_{\rm B} T) \quad . \tag{6.248}$$

If there are two species of charges, with $q = \pm e$, then the local charge density is

$$\rho(\boldsymbol{r}) = e n_+(\boldsymbol{r}) - e n_-(\boldsymbol{r}) + \rho_{\rm ext}(\boldsymbol{r}) \quad , \tag{6.249}$$

where $n_\pm(\boldsymbol{r}) = z_\pm \lambda_\pm^{-d} \exp(\mp e\phi(\boldsymbol{r}) / k_{\rm B} T)$ and where $\rho_{\rm ext}$ is a fixed external charge density[23]. Now we close the loop and acknowledge that the total charge density $\rho(\boldsymbol{r})$ is the *source* of the electric potential $\phi(\boldsymbol{r})$, which entails Poisson's equation. We then arrive at the coupled system of equations,

$$\begin{aligned} \rho(\boldsymbol{r}) &= e z_+ \lambda_+^{-d} \, e^{-e\phi(\boldsymbol{r})/k_{\rm B}T} - e z_- \lambda_-^{-d} \, e^{+e\phi(\boldsymbol{r})/k_{\rm B}T} + \rho_{\rm ext}(\boldsymbol{r}) \\ \nabla^2 \phi(\boldsymbol{r}) &= -4\pi\rho(\boldsymbol{r}) \quad . \end{aligned} \tag{6.250}$$

These are the nonlinear Poisson-Boltzmann equations. In the weak potential limit, their linearization yields what is known as Debye-Hückel theory (see below).

[23] Note that there are two separate chemical potentials $\mu_\pm$ and in principle two independent thermal wavelengths $\lambda_\pm$. But there is one temperature $T$ to rule them all.

**Mean field approach**

Formally, the grand partition function for our plasma is given by

$$\Xi(T,V,\mu_+,\mu_-) = \sum_{N_+=0}^{\infty}\sum_{N_-=0}^{\infty} \frac{(z_+\lambda_+^{-d})^{N_+}}{N_+!}\cdot\frac{(z_-\lambda_-^{-d})^{N_-}}{N_-!}\cdot\int d^dx_1\cdots\int d^dx_N\, e^{-\beta U(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)} \quad , \tag{6.251}$$

where $z_\pm = \exp(\mu_\pm/k_{\rm B}T)$ are the fugacities, $\lambda_\pm = (2\pi\hbar^2/m_\pm k_{\rm B}T)^{1/2}$ are the thermal wavelengths, $N = N_+ + N_-$ is the total number of particles in the $(N_+,N_-)$ sector of the double sum, and

$$U(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_N) = \sum_{i<j}^{N} q_i\,q_j\,u(\boldsymbol{x}_i-\boldsymbol{x}_j) \tag{6.252}$$

is the total potential energy. The many body integral in position space is in general intractable. What's more, the Mayer function

$$f_{ij} = e^{-\beta\,q_iq_ju(r_{ij})} - 1 \tag{6.253}$$

behaves as $f_{ij}\approx -\beta\,q_iq_ju(r_{ij})$ as $r_{ij}\to\infty$, which is not integrable. Thus it would prove a tricky business indeed to attempt to derive a cluster expansion as we did in §6.3.3 for the equation of state[24].

Here we adopt a *mean field* approach, writing

$$\rho(\boldsymbol{r}) = \bar\rho(\boldsymbol{r}) + \delta\rho(\boldsymbol{r}) \qquad , \qquad \phi(\boldsymbol{r}) = \bar\phi(\boldsymbol{r}) + \delta\phi(\boldsymbol{r}) \quad . \tag{6.254}$$

We then have

$$\begin{aligned} U &= \tfrac{1}{2}\int d^dr\,\big[\bar\rho(\boldsymbol{r})+\delta\rho(\boldsymbol{r})\big]\cdot\big[\bar\phi(\boldsymbol{r})+\delta\phi(\boldsymbol{r})\big] \\ &= \tfrac{1}{2}\int d^dr\,\bar\phi(\boldsymbol{r})\,\bar\rho(\boldsymbol{r}) + \overbrace{\tfrac{1}{2}\int d^dr\,\bar\phi(\boldsymbol{r})\,\delta\rho(\boldsymbol{r}) + \tfrac{1}{2}\int d^dr\,\bar\rho(\boldsymbol{r})\,\delta\phi(\boldsymbol{r})}^{\text{these two terms are the same}} + \tfrac{1}{2}\int d^dr\,\delta\rho(\boldsymbol{r})\,\delta\phi(\boldsymbol{r}) \\ &= -\tfrac{1}{2}\int d^dr\,\bar\phi(\boldsymbol{r})\,\bar\rho(\boldsymbol{r}) + \int d^dr\,\bar\phi(\boldsymbol{r})\,\rho(\boldsymbol{r}) + \overbrace{\tfrac{1}{2}\int d^dr\,\delta\rho(\boldsymbol{r})\,\delta\phi(\boldsymbol{r})}^{\text{ignore fluctuation term}} \\ &= -U_0 + \int d^dr\,\bar\phi(\boldsymbol{r})\,\rho(\boldsymbol{r}) + (\text{fluctuations})^2 \quad , \end{aligned} \tag{6.255}$$

where $U_0 \equiv \frac{1}{2}\int d^dr\,\bar\phi(\boldsymbol{r})\,\bar\rho(\boldsymbol{r})$, and where we have used $\delta\rho = \rho - \bar\rho$. Thus we have

$$\Xi = e^{U_0/k_{\rm B}T}\exp\left(z_+\lambda_+^{-d}\int d^dr_+\,e^{-e\,\bar\phi(\boldsymbol{r}_+)/k_{\rm B}T}\right)\exp\left(z_-\lambda_-^{-d}\int d^dr_-\,e^{e\,\bar\phi(\boldsymbol{r}_-)/k_{\rm B}T}\right) \tag{6.256}$$

whence

$$\Omega(T,V,\mu_+,\mu_-) = -U_0 - k_{\rm B}Tz_+\lambda_+^{-d}\int d^dr\,\exp\left(-\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T}\right) - k_{\rm B}Tz_-\lambda_-^{-d}\int d^dr\,\exp\left(+\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T}\right) \quad , \tag{6.257}$$

[24] One can make progress in this direction, but it requires resumming infinite classes of diagrams.

where

$$\lambda_\pm = \left(\frac{2\pi\hbar^2}{m_\pm k_{\rm B} T}\right) \quad , \quad z_\pm = \exp\left(\frac{\mu_\pm}{k_{\rm B} T}\right) \quad . \tag{6.258}$$

Note that since $\bar\phi(\boldsymbol{r}) = \int d^d r'\, u(\boldsymbol{r}-\boldsymbol{r}')\,\bar\rho(\boldsymbol{r}')$ is a linear functional of $\bar\rho(\boldsymbol{r})$, we have

$$\frac{\delta U_0}{\delta\bar\phi(\boldsymbol{r})} = \bar\rho(\boldsymbol{r}) \quad . \tag{6.259}$$

We next demand that the free energy $\Omega$ is extremized with respect to the mean field $\bar\phi(\boldsymbol{r})$, *viz.*

$$0 = \frac{\delta\Omega}{\delta\bar\phi(\boldsymbol{r})} = -\bar\rho(\boldsymbol{r}) + e\,\lambda_+^{-d}\,z_+\,\exp\left(-\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T}\right) - e\,\lambda_-^{-d}\,z_-\,\exp\left(+\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T}\right) \quad . \tag{6.260}$$

This yields a mean field charge density $\bar\rho(\boldsymbol{r})$ which agrees with that in eqn. 6.250. At $r\to\infty$, we assume charge neutrality and $\phi(\infty)=0$. Thus

$$\lambda_+^{-d}\,z_+ = n_+(\infty) = \lambda_-^{-d}\,z_- = n_-(\infty) \equiv n_\infty \quad , \tag{6.261}$$

where $n_\infty$ is the ionic density of either species at infinity. Therefore,

$$\rho(\boldsymbol{r}) = -2e\,n_\infty \sinh\left(\frac{e\,\phi(\boldsymbol{r})}{k_{\rm B}T}\right) \quad , \tag{6.262}$$

where we have dropped the bars on $\bar\phi$ and $\bar\rho$ for convenience. We now invoke Poisson's equation,

$$\nabla^2\phi = 8\pi e n_\infty \sinh(\beta e\phi) - 4\pi\rho_{\rm ext} \quad , \tag{6.263}$$

where $\rho_{\rm ext}$ is an externally imposed charge density.

If $|e\phi| \ll k_{\rm B}T$, we can expand the $\sinh$ function and obtain

$$\nabla^2\phi = \kappa_{\rm D}^2\,\phi - 4\pi\rho_{\rm ext} \quad , \tag{6.264}$$

where

$$\kappa_{\rm D} = \left(\frac{8\pi n_\infty e^2}{k_{\rm B}T}\right)^{1/2} \quad , \quad \lambda_{\rm D} = \left(\frac{k_{\rm B}T}{8\pi n_\infty e^2}\right)^{1/2} \quad . \tag{6.265}$$

The quantity $\lambda_{\rm D}$ is known as the *Debye screening length*. Consider, for example, a point charge $Q$ located at the origin. We then solve Poisson's equation in the weak field limit,

$$\nabla^2\phi = \kappa_{\rm D}^2\,\phi - 4\pi Q\,\delta(\boldsymbol{r}) \quad . \tag{6.266}$$

Fourier transforming, we obtain

$$-\boldsymbol{q}^2\,\hat\phi(\boldsymbol{q}) = \kappa_{\rm D}^2\,\hat\phi(\boldsymbol{q}) - 4\pi Q \qquad \Rightarrow \qquad \hat\phi(\boldsymbol{q}) = \frac{4\pi Q}{\boldsymbol{q}^2+\kappa_{\rm D}^2} \quad . \tag{6.267}$$

Transforming back to real space, we obtain, in three dimensions, the Yukawa potential,

$$\phi(\boldsymbol{r}) = \int\frac{d^3q}{(2\pi)^3}\,\frac{4\pi Q\,e^{i\boldsymbol{q}\cdot\boldsymbol{r}}}{\boldsymbol{q}^2+\kappa_{\rm D}^2} = \frac{Q}{r}\,e^{-\kappa_{\rm D} r} \quad . \tag{6.268}$$

For the charge density,

$$\hat{\rho}(\boldsymbol{q}) = \frac{\boldsymbol{q}^2}{4\pi}\,\hat{\phi}(\boldsymbol{q}) = \frac{Q\,\boldsymbol{q}^2}{\boldsymbol{q}^2 + \kappa_{\rm D}^2} \quad \Rightarrow \quad \rho(\boldsymbol{r}) = Q\left(\delta(\boldsymbol{r}) - \frac{\kappa_{\rm D}^2}{4\pi r}\,e^{-\kappa_{\rm D} r}\right) \quad . \tag{6.269}$$

Note that $\hat{\rho}(0) = 0$, which says there is total charge neutrality: $\int d^3r\,\rho(\boldsymbol{r}) = \hat{\rho}(0) = 0$. Thus, the plasma perfectly screens the external charge. This solution must however break down sufficiently close to $r = 0$, where the assumption $e\phi(r) \ll k_{\rm B}T$ is not valid. For larger $r$, the Yukawa form is increasingly accurate.

For another example, consider an electrolyte held between two conducting plates, one at an electric potential (*i.e.* voltage) $\phi(x = 0) = 0$ and the other at potential $\phi(x = L) = V$, where $\hat{\boldsymbol{x}}$ is normal to the plane of the plates. Again assuming a weak field $e\phi \ll k_{\rm B}T$, we solve $\nabla^2\phi = \kappa_{\rm D}^2\,\phi$ and obtain

$$\phi(x) = A\,e^{\kappa_{\rm D} x} + B\,e^{-\kappa_{\rm D}\,x} \quad . \tag{6.270}$$

We fix the constants $A$ and $B$ by invoking the boundary conditions, which results in

$$\phi(x) = \frac{V\sinh(\kappa_{\rm D}x)}{\sinh(\kappa_{\rm D}L)} \quad . \tag{6.271}$$

Debye-Hückel theory is valid provided $n_\infty\,\lambda_{\rm D}^3 \gg 1$, so that the statistical assumption of many charges in a screening volume is justified.

### 6.5.3 The electron gas : Thomas-Fermi screening

Assuming $k_{\rm B}T \ll \varepsilon_{\rm F}$, thermal fluctuations are unimportant and we may assume $T = 0$. In the same spirit as the Debye-Hückel approach, we assume a slowly varying mean electrostatic potential $\phi(\boldsymbol{r})$. Locally, we can write

$$\varepsilon_{\rm F} = \frac{\hbar^2 k_{\rm F}^2}{2m} - e\phi(\boldsymbol{r}) \quad . \tag{6.272}$$

Thus, the Fermi wavevector $k_{\rm F}$ is spatially varying, according to the relation

$$k_{\rm F}(\boldsymbol{r}) = \left[\frac{2m}{\hbar^2}\Big(\varepsilon_{\rm F} + e\phi(\boldsymbol{r})\Big)\right]^{1/2} \quad . \tag{6.273}$$

The local electron number density is

$$n(\boldsymbol{r}) = \frac{k_{\rm F}^3(\boldsymbol{r})}{3\pi^2} = n_\infty\left(1 + \frac{e\phi(\boldsymbol{r})}{\varepsilon_{\rm F}}\right)^{3/2} \quad . \tag{6.274}$$

In the presence of a uniform compensating positive background charge $\rho_+ = en_\infty$, Poisson's equation takes the form

$$\nabla^2\phi = 4\pi e\,n_\infty\left[\left(1 + \frac{e\phi(\boldsymbol{r})}{\varepsilon_{\rm F}}\right)^{3/2} - 1\right] - 4\pi\rho_{\rm ext}(\boldsymbol{r}) \quad . \tag{6.275}$$

If $|e\phi| \ll \varepsilon_{\rm F}$, we may expand in powers of the ratio, obtaining

$$\nabla^2\phi = \frac{6\pi n_\infty e^2}{\varepsilon_{\rm F}}\,\phi - 4\pi\rho_{\rm ext} \equiv \kappa_{\rm TF}^2\,\phi - 4\pi\rho_{\rm ext} \quad , \tag{6.276}$$

where $\kappa_{\rm TF}$ is the *Thomas-Fermi wavevector*,

$$\kappa_{\rm TF} = \left(\frac{6\pi n_\infty e^2}{\varepsilon_{\rm F}}\right)^{1/2} \quad . \tag{6.277}$$

Thomas-Fermi theory is valid provided $n_\infty\,\lambda_{\rm TF}^3 \gg 1$, where $\lambda_{\rm TF} = \kappa_{\rm TF}^{-1}$, so that the statistical assumption of many electrons in a screening volume is justified.

One important application of Thomas-Fermi screening is to the theory of metals. In a metal, the outer, valence electrons of each atom are stripped away from the positively charged ionic core and enter into itinerant, plane-wave-like states. These states disperse with some $\varepsilon(\boldsymbol{k})$ function (that is periodic in the Brillouin zone, *i.e.* under $\boldsymbol{k} \to \boldsymbol{k} + \boldsymbol{G}$, where $\boldsymbol{G}$ is a reciprocal lattice vector), and at $T = 0$ this *energy band* is filled up to the Fermi level $\varepsilon_{\rm F}$, as Fermi statistics dictates. (In some cases, there may be several bands at the Fermi level, as we saw in the case of yttrium.) The set of ionic cores then acts as a neutralizing positive background. In a perfect crystal, the ionic cores are distributed periodically, and the positive background is approximately uniform. A charged impurity in a metal, such as a zinc atom in a copper matrix, has a different nuclear charge and a different valency than the host. The charge of the ionic core, when valence electrons are stripped away, differs from that of the host ions, and therefore the impurity acts as a *local charge impurity*. For example, copper has an electronic configuration of $[{\rm Ar}]\,3d^{10}\,4s^1$. The 4s electron forms an energy band which contains the Fermi surface. Zinc has a configuration of $[{\rm Ar}]\,3d^{10}\,4s^2$, and in a Cu matrix the Zn gives up its two $4s$ electrons into the $4s$ conduction band, leaving behind a charge $+2$ ionic core. The Cu cores have charge $+1$ since each copper atom contributed only one $4s$ electron to the conduction band. The conduction band electrons neutralize the uniform positive background of the Cu ion cores. What is left is an extra $Q = +e$ nuclear charge at the Zn site, and one extra $4s$ conduction band electron. The $Q = +e$ impurity is, however, *screened* by the electrons, and at distances greater than an atomic radius the potential that a given electron sees due to the Zn core is of the Yukawa form,

$$\phi(\boldsymbol{r}) = \frac{Q}{r}\,e^{-\kappa_{\rm TF} r} \quad . \tag{6.278}$$

We should take care, however, that the dispersion $\varepsilon(\boldsymbol{k})$ for the conduction band in a metal is not necessarily of the free electron form $\varepsilon(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m$. To linear order in the potential, however, the change in the local electronic density is

$$\delta n(\boldsymbol{r}) = e\phi(\boldsymbol{r})\,g(\varepsilon_{\rm F}) \quad , \tag{6.279}$$

where $g(\varepsilon_{\rm F})$ is the density of states at the Fermi energy. Thus, in a metal, we should write

$$\nabla^2\phi = (-4\pi)(-e\,\delta n) = 4\pi e^2 g(\varepsilon_{\rm F})\,\phi = \kappa_{\rm TF}^2\,\phi \quad , \tag{6.280}$$

where

$$\kappa_{\rm TF} = \sqrt{4\pi e^2\,g(\varepsilon_{\rm F})} \quad . \tag{6.281}$$

The value of $g(\varepsilon_{\rm F})$ will depend on the form of the dispersion. For parabolic bands with an effective mass $m^*$, the formula in eqn. 6.276 still applies.

**The Thomas-Fermi atom**

Consider an ion formed of a nucleus of charge $+Ze$ and an electron cloud of charge $-Ne$. The net ionic charge is then $(Z-N)e$. Since we will be interested in atomic scales, we can no longer assume a weak field limit and we must retain the full nonlinear screening theory, for which

$$\nabla^2\phi(\boldsymbol{r}) = 4\pi e\cdot\frac{(2m)^{3/2}}{3\pi^2\hbar^3}\left(\varepsilon_{\rm F}+e\phi(\boldsymbol{r})\right)^{3/2} - 4\pi Ze\,\delta(\boldsymbol{r}) \quad . \tag{6.282}$$

We assume an isotropic solution. It is then convenient to define

$$\varepsilon_{\rm F}+e\phi(\boldsymbol{r}) = \frac{Ze^2}{r}\cdot\chi(r/r_0) \quad , \tag{6.283}$$

where $r_0$ is yet to be determined. As $r\to 0$ we expect $\chi\to 1$ since the nuclear charge is then unscreened. We then have

$$\nabla^2\left\{\frac{Ze^2}{r}\cdot\chi(r/r_0)\right\} = \frac{1}{r_0^2}\,\frac{Ze^2}{r}\,\chi''(r/r_0) \quad , \tag{6.284}$$

thus we arrive at the *Thomas-Fermi equation*,

$$\chi''(t) = \frac{1}{\sqrt{t}}\,\chi^{3/2}(t) \quad , \tag{6.285}$$

with $r = t\,r_0$, provided we take

$$r_0 = \frac{\hbar^2}{2me^2}\left(\frac{3\pi}{4\sqrt{Z}}\right)^{2/3} = 0.885\,Z^{-1/3}\,a_{\rm B} \quad , \tag{6.286}$$

where $a_{\rm B} = \hbar^2/me^2 = 0.529\,\text{Å}$ is the Bohr radius. The TF equation is subject to the following boundary conditions:

- At short distances, the nucleus is unscreened, *i.e.* $\chi(0)=1$.
- For positive ions, with $N<Z$, there is perfect screening at the ionic boundary $R = t^*\,r_0$, where $\chi(t^*)=0$. This requires

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi = \left[-\frac{Ze}{R^2}\,\chi(R/r_0) + \frac{Ze}{R\,r_0}\,\chi'(R/r_0)\right]\hat{\boldsymbol{r}} = \frac{(Z-N)\,e}{R^2}\,\hat{\boldsymbol{r}} \quad . \tag{6.287}$$

  This requires

$$-t^*\,\chi'(t^*) = 1 - \frac{N}{Z} \quad . \tag{6.288}$$

For an atom, with $N=Z$, the asymptotic solution to the TF equation is a power law, and by inspection is found to be $\chi(t)\sim C\,t^{-3}$, where $C$ is a constant. The constant follows from the TF equation, which yields $12\,C = C^{3/2}$, hence $C=144$. Thus, a neutral TF atom has a density with a power law tail, with $\rho\sim r^{-6}$. TF ions with $N>Z$ are unstable.

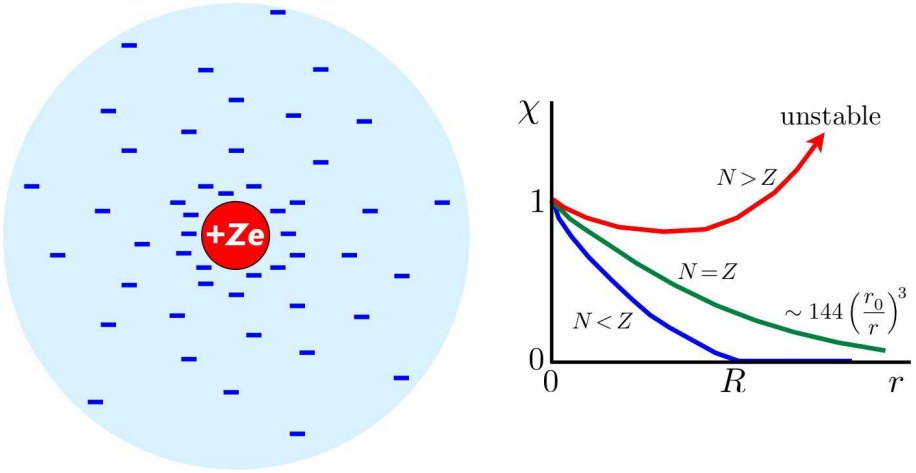



Figure 6.19: The Thomas-Fermi atom consists of a nuclear charge $+Ze$ surrounded by $N$ electrons distributed in a cloud. The electric potential $\phi(\boldsymbol{r})$ felt by any electron at position $\boldsymbol{r}$ is screened by the electrons within this radius, resulting in a self-consistent potential $\phi(\boldsymbol{r}) = \phi_0 + (Ze/r)\,\chi(r/r_0)$.

## 6.6 Polymers

### 6.6.1 Basic concepts

Linear chain polymers are repeating structures with the chemical formula $(A)_x$, where $A$ is the *formula unit* and $x$ is the *degree of polymerization*. In many cases (*e.g.* polystyrene), $x \gtrsim 10^5$ is not uncommon. For a very readable introduction to the subject, see P. G. de Gennes, *Scaling Concepts in Polymer Physics*.

Quite often a given polymer solution will contain a distribution of $x$ values; this is known as *polydispersity*. Various preparation techniques, such as chromatography, can mitigate the degree of polydispersity. Another morphological feature of polymers is *branching*, in which the polymers do not form linear chains.

Polymers exhibit a *static flexibility* which can be understood as follows. Consider a long chain hydrocarbon with a $-\mathrm{C}-\mathrm{C}-\mathrm{C}-$ backbone. The angle between successive $\mathrm{C}-\mathrm{C}$ bonds is fixed at $\theta \approx 68^\circ$, but the azimuthal angle $\varphi$ can take one of three possible low-energy values, as shown in the right panel of fig. 6.21. Thus, the relative probabilities of *gauche* and *trans* orientations are

$$\frac{\text{Prob (gauche)}}{\text{Prob (trans)}} = 2\,e^{-\Delta\varepsilon/k_{\rm B}T} \tag{6.289}$$

where $\Delta\varepsilon$ is the energy difference between *trans* and *gauche* configurations. This means that the polymer chain is in fact a *random coil* with a *persistence length*

$$\ell_{\rm p} = \ell_0\,e^{\Delta\varepsilon/k_{\rm B}T} \tag{6.290}$$

where $\ell_0$ is a microscopic length scale, roughly given by the length of a formula unit, which is approximately a few Ångstroms (see fig. 6.22). Let $L$ be the total length of the polymer when it is stretched into

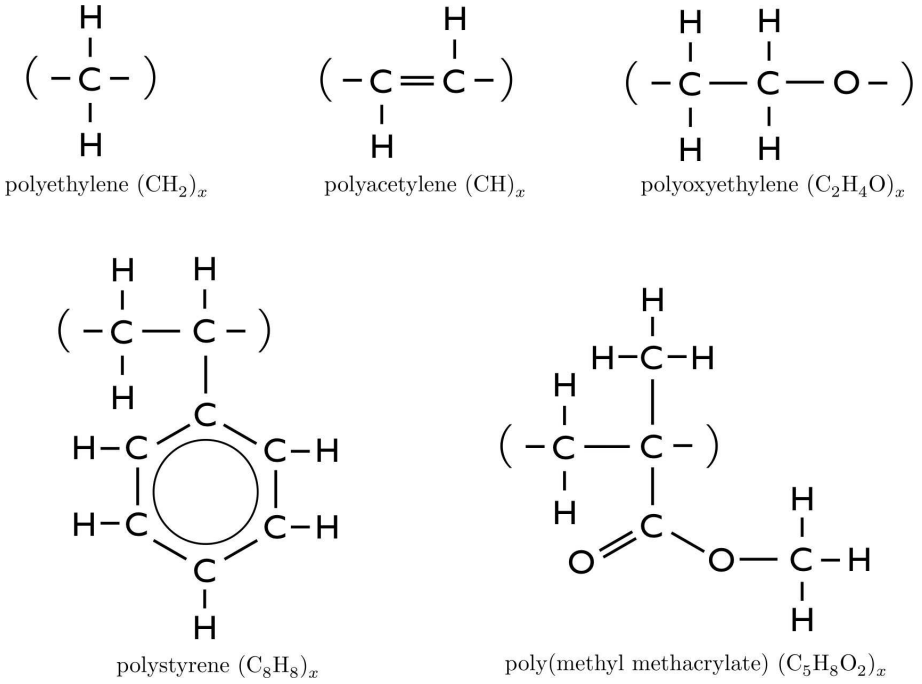



Figure 6.20: Some examples of linear chain polymers.

a straight line. If $\ell_{\rm p} > L$, the polymer is *rigid*. If $\ell_{\rm p} \ll L$, the polymer is rigid on the length scale $\ell_{\rm p}$ but flexible on longer scales. We have

$$\frac{\ell_{\rm p}}{L} = \frac{1}{N}\, e^{\Delta\varepsilon/k_{\rm B}T} \quad , \tag{6.291}$$

where we now use $N$ (rather than $x$) for the degree of polymerization.

In the time domain, the polymer exhibits a *dynamical flexibility* on scales longer than a *persistence time*. The persistence time $\tau_{\rm p}$ is the time required for a *trans-gauche* transition. The rate for such transitions is set by the energy barrier $B$ separating *trans* from *gauche* configurations:

$$\tau_{\rm p} = \tau_0\, e^{B/k_{\rm B}T} \tag{6.292}$$

where $\tau_0 \sim 10^{-11}\,{\rm s}$. On frequency scales $\omega \ll \tau_{\rm p}^{-1}$ the polymer is dynamically flexible. If $\Delta\varepsilon \sim k_{\rm B}T \ll B$ the polymer is flexible from a static point of view, but dynamically rigid. That is, there are many *gauche* orientations of successive carbon bonds which reflect a quenched disorder. The polymer then forms a frozen random coil, like a twisted coat hanger.

### 6.6.2 Polymers as random walks

A polymer can be modeled by a *self-avoiding random walk* (SAW). That is, on scales longer than $\ell_{\rm p}$, it twists about randomly in space subject to the constraint that it doesn't overlap itself. Before we consider the mathematics of SAWs, let's first recall some aspects of ordinary random walks which are not self-avoiding.

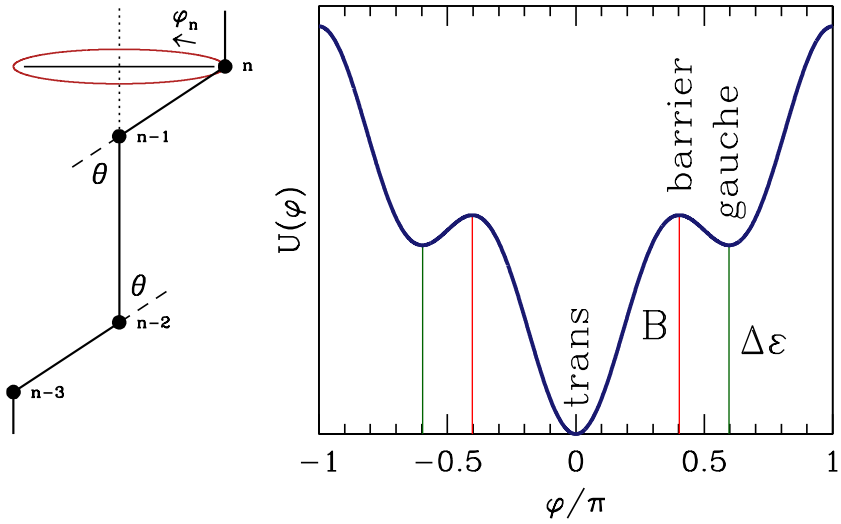



Figure 6.21: Left: *trans* and *gauche* orientations in carbon chains. Right: energy as a function of azimuthal angle $\varphi$. There are three low energy states: *trans* ($\varphi = 0$) and *gauche* ($\varphi = \pm\varphi_0$).

We'll simplify matters further by considering random walks on a hypercubic lattice of dimension $d$. Such a lattice has coordination number $2d$, *i.e.* there are $2d$ nearest neighbor separation vectors, given by $\boldsymbol{\delta} = \pm a\,\hat{\mathbf{e}}_1\,, \pm a\,\hat{\mathbf{e}}_2\,, \ldots\,, \pm a\,\hat{\mathbf{e}}_d$, where $a$ is the lattice spacing. Consider now a random walk of $N$ steps starting at the origin. After $N$ steps the position is $\sum_{j=1}^N \boldsymbol{\delta}_j$, where each $\boldsymbol{\delta}_j$ takes on one of $2d$ possible values. Now $N$ is no longer the degree of polymerization, but something approximating $L/\ell_{\rm p}$, which is the number of persistence lengths in the chain. We assume each step is independent, hence $\langle \delta_j^\alpha\, \delta_{j'}^\beta \rangle = (a^2/d)\,\delta_{jj'}\delta^{\alpha\beta}$ and $\langle \boldsymbol{R}_N^2 \rangle = Na^2$. The full distribution $P_N(\boldsymbol{R})$ is given by

$$\begin{aligned} P_N(\boldsymbol{R}) &= (2d)^{-N} \sum_{\boldsymbol{\delta}_1} \cdots \sum_{\boldsymbol{\delta}_N} \delta_{\boldsymbol{R},\sum_j \boldsymbol{\delta}_j} \\ &= a^d \int\limits_{-\pi/a}^{\pi/a} \frac{dk_1}{2\pi} \cdots \int\limits_{-\pi/a}^{\pi/a} \frac{dk_d}{2\pi}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \left[ \frac{1}{d} \sum_{\mu=1}^{d} \cos(k_\mu a) \right]^N \\ &= a^d \int\limits_{\hat{\Omega}} \frac{d^dk}{(2\pi)^d}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \exp\left[ N \log\left( 1 - \frac{1}{2d}\,\boldsymbol{k}^2 a^2 + \ldots \right) \right] \\ &\approx \left( \frac{a}{2d} \right)^d \int d^dk\, e^{-N\boldsymbol{k}^2 a^2/2d}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} = \left( \frac{d}{2\pi N} \right)^{d/2} e^{-d\boldsymbol{R}^2/2Na^2} \quad . \end{aligned} \tag{6.293}$$

This is a simple Gaussian, with width $\langle \boldsymbol{R}^2 \rangle = d\cdot(Na^2/d) = Na^2$, as we have already computed. The quantity $\boldsymbol{R}$ defined here is the *end-to-end vector* of the chain. The RMS end-to-end distance is then $\langle \boldsymbol{R}^2 \rangle^{1/2} = \sqrt{N}a \equiv R_0$.

A related figure of merit is the *radius of gyration*, $R_{\rm g}$, defined by

$$R_{\rm g}^2 = \frac{1}{N}\Big\langle \sum_{n=1}^{N} \big(\boldsymbol{R}_n - \boldsymbol{R}_{\rm CM}\big)^2 \Big\rangle \quad , \tag{6.294}$$

where $R_{\rm CM} = \frac{1}{N}\sum_{j=1}^{N} \boldsymbol{R}_j$ is the center of mass position. A brief calculation yields

$$R_g^2 = \big(N + 3 - 4N^{-1}\big)\, a^2 \sim \frac{Na^2}{6} \quad , \tag{6.295}$$

in all dimensions.

The total number of random walk configurations with end-to-end vector $\boldsymbol{R}$ is then $(2d)^N P_N(\boldsymbol{R})$, so the entropy of a chain at fixed elongation is

$$S(\boldsymbol{R}, N) = k_{\rm B} \log\Big[(2d)^N P_N(\boldsymbol{R})\Big] = S(0,N) - \frac{dk_{\rm B}\boldsymbol{R}^2}{2Na^2} \quad . \tag{6.296}$$

If we assume that the energy of the chain is conformation independent, then $E = E_0(N)$ and

$$F(\boldsymbol{R}, N) = F(0,N) + \frac{dk_{\rm B}T\boldsymbol{R}^2}{2Na^2} \quad . \tag{6.297}$$

In the presence of an external force $F_{\rm ext}$, the Gibbs free energy is the Legendre transform

$$G(\boldsymbol{F}_{\rm ext}, N) = F(\boldsymbol{R}, N) - \boldsymbol{F}_{\rm ext}\cdot\boldsymbol{R} \quad , \tag{6.298}$$

and $\partial G/\partial \boldsymbol{R} = 0$ then gives the relation

$$\big\langle \boldsymbol{R}(\boldsymbol{F}_{\rm ext}, N)\big\rangle = \frac{Na^2}{dk_{\rm B}T}\,\boldsymbol{F}_{\rm ext} \quad . \tag{6.299}$$

This may be considered an equation of state for the polymer.

Following de Gennes, consider a chain with charges $\pm e$ at each end, placed in an external electric field of magnitude $E = 30,000\,{\rm V/cm}$. Let $N = 10^4$, $a = 2\,$Å, and $d = 3$. What is the elongation? From the above formula, we have

$$\frac{R}{R_0} = \frac{eER_0}{3k_{\rm B}T} = 0.8 \quad , \tag{6.300}$$

with $R_0 = \sqrt{N}a$ as before.

**Structure factor**

We can also compute the structure factor,

$$S(\boldsymbol{k}) = \frac{1}{N}\Big\langle \sum_{m=1}^{N}\sum_{n=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)}\Big\rangle = 1 + \frac{2}{N}\sum_{m=1}^{N}\sum_{n=1}^{m-1}\Big\langle e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)}\Big\rangle \quad . \tag{6.301}$$

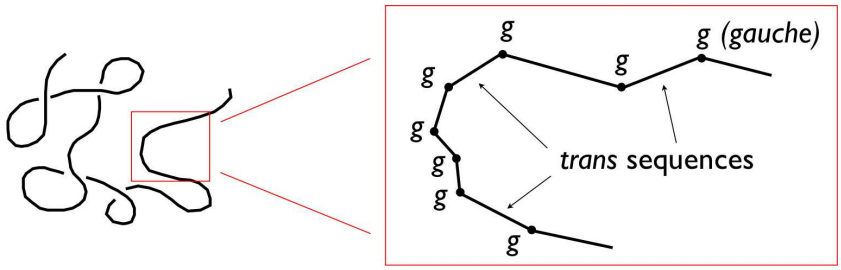



Figure 6.22: The polymer chain as a random coil.

For averages with respect to a Gaussian distribution,

$$\left\langle e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)}\right\rangle = \exp\left\{-\frac{1}{2}\left\langle\big(\boldsymbol{k}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)\big)^2\right\rangle\right\} \quad . \tag{6.302}$$

Now for $m > n$ we have $\boldsymbol{R}_m - \boldsymbol{R}_n = \sum_{j=n+1}^{m} \boldsymbol{\delta}_j$ , and therefore

$$\left\langle\big(\boldsymbol{k}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)\big)^2\right\rangle = \sum_{j=n+1}^{m}\left\langle(\boldsymbol{k}\cdot\boldsymbol{\delta}_j)^2\right\rangle = \frac{1}{d}\,(m-n)\,\boldsymbol{k}^2a^2 \quad , \tag{6.303}$$

since $\langle\delta_j^\alpha\,\delta_{j'}^\beta\rangle = (a^2/d)\,\delta_{jj'}\delta^{\alpha\beta}$ . We then have

$$S(\boldsymbol{k}) = 1 + \frac{2}{N}\sum_{m=1}^{N}\sum_{n=1}^{m-1} e^{-(m-n)\,\boldsymbol{k}^2a^2/2d} = \frac{N\,(e^{2\mu_{\boldsymbol{k}}}-1) - 2\,e^{\mu_{\boldsymbol{k}}}\,(1-e^{-N\mu_{\boldsymbol{k}}})}{N\big(e^{\mu_{\boldsymbol{k}}}-1\big)^2} \quad , \tag{6.304}$$

where $\mu_{\boldsymbol{k}} = \boldsymbol{k}^2a^2/2d$. In the limit where $N\to\infty$ and $a\to 0$ with $Na^2 = R_0^2$ constant, the structure factor has a *scaling form*, $S(\boldsymbol{k}) = Nf(N\mu_{\boldsymbol{k}}) = (R_0/a)^2\,f(k^2R_0^2/2d)$, where

$$f(x) = \frac{2}{x^2}\left(e^{-x}-1+x\right) = 1 - \frac{x}{3} + \frac{x^2}{12} + \ldots \quad . \tag{6.305}$$

### 6.6.3 The Rouse model

Consider next a polymer chain subjected to stochastic forcing. We model the chain as a collection of mass points connected by springs, with a potential energy $U = \frac{1}{2}k\sum_n\big(\boldsymbol{x}_{n+1}-\boldsymbol{x}_n\big)^2$. This reproduces the distribution of eqn. 6.293 if we take the spring constant to be $k = 3k_{\rm B}T/a^2$ and set the equilibrium length of each spring to zero. The equations of motion are then

$$M\ddot{\boldsymbol{x}}_n + \gamma\,\dot{\boldsymbol{x}}_n = -k\big(2\boldsymbol{x}_n - \boldsymbol{x}_{n-1} - \boldsymbol{x}_{n+1}\big) + \boldsymbol{f}_n(t) \quad , \tag{6.306}$$

where $n\in\{1,\ldots,N\}$ and $\{f_n^\mu(t)\}$ a set of Gaussian white noise forcings, each with zero mean, and

$$\left\langle f_n^\mu(t)\,f_{n'}^\nu(t')\right\rangle = 2\gamma k_{\rm B}T\,\delta_{nn'}\,\delta^{\mu\nu}\,\delta(t-t') \quad . \tag{6.307}$$

We define $\boldsymbol{x}_0 \equiv \boldsymbol{x}_1$ and $\boldsymbol{x}_{N+1} \equiv \boldsymbol{x}_N$ so that the end mass points $n = 1$ and $n = N$ experience a restoring force from only one neighbor. We assume the chain is overdamped and set $M \to 0$. We then have

$$\gamma\,\dot{\boldsymbol{x}}_n = -k \sum_{n'=1}^{N} A_{nn'}\,\boldsymbol{x}_{n'} + \boldsymbol{f}_n(t) \quad , \tag{6.308}$$

where

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ 0 & 0 & -1 & \ddots & \cdots & \vdots \\ \vdots & & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{pmatrix} \quad . \tag{6.309}$$

The matrix $A$ is real and symmetric. Its eigenfunctions are labeled $\psi_j(n)$, with $j \in \{0, \ldots, N-1\}$. The $j = 0$ eigenfunction is a constant, $\psi_0(n) = 1/\sqrt{N}$, and the others are given by

$$\psi_j(n) = \sqrt{\frac{2}{N}}\,\cos\!\left(\frac{(2n-1)j\pi}{2N}\right) \qquad , \qquad j \in \{1, \ldots, N-1\} \tag{6.310}$$

The completeness and orthonormality relations are

$$\sum_{j=0}^{N-1} \psi_j(n)\,\psi_j(n') = \delta_{nn'} \qquad , \qquad \sum_{n=1}^{N} \psi_j(n)\,\psi_{j'}(n) = \delta_{jj'} \quad , \tag{6.311}$$

with eigenvalues $\lambda_j = 4\sin^2\!\big(\pi j/2N\big)$. Note that $\lambda_0 = 0$.

We now work in the basis of normal modes $\{\eta_j^\mu\}$, where

$$\eta_j^\mu(t) = \sum_{n=1}^{N} \psi_j(n)\,x_n^\mu(t) \qquad , \qquad x_n^\mu(t) = \sum_{j=0}^{N-1} \psi_j(n)\,\eta_j^\mu(t) \quad . \tag{6.312}$$

We then have

$$\frac{d\boldsymbol{\eta}_j}{dt} = -\frac{1}{\tau_j}\,\boldsymbol{\eta}_j + \boldsymbol{g}_j(t) \quad , \tag{6.313}$$

where the $j^{\rm th}$ relaxation time is

$$\tau_j = \frac{\gamma}{4k\sin^2\!\big(\pi j/2N\big)} \tag{6.314}$$

and

$$g_j^\mu(t) = \gamma^{-1} \sum_{n=1}^{N} \psi_j(n)\,f_n^\mu(t) \quad . \tag{6.315}$$

Note that

$$\big\langle g_j^\mu(t)\, g_{j'}^\nu(t')\big\rangle = 2\gamma^{-1} k_{\rm B}T\, \delta_{jj'}\, \delta^{\mu\nu}\, \delta(t-t') \quad . \tag{6.316}$$

Integrating eqn. 6.313 we have, for $j = 0$,

$$\boldsymbol{\eta}_0(t) = \boldsymbol{\eta}_0(0) + \int_0^t dt'\, \boldsymbol{g}_0(t') \quad . \tag{6.317}$$

For the $j > 0$ modes,

$$\boldsymbol{\eta}_j(t) = \boldsymbol{\eta}_j(0)\, e^{-t/\tau_j} + \int_0^t dt'\, \boldsymbol{g}_j(t')\, e^{(t'-t)/\tau_j} \quad . \tag{6.318}$$

Thus,

$$\begin{aligned} \big\langle \eta_0^\mu(t)\, \eta_0^\nu(t')\big\rangle_{\rm c} &= 2\gamma^{-1} k_{\rm B}T\, \delta^{\mu\nu}\, \min(t,t') \\ \big\langle \eta_j^\mu(t)\, \eta_j^\nu(t')\big\rangle_{\rm c} &= \gamma^{-1} k_{\rm B}T\, \delta^{\mu\nu}\, \tau_j \left(e^{-|t-t'|/\tau_j} - e^{-(t+t')/\tau_j}\right) \quad , \end{aligned} \tag{6.319}$$

where the 'connected average' is defined to be $\langle A(t)\, B(t')\rangle_{\rm c} \equiv \langle A(t)\, B(t')\rangle - \langle A(t)\rangle\langle B(t')\rangle$. Transforming back to the original real space basis, we then have

$$\big\langle x_n^\mu(t)\, x_{n'}^\nu(t')\big\rangle_{\rm c} = \frac{2k_{\rm B}T}{N\gamma}\, \delta^{\mu\nu} \min(t,t') + \frac{k_{\rm B}T}{\gamma}\, \delta^{\mu\nu} \sum_{j=1}^{N-1} \tau_j\, \psi_j(n)\, \psi_j(n') \left(e^{-|t-t'|/\tau_j} - e^{-(t+t')/\tau_j}\right) \quad . \tag{6.320}$$

In particular, the 'connected variance' of $\boldsymbol{x}_n(t)$ is

$$\mathsf{CVar}\big[\boldsymbol{x}_n(t)\big] \equiv \big\langle \big[\boldsymbol{x}_n(t)\big]^2\big\rangle_{\rm c} = \frac{6k_{\rm B}T}{N\gamma}\, t + \frac{3k_{\rm B}T}{\gamma} \sum_{j=1}^{N-1} \tau_j \left[\psi_j(n)\right]^2 \left(1 - e^{-2t/\tau_j}\right) \quad . \tag{6.321}$$

From this we see that at long times, *i.e.* when $t \gg \tau_1$, the motion of $\boldsymbol{x}_n(t)$ is diffusive, with diffusion constant $D = k_{\rm B}T/N\gamma \propto B^{-1}$, which is inversely proportional to the chain length. Recall the Stokes result $\gamma = 6\pi\eta R$ for a sphere of radius $R$ moving in a fluid of dynamical viscosity $\eta$. From $D = k_{\rm B}T/\gamma$, shouldn't we expect the diffusion constant to be $D = k_{\rm B}T/6\pi\eta R \propto N^{-1/2}$, since the radius of gyration of the polymer is $R_{\rm g} \propto N^{1/2}$ ? This argument smuggles in the assumption that the only dissipation is taking place at the *outer surface* of the polymer, modeled as a ball of radius $R_{\rm g}$. In fact, for a Gaussian random walk in three space dimensions, the density for $r < R_{\rm g}$ is $\rho \propto N^{-1/2}$ since there are $N$ monomers inside a region of volume $N^{3/2}$. Accounting for Flory swelling due to steric interactions (see below), the density is $\rho \sim N^{-4/5}$, which is even smaller. So as $N \to \infty$, the density within the $r = R_{\rm g}$ effective sphere gets small, which means water molecules can easily penetrate, in which case the entire polymer chain should be considered to be in a dissipative environment, which is what the Rouse model says – each monomer executes overdamped motion.

A careful analysis of eqn. 6.321 reveals that there is a subdiffusive regime[25] where $\mathsf{CVar}\big[\boldsymbol{x}_n(t)\big] \propto t^{1/2}$. To see this, first take the $N \gg 1$ limit, in which case we may write $\tau_j = N^2\tau_0/j^2$, where $\tau_0 \equiv \gamma/\pi^2 k$ and

[25] I am grateful to Jonathan Lam and Olga Dudko for explaining this to me.

$j \in \{1, \ldots, N-1\}$. Let $s \equiv (n - \frac{1}{2})/N \in [0,1]$ be the scaled coordinate along the chain. The second term in eqn. 6.321 is then

$$S(s,t) \equiv \frac{6k_{\rm B}T}{\gamma} \cdot \frac{\tau_1}{N} \sum_{j=1}^{N-1} \frac{\cos^2(\pi j s)}{j^2} \left(1 - e^{-2j^2 t/\tau_1}\right) \quad . \tag{6.322}$$

Let $\sigma \equiv (t/\tau_1)^{1/2}$. When $t \ll \tau_1$, *i.e.* $\sigma \ll 1$, we have

$$S(s,t) \simeq \frac{6k_{\rm B}T}{\gamma} \cdot \frac{\tau_1}{N} \, \sigma \int_0^{N\sigma} du \, \frac{\cos^2(\pi u s/\sigma)}{u^2} \left(1 - e^{-2u^2}\right) \quad . \tag{6.323}$$

Since $s/\sigma \gg 1$, we may replace the cosine squared term by its average $\frac{1}{2}$. If we further assume $N\sigma \gg 1$, which means we are in the regime $1 \ll t/\tau_0 \ll N^2$, after performing the integral we obtain the result

$$S(s,t) = \frac{3k_{\rm B}T}{\gamma} \sqrt{2\pi\tau_0 t} \quad , \tag{6.324}$$

provided $s = \mathcal{O}(1)$, *i.e.* the site $n$ is not on either end of the chain. The result in eqn. 6.324 dominates the first term on the RHS of eqn. 6.321 since $\tau_0 \ll t \ll \tau_1$. This is the subdiffusive regime.

When $t \gg \tau_1 = N^2\tau_0$, the exponential on the RHS of eqn. 6.322 is negligible, and if we again approximate $\cos^2(\pi j s) \simeq \frac{1}{2}$, and we extend the upper limit on the sum to infinity, we obtain

$$S(t) = \frac{3k_{\rm B}T}{\gamma} \cdot \frac{\tau_1}{N} \cdot \frac{\pi^2}{6} \propto t^0 \quad , \tag{6.325}$$

which is dominated by the leading term on the RHS of eqn. 6.321. This is the diffusive regime, with $D = k_{\rm B}T/N\gamma$.

Finally, when $t \ll \tau_0$, the factor $1 - \exp(-2t/\tau_j)$ may be expanded to first order in $t$. One then obtains $\mathsf{CVar}\big[\boldsymbol{x}_n(t)\big] = (6k_{\rm B}T/\gamma)\, t$, which is independent of the force constant $k$. In this regime, the monomers don't have time to respond to the force from their neighbors, hence they each diffuse independently. On such short time scales, however, one should check to make sure that inertial effects can be ignored, *i.e.* that $t \gg M/\gamma$.

One serious defect of the Rouse model is its prediction of the relaxation time of the $j = 1$ mode, $\tau_1 \propto N^2$. The experimentally observed result is $\tau_1 \propto N^{3/2}$. We should stress here that the Rouse model applies to ideal chains. In the theory of polymer solutions, a *theta solvent* is one in which polymer coils act as ideal chains. An extension of the Rouse model, due to my former UCSD colleague Bruno Zimm, accounts for hydrodynamically-mediated interactions between any pair of 'beads' along the chain. Specifically, the Zimm model is given by

$$\frac{dx_n^\mu}{dt} = \sum_{n'} H^{\mu\nu}(\boldsymbol{x}_n - \boldsymbol{x}_{n'}) \Big[ k\big(x_{n'+1}^\nu + x_{n'-1}^\nu - 2x_{n'}^\nu\big) + f_{n'}^\nu(t) \Big] \quad , \tag{6.326}$$

where

$$H^{\mu\nu}(\boldsymbol{R}) = \frac{1}{6\pi\eta R} \big(\delta^{\mu\nu} + \hat{R}^\mu \hat{R}^\nu\big) \tag{6.327}$$

is known as the Oseen hydrodynamic tensor (1927) and arises when computing the velocity in a fluid at position $\boldsymbol{R}$ when a point force $\boldsymbol{F} = \boldsymbol{f}\,\delta(\boldsymbol{r})$ is applied at the origin. Typically one replaces $H(\boldsymbol{R})$ by its average over the equilibrium distribution of polymer configurations. Zimm's model more correctly reproduces the behavior of polymers in $\theta$-solvents.

### 6.6.4 Flory theory of self-avoiding walks

What is missing from the random walk free energy is the effect of *steric interactions*. An argument due to Flory takes these interactions into account in a mean field treatment. Suppose we have a chain of radius $R$. Then the average monomer density within the chain is $c = N/R^d$. Assuming short-ranged interactions, we should then add a term to the free energy which effectively counts the number of near self-intersections of the chain. This number should be roughly $Nc$. Thus, we write

$$F(\boldsymbol{R}, N) = F_0 + u(T)\,\frac{N^2}{R^d} + \tfrac{1}{2}dk_{\rm B}T\,\frac{R^2}{Na^2} \quad . \tag{6.328}$$

The effective interaction $u(T)$ is positive in the case of a so-called 'good solvent'.

The free energy is minimized when

$$0 = \frac{\partial F}{\partial R} = -\frac{duN^2}{R^{d+1}} + \frac{dRk_{\rm B}T}{Na^2} \quad , \tag{6.329}$$

which yields the result

$$R_{\rm F}(N) = \left(\frac{ua^2}{k_{\rm B}T}\right)^{1/(d+2)} N^{3/(d+2)} \propto N^{\nu} \quad . \tag{6.330}$$

Thus, we obtain $\nu = 3/(d+2)$. In $d = 1$ this says $\nu = 1$, which is exactly correct because a SAW in $d = 1$ has no option but to keep going in the same direction. In $d = 2$, Flory theory predicts $\nu = \frac{3}{4}$, which is also exact. In $d = 3$, we have $\nu_{d=3} = \frac{3}{5}$, which is extremely close to the numerical value $\nu = 0.5880$. Flory theory is again exact at the SAW upper critical dimension, which is $d = 4$, where $\nu = \frac{1}{2}$, corresponding to a Gaussian random walk[26]. Best. Mean. Field. Theory. Ever.

How well are polymers described as SAWs? Fig. 6.23 shows the radius of gyration $R_{\rm g}$ *versus* molecular weight $M$ for polystyrene chains in a toluene and benzene solvent. The slope is $\nu = d\log R_{\rm g}/d\log M = 0.5936$. Experimental results can vary with concentration and temperature, but generally confirm the validity of the SAW model.

For a SAW under an external force, we compute the Gibbs partition function,

$$Y(\boldsymbol{F}_{\rm ext}, N) = \int d^dR\, P_N(\boldsymbol{R})\, e^{\boldsymbol{F}_{\rm ext}\cdot\boldsymbol{R}/k_{\rm B}T} = \int d^dx\, f(x)\, e^{s\hat{\boldsymbol{n}}\cdot\boldsymbol{x}} \quad , \tag{6.331}$$

where $x = R/R_{\rm F}$ and $s = R_{\rm F}F_{\rm ext}/k_{\rm B}T$ and $\hat{\boldsymbol{n}} = \hat{\boldsymbol{F}}_{\rm ext}$. One then has $R(F_{\rm ext}) = R_{\rm F}\,\Phi(R_{\rm F}/\xi)$, where $\xi = k_{\rm B}T/F_{\rm ext}$ and, in the small $F_{\rm ext}$ limit, $R(F_{\rm ext}) = F_{\rm ext}R_{\rm F}^2/k_{\rm B}T$. For small values of its argument one has $\Phi(u) \propto u$. For large $u$ it can be shown that $R(F_{\rm ext}) \propto (F_{\rm ext}R_{\rm F}/k_{\rm B}T)^{2/3}$.

[26] There are logarithmic corrections to the SAW result exactly at $d = 4$, but for all $d > 4$ one has $\nu = \frac{1}{2}$.

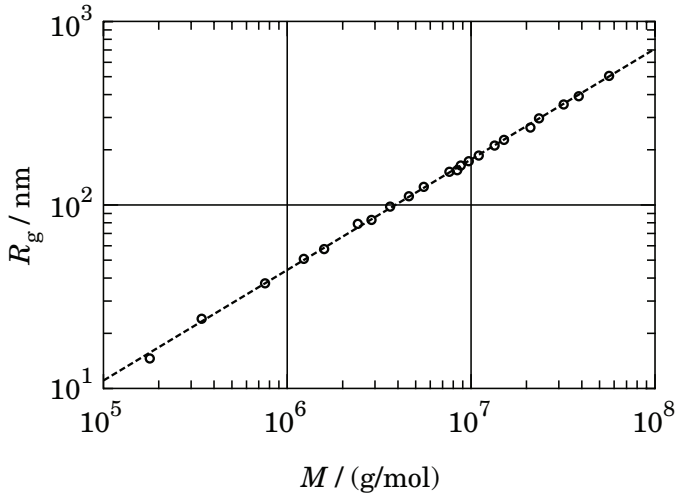



Figure 6.23: Radius of gyration $R_{\rm g}$ of polystyrene in a toluene and benzene solvent, plotted as a function of molecular weight of the polystyrene. The best fit corresponds to a power law $R_{\rm g} \propto M^\nu$ with $\nu = 0.5936$. From J. Des Cloizeaux and G. Jannink, *Polymers in Solution: Their Modeling and Structure* (Oxford, 1990).

On a lattice of coordination number $z$, the number of $N$-step random walks starting from the origin is $\Omega_N = z^N$. If we constrain our random walks to be self-avoiding, the number is reduced to

$$\Omega_N^{\rm SAW} = \mathcal{C}\, N^{\gamma-1}\, y^N \quad , \tag{6.332}$$

where $\mathcal{C}$ and $\gamma$ are dimension-dependent constants, and we expect $y \lesssim z-1$, since at the very least a SAW cannot immediately double back on itself. In fact, on the cubic lattice one has $z = 6$ but $y = 4.68$, slightly less than $z - 1$. One finds $\gamma_{d=2} \simeq \frac{4}{3}$ and $\gamma_{d=3} \simeq \frac{7}{6}$. The RMS end-to-end distance of the SAW is

$$R_{\rm F} = a\, N^\nu \quad , \tag{6.333}$$

where $a$ and $\nu$ are $d$-dependent constants, with $\nu_{d=1} = 1$, $\nu_{d=2} \simeq \frac{3}{4}$, and $\nu_{d=3} \simeq \frac{3}{5}$. The distribution $P_N(\boldsymbol{R})$ has a scaling form,

$$P_N(R) = \frac{1}{R_{\rm F}^d}\, f\!\left(\frac{R}{R_{\rm F}}\right) \qquad (a \ll R \ll Na) \quad . \tag{6.334}$$

One finds

$$f(x) \sim \begin{cases} x^g & x \ll 1 \\ \exp(-x^\delta) & x \gg 1 \end{cases} \quad , \tag{6.335}$$

with $g = (\gamma - 1)/\nu$ and $\delta = 1/(1-\nu)$.

### 6.6.5 Polymers and solvents

Consider a solution of monodisperse polymers of length $N$ in a solvent. Let $\phi$ be the dimensionless monomer concentration, so $\phi/N$ is the dimensionless polymer concentration and $\phi_{\rm s} = 1 - \phi$ is the

dimensionless solvent concentration. (Dimensionless concentrations are obtained by dividing the corresponding dimensionful concentration by the overall density.) The entropy of mixing for such a system is given by eqn. 2.352. We have

$$S_{\rm mix} = -\frac{V k_{\rm B}}{v_0} \cdot \left\{ \frac{1}{N}\phi \log \phi + (1-\phi)\log(1-\phi) \right\} \quad , \tag{6.336}$$

where $v_0 \propto a^3$ is the volume per monomer. Accounting for an interaction between the monomer and the solvent, we have that the free energy of mixing is

$$\frac{v_0\, F_{\rm mix}}{V k_{\rm B} T} = \frac{1}{N}\,\phi \log \phi + (1-\phi)\log(1-\phi) + \chi\,\phi(1-\phi) \quad . \tag{6.337}$$

where $\chi$ is the dimensionless polymer-solvent interaction, called the *Flory parameter*. This provides a mean field theory of the polymer-solvent system.

The osmotic pressure $\Pi$ is defined by

$$\Pi = -\left.\frac{\partial F_{\rm mix}}{\partial V}\right|_{N_{\rm p}} \quad , \tag{6.338}$$

which is the variation of the free energy of mixing with respect to volume *holding the number of polymers constant*. The monomer concentration is $\phi = N N_{\rm p} v_0 / V$, so

$$\left.\frac{\partial}{\partial V}\right|_{N_{\rm p}} = -\frac{\phi^2}{N N_{\rm p}\, v_0}\left.\frac{\partial}{\partial \phi}\right|_{N_{\rm p}} \quad . \tag{6.339}$$

Now we have

$$F_{\rm mix} = N N_{\rm p}\, k_{\rm B} T \left\{ \frac{1}{N}\,\log\phi + (\phi^{-1}-1)\log(1-\phi) + \chi\,(1-\phi) \right\} \quad , \tag{6.340}$$

and therefore

$$\Pi = \frac{k_{\rm B} T}{v_0}\Big[ (N^{-1}-1)\,\phi - \log(1-\phi) - \chi\,\phi^2 \Big] \quad . \tag{6.341}$$

In the limit of vanishing monomer concentration $\phi \to 0$, we recover

$$\Pi = \frac{\phi\, k_{\rm B} T}{N v_0} \quad , \tag{6.342}$$

which is the ideal gas law for polymers. For $N^{-1} \ll \phi \ll 1$, we expand the logarithm and obtain

$$\begin{aligned} \frac{v_0\,\Pi}{k_{\rm B} T} &= \frac{1}{N}\phi + \tfrac{1}{2}(1-2\chi)\,\phi^2 + \mathcal{O}(\phi^3) \\ &\approx \tfrac{1}{2}(1-2\chi)\,\phi^2 \quad . \end{aligned} \tag{6.343}$$

Note that $\Pi > 0$ only if $\chi < \frac{1}{2}$, which is the condition for a 'good solvent'.

In fact, eqn. 6.343 is only qualitatively correct. In the limit where $\chi \ll \frac{1}{2}$, Flory showed that the individual polymer coils behave much as hard spheres of radius $R_{\rm F}$. The osmotic pressure then satisfies

something analogous to a virial equation of state:

$$\begin{aligned}\frac{\Pi}{k_{\rm B}T} &= \frac{\phi}{Nv_0} + A\left(\frac{\phi}{Nv_0}\right)^2 R_{\rm F}^3 + \dots \\ &= \frac{\phi}{Nv_0}\,h(\phi/\phi^*) \quad .\end{aligned} \tag{6.344}$$

This is generalized to a scaling form in the second line, where $h(x)$ is a scaling function, and the quantity $\phi^*$ is given by $\phi^* = Nv_0/R_{\rm F}^3 \propto N^{-4/5}$, assuming $d = 3$ and $\nu = \frac{3}{5}$ from Flory theory. As $x = \phi/\phi^* \to 0$, we must recover the ideal gas law, so $h(x) = 1 + \mathcal{O}(x)$ in this limit. For $x \to \infty$, we require that the result be independent of the degree of polymerization $N$. This means $h(x) \propto x^p$ with $\frac{4}{5}p = 1$, *i.e.* $p = \frac{5}{4}$. The result is known as the des Cloiseaux law:

$$\frac{v_0\Pi}{k_{\rm B}T} = C\,\phi^{9/4} \quad , \tag{6.345}$$

where $C$ is a constant. This is valid for what is known as semi-dilute solutions, where $\phi^* \ll \phi \ll 1$. In the dense limit $\phi \sim 1$, the results do not exhibit this universality, and we must appeal to liquid state theory, which is no fun at all.

## 6.7 Lee-Yang Theory

### 6.7.1 Analytic properties of the partition function

How can statistical mechanics describe phase transitions? This question was addressed in some beautiful mathematical analysis by Lee and Yang[27]. Consider the grand partition function $\Xi$,

$$\Xi(T, V, z) = \sum_{N=0}^{\infty} z^N\, Q_N(T, V)\, \lambda_T^{-dN} \quad , \tag{6.346}$$

where

$$Q_N(T, V) = \frac{1}{N!}\int d^d x_1 \cdots \int d^d x_N\, e^{-U(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N)/k_{\rm B}T} \tag{6.347}$$

is the contribution to the $N$-particle partition function from the potential energy $U$ (assuming no momentum-dependent potentials). For two-body central potentials, we have

$$U(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) = \sum_{i<j} u\big(|\boldsymbol{x}_i - \boldsymbol{x}_j|\big) \quad . \tag{6.348}$$

Suppose further that these classical particles have hard cores. Then for any *finite* volume, there must be some maximum number $N_V$ such that $Q_N(T, V)$ vanishes for $N > N_V$. This is because if $N > N_V$ at least two spheres must overlap, in which case the potential energy is infinite. One defines the *sphere packing fraction* $\varphi \equiv \frac{1}{6}\pi a^3 n$, where $n$ is the number density of spheres and $a$ is the sphere diameter. Equivalently,

[27] See C. N. Yang and R. D. Lee, *Phys. Rev.* **87**, 404 (1952) and *ibid*, p. 410

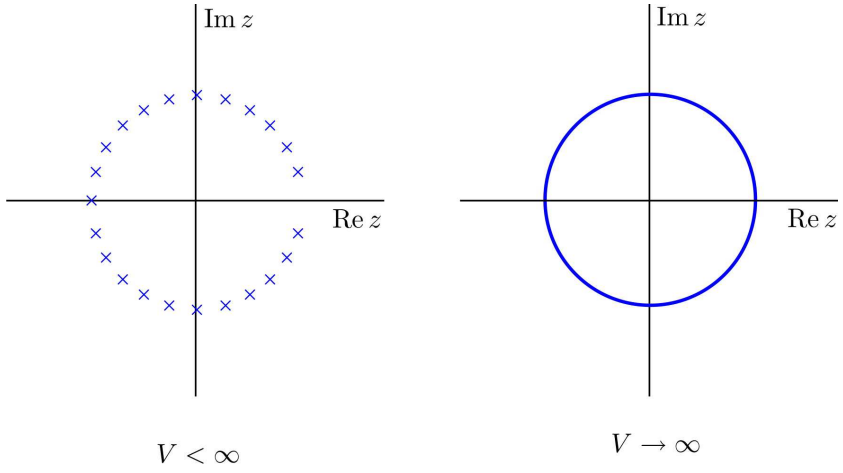



Figure 6.24: In the thermodynamic limit, the grand partition function can develop a singularity at positive real fugacity $z$. The set of discrete zeros fuses into a branch cut.

$\varphi = v_0/v$ where $v_0$ is the sphere volume and $v = n^{-1} = V/N$. The maximum packing density for hard spheres is achieved for a hexagonal close packed (HCP) lattice[28], for which $\varphi_{\rm HCP} = \pi/3\sqrt{2} = 0.74048$, in which case $N_V = \sqrt{2}\,V/a^3$ is the maximum particle number.

Thus, if $V$ itself is finite, then $\Xi(T,V,z)$ is a *finite* degree polynomial in $z$, and may be factorized as

$$\Xi(T,V,z) = \sum_{N=0}^{N_V} z^N\, Q_N(T,V)\,\lambda_T^{-dN} = \prod_{k=1}^{N_V}\left(1 - \frac{z}{z_k}\right) \quad , \tag{6.349}$$

where $z_k(T,V)$ is one of the $N_V$ zeros of the grand partition function. Note that the $\mathcal{O}(z^0)$ term is fixed to be unity. Note also that since the configuration integrals $Q_N(T,V)$ are all positive, $\Xi(z)$ is an increasing function along the positive real $z$ axis. In addition, since the coefficients of $z^N$ in the polynomial $\Xi(z)$ are all real, then $\Xi(z) = 0$ implies $\overline{\Xi(z)} = \Xi(\bar z) = 0$, so the zeros of $\Xi(z)$ are either real and negative or else come in complex conjugate pairs.

For finite $N_V$, the situation is roughly as depicted in the left panel of fig. 6.24, with a set of $N_V$ zeros arranged in complex conjugate pairs (or negative real values). The zeros aren't necessarily distributed along a circle as shown in the figure, though. They could be anywhere, so long as they are symmetrically distributed about the $\mathrm{Re}(z)$ axis, and no zeros occur for $z$ real and nonnegative.

Lee and Yang proved the existence of the limits

$$\begin{aligned} \frac{p}{k_{\rm B}T} &= \lim_{V\to\infty}\frac{1}{V}\,\log\Xi(T,V,z) \\ n &= \lim_{V\to\infty} z\,\frac{\partial}{\partial z}\left[\frac{1}{V}\,\log\Xi(T,V,z)\right] \quad , \end{aligned} \tag{6.350}$$

[28] See *e.g.* http://en.wikipedia.org/wiki/Close-packing. For *randomly* close-packed hard spheres, one finds, from numerical simulations, $\varphi_{\rm RCP} = 0.644$.

and notably the result

$$n = z\,\frac{\partial}{\partial z}\left(\frac{p}{k_{\rm B}T}\right) \quad , \tag{6.351}$$

which amounts to the commutativity of the thermodynamic limit $V \to \infty$ with the differential operator $z\,\frac{\partial}{\partial z}$. In particular, $p(T,z)$ is a smooth function of $z$ in regions free of roots. If the roots do coalesce and pinch the positive real axis, then the density $n$ can be discontinuous, as in a first order phase transition, or a higher derivative $\partial^j p/\partial z^j$ can be discontinuous or divergent, as in a second order phase transition.

### 6.7.2 Electrostatic analogy

There is a beautiful analogy to the theory of two-dimensional electrostatics. We write

$$\begin{aligned}\frac{p}{k_{\rm B}T} &= \frac{1}{V}\sum_{k=1}^{N_V}\log\left(1-\frac{z}{z_k}\right) \\ &= -\sum_{k=1}^{N_V}\Big[\phi(z-z_k)-\phi(0-z_k)\Big] \quad ,\end{aligned} \tag{6.352}$$

where $\phi(z) = -V^{-1}\log(z)$ is the complex potential due to a line charge of linear density $\lambda = V^{-1}$ located at origin. The number density is then

$$n = z\,\frac{\partial}{\partial z}\left(\frac{p}{k_{\rm B}T}\right) = -z\,\frac{\partial}{\partial z}\sum_{k=1}^{N_V}\phi(z-z_k) \quad , \tag{6.353}$$

to be evaluated for physical values of $z$, *i.e.* $z \in \mathbb{R}^+$. Since $\phi(z)$ is analytic,

$$\frac{\partial\phi}{\partial\bar z} = \frac{1}{2}\,\frac{\partial\phi}{\partial x} + \frac{i}{2}\,\frac{\partial\phi}{\partial y} = 0 \quad . \tag{6.354}$$

If we decompose the complex potential $\phi = \phi_1 + i\phi_2$ into real and imaginary parts, the condition of analyticity is recast as the Cauchy-Riemann equations,

$$\frac{\partial\phi_1}{\partial x} = \frac{\partial\phi_2}{\partial y} \qquad , \qquad \frac{\partial\phi_1}{\partial y} = -\frac{\partial\phi_2}{\partial x} \quad . \tag{6.355}$$

Thus,

$$\begin{aligned}-\frac{\partial\phi}{\partial z} &= -\frac{1}{2}\,\frac{\partial\phi}{\partial x} + \frac{i}{2}\,\frac{\partial\phi}{\partial y} \\ &= -\frac{1}{2}\left(\frac{\partial\phi_1}{\partial x}+\frac{\partial\phi_2}{\partial y}\right) + \frac{i}{2}\left(\frac{\partial\phi_1}{\partial y}-\frac{\partial\phi_2}{\partial x}\right) \\ &= -\frac{\partial\phi_1}{\partial x} + i\,\frac{\partial\phi_1}{\partial y} = E_x - iE_y \quad ,\end{aligned} \tag{6.356}$$

where $\boldsymbol{E} = -\boldsymbol{\nabla}\phi_1$ is the electric field. Suppose, then, that as $V \to \infty$ a continuous charge distribution develops, which crosses the positive real $z$ axis at a point $x \in \mathbb{R}^+$. Then

$$\frac{n_+ - n_-}{x} = E_x(x^+) - E_x(x^-) = 2\pi\sigma(x) \quad , \tag{6.357}$$

where $\sigma$ is the linear charge density (assuming logarithmic two-dimensional potentials), or the two-dimensional charge density (if we extend the distribution along a third axis).

### 6.7.3 Example

As an example, consider the function

$$\begin{aligned} \Xi(z) &= \frac{(1+z)^M\,(1-z^K)}{1-z} \\ &= (1+z)^M\left(1 + z + z^2 + \ldots + z^{K-1}\right) \quad . \end{aligned} \tag{6.358}$$

The $(M + K - 1)$ degree polynomial has an $M^{\text{th}}$ order zero at $z = -1$ and $(K - 1)$ simple zeros at $z = e^{2\pi ij/K}$, where $j \in \{1, \ldots, K-1\}$. The maximum particle number is the order of $\Xi(z)$, which is $N_V = M + K - 1$. We define $v_0 \equiv V/N_V$ and $r \equiv M/N_V \in [0, 1]$. The thermodynamic limit is then taken as $N_V \to \infty$ with $r$ finite, in which case $K/N_V = 1 - r$. Thus,

$$\begin{aligned} \frac{p v_0}{k_{\text{B}}T} = \lim_{N_V\to\infty}\frac{1}{N_V}\log\Xi(z) &= \lim_{N_V\to\infty}\frac{1}{N_V}\log\Xi(z) \\ &= \lim_{N_V\to\infty}\frac{1}{N_V}\left[M\log(1+z) + \log\left(\frac{1-z^K}{1-z}\right)\right] \quad . \end{aligned} \tag{6.359}$$

The limit depends on whether $|z| > 1$ or $|z| < 1$. For $z$ real and nonnegative we have

$$\frac{p v_0}{k_{\text{B}}T} = \begin{cases} r\log(1+z) & \text{if} \quad 0 < z \le 1 \\ r\log(1+z) + (1-r)\log z & \text{if} \quad 1 < z < \infty \quad . \end{cases} \tag{6.360}$$

Thus,

$$\frac{v_0}{v} = n v_0 = z\,\frac{\partial}{\partial z}\left(\frac{p v_0}{k_{\text{B}}T}\right) = \begin{cases} \frac{rz}{1+z} & \text{if} \quad 0 < z \le 1 \\ \frac{rz}{1+z} + 1 - r & \text{if} \quad 1 < z < \infty \quad . \end{cases} \tag{6.361}$$

with $v = n^{-1}$. Solving for $z(v)$, we find

$$z = \begin{cases} \frac{v_0 - (1-r)v}{v - v_0} & \text{if} \quad v_0 < v \le \frac{2v_0}{2-r} \\ 1 & \text{if } \frac{2v_0}{2-r} \le v \le \frac{2v_0}{r} \\ \frac{v_0}{rv - v_0} & \text{if } \frac{2v_0}{r} \le v < \infty \quad . \end{cases} \tag{6.362}$$

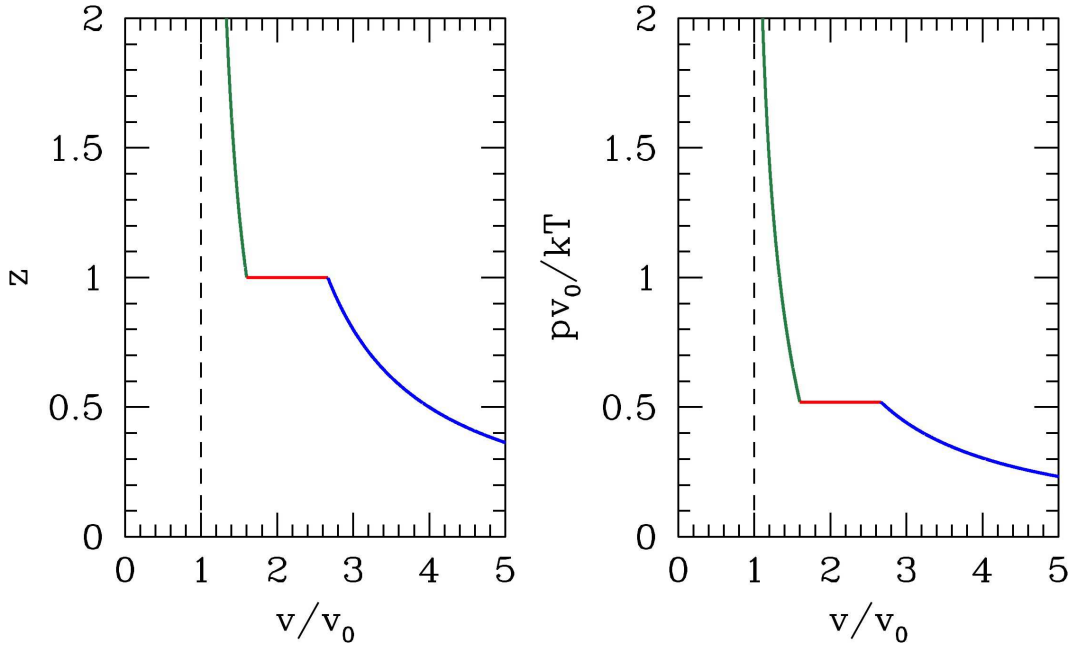



Figure 6.25: Fugacity $z$ and $pv_0/k_{\rm B}T$ *versus* dimensionless specific volume $v/v_0$ for the example problem discussed in the text. The case $r = 0.75$ is shown.

We then obtain the equation of state,

$$\frac{pv_0}{k_{\rm B}T} = \begin{cases} r\log\left(\frac{rv}{v-v_0}\right) + (1-r)\log\left(\frac{v_0-(1-r)v}{v-v_0}\right) & \text{if} \quad v_0 < v \le \frac{2v_0}{2-r} \\ r\log 2 & \text{if} \quad \frac{2v_0}{2-r} \le v \le \frac{2v_0}{r} \\ r\log\left(\frac{rv}{rv-v_0}\right) & \text{if} \quad \frac{2v_0}{r} \le v < \infty \quad . \end{cases} \tag{6.363}$$

Note that the smallest possible volume is $v_{\rm min} = v_0$. Thus, the specific volume jumps discontinuously from $v_1 = \frac{2v_0}{2-r}$ to $v_2 = \frac{2v_0}{r}$ at a fixed $p/k_{\rm B}T$. This is reminiscent of the liquid-gas transition in §2.11 and in §§7.2,3. Here the interval $v \in [v_1, v_2]$ would correspond to the coexistence region, with $v_{1,2}$ being the specific volumes for the liquid and gas endpoints, respectively.

## 6.8 Appendix: Potts Model in One Dimension

### 6.8.1 Hamiltonian

The Potts model is defined by the Hamiltonian

$$H = -J\sum_{\langle ij\rangle}\delta_{\sigma_i,\sigma_j} - h\sum_i \delta_{\sigma_i,1} \quad . \tag{6.364}$$

Here, the spin variables $\sigma_i$ take values in the set $\{1, 2, \ldots, q\}$ on each site. The equivalent of an external magnetic field in the Ising case is a field $h$ which prefers a particular value of $\sigma$ ($\sigma = 1$ in the

above Hamiltonian). Once again, it is not possible to compute the partition function on general lattices, however in one dimension we may once again find $Z$ using the transfer matrix method.

### 6.8.2 Transfer matrix

On a ring of $N$ sites, we have

$$Z = \mathrm{Tr}\, e^{-\beta H} = \sum_{\{\sigma_n\}} e^{\beta h \delta_{\sigma_1,1}}\, e^{\beta J \delta_{\sigma_1,\sigma_2}} \cdots e^{\beta h \delta_{\sigma_N,1}}\, e^{\beta J \delta_{\sigma_N,\sigma_1}} = \mathrm{Tr}\left(R^N\right) \quad , \tag{6.365}$$

where the $q \times q$ transfer matrix $R$ is given by

$$R_{\sigma\sigma'} = e^{\beta J \delta_{\sigma\sigma'}}\, e^{\frac{1}{2}\beta h \delta_{\sigma,1}}\, e^{\frac{1}{2}\beta h \delta_{\sigma',1}} = \begin{cases} e^{\beta(J+h)} & \text{if } \sigma = \sigma' = 1 \\ e^{\beta J} & \text{if } \sigma = \sigma' \neq 1 \\ e^{\beta h/2} & \text{if } \sigma = 1 \text{ and } \sigma' \neq 1 \\ e^{\beta h/2} & \text{if } \sigma \neq 1 \text{ and } \sigma' = 1 \\ 1 & \text{if } \sigma \neq 1 \text{ and } \sigma' \neq 1 \text{ and } \sigma \neq \sigma' \end{cases} . \tag{6.366}$$

In matrix form,

$$R = \begin{pmatrix} e^{\beta(J+h)} & e^{\beta h/2} & e^{\beta h/2} & \cdots & & e^{\beta h/2} \\ e^{\beta h/2} & e^{\beta J} & 1 & \cdots & & 1 \\ e^{\beta h/2} & 1 & e^{\beta J} & \cdots & & 1 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ e^{\beta h/2} & 1 & 1 & \cdots & e^{\beta J} & 1 \\ e^{\beta h/2} & 1 & 1 & \cdots & 1 & e^{\beta J} \end{pmatrix} \tag{6.367}$$

The matrix $R$ has $q$ eigenvalues $\lambda_j$, with $j = 1, \ldots, q$. The partition function for the Potts chain is then

$$Z = \sum_{j=1}^{q} \lambda_j^N \quad . \tag{6.368}$$

We can actually find the eigenvalues of $R$ analytically. To this end, consider the vectors

$$\phi = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad , \qquad \psi = \left(q - 1 + e^{\beta h}\right)^{-1/2} \begin{pmatrix} e^{\beta h/2} \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad . \tag{6.369}$$

Then $R$ may be written as

$$R = \left(e^{\beta J} - 1\right) \mathbb{I} + \left(q - 1 + e^{\beta h}\right) |\,\psi\,\rangle\langle\,\psi\,| + \left(e^{\beta J} - 1\right)\left(e^{\beta h} - 1\right) |\,\phi\,\rangle\langle\,\phi\,| \quad , \tag{6.370}$$

where $\mathbb{I}$ is the $q \times q$ identity matrix. When $h = 0$, we have a simpler form,

$$R = \big(e^{\beta J} - 1\big)\,\mathbb{I} + q\,|\,\psi\,\rangle\langle\,\psi\,| \quad . \tag{6.371}$$

From this we can read off the eigenvalues:

$$\begin{aligned} \lambda_1 &= e^{\beta J} + q - 1 \\ \lambda_j &= e^{\beta J} - 1 \quad , \quad j \in \{2, \ldots, q\} \quad , \end{aligned} \tag{6.372}$$

since $|\,\psi\,\rangle$ is an eigenvector with eigenvalue $\lambda = e^{\beta J} + q - 1$, and any vector orthogonal to $|\,\psi\,\rangle$ has eigenvalue $\lambda = e^{\beta J} - 1$. The partition function is then

$$Z = \big(e^{\beta J} + q - 1\big)^N + (q-1)\big(e^{\beta J} - 1\big)^N \quad . \tag{6.373}$$

In the thermodynamic limit $N \to \infty$, only the $\lambda_1$ eigenvalue contributes, and we have

$$F(T, N, h = 0) = -N k_{\rm B} T \log\big(e^{J/k_{\rm B}T} + q - 1\big) \qquad \text{for } N \to \infty \quad . \tag{6.374}$$

When $h$ is nonzero, the calculation becomes somewhat more tedious, but still relatively easy. The problem is that $|\,\psi\,\rangle$ and $|\,\phi\,\rangle$ are not orthogonal, so we define

$$|\,\chi\,\rangle = \frac{|\,\phi\,\rangle - |\,\psi\,\rangle\langle\,\psi\,|\,\phi\,\rangle}{\sqrt{1 - \langle\,\phi\,|\,\psi\,\rangle^2}} \quad , \tag{6.375}$$

where

$$x \equiv \langle\,\phi\,|\,\psi\,\rangle = \left(\frac{e^{\beta h}}{q - 1 + e^{\beta h}}\right)^{1/2} \quad . \tag{6.376}$$

Now we have $\langle\,\chi\,|\,\psi\,\rangle = 0$, with $\langle\,\chi\,|\,\chi\,\rangle = 1$ and $\langle\,\psi\,|\,\psi\,\rangle = 1$, with

$$|\,\phi\,\rangle = \sqrt{1 - x^2}\;|\,\chi\,\rangle + x\,|\,\psi\,\rangle \quad . \tag{6.377}$$

and the transfer matrix is then

$$\begin{aligned} R &= \big(e^{\beta J} - 1\big)\,\mathbb{I} + \big(q - 1 + e^{\beta h}\big)\,|\,\psi\,\rangle\langle\,\psi\,| \\ &\quad + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big)\left[(1 - x^2)\,|\,\chi\,\rangle\langle\,\chi\,| + x^2\,|\,\psi\,\rangle\langle\,\psi\,| + x\sqrt{1 - x^2}\,\Big(|\,\chi\,\rangle\langle\,\psi\,| + |\,\psi\,\rangle\langle\,\chi\,|\Big)\right] \\ &= \big(e^{\beta J} - 1\big)\,\mathbb{I} + \left[\big(q - 1 + e^{\beta h}\big) + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big)\left(\frac{e^{\beta h}}{q - 1 + e^{\beta h}}\right)\right]|\,\psi\,\rangle\langle\,\psi\,| \\ &\quad + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big)\left(\frac{q-1}{q - 1 + e^{\beta h}}\right)|\,\chi\,\rangle\langle\,\chi\,| \\ &\quad + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big)\left(\frac{(q-1)\,e^{\beta h}}{q - 1 + e^{\beta h}}\right)^{1/2}\Big(|\,\chi\,\rangle\langle\,\psi\,| + |\,\psi\,\rangle\langle\,\chi\,|\Big) \quad , \end{aligned} \tag{6.378}$$

which in the two-dimensional subspace spanned by $|\,\chi\,\rangle$ and $|\,\psi\,\rangle$ is of the form

$$R = \begin{pmatrix} a & c \\ c & b \end{pmatrix} \quad . \tag{6.379}$$

Recall that for any $2 \times 2$ Hermitian matrix,

$$M = a_0 \, \mathbb{I} + \boldsymbol{a} \cdot \boldsymbol{\tau} = \begin{pmatrix} a_0 + a_3 & a_1 - i a_2 \\ a_1 + i a_2 & a_0 - a_3 \end{pmatrix} \quad , \tag{6.380}$$

the characteristic polynomial is

$$P(\lambda) = \det\big(\lambda \, \mathbb{I} - M\big) = (\lambda - a_0)^2 - a_1^2 - a_2^2 - a_3^2 \quad , \tag{6.381}$$

and hence the eigenvalues are

$$\lambda_\pm = a_0 \pm \sqrt{a_1^2 + a_2^2 + a_3^2} \quad . \tag{6.382}$$

For the transfer matrix of eqn. 6.378, we obtain, after a little work,

$$\begin{aligned} \lambda_{1,2} = e^{\beta J} - 1 + \tfrac{1}{2}\Big[q - 1 + e^{\beta h} + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big)\Big] \\ \pm \tfrac{1}{2}\sqrt{\Big[q - 1 + e^{\beta h} + \big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big)\Big]^2 - 4(q-1)\big(e^{\beta J} - 1\big)\big(e^{\beta h} - 1\big)} \quad . \end{aligned} \tag{6.383}$$

There are $q-2$ other eigenvalues, however, associated with the $(q-2)$-dimensional subspace orthogonal to $|\,\chi\,\rangle$ and $|\,\psi\,\rangle$. Clearly all these eigenvalues are given by

$$\lambda_j = e^{\beta J} - 1 \qquad , \quad j \in \{3\,, \ldots\,, q\} \quad . \tag{6.384}$$

The partition function is then

$$Z = \lambda_1^N + \lambda_2^N + (q-2)\,\lambda_3^N \quad , \tag{6.385}$$

and in the thermodynamic limit $N \to \infty$ the maximum eigenvalue $\lambda_1$ dominates. Note that we recover the correct limit as $h \to 0$.

# Chapter 7

# Mean Field Theory of Phase Transitions

## 7.1 References

- M. Kardar, *Statistical Physics of Particles* (Cambridge, 2007)
  A superb modern text, with many insightful presentations of key concepts.
- M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* ($3^{\rm rd}$ edition, World Scientific, 2006)
  An excellent graduate level text. Good discussion of mean field theory.
- G. Parisi, *Statistical Field Theory* (Addison-Wesley, 1988)
  An advanced text focusing on field theoretic approaches, covering mean field and Landau-Ginzburg theories before moving on to renormalization group and beyond.
- J. P. Sethna, *Entropy, Order Parameters, and Complexity* (Oxford, 2006)
  An excellent introductory text with a very modern set of topics and exercises. Available online at
  http://www.physics.cornell.edu/sethna/StatMech

## 7.2 The van der Waals system

### 7.2.1 Equation of state

Recall the van der Waals equation of state,

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT \quad , \tag{7.1}$$

where $v = N_{\rm A}V/N$ is the molar volume. Solving for $p(v, T)$, we have

$$p = \frac{RT}{v - b} - \frac{a}{v^2} \quad . \tag{7.2}$$

Let us fix the temperature $T$ and examine the function $p(v)$. Clearly $p(v)$ is a decreasing function of volume for $v$ just above the minimum allowed value $v = b$, as well as for $v \to \infty$. But is $p(v)$ a monotonic function for all $v \in [b, \infty]$?

We can answer this by computing the derivative,

$$\left(\frac{\partial p}{\partial v}\right)_T = \frac{2a}{v^3} - \frac{RT}{(v - b)^2} \quad . \tag{7.3}$$

Setting this expression to zero for finite $v$, we obtain the equation[1]

$$\frac{2a}{bRT} = \frac{u^3}{(u - 1)^2} \quad , \tag{7.4}$$

where $u \equiv v/b$ is dimensionless. It is easy to see that the function $f(u) = u^3/(u - 1)^2$ has a unique minimum for $u > 1$. Setting $f'(u^*) = 0$ yields $u^* = 3$, and so $f_{\rm min} = f(3) = \frac{27}{4}$. Thus, for $T > T_{\rm c} = 8a/27bR$, the LHS of eqn. 7.4 lies below the minimum value of the RHS, and there is no solution. This means that $p(v, T > T_{\rm c})$ is a monotonically decreasing function of $v$.

At $T = T_{\rm c}$ there is a saddle-node bifurcation[2]. Setting $v_{\rm c} = bu^* = 3b$ and evaluating $p_{\rm c} = p(v_{\rm c}, T_{\rm c})$, we have that the location of the *critical point* for the van der Waals system is[3]

$$p_{\rm c} = \frac{a}{27\,b^2} \qquad , \qquad v_{\rm c} = 3b \qquad , \qquad T_{\rm c} = \frac{8a}{27\,bR} \quad . \tag{7.5}$$

For $T < T_{\rm c}$, there are two solutions to eqn. 7.4, corresponding to a local minimum and a local maximum of the function $p(v)$. The locus of points in the $(v, p)$ plane for which $(\partial p/\partial v)_T = 0$ is obtained by setting eqn. 7.3 to zero and solving for $T$, then substituting this into eqn. 7.2. The result is

$$p^*(v) = \frac{a}{v^2} - \frac{2ab}{v^3} \quad . \tag{7.6}$$

[1] There is always a solution to $(\partial p/\partial v)_T = 0$ at $v = \infty$.

[2] See the appendix in §7.10

[3] Equivalently, one can obtain the location of the saddle-node bifurcation in the local extrema of $p(T, v)$ by setting $(\partial p/\partial v)_T = 0$ and $(\partial^2 p/\partial v^2)_T = 0$. This yields two equations for the two unknowns $(T_{\rm c}, v_{\rm c})$. Invoking the van der Waals equation of state then yields $p_{\rm c} = p(T_{\rm c}, v_{\rm c})$.

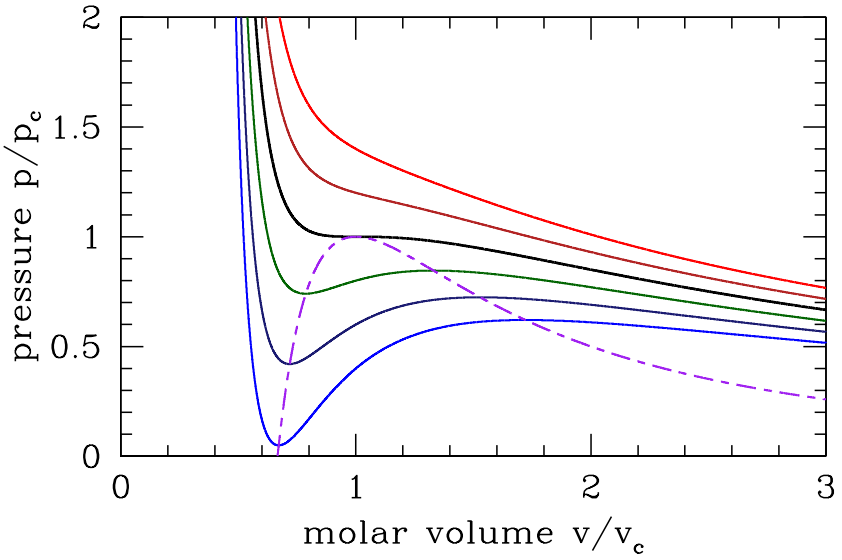



Figure 7.1: Pressure *versus* molar volume for the van der Waals gas at temperatures in equal intervals from $T = 1.10\,T_{\rm c}$ (red) to $T = 0.85\,T_{\rm c}$ (blue). The purple curve is $\bar{p}^*(\bar{v})$.

Expressed in terms of dimensionless quantities $\bar{p} = p/p_{\rm c}$ and $\bar{v} = v/v_{\rm c}$, this equation becomes

$$\bar{p}^*(\bar{v}) = \frac{3}{\bar{v}^2} - \frac{2}{\bar{v}^3} \quad . \tag{7.7}$$

Along the curve $p = p^*(v)$, the isothermal compressibility, $\kappa_T = -v^{-1}\big(\partial v/\partial p\big)_T$ diverges, heralding a *thermodynamic instability*. To understand better, let us compute the free energy of the van der Waals system, $F = E - TS$. Regarding the energy $E$, we showed back in chapter 2 that

$$\left(\frac{\partial \varepsilon}{\partial v}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_v - p = \frac{a}{v^2} \quad , \tag{7.8}$$

which entails

$$\varepsilon(T,v) = \tfrac{1}{2}\mathfrak{f}RT - \frac{a}{v} \quad , \tag{7.9}$$

where $\varepsilon = E/\nu$ is the molar internal energy. The first term is the molar energy of an ideal gas, where $\mathfrak{f}$ is the number of molecular degrees of freedom, which is the appropriate low density limit. The molar specific heat is then $c_V = \big(\partial\varepsilon/\partial T\big)_v = \frac{1}{2}\mathfrak{f}R$, which means that the molar entropy is

$$s(T,v) = \int^T\!\! dT'\,\frac{c_V}{T'} = \tfrac{1}{2}\mathfrak{f}R\log(T/T_{\rm c}) + s_1(v) \quad . \tag{7.10}$$

We then write $f = \varepsilon - Ts$, and we fix the function $s_1(v)$ by demanding that $p = -\big(\partial f/\partial v\big)_T$. This yields $s_1(v) = R\log(v-b) + s_0$, where $s_0$ is a constant. Thus[4],

$$f(T,v) = \tfrac{1}{2}\mathfrak{f}RT\left(1 - \log\big(T/T_{\rm c}\big)\right) - \frac{a}{v} - RT\log(v-b) - Ts_0 \quad . \tag{7.11}$$

| gas | $a\ \left(\frac{\mathrm{L}^2\cdot\mathrm{bar}}{\mathrm{mol}^2}\right)$ | $b\ \left(\frac{\mathrm{L}}{\mathrm{mol}}\right)$ | $p_c$ (bar) | $T_c$ (K) | $v_c$ (L/mol) |
|---|---|---|---|---|---|
| Acetone | 14.09 | 0.0994 | 52.82 | 505.1 | 0.2982 |
| Argon | 1.363 | 0.03219 | 48.72 | 150.9 | 0.0966 |
| Carbon dioxide | 3.640 | 0.04267 | 74.04 | 304.0 | 0.1280 |
| Ethanol | 12.18 | 0.08407 | 63.83 | 516.3 | 0.2522 |
| Freon | 10.78 | 0.0998 | 40.09 | 384.9 | 0.2994 |
| Helium | 0.03457 | 0.0237 | 2.279 | 5.198 | 0.0711 |
| Hydrogen | 0.2476 | 0.02661 | 12.95 | 33.16 | 0.0798 |
| Mercury | 8.200 | 0.01696 | 1055 | 1723 | 0.0509 |
| Methane | 2.283 | 0.04278 | 46.20 | 190.2 | 0.1283 |
| Nitrogen | 1.408 | 0.03913 | 34.06 | 128.2 | 0.1174 |
| Oxygen | 1.378 | 0.03183 | 50.37 | 154.3 | 0.0955 |
| Water | 5.536 | 0.03049 | 220.6 | 647.0 | 0.0915 |

Table 7.1: van der Waals parameters for some common gases. (Source: *Wikipedia*)

We know that under equilibrium conditions, $f$ is driven to a minimum by spontaneous processes. Now suppose that $\left(\partial^2 f/\partial v^2\right)_T < 0$ over some range of $v$ at a given temperature $T$. This would mean that one mole of the system at volume $v$ and temperature $T$ could lower its energy by rearranging into two half-moles, with respective molar volumes $v \pm \delta v$, each at temperature $T$, where $\delta v$ is a positive infinitesimal. The total volume and temperature thus remain fixed, but the free energy changes by an amount $\Delta f = \frac{1}{2}\left(\partial^2 f/\partial v^2\right)_T (\delta v)^2 < 0$. This means that the system is unstable – it can lower its energy by dividing up into two subsystems each with different densities (*i.e.* molar volumes). Note that the onset of stability occurs when

$$\left.\frac{\partial^2 f}{\partial v^2}\right|_T = -\left.\frac{\partial p}{\partial v}\right|_T = \frac{1}{v\kappa_T} = 0 \quad , \tag{7.12}$$

which is to say when $\kappa_T = \infty$. As we saw, this occurs at $p = p^*(v)$, given in eqn. 7.6.

However, this condition, $\left(\partial^2 f/\partial v^2\right)_T < 0$, is in fact too strong. We need not limit our analysis to cases where $\delta v$ is infinitesimal, in which case the system can be unstable even at molar volumes where $\left(\partial^2 f/\partial v^2\right)_T > 0$. The reason is shown graphically in fig. 7.2. At the fixed temperature $T$, for any molar volume $v$ between $v_{\rm liquid} \equiv v_1$ and $v_{\rm gas} \equiv v_2$, the system can lower its free energy by *phase separating* into regions of different molar volumes. In general we can write

$$v = (1-x)\,v_1 + x\,v_2 \quad , \tag{7.13}$$

so $v = v_1$ when $x = 0$ and $v = v_2$ when $x = 1$. The free energy upon phase separation is simply

$$f = (1-x)\,f_1 + x\,f_2 \quad , \tag{7.14}$$

[4] Don't confuse the molar free energy ($f$) with the number of molecular degrees of freedom (ƒ)!

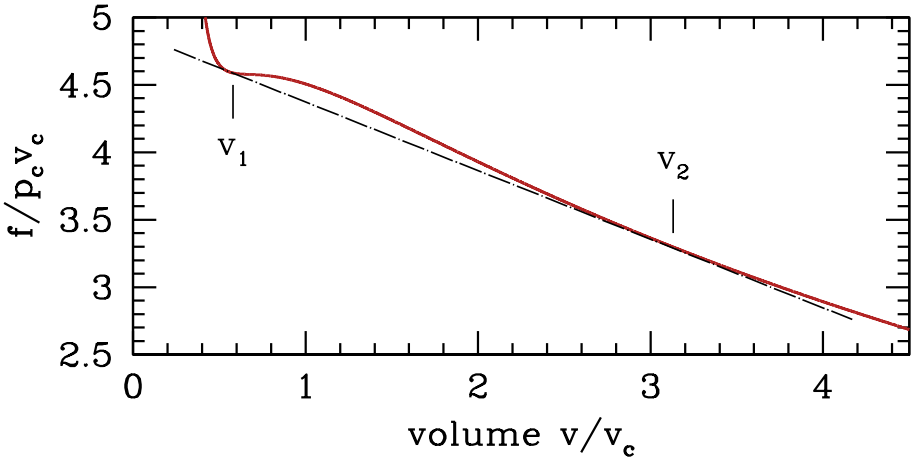



Figure 7.2: The molar free energy $f(T,v)$ of the van der Waals system at $T = 0.85\,T_{\rm c}$. The dot-dashed black line shows the Maxwell construction connecting molar volumes $v_{1,2}$ on opposite sides of the coexistence curve.

where $f_j = f(v_j, T)$. This function is given by the straight black line connecting the points at volumes $v_1$ and $v_2$ in fig. 7.2.

The two equations which give us $v_1$ and $v_2$ are

$$\left.\frac{\partial f}{\partial v}\right|_{v_1,T} = \left.\frac{\partial f}{\partial v}\right|_{v_2,T} = \frac{f(T,v_2) - f(T,v_1)}{(v_2 - v_1)} \quad . \tag{7.15}$$

Equivalently, in terms of the pressure, $p = -\big(\partial f/\partial v\big)_T$, these equations are equivalent to

$$p(T,v_1) = p(T,v_2) = \frac{1}{v_2 - v_1}\int\limits_{v_1}^{v_2} dv\, p(T,v) \quad . \tag{7.16}$$

This procedure is known as the *Maxwell construction*, and is depicted graphically in fig. 7.4. When the Maxwell construction is enforced, the isotherms resemble the curves in fig. 7.3. In this figure, all points within the purple shaded region have $\partial^2 f/\partial v^2 < 0$, hence this region is unstable to infinitesimal fluctuations. The boundary of this region is called the *spinodal*, and the spontaneous phase separation into two phases is a process known as *spinodal decomposition*. The dot-dashed orange curve, called the *coexistence curve*, marks the instability boundary for *nucleation*. In a nucleation process, an energy barrier must be overcome in order to achieve the lower free energy state. There is no energy barrier for spinodal decomposition – it is a spontaneous process.

Suppose we follow along an isotherm starting from the high molar volume (gas) phase. If $T > T_{\rm c}$, the volume $v$ decreases continuously as the pressure $p$ increases. If $T < T_{\rm c}$, then at the instant the isotherm first intersects the orange boundary curve in fig. 7.3, there is a discontinuous change in the molar volume from high (gas) to low (liquid). This discontinuous change is the hallmark of a *first order phase transition*. The volume discontinuity is $\Delta v \propto (T_{\rm c} - T)^{1/2}$. This is an example of a *critical behavior* in which the *order*

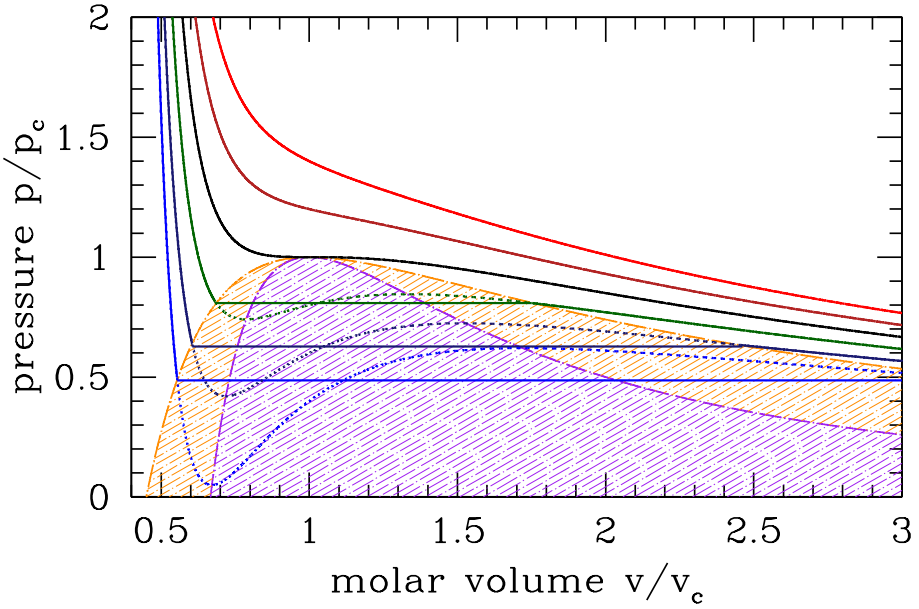



Figure 7.3: Pressure-volume isotherms for the van der Waals system, as in fig. 7.1, but corrected to account for the Maxwell construction. The boundary of the purple shaded region is the spinodal line $\bar{p}^*(\bar{v})$. The boundary of the orange shaded region is the stability boundary with respect to phase separation, and is called the *coexistence curve*.

*parameter* $\phi$, which in this case may be taken to be the difference $\phi = v_{\rm G} - v_{\rm L}$, behaves as a power law in $\big|T - T_{\rm c}\big|$, where $T_{\rm c}$ is the *critical temperature*. In this case, we have $\phi(T) \propto (T_{\rm c} - T)_+^\beta$, where $\beta = \frac{1}{2}$ is the exponent, and where $(T_{\rm c} - T)_+$ is defined to be $T_{\rm c} - T$ if $T < T_{\rm c}$ and 0 otherwise. Recall the isothermal compressibility is $\kappa_T = -v^{-1}\big(\partial v/\partial p\big)_T$. This is finite along the coexistence curve – it diverges only along the spinodal. It therefore diverges at the critical point, which lies at the intersection of the spinodal and the coexistence curve.

It is convenient to express the equation of state and the coexistence curve in terms of dimensionless variables. Writing

$$\bar{p} = \frac{p}{p_{\rm c}} \quad , \quad \bar{v} = \frac{v}{v_{\rm c}} \quad , \quad \bar{T} = \frac{T}{T_{\rm c}} \quad , \tag{7.17}$$

the dimensionless van der Waals equation of state takes the form

$$\bar{p} = \frac{8\bar{T}}{3\bar{v} - 1} - \frac{3}{\bar{v}^2} \quad . \tag{7.18}$$

### 7.2.2 Analytic form of the coexistence curve near the critical point

We write $v_{\rm L} = v_{\rm c} + w_{\rm L}$ and $v_{\rm G} = v_{\rm c} + w_{\rm G}$. One of our equations is $p(v_{\rm c} + w_{\rm L}, T) = p(v_{\rm c} + w_{\rm G}, T)$. Taylor expanding in powers of $w_{\rm L}$ and $w_{\rm G}$, we have

$$0 = p_v(v_{\rm c}, T)\,(w_{\rm G} - w_{\rm L}) + \tfrac{1}{2}\,p_{vv}(v_{\rm c}, T)\,\big(w_{\rm G}^2 - w_{\rm L}^2\big) + \tfrac{1}{6}\,p_{vvv}(v_{\rm c}, T)\,\big(w_{\rm G}^3 - w_{\rm L}^3\big) + \ldots \quad , \tag{7.19}$$

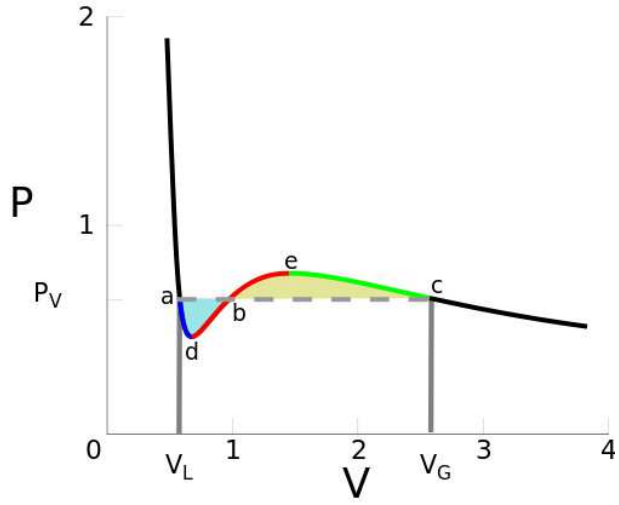



Figure 7.4: Maxwell construction in the $(v, p)$ plane. The system is absolutely unstable between volumes $v_{\rm d}$ and $v_{\rm e}$. For $v \in [v_{\rm a}, v_{\rm d}]$ or $v \in [v_{\rm e}, v_{\rm c}]$, the solution is unstable with respect to phase separation. Source: *Wikipedia*.

where

$$p_v \equiv \frac{\partial p}{\partial v} \quad , \quad p_{vv} \equiv \frac{\partial^2 p}{\partial v^2} \quad , \quad p_{vvv} \equiv \frac{\partial^3 p}{\partial v^3} \quad , \quad p_{vT} \equiv \frac{\partial^2 p}{\partial v\, \partial T} \quad , \quad \text{etc.} \tag{7.20}$$

The second equation we write as

$$\int\limits_{w_{\rm L}}^{w_{\rm G}} dw\, p(v_{\rm c}+w,T) = \tfrac{1}{2}(w_{\rm G}-w_{\rm L})\Big(p(v_{\rm c}+w_{\rm L},T)+p(v_{\rm c}+w_{\rm G},T)\Big) \quad . \tag{7.21}$$

Expanding in powers of $w_{\rm L}$ and $w_{\rm G}$, this becomes

$$\begin{aligned} &p(v_{\rm c},T)\,(w_{\rm G}-w_{\rm L}) + \tfrac{1}{2}\,p_v(v_{\rm c},T)\,\big(w_{\rm G}^2-w_{\rm L}^2\big) + \tfrac{1}{6}\,p_{vv}(v_{\rm c},T)\,\big(w_{\rm G}^3-w_{\rm L}^3\big) \\ &\qquad\qquad + \tfrac{1}{24}\,p_{vvv}(v_{\rm c},T)\,\big(w_{\rm G}^4-w_{\rm L}^4\big) + \tfrac{1}{120}\,p_{vvvv}(v_{\rm c},T)\,\big(w_{\rm G}^5-w_{\rm L}^5\big) + \ldots \\ &= \tfrac{1}{2}(w_{\rm G}-w_{\rm L})\Big\{2\,p(v_{\rm c},T) + p_v(v_{\rm c},T)\,(w_{\rm G}+w_{\rm L}) + \tfrac{1}{2}\,p_{vv}(v_{\rm c},T)\,\big(w_{\rm G}^2+w_{\rm L}^2\big) \\ &\qquad\qquad + \tfrac{1}{6}\,p_{vvv}(v_{\rm c},T)\,\big(w_{\rm G}^3+w_{\rm L}^3\big) + \tfrac{1}{24}\,p_{vvvv}(v_{\rm c},T)\,\big(w_{\rm G}^4+w_{\rm L}^4\big) + \ldots\Big\} \end{aligned} \tag{7.22}$$

Subtracting the LHS from the RHS, we find that we can then divide by $\frac{1}{6}\big(w_{\rm G}^2-w_{\rm L}^2\big)$, resulting in

$$0 = p_{vv}(v_{\rm c},T) + \tfrac{1}{2}\,p_{vvv}(v_{\rm c},T)\,(w_{\rm G}+w_{\rm L}) + \tfrac{1}{20}\,p_{vvvv}(v_{\rm c},T)\,\big(3w_{\rm G}^2+4w_{\rm G}w_{\rm L}+3w_{\rm L}^2\big) + \mathcal{O}\big(w_{\rm G,L}^3\big) \quad . \tag{7.23}$$

We now define $w_\pm \equiv w_{\rm G} \pm w_{\rm L}$. In terms of these variables, eqns. 7.19 and 7.23 become

$$\begin{aligned} 0 &= p_v(v_{\rm c},T) + \tfrac{1}{2}\,p_{vv}(v_{\rm c},T)\,w_+ + \tfrac{1}{8}\,p_{vvv}(v_{\rm c},T)\,\big(w_+^2+\tfrac{1}{3}\,w_-^2\big) + \mathcal{O}\big(w_\pm^3\big) \\ 0 &= p_{vv}(v_{\rm c},T) + \tfrac{1}{2}\,p_{vvv}(v_{\rm c},T)\,w_+ + \tfrac{1}{8}\,p_{vvvv}(v_{\rm c},T)\,\big(w_+^2+\tfrac{1}{5}\,w_-^2\big) + \mathcal{O}\big(w_\pm^3\big) \quad . \end{aligned} \tag{7.24}$$

We now write $T = T_{\rm c} + \Theta$ and evaluate $w_\pm$ as expansions in $\Theta$. Note that $p_v(v_{\rm c}, T_{\rm c}) = p_{vv}(v_{\rm c}, T_{\rm c}) = 0$, since the critical point is an inflection point in the $(v, p)$ plane. Thus, we have $p_v(v_{\rm c}, T) = p^0_{vT}\,\Theta + \mathcal{O}(\Theta^2)$, where $p^0_{vT} = p_{vT}(v_{\rm c}, T_{\rm c})$. We can then see that $w_- \propto \sqrt{-\Theta}$, while $w_+ \propto \Theta$ as $\Theta \to 0^-$, and we have

$$\begin{aligned} 0 &= p^0_{vT}\,\Theta + \tfrac{1}{24}\,p^0_{vvv}\,w_-^2 + \mathcal{O}(\Theta^2) \\ 0 &= p^0_{vvT}\,\Theta + \tfrac{1}{2}\,p^0_{vvv}\,w_+ + \tfrac{1}{40}\,p^0_{vvvv}\,w_-^2 + \mathcal{O}(\Theta^2) \quad . \end{aligned} \tag{7.25}$$

Thus,

$$\begin{aligned} w_- &= \left(\frac{24\,p^0_{vT}}{p^0_{vvv}}\right)^{1/2}\sqrt{-\Theta} + \ldots \\ w_+ &= \left(\frac{6\,p^0_{vT}\,p^0_{vvvv} - 10\,p^0_{vvv}\,p^0_{vvT}}{5\,(p^0_{vvv})^2}\right)\Theta + \ldots \quad . \end{aligned} \tag{7.26}$$

We then have

$$\begin{aligned} w_{\rm L} &= -\left(\frac{6\,p^0_{vT}}{p^0_{vvv}}\right)^{1/2}\sqrt{-\Theta} + \left(\frac{3\,p^0_{vT}\,p^0_{vvvv} - 5\,p^0_{vvv}\,p^0_{vvT}}{5\,(p^0_{vvv})^2}\right)\Theta + \mathcal{O}\big(\Theta^{3/2}\big) \\ w_{\rm G} &= \left(\frac{6\,p^0_{vT}}{p^0_{vvv}}\right)^{1/2}\sqrt{-\Theta} + \left(\frac{3\,p^0_{vT}\,p^0_{vvvv} - 5\,p^0_{vvv}\,p^0_{vvT}}{5\,(p^0_{vvv})^2}\right)\Theta + \mathcal{O}\big(\Theta^{3/2}\big) \quad . \end{aligned} \tag{7.27}$$

Close to the critical point, the dimensionless equation of state may be written as $\pi = \pi(\epsilon, t)$, where

$$\bar p = 1 + \pi \quad , \quad \bar v = 1 + \epsilon \quad , \quad \bar T = 1 + t \quad , \tag{7.28}$$

where $\pi(0,0) = 0$. Equivalently,

$$\pi = \frac{p - p_{\rm c}}{p_{\rm c}} \quad , \quad \epsilon = \frac{v - v_{\rm c}}{v_{\rm c}} \quad , \quad t = \frac{T - T_{\rm c}}{T_{\rm c}} = \frac{\Theta}{T_{\rm c}} \quad . \tag{7.29}$$

Here $\pi$, $\epsilon$, and $t$ are, respectively, the dimensionless deviations of pressure, molar volume, and temperature from their critical point values. For the van der Waals equation of state in eqn. 7.18, we have

$$\begin{aligned} \pi(\epsilon, t) &= \frac{8(1+t)}{2+3\epsilon} - \frac{3}{(1+\epsilon)^2} - 1 \\ &= 4t - 6t\epsilon + 9\epsilon^2 t - \tfrac{3}{2}\epsilon^3 - \tfrac{27}{2}\epsilon^3 t + \tfrac{21}{4}\epsilon^4 + \tfrac{81}{4}\epsilon^4 t - \tfrac{99}{8}\epsilon^5 + \ldots \quad . \end{aligned} \tag{7.30}$$

Expressed in these dimensionless quantities, eqns. 7.27 take the form

$$\epsilon_{\rm L,G} = \mp\left(\frac{6\,\pi^0_{\epsilon t}}{\pi^0_{\epsilon\epsilon\epsilon}}\right)^{1/2}(-t)^{1/2} + \left(\frac{3\,\pi^0_{\epsilon t}\,\pi^0_{\epsilon\epsilon\epsilon\epsilon} - 5\,\pi^0_{\epsilon\epsilon\epsilon}\,\pi^0_{\epsilon\epsilon t}}{5\,(\pi^0_{\epsilon\epsilon\epsilon})^2}\right)t + \mathcal{O}\big((-t)^{3/2}\big) \quad . \tag{7.31}$$

For the van der Waals system , we have

$$\pi^0_{\epsilon t} = -6 \quad , \quad \pi^0_{\epsilon\epsilon t} = 18 \quad , \quad \pi^0_{\epsilon\epsilon\epsilon} = -9 \quad , \quad \pi^0_{\epsilon\epsilon\epsilon\epsilon} = 126 \quad , \tag{7.32}$$

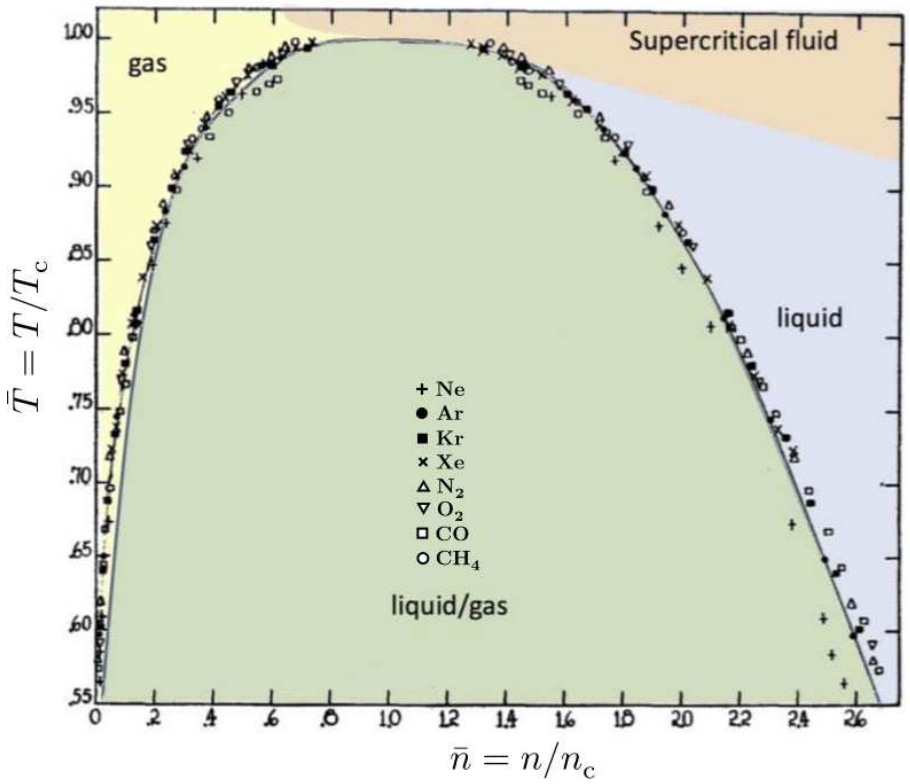



Figure 7.5: Universality of the liquid-gas transition for eight different atomic and molecular fluids, from E. A. Guggenheim, *J. Chem. Phys.* **13**, 253 (1945). Dimensionless temperature $\bar{T} = T/T_c$ *versus* dimensionless number density $\bar{n} = n/n_c = v_c/v$ is shown. The van der Waals / mean field theory gives $\Delta n = n_{\rm liquid} - n_{\rm gas} \propto (-t)^{1/2}$, while experiments show a result closer to $\Delta n \propto (-t)^{1/3}$. Here $t \equiv \bar{T} - 1 = (T - T_c)/T_c$ is the dimensionless temperature deviation with respect to the critical point. (Image adapted from Matthew Schwartz's Harvard lecture notes, adapted from Guggenheim 1945.)

where the derivatives are evaluated at critical point, where $\epsilon = t = \pi = 0$. Thus, for van der Waals,

$$\epsilon_{\mathsf{L,G}} = \mp 2\sqrt{-t} - \tfrac{5}{18}\,t + \ldots \quad . \tag{7.33}$$

We identify the difference $\Delta v \equiv v_{\mathsf{G}} - v_{\mathsf{L}}$ as the *order parameter* for the transition which occurs at $T = T_c$. We see that the order parameter behaves as a power law for $T$ just below the critical point, with $\Delta v \propto (-t)^\beta$ and $\beta = \frac{1}{2}$, which is the *order parameter critical exponent*.

The spinodal boundary for the vdW system is then given by the solution to

$$0 = \frac{\partial \pi}{\partial \epsilon} = -6t + 18\epsilon t - \tfrac{9}{2}\epsilon^2 - \tfrac{81}{2}\epsilon^2 t + 21\epsilon^3 + \ldots \quad . \tag{7.34}$$

For the spinodal, it is easy to see that the lowest order nontrivial solution is $\epsilon = \mp\frac{2}{\sqrt{3}}\sqrt{-t}$. For the coexistence curve, we found $\epsilon_{\mathsf{L,G}} = \mp 2\sqrt{-t}$.

Fig. 7.5 shows the universality of the liquid-gas transition for eight different fluids: Ne, Ar, Kr, Xe, $N_2$, $O_2$, CO, and $CH_4$. The experimental coexistence curve expressed in dimensionless variables $\bar{n} = 1/\bar{v}$

and $\bar{T} = T/T_{\rm c}$ is fairly well-fit to the curve[5]

$$\bar{n}_{\rm L,G}(t) = 1 \pm \tfrac{7}{4}(1-\bar{T})^{1/3} + \tfrac{3}{4}(1-\bar{T}) \tag{7.35}$$

which shows that the critical exponent $\beta$ is much closer to $\beta = \frac{1}{3}$ than to the vdW value $\beta^{\rm vdW} = \frac{1}{2}$. The van der Waals equation is in essence a *mean field theory* of the liquid-gas transition.

### 7.2.3 History of the van der Waals equation

The van der Waals equation of state first appears in van der Waals' 1873 PhD thesis[6], *"Over de Continuïteit van den Gas - en Vloeistoftoestand"* ("On the continuity of the gas and liquid state"). In his Nobel lecture[7], van der Waals writes of how he was inspired by Rudolf Clausius' 1857 treatise on the nature of heat, where it is posited that a gas in fact consists of microscopic particles whizzing around at high velocities. van der Waals reasoned that liquids, which result when gases are compressed, also consist of 'small moving particles': *"Thus I conceived the idea that there is no essential difference between the gaseous and the liquid state of matter..."*

Clausius' treatise showed how his kinetic theory of heat was consistent with Boyle's law for gases ($pV = \text{constant}$ at fixed temperature). van der Waals pondered why this might fail for the non-dilute liquid phase, and he reasoned that there were two principal differences: inter-particle attraction and excluded volume. These considerations prompted him to posit his famous equation,

$$p = \frac{RT}{v-b} - \frac{a}{v^2} \quad . \tag{7.36}$$

The first term on the RHS accounts for excluded volume effects, and the second for mutual attractions.

In the limiting case of $p \to \infty$, the molar volume approaches $v = b$. On physical grounds, one might expect $b = v_0/\zeta$, where $v_0 = N_{\rm A}\,\omega_0$ is $N_{\rm A}$ times the volume $\omega_0$ of a single molecule, and the *packing fraction* is $\zeta = N\omega_0/V = v_0/v$, which is the ratio of the total molecular volume to the total system volume. In three dimensions, the maximum possible packing fraction is for fcc and hcp lattices, each of which have coordination number 12, with $\zeta_{\rm max} = \frac{\pi}{3\sqrt{2}} = 0.74048$. Dense random packing results in $\zeta_{\rm drp} = 0.634$. Expanding the vdW equation of state in inverse powers of $v$ yields

$$p = \frac{RT}{v} + \left(b - \frac{a}{RT}\right)\cdot\frac{RT}{v^2} + \mathcal{O}\big(v^{-3}\big) \quad , \tag{7.37}$$

and we read off the second virial coefficient $B_2 = \big(b - \frac{a}{RT}\big)/N_{\rm A}$. For hard spheres, $a = 0$, and the result $B_2 = 4\omega_0$ from the Mayer cluster expansion corresponds to $b_{\rm Mayer} = 4v_0$ , which is larger than the result from even the loosest regular sphere packing, *i.e.* that for a cubic lattice, with $\zeta_{\rm cub} = \frac{\pi}{6}$.

[5] See M. Schwartz, https://scholar.harvard.edu/files/schwartz/files/9-phases.pdf.

[6] Johannes Diderik van der Waals, the eldest of ten children, was the son of a carpenter. As a child he received only a primary school education. He worked for a living until age 25, and was able to enroll in a three-year industrial evening school for working class youth. Afterward he continued his studies independently, in his spare time, working as a teacher. By the time he obtained his PhD, he was 36 years old. He received the Nobel Prize for Physics in 1910.

[7] See http://www.nobelprize.org/nobel_prizes/physics/laureates/1910/waals-lecture.pdf

**The law of corresponding states**

Another of van der Waals' great achievements was his articulation of the *law of corresponding states*. Recall that the van der Waals equation of state, when written in terms of dimensionless quantities $\bar{p} = p/p_{\rm c}$, $\bar{v} = v/v_{\rm c}$, and $\bar{T} = T/T_{\rm c}$, takes the form of eqn. 7.18. Thus, while the $a$ and $b$ parameters are specific to each fluid – see Tab. 7.1 – when written in terms of these scaled dimensionless variables, the equation of state and all its consequent properties (*i.e.* the liquid-gas phase transition) are *universal*.

The van der Waals equation is best viewed as semi-phenomenological. Interaction and excluded volume effects surely are present, but the van der Waals equation itself only captures them in a very approximate way. It is applicable to gases, where it successfully predicts features that are not present in ideal systems (*e.g.* throttling). It is of only qualitative and pedagogical use in the study of fluids, the essential physics of which lies in the behavior of quantities like the pair distribution function $g(r)$. As we saw in chapter 6, any adequate first principles derivation of $g(r)$ - a function which can be measured in scattering experiments - involves rather complicated approximation schemes to close the BBGKY hierarchy. Else one must resort to numerical simulations such as the Monte Carlo method. Nevertheless, the lessons learned from the van der Waals system are invaluable and they provide us with a first glimpse of what is going on in the vicinity of a phase transition, and how nonanalytic behavior, such as $v_{\rm G} - v_{\rm L} \propto (T_{\rm c} - T)^\beta$ with noninteger exponent $\beta$ may result due to singularities in the free energy at the critical point.

## 7.3 Fluids, Magnets, and the Ising Model

### 7.3.1 Lattice gas description of a fluid

The usual description of a fluid follows from a continuum Hamiltonian of the form

$$\hat{H}(\boldsymbol{p}, \boldsymbol{x}) = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} + \sum_{i<j} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \quad . \tag{7.38}$$

The potential $u(\boldsymbol{r})$ is typically central, depending only on the magnitude $|\boldsymbol{r}|$, and short-ranged. Now consider a discretized version of the fluid, in which we divide up space into cells (cubes, say), each of which can accommodate at most one fluid particle (due to excluded volume effects). That is, each cube has a volume on the order of $a^3$, where $a$ is the diameter of the fluid particles. In a given cube $i$ we set the occupancy $n_i = 1$ if a fluid particle is present and $n_i = 0$ if there is no fluid particle present. We then have that the potential energy is

$$U = \sum_{i<j} u(\boldsymbol{x}_i - \boldsymbol{x}_j) = \tfrac{1}{2} \sum_{\boldsymbol{R} \neq \boldsymbol{R}'} V_{\boldsymbol{R},\boldsymbol{R}'}\, n_{\boldsymbol{R}}\, n_{\boldsymbol{R}'} \quad , \tag{7.39}$$

where $V_{\boldsymbol{R},\boldsymbol{R}'} \approx u(\boldsymbol{R} - \boldsymbol{R}')$, where $R_k$ is the position at the center of cube $k$. The grand partition function is then approximated as

$$\Xi(T, V, \mu) \approx \sum_{\{n_{\boldsymbol{R}}\}} \left( \prod_{\boldsymbol{R}} \xi^{n_{\boldsymbol{R}}} \right) \exp\left( -\frac{1}{2k_{\rm B}T} \sum_{\boldsymbol{R} \neq \boldsymbol{R}'} V_{\boldsymbol{R},\boldsymbol{R}'}\, n_{\boldsymbol{R}}\, n_{\boldsymbol{R}'} \right) \quad , \tag{7.40}$$

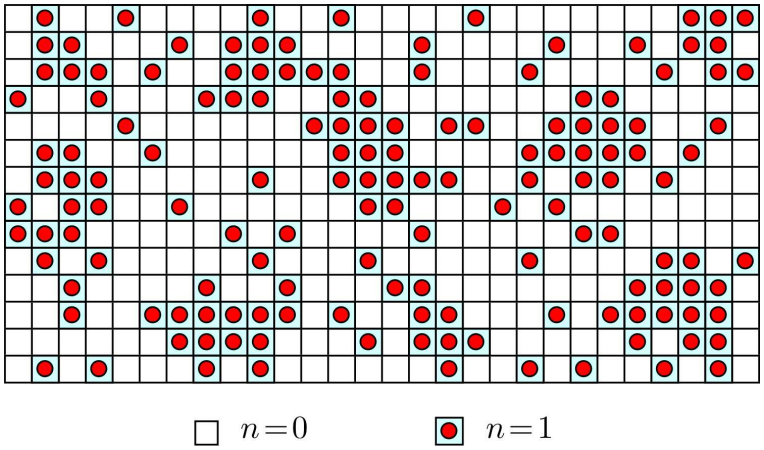



Figure 7.6: The lattice gas model. An occupied cell corresponds to $n = 1$ ($\sigma = +1$), and a vacant cell to $n = 0$ ($\sigma = -1$).

where $\xi = e^{\beta\mu}\,\lambda_T^{-d}\,a^d$, and where $a$ is the side length of each cube (chosen to be on the order of the hard sphere diameter). The $\lambda_T^{-d}$ factor arises from the integration over the momenta. Note $N = \sum_{\boldsymbol{R}} n_{\boldsymbol{R}}$ is the total number of fluid particles, so $\prod_{\boldsymbol{R}} \xi^{\,n_{\boldsymbol{R}}} = \xi^N$ where $\xi = e^{\beta\mu}(a/\lambda_T)^d$.

Thus, we arrive at $\Xi = \mathsf{Tr}\,\exp(-\beta\hat{H}_{\mathsf{LG}})$ where $\hat{H}_{\mathsf{LG}}$ is the *lattice gas Hamiltonian*,

$$\begin{aligned}\hat{H}_{\mathsf{LG}} &= \tfrac{1}{2}\sum_{\boldsymbol{R}\neq\boldsymbol{R}'} V_{\boldsymbol{R},\boldsymbol{R}'}\,n_{\boldsymbol{R}}\,n_{\boldsymbol{R}'} - k_{\rm B}T\log\xi\sum_{\boldsymbol{R}} n_{\boldsymbol{R}} \\ &= -\tfrac{1}{2}\sum_{\boldsymbol{R}\neq\boldsymbol{R}'} J_{\boldsymbol{R},\boldsymbol{R}'}\,\sigma_{\boldsymbol{R}}\,\sigma_{\boldsymbol{R}'} - H\sum_{\boldsymbol{R}}\sigma_{\boldsymbol{R}} + E_0 \quad ,\end{aligned} \tag{7.41}$$

where $\sigma_{\boldsymbol{R}} \equiv 2n_{\boldsymbol{R}} - 1$ is a spin variable taking the possible values $\{-1, +1\}$, and

$$J_{\boldsymbol{R},\boldsymbol{R}'} = -\tfrac{1}{4}u(\boldsymbol{R}-\boldsymbol{R}') \qquad , \qquad H = \mu + dk_{\rm B}T\log(a/\lambda_T) - \tfrac{1}{4}{\sum_{\boldsymbol{R}}}'u(\boldsymbol{R}) \tag{7.42}$$

and $E_0 = \frac{1}{8}N_{\rm c}\,{\sum_{\boldsymbol{R}}}'u(\boldsymbol{R})$, where $N_{\rm c} = V/a^d$ is the number of unit cells, and where the prime on the sum indicates that $\boldsymbol{R} = 0$ is to be excluded. For the Lennard-Jones system, $u(\boldsymbol{R}-\boldsymbol{R}') < 0$ is due to the attractive tail of the potential, hence $J_{\boldsymbol{R},\boldsymbol{R}'} > 0$, which prefers alignment of the spins $\sigma_{\boldsymbol{R}}$ and $\sigma_{\boldsymbol{R}'}$. This interaction is therefore *ferromagnetic*. The spin Hamiltonian in eqn. 7.41 is recognized as the Ising model.

### 7.3.2 Phase diagrams and critical exponents

The physics of the liquid-gas transition in fact has a great deal in common with that of the transition between a magnetized and unmagnetized state of a magnetic system. The correspondences are[8] $p \leftrightarrow H$

[8] $H$ is more properly analogous to $\mu$. However, since $\mu = \mu(p,T)$, $H$ can also be regarded as analogous to $p$. Note also that $\beta p = z\lambda_T^{-d}$ for the ideal gas, in which case $\xi = z\,(a/\lambda_T)^d$ is proportional to $p/k_{\rm B}T$.

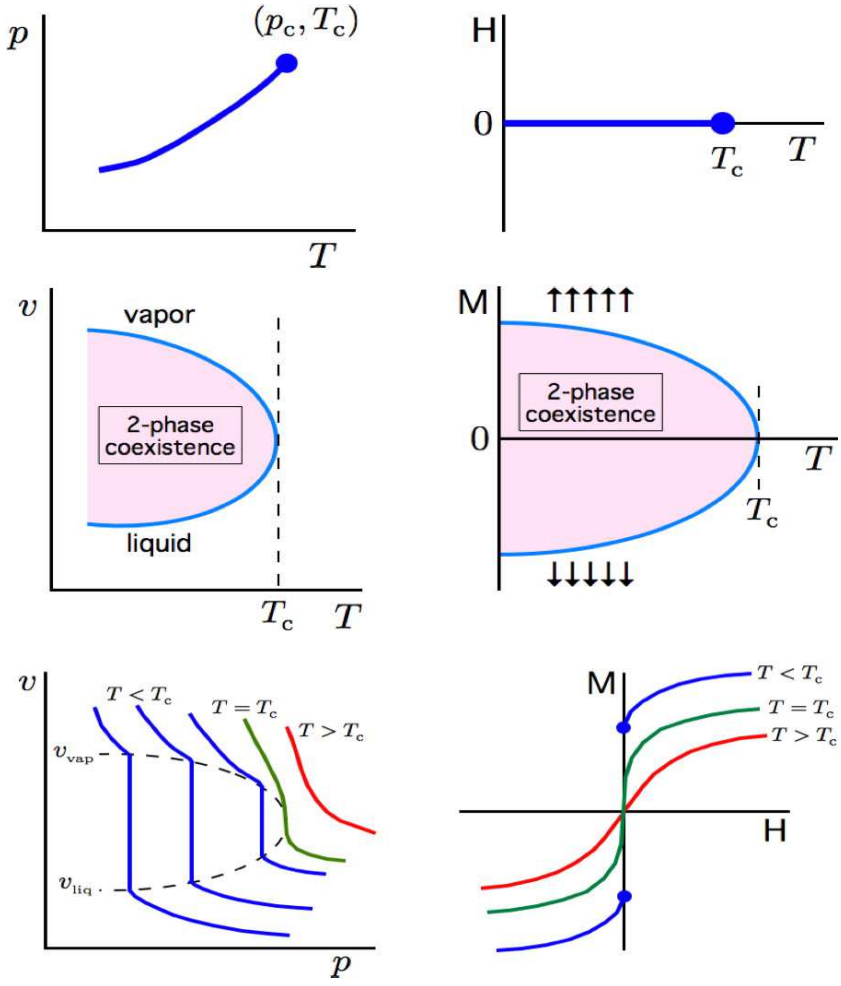



Figure 7.7: Comparison of the liquid-gas phase diagram with that of the Ising ferromagnet.

and $v \leftrightarrow m$, where $m$ is the magnetization density, defined here to be the total magnetization $M$ divided by the number of lattice sites[9] $N_c$ , *viz.*

$$m = \frac{M}{N_c} = \frac{1}{N_c}\sum_{\boldsymbol{R}} \langle \sigma_{\boldsymbol{R}} \rangle \quad . \tag{7.43}$$

Sketches of the phase diagrams are reproduced in fig. 7.7. Of particular interest is the *critical point*, which occurs at $(T_c, p_c)$ in the fluid system and $(T_c, H_c)$ in the magnetic system, with $H_c = 0$ by symmetry.

In the fluid, the coexistence curve in the $(p, T)$ plane separates high density (liquid) and low density (vapor) phases. The specific volume $v$ (or the density $n = v^{-1}$) jumps discontinuously across the coexistence curve. In the magnet, the coexistence curve in the $(H, T)$ plane separates positive magnetization and negative magnetization phases. The magnetization density $m$ jumps discontinuously across the coexistence curve. For $T > T_c$, the latter system is a *paramagnet*, in which the magnetization varies

[9]Note the distinction between the number of lattice sites $N_c$ and the number of occupied cells $N$. According to our definitions, $N = \frac{1}{2}(M + N_c)$.

smoothly as a function of $H$. This behavior is most apparent in the bottom panel of the figure, where $v(p)$ and $m(H)$ curves are shown.

For $T < T_{\rm c}$, the fluid exists in a *two phase region*, which is spatially inhomogeneous, supporting local regions of high and low density. There is no stable homogeneous thermodynamic phase for $(T, v)$ within the two phase region shown in the middle left panel. Similarly, for the magnet, there is no stable homogeneous thermodynamic phase at fixed temperature $T$ and magnetization $m$ if $(T, m)$ lies within the coexistence region. Rather, the system consists of blobs where the spin is predominantly up, and blobs where the spin is predominantly down.

Note also the analogy between the isothermal compressibility $\kappa_T$ and the isothermal susceptibility $\chi_T$:

$$\kappa_T = -\frac{1}{v}\left(\frac{\partial v}{\partial p}\right)_T \quad , \quad \chi_T = \left(\frac{\partial m}{\partial H}\right)_T \quad , \tag{7.44}$$

with $\kappa_T(T_{\rm c}, p_{\rm c}) = \infty$ and $\chi_T(T_{\rm c}, H_{\rm c}) = \infty$.

The *order parameter* for a second order phase transition is a quantity which vanishes in the disordered phase and is finite in the ordered phase. For the fluid, the order parameter can be chosen to be $\Psi \propto (v_{\rm vap} - v_{\rm liq})$, the difference in the specific volumes of the vapor and liquid phases. In the vicinity of the critical point, the system exhibits power law behavior in many physical quantities, *viz.*

$$m(T, H_{\rm c}) \sim \big(T_{\rm c} - T\big)_+^{\beta} \quad , \quad \chi(T, H_{\rm c}) \sim |T - T_{\rm c}|^{-\gamma} \quad , \quad C(T, H_{\rm c}) \sim |T - T_{\rm c}|^{-\alpha} \quad , \quad m(T_{\rm c}, H) \sim \pm |H|^{1/\delta} \quad . \tag{7.45}$$

The quantities $\alpha$, $\beta$, $\gamma$, and $\delta$ are the *critical exponents* associated with the transition. These exponents satisfy certain equalities, such as the Rushbrooke and Griffiths relations:

$$\alpha + 2\beta + \gamma = 2 \quad \text{(Rushbrooke)} \qquad , \qquad \beta + \gamma = \beta\delta \quad \text{(Griffiths)} \quad . \tag{7.46}$$

Originally such relations were derived as inequalities, and only after the advent of scaling and renormalization group theories it was realized that they held as equalities.

In addition to the exponents $\alpha$, $\beta$, $\gamma$, and $\delta$, one defines the *correlation length exponent* $\nu$ from the behavior of the two-point correlation function $C(\boldsymbol{r}, T) = \langle \psi(0)\, \psi(\boldsymbol{r}) \rangle$, where $\psi(\boldsymbol{r})$ is a local operator, such as the local density in a fluid or the local spin polarization in a magnet[10]. In the limits $T \to T_{\rm c}$ and $H \to H_{\rm c} = 0$, one has

$$C(\boldsymbol{r}, T, H) = r^{-(d-2+\eta)}\, \phi\big(r/\xi(T), H/\xi_H(T)\big) \quad , \tag{7.47}$$

where $\eta$ is the *anomalous exponent*, $\phi(r/\xi, H/\xi_H)$ is a *scaling function*, $\xi(T) \propto |T - T_{\rm c}|^{-\nu}$ is the *correlation length*, and $\xi_H(T) \propto |T - T_{\rm c}|^{\Delta}$, with $\Delta = \beta\delta$, is a field scale. As we have seen, for the lattice gas system the effective magnetic field $H$ is a proxy for the pressure or the chemical potential. Along with the new exponents $\eta$ and $\nu$ come additional exponent relations,

$$(2 - \eta)\nu = \gamma \qquad , \qquad d\,\nu = 2 - \alpha \quad \text{(hyperscaling)} \quad . \tag{7.48}$$

Thus there are three relations among the six critical exponents $\alpha$, $\beta$, $\gamma$, $\delta$, $\eta$, and $\nu$, which entails that there are three independent values among the six.

[10] The local 'order parameter field' $\psi(\boldsymbol{r})$ may carry vector or tensor indices. In general, it transforms as the fundamental representation of the global symmetry group $G$.

### 7.3.3 Gibbs-Duhem relation for magnetic systems

Homogeneity of $E(S, M, N)$ means $E = TS + HM + \mu N$, by Euler's theorem. After invoking the First Law $dE = T\,dS + H\,dM + \mu\,dN$, we have

$$S\,dT + M\,dH + N\,d\mu = 0 \quad . \tag{7.49}$$

Now consider two magnetic phases in coexistence. We must have $d\mu_1 = d\mu_2$, hence

$$d\mu_1 = -s_1\,dT - m_1\,dH = -s_2\,dT - m_2\,dH = d\mu_2 \quad , \tag{7.50}$$

where $m = M/N$ is the magnetization per site and $s = S/N$ is the specific entropy. Thus, we obtain the Clapeyron equation for magnetic systems,

$$\left(\frac{dH}{dT}\right)_{\rm coex} = -\frac{s_1 - s_2}{m_1 - m_2} \quad . \tag{7.51}$$

Thus, if $m_1 \neq m_2$ and $\big(dH/dT\big)_{\rm coex} = 0$, then we must have $s_1 = s_2$, which says that there is *no latent heat associated with the transition*. This absence of latent heat is a consequence of the *symmetry* which guarantees that $G(T, H, N) = G(T, -H, N)$.

Recall our discussion in §2.11.2 of the Clausius-Clapeyron relation for liquid-gas systems. From $G = E - TS + pV = G(T, p, N)$, the differential of the Gibbs free energy per particle, $\mu = G/N$, is given by $d\mu = -s\,dT + v\,dp$, where $v = V/N$ is the volume per particle[11]. This leads to the Clapeyron relation,

$$\left(\frac{dp}{dT}\right)_{\rm coex} = \frac{s_2 - s_1}{v_2 - v_1} = \frac{\ell}{T\,\Delta v} \quad , \tag{7.52}$$

which determines the slope of the coexistence curve in the $(T, p)$ plane.

### 7.3.4 Order-disorder transitions

Another application of the Ising model lies in the theory of order-disorder transitions in alloys. Examples include $Cu_3Au$, CuZn, and other compounds. In CuZn, the Cu and Zn atoms occupy sites of a body centered cubic (BCC) lattice, forming an alloy known as $\beta$-brass. Below $T_{\rm c} \simeq 740\,{\rm K}$, the atoms are ordered, with the Cu preferentially occupying one simple cubic sublattice and the Zn preferentially occupying the other.

The energy is a sum of pairwise interactions, with a given link contributing $\varepsilon_{\rm AA}$, $\varepsilon_{\rm BB}$, or $\varepsilon_{\rm AB}$, depending on whether it is an A-A, B-B, or A-B/B-A link. Here A and B represent Cu and Zn, respectively. Thus, we can write the energy of the link $\langle ij \rangle$ as

$$E_{ij} = \varepsilon_{\rm AA}\,P_i^{\rm A}\,P_j^{\rm A} + \varepsilon_{\rm BB}\,P_i^{\rm B}\,P_j^{\rm B} + \varepsilon_{\rm AB}\left(P_i^{\rm A}\,P_j^{\rm B} + P_i^{\rm B}\,P_j^{\rm A}\right) \quad , \tag{7.53}$$

[11] In §2.11.2 we considered, equivalently, the differential of the molar free energy $g = N_{\rm A}\mu$, with $v$ the molar volume.

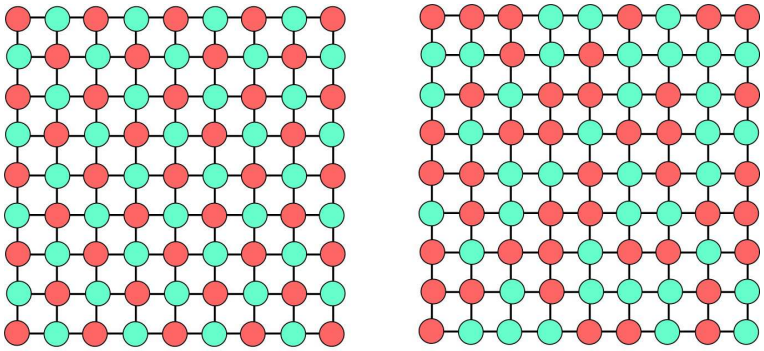



Figure 7.8: Order-disorder transition on the square lattice. Below $T = T_{\rm c}$, order develops spontaneously on the two $\sqrt{2} \times \sqrt{2}$ sublattices. There is perfect sublattice order at $T = 0$ (left panel).

where

$$P_i^{\rm A} = \tfrac{1}{2}(1+\sigma_i) = \begin{cases} 1 & \text{if site } i \text{ contains Cu} \\ 0 & \text{if site } i \text{ contains Zn} \end{cases} \quad , \quad P_i^{\rm B} = \tfrac{1}{2}(1-\sigma_i) = \begin{cases} 1 & \text{if site } i \text{ contains Zn} \\ 0 & \text{if site } i \text{ contains Cu} \end{cases} \quad .$$

The Hamiltonian is then

$$\begin{aligned} \hat{H} = \sum_{\langle ij \rangle} E_{ij} &= \sum_{\langle ij \rangle} \left\{ \tfrac{1}{4}\big(\varepsilon_{\rm AA} + \varepsilon_{\rm BB} - 2\varepsilon_{\rm AB}\big)\,\sigma_i\,\sigma_j + \tfrac{1}{4}\big(\varepsilon_{\rm AA} - \varepsilon_{\rm BB}\big)\,(\sigma_i + \sigma_j) + \tfrac{1}{4}\big(\varepsilon_{\rm AA} + \varepsilon_{\rm BB} + 2\varepsilon_{\rm AB}\big) \right\} \\ &= -J \sum_{\langle ij \rangle} \sigma_i\,\sigma_j - H \sum_{i=1}^{N} \sigma_i + E_0 \quad , \end{aligned} \tag{7.54}$$

where the exchange constant $J$ and the magnetic field $H$ are given by

$$J = \tfrac{1}{4}\big(2\varepsilon_{\rm AB} - \varepsilon_{\rm AA} - \varepsilon_{\rm BB}\big) \qquad , \qquad H = \tfrac{1}{4}\big(\varepsilon_{\rm BB} - \varepsilon_{\rm AA}\big) \quad , \tag{7.55}$$

and $E_0 = \frac{1}{8}Nz\big(\varepsilon_{\rm AA} + \varepsilon_{\rm BB} + 2\varepsilon_{\rm AB}\big)$, where $N$ is the total number of lattice sites and $z = 8$ is the *lattice coordination number*, which is the number of nearest neighbors of any given site.

Note that

$$\begin{aligned} 2\varepsilon_{\rm AB} > \varepsilon_{\rm AA} + \varepsilon_{\rm BB} \quad &\Longrightarrow \quad J > 0 \quad \text{(ferromagnetic)} \\ 2\varepsilon_{\rm AB} < \varepsilon_{\rm AA} + \varepsilon_{\rm BB} \quad &\Longrightarrow \quad J < 0 \quad \text{(antiferromagnetic)} \quad . \end{aligned} \tag{7.56}$$

The antiferromagnetic case is depicted in fig. 7.8.

## 7.4 Mean Field Theory

### 7.4.1 The mean field *Ansatz*

Consider the Ising model Hamiltonian,

$$\hat{H} = -J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j - H \sum_{i=1}^{N} \sigma_i \quad , \tag{7.57}$$

where the first sum on the RHS is over all links of the lattice. Each spin can be either 'up' ($\sigma = +1$) or 'down' ($\sigma = -1$). We further assume that the spins are located on a Bravais lattice[12] and that the coupling between sites $i$ and $j$ is $J_{ij} = J\big(|\boldsymbol{R}_i - \boldsymbol{R}_j|\big)$, where $\boldsymbol{R}_i$ is the position of the $i^{\rm th}$ spin. Here we assume $J_{ij} = 0$ beyond nearest neighbors; later we will assume a more general set of couplings. We also set the self-interaction terms to zero: $J_{ii} = 0$ for all $i$.

On each site $i$ we decompose $\sigma_i$ into a contribution from its thermodynamic average and a fluctuation:

$$\sigma_i = \langle \sigma_i \rangle + \delta\sigma_i \quad . \tag{7.58}$$

We will write $\langle \sigma_i \rangle \equiv m$, the local magnetization (dimensionless), and assume that $m$ is independent of position $i$. Then

$$\begin{aligned} \sigma_i \, \sigma_j &= (m + \delta\sigma_i)\,(m + \delta\sigma_j) \\ &= m^2 + m\,(\delta\sigma_i + \delta\sigma_j) + \delta\sigma_i \, \delta\sigma_j \\ &= -m^2 + m\,(\sigma_i + \sigma_j) + \delta\sigma_i \, \delta\sigma_j \quad . \end{aligned} \tag{7.59}$$

The last term on the RHS of the second equation above is quadratic in the fluctuations, and we assume this to be negligibly small. This neglect of the fluctuations is the *mean field Ansatz*, and results in the *mean field Hamiltonian*

$$\hat{H}_{\rm MF} = \tfrac{1}{2} N z J \, m^2 - \big(H + zJm\big) \sum_{i=1}^{N} \sigma_i \quad , \tag{7.60}$$

where $N$ is the total number of lattice sites. Note that the number of nearest-neighbor links is given by $N_{\rm links} = \frac{1}{2} zN$, where $z$ is the lattice coordination number, *i.e.* the number of nearest neighbors for any given site, and $N$ is the number of lattice sites[13]. The first term is a constant, although the value of $m$ is yet to be determined. The Boltzmann weights are then completely determined by the second term, which is just what we would write down for a Hamiltonian of *noninteracting* spins in an effective 'mean field'

$$H_{\rm eff} = H + zJm \quad . \tag{7.61}$$

[12] A Bravais lattice is one in which any site is equivalent to any other site through an appropriate discrete translation. Examples of Bravais lattices include the linear chain, square, triangular, simple cubic, face-centered cubic, *etc.* lattices. The honeycomb lattice is not a Bravais lattice, because there are two sets of inequivalent sites – those in the center of a Y and those in the center of an upside down Y.

[13] Here we consider a finite lattice in a periodic geometry. Note that if we count each link twice, then each site is counted $z$ times because there are $z$ links which connect to any given site. Thus $2N_{\rm links} = zN$, which is to say $N_{\rm links} = \frac{1}{2} zN$.

In other words, $H_{\rm eff} = H_{\rm ext} + H_{\rm int}$, where the external field is applied field $H_{\rm ext} = H$, and the 'internal field' is $H_{\rm int} = zJm$. The internal field accounts for the interaction with the *average* values of all other spins coupled to a spin at a given site, hence it is often called the 'mean field'. Since the spins are noninteracting, we have

$$m = \frac{e^{\beta H_{\rm eff}} - e^{-\beta H_{\rm eff}}}{e^{\beta H_{\rm eff}} + e^{-\beta H_{\rm eff}}} = \tanh\left(\frac{H + zJm}{k_{\rm B}T}\right) \quad . \tag{7.62}$$

It is a simple matter to solve for the free energy, given the noninteracting Hamiltonian $\hat{H}_{\rm MF}$. The partition function is

$$Z = \mathsf{Tr}\; e^{-\beta \hat{H}_{\rm MF}} = e^{-\frac{1}{2}\beta N z J\, m^2} \left( \sum_\sigma e^{\beta(H + zJm)\sigma} \right)^N = e^{-\beta F} \quad . \tag{7.63}$$

We now define dimensionless variables:

$$f \equiv \frac{F}{NzJ} \qquad , \qquad \theta \equiv \frac{k_{\rm B}T}{zJ} \qquad , \qquad h \equiv \frac{H}{zJ} \quad , \tag{7.64}$$

and obtain the dimensionless free energy

$$f(m, \theta, h) = \tfrac{1}{2}m^2 - \theta\, \log\cosh\left(\frac{m+h}{\theta}\right) - \theta \log 2 \quad . \tag{7.65}$$

Differentiating with respect to $m$ gives the mean field equation,

$$m = \tanh\left(\frac{m+h}{\theta}\right) \quad , \tag{7.66}$$

which is equivalent to the self-consistency requirement, $m = \langle \sigma_i \rangle$. In terms of the dimensionless variables $\theta$ and $h$, the physics is universal, and independent of details such as the magnitude of $J$, the value of $z$, and Boltzmann's constant.

### 7.4.2 Zero external field

When $h = 0$ the mean field equation becomes $m = \tanh(m/\theta)$. This nonlinear equation can be solved graphically, as in the top panel of fig. 7.9. The RHS is a tanh function which gets steeper with decreasing dimensionless temperature $\theta$. If, at $m = 0$, the slope of $\tanh(m/\theta)$ is smaller than unity, then the curve $y = \tanh(m/\theta)$ will intersect $y = m$ only at $m = 0$. However, if the slope is larger than unity, there will be three such intersections. Since the slope is $1/\theta$, we identify $\theta_{\rm c} = 1$ as the *mean field transition temperature*.

The mean field free energies are plotted in the bottom panel of fig. 7.9. It is possible to make analytical progress by assuming $m$ is small and Taylor expanding the free energy $f(m, \theta)$ in powers of $m$ when we are very close to the critical point, *i.e.* when $|\theta - \theta_{\rm c}| \ll 1$. Then we have[14]

$$f(m, \theta) = -\theta \log 2 + \tfrac{1}{2}m^2 - \theta\, \log\cosh\left(\frac{m}{\theta}\right) = f_0 + \tfrac{1}{2}(\theta - \theta_{\rm c})\, m^2 + \frac{m^4}{12} + \ldots \quad , \tag{7.67}$$

[14] We invoke the Taylor series expansion $\log\cosh x = \frac{1}{2}x^2 - \frac{1}{12}x^4 + \frac{1}{45}x^6 - \frac{17}{2520}x^8 + \ldots$. With $\varepsilon \equiv \theta - \theta_{\rm c}$, we drop terms which are $\mathcal{O}(\varepsilon^2 m^2, \varepsilon m^4, m^6)$. As we shall see, for $\varepsilon < 0$ we have $m^2 \propto -\varepsilon$, in which case the singular part of the free energy (*i.e.* neglecting $f_0$) is proportional to $\varepsilon^2$, while the terms we are dropping are $\mathcal{O}(\varepsilon^3)$ and smaller.

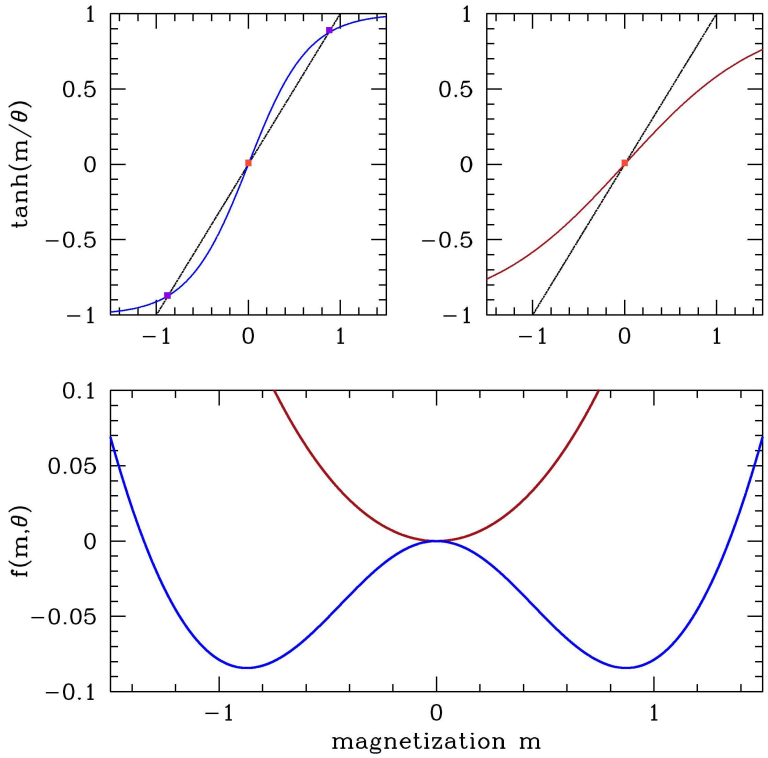



Figure 7.9: Results for $h = 0$. Upper panels: graphical solution to self-consistency equation $m = \tanh(m/\theta)$ at temperatures $\theta = 0.65$ (blue) and $\theta = 1.5$ (dark red). Lower panel: mean field free energy, with energy shifted by $\theta \log 2$ so that $f(m = 0, \theta) = 0$.

with $f_0 = -\theta \log 2$. Note that the sign of the quadratic term is positive for $\theta > \theta_{\rm c}$ and negative for $\theta < \theta_{\rm c}$. Thus, the shape of the free energy $f(m, \theta)$ as a function of $m$ qualitatively changes at $\theta = \theta_{\rm c} = 1$, the mean field transition temperature, also known as the (dimensionless) *critical temperature*. Within our mean field theory, the predicted critical temperature is $T_{\rm c} = zJ\theta_{\rm c}/k_{\rm B} = zJ/k_{\rm B}$.

In the high temperature phase, $\theta > \theta_{\rm c}$, and there is a unique minimum to $f(\theta, m)$ lying at $m = 0$. This is the disordered phase, where the *order parameter* $m$ vanishes. By contrast, in the low temperature phase $\theta < \theta_{\rm c}$, there are three solutions to the mean field equations. One solution is always at $m = 0$. The other two solutions must be related by the $\mathbb{Z}_2$ symmetry of the free energy ($m \to -m$ with $h = 0$). For $\theta > \theta_{\rm c}$, the free energy $f(m, \theta)$ has a single minimum at $m = 0$. Below $\theta_{\rm c}$, the curvature at $m = 0$ reverses, and $m = 0$ becomes a local maximum. There are then two equivalent minima symmetrically displaced on either side of $m = 0$. Differentiating with respect to $m$, we find these additional local minima to lie at $m^2 = 3(\theta_{\rm c} - \theta) + \mathcal{O}\big((\Delta\theta)^2\big)$. Thus, we find for $|\theta - \theta_{\rm c}| \ll 1$,

$$m(\theta, h = 0) = \pm\sqrt{3}\,\big(\theta_{\rm c} - \theta\big)_+^{1/2} \quad , \tag{7.68}$$

where the $+$ subscript indicates that this solution is only for $\theta_{\rm c} - \theta > 0$. As the blue curve in fig. 7.9 shows, these nonzero solutions for $m$ in the low temperature phase lie at a lower value of the free energy $f$ than $m = 0$.

Again, for $\theta > \theta_{\rm c}$ the only solution is $m = 0$. The high temperature phase is thus one where the $\mathbb{Z}_2$ (*i.e.*

$m \to -m$) symmetry is *unbroken*. In the low temperature phase, the magnetization $m$ is nonzero, and takes on one of two possible values which are degenerate in free energy. The degeneracy is guaranteed by the $\mathbb{Z}_2$ symmetry present when $h = 0$. But the system must somehow choose! This is the phenomenon of *spontaneous symmetry breaking* (SSB). The exponent with which $m(\theta)$ vanishes as $\theta \to \theta_{\rm c}^-$ is denoted as $\beta$. That is, $m(\theta, h = 0) \propto (\theta_{\rm c} - \theta)_+^\beta$ with $\beta = \frac{1}{2}$ for our mean field theory.

**Specific heat**

We can now expand the free energy $f(\theta, h = 0)$. We find

$$f(\theta, h = 0) = \begin{cases} -\theta \log 2 & \text{if } \theta > \theta_{\rm c} \\ -\theta \log 2 - \frac{3}{4}(\theta_{\rm c} - \theta)^2 + \mathcal{O}\big((\theta_{\rm c} - \theta)^4\big) & \text{if } \theta < \theta_{\rm c} \end{cases} \quad . \tag{7.69}$$

Thus, if we compute the heat capacity, we find in the vicinity of $\theta = \theta_{\rm c}$

$$c_V = -\theta \, \frac{\partial^2 f}{\partial \theta^2} = \begin{cases} 0 & \text{if } \theta > \theta_{\rm c} \\ \frac{3}{2} & \text{if } \theta < \theta_{\rm c} \end{cases} \quad . \tag{7.70}$$

Thus, the specific heat is *discontinuous* at $\theta = \theta_{\rm c}$. We emphasize that our results here are only valid for $|\theta - \theta_{\rm c}| \ll 1$. The general result valid for all $\theta$ (within our mean field theory) is[15]

$$c_V(\theta) = \frac{1}{\theta} \cdot \frac{m^2(\theta) - m^4(\theta)}{\theta - 1 + m^2(\theta)} \quad , \tag{7.71}$$

With this expression one can check both limits $\theta \to 0$ and $\theta \to \theta_{\rm c}$. As $\theta \to 0$ the magnetization saturates and one has $m^2(\theta) \simeq 1 - 4\, e^{-2/\theta}$. The numerator then vanishes as $e^{-2/\theta}$, which overwhelms the denominator that itself vanishes as $\theta^2$. As a result, $c_V(\theta \to 0) = 0$, as expected. As $\theta \to \theta_{\rm c}^-$, we have $m^2 = 3(\theta_{\rm c} - \theta) + \ldots$ and we recover $c_V(\theta_{\rm c}^-) = \frac{3}{2}$. In the theory of critical phenomena, $c_V(\theta) \propto |\theta - \theta_{\rm c}|^{-\alpha}$ as $\theta \to \theta_{\rm c}$. We see that mean field theory yields $\alpha = 0$.

### 7.4.3 Finite external field

Let us first assume $|h| \ll |\theta - 1| \ll 1$, *i.e.* that we are very close to the critical point $(\theta_{\rm c}, h_{\rm c}) = (1, 0)$. The mean field solution for $m(\theta, h)$ then be small, and we may expand the free energy from eqn. 7.65 in $m$ and $h$, *viz.*

$$\begin{aligned} f(m, \theta, h) &= -\theta \, \log 2 + \tfrac{1}{2}(1 - \theta^{-1})\, m^2 + \frac{m^4}{12\,\theta^3} - \frac{hm}{\theta} \\ &= f_0 + \tfrac{1}{2}\,(\theta - \theta_{\rm c})\, m^2 + \tfrac{1}{12} m^4 - hm + \ldots \quad . \end{aligned} \tag{7.72}$$

Note that we have only gone to linear order in $h$. Setting $\partial f / \partial m = 0$, we obtain

$$\tfrac{1}{3} m^3 + (\theta - \theta_{\rm c}) \cdot m - h = 0 \quad . \tag{7.73}$$

[15] To obtain this result, one writes $f = f\big(\theta, m(\theta)\big)$ and then differentiates twice with respect to $\theta$, using the chain rule. Along the way, any naked (*i.e.* undifferentiated) term proportional to $\partial f / \partial m$ may be dropped, since this vanishes at any $\theta$ by the mean field equation.

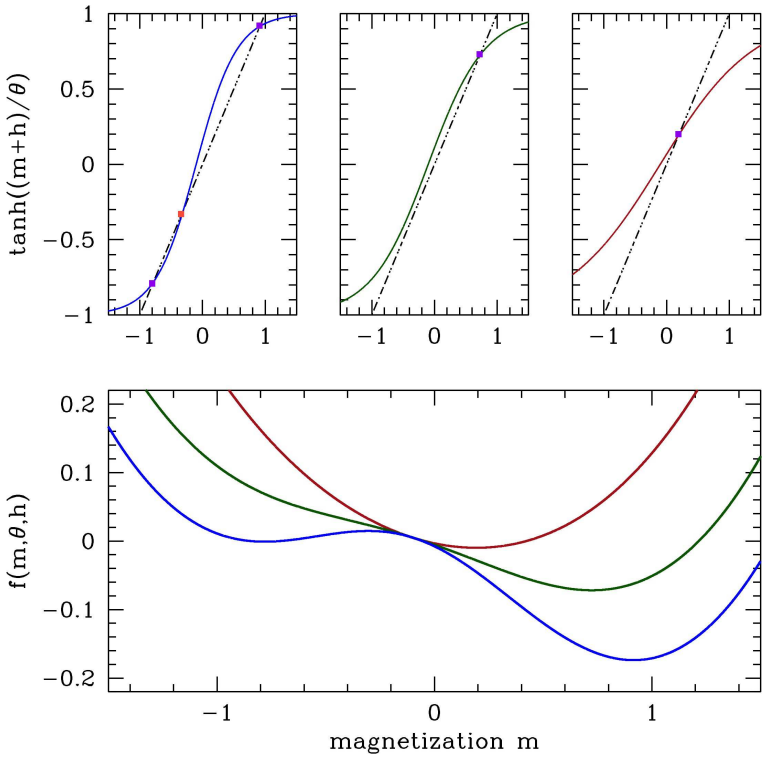



Figure 7.10: $f(m,\theta,h)$ for $h = 0.1$. Upper panels: graphical solution to the self-consistency equation $m = \tanh\big((m+h)/\theta\big)$ at temperatures $\theta = 0.65$ (blue), $\theta = 0.9$ (dark green), and $\theta = 1.5$ (dark red). Lower panel: mean field free energy, with energy shifted by $\theta \log 2$ so that $f(m=0,\theta) = 0$.

If $\theta > \theta_{\rm c}$ then we have a solution $m = h/(\theta - \theta_{\rm c})$. The $m^3$ term can be ignored because it is higher order in $h$, and we have assumed $|h| \ll |\theta - \theta_{\rm c}| \ll 1$. This is known as the *Curie-Weiss law*[16]. The zero field magnetic susceptibility is then given by

$$\chi(\theta > \theta_{\rm c}) = \left.\frac{\partial m}{\partial h}\right|_{h=0} = \frac{1}{\theta - \theta_{\rm c}} \quad , \tag{7.74}$$

where the susceptibility critical exponent $\gamma$ is $\gamma = 1$. If $\theta < \theta_{\rm c}$ then while there is still a solution at $m = h/(\theta - \theta_{\rm c})$, it lies at a local maximum of the free energy, as shown in fig. 7.10. The minimum of the

[16]Pierre Curie was a pioneer in the fields of crystallography, magnetism, and radiation physics. In 1880, Pierre and his older brother Jacques discovered piezoelectricity. He was 21 years old at the time. It was in 1895 that Pierre made the first systematic studies of the effects of temperature on magnetic materials, and he formulated what is known as *Curie's Law*, $\chi = C/T$, where $C$ is a constant. Curie married Marie Sklodowska in the same year. Their research turned toward radiation, recently discovered by Becquerel and Röntgen. In 1898, Pierre and Marie Curie discovered radium. They shared the 1903 Nobel Prize in Physics with Becquerel. Marie went on to win the 1911 Nobel Prize in Chemistry and was the first person ever awarded two Nobel Prizes. Their daughter Irène Joliot Curie shared the 1935 Prize in Chemistry (with her husband), also for work on radioactivity. Pierre Curie met an untimely and unfortunate end in the Spring of 1906. Walking across the Place Dauphine, he slipped and fell under a heavy horse-drawn wagon carrying military uniforms. His skull was crushed by one of the wagon wheels, killing him instantly. Later on that year, Pierre-Ernest Weiss proposed a modification of Curie's Law to account for ferromagnetism. This became known as the Curie-Weiss law, $\chi = C/(T - T_{\rm c})$.

| Exponent | MFT | 2D Ising (exact) | 3D Ising (numerical) | binary liquid (3D) |
|---|---|---|---|---|
| $\alpha$ | 0 | 0 | 0.11008(1) | 0.113(5) |
| $\beta$ | 1/2 | 1/8 | 0.326419(3) | 0.316(8) |
| $\gamma$ | 1 | 7/4 | 1.237075(10) | 1.240(7) |
| $\delta$ | 3 | 15 | 4.78984(1) | – |
| $\eta$ | 0 | 1/4 | 0.036298(2) | 0.016(7) |
| $\nu$ | 1/2 | 1 | 0.629971(4) | 0.625(5) |

Table 7.2: Critical exponents from mean field theory as compared with exact results for the two-dimensional Ising model, numerical results for the three-dimensional Ising model, and experiments on the liquid-gas transition in the binary fluids triethylamine and water ($\alpha$ : D. Beysens and A. Bourgou, *Phys. Rev. A* **19**, 2407 (1979)), isobutyric acid and water ($\beta$ : S. C. Greer, *Phys. Rev. A* **14**, 1770 (1976)), and 3-methylpentane-nitroethane ($\gamma$, $\eta$, $\nu$ : R. F. Chang *et al.*, *Phys. Rev. Lett.* **37**, 1481 (1976)).

free energy occurs close to the $h = 0$ solution $m = m_0(\theta) \equiv \pm\sqrt{3}\,(\theta_{\rm c} - \theta)$, and writing $m = m_0 + \delta m$ we find $\delta m$ to linear order in $h$ as $\delta m(\theta, h) = h/2(\theta_{\rm c} - \theta)$. Thus,

$$m(\theta < \theta_{\rm c}, h) = \pm\sqrt{3}\,(\theta_{\rm c} - \theta)^{1/2} + \frac{h}{2(\theta_{\rm c} - \theta)} \quad \Rightarrow \quad \chi(\theta < \theta_{\rm c}) = \frac{1}{2(\theta_{\rm c} - \theta)} \quad . \tag{7.75}$$

Both above and below the critical temperature we find that $\chi(\theta)$ diverges as $A_\pm|\theta - \theta_{\rm c}|^{-\gamma}$ where the exponent on either side of the transition is $\gamma = 1$, but the *critical amplitudes* differ: $A_+ = 1$ and $A_- = \frac{1}{2}$.

Finally, we can set $\theta = \theta_{\rm c}$ and examine $m(h)$. We find, from eqn. 7.73,

$$m(\theta = \theta_{\rm c}, h) = (3h)^{1/3} \propto h^{1/\delta} \quad , \tag{7.76}$$

where $\delta$ is a new critical exponent. Mean field theory gives $\delta = 3$. Note that at $\theta = \theta_{\rm c} = 1$ we have $m = \tanh(m + h)$, and inverting we find

$$h(m, \theta = \theta_{\rm c}) = \tfrac{1}{2}\log\left(\frac{1+m}{1-m}\right) - m = \frac{m^3}{3} + \frac{m^5}{5} + \ldots \quad , \tag{7.77}$$

which is consistent with what we just found for $m(h, \theta = \theta_{\rm c})$.

How well does mean field theory do in describing the phase transition of the Ising model? In table 7.2 we compare our mean field results for the exponents $\alpha$, $\beta$, $\gamma$, and $\delta$ with exact values for the two-dimensional Ising model, numerical work on the three-dimensional Ising model, and experiments on the liquid-gas transition in $CO_2$. The first thing to note is that the exponents are dependent on the dimension of space, and this is something that mean field theory completely misses. In fact, it turns out that the mean field exponents are exact provided $d > d_{\rm u}$, where $d_{\rm u}$ is the *upper critical dimension* of the theory. For the Ising model, $d_{\rm u} = 4$, and above four dimensions (which is of course unphysical) the mean field exponents are in fact exact. We see that all in all the MFT results compare better with the

three dimensional exponent values than with the two-dimensional ones – this makes sense since MFT does better in higher dimensions. The reason for this is that higher dimensions means more nearest neighbors, which effectively reduces the relative importance of the fluctuations we neglected to include.

**Metastable states at $h \neq 0$**

Consider the free energy $f(m, \theta, h)$ in eqn. 7.65, now for general $\theta$ not restricted to the immediate vicinity of $\theta_{\rm c}$. When $\theta < \theta_{\rm c}$ and $h$ is sufficiently small – just how small we are about to find out – the free energy as a function of $m$ has one local maximum and two local minima, one of which is the thermodynamically stable state (*i.e.* the one for which $mh > 0$), and the other (with $mh < 0$) is metastable. Consider the case $h > 0$. As the temperature is raised, the metastable local minimum at $m < 0$ eventually vanishes, annihilating with the local maximum in a saddle-node bifurcation. To find where this happens, one sets $\partial f/\partial m = 0$ and $\partial^2 f/\partial m^2 = 0$ simultaneously. From eqn. 7.65, we have

$$\frac{\partial f}{\partial m} = m - \tanh\Big(\frac{m+h}{\theta}\Big) \quad , \quad \frac{\partial^2 f}{\partial m^2} = 1 - \theta^{-1}\mathrm{sech}^2\Big(\frac{m+h}{\theta}\Big) \quad . \tag{7.78}$$

Thus $h = \theta\,\tanh^{-1}(m) - m$, and using $\mathrm{sech}^2 x = 1 - \tanh^2 x$, we have $m^2 = 1 - \theta$ and

$$h^*(\theta) = \sqrt{1-\theta} - \frac{\theta}{2}\log\left(\frac{1+\sqrt{1-\theta}}{1-\sqrt{1-\theta}}\right). \tag{7.79}$$

The solutions lie at $h = \pm h^*(\theta)$. For $\theta < \theta_{\rm c} = 1$ and $h \in \big[-h^*(\theta)\,,\,+h^*(\theta)\big]$, there are three solutions to the mean field equation. Equivalently we could in principle invert the above expression to obtain $\theta^*(h)$. For $\theta > \theta^*(h)$, there is only a single global minimum in the free energy $f(m)$ and there is no local minimum. Note $\theta^*(h=0) = 1$. Note that we could in principle invert the above relation to obtain $\theta^*(h)$, but alas this is not analytically possible.

### 7.4.4 Magnetization dynamics

Dissipative processes drive physical systems to minimum energy states. We can crudely model the dissipative dynamics of a magnet by writing the phenomenological equation

$$\frac{1}{\Gamma}\frac{dm}{dt} = -\frac{\partial f}{\partial m} = \tanh\left(\frac{m+h}{\theta}\right) - m \quad , \tag{7.80}$$

where $\Gamma$ has the dimensions of frequency. We may define $s = \Gamma t$ to be a dimensionless rescaled time. Under these dynamics, the free energy is never increasing:

$$\frac{df}{ds} = \frac{\partial f}{\partial m}\frac{dm}{ds} = -\left(\frac{\partial f}{\partial m}\right)^2 \leq 0 \quad . \tag{7.81}$$

Clearly the *fixed point* of these dynamics, where $\dot{m} = 0$, is a solution to the mean field equation $\frac{\partial f}{\partial m} = 0$. The time dependent $m(s)$ thus evolves until it reaches the first fixed point encountered, which is to say the first local extremum of the function $f(m)$. This extremum could be a global minimum, but it could also be a local minimum or even an inflection point.

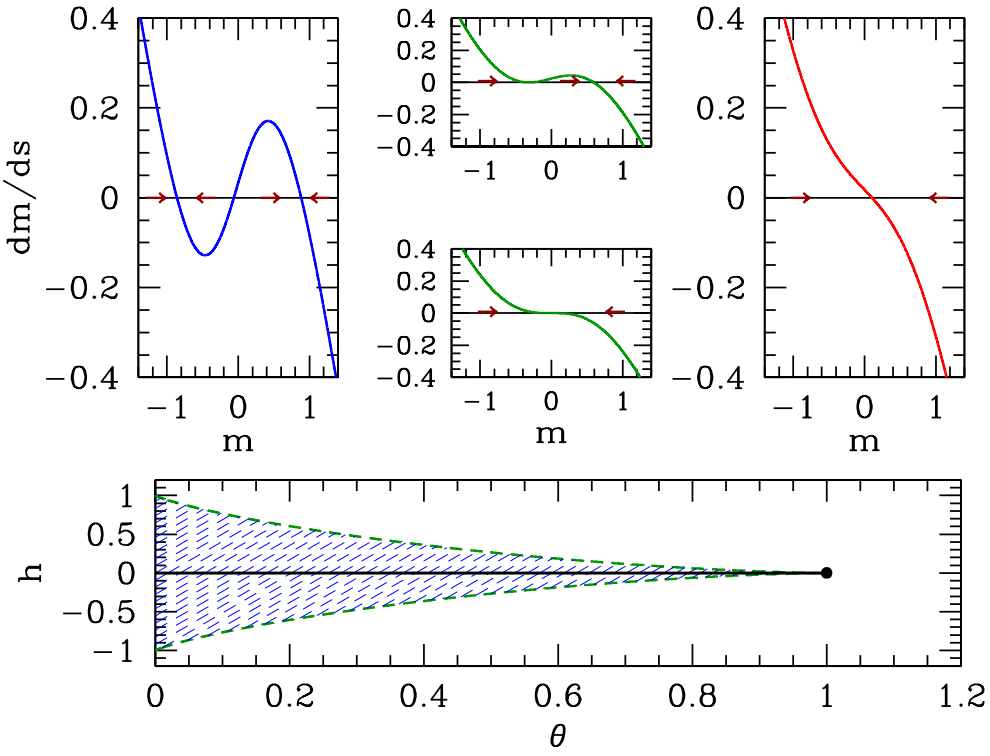



Figure 7.11: Dissipative magnetization dynamics $\dot{m} = -f'(m)$. Bottom panel shows $h^*(\theta)$ from eqn. 7.79. For $(\theta, h)$ within the blue shaded region, the free energy $f(m)$ has a global minimum plus a local minimum and a local maximum. Otherwise $f(m)$ has only a single global minimum. Top panels show an imperfect bifurcation in the magnetization dynamics at $h = 0.0215$ , for which $\theta^* = 0.90$. Temperatures shown: $\theta = 0.65$ (blue), $\theta = \theta^*(h) = 0.90$ (green), and $\theta = 1.2$. The rightmost stable fixed point corresponds to the global minimum of the free energy. The bottom of the middle two upper panels shows $h = 0$, where both of the attractive fixed points and the repulsive fixed point coalesce into a single attractive fixed point (supercritical pitchfork bifurcation).

A first order (generally nonlinear) ordinary differential equation of the form $\dot{u} = g(u)$ is called an $N = 1$ component *dynamical system*. For a general $N$-component dynamical system, $\boldsymbol{u} = (u_1, \ldots, u_N)$ is an $N$-component vector and $\boldsymbol{V}(\boldsymbol{u})$ is an $N$-component *vector field*, *i.e.* a set of $N$ functions $V_j(u_1, \ldots, u_N)$ which specify the velocity vector at each point in $\boldsymbol{u}$ space, which is $\mathbb{R}^N$ (or some other $N$-dimensional manifold, on which we might have to define an *atlas* of coordinate patches). We then have $\dot{\boldsymbol{u}} = \boldsymbol{V}(\boldsymbol{u})$, which is a set of $N$ coupled first-order ODEs. The $N = 1$ case is particularly simple to understand: $u(t)$ flows to the right in regions where $g(u) > 0$ and to the left where $g(u) < 0$. The case $g(u^*) = 0$ corresponds to a *fixed point*. Linearizing about $u^*$, *i.e.* writing $u(t) = u^* + \eta(t)$, the linearized dynamics are $\dot{\eta} = g'(u^*)\eta$, with $\eta(t) = \eta(0)\exp\big[g'(u^*)t\big]$. Thus for $g'(u^*) > 0$, corresponding to an *unstable fixed point* (UFP), the linearized dynamics take us further and further from $u^*$, and the linearization soon fails. When $g'(u^*) < 0$, corresponding to a *stable fixed point* (SFP), $u(t)$ approaches $u^*$ and the linearization gets better and better. Generically the approach to a SFP is exponential in time, *i.e.* $t = \log\big[\eta(0)/\eta(t)\big]/|g'(u^*)|$. Given the initial condition $u(0)$, the fate of an $N = 1$ dynamical system is to flow to the closest accessible

stable fixed point[17].

Our phenomenological magnetization dynamics is described by $\dot{m} = -f'(m)$, and is thus an $N = 1$ dynamical system with $g(m) = -f'(m)$. Fixed points satisfy $f'(m^*) = 0$. Thus, local minima of the free energy as a function of $m$, where $f''(m^*) > 0$, correspond to stable fixed points, while local maxima, where $f''(m^*) < 0$, correspond to unstable fixed points. This phase flow is sketched in fig. 7.11. As we have seen, for any value of $h$ there is a temperature $\theta^*$ below which the free energy $f(m)$ has two local minima and one local maximum. When $h = 0$ the minima are degenerate, but at finite $h$ one of the minima is a global minimum. Thus, for $\theta < \theta^*(h)$ there are three solutions to the mean field equations. In the language of dynamical systems, under the dynamics of eqn. 7.80, minima of $f(m)$ correspond to attractive fixed points and maxima to repulsive fixed points. If $h > 0$, the rightmost of these fixed points corresponds to the global minimum of the free energy. As $\theta$ is increased, this fixed point evolves smoothly. At $\theta = \theta^*$, the (metastable) local minimum and the local maximum coalesce and annihilate in a saddle-node bifurcation. However at $h = 0$ all three fixed points coalesce at $\theta = \theta_{\rm c}$ and the bifurcation is a supercritical pitchfork. As a function of $\theta$ at finite $h$, the dynamics are said to exhibit an *imperfect bifurcation*, which is a deformed supercritical pitchfork.

The solution set for the mean field equation is simply expressed by inverting the tanh function to obtain $h(\theta, m)$. One readily finds

$$h(\theta, m) = \frac{\theta}{2} \log\left(\frac{1+m}{1-m}\right) - m \quad . \tag{7.82}$$

As we see in the bottom panel of fig. 7.12, $m(h)$ becomes multivalued for $h \in \big[ -h^*(\theta)\,,\, +h^*(\theta) \big]$, where $h^*(\theta)$ is given in eqn. 7.79. Now imagine that $\theta < \theta_{\rm c}$ and we slowly ramp the field $h$ from a large negative value to a large positive value, and then slowly back down to its original value. On the time scale of the magnetization dynamics, we can regard $h(s)$ as a constant. (Remember the time variable is $s$ here.) Thus, $m(s)$ will flow to the nearest stable fixed point. Initially the system starts with $m = -1$ and $h$ large and negative, and there is only one fixed point, at $m^* \approx -1$. As $h$ slowly increases, the fixed point value $m^*$ also slowly increases. As $h$ exceeds $-h^*(\theta)$, a saddle-node bifurcation occurs, and two new fixed points are created at positive $m$, one stable and one unstable. The global minimum of the free energy still lies at the fixed point with $m^* < 0$. However, when $h$ crosses $h = 0$, the global minimum of the free energy lies at the most positive fixed point $m^*$. The dynamics, however, keep the system stuck in what is a metastable phase. This persists until $h = +h^*(\theta)$, at which point another saddle-note bifurcation occurs, and the attractive fixed point at $m^* < 0$ annihilates with the repulsive fixed point. The dynamics then act quickly to drive $m$ to the only remaining fixed point. This process is depicted in the top panel of fig. 7.12. As one can see from the figure, the system follows a stable fixed point until the fixed point disappears, even though that fixed point may not always correspond to a global minimum of the free energy. The resulting $m(h)$ curve is then not reversible as a function of time, and it possesses a characteristic shape known as a *hysteresis loop*. Etymologically, the word *hysteresis* derives from the Greek $\upsilon\sigma\tau\varepsilon\rho\eta\sigma\iota\varsigma$, which means 'lagging behind' (and not from $\iota\sigma\tau o\rho\iota\alpha$, which means 'inquiry'). Systems which are hysteretic exhibit a *history-dependence* to their status, which is not uniquely determined by external conditions. Hysteresis may be exhibited with respect to changes in applied magnetic field, changes in temperature,

[17]For the nongeneric case where $g(u^*) = 0$ and $g'(u^*) = 0$ vanish concomitantly, $u^*$ is a *half-stable* fixed point. For example, if $g(u) = Au^2$ then $u^* = 0$ is half -stable, and for initial conditions $u(0) < 0$ the velocity $\dot{u}(t)$ is positive and the flow asymptotically approaches $u(\infty) = 0$, while for $u(0) > 0$ the velocity is also positive and the flow is to infinity, which is reached at a finite time $t_\infty = 1/Au(0)$.

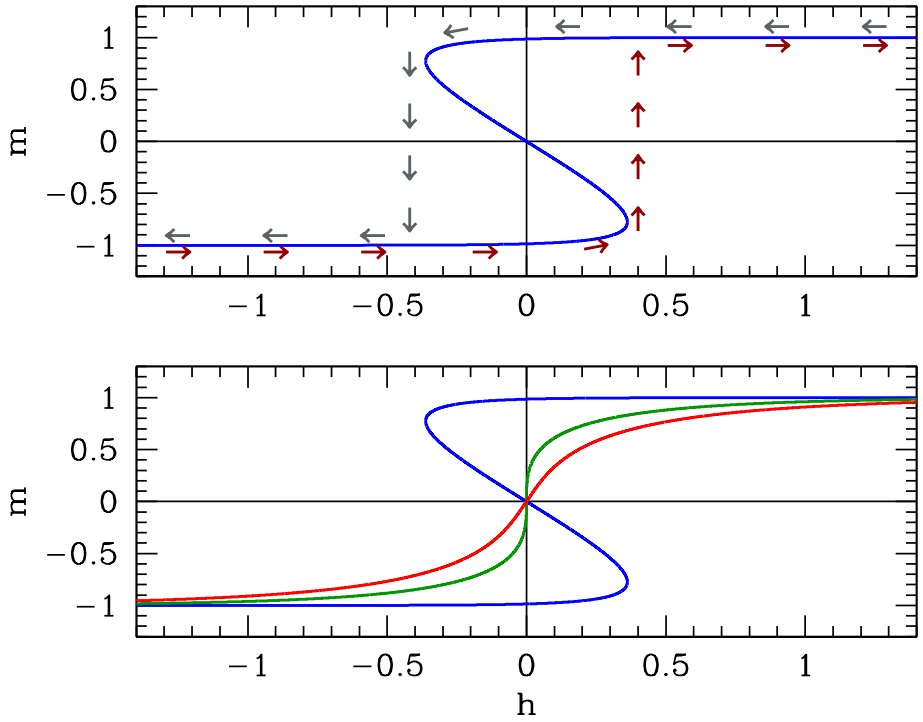



Figure 7.12: Top panel: hysteresis as a function of ramping the dimensionless magnetic field $h$ at $\theta = 0.40$. Dark red arrows below the curve track the evolution of $m$ upon a slow increase of $h$. Dark grey arrows above the curve track a slow decrease of $h$. Bottom panel : solution set for $m(\theta, h)$ as a function of $h$ at temperatures $\theta = 0.40$ (blue), $\theta = \theta_c = 1.0$ (dark green), and $\theta = 1.25$ (red).

or changes in other externally determined parameters.

### 7.4.5 Beyond nearest neighbors

Up to this point we have assumed nearest-neighbor interactions on a lattice of coordination number $z$. Suppose, instead, that we had started with the more general model,

$$\hat{H} = -\tfrac{1}{2}\sum_{i\neq j} J_{ij}\,\sigma_i\,\sigma_j - H\sum_i \sigma_i \quad , \tag{7.83}$$

where $J_{ij} = J\big(|\boldsymbol{R}_i - \boldsymbol{R}_j|\big)$ is the coupling between spins on sites $i$ and $j$[18]. In the top equation above, each pair $(ij)$ is counted once in the interaction term; this may be replaced by a sum over all $i$ and $j$ if

[18]Note that we may write $\frac{1}{2}\sum_{i\neq j} A_{ij}$ as $\sum_{i<j} A_{ij}$ provided $A = A^{\intercal}$ is symmetric. Clearly $A_{ij} = \sigma_i\,\sigma_j$ is symmetric under interchange of indices $i$ and $j$. Each version of the sum counts all unique *pairs* $(i,j) = (j,i)$ exactly once.

we include a factor of $\frac{1}{2}$.[19] The resulting mean field Hamiltonian is then

$$\hat{H}_{\rm MF} = \tfrac{1}{2} N \hat{J}(0)\, m^2 - \big(H + \hat{J}(0)\, m\big) \sum_{i=1}^{N} \sigma_i \quad . \tag{7.84}$$

Here, $\hat{J}(\boldsymbol{q})$ is the lattice Fourier transform of the interaction function $J(\boldsymbol{R})$:[20]

$$\hat{J}(\boldsymbol{q}) = \sum_{\boldsymbol{R}} J(\boldsymbol{R})\, e^{-i\boldsymbol{q}\cdot\boldsymbol{R}} \quad . \tag{7.85}$$

For nearest neighbor interactions only, one has $\hat{J}(0) = zJ$, where $z$ is the *lattice coordination number*, *i.e.* the number of nearest neighbors of any given site. The scaled free energy is as in eqn. 7.65, with $f = F/N\hat{J}(0)$, $\theta = k_{\rm B}T/\hat{J}(0)$, and $h = H/\hat{J}(0)$. The analysis proceeds precisely as before, and we conclude $\theta_{\rm c} = 1$, *i.e.* $k_{\rm B}T_{\rm c}^{\rm MF} = \hat{J}(0)$. For example, on the simple cubic lattice there are six nearest neighbors and twelve next-nearest neighbors. Thus, if $J$ is the nearest neighbor coupling and $J'$ the next-nearest neighbor coupling, we have $\hat{J}(0) = 6J + 12J'$.

### 7.4.6 Ising model with long-ranged forces

Consider an Ising model where $J_{ij} = J/N$ for all $i$ and $j$, so that there is a very weak interaction between every pair of spins. The Hamiltonian is then

$$\hat{H} = -\frac{J}{2N}\left(\sum_{i=1}^{N}\sigma_i\right)^2 - H\sum_{i=1}^{N}\sigma_i \quad . \tag{7.86}$$

The partition function is

$$Z = \operatorname*{Tr}_{\{\sigma_i\}} \exp\left[\frac{\beta J}{2N}\left(\sum_{i=1}^{N}\sigma_i\right)^2 + \beta H\sum_{i=1}^{N}\sigma_i\right] \quad . \tag{7.87}$$

We now invoke the Gaussian integral,

$$\int_{-\infty}^{\infty} dx\, e^{-\alpha x^2 - \beta x} = \sqrt{\frac{\pi}{\alpha}}\, e^{\beta^2/4\alpha} \quad . \tag{7.88}$$

Thus,

$$\exp\left[\frac{\beta J}{2N}\left(\sum_{i=1}^{N}\sigma_i\right)^2\right] = \left(\frac{N\beta J}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} dm\, e^{-\frac{1}{2}N\beta J m^2 + \beta J m \sum_{i=1}^{N}\sigma_i} \quad , \tag{7.89}$$

[19] The self-interaction terms with $i = j$ contribute a constant to $\hat{H}$ and may be either included or excluded. However, this property only pertains to the $\sigma_i = \pm 1$ model. For higher spin versions of the Ising model, say where $S_i \in \{-1, 0, +1\}$, then $S_i^2$ is not constant and we should explicitly exclude the self-interaction terms.

[20] The sum in the discrete Fourier transform is over all 'direct Bravais lattice vectors' and the wavevector $\boldsymbol{q}$ may be restricted to the 'first Brillouin zone'. These terms are familiar from elementary solid state physics.

and we can write the partition function as

$$Z = \left(\frac{N\beta J}{2\pi}\right)^{1/2} \int\limits_{-\infty}^{\infty} dm\, e^{-\frac{1}{2}N\beta J m^2} \left(\sum_\sigma e^{\beta(H+Jm)\sigma}\right)^N = \left(\frac{N}{2\pi\theta}\right)^{1/2} \int\limits_{-\infty}^{\infty} dm\, e^{-NA(m)/\theta} \quad , \tag{7.90}$$

where $\theta = k_{\rm B}T/J$, $h = H/J$, and

$$A(m) = \tfrac{1}{2}m^2 - \theta \log\left[2\cosh\left(\frac{h+m}{\theta}\right)\right] \quad . \tag{7.91}$$

Since $N \to \infty$, we can perform the integral using the method of steepest descents. Thus, we must set

$$\left.\frac{dA}{dm}\right|_{m^*} = 0 \quad \Longrightarrow \quad m^* = \tanh\left(\frac{m^*+h}{\theta}\right) \quad . \tag{7.92}$$

Expanding about $m = m^*$, we write

$$A(m) = A(m^*) + \tfrac{1}{2}A''(m^*)\,(m-m^*)^2 + \tfrac{1}{6}\,A'''(m^*)\,(m-m^*)^3 + \ldots \quad . \tag{7.93}$$

Writing $\nu = m - m^*$ and performing the integration, we obtain

$$\begin{aligned} Z &= \left(\frac{N}{2\pi\theta}\right)^{1/2} e^{-NA(m^*)/\theta} \int\limits_{-\infty}^{\infty} d\nu\, \exp\left[-\frac{NA''(m^*)}{2\theta}\,\nu^2 - \frac{NA'''(m^*)}{6\theta}\,\nu^3 + \ldots\right] \\ &= \frac{1}{\sqrt{A''(m^*)}}\, e^{-NA(m^*)/\theta} \cdot \Big\{1 + \mathcal{O}(N^{-1})\Big\} \quad . \end{aligned} \tag{7.94}$$

The corresponding free energy per site

$$f = \frac{F}{NJ} = A(m^*) + \frac{\theta}{2N}\,\log A''(m^*) + \mathcal{O}(N^{-2}) \quad , \tag{7.95}$$

where $m^*$ is the solution to the mean field equation which minimizes $A(m)$. Mean field theory is *exact* for this model!

## 7.5 Landau Theory of Phase Transitions

Landau's theory of phase transitions is based on an expansion of the free energy of a thermodynamic system in terms of an *order parameter*, which is nonzero in an ordered phase and zero in a disordered phase. For example, the magnetization density $m$ of a ferromagnet in zero external field but at finite temperature typically vanishes for temperatures $T > T_{\rm c}$, where $T_{\rm c}$ is the *critical temperature*, also called the *Curie temperature* in a ferromagnet. A low order expansion in powers of the order parameter is appropriate sufficiently close to the phase transition, *i.e.* at temperatures such that the order parameter, if nonzero, is still small.

The order parameter $m$ might be a scalar or a vector or even a matrix. In many cases the free energy $f(m)$ is invariant under the operations of a *symmetry group* $\mathsf{G}$, such that $f(m) = f(\mathsf{g}m)$ for all $\mathsf{g} \in \mathsf{G}$, where $\mathsf{g}m$ is the action of the group element $\mathsf{g}$ on $m$. For example, we might have $\mathsf{G} = \mathbb{Z}_2$ with elements $\mathsf{g} \in \{+1, -1\}$ and $m \in \mathbb{R}$. As we saw previously, in the high temperature phase $f(m)$ has a global minimum at $m = 0$. But in the low temperature phase, $f(m)$ is minimized at $m = \pm m_0 \neq 0$, with $m = 0$ a local maximum. The free energy is symmetric under the action of $\mathbb{Z}_2$, but the system has to choose one from among the two degenerate 'vacua' $m = \pm m_0$. This is the phenomenon of *spontaneous symmetry breaking* (SSB) described in §6.2.1. The order parameter or one of its derivatives exhibits a singularity or discontinuity as a function of thermodynamic parameters, such as the temperature $T$. This singular behavior occurs along a subset of parameter space, identified as a *critical point* or *critical surface*. The imposition of a finite external field $h$ *explicitly breaks* the $\mathsf{G}$-symmetry, hence SSB can occur only at $h = 0$.

### 7.5.1 Quartic free energy with Ising symmetry

The simplest example is the quartic free energy,

$$f(m, \theta, h = 0) = f_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 \quad , \tag{7.96}$$

where $f_0 = f_0(\theta)$, $a = a(\theta)$, and $b = b(\theta)$. Here, $\theta$ is a dimensionless measure of the temperature[21]. We assume $b > 0$, which is necessary if the free energy is to be bounded from below[22].

The equation of state, which relates the intensive quantities $m$ and $\theta$ at $h = 0$, is then

$$\frac{\partial f}{\partial m} = 0 = am + bm^3 \quad , \tag{7.97}$$

has three solutions in the complex $m$ plane: (i) $m = 0$, (ii) $m = \sqrt{-a/b}$, and (iii) $m = -\sqrt{-a/b}$. The latter two solutions lie along the (physical) real axis provided $a < 0$. We assume that there exists a unique temperature $\theta_{\rm c}$ where $a(\theta_{\rm c}) = 0$. Minimizing $f$, we find

$$\begin{aligned} a < 0 \quad (\theta < \theta_{\rm c}) \quad &: \quad f(\theta) = f_0 - \frac{a^2}{4b} \\ a > 0 \quad (\theta > \theta_{\rm c}) \quad &: \quad f(\theta) = f_0 \quad . \end{aligned} \tag{7.98}$$

Thus $a(\theta)$ changes sign at $\theta = \theta_{\rm c}$, where the free energy is continuous, since $a(\theta_{\rm c}) = 0$. The specific heat, however, is discontinuous across the transition, with

$$c\big(\theta_{\rm c}^+\big) - c\big(\theta_{\rm c}^-\big) = -\theta_{\rm c}\,\frac{\partial^2}{\partial\theta^2}\bigg|_{\theta=\theta_{\rm c}}\left(\frac{a^2}{4b}\right) = -\frac{\theta_{\rm c}\,\big[a'(\theta_{\rm c})\big]^2}{2b(\theta_{\rm c})} \quad . \tag{7.99}$$

The presence of a magnetic field $h$ breaks the $\mathbb{Z}_2$ symmetry of $m \to -m$. The free energy becomes

$$f(m, \theta, h) = f_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 - hm \quad , \tag{7.100}$$

[21] For example in an Ising ferromagnet we might define $\theta = k_{\rm B}T/\hat{J}(0)$, as before.

[22] It is always the case that $f$ is bounded from below, on physical grounds. Were $b$ negative, we'd have to consider higher order terms in the Landau expansion.

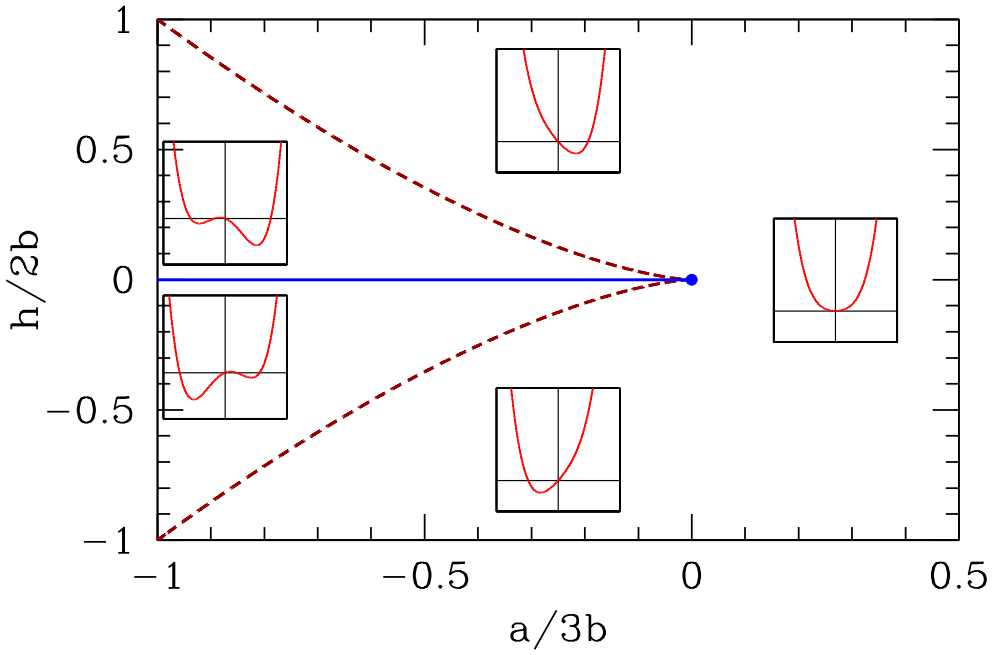



Figure 7.13: Phase diagram for the quartic Landau free energy $f = f_0 + \frac{1}{2}am^2 + \frac{1}{4}bm^4 - hm$, with $b > 0$. There is a first order line at $h = 0$ extending from $a = -\infty$ and terminating in a critical point at $a = 0$. For $|h| < h^*(a)$ (dashed red line) there are three solutions to the mean field equation, corresponding to one global minimum, one local minimum, and one local maximum. Insets show behavior of the free energy $f(m)$.

and the mean field equation is

$$bm^3 + am - h = 0 \quad . \tag{7.101}$$

This is a cubic equation for $m$ with real coefficients, and as such it can either have three real solutions or one real solution and two complex solutions related by complex conjugation. Clearly we must have $a < 0$ in order to have three real roots, since $bm^3 + am$ is monotonically increasing otherwise. The boundary between these two classes of solution sets occurs when two roots coincide, which means $f''(m) = 0$ as well as $f'(m) = 0$. Simultaneously solving these two equations, we find

$$h^*(a) = \pm\frac{2}{3^{3/2}}\,\frac{(-a)^{3/2}}{b^{1/2}} \quad , \tag{7.102}$$

or, equivalently,

$$a^*(h) = -\frac{3}{2^{2/3}}\,b^{1/3}\,|h|^{2/3}. \tag{7.103}$$

If, for fixed $h$, we have $a < a^*(h)$, then there will be three real solutions to the mean field equation $f'(m) = 0$, one of which is a global minimum (the one for which $m \cdot h > 0$). For $a > a^*(h)$ there is only a single global minimum, at which $m$ also has the same sign as $h$. If we solve the mean field equation

perturbatively in $h/a$, we find

$$\begin{aligned} m(a,h) &= \frac{h}{a} - \frac{b}{a^4}\,h^3 + \mathcal{O}(h^5) \qquad\qquad (a>0) \\ &= \pm\frac{|a|^{1/2}}{b^{1/2}} + \frac{h}{2\,|a|} \mp \frac{3\,b^{1/2}}{8\,|a|^{5/2}}\,h^2 + \mathcal{O}(h^3) \qquad (a<0) \quad . \end{aligned} \tag{7.104}$$

### 7.5.2 Cubic terms in Landau theory : first order transitions

Next, consider a free energy with a cubic term,

$$f = f_0 + \tfrac{1}{2}am^2 - \tfrac{1}{3}ym^3 + \tfrac{1}{4}bm^4 \quad , \tag{7.105}$$

with $b>0$ for stability. Without loss of generality, we may assume $y>0$ (else send $m\to -m$). Note that we no longer have $m\to -m$ (*i.e.* $\mathbb{Z}_2$) symmetry. The cubic term favors positive $m$. What is the phase diagram in the $(a,y)$ plane?

Extremizing the free energy with respect to $m$, we obtain

$$\frac{\partial f}{\partial m} = 0 = am - ym^2 + bm^3 \quad . \tag{7.106}$$

This cubic equation factorizes into a linear and quadratic piece, and hence may be solved simply. The three solutions are $m=0$ and

$$m = m_\pm \equiv \frac{y}{2b} \pm \sqrt{\Big(\frac{y}{2b}\Big)^2 - \frac{a}{b}} \quad . \tag{7.107}$$

We now see that for $y^2 < 4ab$ there is only one real solution, at $m=0$, while for $y^2>4ab$ there are three real solutions. Which solution has lowest free energy? To find out, we compare the energy $f(0)$ with $f(m_+)$[23]. Thus, we set

$$f(m) = f(0) \quad\Longrightarrow\quad \tfrac{1}{2}am^2 - \tfrac{1}{3}ym^3 + \tfrac{1}{4}bm^4 = 0 \quad , \tag{7.108}$$

and we now have two quadratic equations to solve simultaneously:

$$0 = a - ym + bm^2 \qquad , \qquad 0 = \tfrac{1}{2}a - \tfrac{1}{3}ym + \tfrac{1}{4}bm^2 = 0 \quad . \tag{7.109}$$

Eliminating the quadratic term gives $m = 3a/y$. Finally, substituting $m=m_+$ gives us a relation between $a$, $b$, and $y$:

$$y^2 = \tfrac{9}{2}\,ab \quad . \tag{7.110}$$

Thus, we have the following:

$$\begin{aligned} a > \frac{y^2}{4b} \quad &: \quad \text{1 real root } m=0 \\ \frac{y^2}{4b} > a > \frac{2y^2}{9b} \quad &: \quad \text{3 real roots; minimum at } m=0 \\ \frac{2y^2}{9b} > a \quad &: \quad \text{3 real roots; minimum at } m = \frac{y}{2b} + \sqrt{\Big(\frac{y}{2b}\Big)^2 - \frac{a}{b}} \end{aligned} \tag{7.111}$$

[23] We needn't waste our time considering the $m=m_-$ solution, since the cubic term prefers positive $m$.

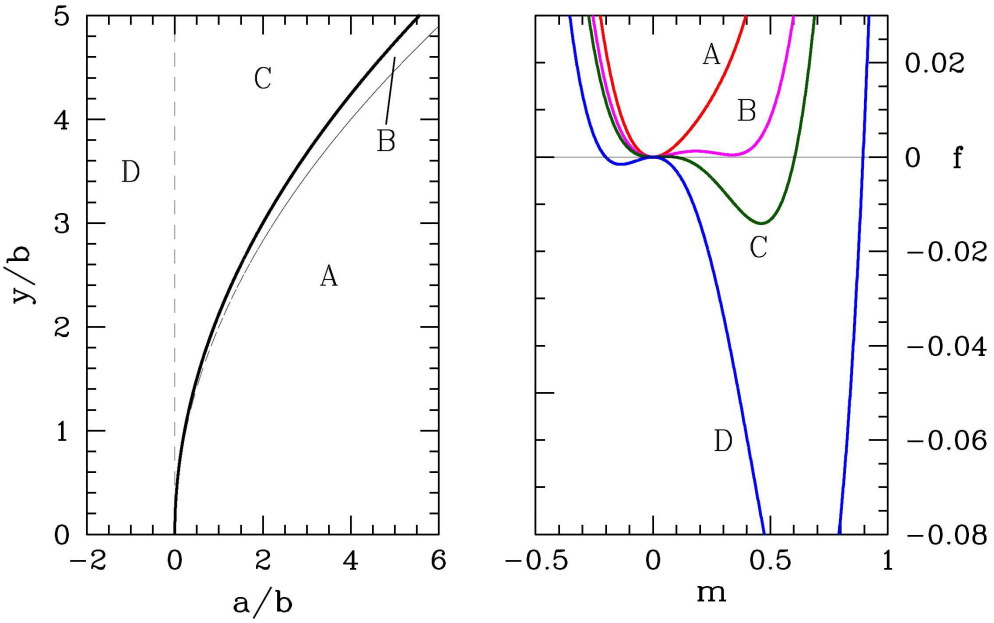



Figure 7.14: Behavior of the quartic free energy $f(m) = \frac{1}{2}am^2 - \frac{1}{3}ym^3 + \frac{1}{4}bm^4$. A: $y^2 < 4ab$ ; B: $4ab < y^2 < \frac{9}{2}ab$ ; C and D: $y^2 > \frac{9}{2}ab$. The thick black line denotes a line of first order transitions, where the order parameter is discontinuous across the transition.

The solution $m = 0$ lies at a local minimum of the free energy for $a > 0$ and at a local maximum for $a < 0$. Over the range $y^2/4b > a > 2y^2/9b$, then, there is a global minimum at $m = 0$, a local minimum at $m = m_+$, and a local maximum at $m = m_-$, with $m_+ > m_- > 0$. For $2y^2/9b > a > 0$, there is a local minimum at $m = 0$, a global minimum at $m = m_+$, and a local maximum at $m = m_-$, again with $m_+ > m_- > 0$. For $a < 0$, there is a local maximum at $m = 0$, a local minimum at $m = m_-$, and a global minimum at $m = m_+$, with $m_+ > 0 > m_-$. See fig. 7.14.

With $y = 0$, we have a second order transition at $a = 0$. With $y \neq 0$, there is a discontinuous (first order) transition at $a_{\rm c} = 2y^2/9b > 0$ and $m_{\rm c} = 2y/3b$. This occurs before $a$ reaches the value $a = 0$ where the curvature at $m = 0$ turns negative. If we write $a = \alpha(T - T_0)$, then the expected second order transition at $T = T_0$ is *preempted* by a first order transition at $T_{\rm c} = T_0 + 2y^2/9\alpha b$.

### 7.5.3 Order parameter dynamics

Suppose we now impose some dynamics on the system, of the simple relaxational type

$$\frac{dm}{dt} = -\Gamma \frac{\partial f}{\partial m} \quad , \tag{7.112}$$

where $\Gamma$ is a phenomenological kinetic coefficient. Assuming $y > 0$ and $b > 0$, it is convenient to adimensionalize by writing

$$m \equiv \frac{y}{b} \cdot u \qquad , \qquad a \equiv \frac{y^2}{b} \cdot r \qquad , \qquad t \equiv \frac{b}{\Gamma y^2} \cdot s \quad . \tag{7.113}$$

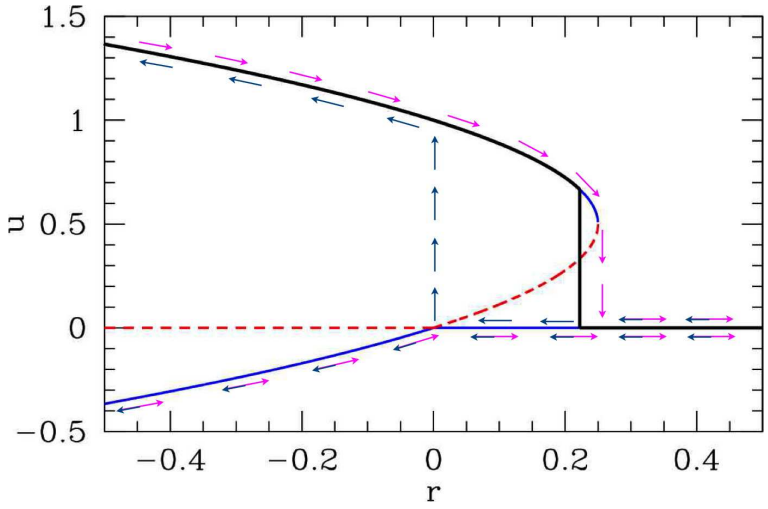



Figure 7.15: Fixed points for $\varphi(u) = \frac{1}{2}ru^2 - \frac{1}{3}u^3 + \frac{1}{4}u^4$ and flow under the dynamics $\dot{u} = -\varphi'(u)$. Solid curves represent stable fixed points and dashed curves unstable fixed points. Magenta arrows show behavior under slowly increasing control parameter $r$ and dark blue arrows show behavior under slowly decreasing $r$. For $u > 0$ there is a hysteresis loop. The thick black curve shows the equilibrium thermodynamic value of $u(r)$, *i.e.* that value which minimizes the free energy $\varphi(u)$. There is a first order phase transition at $r = \frac{2}{9}$, where the thermodynamic value of $u$ jumps from $u = 0$ to $u = \frac{2}{3}$.

Then we obtain

$$\frac{du}{ds} = -\frac{\partial\varphi}{\partial u} \quad , \tag{7.114}$$

where the dimensionless free energy function is

$$\varphi(u) = \tfrac{1}{2}ru^2 - \tfrac{1}{3}u^3 + \tfrac{1}{4}u^4 \quad . \tag{7.115}$$

We see that there is a single control parameter, $r = ab/y^2$. The fixed points of the dynamics are then the stationary points of $\varphi(u)$, where $\varphi'(u) = 0$, with $\varphi'(u) = u\,(r - u + u^2)$. The solutions to $\varphi'(u) = 0$ are then given by

$$u^* = 0 \qquad , \qquad u^* = \tfrac{1}{2} \pm \left(\tfrac{1}{4} - r\right)^{1/2} \quad . \tag{7.116}$$

The second of these is a parabola $r = u(1-u)$ opening to the left in the $(r, u)$ plane. For $r > \frac{1}{4}$ there is one fixed point at $u = 0$, which is attractive under the dynamics $\dot{u} = -\varphi'(u)$ since $\varphi''(0) = r$. At $r = \frac{1}{4}$ there occurs a saddle-node bifurcation and a pair of fixed points is generated, one stable and one unstable. As we see from fig. 7.13, the interior fixed point is always unstable and the two exterior fixed points are always stable. At $r = 0$ there is a transcritical bifurcation where two fixed points of opposite stability collide and bounce off one another (metaphorically speaking).

At the saddle-node bifurcation, $r = \frac{1}{4}$ and $u = \frac{1}{2}$, and we find $\varphi(u = \frac{1}{2}; r = \frac{1}{4}) = \frac{1}{192}$, which is positive. Thus, the thermodynamic state of the system remains at $u = 0$ until the value of $\varphi(u_+)$ crosses zero. This occurs when $\varphi(u) = 0$ and $\varphi'(u) = 0$, the simultaneous solution of which yields $r = \frac{2}{9}$ and $u = \frac{2}{3}$.

Suppose we slowly ramp the control parameter $r$ up and down as a function of the dimensionless time $s$. Under the dynamics of eqn. 7.114, $u(s)$ flows to the first stable fixed point encountered – this is always

the case for a dynamical system with a one-dimensional phase space. Then as $r$ is further varied, $u$ follows the position of whatever locally stable fixed point it initially encountered. Thus, $u\big(r(s)\big)$ evolves smoothly until a bifurcation is encountered. The situation is depicted by the arrows in fig. 7.15. The equilibrium thermodynamic value for $u(r)$ is discontinuous; there is a first order phase transition at $r = \frac{2}{9}$, as we've already seen. As $r$ is increased, $u(r)$ follows a trajectory indicated by the magenta arrows. For an negative initial value of $u$, the evolution as a function of $r$ will be *reversible*. However, if $u(0)$ is initially positive, then the system exhibits *hysteresis*, as shown. Starting with a large positive value of $r$, $u(s)$ quickly evolves to $u = 0^+$, which means a positive infinitesimal value. Then as $r$ is decreased, the system remains at $u = 0^+$ even through the first order transition, because $u = 0$ is an attractive fixed point. However, once $r$ begins to go negative, the $u = 0$ fixed point becomes repulsive, and $u(s)$ quickly flows to the stable fixed point $u_+ = \frac{1}{2} + \left(\frac{1}{4} - r\right)^{1/2}$. Further decreasing $r$, the system remains on this branch. If $r$ is later increased, then $u(s)$ remains on the upper branch past $r = 0$, until the $u_+$ fixed point annihilates with the unstable fixed point at $u_- = \frac{1}{2} - \left(\frac{1}{4} - r\right)^{1/2}$, at which time $u(s)$ quickly flows down to $u = 0^+$ again.

### 7.5.4 Sixth order Landau theory : tricritical point

Finally, consider a model with $\mathbb{Z}_2$ symmetry, with the Landau free energy

$$f = f_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 + \tfrac{1}{6}cm^6 \quad , \tag{7.117}$$

with $c > 0$ for stability. We seek the phase diagram in the $(a, b)$ plane. Extremizing $f$ with respect to $m$,

$$\frac{\partial f}{\partial m} = 0 = m\left(a + bm^2 + cm^4\right) \quad , \tag{7.118}$$

which is a quintic with five solutions over the complex $m$ plane. One solution is obviously $m = 0$. The other four are

$$m = \pm\sqrt{-\frac{b}{2c} \pm \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a}{c}}} \quad . \tag{7.119}$$

For each $\pm$ symbol in the above equation, there are two options, hence four roots in all.

If $a > 0$ and $b > 0$, then four of the roots are imaginary and there is a unique minimum at $m = 0$.

For $a < 0$, there are only three solutions to $f'(m) = 0$ for real $m$, regardless of the sign of $b$, since the minus choice for the $\pm$ sign under the radical leads to imaginary roots. One of the solutions is $m = 0$, which is a local maximum of $f(m)$. The other two are degenerate minima at

$$m = \pm\sqrt{-\frac{b}{2c} + \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a}{c}}} \quad . \tag{7.120}$$

The most interesting situation is $a > 0$ and $b < 0$. If $a > 0$ and $b < -2\sqrt{ac}$, all five roots are real. There must be three minima, separated by two local maxima. Clearly if $m^*$ is a solution, then so is $-m^*$. Thus, the only question is whether the outer minima are of lower energy than the minimum at $m = 0$. We

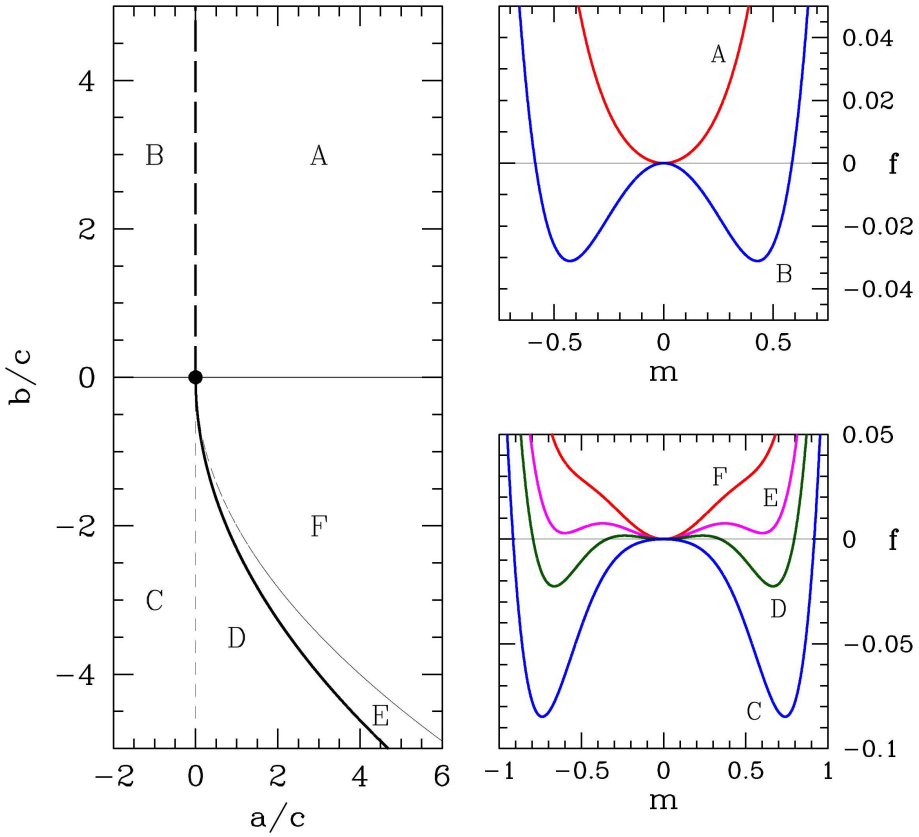



Figure 7.16: Behavior of the sextic free energy $f(m) = \frac{1}{2}am^2 + \frac{1}{4}bm^4 + \frac{1}{6}cm^6$. A: $a > 0$ and $b > 0$ ; B: $a < 0$ and $b > 0$ ; C: $a < 0$ and $b < 0$ ; D: $a > 0$ and $b < -\frac{4}{\sqrt{3}}\sqrt{ac}$ ; E: $a > 0$ and $-\frac{4}{\sqrt{3}}\sqrt{ac} < b < -2\sqrt{ac}$ ; F: $a > 0$ and $-2\sqrt{ac} < b < 0$. The thick dashed line is a line of second order transitions, which meets the thick solid line of first order transitions at the tricritical point, $(a, b) = (0, 0)$.

assess this by demanding $f(m^*) = f(0)$, where $m^*$ is the position of the largest root (*i.e.* the rightmost minimum). This gives a second quadratic equation,

$$0 = \tfrac{1}{2}a + \tfrac{1}{4}bm^2 + \tfrac{1}{6}cm^4 \quad , \tag{7.121}$$

which together with equation 7.118 gives $b = -\frac{4}{\sqrt{3}}\sqrt{ac}$. Thus, we have the following, for fixed $a > 0$:

$$\begin{aligned} b > -2\sqrt{ac} \quad &: \quad \text{1 real root } m = 0 \\ -2\sqrt{ac} > b > -\tfrac{4}{\sqrt{3}}\sqrt{ac} \quad &: \quad \text{5 real roots; minimum at } m = 0 \\ -\tfrac{4}{\sqrt{3}}\sqrt{ac} > b \quad &: \quad \text{5 real roots; minima at } m = \pm\sqrt{-\frac{b}{2c} + \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a}{c}}} \end{aligned} \tag{7.122}$$

The point $(a, b) = (0, 0)$, which lies at the confluence of a first order line and a second order line, is known as a *tricritical point*.

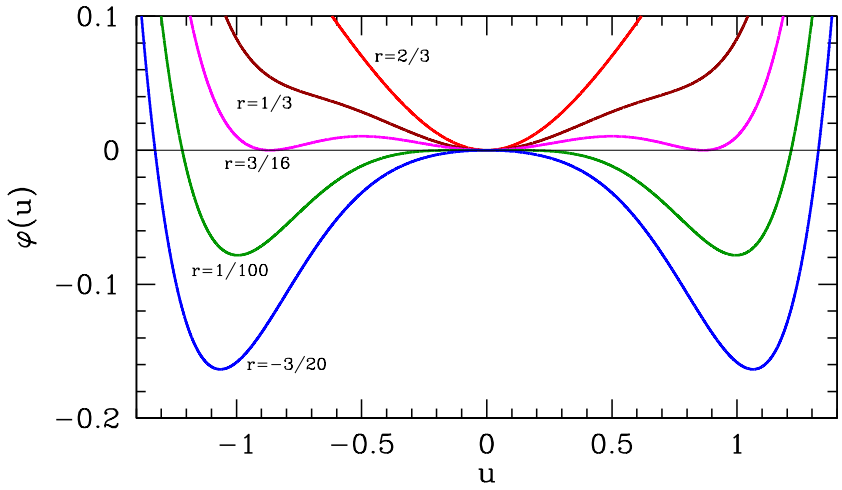



Figure 7.17: Free energy $\varphi(u) = \frac{1}{2}ru^2 - \frac{1}{4}u^4 + \frac{1}{6}u^6$ for different values of the control parameter $r$.

### 7.5.5 Hysteresis for the sextic potential

Once again, we consider the dissipative dynamics $\dot{m} = -\Gamma f'(m)$. We adimensionalize by writing

$$m \equiv \sqrt{\frac{|b|}{c}} \cdot u \quad , \quad a \equiv \frac{b^2}{c} \cdot r \quad , \quad t \equiv \frac{c}{\Gamma b^2} \cdot s \quad . \tag{7.123}$$

Then we obtain once again the dimensionless equation $du/ds = -\partial\varphi/\partial u$, where

$$\varphi(u) = \tfrac{1}{2}ru^2 \pm \tfrac{1}{4}u^4 + \tfrac{1}{6}u^6 \quad . \tag{7.124}$$

Once again, as in the case of the cubic Landau theory examined in §7.5.2, there is a single control parameter, which here is given by $r = ac/b^2$, but there are two distinct sub-cases to consider. In eqn. 7.124, the coefficient of the quartic term is positive if $b > 0$ and negative if $b < 0$. That is, the coefficient is $\mathrm{sgn}(b)$. When $b > 0$ we can ignore the sextic term for sufficiently small $u$, and we recover the quartic free energy studied earlier. There is then a second order transition at $r = 0$. The free energy curves for various values of $r$ are plotted in fig. 7.17.

New and interesting behavior occurs for $b > 0$. The fixed points of the dynamics are obtained by setting $\varphi'(u) = 0$. We have

$$\varphi(u) = \tfrac{1}{2}ru^2 - \tfrac{1}{4}u^4 + \tfrac{1}{6}u^6 \quad , \quad \varphi'(u) = u\left(r - u^2 + u^4\right) \quad . \tag{7.125}$$

Thus, the equation $\varphi'(u) = 0$ factorizes into a linear factor $u$ and a quartic factor $u^4 - u^2 + r$ which is quadratic in $u^2$. Thus, we can easily obtain the roots:

$$\begin{aligned} r < 0 \quad &: \quad u^* = 0 \quad , \; u^* = \pm\sqrt{\tfrac{1}{2} + \sqrt{\tfrac{1}{4} - r}} \\ 0 < r < \tfrac{1}{4} \quad &: \quad u^* = 0 \quad , \; u^* = \pm\sqrt{\tfrac{1}{2} + \sqrt{\tfrac{1}{4} - r}} \quad , \; u^* = \pm\sqrt{\tfrac{1}{2} - \sqrt{\tfrac{1}{4} - r}} \\ r > \tfrac{1}{4} \quad &: \quad u^* = 0 \quad . \end{aligned} \tag{7.126}$$

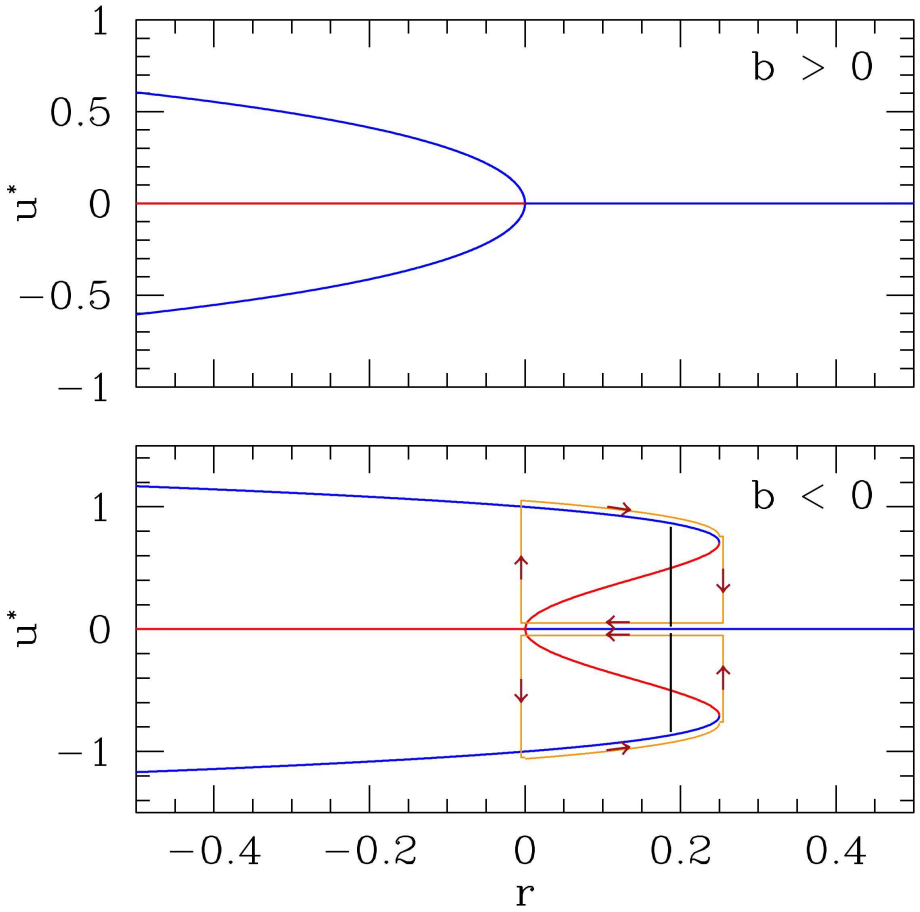



Figure 7.18: Fixed points branches $\varphi'(u^*) = 0$ for the sextic potential $\varphi(u) = \frac{1}{2}ru^2 + \frac{\sigma}{4}u^4 + \frac{1}{6}u^6$ with $\sigma = \mathrm{sgn}(b)$ (top for $\sigma = +1$, bottom for $\sigma = -1$). Blue curves indicate stable branches and red unstable branches under the dynamics $\dot{u} = -\varphi'(u)$. For $b > 0$ there is a supercritical pitchfork bifurcation at $(0, 0)$ and no hysteresis. For $b > 0$ there is a subcritical pitchfork bifurcation at $(0, 0)$ and two saddle-node bifurcations at $(\frac{1}{4}, \pm\frac{1}{\sqrt{2}})$. Solid curves show stable fixed points and dashed curves show unstable fixed points. The thick solid black segment indicates the order parameter discontinuity in $u$ at the thermodynamic first order phase transition. Note the overall $u \to -u$ symmetry. Within the region $r \in [0, \frac{1}{4}]$ the dynamics are irreversible and the system exhibits the phenomenon of hysteresis. There is a first order phase transition at $r = \frac{3}{16}$, indicated by the thick black line segment which shows the discontinuous jump of the order parameter $u$ at the transition.

In fig. 7.18, we plot the fixed points and the hysteresis loops for this system. At $r = \frac{1}{4}$, there are two symmetrically located saddle-node bifurcations at $u = \pm\frac{1}{\sqrt{2}}$. We find $\varphi(u = \pm\frac{1}{\sqrt{2}}, r = \frac{1}{4}) = \frac{1}{48}$, which is positive, indicating that the stable fixed point $u^* = 0$ remains the thermodynamic minimum for the free energy $\varphi(u)$ as $r$ is decreased through $r = \frac{1}{4}$. Setting $\varphi(u) = 0$ and $\varphi'(u) = 0$ simultaneously, we obtain $r = \frac{3}{16}$ and $u = \pm\frac{\sqrt{3}}{2}$. The thermodynamic value for $u$ therefore jumps discontinuously from $u = 0$ to $u = \pm\frac{\sqrt{3}}{2}$ (either branch) at $r = \frac{3}{16}$; this is a first order transition.

Under the dissipative dynamics considered here, the system exhibits hysteresis, as indicated in the figure, where the arrows show the evolution of $u(s)$ for very slowly varying $r(s)$. When the control parameter $r$ is large and positive, the flow is toward the sole fixed point at $u^* = 0$. At $r = \frac{1}{4}$, two simultaneous saddle-node bifurcations take place at $u^* = \pm\frac{1}{\sqrt{2}}$; the outer branch is stable and the inner branch unstable in both cases. At $r = 0$ there is a subcritical pitchfork bifurcation, and the fixed point at $u^* = 0$ becomes unstable.

Suppose one starts off with $r \gg \frac{1}{4}$ with some value $u > 0$. The flow $\dot{u} = -\varphi'(u)$ then rapidly results in $u \to 0^+$. This is the 'high temperature phase' in which there is no magnetization. Now let $r$ increase slowly, using $s$ as the dimensionless time variable. The scaled magnetization $u(s) = u^*\big(r(s)\big)$ will remain pinned at the fixed point $u^* = 0^+$. As $r$ passes through $r = \frac{1}{4}$, two new stable values of $u^*$ appear, but our system remains at $u = 0^+$, since $u^* = 0$ is a stable fixed point. But after the subcritical pitchfork, $u^* = 0$ becomes unstable. The magnetization $u(s)$ then flows rapidly to the stable fixed point at $u^* = \frac{1}{\sqrt{2}}$, and follows the curve $u^*(r) = \sqrt{\frac{1}{2} + (\frac{1}{4} - r)^{1/2}}$ for all $r < 0$.

Now suppose we start increasing $r$ (*i.e.* increasing temperature). The magnetization follows the stable fixed point $u^*(r) = \sqrt{\frac{1}{2} + (\frac{1}{4} - r)^{1/2}}$ past $r = 0$, beyond the first order phase transition point at $r = \frac{3}{16}$, and all the way up to $r = \frac{1}{4}$, at which point this fixed point is annihilated at a saddle-node bifurcation. The flow then rapidly takes $u \to u^* = 0^+$, where it remains as $r$ continues to be increased further. Within the region $r \in \big[0, \frac{1}{4}\big]$ of control parameter space, the dynamics are said to be *irreversible* and the behavior of $u(s)$ is said to be *hysteretic*.

## 7.6 Variational Density Matrix Method

### 7.6.1 The variational principle

Suppose we are given a Hamiltonian $\hat{H}$. From this we construct the free energy, $F$:

$$F = E - TS = \mathrm{Tr}\,(\varrho\,\hat{H}) + k_{\rm B}T\,\mathrm{Tr}\,(\varrho\,\log\varrho) \quad . \tag{7.127}$$

Here, $\varrho$ is the *density matrix*[24]. A physical density matrix must be (i) normalized (*i.e.* $\mathrm{Tr}\,\varrho = 1$), (ii) Hermitian, and (iii) non-negative definite (*i.e.* all the eigenvalues of $\varrho$ must be non-negative).

Our goal is to extremize the free energy subject to the various constraints on $\varrho$. Let us assume that $\varrho$ is diagonal in the basis of eigenstates of $\hat{H}$, *i.e.*

$$\varrho = \sum_\gamma P_\gamma\,|\,\gamma\,\rangle\langle\,\gamma\,| \quad , \tag{7.128}$$

where $P_\gamma$ is the probability that the system is in state $\big|\,\gamma\,\big\rangle$. Then

$$F = \sum_\gamma E_\gamma\,P_\gamma + k_{\rm B}T\sum_\gamma P_\gamma\log P_\gamma \quad . \tag{7.129}$$

[24]How do we take the logarithm of a matrix? The rule is this: $A = \log B$ if $B = \exp(A)$. The exponential of a matrix may be evaluated via its Taylor expansion.

Thus, the free energy is a function of the set $\{P_\gamma\}$. We now extremize $F$ subject to the normalization constraint. This means we form the extended function

$$F^*\big(\{P_\gamma\},\lambda\big) = F\big(\{P_\gamma\}\big) + \lambda\Big(\sum_\gamma P_\gamma - 1\Big) \quad , \tag{7.130}$$

and then freely extremize over both the probabilities $\{P_\gamma\}$ as well as the Lagrange multiplier $\lambda$. This yields the Boltzmann distribution,

$$P_\gamma^{\rm eq} = \frac{1}{Z}\,\exp(-E_\gamma/k_{\rm B}T) \quad , \tag{7.131}$$

where $Z = \sum_\gamma e^{-E_\gamma/k_{\rm B}T} = {\rm Tr}\, e^{-\hat{H}/k_{\rm B}T}$ is the canonical partition function, which is related to $\lambda$ through

$$\lambda = k_{\rm B}T\,(\log Z - 1) \quad . \tag{7.132}$$

Note that the Boltzmann weights are, appropriately, all positive.

If the spectrum of $\hat{H}$ is bounded from below, our extremum should in fact yield a minimum for the free energy $F$. Furthermore, since we have freely minimized over all the probabilities, subject to the single normalization constraint, *any distribution $\{P_\gamma\}$ other than the equilibrium one must yield a greater value of $F$.*

Alas, the Boltzmann distribution, while exact, is often intractable to evaluate. For one-dimensional systems, there are general methods such as the transfer matrix approach which do permit an exact evaluation of the free energy. However, beyond one dimension the situation is in general hopeless. A family of solvable ("integrable") models exists in two dimensions, but their solutions require specialized techniques and are extremely difficult. The idea behind the variational density matrix approximation is to construct a tractable *trial* density matrix $\varrho$ which depends on a set of variational parameters $\{x_\alpha\}$, and to minimize $F$ with respect to this set.

### 7.6.2 Variational density matrix for the Ising model

Consider once again the Ising model Hamiltonian,

$$\hat{H} = -\frac{1}{2}\sum_{i,j=1}^N J_{ij}\,\sigma_i\,\sigma_j - H\sum_{i=1}^N \sigma_i \quad . \tag{7.133}$$

The states of the system $|\boldsymbol{\sigma}\rangle$ may be labeled by the values of the spin variables: $|\boldsymbol{\sigma}\rangle = |\sigma_1,\sigma_2,\ldots,\sigma_N\rangle$. We assume the density matrix is diagonal in this basis, *i.e.*

$$\varrho_N(\boldsymbol{\sigma}\,|\,\boldsymbol{\sigma}') \equiv \varrho_N(\boldsymbol{\sigma})\,\delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} \quad , \tag{7.134}$$

where $\delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} = \prod_{i=1}^N \delta_{\sigma_i,\sigma_i'}$ and where $N$ is the number of sites. Indeed, this is the case for the exact density matrix, which is to say the Boltzmann weight, $\varrho_N(\boldsymbol{\sigma}) = Z_N^{-1} e^{-\beta\hat{H}(\boldsymbol{\sigma})}$.

We now write a *trial density matrix* which is a product over contributions from independent single sites:

$$\varrho_N(\boldsymbol{\sigma}) = \prod_{i=1}^N \varrho(\sigma_i) \quad , \tag{7.135}$$

where

$$\varrho(\sigma) = \left(\frac{1+m}{2}\right)\delta_{\sigma,1} + \left(\frac{1-m}{2}\right)\delta_{\sigma,-1} \quad . \tag{7.136}$$

Note that we've changed our notation slightly. We are denoting by $\varrho(\sigma)$ the corresponding diagonal element of the matrix

$$\varrho = \frac{1}{2}\begin{pmatrix} 1+m & 0 \\ 0 & 1-m \end{pmatrix} \quad , \tag{7.137}$$

and the full density matrix is a tensor product over the single site matrices: $\varrho_N = \varrho \otimes \varrho \otimes \cdots \otimes \varrho$. Note also that $\varrho$ and $\varrho_N$ are normalized with unit trace.

Note that $\varrho$ and hence $\varrho_N$ are appropriately normalized. The variational parameter here is $m$, which, if $\rho$ is to be non-negative definite, must satisfy $-1 \le m \le 1$. The quantity $m$ has the physical interpretation of the average spin on any given site, since

$$\langle \sigma_i \rangle = \sum_\sigma \varrho(\sigma)\,\sigma = m. \tag{7.138}$$

We may now evaluate the average energy:

$$\begin{aligned} E = \mathsf{Tr}\,(\varrho_N \hat{H}) &= -\sum_{i<j} J_{ij}\, m^2 - H\sum_i m \\ &= -\tfrac{1}{2} N \hat{J}(0)\, m^2 - NHm \quad , \end{aligned} \tag{7.139}$$

where once again $\hat{J}(0)$ is the discrete Fourier transform of $J(\boldsymbol{R})$ at wavevector $\boldsymbol{q} = 0$. The entropy is given by

$$\begin{aligned} S &= -k_{\rm B}\,\mathsf{Tr}\,(\varrho_N \log \varrho_N) = -Nk_{\rm B}\,\mathsf{Tr}\,(\varrho \log \varrho) \\ &= -Nk_{\rm B}\left\{\left(\frac{1+m}{2}\right)\log\left(\frac{1+m}{2}\right) + \left(\frac{1-m}{2}\right)\log\left(\frac{1-m}{2}\right)\right\} \quad . \end{aligned} \tag{7.140}$$

We now define the dimensionless free energy per site: $f \equiv F/N\hat{J}(0)$. We have

$$f(m,\theta,h) = -\tfrac{1}{2}\,m^2 - hm + \theta\left\{\left(\frac{1+m}{2}\right)\log\left(\frac{1+m}{2}\right) + \left(\frac{1-m}{2}\right)\log\left(\frac{1-m}{2}\right)\right\} \quad , \tag{7.141}$$

where $\theta \equiv k_{\rm B}T/\hat{J}(0)$ is the dimensionless temperature, and $h \equiv H/\hat{J}(0)$ the dimensionless magnetic field, as before. We extremize $f(m)$ by setting

$$\frac{\partial f}{\partial m} = 0 = -m - h + \frac{\theta}{2}\log\left(\frac{1+m}{1-m}\right) \quad . \tag{7.142}$$

Solving for $m$, we obtain

$$m = \tanh\left(\frac{m+h}{\theta}\right) \quad , \tag{7.143}$$

which is precisely what we found in eqn. 7.66. There is a second order phase transition at a critical temperature $\theta_{\rm c} = 1$.

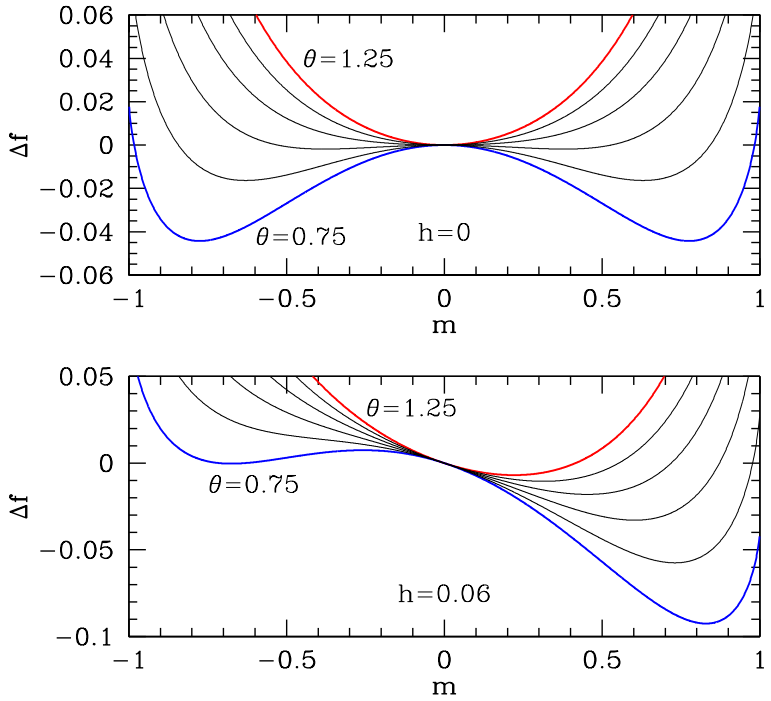



Figure 7.19: Variational field free energy $\Delta f = f(m,\theta,h) + \theta\,\log 2$ *versus* magnetization $m$ at six equally spaced temperatures interpolating between 'high' ($\theta = 1.25$, red) and 'low' ($\theta = 0.75$, blue) values. Top panel: $h = 0$. Bottom panel: $h = 0.06$.

Note that the optimal value of $m$ indeed satisfies the requirement $|m| \le 1$ of non-negative probability. This nonlinear equation may be solved graphically. For $h = 0$, the unmagnetized solution $m = 0$ always applies. However, for $\theta < 1$ there are two additional solutions at $m = \pm m_{\rm A}(\theta)$, with $m_{\rm A}(\theta) = \sqrt{3(1-\theta)} + \mathcal{O}\big((1-\theta)^{3/2}\big)$ for $t$ close to (but less than) one. These solutions, which are related by the $\mathbb{Z}_2$ symmetry of the $h = 0$ model, are in fact the low energy solutions. This is shown clearly in figure 7.19, where the variational free energy $f(m,\theta)$ is plotted as a function of $m$ for a range of temperatures interpolating between 'high' and 'low' values. At the *critical temperature* $\theta_{\rm c} = 1$, the lowest energy state changes from being unmagnetized (high temperature) to magnetized (low temperature).

For $h > 0$, there is no longer a $\mathbb{Z}_2$ symmetry (*i.e.* $\sigma_i \to -\sigma_i\ \forall\ i$). The high temperature solution now has $m > 0$ (or $m < 0$ if $h < 0$), and this smoothly varies as $t$ is lowered, approaching the completely polarized limit $m = 1$ as $\theta \to 0$. At very high temperatures, the argument of the $\tanh$ function is small, and we may approximate $\tanh(x) \simeq x$, in which case

$$m(h,\theta) = \frac{h}{\theta - \theta_{\rm c}} \quad . \tag{7.144}$$

This is called the *Curie-Weiss law.* One can infer $\theta_{\rm c}$ from the high temperature susceptibility $\chi(\theta) = (\partial m/\partial h)_{h=0}$ by plotting $\chi^{-1}$ *versus* $\theta$ and extrapolating to obtain the $\theta$-intercept. In our case, $\chi(\theta) = (\theta - \theta_{\rm c})^{-1}$. For low $\theta$ and weak $h$, there are two inequivalent minima in the free energy.

When $m$ is small, it is appropriate to expand $f(m,\theta,h)$, obtaining

$$f(m,\theta,h) = -\theta\,\log 2 - hm + \tfrac{1}{2}\,(\theta-1)\,m^2 + \tfrac{\theta}{12}\,m^4 + \tfrac{\theta}{30}\,m^6 + \tfrac{\theta}{56}\,m^8 + \ldots \quad . \tag{7.145}$$

This is known as the *Landau expansion* of the free energy in terms of the *order parameter* $m$. An order parameter is a thermodynamic variable $\phi$ which distinguishes ordered and disordered phases. Typically $\phi = 0$ in the disordered (high temperature) phase, and $\phi \neq 0$ in the ordered (low temperature) phase. When the order sets in continuously, *i.e.* when $\phi$ is continuous across $\theta_{\rm c}$, the phase transition is said to be *second order.* When $\phi$ changes abruptly, the transition is *first order.* It is also quite commonplace to observe phase transitions between two ordered states. For example, a crystal, which is an ordered state, may change its lattice structure, say from a high temperature tetragonal phase to a low temperature orthorhombic phase. When the high $T$ phase possesses the same symmetries as the low $T$ phase, as in the tetragonal-to-orthorhombic example, the transition may be second order. When the two symmetries are completely unrelated, for example in a hexagonal-to-tetragonal transition, or in a transition between a ferromagnet and an antiferromagnet, the transition is in general first order.

Throughout this discussion, we have assumed that the interactions $J_{ij}$ are predominantly *ferromagnetic*, *i.e.* $J_{ij} > 0$, so that all the spins prefer to align. When $J_{ij} < 0$, the interaction is said to be *antiferromagnetic* and prefers anti-alignment of the spins (*i.e.* $\sigma_i\,\sigma_j = -1$). Clearly not every pair of spins can be anti-aligned – there are two possible spin states and a thermodynamically extensive number of spins. But on the square lattice, for example, if the only interactions $J_{ij}$ are between nearest neighbors and the interactions are antiferromagnetic, then the lowest energy configuration ($T = 0$ ground state) will be one in which spins on opposite sublattices are anti-aligned. The square lattice is *bipartite* – it breaks up into two interpenetrating sublattices A and B (which are themselves square lattices, rotated by 45° with respect to the original, and with a larger lattice constant by a factor of $\sqrt{2}$), such that any site in A has nearest neighbors in B, and *vice versa*. The honeycomb lattice is another example of a bipartite lattice. So is the simple cubic lattice. The triangular lattice, however, is not bipartite (it is *tripartite*). Consequently, with nearest neighbor antiferromagnetic interactions, the triangular lattice Ising model is highly *frustrated*. The moral of the story is this: antiferromagnetic interactions can give rise to complicated magnetic ordering, and, when frustrated by the lattice geometry, may have finite specific entropy even at $T = 0$.

### 7.6.3 $q$-state Potts model

The Hamiltonian for the Potts model is

$$\hat{H} = -\frac{1}{2}\sum_{i,j=1}^{N} J_{ij}\,\delta_{\sigma_i,\sigma_j} - H\sum_{i=1}^{N}\delta_{\sigma_i,1} \quad . \tag{7.146}$$

Here, $\sigma_i \in \{1,\ldots,q\}$, with integer $q$. This is the so-called '$q$-state Potts model'. The quantity $H$ is analogous to an external magnetic field, and preferentially aligns (for $H > 0$) the local spins in the $\sigma = 1$ direction. We will assume $H \geq 0$.

The $q$-component set is conveniently taken to be the set $\{1,\ldots,q\}$, but it could be anything, such as

$$\sigma_i \in \{\text{tomato, penny, ostrich, Grateful Dead ticket from 1987, } \ldots\} \quad . \tag{7.147}$$

The interaction energy is $-J_{ij}$ if sites $i$ and $j$ contain the same object ($q$ possibilities), and $0$ if $i$ and $j$ contain different objects ($q^2 - q$ possibilities).

The two-state Potts model is equivalent to the Ising model. Let the allowed values of $\sigma$ be $\pm 1$. Then the quantity

$$\delta_{\sigma,\sigma'} = \tfrac{1}{2} + \tfrac{1}{2}\,\sigma\sigma' \tag{7.148}$$

equals $1$ if $\sigma = \sigma'$, and is zero otherwise. The three-state Potts model cannot be written as a simple three-state Ising model, *i.e.* one with a bilinear interaction $\sigma\,\sigma'$ where $\sigma \in \{-1, 0, +1\}$. However, it is straightforward to verify the identity

$$\delta_{\sigma,\sigma'} = 1 + \tfrac{1}{2}\,\sigma\sigma' + \tfrac{3}{2}\,\sigma^2\sigma'^2 - (\sigma^2 + \sigma'^2) \quad . \tag{7.149}$$

Thus, the $q = 3$-state Potts model is equivalent to a $S = 1$ (three-state) Ising model which includes both bilinear ($\sigma\sigma'$) and biquadratic ($\sigma^2\sigma'^2$) interactions, as well as a local field term which couples to the square of the spin, $\sigma^2$. In general one can find such correspondences for higher $q$ Potts models, but, as should be expected, the interactions become increasingly complex, with bi-cubic, bi-quartic, bi-quintic, *etc.* terms. Such a formulation, however, obscures the beautiful $S_q$ symmetry inherent in the model, where $S_q$ is the permutation group on $q$ symbols, which has $q!$ elements.

Getting back to the mean field theory, we write the single site variational density matrix $\varrho$ as a diagonal matrix with entries

$$\varrho(\sigma) = x\,\delta_{\sigma,1} + \left(\frac{1-x}{q-1}\right)\big(1 - \delta_{\sigma,1}\big) \quad , \tag{7.150}$$

with $\varrho_N(\sigma_1, \ldots, \sigma_N) = \varrho(\sigma_1) \cdots \varrho(\sigma_N)$. Note that $\mathsf{Tr}\,(\varrho) = 1$. The variational parameter is $x$. When $x = q^{-1}$, all states are equally probable. But for $x > q^{-1}$, the state $\sigma = 1$ is preferred, and the other $(q-1)$ states have identical but smaller probabilities. It is a simple matter to compute the energy and entropy:

$$\begin{aligned} E &= \mathsf{Tr}\,(\varrho_N \hat{H}) = -\tfrac{1}{2} N \hat{J}(0) \left\{ x^2 + \frac{(1-x)^2}{q-1} \right\} - NHx \\ S &= -k_{\rm B}\,\mathsf{Tr}\,(\varrho_N \log \varrho_N) = -N k_{\rm B} \left\{ x \log x + (1-x) \log\left(\frac{1-x}{q-1}\right) \right\} \quad . \end{aligned} \tag{7.151}$$

The dimensionless free energy per site is then

$$f(x, \theta, h) = -\tfrac{1}{2} \left\{ x^2 + \frac{(1-x)^2}{q-1} \right\} + \theta \left\{ x \log x + (1-x) \log\left(\frac{1-x}{q-1}\right) \right\} - hx \quad , \tag{7.152}$$

where $\theta = k_{\rm B}T/\hat{J}(0)$ and $h = H/\hat{J}(0)$. We now extremize with respect to $x$ to obtain the mean field equation,

$$\frac{\partial f}{\partial x} = 0 = -x + \frac{1-x}{q-1} + \theta\,\log x - \theta\,\log\left(\frac{1-x}{q-1}\right) - h \quad . \tag{7.153}$$

For reference, the second derivative of $f$ with respect to $x$ is given by

$$\frac{\partial^2 f}{\partial x^2} = -\frac{q}{q-1} + \frac{\theta}{x(1-x)} \quad , \tag{7.154}$$

and solving for $f''(x) = 0$, we obtain

$$x_\pm = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{\theta}{\theta_0}} \quad , \qquad \theta_0 = \frac{q}{4(q-1)} \quad . \tag{7.155}$$

Note that when $h = 0$, $x = q^{-1}$ is a solution to $f'(x) = 0$, corresponding to a $S_q$-symmetric disordered state in which all states are equally probable. At high temperatures, for small $h$, we expect $x - q^{-1} \propto h$. Using `Mathematica`, one can set $x \equiv q^{-1} + u$ and expand the mean field equation in powers of $u$. One obtains

$$h = \frac{q\,(q\theta - 1)}{q-1}\,u + \frac{q^3\,(q-2)\,\theta}{2\,(q-1)^2}\,u^2 + \mathcal{O}(u^3) \quad . \tag{7.156}$$

For weak fields, $|h| \ll 1$, and in the high temperature phase we have

$$u(\theta) = \frac{(q-1)\,h}{q\,(q\theta - 1)} + \mathcal{O}(h^2) \quad , \tag{7.157}$$

which again is of the Curie-Weiss form, with a Curie temperature $\theta_{\rm C} = q^{-1}$. The difference $u = x - q^{-1}$ is the order parameter for the transition. However, for $q > 2$ we expect the transition to be first order, which we can see by expanding the free energy in powers of $u$, obtaining the Landau expansion. Expressed in terms of the order parameter $u$ (rather than $x = q^{-1} + u$), the free energy $f(u, \theta, h)$ is

$$f(u,\theta,h) = f_0 - \frac{qu^2}{2(q-1)} - hu + \theta\,(q^{-1} + u)\log(1 + qu) - \theta\,(1 - q^{-1} - u)\log\left(1 - \frac{qu}{q-1}\right) \quad , \tag{7.158}$$

with $f_0 = -\theta \log q - (h + \frac{1}{2})\,q^{-1}$; note that $f(u = 0, \theta, h) = f_0$. Expanding in powers of $u$, we find

$$\begin{aligned} f(u,\theta,h) = f_0 - hu + \frac{q^2\,(\theta - q^{-1})}{2\,(q-1)}\,u^2 - \frac{(q-2)\,q^3\,\theta}{6\,(q-1)^2}\,u^3 \\ + \frac{q^3\theta}{12}\Big[1 + (q-1)^{-3}\Big]u^4 - \frac{q^4\theta}{20}\Big[1 - (q-1)^{-4}\Big]u^5 + \frac{q^5\theta}{30}\Big[1 + (q-1)^{-5}\Big]u^6 + \ldots \quad . \end{aligned} \tag{7.159}$$

Note that, for $q = 2$, the coefficients of $u^3$, $u^5$, and higher order odd powers of $u$ vanish in the Landau expansion. This is consistent with what we found for the Ising model, and is related to the $\mathbb{Z}_2$ symmetry of that model. For $q \geq 3$, there is a cubic term in the mean field free energy, and thus we generically expect a first order transition, as we found in §7.5.2. It is important to understand that while the mean field theory which yields the free energy function $f(u, \theta, h)$ is an approximation to a full interacting theory for which no analytic solutions may exist, truncating the MF free energy to yield a low-order polynomial in $u$ and $h$, *i.e.* a Landau free energy, constitutes a further approximation to the mean field theory itself. In particular, if at $h = 0$ we include only terms in $f(u, \theta)$ up to fourth order in $u$, we will obtain a free energy of the form $f(u) = \frac{1}{2}au^2 - \frac{1}{3}yu^3 + \frac{1}{4}bu^4$, where $\{a, y, b\}$ may be read off from eqn. 7.159. We found earlier that such a free energy results in a first order transition at $a = 2y^2/9b$. One can thus compute expressions for the critical temperature $\theta_{\rm c}^{\rm L}$ and order parameter discontinuity $u(\theta_{\rm c}^{\rm L-})$ as functions of $q$, but these will not be the same as what one obtains by working with the untruncated mean field free energy. For $q = 2$ there is a $\mathbb{Z}_2$ symmetry and our Potts model mean field theory predicts a

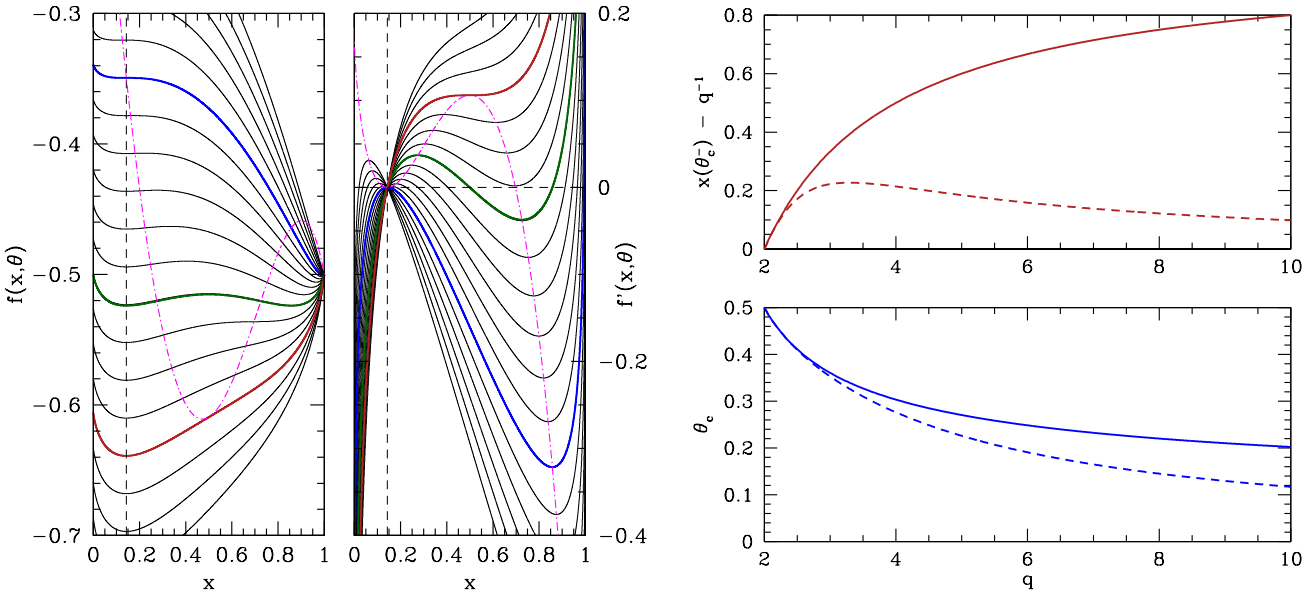



Figure 7.20: The left panels show the variational free energy $f(x,\theta)$ of the $q = 7$ state Potts model and its $x$-derivative $f'(x,\theta)$ *versus* the variational parameter $x$, for a set of equally spaced temperatures. The vertical dashed black line is at $x = q^{-1} = \frac{1}{7}$. The dot-dash magenta curve in both cases is the locus of points for which the second derivative $f''(x,\theta)$ with respect to $x$ vanishes. Three characteristic temperatures are marked: (i) $\theta = q^{-1}$ (blue), where the coefficient of the quadratic term in the Landau expansion changes sign, (ii) $\theta = \theta_0$ (red), where there is a saddle-node bifurcation and above which the free energy has only one minimum at $x = q^{-1}$ (symmetric phase), and (iii) $\theta = \theta_c$ (green), where the first order transition occurs. The right panels show comparisons of the order parameter jump at $\theta_c$ (top) and critical temperature $\theta_c$ (bottom) *versus* $q$ for untruncated (solid lines) and truncated (dashed lines) expansions of the mean field free energy. Note the agreement as $q \to 2$, where the jump is small and a truncated expansion is then valid.

second order transition at $\theta_c = \frac{1}{2}$.[25] For $q > 2$, the mean field critical point is obtained by simultaneously solving the equations $f'(u) = 0$ and $f(0) = f(u)$, both at $h = 0$. These are two equations for the two unknowns $u_c$ and $\theta_c$, with $u_c > 0$ the value of the order parameter just inside the ordered phase. This can be done numerically, but there is an analytic solution (see fig. 7.20):

$$\theta_c = \frac{q-2}{2(q-1)\log(q-1)} \quad , \qquad u_c = u(\theta_c^-) = 1 - \frac{2}{q} \quad . \tag{7.160}$$

In fig. 7.20 we plot the mean field free energy $f(x,\theta)$ and its $x$-derivative $f'(x,\theta)$ *versus* $x$ for an equally spaced set of temperatures. Also plotted are the critical temperature $\theta_c$ and the order parameter discontinuity $u(\theta_c^-)$ *versus* $q$ for the $q$-state Potts model MFT. For comparison, $\theta_c$ and $u(\theta_c^-)$ are also plotted for the truncated quartic free energy Landau theory.

[25] In §7.6.2 we found $\theta_c = 1$ for the Ising model. So why do we find $\theta_c = \frac{1}{2}$ for the $q = 2$ Potts model?

### 7.6.4 $XY$ Model

**Variational density matrix**

Consider the so-called $XY$ model, in which each site contains a continuous planar spin, represented by an angular variable $\phi_i \in [-\pi, \pi]$:

$$\hat{H} = -\frac{1}{2}\sum_{i\neq j}^{N} J_{ij}\cos\big(\phi_i - \phi_j\big) - H\sum_{i=1}^{N}\cos\phi_i \quad . \tag{7.161}$$

We write the (diagonal elements of the) full density matrix once again as a product:

$$\varrho_N(\boldsymbol{\phi}) = \prod_{i=1}^{N}\varrho(\phi_i) \quad . \tag{7.162}$$

Our goal will be to extremize the free energy with respect to the *function* $\varrho(\phi)$. To this end, we compute

$$E = \mathsf{Tr}\,(\varrho_N\,\hat{H}) = -\tfrac{1}{2}N\hat{J}(0)\,\Big|\mathsf{Tr}\,\big(\varrho\,e^{i\phi}\big)\Big|^2 - NH\,\mathsf{Tr}\,\big(\varrho\,\cos\phi\big) \quad . \tag{7.163}$$

The entropy is $S = -Nk_{\rm B}\,\mathsf{Tr}\,(\varrho\,\log\varrho)$. Note that for any function $A(\phi)$, we have[26]

$$\mathsf{Tr}\,(\varrho\,A) \equiv \int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\varrho(\phi)\,A(\phi) \quad . \tag{7.164}$$

We now extremize the *functional* $F\big[\varrho(\phi)\big] = E - TS$ with respect to $\varrho(\phi)$, under the condition that $\mathsf{Tr}\;\varrho = 1$. We therefore use Lagrange's method of undetermined multipliers, writing

$$F^* = F - Nk_{\rm B}T\,\lambda\,\Big(\mathsf{Tr}\,\varrho - 1\Big) \quad . \tag{7.165}$$

Note that $F^*$ is a *function* of the Lagrange multiplier $\lambda$ and a *functional* of the density matrix $\varrho(\phi)$. The prefactor $Nk_{\rm B}T$ which multiplies $\lambda$ is of no mathematical consequence – we could always redefine the multiplier to be $\lambda' \equiv Nk_{\rm B}T\lambda$. It is present only to maintain homogeneity and proper dimensionality of $F^*$ with $\lambda$ itself dimensionless and of order $N^0$. We now have

$$\begin{aligned}\frac{\delta F^*}{\delta\varrho(\phi)} = \frac{\delta}{\delta\varrho(\phi)}\Bigg\{&-\tfrac{1}{2}N\hat{J}(0)\,\Big|\mathsf{Tr}\big(\varrho\,e^{i\phi}\big)\Big|^2 - NH\,\mathsf{Tr}\big(\varrho\,\cos\phi\big)\\ &+ Nk_{\rm B}T\;\mathsf{Tr}\,\big(\varrho\,\log\varrho\big) - Nk_{\rm B}T\,\lambda\,\Big(\mathsf{Tr}\,\varrho - 1\Big)\Bigg\} \quad .\end{aligned} \tag{7.166}$$

[26]The denominator of $2\pi$ in the measure is not necessary, and in fact it is even slightly cumbersome. It divides out whenever we take a ratio to compute a thermodynamic average. I introduce this factor to preserve the relation $\mathsf{Tr}\;1 = 1$. I personally find unnormalized traces to be profoundly unsettling on purely aesthetic grounds.

To this end, we note that

$$\frac{\delta}{\delta\varrho(\phi)}\,\mathsf{Tr}\,(\varrho\,A) = \frac{\delta}{\delta\varrho(\phi)}\int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\varrho(\phi)\,A(\phi) = \frac{1}{2\pi}\,A(\phi) \quad . \tag{7.167}$$

Thus, we have

$$\frac{\delta F^*}{\delta\varrho(\phi)} = -\frac{N\hat{J}(0)}{4\pi}\left[\mathop{\mathsf{Tr}}_{\phi'}\big(\varrho\,e^{i\phi'}\big)\,e^{-i\phi} + \mathop{\mathsf{Tr}}_{\phi'}\big(\varrho\,e^{-i\phi'}\big)\,e^{i\phi}\right] - \frac{NH}{2\pi}\cos\phi + \frac{Nk_{\rm B}T}{2\pi}\Big[\log\varrho(\phi) + 1 - \lambda\Big] \quad . \tag{7.168}$$

Now let us define

$$\mathop{\mathsf{Tr}}_{\phi}\big(\varrho\,e^{i\phi}\big) = \int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\varrho(\phi)\,e^{i\phi} \equiv m\,e^{i\phi_0} \quad . \tag{7.169}$$

We then have

$$\log\varrho(\phi) = \frac{\hat{J}(0)}{k_{\rm B}T}\,m\,\cos(\phi-\phi_0) + \frac{H}{k_{\rm B}T}\,\cos\phi + \lambda - 1. \tag{7.170}$$

Clearly the free energy will be reduced if $\phi_0 = 0$ so that the mean field is maximal and aligns with the external field, which prefers $\phi = 0$. Thus, we conclude

$$\varrho(\phi) = \mathcal{C}\,\exp\!\left(\frac{H_{\rm eff}}{k_{\rm B}T}\,\cos\phi\right) \quad , \tag{7.171}$$

where $H_{\rm eff} = \hat{J}(0)\,m + H$ and $\mathcal{C} = e^{\lambda-1}$. The value of $\lambda$ is then determined by invoking the constraint,

$$\mathsf{Tr}\,\varrho = 1 = \mathcal{C}\int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\exp\!\left(\frac{H_{\rm eff}}{k_{\rm B}T}\,\cos\phi\right) = \mathcal{C}\,I_0(H_{\rm eff}/k_{\rm B}T) \quad , \tag{7.172}$$

where $I_0(z)$ is the Bessel function. We are free to define $\varepsilon \equiv H_{\rm eff}/k_{\rm B}T$, and treat $\varepsilon$ as our single variational parameter. We then have the normalized single site density matrix

$$\varrho(\phi) = \frac{\exp(\varepsilon\cos\phi)}{\int\limits_{-\pi}^{\pi}\frac{d\phi'}{2\pi}\,\exp(\varepsilon\cos\phi')} = \frac{\exp(\varepsilon\cos\phi)}{I_0(\varepsilon)} \quad . \tag{7.173}$$

We next compute the following averages:

$$\big\langle\,e^{\pm i\phi}\,\big\rangle = \int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\varrho(\phi)\,e^{\pm i\phi} = \frac{I_1(\varepsilon)}{I_0(\varepsilon)} \tag{7.174}$$

$$\big\langle\cos(\phi-\phi')\big\rangle = \mathsf{Re}\,\big\langle e^{i\phi}\,e^{-i\phi'}\big\rangle = \left(\frac{I_1(\varepsilon)}{I_0(\varepsilon)}\right)^2 \quad , \tag{7.175}$$

as well as

$$\mathsf{Tr}\,(\varrho\,\log\varrho) = \int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\frac{e^{\varepsilon\cos\phi}}{I_0(\varepsilon)}\Big\{\varepsilon\cos\phi - \log I_0(\varepsilon)\Big\} = \varepsilon\,\frac{I_1(\varepsilon)}{I_0(\varepsilon)} - \log I_0(\varepsilon) \quad . \tag{7.176}$$

The dimensionless free energy per site is therefore

$$f(\varepsilon,\theta,h) = -\frac{1}{2}\left(\frac{I_1(\varepsilon)}{I_0(\varepsilon)}\right)^2 + (\theta\varepsilon - h)\,\frac{I_1(\varepsilon)}{I_0(\varepsilon)} - \theta\,\log I_0(\varepsilon) \quad , \tag{7.177}$$

with $\theta = k_{\rm B}T/\hat{J}(0)$ and $h = H/\hat{J}(0)$ and $f = F/N\hat{J}(0)$ as before. Note that the mean field equation is $m = \theta\varepsilon - h = \langle\, e^{i\phi}\,\rangle$, *i.e.*

$$\theta\varepsilon - h = \frac{I_1(\varepsilon)}{I_0(\varepsilon)} \quad . \tag{7.178}$$

For small $\varepsilon$, we may expand the Bessel functions, using

$$I_\nu(z) = (\tfrac{1}{2}z)^\nu \sum_{k=0}^{\infty} \frac{(\frac{1}{4}z^2)^k}{k!\,\Gamma(k+\nu+1)} \quad , \tag{7.179}$$

to obtain

$$f(\varepsilon,\theta,h) = \tfrac{1}{4}\left(\theta - \tfrac{1}{2}\right)\varepsilon^2 + \tfrac{1}{64}\left(2 - 3\theta\right)\varepsilon^4 - \tfrac{1}{2}h\varepsilon + \tfrac{1}{16}h\varepsilon^3 + \ldots \quad . \tag{7.180}$$

This predicts a second order phase transition at $\theta_{\rm c} = \frac{1}{2}$.[27] Note also the Curie-Weiss form of the susceptibility at high $\theta$:

$$\frac{\partial f}{\partial \varepsilon} = 0 \quad \Longrightarrow \quad \varepsilon = \frac{h}{\theta - \theta_{\rm c}} + \ldots \quad . \tag{7.181}$$

**Neglect of fluctuations method**

Consider again the Hamiltonian of eqn. 7.161. Define $z_i \equiv \exp(i\phi_i)$ and write $z_i = w + \delta z_i$, where $w \equiv \langle z_i\rangle$ and $\delta z_i \equiv z_i - w$. Of course we also have the complex conjugate relations $z_i^* = w^* + \delta z_i^*$ and $w^* = \langle z_i^*\rangle$. Writing $\cos(\phi_i - \phi_j) = \mathsf{Re}\,(z_i^* z_j)$, by neglecting the terms proportional to $\delta z_i^*\,\delta z_j$ in $\hat{H}$ we arrive at the mean field Hamiltonian,

$$\hat{H}^{\rm MF} = \tfrac{1}{2}N\hat{J}(0)\,|w|^2 - \tfrac{1}{2}\hat{J}(0)\,|w| \sum_{i=1}^{N}\left(w^* z_i + w z_i^*\right) - \tfrac{1}{2}H\sum_{i=1}^{N}\left(z_i^* + z_i\right) \tag{7.182}$$

It is clear that the free energy will be minimized if the mean field $w$ breaks the $\mathsf{O}(2)$ symmetry in the same direction as the external field $H$, which means $w \in \mathbb{R}$ and

$$\hat{H}^{\rm MF} = \tfrac{1}{2}N\hat{J}(0)\,|w|^2 - \left(H + \hat{J}(0)\,|w|\right)\sum_{i=1}^{N}\cos\phi_i \quad . \tag{7.183}$$

The dimensionless free energy per site is then

$$f = \tfrac{1}{2}|w|^2 - \theta\log I_0\!\left(\frac{h+|w|}{\theta}\right) \quad . \tag{7.184}$$

[27] Note that the coefficient of the quartic term in $\varepsilon$ is negative for $\theta > \frac{2}{3}$. At $\theta = \theta_{\rm c} = \frac{1}{2}$, the coefficient is positive, but for larger $\theta$ one must include higher order terms in the Landau expansion.

Differentiating with respect to $|w|$, one obtains

$$|w| \equiv m = I_1\left(\frac{h+m}{\theta}\right) \Big/ I_0\left(\frac{h+m}{\theta}\right) \quad , \tag{7.185}$$

which is the same equation as eqn. 7.178. The two mean field theories yield the same results in every detail (see §7.11).

## 7.7 Mean Field Theory of Fluctuations

### 7.7.1 Correlation and response in mean field theory

We now consider the matter of correlation and response functions within mean field theory. A correlation function is a thermodynamic average of fields at various sites, such as $\langle \sigma_i \sigma_j \rangle$ (two-point correlation function), $\langle \sigma_i \sigma_j \sigma_k \sigma_l \rangle$ (four-point correlation function), *etc.* A response function describes the influence of external fields on thermodynamic averages, such as $\partial m_i / \partial H_j$, the variation of the local magnetization $m_i = \langle \sigma_i$ at site $i$ due to the presence of a local field $H_j$ at site $j$. We shall see that there is a direct relation between correlation and response functions.

Consider the Ising model,

$$\hat{H} = -\tfrac{1}{2} \sum_{i,j=1}^{N} J_{ij}\, \sigma_i\, \sigma_j - \sum_{i=1}^{N} H_i\, \sigma_i \quad , \tag{7.186}$$

where the local magnetic field on site $k$ is now $H_k$. We assume without loss of generality that the diagonal terms vanish: $J_{ii} = 0$. Now consider the partition function $Z = \mathsf{Tr}\; e^{-\beta \hat{H}}$ as a function of the temperature $T$ and the local field values $\{H_i\}$. We have

$$\begin{aligned} \frac{\partial Z}{\partial H_i} &= \beta\, \mathsf{Tr}\left[\sigma_i\, e^{-\beta \hat{H}}\right] = \beta Z \langle \sigma_i \rangle \\ \frac{\partial^2 Z}{\partial H_i\, \partial H_j} &= \beta^2\, \mathsf{Tr}\left[\sigma_i\, \sigma_j\, e^{-\beta \hat{H}}\right] = \beta^2 Z \langle \sigma_i\, \sigma_j \rangle \quad . \end{aligned} \tag{7.187}$$

Thus,

$$\begin{aligned} m_i &= -\frac{\partial F}{\partial H_i} = \langle \sigma_i \rangle \\ \chi_{ij} &= \frac{\partial m_i}{\partial H_j} = -\frac{\partial^2 F}{\partial H_i\, \partial H_j} = \frac{1}{k_{\rm B} T} \cdot \Big\{ \langle \sigma_i\, \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle \Big\} \quad . \end{aligned} \tag{7.188}$$

Expressions such as $\langle \sigma_i \rangle$, $\langle \sigma_i\, \sigma_j \rangle$, *etc.* are in general called *correlation functions*. For example, we define the *spin-spin correlation function* $C_{ij}$ as

$$C_{ij} \equiv \langle \sigma_i\, \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle \quad . \tag{7.189}$$

Expressions such as $\partial F/\partial H_i$ and $\partial^2 F/\partial H_i\,\partial H_j$ are called *response functions*. The above relation between correlation functions and response functions, $C_{ij} = k_{\rm B}T\,\chi_{ij}$, is valid *only for the equilibrium distribution*. In particular, this relationship is *invalid* if one uses an approximate distribution, such as the variational density matrix formalism of mean field theory.

The question then arises: within mean field theory, which is more accurate – correlation functions or response functions? A simple argument suggests that the *response functions* are more accurate representations of the real physics. To see this, let's write the variational density matrix $\varrho^{\rm var}$ as the sum of the exact equilibrium (Boltzmann) distribution $\varrho^{\rm eq} = Z^{-1}\exp(-\beta\hat{H})$ plus a deviation $\delta\varrho$:

$$\varrho^{\rm var} = \varrho^{\rm eq} + \delta\varrho \quad . \tag{7.190}$$

Then if we calculate a correlator using the variational distribution, we have

$$\langle\sigma_i\,\sigma_j\rangle_{\rm var} = {\rm Tr}\left[\varrho^{\rm var}\,\sigma_i\,\sigma_j\right] = {\rm Tr}\left[\varrho^{\rm eq}\,\sigma_i\,\sigma_j\right] + {\rm Tr}\left[\delta\varrho\,\sigma_i\,\sigma_j\right] \quad . \tag{7.191}$$

Thus, the variational density matrix gets the correlator right to first order in $\delta\varrho$. On the other hand, the free energy is given by

$$F^{\rm var} = F^{\rm eq} + \sum_{\boldsymbol{\sigma}} \frac{\partial F}{\partial \varrho_{\boldsymbol{\sigma}}}\bigg|_{\varrho^{\rm eq}} \delta\varrho_{\boldsymbol{\sigma}} + \frac{1}{2}\sum_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} \frac{\partial^2 F}{\partial\varrho_{\boldsymbol{\sigma}}\partial\varrho_{\boldsymbol{\sigma}'}}\bigg|_{\varrho^{\rm eq}} \delta\varrho_{\boldsymbol{\sigma}}\,\delta\varrho_{\boldsymbol{\sigma}'} + \ldots \quad . \tag{7.192}$$

Here $\boldsymbol{\sigma}$ denotes a state of the system, *i.e.* $|\,\boldsymbol{\sigma}\,\rangle = |\,\sigma_1,\ldots,\sigma_N\,\rangle$, where every spin polarization is specified. Since the free energy is an extremum (and in fact an absolute minimum) with respect to the distribution, the second term on the RHS vanishes. This means that the free energy is accurate to second order in the deviation $\delta\varrho$.

### 7.7.2 Calculation of the response functions

Consider the variational density matrix $\varrho(\boldsymbol{\sigma}) = \prod_{i=1}^N \varrho_i(\sigma_i)$, where

$$\varrho_i(\sigma_i) = \left(\frac{1+m_i}{2}\right)\delta_{\sigma_i,1} + \left(\frac{1-m_i}{2}\right)\delta_{\sigma_i,-1} \quad . \tag{7.193}$$

The variational energy $E = {\rm Tr}\,(\varrho\,\hat{H})$ is

$$E = -\tfrac{1}{2}\sum_{i,j=1^N} J_{ij}\,m_i\,m_j - \sum_{i=1}^N H_i\,m_i \tag{7.194}$$

and the entropy $S = -k_{\rm B}\,{\rm Tr}\,(\varrho\log\varrho)$ is

$$S = -k_{\rm B}\sum_{i=1}^N \left\{\left(\frac{1+m_i}{2}\right)\log\left(\frac{1+m_i}{2}\right) + \left(\frac{1-m_i}{2}\right)\log\left(\frac{1-m_i}{2}\right)\right\} \quad . \tag{7.195}$$

Setting the variation $\partial F/\partial m_i = 0$, with $F = E - TS$, we obtain the mean field equations,

$$m_i = \tanh\left(\beta J_{ij}\,m_j + \beta H_i\right) \quad , \tag{7.196}$$

where we use the summation convention: $J_{ij}\,m_j \equiv \sum_j J_{ij}\,m_j$. Suppose $T > T_{\rm c}$ and $m_i$ is small. Then we can expand the RHS of the above mean field equations, obtaining

$$\big(\delta_{ij} - \beta J_{ij}\big)\,m_j = \beta H_i \quad . \tag{7.197}$$

Thus, the susceptibility tensor $\chi$ is the inverse of the matrix $(k_{\rm B}T\,\mathbb{1} - \mathbb{J})$:

$$\chi_{ij} = \frac{\partial m_i}{\partial H_j} = \big(k_{\rm B}T\,\mathbb{1} - \mathbb{J}\big)^{-1}_{ij} \quad , \tag{7.198}$$

where $\mathbb{1}$ is the identity. Note also that so-called *connected averages* of the kind in eqn. 7.189 vanish identically if we compute them using our variational density matrix, since all the sites are independent, hence

$$\langle\sigma_i\,\sigma_j\rangle = {\rm Tr}\,\big(\varrho^{\rm var}\,\sigma_i\,\sigma_j\big) = {\rm Tr}\,\big(\varrho_i^{\rm var}\,\sigma_i\big)\,{\rm Tr}\,\big(\varrho_j^{\rm var}\,\sigma_j\big) = \langle\sigma_i\rangle\langle\sigma_j\rangle \quad , \tag{7.199}$$

and therefore $\chi_{ij} = 0$ if we compute the correlation functions themselves from the variational density matrix, rather than from the free energy $F$. As we have argued above, the latter approximation is more accurate.

Assuming $J_{ij} = J(\boldsymbol{R}_i - \boldsymbol{R}_j)$, where $\boldsymbol{R}_i$ is a Bravais lattice site, we can Fourier transform the above equation, resulting in

$$\hat{m}(\boldsymbol{q}) = \frac{\hat{H}(\boldsymbol{q})}{k_{\rm B}T - \hat{J}(\boldsymbol{q})} \equiv \hat{\chi}(\boldsymbol{q})\,\hat{H}(\boldsymbol{q}) \quad . \tag{7.200}$$

Once again, our definition of lattice Fourier transform of a function $\phi(\boldsymbol{R})$ is

$$\hat{\phi}(\boldsymbol{q}) \equiv \sum_{\boldsymbol{R}} \phi(\boldsymbol{R})\,e^{-i\boldsymbol{q}\cdot\boldsymbol{R}} \qquad , \qquad \phi(\boldsymbol{R}) = \Omega\int\limits_{\hat{\Omega}} \frac{d^dq}{(2\pi)^d}\,\hat{\phi}(\boldsymbol{q})\,e^{i\boldsymbol{q}\cdot\boldsymbol{R}} \quad , \tag{7.201}$$

where $\Omega$ is the unit cell in real space, called the *Wigner-Seitz cell*, and $\hat{\Omega}$ is the first Brillouin zone, which is the unit cell in *reciprocal space*. Similarly, we have

$$\begin{aligned}\hat{J}(\boldsymbol{q}) &= \sum_{\boldsymbol{R}} J(\boldsymbol{R})\Big(1 - i\boldsymbol{q}\cdot\boldsymbol{R} - \tfrac{1}{2}(\boldsymbol{q}\cdot\boldsymbol{R})^2 + \ldots\Big) \\ &= \hat{J}(0)\cdot\Big\{1 - q^2R_*^2 + \mathcal{O}(q^4)\Big\} \quad ,\end{aligned} \tag{7.202}$$

where

$$R_*^2 = \frac{\sum_{\boldsymbol{R}} \boldsymbol{R}^2 J(\boldsymbol{R})}{2d\sum_{\boldsymbol{R}} J(\boldsymbol{R})} \quad . \tag{7.203}$$

Here we have assumed inversion symmetry for the lattice, in which case

$$\sum_{\boldsymbol{R}} R^\mu R^\nu J(\boldsymbol{R}) = \frac{1}{d}\cdot\delta^{\mu\nu}\sum_{\boldsymbol{R}} \boldsymbol{R}^2 J(\boldsymbol{R}) \quad . \tag{7.204}$$

On cubic lattices with nearest neighbor interactions only, one has $R_* = a/\sqrt{2d}$, where $a$ is the lattice constant and $d$ is the dimension of space.

Thus, with the identification $k_{\rm B}T_{\rm c} = \hat{J}(0)$, we have

$$\hat{\chi}(\boldsymbol{q}) = \frac{1}{k_{\rm B}(T - T_{\rm c}) + k_{\rm B}T_{\rm c}\, R_*^2\, \boldsymbol{q}^2 + \mathcal{O}(q^4)} = \frac{1}{k_{\rm B}T_{\rm c}\, R_*^2} \cdot \frac{1}{\xi^{-2} + q^2 + \mathcal{O}(q^4)} \quad , \tag{7.205}$$

where

$$\xi(T) = R_* \cdot \left(\frac{T - T_{\rm c}}{T_{\rm c}}\right)^{-1/2} \tag{7.206}$$

is the *correlation length*. With the definition $\xi(T) \propto |T - T_{\rm c}|^{-\nu}$ as $T \to T_{\rm c}$, we obtain the mean field correlation length exponent $\nu = \frac{1}{2}$. The exact result for the two-dimensional Ising model is $\nu = 1$, whereas $\nu \approx 0.6$ for the $d = 3$ Ising model. Note that $\hat{\chi}(\boldsymbol{q} = 0, T)$ diverges as $(T - T_{\rm c})^{-1}$ for $T > T_{\rm c}$.

In real space, we have

$$m_i = \sum_j \chi_{ij}\, H_j \qquad , \qquad \chi_{ij} = \Omega \int\limits_{\hat{\Omega}} \frac{d^d q}{(2\pi)^d}\, \hat{\chi}(\boldsymbol{q})\, e^{i\boldsymbol{q}\cdot(\boldsymbol{R}_i - \boldsymbol{R}_j)} \quad . \tag{7.207}$$

Note that $\hat{\chi}(\boldsymbol{q})$ is properly periodic under $\boldsymbol{q} \to \boldsymbol{q} + \boldsymbol{G}$, where $\boldsymbol{G}$ is a reciprocal lattice vector, which satisfies $e^{i\boldsymbol{G}\cdot\boldsymbol{R}} = 1$ for any direct Bravais lattice vector $\boldsymbol{R}$. Indeed, we have

$$\hat{\chi}^{-1}(\boldsymbol{q}) = k_{\rm B}T - \hat{J}(\boldsymbol{q}) = k_{\rm B}T - J \sum_{\boldsymbol{\delta}} e^{i\boldsymbol{q}\cdot\boldsymbol{\delta}} \quad , \tag{7.208}$$

where $\boldsymbol{\delta}$ is a nearest neighbor separation vector, and where in the second line we have assumed nearest neighbor interactions only. On cubic lattices in $d$ dimensions, there are $2d$ nearest neighbor separation vectors, $\boldsymbol{\delta} = \pm a\, \hat{\mathbf{e}}_\mu$, where $\mu \in \{1, \ldots, d\}$. The real space susceptibility is then

$$\chi(\boldsymbol{R}, H = 0) = \int\limits_{-\pi}^{\pi} \frac{d\theta_1}{2\pi} \cdots \int\limits_{-\pi}^{\pi} \frac{d\theta_d}{2\pi}\, \frac{e^{in_1\theta_1} \cdots e^{in_d\theta_d}}{k_{\rm B}T - (2J\cos\theta_1 + \ldots + 2J\cos\theta_d)} \quad , \tag{7.209}$$

where $\boldsymbol{R} = a\sum_{\mu=1}^d n_\mu\, \hat{\mathbf{e}}_\mu$ is a general direct lattice vector for the cubic Bravais lattice in $d$ dimensions, and the $\{n_\mu\}$ are integers.

The long distance behavior was discussed in chapter 6 (see §6.4.7 on Ornstein-Zernike theory[28]). For convenience we reiterate those results:

- In $d = 1$ dimension,

$$\chi_{d=1}(x) = \left(\frac{\xi}{2k_{\rm B}T_{\rm c}R_*^2}\right) e^{-|x|/\xi} \quad . \tag{7.210}$$

- In $d = 2$ dimensions,

$$\chi_{d=2}(\boldsymbol{x}) = \left(\frac{1}{2\pi k_{\rm B}T_{\rm c}R_*^2}\right) K_0(|\boldsymbol{x}|/\xi) \quad . \tag{7.211}$$

[28]There is a sign difference between the particle susceptibility defined in chapter 6 and the spin susceptibility defined here. The origin of the difference is that the single particle potential $v$ as defined was repulsive for $v > 0$, meaning the local density response $\delta n$ should be negative, while in the current discussion a positive magnetic field $H$ prefers $m > 0$.

- In $d > 2$, with $r \to \infty$ and $\xi$ fixed,

$$\chi_d^{\rm OZ}(\boldsymbol{r}) \simeq C_d \cdot \frac{\xi^{(3-d)/2}}{k_{\rm B}T\,R_*^2} \cdot \frac{e^{-r/\xi}}{r^{(d-1)/2}} \cdot \left\{1 + \mathcal{O}\left(\frac{d-3}{r/\xi}\right)\right\} \quad , \tag{7.212}$$

  where the $C_d$ are dimensionless constants.

- In $d > 2$, with $\xi \to \infty$ and $r$ fixed (*i.e.* $T \to T_{\rm c}$ at fixed separation $\boldsymbol{r}$),

$$\chi_d(\boldsymbol{r}) \simeq \frac{C_d'}{k_{\rm B}TR_*^2} \cdot \frac{e^{-r/\xi}}{r^{d-2}} \cdot \left\{1 + \mathcal{O}\left(\frac{d-3}{r/\xi}\right)\right\} \quad . \tag{7.213}$$

Close to the critical point the spatial dependence of the two-point correlation $C(\boldsymbol{r},T) = k_{\rm B}T\,\chi(\boldsymbol{r},T)$ is given by

$$C(\boldsymbol{r},T) = r^{-(d-2+\eta)}\,\phi\big(r/\xi(T)\big) \quad , \tag{7.214}$$

where $\eta$ is the *anomalous critical exponent* and $\phi(r/\xi)$ is a *scaling function*. The condition $T \approx T_{\rm c}$ means that $\xi(T) \gg a$, where $a$ is a microscopic length, such as a lattice constant[29]. Thus, we've encountered in this section two additional critical exponents, the anomalous exponent $\eta$ and the correlation length exponent $\nu$, which we first met in §7.3.2, whose mean field values are $\eta = 0$ and $\nu = \frac{1}{2}$.

### 7.7.3 Beyond the two-state Ising model

Consider a general spin model, and a variational density matrix $\varrho^{\rm var}$ which is a product of single site density matrices:

$$\varrho^{\rm var}\big[\{\boldsymbol{S}_i\}\big] = \prod_{i=1}^N \varrho_i(\boldsymbol{S}_i) \quad , \tag{7.215}$$

where $\boldsymbol{m}_i = {\rm Tr}\left(\varrho^{\rm var}\boldsymbol{S}_i\right)$ is the local magnetization and $\boldsymbol{S}_i$, which may be a scalar (*e.g.*, $\sigma_i$ in the Ising model previously discussed), is the local spin operator. Note that $\varrho_i(\boldsymbol{S}_i)$ depends parametrically on the variational parameter(s) $\boldsymbol{m}_i$ (*e.g.*, as in eqn. 7.137). Let the Hamiltonian be

$$\hat{H} = -\frac{1}{2}\sum_{i,j=1}^N J_{i,j}^{\mu\nu}\,S_i^\mu\,S_j^\nu + \sum_{i=1}^N \hat{h}_i(\boldsymbol{S}_i) - \sum_{i=1}^N \boldsymbol{H}_i\cdot\boldsymbol{S}_i \quad , \tag{7.216}$$

with an implied sum over repeated Greek indices $\mu$ and $\nu$. Here we separate the external applied fields $\boldsymbol{H}_i$ from the local Hamiltonians $\hat{h}_i(\boldsymbol{S}_i)$, which are presumed time-reversal symmetric, *i.e.* ${\rm Tr}\left[\boldsymbol{S}\exp(-\beta\hat{h}(\boldsymbol{S})\right] = 0$. The variational free energy is

$$F^{\rm var} = -\frac{1}{2}\sum_{i,j=1}^N J_{ij}^{\mu\nu}\,m_i^\mu\,m_j^\nu + \sum_{i=1}^N \varphi(\boldsymbol{m}_i,T) - \sum_{i=1}^N H_i^\mu\,m_i^\mu \quad , \tag{7.217}$$

[29] The scaling functions on the high and low temperature sides of the transition may be different, and are denoted as $\phi_\pm(r/\xi)$, respectively.

where the single site free energy $\varphi(\boldsymbol{m}_i, T)$ in the absence of an external field is given by

$$\varphi(\boldsymbol{m}_i, T) = \mathsf{Tr}\Big[\varrho_i(\boldsymbol{S})\,\hat{h}(\boldsymbol{S})\Big] + k_{\rm B}T\,\mathsf{Tr}\Big[\varrho_i(\boldsymbol{S})\log\varrho_i(\boldsymbol{S})\Big] \quad , \tag{7.218}$$

where $\hat{h}(\boldsymbol{S})$ is the single site Hamiltonian. We then have

$$\frac{\partial F^{\rm var}}{\partial m_i^\mu} = -\sum_j J_{ij}^{\mu\nu}\,m_j^\nu - H_i^\mu + \frac{\partial\varphi(\boldsymbol{m}_i,T)}{\partial m_i^\mu} \quad . \tag{7.219}$$

For the noninteracting system, we have $J_{ij}^{\mu\nu} = 0$, and the weak field response must be linear. In this limit we may write $m_i^\mu = \chi^0_{\mu\nu}(T)\,H_i^\nu + \mathcal{O}(\boldsymbol{H}_i^3)$, and we conclude

$$\frac{\partial\varphi(\boldsymbol{m}_i,T)}{\partial m_i^\mu} = \big[\chi^0(T)\big]^{-1}_{\mu\nu}\,m_i^\nu + \mathcal{O}\big(m_i^3\big) \quad . \tag{7.220}$$

Note that this entails the following expansion for the single site free energy:

$$\varphi(\boldsymbol{m}_i,T) = \varphi(\boldsymbol{0},T) + \tfrac{1}{2}\big[\chi^0(T)\big]^{-1}_{\mu\nu}\,m_i^\mu\,m_i^\nu + \mathcal{O}(m^4) \quad . \tag{7.221}$$

Finally, we restore the interaction term and extremize $F^{\rm var}$ by setting $\partial F^{\rm var}/\partial m_i^\mu = 0$. To linear order, then,

$$m_i^\mu = \chi^0_{\mu\nu}(T)\left(H_i^\nu + \sum_j J_{ij}^{\nu\lambda}\,m_j^\lambda\right) \quad . \tag{7.222}$$

Typically the local susceptibility is a scalar in the internal spin space, *i.e.* $\chi^0_{\mu\nu}(T) = \chi^0(T)\,\delta_{\mu\nu}$, in which case we obtain

$$\big(\delta^{\mu\nu}\,\delta_{ij} - \chi^0(T)\,J_{ij}^{\mu\nu}\big)\,m_j^\nu = \chi^0(T)\,H_i^\mu \quad . \tag{7.223}$$

In Fourier space, then,

$$\hat{\chi}_{\mu\nu}(\boldsymbol{q},T) = \chi^0(T)\left(\mathbb{1} - \chi^0(T)\,\hat{\mathbb{J}}(\boldsymbol{q})\right)^{-1}_{\mu\nu} \quad , \tag{7.224}$$

where $\hat{\mathbb{J}}(\boldsymbol{q})$ is the matrix whose elements are $\hat{J}^{\mu\nu}(\boldsymbol{q})$. If $\hat{J}^{\mu\nu}(\boldsymbol{q}) = \hat{J}(\boldsymbol{q})\,\delta^{\mu\nu}$, then the susceptibility is isotropic in spin space, with

$$\hat{\chi}(\boldsymbol{q},T) = \frac{1}{\big[\chi^0(T)\big]^{-1} - \hat{J}(\boldsymbol{q})} \quad . \tag{7.225}$$

Consider now the following examples:

(i) Spin-$S$ with $\mathbb{Z}_2$ symmetry and $\hat{h}(\hat{S}) = 0$: Write $\varrho_i(S_i) = \zeta^{-1}\,\exp(\mathfrak{h}S_i/k_{\rm B}T)$, where $\mathfrak{h}$ acts as a local field. The normalization is[30]

$$\zeta(\mathfrak{h},T) = \mathsf{Tr}\,e^{\mathfrak{h}\hat{S}/k_{\rm B}T} = \sum_{k=-S}^{S} e^{k\mathfrak{h}/k_{\rm B}T} = \frac{\sinh\big[(S+\frac{1}{2})\,\mathfrak{h}/k_{\rm B}T\big]}{\sinh\big[\mathfrak{h}/2k_{\rm B}T\big]} \quad . \tag{7.226}$$

[30]Here we distinguish the spin 'operator' $\hat{S}$ from its maximal value $S$.

The local magnetization $m_i$ is then given by $m_i = m(\mathfrak{h}_i, T) = \mathsf{Tr}(\varrho_i S_i)$. The relation between $m$, $\mathfrak{h}$, and $T$ is then given by

$$\begin{aligned} m = \langle \hat{S} \rangle &= k_{\rm B} T \, \frac{\partial \log \zeta}{\partial \mathfrak{h}} = (S + \tfrac{1}{2}) \, \mathrm{ctnh}\big[(S + \tfrac{1}{2}) \, \mathfrak{h}/k_{\rm B}T\big] - \tfrac{1}{2} \, \mathrm{ctnh}\big[\mathfrak{h}/2k_{\rm B}T\big] \\ &= \frac{S(S+1)}{3k_{\rm B}T} \, \mathfrak{h} + \mathcal{O}(\mathfrak{h}^3) \end{aligned} \tag{7.227}$$

at each site. We identify the single-site susceptibility as

$$\chi^0(T) = \frac{S(S+1)}{3k_{\rm B}T} \quad , \tag{7.228}$$

which is the Curie result. The free-field single-site free energy (see eqn. 7.218) is then

$$\varphi(m_i, T) = k_{\rm B}T \, \mathsf{Tr} \left( \varrho_i \log \varrho_i \right) = \mathfrak{h}_i \, m_i - k_{\rm B}T \log \zeta_i \quad , \tag{7.229}$$

whence for each $i$ we have

$$\frac{\partial \varphi}{\partial m} = \mathfrak{h} + m \, \frac{\partial \mathfrak{h}}{\partial m} - \overbrace{k_{\rm B}T \, \frac{\partial \log \zeta}{\partial \mathfrak{h}}}^{= m} \, \frac{\partial \mathfrak{h}}{\partial m} = \mathfrak{h} \quad \Rightarrow \quad \varphi(m, T) = \frac{m^2}{2\chi^0(T)} + \mathcal{O}(m^4) \tag{7.230}$$

and $m = \chi^0(T) \, \mathfrak{h}$.

(ii) Spin-$S$ with $\mathbb{Z}_2$ symmetry and $\hat{h}(\hat{S}) = K\hat{S}^2$: We take $\varrho_i(S_i) = \zeta^{-1} \exp(\beta \mathfrak{h} S_i - \beta K S_i^2)$, and one obtains the partition function

$$\zeta(\mathfrak{h}, T) = \sum_{j=-S}^{S} e^{\mathfrak{h} j / k_{\rm B}T} e^{-Kj^2/k_{\rm B}T} \quad . \tag{7.231}$$

The single site susceptibility is then

$$\chi^0(T) = k_{\rm B}T \, \frac{\partial^2 \log \zeta}{\partial \mathfrak{h}^2} \bigg|_{\mathfrak{h}=0} = \frac{1}{k_{\rm B}T} \cdot \frac{\sum_{j=-S}^{S} j^2 \, e^{-Kj^2/k_{\rm B}T}}{\sum_{l=-S}^{S} e^{-Kl^2/k_{\rm B}T}} \quad . \tag{7.232}$$

When $K \to 0$ we recover $\chi^0(T) = S(S+1)/3k_{\rm B}T$. In the large $K$ limit we obtain

$$\chi^0(T) = \begin{cases} \frac{2}{k_{\rm B}T} \, e^{-K/k_{\rm B}T} + \mathcal{O}(e^{-2K/k_{\rm B}T}) & \text{for } S \in \mathbb{Z} \\ \frac{1}{4\,k_{\rm B}T} & \text{for } S \in \mathbb{Z} + \frac{1}{2} \quad . \end{cases} \tag{7.233}$$

Thus for $K \to \infty$ we effectively recover the $S = 0$ and $S = \frac{1}{2}$ cases for $S \in \mathbb{Z}$ and $S \in \mathbb{Z} + \frac{1}{2}$, respectively. *What happens for $K \to -\infty$?*

(iii) Classical spin $\boldsymbol{S} = S\hat{\boldsymbol{n}}$ with $\hat{h}(\boldsymbol{S}) = 0$ and $\hat{\boldsymbol{n}}$ an $N$-component unit vector: We take the single site density matrix to be $\varrho_i(\boldsymbol{S}_i) = z^{-1}\exp(\mathfrak{h}_i \cdot \boldsymbol{S}_i/k_{\rm B}T)$. The single site field-free partition function is then

$$\zeta(\mathfrak{h}, T) = \int \frac{d\hat{\boldsymbol{n}}}{\Omega_N}\, \exp(\mathfrak{h}\cdot\boldsymbol{S}/k_{\rm B}T) = 1 + \frac{S^2\mathfrak{h}^2}{2N(k_{\rm B}T)^2} + \mathcal{O}(\mathfrak{h}^4) \quad , \tag{7.234}$$

with $\Omega_N = 2\pi^{N/2}/\Gamma(N/2)$ the total solid angle, and therefore

$$\boldsymbol{m} = k_{\rm B}T\,\frac{\partial \log\zeta}{\partial\mathfrak{h}} = \frac{S^2\mathfrak{h}}{2Nk_{\rm B}T} + \mathcal{O}(\mathfrak{h}^3) \quad , \tag{7.235}$$

from which we read off $\chi^0(T) = S^2/Nk_{\rm B}T$. Note that this agrees in the classical ($S \to \infty$) limit, for $N = 3$, with our previous result.

### 7.7.4 Magnetic ordering

Recall our mean field expression for the susceptibility $\hat{\chi}_{\mu\nu}(\boldsymbol{q}, T)$,

$$\hat{\chi}_{\mu\nu}(\boldsymbol{q}, T) = \chi^0(T)\,\Big(\mathbb{1} - \chi^0(T)\,\hat{\mathbb{J}}(\boldsymbol{q})\Big)^{-1}_{\mu\nu} \quad , \tag{7.236}$$

where $\chi^0(T)$ is the local susceptibility for the noninteracting model with all couplings $J_{ij} = 0$. Physically, $\chi^0(T) \to 0$ as $T \to \infty$ and therefore $\hat{\chi}_{\mu\nu}(\boldsymbol{q}, T) \sim \chi^0(T)\,\delta_{\mu\nu}$ at high temperatures.

Consider the case where there is isotropy in spin space. The susceptibility $\chi(\boldsymbol{q}, T)$ then diverges when $\chi^0(T)\hat{J}(\boldsymbol{q}) = 1$, where $\chi^0(T)$ is the local susceptibility. As we know, such a divergence heralds the onset of a *phase transition* where there is a *spontaneous magnetization* in the ordered (*i.e.* low temperature) phase. Typically this happens at a particular wavevector $\boldsymbol{Q}$, or a set of symmetry related wavevectors $\{\boldsymbol{Q}_1, \boldsymbol{Q}_2, \ldots\}$. The ordering wavevector is that value of $\boldsymbol{q}$ which results in a *maximum* of $\hat{J}(\boldsymbol{q})$: $\max_{\boldsymbol{q}}\big\{\hat{J}(\boldsymbol{q})\big\} = \hat{J}(\boldsymbol{Q})$. The susceptibility, for isotropic systems, can be written

$$\hat{\chi}(\boldsymbol{q}) = \frac{\chi^0(T)}{\big[1 - \chi^0(T)\,\hat{J}(\boldsymbol{Q})\big] + \chi^0(T)\big[\hat{J}(\boldsymbol{Q}) - \hat{J}(\boldsymbol{q})\big]} \quad . \tag{7.237}$$

The critical temperature $T_{\rm c}$ is determined by the relation

$$\chi^0(T_{\rm c})\,\hat{J}(\boldsymbol{Q}) = 1. \tag{7.238}$$

Expanding about $T = T_{\rm c}$ and about $\boldsymbol{q} = \boldsymbol{Q}$, where

$$\hat{J}(\boldsymbol{q}) = \hat{J}(\boldsymbol{Q})\Big\{1 - (\boldsymbol{q}-\boldsymbol{Q})^2 R_*^2 + \ldots\Big\} \quad , \tag{7.239}$$

we have

$$\hat{\chi}(\boldsymbol{q}) \approx \frac{\chi^0(T_{\rm c})/R_*^2}{\xi^{-2}(T) + (\boldsymbol{q}-\boldsymbol{Q})^2} \quad , \tag{7.240}$$

where

$$\xi^{-2}(T) = -\frac{\chi^{0\prime}(T_{\rm c})}{\chi^0(T_{\rm c})}\cdot R_*^{-2}\cdot(T - T_{\rm c}) \quad . \tag{7.241}$$

Thus, $\xi(T) \propto (T - T_{\rm c})^{-1/2}$. The real space susceptibility $\chi(\boldsymbol{R}, T)$ oscillates with wavevector $\boldsymbol{Q}$ and decays on the scale of the correlation length $\xi(T)$.

- *Ferromagnet:* $J_{ij} = +J > 0$ if $i$ and $j$ are nearest neighbors; otherwise $J_{ij} = 0$. On a hypercubic lattice ($d$ dimensions, $2d$ nearest neighbor vectors $\boldsymbol{\delta} = \pm a\,\hat{\mathbf{e}}_\nu$), we then have

$$\hat{J}(\boldsymbol{q}) = J\sum_{\boldsymbol{\delta}} e^{-i\boldsymbol{q}\cdot\boldsymbol{\delta}} = 2J\big\{\cos(q_1 a) + \cos(q_2 a) + \ldots + \cos(q_d a)\big\} \quad . \tag{7.242}$$

The ordering wavevector is $\boldsymbol{Q} = 0$, and $\hat{J}(\boldsymbol{Q}) = 2dJ$. For the spin-$S$ Heisenberg model, then, $k_{\rm B}T_{\rm C} = \frac{2}{3}d\,S(S+1)\,J$, and the susceptibility is

$$\hat{\chi}(\boldsymbol{q}, T) = \frac{\frac{1}{3}S(S+1)/k_{\rm B}}{(T - T_{\rm C}) + T_{\rm C}\,d^{-1}\sum_{\nu=1}^{d}\big[1 - \cos(q_\nu a)\big]} \quad . \tag{7.243}$$

The uniform susceptibility $\chi = \hat{\chi}(\boldsymbol{q} = 0)$ is then

$$\chi(T) = \frac{S(S+1)}{3k_{\rm B}(T - T_{\rm C})} \quad , \tag{7.244}$$

where $T_{\rm C}$ is the *Curie temperature*.

Ferromagnetic insulators: ferrites, EuO, TDAE-$C_{60}$.

- *Antiferromagnet:* $J_{ij} = -J < 0$ if $i$ and $j$ are nearest neighbors; otherwise $J_{ij} = 0$. On a hypercubic lattice ($d$ dimensions, $2d$ nearest neighbors), we then have

$$\hat{J}(\boldsymbol{q}) = -J\sum_{\boldsymbol{\delta}} e^{-i\boldsymbol{q}\cdot\boldsymbol{\delta}} = -2J\big\{\cos(q_1 a) + \cos(q_2 a) + \ldots + \cos(q_d a)\big\} \quad . \tag{7.245}$$

The ordering wavevector is $\boldsymbol{Q} = (\pi/a, \ldots, \pi/a)$, at the zone corner, where $\hat{J}(\boldsymbol{Q}) = 2dJ$. For the spin-$S$ Heisenberg model, then, $k_{\rm B}T_{\rm N} = \frac{2}{3}d\,S(S+1)J$, and the susceptibility is

$$\hat{\chi}(\boldsymbol{q}, T) = \frac{S(S+1)/3k_{\rm B}}{(T - T_{\rm N}) + T_{\rm N}\,d^{-1}\sum_{\nu=1}^{d}\big[1 + \cos(q_\nu a)\big]} \quad . \tag{7.246}$$

The uniform susceptibility $\chi_{\rm unif}(T) \equiv \hat{\chi}(\boldsymbol{q} = 0, T)$ is then

$$\chi(\boldsymbol{q} = 0, T) = \frac{S(S+1)}{3k_{\rm B}(T + T_{\rm N})} \quad , \tag{7.247}$$

where $T_{\rm N}$ is the *Néel temperature*. Note that $\chi(\boldsymbol{q} = 0, T)$ *does not diverge*. Indeed, plotting the inverse uniform susceptibility $\chi^{-1}((\boldsymbol{q} = 0, T)$ *versus* $T$, one obtains an intercept along the $T$-axis at $T = -T_{\rm N}$. This is one crude way of estimating the Néel temperature. What *does* diverge is the *staggered susceptibility* $\chi_{\rm stag}(T) \equiv \hat{\chi}(\boldsymbol{Q}, T)$, *i.e.* the susceptibility at the ordering wavevector:

$$\chi(\boldsymbol{Q}, T) = \frac{S(S+1)}{3k_{\rm B}(T - T_{\rm N})} \quad . \tag{7.248}$$

Antiferromagnetic insulators: MnO, CoO, FeO, NiO, $La_2CuO_4$.

- *Frustrated Antiferromagnet:* On the triangular lattice, the antiferromagnetic state is frustrated. What does mean field theory predict? We begin by writing primitive direct lattice vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ and primitive reciprocal lattice vectors $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$, *viz.*

$$\begin{aligned} \boldsymbol{a}_1 &= a\,(1,0) & \boldsymbol{b}_1 &= \frac{4\pi}{a\sqrt{3}}\,\big(\tfrac{\sqrt{3}}{2}, -\tfrac{1}{2}\big) \\ \boldsymbol{a}_2 &= a\,\big(\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}\big) & \boldsymbol{b}_2 &= \frac{4\pi}{a\sqrt{3}}\,(0,1) \quad , \end{aligned} \tag{7.249}$$

where $a$ is the lattice constant. Note that $\boldsymbol{a}_i \cdot \boldsymbol{b}_j = 2\pi\,\delta_{ij}$: The six nearest neighbor vectors are then

$$\boldsymbol{\delta} \in \Big\{ \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_2 - \boldsymbol{a}_1, -\boldsymbol{a}_1, -\boldsymbol{a}_2, \boldsymbol{a}_1 - \boldsymbol{a}_2 \Big\} \quad , \tag{7.250}$$

and writing $\boldsymbol{q} \equiv \theta_1 \boldsymbol{b}_1/2\pi + \theta_2\,\boldsymbol{b}_2/2\pi$, we find

$$\hat{J}(\boldsymbol{q}) = -2J\big\{\cos(\theta_1) + \cos(\theta_2) + \cos(\theta_1 - \theta_2)\big\} \quad . \tag{7.251}$$

We suspect that this should be maximized somewhere along the perimeter of the Brillouin zone. The face center lies at $(\theta_1, \theta_2) = 2\pi(\frac{1}{2}, \frac{1}{2})$, where $\hat{J}(\boldsymbol{q}) = +2J$. However, an even greater value is obtained either of the two inequivalent zone corners, $(\theta_1, \theta_2) = 2\pi(\frac{2}{3}, \frac{1}{3})$ and $(\theta_1, \theta_2) = 2\pi(\frac{1}{3}, \frac{2}{3})$, where $\hat{J}(\boldsymbol{q}) = +3J$. Each of these corresponds to a *tripartite* division of the triangular lattice into three $\sqrt{3} \times \sqrt{3}$ triangular sublattices.

- *Helimagnet:* Consider a cubic lattice system with mixed ferromagnetic and antiferromagnetic interactions:

$$J_{ij} = \begin{cases} +J_1 > 0 & \text{6 nearest neighbors} \\ -J_2 < 0 & \text{12 next-nearest neighbors} \\ 0 & \text{otherwise} \quad . \end{cases} \tag{7.252}$$

Then

$$\begin{aligned} \hat{J}(\boldsymbol{q}) = 2J_1\big[\,&\cos(q_x a) + \cos(q_y a) + \cos(q_z a)\big] \\ &- 4J_2\big[\,\cos(q_x a)\cos(q_y a) + \cos(q_x a)\cos(q_z a) + \cos(q_y a)\cos(q_z a)\big] \quad . \end{aligned} \tag{7.253}$$

The ordering wavevector is then

$$\boldsymbol{Q} = \begin{cases} a^{-1}\cos^{-1}(J_1/4J_2)\,(\pm\hat{\boldsymbol{x}} \pm \hat{\boldsymbol{y}} \pm \hat{\boldsymbol{z}}) & \text{if } J_1 < 4J_2 \\ 0 & \text{if } J_1 \geq 4J_2 \quad . \end{cases} \tag{7.254}$$

Thus, for $J_1 < 4J_2$ the order is *incommensurate* with the lattice. Note also that in this incommensurate regime there are eight symmetry-related ordering wavevectors $\boldsymbol{Q} = Q_0\,(\pm\hat{\boldsymbol{x}} \pm \hat{\boldsymbol{y}} \pm \hat{\boldsymbol{z}})/\sqrt{3}$, with $Q_0 = \sqrt{3}\,a^{-1}\cos^{-1}(J_1/4J_2)$, which point from the center of a cube toward each of its eight vertices. The maximum value of $\hat{J}(\boldsymbol{q})$ is

$$\hat{J}(\boldsymbol{Q}) = \begin{cases} 3J_1^2/4J_2 & \text{if } J_1 < 4J_2 \\ 6(J_1 - 2J_2) & \text{if } J_1 \geq 4J_2 \quad , \end{cases} \tag{7.255}$$

hence incommensurate order sets in at $T_{\rm I} = S(S+1)J_1^2/4k_{\rm B}J_2$. The uniform susceptibility is

$$\hat{\chi}(\boldsymbol{q}=0,T) = \frac{S(S+1)/3k_{\rm B}}{T - 8\,T_{\rm I}\,\frac{J_2}{J_1}\Big(1-\frac{2J_2}{J_1}\Big)} \quad . \tag{7.256}$$

Thus,

$$\chi(\boldsymbol{q}=0,T) \simeq \begin{cases} \mathcal{C}/(T+T^*) \quad , \qquad 0 < J_1 < 2J_2 \text{ (AFM-like)} \\ \\ \mathcal{C}/(T-T^*) \quad , \quad 2J_2 < J_1 < 4J_2 \text{ (FM-like)} \quad . \end{cases} \tag{7.257}$$

## 7.8 Global Symmetries

### 7.8.1 Symmetries and symmetry groups

Interacting systems can be broadly classified according to their *global symmetry group*. Consider the following five examples:

$$\hat{H}_{\rm Ising} = -\frac{1}{2}\sum_{i,j=1}^{N} J_{ij}\,\sigma_i\,\sigma_j \quad , \quad \sigma_i \in \{-1,+1\} \quad , \tag{7.258}$$

which is the Ising model,

$$\hat{H}_{p-{\rm clock}} = -\frac{1}{2}\sum_{i,j=1}^{N} J_{ij}\,\cos\left(\frac{2\pi(l_i-l_j)}{p}\right) \quad , \quad l_i \in \{1,2,\ldots,p\} \quad , \tag{7.259}$$

which is the $p$-state clock model,

$$\hat{H}_{q-{\rm Potts}} = -\frac{1}{2}\sum_{i,j=1}^{N} J_{ij}\,\delta_{\sigma_i,\sigma_j} \quad , \quad \sigma_i \in \{1,2,\ldots,q\} \quad , \tag{7.260}$$

which is the $q$-state Potts model,

$$\hat{H}_{XY} = -\frac{1}{2}\sum_{i,j=1}^{N} J_{ij}\,\cos(\phi_i-\phi_j) \quad , \quad \phi_i \in \big[0,2\pi\big] \quad , \tag{7.261}$$

which is the XY model (also called the $\mathsf{O}(2)$ model), and

$$\hat{H}_{\mathsf{O}(N)} = -\frac{1}{2}\sum_{i,j=1}^{N} J_{ij}\,\hat{\boldsymbol{n}}_i\cdot\hat{\boldsymbol{n}}_j \quad , \quad \hat{\boldsymbol{n}}_i \in S^{N-1} \quad , \tag{7.262}$$

which is the $\mathsf{O}(N)$ model.

The Ising Hamiltonian is left invariant by the global symmetry group $\mathbb{Z}_2$. This group has two elements, which we write as $E$ and $r$. $E$ is the identity operation, *i.e.* $E\,\sigma_i = \sigma_i$ for all $i$. The $r$ operation reverses the

spins: $r\,\sigma_i = -\sigma_i$ . By simultaneously reversing *all* the spins $\sigma_i \to -\sigma_i$, the interactions remain invariant. Note that $r^2 = E$.

The degrees of freedom of the *p-state clock model* are integer variables $l_i$ each of which ranges from 1 to $p$. The Hamiltonian is invariant under the discrete group $\mathbb{Z}_p$, whose $p$ elements are generated by the single operation $r$, where

$$r\,l_i = \begin{cases} l_i + 1 & \text{if} \quad l_i \in \{1, 2, \ldots, p-1\} \\ 1 & \text{if} \quad l_i = p \quad . \end{cases} \tag{7.263}$$

Think of a clock with one hand and $p$ 'hour' markings consecutively spaced by an angle $2\pi/p$. In each site $i$, a hand points to one of the $p$ hour marks; this determines $l_i$. The operation $r$ simply advances *all* the hours by one tick, with hour $p$ advancing to hour 1, just as 23:00 military time is followed one hour later by 00:00. The interaction $\cos\big(2\pi(l_i - l_j)/p\big)$ is invariant under such an operation. The $p$ elements of the group $\mathbb{Z}_p$ are then $\{E,\, r,\, r^2,\, \ldots\,,\, r^{p-1}\}$.

We've already met up with the $q$-state Potts model, where each site supports a 'spin' $\sigma_i$ which can be in any of $q$ possible states, which we may label by integers $\{1\,,\,\ldots\,,\,q\}$. The energy of two interacting sites $i$ and $j$ is $-J_{ij}$ if $\sigma_i = \sigma_j$ and zero otherwise. This energy function is invariant under global operations of the *symmetric group on q characters*, $S_q$ , which is the group of permutations of the sequence $\{1\,,\,2\,,\,3\,,\,\ldots\,,\,q\}$. The group $S_q$ has $q!$ elements. Note the difference between a $\mathbb{Z}_q$ symmetry and an $S_q$ symmetry. In the former case, the Hamiltonian is invariant only under the $q$-element cyclic permutations, *e.g.*

$$r \equiv \begin{pmatrix} 1\,2\cdots q-1\;q \\ 2\,3\cdots \quad q \quad 1 \end{pmatrix}$$

and its powers $r^l$ with $l = 0, \ldots, q-1$.

All these models – the Ising, $p$-state clock, and $q$-state Potts models – possess a global symmetry group which is *discrete*. That is, each of the symmetry groups $\mathbb{Z}_2$, $\mathbb{Z}_p$, and $S_q$ is a discrete group, with a finite number of elements. The $XY$ Hamiltonian $\hat{H}_{XY}$ on the other hand is invariant under a *continuous* group of transformations $\phi_i \to \phi_i + \alpha$, where $\phi_i$ is the angle variable on site $i$. More to the point, we could write the interaction term $\cos(\phi_i - \phi_j)$ as $\frac{1}{2}\big(z_i^* z_j + z_i z_j^*\big)$, where $z_i = e^{i\phi_i}$ is a phase which lives on the unit circle, and $z_i^*$ is the complex conjugate of $z_i$. The model is then invariant under the global transformation $z_i \to e^{i\alpha} z_i$. The phases $e^{i\alpha}$ form a group under multiplication, called $\mathsf{U}(1)$, which is the same as $\mathsf{SO}(2)$. Equivalently, we could write the interaction as $\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j$, where $\hat{\boldsymbol{n}}_i = (\cos\phi_i\,,\,\sin\phi_i)$, which explains the $\mathsf{O}(2) = \mathsf{SO}(2) \rtimes \mathbb{Z}_2$ symmetry, since the symmetry operations are global rotations in the plane plus reflections, which is to say the two-dimensional orthogonal group. This last representation generalizes nicely to unit vectors in $n$ dimensions, where $\hat{\boldsymbol{n}} = (n^1\,,\,n^2\,,\,\ldots\,,\,n^N)$. with $\hat{\boldsymbol{n}}^2 = 1$. The dot product $\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j$ is then invariant under global rotations in this $n$-dimensional space, which is the group $\mathsf{O}(N)$.

### 7.8.2 Domain walls and lower critical dimension

Depending on whether the global symmetry group of a model is discrete or continuous, there exists a *lower critical dimension* $d_\ell$ at or below which no phase transition may take place at finite temperature. That is, for $d \le d_\ell$, the critical temperature is $T_{\rm c} = 0$. Owing to its neglect of fluctuations, mean field

theory generally *overestimates* the value of $T_{\rm c}$ because it overestimates the stability of the ordered phase[31]. Indeed, there are many examples where mean field theory predicts a finite $T_{\rm c}$ when the actual critical temperature is $T_{\rm c} = 0$. This happens whenever $d \le d_\ell$.

Let's test the stability of the ordered (ferromagnetic) state of the one-dimensional Ising model at low temperatures. We consider order-destroying *domain wall* excitations which interpolate between regions of degenerate, symmetry-related ordered phase, *i.e.* $\uparrow\uparrow\uparrow\uparrow\uparrow$ and $\downarrow\downarrow\downarrow\downarrow\downarrow$[32]. For a system with a discrete symmetry at low temperatures, the domain wall is abrupt, on the scale of a single lattice spacing. If the exchange energy is $J$, then the energy of a single domain wall is $2J$, since a link of energy $-J$ is replaced with one of energy $+J$. However, in a system of length $L$ (with *anti*periodic boundary conditions), there are $L$ possible locations for the domain wall, hence its entropy is $k_{\rm B} \log L$. For a system with $M$ domain walls, the free energy is[33]

$$\begin{aligned} F &= 2MJ - k_{\rm B}T \log \binom{L}{M} \\ &= N \cdot \left\{ 2Jx + k_{\rm B}T \Big[ x \log x + (1-x) \log(1-x) \Big] \right\} \quad , \end{aligned} \tag{7.264}$$

where $x = M/L$ is the density of domain walls, and where we have used Stirling's approximation for $k!$ when $k$ is large. Extremizing with respect to $x$, we find

$$\frac{x}{1-x} = e^{-2J/k_{\rm B}T} \qquad \Longrightarrow \qquad x = \frac{1}{e^{2J/k_{\rm B}T} + 1} \quad . \tag{7.265}$$

The average distance between domain walls is $x^{-1}$, which is finite for finite $T$. Thus, the thermodynamic state of the system is *disordered*, with no net average magnetization.

Consider next an Ising domain wall in $d$ dimensions. Let the linear dimension of the system be $L \cdot a$, where $L$ is a real number and $a$ is the lattice constant. Then the energy of a single domain wall which partitions the entire system is $2J \cdot L^{d-1}$. The domain wall entropy is difficult to compute, because the wall can fluctuate significantly, but for a single domain wall we have $S \gtrsim k_{\rm B} \log L$. Thus, the free energy $F = 2JL^{d-1} - k_{\rm B}T \log L$ is dominated by the energy term if $d > 1$, suggesting that the system *may* be ordered. We can do a slightly better job in $d = 2$ by writing

$$Z \approx \exp\left( L^d \sum_P N_P \, e^{-2PJ/k_{\rm B}T} \right) \quad , \tag{7.266}$$

[31] It is a simple matter to concoct models for which the mean field transition temperature *underestimates* the actual critical temperature. Consider for example an Ising model with interaction $u(\sigma,\sigma') = -\epsilon^{-1}\log(1+\epsilon\sigma\sigma')$, where the spins take values $\sigma,\sigma' = \pm 1$, and where $0 < \epsilon < 1$. If we write $\sigma = \langle\sigma\rangle + \delta\sigma$ at each site and neglect terms quadratic in fluctuations, the resulting mean field Hamiltonian is equivalent to a set of decoupled spins in an external field $h = zm/(1+\epsilon m^2)$. The mean field transition temperature is $T_{\rm c}^{\rm MF} = z$, the lattice coordination number, independent of $\epsilon$. On the other hand, we may also write $u(\sigma,\sigma') = u_\epsilon - J_\epsilon\,\sigma\sigma'$, where $u_\epsilon = -\log(1-\epsilon^2)/2\epsilon$ and $J_\epsilon = \epsilon^{-1}\tanh^{-1}(\epsilon)$. On the square lattice, where $z = 4$, one has the exact result $T_{\rm c}(\epsilon) = 2J_\epsilon/\sinh^{-1}(1)$, which diverges as $\epsilon \to 1$, while $T_{\rm c}^{\rm MF} = 4$ remains finite. For $\epsilon > 0.9265$, one has $T_{\rm c}(\epsilon) > T_{\rm c}^{\rm MF}$.

[32] Here and below we reiterate some material from §6.2.3.

[33] With periodic boundary conditions, $M$ must be even. With antiperiodic boundary conditions, $M$ must be odd. The thermodynamic conclusions are the same, however.

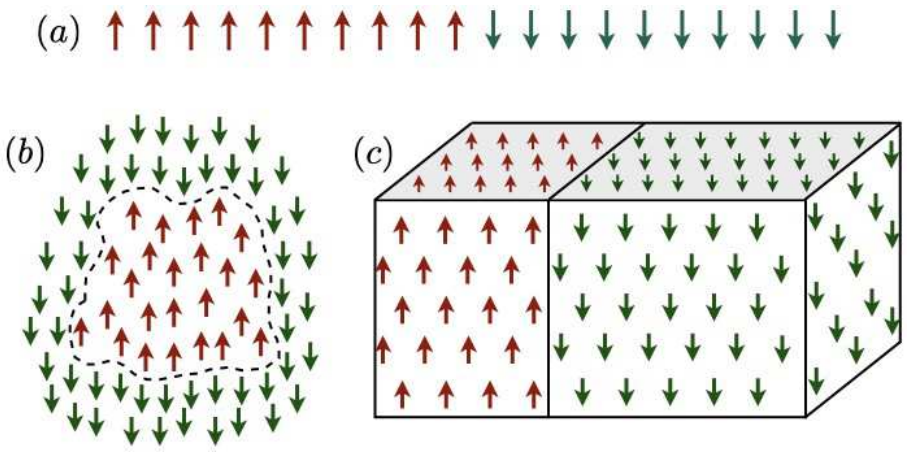



Figure 7.21: Domain walls in the (a) $d = 1$, (b) $d = 2$, and (c) $d = 3$ Ising model.

where the sum is over all closed loops of perimeter $P$, and $N_P$ is the number of such loops. An example of such a loop circumscribing a domain is depicted in the left panel of fig. 7.21. It turns out that

$$N_P \simeq \kappa^P P^{-\theta} \cdot \Big\{1 + \mathcal{O}(P^{-1})\Big\} \quad , \tag{7.267}$$

where $\kappa \approx z - 1$ with $z$ the lattice coordination number, and $\theta$ is some exponent. We can understand the $\kappa^P$ factor in the following way. At each step along the perimeter of the loop, there are $\kappa = z-1$ possible directions to go (since one doesn't backtrack). The fact that the loop must avoid overlapping itself and must return to its original position to be closed leads to the power law term $P^{-\theta}$, which is subleading since $\kappa^P P^{-\theta} = \exp(P \log \kappa - \theta \log P)$ and $P \gg \log P$ for $P \gg 1$. Thus,

$$F \approx -\frac{1}{\beta} L^d \sum_P P^{-\theta}\, e^{(\log\kappa - 2\beta J)P} \quad , \tag{7.268}$$

which diverges if $\log \kappa > 2\beta J$, *i.e.* if $T > 2J/k_{\rm B} \log(z-1)$. We identify this singularity with the phase transition. The high temperature phase involves a proliferation of such loops. The excluded volume effects between the loops, which we have not taken into account, then enter in an essential way so that the sum converges. Thus, we have the following picture:

$$\begin{aligned} &\log\kappa < 2\beta J \ : \text{ large loops suppressed ; ordered phase} \\ &\log\kappa > 2\beta J \ : \text{ large loops proliferate ; disordered phase} \quad . \end{aligned}$$

On the square lattice, we obtain

$$k_{\rm B} T_{\rm c}^{\rm approx} = \frac{2J}{\log 3} = 1.82\, J \qquad , \qquad k_{\rm B} T_{\rm c}^{\rm exact} = \frac{2J}{\sinh^{-1}(1)} = 2.27\, J \quad .$$

The agreement is better than we should reasonably expect from such a crude argument.

*Nota bene* : Beware of arguments which allegedly prove the existence of an ordered phase. Generally speaking, any approximation will *underestimate* the entropy, and thus will overestimate the stability of the putative ordered phase.

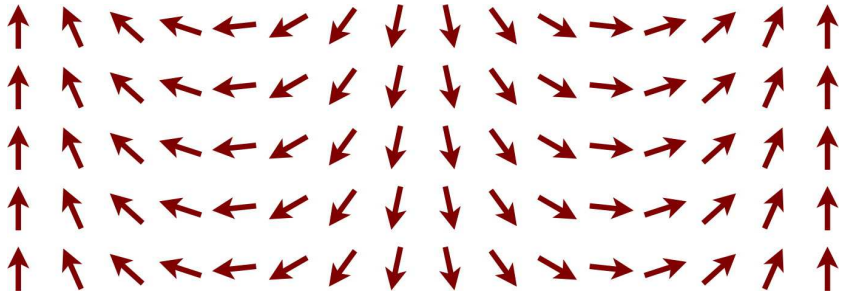


Figure 7.22: A domain wall in an $XY$ ferromagnet.

### 7.8.3 Continuous symmetries

When the global symmetry group is continuous, the domain walls interpolate smoothly between ordered phases. The energy generally involves a stiffness term,

$$E = \tfrac{1}{2}\rho_{\rm s}\int d^dr\ (\boldsymbol{\nabla}\theta)^2 \quad , \tag{7.269}$$

where $\theta(\boldsymbol{r})$ is the angle of a local rotation about a single axis and where $\rho_{\rm s}$ is the *spin stiffness*. Of course, in $\mathsf{O}(n)$ models, the rotations can be with respect to several different axes simultaneously.

In the ordered phase, we have $\theta(\boldsymbol{r}) = \theta_0$, a constant. Now imagine a domain wall in which $\theta(\boldsymbol{r})$ rotates by $2\pi$ across the width of the sample. We write $\theta(\boldsymbol{r}) = 2\pi n x/L$, where $L$ is the linear size of the sample (here with dimensions of length) and $n$ is an integer telling us how many complete twists the order parameter field makes. The domain wall then resembles that in fig. 7.22. The gradient energy is

$$E = \tfrac{1}{2}\rho_{\rm s}\,L^{d-1}\int\limits_0^L dx\left(\frac{2\pi n}{L}\right)^2 = 2\pi^2 n^2\rho_{\rm s}\,L^{d-2} \quad . \tag{7.270}$$

Recall that in the case of discrete symmetry, the domain wall energy scaled as $E \propto L^{d-1}$. Thus, with $S \gtrsim k_{\rm B}\log L$ for a single wall, we see that the entropy term dominates if $d \le 2$, in which case there is no finite temperature phase transition. Thus, the lower critical dimension $d_\ell$ depends on whether the global symmetry is discrete or continuous, with

$$\begin{aligned} \text{discrete global symmetry} \quad &\Longrightarrow \quad d_\ell = 1 \\ \text{continuous global symmetry} \quad &\Longrightarrow \quad d_\ell = 2 \quad . \end{aligned}$$

Note that all along we have assumed local, *short-ranged* interactions. Long-ranged interactions can enhance order and thereby suppress $d_\ell$.

Thus, we expect that for models with discrete symmetries, $d_\ell = 1$ and there is no finite temperature phase transition for $d \le 1$. For models with continuous symmetries, $d_\ell = 2$, and we expect $T_{\rm c} = 0$ for $d \le 2$. In this context we should emphasize that the two-dimensional $XY$ model *does* exhibit a phase transition at finite temperature, called the *Kosterlitz-Thouless* transition. However, this phase transition is *not* associated with the breaking of the continuous global $\mathsf{O}(2)$ symmetry and rather has to do with the unbinding of vortices and antivortices. So there is still no true long-ranged order below the critical temperature $T_{\rm KT}$, even though there is a phase transition!

### 7.8.4 Random systems : Imry-Ma argument

Oftentimes, particularly in condensed matter systems, intrinsic randomness exists due to quenched impurities, grain boundaries, immobile vacancies, *etc.* How does this quenched randomness affect a system's attempt to order at $T = 0$? This question was taken up in a beautiful and brief paper by J. Imry and S.-K. Ma, *Phys. Rev. Lett.* **35**, 1399 (1975). Imry and Ma considered models in which there are short-ranged interactions and a random local field coupling to the local order parameter:

$$\hat{H}_{\rm RFI} = -J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j - \sum_{i=1}^{N} H_i \, \sigma_i \tag{7.271}$$

$$\hat{H}_{\rm RFO(\mathit{N})} = -J \sum_{\langle ij \rangle} \hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j - \sum_{i=1}^{N} H_i^\alpha \, n_i^\alpha \quad , \tag{7.272}$$

where

$$\langle\!\langle \, H_i^\alpha \, \rangle\!\rangle = 0 \qquad , \qquad \langle\!\langle \, H_i^\alpha \, H_j^\beta \, \rangle\!\rangle = \Gamma \, \delta^{\alpha\beta} \, \delta_{ij} \quad , \tag{7.273}$$

where $\langle\!\langle \, \bullet \, \rangle\!\rangle$ denotes a configurational average over the disorder. Imry and Ma reasoned that a system could try to lower its free energy by forming *domains* in which the order parameter takes advantage of local fluctuations in the random field. The size of these domains is assumed to be $L_{\rm d}$, a length scale to be determined. See the sketch in the left panel of fig. 7.23.

There are two contributions to the energy of a given domain: bulk and surface terms. The bulk energy is given by

$$E_{\rm bulk} = -H_{\rm rms} \, (L_{\rm d}/a)^{d/2} \quad , \tag{7.274}$$

where $a$ is the lattice spacing. This is because when we add together $(L_{\rm d}/a)^d$ random fields, the magnitude of the result is proportional to the square root of the number of terms, *i.e.* to $(L_{\rm d}/a)^{d/2}$. The quantity $H_{\rm rms} = \sqrt{\Gamma}$ is the root-mean-square fluctuation in the random field at a given site. The surface energy is

$$E_{\rm surface} \propto \begin{cases} J \, (L_{\rm d}/a)^{d-1} & \text{(discrete symmetry)} \\ J \, (L_{\rm d}/a)^{d-2} & \text{(continuous symmetry)} \end{cases} . \tag{7.275}$$

We compute the critical dimension $d_{\rm c}$ by balancing the bulk and surface energies,

$$\begin{aligned} d - 1 = \tfrac{1}{2} d \quad &\Longrightarrow \quad d_{\rm c} = 2 \quad \text{(discrete)} \\ d - 2 = \tfrac{1}{2} d \quad &\Longrightarrow \quad d_{\rm c} = 4 \quad \text{(continuous)} \end{aligned} .$$

The total free energy is $F = (V/L_{\rm d}^d) \cdot \Delta E$, where $\Delta E = E_{\rm bulk} + E_{\rm surf}$. The free energy per unit cell is

$$f = \frac{F}{V/a^d} \approx J \left( \frac{a}{L_{\rm d}} \right)^{d_{\rm c}/2} - H_{\rm rms} \left( \frac{a}{L_{\rm d}} \right)^{d/2} \quad . \tag{7.276}$$

If $d < d_{\rm c}$, the surface term dominates for small $L_{\rm d}$ and the bulk term dominates for large $L_{\rm d}$ There is a global minimum at

$$\frac{L_{\rm d}}{a} = \left( \frac{d_{\rm c}}{d} \cdot \frac{J}{H_{\rm rms}} \right)^{2/(d_{\rm c} - d)} \quad . \tag{7.277}$$

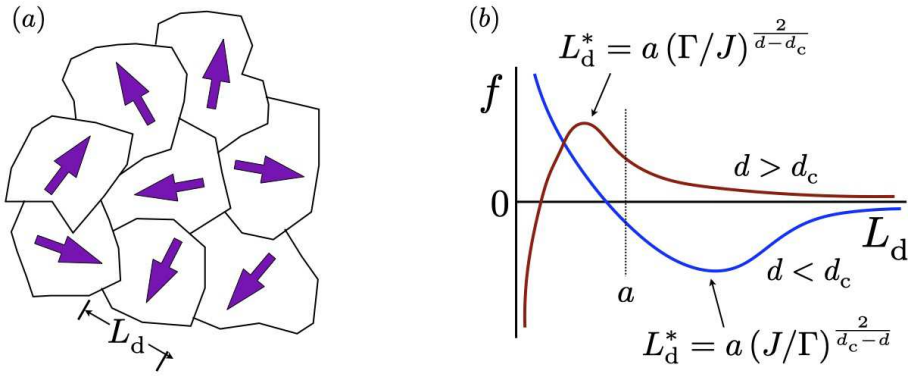



Figure 7.23: (a) Imry-Ma domains for an O(2) model. The arrows point in the direction of the local order parameter field $\langle \hat{\boldsymbol{n}}(\boldsymbol{r}) \rangle$. (b) Free energy density as a function of domain size $L_{\rm d}$. Keep in mind that the minimum possible value for $L_{\rm d}$ is the lattice spacing $a$.

For $d > d_{\rm c}$, the relative dominance of the bulk and surface terms is reversed, and there is a global maximum at this value of $L_{\rm d}$.

Sketches of the free energy $f(L_{\rm d})$ in both cases are provided in the right panel of fig. 7.23. We must keep in mind that the domain size $L_{\rm d}$ cannot become smaller than the lattice spacing $a$. Hence we should draw a vertical line on the graph at $L_d = a$ and discard the portion $L_d < a$ as unphysical. For $d < d_{\rm c}$, we see that the state with $L_d = \infty$, *i.e.* the ordered state, is never the state of lowest free energy. *In dimensions $d < d_{\rm c}$, the ordered state is always unstable to domain formation in the presence of a random field.*

For $d > d_{\rm c}$, there are two possibilities, depending on the relative size of $J$ and $H_{\rm rms}$. We can see this by evaluating $f(L_d = a) = J - H_{\rm rms}$ and $f(L_{\rm d} = \infty) = 0$. Thus, if $J > H_{\rm rms}$, the minimum energy state occurs for $L_{\rm d} = \infty$. In this case, the system has an ordered ground state, and we expect a finite temperature transition to a disordered state at some critical temperature $T_{\rm c} > 0$. If, on the other hand, $J < H_{\rm rms}$, then the fluctuations in $H$ overwhelm the exchange energy at $T = 0$, and the ground state is disordered down to the very smallest length scale (*i.e.* the lattice spacing $a$).

Please read the essay, Memories of Shang-Keng Ma.

## 7.9 Ginzburg-Landau Theory

### 7.9.1 Ginzburg-Landau free energy

Including gradient terms in the free energy, we write[34]

$$F\big[m(\boldsymbol{x})\,,\,h(\boldsymbol{x})\big] = \int d^dx \left\{ f_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 + \tfrac{1}{6}cm^6 \;\; - hm + \tfrac{1}{2}\kappa\,(\boldsymbol{\nabla} m)^2 + \ldots \right\} \quad . \tag{7.278}$$

In principle, any term which does not violate the appropriate global symmetry will turn up in such an expansion of the free energy, with some coefficient. Examples include $hm^3$ (both $m$ and $h$ are odd

[34]For a systematic derivation of the Ginzburg-Landau free energy, see the appendix in §7.13.

under time reversal), $m^2(\boldsymbol{\nabla} m)^2$, *etc.* We now ask: what function $m(\boldsymbol{x})$ extremizes the free energy functional $F\big[m(\boldsymbol{x})\,,\,h(\boldsymbol{x})\big]$? The answer is that $m(\boldsymbol{x})$ must satisfy the corresponding Euler-Lagrange equation, which for the above functional is

$$a\,m + b\,m^3 + c\,m^5 - h - \kappa\,\nabla^2 m = 0 \quad . \tag{7.279}$$

If $a > 0$ and $h$ is small (we assume $b > 0$ and $c > 0$), we may neglect the $m^3$ and $m^5$ terms. The above equation then simplifies to $\big(a - \kappa\,\nabla^2\big)m = h$. We solve by Fourier transform, obtaining

$$\hat{m}(\boldsymbol{q}) = \frac{\hat{h}(\boldsymbol{q})}{a + \kappa \boldsymbol{q}^2} \qquad \Rightarrow \qquad \hat{\chi}(\boldsymbol{q}) = \frac{\partial \hat{m}(\boldsymbol{q})}{\partial \hat{h}(\boldsymbol{q})} = \frac{1}{a + \kappa \boldsymbol{q}^2} \quad . \tag{7.280}$$

Thus,

$$m(\boldsymbol{x}) = \int d^d x'\,\chi(\boldsymbol{x} - \boldsymbol{x}')\,h(\boldsymbol{x}') \quad , \tag{7.281}$$

where

$$\chi(\boldsymbol{x} - \boldsymbol{x}') = \kappa^{-1} \int \frac{d^d q}{(2\pi)^d}\,\frac{e^{i\boldsymbol{q}\cdot(\boldsymbol{x}-\boldsymbol{x}')}}{\boldsymbol{q}^2 + \xi^{-2}} \quad , \tag{7.282}$$

where the correlation length is $\xi = \sqrt{\kappa/a} \propto (T - T_{\rm c})^{-1/2}$, as before.

If $a < 0$ then there is a spontaneous magnetization, given by the solution to $b m_0^2 + c m_0^4 = |a|$, and we write $m(\boldsymbol{x}) = m_0 + \delta m(\boldsymbol{x})$. Assuming $h$ is weak, we then have two equations

$$\begin{aligned} a + b\,m_0^2 + c\,m_0^4 &= 0 \\ (a + 3b\,m_0^2 + 5c\,m_0^4 - \kappa\,\nabla^2)\,\delta m &= h \quad . \end{aligned} \tag{7.283}$$

Assuming $|a|$ is small, we have $m_0^2 = |a|/b$ and

$$\delta\hat{m}(\boldsymbol{q}) = \frac{\hat{h}(\boldsymbol{q})}{2|a| + \kappa \boldsymbol{q}^2} \quad , \tag{7.284}$$

### 7.9.2 Domain wall profile

A particularly interesting application of Ginzburg-Landau theory is its application toward modeling the spatial profile of defects such as vortices and domain walls. Consider, for example, the case of Ising ($\mathbb{Z}_2$) symmetry with $h = 0$. We expand the free energy density to order $m^4$:

$$F\big[m(\boldsymbol{x})\big] = \int d^d x \left\{ f_0 + \tfrac{1}{2} a m^2 + \tfrac{1}{4} b m^4 + \tfrac{1}{2}\kappa\,(\boldsymbol{\nabla} m)^2 \right\} \quad . \tag{7.285}$$

We assume $a < 0$, corresponding to $T < T_{\rm c}$. Consider now a domain wall, where $m(x \to -\infty) = -m_0$ and $m(x \to +\infty) = +m_0$, where $m_0$ is the equilibrium magnetization, which we obtain from the Euler-Lagrange equation,

$$a\,m + b\,m^3 - \kappa\,\nabla^2 m = 0 \quad , \tag{7.286}$$

assuming a uniform solution where $\boldsymbol{\nabla} m = 0$. This gives $m_0 = \sqrt{|a|/b}$. It is useful to scale $m(\boldsymbol{x})$ by $m_0$, writing $m(\boldsymbol{x}) = m_0\,\phi(\boldsymbol{x})$. The scaled order parameter function $\phi(\boldsymbol{x})$ interpolates between $\phi(-\infty) = -1$ and $\phi(+\infty) = 1$.

Thus, we have

$$\xi^2\,\boldsymbol{\nabla}^2\phi = -\phi + \phi^3 \quad , \tag{7.287}$$

where $\xi = \sqrt{\kappa/|a|}$. We assume $\phi(\boldsymbol{x}) = \phi(x_1)$ is only a function of the first coordinate. Then the Euler-Lagrange equation becomes

$$\xi^2\,\frac{d^2\phi}{dx_1^2} = -\phi + \phi^3 \equiv -\frac{dU}{d\phi} \quad , \tag{7.288}$$

where $U(\phi) = -\frac{1}{4}\big(\phi^2 - 1\big)^2$. The 'potential' $U(\phi)$ is an inverted double well, with maxima at $\phi = \pm 1$. The equation $\ddot{\phi} = -U'(\phi)$, where dot denotes differentiation with respect to $\zeta$, is simply Newton's second law with time replaced by space. In order to have a stationary solution at $\zeta \to \pm\infty$ where $\phi = \pm 1$, the total energy must be $E = U(\phi = \pm 1) = 0$, where $E = \frac{1}{2}\dot{\phi}^2 + U(\phi)$. This leads to the first order differential equation

$$\sqrt{2}\,\xi\,\frac{d\phi}{dx_1} = 1 - \phi^2 \quad , \tag{7.289}$$

with solution

$$\phi(\boldsymbol{x}) = \tanh\big(x_1/\sqrt{2}\,\xi\big) \qquad \Rightarrow \qquad m(\boldsymbol{x}) = m_0\,\tanh\big(x_1/\sqrt{2}\,\xi\big) \quad . \tag{7.290}$$

Note that the correlation length $\xi$ diverges at the Ising transition.

### 7.9.3 Ginzburg criterion

Let us define $A(T, H, V, N)$ to be the usual (*i.e.* thermodynamic) Helmholtz free energy. Then

$$e^{-\beta A} = \int Dm\; e^{-\beta F[m(\boldsymbol{x})]} \quad , \tag{7.291}$$

where the functional $F[m(\boldsymbol{x})]$ is of the Ginzburg-Landau form, given in eqn. 7.285. The integral above is a *functional integral*. We can give it a more precise meaning by defining its measure in the case of periodic functions $m(\boldsymbol{x})$ confined to a rectangular box. Then we can expand

$$m(\boldsymbol{x}) = \frac{1}{\sqrt{V}}\sum_{\boldsymbol{q}} \hat{m}_{\boldsymbol{q}}\,e^{i\boldsymbol{q}\cdot\boldsymbol{x}} \quad , \tag{7.292}$$

and we define the measure

$$Dm \equiv dm_0 \prod_{\substack{\boldsymbol{q} \\ q_x > 0}} d\,\mathsf{Re}\;\hat{m}_{\boldsymbol{q}}\; d\,\mathsf{Im}\;\hat{m}_{\boldsymbol{q}} \quad . \tag{7.293}$$

Note that the fact that $m(\boldsymbol{x}) \in \mathbb{R}$ means that $\hat{m}_{-\boldsymbol{q}} = \hat{m}^*_{\boldsymbol{q}}$. We'll assume $T > T_{\rm c}$ and $H = 0$ and we'll explore limit $T \to T_{\rm c}^+$ from above to analyze the properties of the critical region close to $T_{\rm c}$. In this limit we can ignore all but the quadratic terms in $m$, and we have

$$e^{-\beta A} = \int Dm\;\exp\left(-\tfrac{1}{2}\beta\sum_{\boldsymbol{q}}(a + \kappa\,\boldsymbol{q}^2)\,|\hat{m}_{\boldsymbol{q}}|^2\right) = \prod_{\boldsymbol{q}}\left(\frac{\pi k_{\rm B}T}{a + \kappa\,\boldsymbol{q}^2}\right)^{1/2} \quad . \tag{7.294}$$

Thus,

$$A = \tfrac{1}{2} k_{\rm B} T \sum_{\boldsymbol{q}} \log\left( \frac{a + \kappa\, \boldsymbol{q}^2}{\pi k_{\rm B} T} \right) \quad . \tag{7.295}$$

Close to the critical point, we write $a(T) = \alpha t$, where $t$ is the dimensionless quantity $t \equiv (T - T_{\rm c})/T_{\rm c}$, known as the *reduced temperature*.

We now compute the heat capacity $C_V = -T\,\partial^2 A/\partial T^2$. We are really only interested in the singular contributions to $C_V$, which means that we're only interested in differentiating with respect to $T$ as it appears in $a(T)$. We divide by $N k_{\rm B}$ where $N$ is the number of unit cells of our system, which we presume is a lattice-based model. Note $N \sim V/\mathsf{a}^d$ where $V$ is the volume and $\mathsf{a}$ the lattice constant. The dimensionless heat capacity per lattice site is then

$$c \equiv \frac{C_V}{N} = \frac{\alpha^2 \mathsf{a}^d}{2\kappa^2} \int^{\Lambda} \frac{d^d q}{(2\pi)^d}\, \frac{1}{(\xi^{-2} + \boldsymbol{q}^2)^2} \quad , \tag{7.296}$$

where $\xi = (\kappa/\alpha t)^{1/2} \propto |t|^{-1/2}$ is the correlation length, and where $\Lambda \sim \mathsf{a}^{-1}$ is an ultraviolet cutoff. We define $R_* \equiv (\kappa/\alpha)^{1/2}$, in which case

$$c = R_*^{-4}\, \mathsf{a}^d\, \xi^{4-d} \cdot \tfrac{1}{2} \int^{\Lambda\xi} \frac{d^d \bar{q}}{(2\pi)^d}\, \frac{1}{(1 + \bar{q}^2)^2} \quad , \tag{7.297}$$

where $\bar{\boldsymbol{q}} \equiv \boldsymbol{q}\xi$. Thus,

$$c(t) \sim \begin{cases} \text{const.} & \text{if } d > 4 \\ -\log t & \text{if } d = 4 \\ t^{\frac{d}{2}-2} & \text{if } d < 4 \quad . \end{cases} \tag{7.298}$$

For $d > 4$, mean field theory is qualitatively accurate, with finite corrections. In dimensions $d \le 4$, the mean field result is overwhelmed by fluctuation contributions as $t \to 0^+$ (*i.e.* as $T \to T_{\rm c}^+$). We see that MFT is sensible provided the fluctuation contributions are small, *i.e.* provided $R_*^{-4}\, \mathsf{a}^d\, \xi^{4-d} \ll 1$, which entails $t \gg t_{\rm G}$, where

$$t_{\rm G} = \left( \frac{\mathsf{a}}{R_*} \right)^{2d/(4-d)} \tag{7.299}$$

is the *Ginzburg reduced temperature*. The criterion for the sufficiency of mean field theory, namely $t \gg t_{\rm G}$, is known as the *Ginzburg criterion*. The region $|t| < t_{\rm G}$ is known as the *critical region*.

In a lattice ferromagnet, as we have seen, $R_* \sim \mathsf{a}$ is on the scale of the lattice spacing itself, hence $t_{\rm G} \sim 1$ and the critical regime is very large. Mean field theory then fails quickly as $T \to T_{\rm c}$. In a (conventional) three-dimensional superconductor, $R_*$ is on the order of the Cooper pair size, and $R_*/\mathsf{a} \sim 10^2 - 10^3$, hence $t_{\rm G} = (a/R_*)^6 \sim 10^{-18} - 10^{-12}$ is negligibly narrow. The mean field theory of the superconducting transition – BCS theory – is then valid essentially all the way to $T = T_{\rm c}$.

## 7.10 Appendix I: Bifurcations

### 7.10.1 $N = 1$ dynamical systems

A dynamical system with an $N = 1$ dimensional phase space is a first order ODE $\dot{u} = g(u)$, where $g(u)$ is in general a nonlinear function of $u$. Thus $\dot{u} > 0$ for values of $u$ where $g(u) > 0$ and $\dot{u} < 0$ for values of $u$ where $g(u) < 0$. For $u = u^*$ where $g(u^*) = 0$ we have a *fixed point*. Linearizing about the fixed point, we write $u = u^* + \eta$, in which case $\dot{\eta} = g'(u^*)\,\eta$. Thus $\eta(t) = \eta(0)\,\exp\big[g'(u^*)\,t\big]$, which says that $|\eta(t)|$ grows exponentially if $g'(u^*) > 0$ and $|\eta(t)|$ shrinks exponentially if $g'(u^*) < 0$. Thus, for $g'(u^*) > 0$ we say that $u^*$ is an *unstable fixed point* (UFP), while for $g'(u^*) < 0$ we say that $u^*$ is a *stable fixed point* (SFP). The linearization about a UFP rapidly fails due to the exponential increase of the distance from $u^*$, while the linearization about a SFP gets ever more accurate as $\eta(t) \to 0$.

### 7.10.2 Saddle-node bifurcation

Generically, $g'(u^*) \neq 0$ at points where $g(u^*) = 0$, since two equations in a single unknown is an overdetermined set. However, consider the function $G(x, \alpha)$, where $\alpha$ is a control parameter. If we demand $G(x, \alpha) = 0$ and $\partial_x G(x, \alpha) = 0$, we have two equations in two unknowns, and in general there will be a zero-dimensional solution set consisting of points $(x_{\rm c}, \alpha_{\rm c})$. The situation is depicted in Fig. 7.24.

Let's expand $G(x, \alpha)$ in the vicinity of such a point $(x_{\rm c}, \alpha_{\rm c})$:

$$\begin{aligned} G(x,\alpha) &= G(x_{\rm c},\alpha_{\rm c}) + \left.\frac{\partial G}{\partial x}\right|_{(x_{\rm c},\alpha_{\rm c})} (x - x_{\rm c}) + \left.\frac{\partial G}{\partial \alpha}\right|_{(x_{\rm c},\alpha_{\rm c})} (\alpha - \alpha_{\rm c}) + \frac{1}{2}\left.\frac{\partial^2 G}{\partial x^2}\right|_{(x_{\rm c},\alpha_{\rm c})} (x - x_{\rm c})^2 \\ &\qquad + \left.\frac{\partial^2 G}{\partial x\,\partial \alpha}\right|_{(x_{\rm c},\alpha_{\rm c})} (x - x_{\rm c})\,(\alpha - \alpha_{\rm c}) + \frac{1}{2}\left.\frac{\partial^2 G}{\partial \alpha^2}\right|_{(x_{\rm c},\alpha_{\rm c})} (\alpha - \alpha_{\rm c})^2 + \ldots \\ &= A\,(\alpha - \alpha_{\rm c}) + B\,(x - x_{\rm c})^2 + \ldots \quad , \end{aligned} \tag{7.300}$$

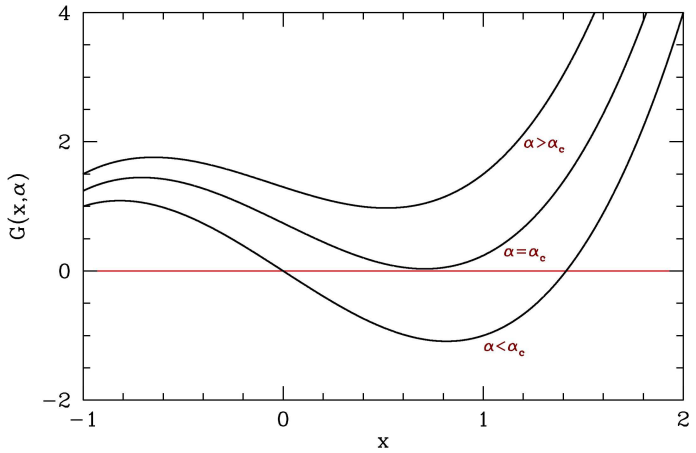



Figure 7.24: Evolution of $G(x, \alpha)$ as a function of the control parameter $\alpha$.

where we keep terms of lowest order in the deviations $\delta x$ and $\delta\alpha$. Note that we can separately change the signs of $A$ and $B$ by redefining $\alpha \to -\alpha$ and/or $x \to -x$, so without loss of generality we may assume both $A$ and $B$ are positive. If we now rescale $u \equiv \sqrt{B/A}\,(x - x_{\rm c})$, $r \equiv \alpha - \alpha_{\rm c}$, and $\tau = \sqrt{AB}\,t$, we have, neglecting the higher order terms, we obtain the 'normal form' of the saddle-node bifurcation,

$$\frac{du}{d\tau} = r + u^2 \quad . \tag{7.301}$$

The evolution of the flow is depicted in Fig. 7.25. For $r < 0$ there are two fixed points – one stable ($u^* = -\sqrt{-r}$) and one unstable ($u = +\sqrt{-r}$). At $r = 0$ these two nodes coalesce and annihilate each other. (The point $u^* = 0$ is half-stable precisely at $r = 0$.) For $r > 0$ there are no longer any fixed points in the vicinity of $u = 0$. In the left panel of Fig. 7.26 we show the flow in the extended $(r, u)$ plane. The unstable and stable nodes annihilate at $r = 0$.

### 7.10.3 Transcritical bifurcation

Another situation which arises frequently is the *transcritical bifurcation*. Consider the equation $\dot{x} = g(x)$ in the vicinity of a fixed point $x^*$.

$$\frac{dx}{dt} = g'(x^*)\,(x - x^*) + \tfrac{1}{2}\,g''(x^*)(x - x^*)^2 + \ldots \quad . \tag{7.302}$$

We rescale $u \equiv \beta\,(x - x^*)$ with $\beta = -\frac{1}{2}\,g''(x^*)$ and define $r \equiv g'(x^*)$ as the control parameter, to obtain, to order $u^2$,

$$\frac{du}{dt} = ru - u^2 \quad . \tag{7.303}$$

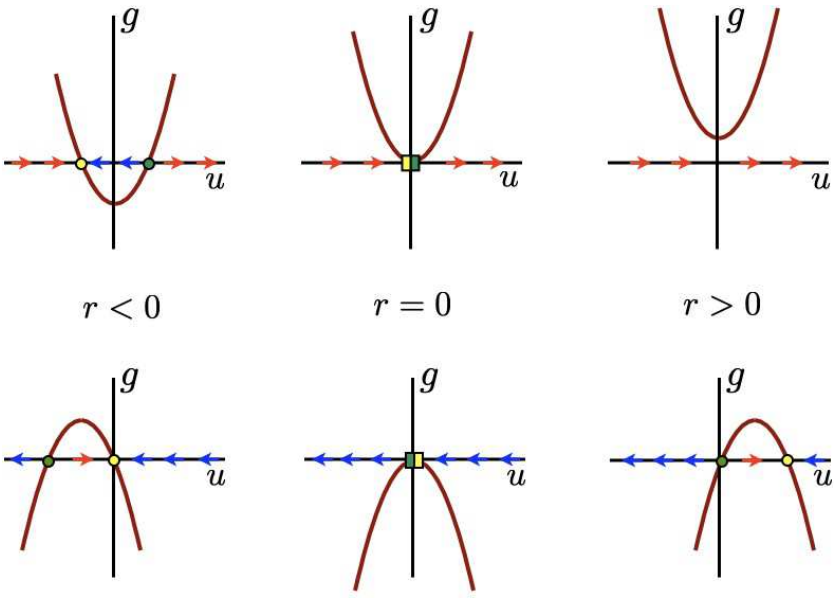



Figure 7.25: Flow diagrams for the saddle-node bifurcation $\dot{u} = r + u^2$ (top) and the transcritical bifurcation $\dot{u} = ru - u^2$ (bottom).

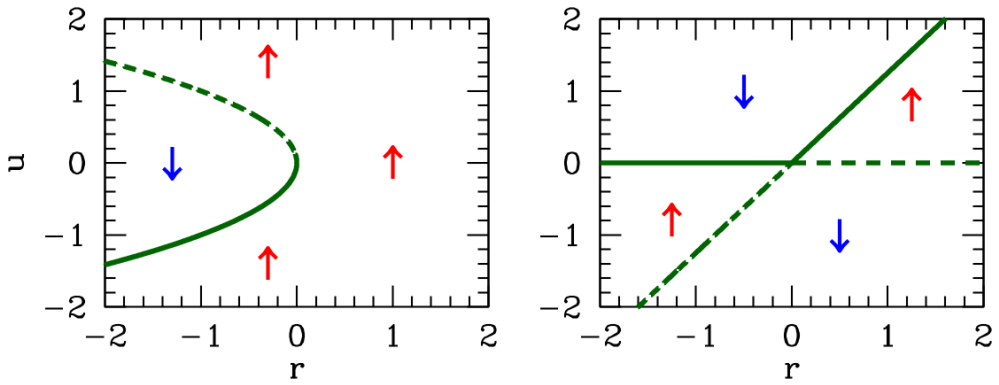



Figure 7.26: Extended phase space $(r, u)$ flow diagrams for the saddle-node bifurcation $\dot{u} = r + u^2$ (left) and the transcritical bifurcation $\dot{u} = ru - u^2$ (right).

Note that the sign of the $u^2$ term can be reversed relative to the others by sending $u \to -u$.

Consider a crude model of a laser threshold. Let $n$ be the number of photons in the laser cavity, and $N$ the number of excited atoms in the cavity. The dynamics of the laser are approximated by the equations

$$\begin{aligned} \dot{n} &= \gamma N n - kn \\ N &= N_0 - \alpha n \quad . \end{aligned} \tag{7.304}$$

Here $\gamma$ is the gain coefficient and $k$ the photon decay rate. $N_0$ is the pump strength, and $\alpha$ is a numerical factor. The first equation tells us that the number of photons in the cavity grows with a rate $\gamma N - k$, *i.e.* gain is proportional to the number of excited atoms, and the loss rate is a constant cavity-dependent quantity (typically through the ends, which are semi-transparent). The second equation says that the number of excited atoms is equal to the pump strength minus a term proportional to the number of photons (since the presence of a photon means an excited atom has decayed). Putting them together,

$$\dot{n} = (\gamma N_0 - k)\, n - \alpha\gamma n^2 \quad , \tag{7.305}$$

which exhibits a transcritical bifurcation at pump strength $N_0 = k/\gamma$. For $N_0 < k/\gamma$ the system acts as a lamp; for $N_0 > k/\gamma$ the system acts as a laser.

What happens in the transcritical bifurcation is an exchange of stability of the fixed points at $u^* = 0$ and $u^* = r$ as $r$ passes through zero. This is depicted graphically in the bottom panel of Fig. 7.25.

### 7.10.4 Pitchfork bifurcation

The pitchfork bifurcation is commonly encountered in systems in which there is an overall parity symmetry ($u \to -u$). There are two classes of pitchfork: supercritical and subcritical. The normal form of the supercritical bifurcation is

$$\dot{u} = ru - u^3 \quad , \tag{7.306}$$

which has fixed points at $u^* = 0$ and $u^* = \pm\sqrt{r}$. Thus, the situation is as depicted in fig. 7.27 (top panel). For $r < 0$ there is a single stable fixed point at $u^* = 0$. For $r > 0$, $u^* = 0$ is unstable, and flanked by two stable fixed points at $u^* = \pm\sqrt{r}$.

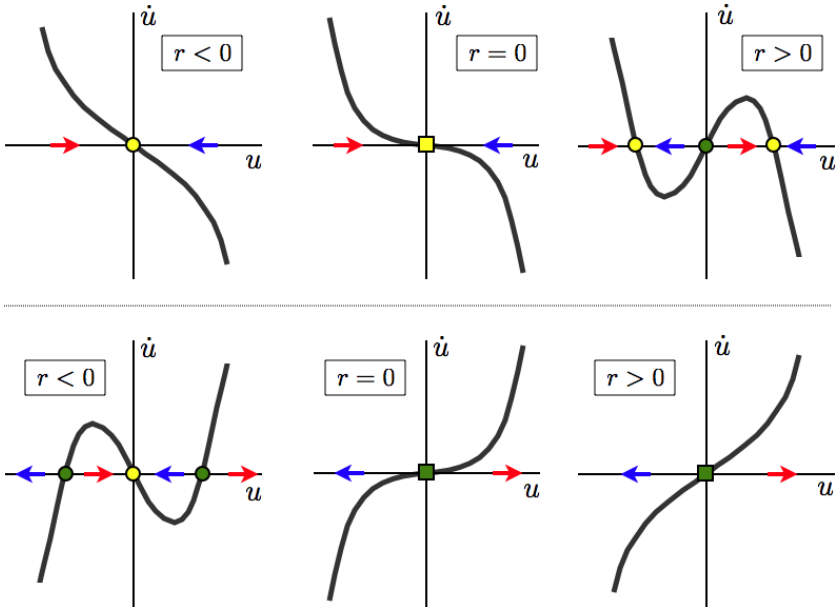



Figure 7.27: Top: supercritical pitchfork bifurcation $\dot{u} = ru - u^3$. Bottom: subcritical pitchfork bifurcation $\dot{u} = ru + u^3$.

If we send $u \to -u$, $r \to -r$, and $t \to -t$, we obtain the *subcritical pitchfork*, depicted in the bottom panel of fig. 7.27. The normal form of the subcritical pitchfork bifurcation is $\dot{u} = ru + u^3$. The fixed point structure in both supercritical and subcritical cases is shown in Fig. 7.28.

### 7.10.5 Imperfect bifurcation

The imperfect bifurcation occurs when a symmetry-breaking term is added to the pitchfork. The normal form contains two control parameters: $\dot{u} = h + ru - u^3$. The constant $h$ explicitly breaks the $\mathbb{Z}_2$ parity symmetry $u \to -u$.

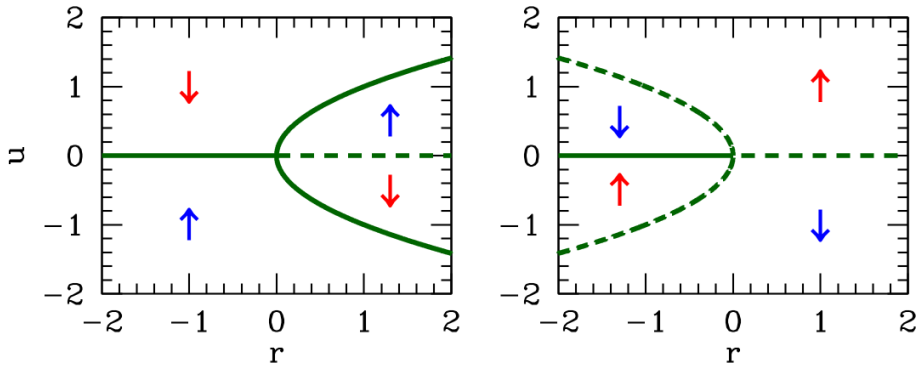



Figure 7.28: Extended phase space $(r, u)$ flow diagrams for the pitchfork bifurcation. Left: supercritical pitchfork $\dot{u} = ru - u^3$. Right: subcritical pitchfork $\dot{u} = ru + u^3$.

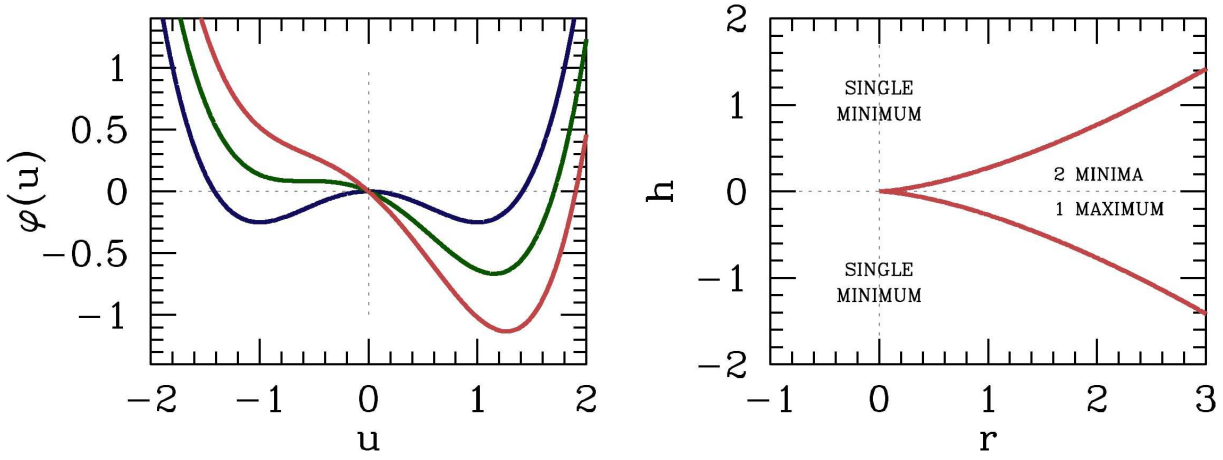



Figure 7.29: Left: scaled free energy $\varphi(u) = -\frac{1}{2}ru^2 + \frac{1}{4}u^4 - hu$, with $h = 0$ (blue), $h = h_{\rm c}$ (green), and $h = 2\,h_{\rm c}$ (red), where $h_{\rm c} = \frac{2}{3\sqrt{3}}\,r^{3/2}$. Right: phase diagram for the imperfect bifurcation $\dot{u} = -\varphi'(u) = h + ru - u^3$ in the $(r, h)$ plane.

This equation arises from a crude model of magnetization dynamics. Let $m$ be the magnetization density of a sample, and $f(m)$ the free energy density. Assuming $m$ is small, we can expand $f(m)$ as

$$f(m) = \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 - Hm + \ldots \quad , \tag{7.307}$$

where $H$ is the external magnetic field, and $a$ and $b$ are temperature-dependent constants. This is called the *Landau expansion* of the free energy. We assume $b > 0$ in order that the minimum of $f(m)$ not lie at infinity. The dynamics of $m(t)$ are modeled by

$$\frac{dm}{dt} = -\Gamma\,\frac{\partial f}{\partial m} \quad , \tag{7.308}$$

with $\Gamma > 0$. Thus, the magnetization evolves toward a local minimum in the free energy. Note that the free energy is a decreasing function of time:

$$\frac{df}{dt} = \frac{\partial f}{\partial m}\,\frac{dm}{dt} = -\Gamma\left(\frac{\partial f}{\partial m}\right)^2 \quad . \tag{7.309}$$

By rescaling $\tau \equiv \Gamma t$, $m \equiv b^{-1/2}u$, $h \equiv b^{1/2}H$, and defining $r \equiv -a$, we obtain the normal form of the bifurcation,

$$\dot{u} = h + ru - u^3 = -\frac{\partial\varphi}{\partial u} \quad , \tag{7.310}$$

where

$$\varphi(u) = -\tfrac{1}{2}ru^2 + \tfrac{1}{4}u^4 - hu \tag{7.311}$$

is a scaled version of the free energy. Fixed points satisfy the equation

$$u^3 - ru - h = 0 \quad , \tag{7.312}$$

and correspond to extrema in $\varphi(u)$. By the fundamental theorem of algebra, this cubic polynomial may be uniquely factorized over the complex plane. Since the coefficients are real, the complex conjugate $\bar{u}$

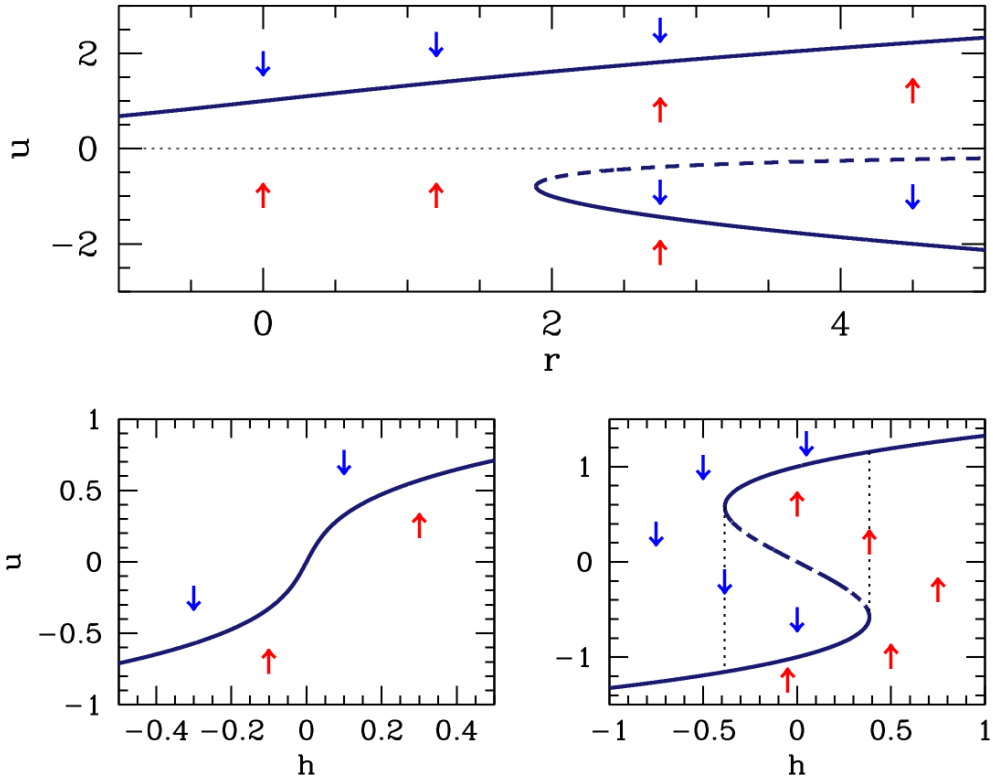



Figure 7.30: Top: extended phase space $(r, u)$ flow diagram for $\dot{u} = h + ru - u^3$, the imperfect pitchfork bifurcation, at $h = 1$. This is in a sense a deformed supercritical pitchfork. The saddle-node bifurcation occurs at $r_{\rm c} = (3/2^{2/3})|h|^{2/3} = 1.8899$. Bottom: extended phase space $(h, u)$ flow diagram for the imperfect pitchfork bifurcation $r = -0.2$ (left panel) and $r = 1$ (right panel). For $r < 0$ the behavior is completely reversible. For $r > 0$, a regime of irreversibility sets in between $-h_{\rm c}$ and $+h_{\rm c}$, where $h_{\rm c} = 2(r/3)^{3/2}$. The system then exhibits the phenomenon of hysteresis. The dotted vertical lines show the boundaries of the hysteresis loop.

satisfies the same equation as $u$, hence there are two possibilities for the roots: either (i) all three roots are real, or (ii) one root is real and the other two are a complex conjugate pair. Clearly for $r < 0$ we are in situation (ii) since $u^3 - ru$ is then monotonically increasing for $u \in \mathbb{R}$, and therefore takes the value $h$ precisely once for $u$ real. For $r > 0$, there is a region $h \in \big[ -h_{\rm c}(r), h_{\rm c}(r) \big]$ over which there are three real roots. To find $h_{\rm c}(r)$, we demand $\varphi''(u) = 0$ as well as $\varphi'(u) = 0$, which says that two roots have merged in a saddle-node bifurcation, forming an inflection point. Thus $\varphi''(u) = 3u^2 - r = 0$ yields $u = \pm(r/3)^{1/2}$, which requires $r > 0$, and using this to eliminate $u$ from the equation $\varphi'(u) = 0$ yields the critical value of $h$ as a function of $r$, *viz.* $h_{\rm c}(r) = \frac{2}{3\sqrt{3}}\, r^{3/2}\, \Theta(r)$.

Examples of the function $\varphi(u)$ for $r > 0$ are shown in the left panel of Fig. 7.29 for three different values of $h$. For $|h| < h_{\rm c}(r)$ there are three extrema satisfying $\varphi'(u^*) = 0$: $u_1^* < u_2^* < 0 < u_3^*$, assuming (without loss of generality) that $h > 0$. Clearly $u_1^*$ is a local minimum, $u_2^*$ a local maximum, and $u_3^*$ the global minimum of the function $f(u)$. The 'phase diagram' for this system, plotted in the $(r, h)$ control parameter space, is shown in the right panel of Fig. 7.29.

In Fig. 7.30 we plot the fixed points $u^*(r)$ for fixed $h$. A saddle-node bifurcation occurs at $r = r_{\rm c}(h) = \frac{3}{2^{2/3}}\,|h|^{2/3}$. For $h = 0$ this reduces to the supercritical pitchfork; for finite $h$ the pitchfork is deformed and even changed topologically. Finally, in Fig. 7.30 we show the behavior of $u^*(h)$ for fixed $r$. When $r < 0$ the curve retraces itself as $h$ is ramped up and down, but for $r > 0$ the system exhibits the phenomenon of *hysteresis*, *i.e.* there is an irreversible aspect to the behavior. Fig. 7.30 shows a *hysteresis loop* when $r > 0$.

## 7.11 Appendix II: Equivalence of the Mean Field Descriptions

In both the variational density matrix and mean field Hamiltonian methods as applied to the Ising model, we obtained the same result $m = \tanh\big((m+h)/\theta\big)$. What is perhaps not obvious is whether these theories are in fact the same, *i.e.* if their respective free energies agree. Indeed, the two free energy functions,

$$\begin{aligned} f_{\rm A}(m,h,\theta) &= -\tfrac{1}{2}\,m^2 - hm + \theta\left\{\left(\frac{1+m}{2}\right)\log\left(\frac{1+m}{2}\right) + \left(\frac{1-m}{2}\right)\log\left(\frac{1-m}{2}\right)\right\} \\ f_{\rm B}(m,h,\theta) &= +\tfrac{1}{2}\,m^2 - \theta\,\log\left(e^{+(m+h)/\theta} + e^{-(m+h)/\theta}\right) \quad , \end{aligned} \tag{7.313}$$

where $f_{\rm A}$ is the variational density matrix result and $f_{\rm B}$ is the mean field Hamiltonian result, clearly are different functions of their arguments. However, it turns out that upon minimizing with respect to $m$ in each case, the resulting free energies obey $f_{\rm A}(h,\theta) = f_{\rm B}(h,\theta)$. This agreement may seem surprising. The first method utilizes an approximate (variational) density matrix applied to the exact Hamiltonian $\hat{H}$. The second method approximates the Hamiltonian as $\hat{H}_{\rm MF}$, but otherwise treats it exactly. The two Landau expansions seem hopelessly different:

$$\begin{aligned} f_{\rm A}(m,h,\theta) &= -\theta\,\log 2 - hm + \tfrac{1}{2}\,(\theta-1)\,m^2 + \tfrac{\theta}{12}\,m^4 + \tfrac{\theta}{30}\,m^6 + \ldots \\ f_{\rm B}(m,h,\theta) &= -\theta\,\log 2 + \tfrac{1}{2}m^2 - \frac{(m+h)^2}{2\,\theta} + \frac{(m+h)^4}{12\,\theta^3} - \frac{(m+h)^6}{45\,\theta^5} + \ldots \quad . \end{aligned} \tag{7.314}$$

We shall now prove that these two methods, the variational density matrix and the mean field approach, are in fact equivalent, and yield the *same* free energy $f(\theta,h)$.

Let us generalize the Ising model and write

$$\hat{H} = -\sum_{i<j} J_{ij}\,\varepsilon(\sigma_i,\sigma_j) - \sum_i \Phi(\sigma_i) \quad . \tag{7.315}$$

Here, each 'spin' $\sigma_i$ may take on any of $K$ possible values, $\{s_1,\ldots,s_K\}$. For the $S = 1$ Ising model, we would have $K = 3$ possibilities, with $s_1 = -1$, $s_2 = 0$, and $s_3 = +1$. But the set $\{s_\alpha\}$, with $\alpha \in \{1,\ldots,K\}$, is completely arbitrary[35]. The 'local field' term $\Phi(\sigma)$ is also a completely arbitrary function. It may be linear, with $\Phi(\sigma) = H\sigma$, for example, but it could also contain terms quadratic in $\sigma$, or whatever one desires.

[35] It needn't be an equally spaced sequence, for example.

The symmetric, dimensionless interaction function $\varepsilon(\sigma,\sigma') = \varepsilon(\sigma',\sigma)$ is a real symmetric $K \times K$ matrix. According to the singular value decomposition theorem, any such matrix may be written in the form

$$\varepsilon(\sigma,\sigma') = \sum_{p=1}^{N_{\rm s}} A_p \, \lambda_p(\sigma) \, \lambda_p(\sigma') \quad , \tag{7.316}$$

where $\{A_p\}$ are the eigenvalues and $\big\{\lambda_p(\sigma)\big\}$ the eigenectors. The number of terms $N_{\rm s}$ in this decomposition is such that $N_{\rm s} \le K$. This treatment can be generalized to account for continuous $\sigma$.

### 7.11.1 Variational density matrix

The most general single-site variational density matrix is written

$$\varrho(\sigma) = \sum_{\alpha=1}^{K} x_\alpha \, \delta_{\sigma,s_\alpha} \quad . \tag{7.317}$$

Thus, $x_\alpha$ is the probability for a given site to be in state $\alpha$, with $\sigma = s_\alpha$. *The $\{x_\alpha\}$ are the $K$ variational parameters, subject to the single normalization constraint,* $\sum_\alpha x_\alpha = 1$. We now have

$$\begin{aligned} f &= \frac{1}{N\hat{J}(0)} \Big\{ \, \mathsf{Tr}\,(\varrho \hat{H}) + k_{\rm B}T \; \mathsf{Tr}\,(\varrho \log \varrho) \Big\} \\ &= -\tfrac{1}{2} \sum_p \sum_{\alpha,\alpha'} A_p \, \lambda_p(s_\alpha) \, \lambda_p(s_{\alpha'}) \, x_\alpha \, x_{\alpha'} - \sum_\alpha \varphi(s_\alpha) \, x_\alpha + \theta \sum_\alpha x_\alpha \log x_\alpha \quad , \end{aligned} \tag{7.318}$$

where $\varphi(\sigma) = \Phi(\sigma)/\hat{J}(0)$. We extremize in the usual way, introducing a Lagrange undetermined multiplier $\zeta$ to enforce the constraint. This means we extend the function $f\big(\{x_\alpha\}\big)$, writing

$$f^*(x_1,\ldots,x_K,\zeta) = f(x_1,\ldots,x_K) + \zeta \left( \sum_{\alpha=1}^{K} x_\alpha - 1 \right) \quad , \tag{7.319}$$

and freely extremizing with respect to the $(K+1)$ parameters $\{x_1,\ldots,x_K,\zeta\}$. This yields $K$ nonlinear equations,

$$0 = \frac{\partial f^*}{\partial x_\alpha} = -\sum_p \sum_{\alpha'} A_p \, \lambda_p(s_\alpha) \, \lambda_p(s_{\alpha'}) \, x_{\alpha'} - \varphi(s_\alpha) + \theta \, \log x_\alpha + \zeta + \theta \quad , \tag{7.320}$$

for each $\alpha$, and one linear equation, which is the normalization condition,

$$0 = \frac{\partial f^*}{\partial \zeta} = \sum_\alpha x_\alpha - 1 \quad . \tag{7.321}$$

We cannot solve these nonlinear equations analytically, but they may be recast, by exponentiating them, as

$$x_\alpha = \frac{1}{Z} \, \exp\left\{ \frac{1}{\theta} \left[ \sum_p \sum_{\alpha'} A_p \, \lambda_p(s_\alpha) \, \lambda_p(s_{\alpha'}) \, x_{\alpha'} + \varphi(s_\alpha) \right] \right\} \quad , \tag{7.322}$$

with

$$Z = e^{(\zeta/\theta)+1} = \sum_{\alpha} \exp\left\{ \frac{1}{\theta} \left[ \sum_{p} \sum_{\alpha'} A_p \, \lambda_p(s_\alpha) \, \lambda_p(s_{\alpha'}) \, x_{\alpha'} + \varphi(s_\alpha) \right] \right\} \quad . \tag{7.323}$$

From the logarithm of $x_\alpha$, we may compute the entropy, and, finally, the free energy:

$$f(\theta) = \tfrac{1}{2} \sum_{p} \sum_{\alpha,\alpha'} A_p \, \lambda_p(s_\alpha) \, \lambda_p(s_{\alpha'}) \, x_\alpha \, x_{\alpha'} - \theta \log Z \quad , \tag{7.324}$$

which is to be evaluated at the solution of 7.320, $\{x^*_\alpha(h,\theta)\}$.

### 7.11.2 Mean field approximation

We now derive a mean field approximation in the spirit of that used in the Ising model above. We write

$$\lambda_p(\sigma) = \big\langle \lambda_p(\sigma) \big\rangle + \delta\lambda_p(\sigma) \quad , \tag{7.325}$$

and abbreviate $\bar\lambda_p = \big\langle \lambda_p(\sigma) \big\rangle$, the thermodynamic average of $\lambda_p(\sigma)$ on any given site. We then have

$$\begin{aligned} \lambda_p(\sigma) \, \lambda_p(\sigma') &= \bar\lambda_p^2 + \bar\lambda_p \, \delta\lambda_p(\sigma) + \bar\lambda_p \, \delta\lambda_p(\sigma') + \delta\lambda_p(\sigma) \, \delta\lambda_p(\sigma') \\ &= -\bar\lambda_p^2 + \bar\lambda_p \left( \lambda_p(\sigma) + \lambda_p(\sigma') \right) + \delta\lambda_p(\sigma) \, \delta\lambda_p(\sigma') \quad . \end{aligned} \tag{7.326}$$

The product $\delta\lambda_p(\sigma) \, \delta\lambda_p(\sigma')$ is of second order in fluctuations, and we neglect it. This leads us to the mean field Hamiltonian,

$$\hat H_{\rm MF} = +\tfrac{1}{2} N \hat J(0) \sum_p A_p \, \bar\lambda_p^2 - \sum_{i=1}^{N} \left[ \hat J(0) \sum_p A_p \, \bar\lambda_p \, \lambda_p(\sigma_i) + \Phi(\sigma_i) \right] \quad . \tag{7.327}$$

The free energy is then

$$f\big(\{\bar\lambda_p\},\theta\big) = \tfrac{1}{2} \sum_p A_p \, \bar\lambda_p^2 - \theta \, \log \sum_\alpha \exp\left\{ \frac{1}{\theta} \left[ \sum_p A_p \, \bar\lambda_p \, \lambda_p(s_\alpha) + \varphi(s_\alpha) \right] \right\} \quad . \tag{7.328}$$

*The variational parameters are the mean field values* $\{\bar\lambda_p\}$.

The single site probabilities $\{x_\alpha\}$ are then

$$x_\alpha = \frac{1}{Z} \, \exp\left\{ \frac{1}{\theta} \left[ \sum_p A_p \, \bar\lambda_p \, \lambda_p(s_\alpha) + \varphi(s_\alpha) \right] \right\} \quad , \tag{7.329}$$

with $Z$ implied by the normalization $\sum_\alpha x_\alpha = 1$. These results reproduce exactly what we found in eqn. 7.320, since the mean field equation here, $\partial f / \partial \bar\lambda_p = 0$, yields

$$\bar\lambda_p = \sum_{\alpha=1}^{K} \lambda_p(s_\alpha) \, x_\alpha \quad . \tag{7.330}$$

The free energy is immediately found to be

$$f(\theta) = \tfrac{1}{2}\sum_p A_p\,\bar{\lambda}_p^2 - \theta\,\log Z \quad , \tag{7.331}$$

which again agrees with what we found using the variational density matrix.

Thus, whether one extremizes with respect to the set $\{x_1, \ldots, x_K, \zeta\}$, or with respect to the set $\{\bar{\lambda}_p\}$, the results are the same, in terms of all these parameters, as well as the free energy $f(\theta)$. Generically, both approaches may be termed 'mean field theory' since the variational density matrix corresponds to a mean field which acts on each site independently[36].

## 7.12 Appendix III: Additional Examples of Mean Field Theory

### 7.12.1 Blume-Capel model

The Blume-Capel model provides a simple and convenient way to model systems with vacancies. The simplest version of the model is written

$$\hat{H} = -\tfrac{1}{2}\sum_{i,j=1}^{N} J_{ij}\,S_i\,S_j + \Delta\sum_{i=1}^{N} S_i^2 \quad . \tag{7.332}$$

The spin variables $S_i$ range over the values $\{-1\,,\,0\,,\,+1\}$, so this is an extension of the $S = 1$ Ising model. We explicitly separate out the diagonal terms, writing $J_{ii} \equiv 0$, and placing them in the second term on the RHS above. We say that site $i$ is occupied if $S_i = \pm 1$ and vacant if $S_i = 0$, and we identify $-\Delta$ as the vacancy creation energy, which may be positive or negative, depending on whether vacancies are disfavored or favored in our system.

We make the mean field *Ansatz*, writing $S_i = m + \delta S_i$. This results in the mean field Hamiltonian,

$$\hat{H}_{\rm MF} = \tfrac{1}{2}N\hat{J}(0)\,m^2 - \hat{J}(0)\,m\sum_{i=1}^{N} S_i + \Delta\sum_{i=1}^{N} S_i^2 \quad . \tag{7.333}$$

Once again, we adimensionalize, writing $f \equiv F/N\hat{J}(0)$, $\theta = k_{\rm B}T/\hat{J}(0)$, and $\delta = \Delta/\hat{J}(0)$. We assume $\hat{J}(0) > 0$. The free energy per site is then

$$f(m,\theta,\delta) = \tfrac{1}{2}m^2 - \theta\log\Big(1 + 2e^{-\delta/\theta}\,\cosh(m/\theta)\Big) \quad . \tag{7.334}$$

Extremizing with respect to $m$, we obtain the mean field equation,

$$m = \frac{2\sinh(m/\theta)}{\exp(\delta/\theta) + 2\cosh(m/\theta)} \quad . \tag{7.335}$$

[36] The function $\Phi(\sigma)$ may involve one or more adjustable parameters which could correspond, for example, to an external magnetic field $h$. We suppress these parameters when we write the free energy as $f(\theta)$.

Note that $m = 0$ is always a solution. Finding the slope of the RHS at $m = 0$ and setting it to unity gives at $\theta = \theta^*$ yields

$$\theta^*(\delta) = \frac{2}{\exp(\delta/\theta^*) + 2} \quad . \tag{7.336}$$

This is an implicit equation for $\theta^*$ in terms of the vacancy energy $\delta$. Equivalently, we can write

$$\delta^*(\theta) = \theta \log\left(\frac{2}{\theta} - 2\right) \quad . \tag{7.337}$$

Is $\theta^*(\delta)$ the critical temperature? We need to investigate further.

Let's expand the free energy in terms of the magnetization $m$. We find, to fourth order,

$$f = -\theta \log\left(1 + 2e^{-\delta/\theta}\right) + \frac{1}{2\theta}\left(\theta - \frac{2}{2 + \exp(\delta/\theta)}\right)m^2 + \frac{1}{12\left(2 + \exp(\delta/\theta)\right)\theta^3}\left(\frac{6}{2 + \exp(\delta/\theta)} - 1\right)m^4 + \ldots \quad . \tag{7.338}$$

Note that setting the coefficient of the $m^2$ term to zero yields the equation for $\theta^*(\delta)$. However, upon further examination, we see that the coefficient of the $m^4$ term can *also* vanish. As we have seen, when the coefficients of the $m^2$ and the $m^4$ terms vanish concomitantly, we have a *tricritical point*[37]. Setting both coefficients to zero, we obtain

$$\theta_{\rm t} = \tfrac{1}{3} \quad , \quad \delta_{\rm t} = \tfrac{2}{3}\log 2 \quad . \tag{7.339}$$

At $\theta = 0$, it is easy to see we have a first order transition at $\delta = \frac{1}{2}$, simply by comparing the energies of the paramagnetic ($S_i = 0$) and ferromagnetic ($S_i = +1$ or $S_i = -1$) states. We have

$$\frac{E_{\rm MF}}{N\hat{J}(0)} = \begin{cases} 0 & \text{if } m = 0 \\ -\frac{1}{2} + \delta & \text{if } m = \pm 1 \end{cases} \quad . \tag{7.340}$$

These results are exact, and not only valid for the mean field theory. Mean field theory is approximate because it neglects fluctuations, but at zero temperature, there are no fluctuations to neglect!

How about the first order line in the phase diagram? To fix this, we seek the boundary between what are the analogs to regions D and E in fig. 7.117. This requires that $\partial^2 f/\partial m^2 > 0$ for $m = 0$, which means $\delta > \delta^*(\theta)$. We then invoke the mean field equation eqn. 7.335 as well as the condition $f(m) = f(0)$. This latter condition is found to be

$$m^2 = 2\theta \log\left(\frac{1 + 2\,e^{-\delta/\theta}\cosh(m/\theta)}{1 + 2\,e^{-\delta/\theta}}\right) \quad . \tag{7.341}$$

The simultaneous solution of this equation and the mean field equation yield the equation of the first order portion of the critical line $\delta_{\rm c}(\theta)$, as well as the order parameter discontinuity $m^*(\theta)$ at the first order transition. As $\theta \to \theta_{\rm t}^-$, we have $\delta^*(\theta_{\rm t}^-) = \delta_{\rm t} = \frac{2}{3}\log 2$ and $m^*(\theta_{\rm t}^-) = 0$.

[37] We should really check that the coefficient of the sixth order term is positive, but that is left as an exercise to the eager student.

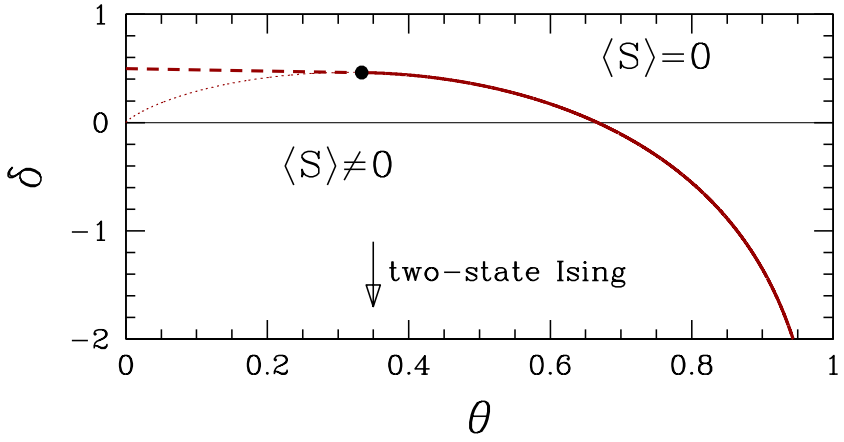



Figure 7.31: Mean field phase diagram for the Blume-Capel model. The black dot signifies a tricritical point, where the coefficients of $m^2$ and $m^4$ in the Landau free energy expansion both vanish. The dashed curve denotes a first order transition, and the solid curve a second order transition. The thin dotted line is the continuation of the $\delta^*(\theta)$ relation to zero temperature.

The phase diagram is shown in fig. 7.31. Note that for $\delta$ large and negative, vacancies are strongly disfavored, hence the only allowed states on each site have $S_i = \pm 1$, which is our old friend the two-state Ising model. Accordingly, the phase boundary there approaches the vertical line $\theta_{\rm c} = 1$, which is the mean field transition temperature for the two-state Ising model.

### 7.12.2 Ising antiferromagnet in an external field

Consider the following model:

$$\hat{H} = J \sum_{\langle ij \rangle} \sigma_i \, \sigma_j - H \sum_{i=1}^{N} \sigma_i \quad , \tag{7.342}$$

with $J > 0$ and $\sigma_i = \pm 1$. We've solved for the mean field phase diagram of the Ising ferromagnet; what happens if the interactions are antiferromagnetic?

It turns out that under certain circumstances, the ferromagnet and the antiferromagnet behave exactly the same in terms of their phase diagram, response functions, *etc.* This occurs when $H = 0$, and when the interactions are between nearest neighbors on a *bipartite lattice*. A bipartite lattice is one which can be divided into two sublattices, which we call A and B, such that an A site has only B neighbors, and a B site has only A neighbors. The square, honeycomb, and body centered cubic (BCC) lattices are bipartite. The triangular and face centered cubic lattices are non-bipartite. Now if the lattice is bipartite and the interaction matrix $J_{ij}$ is nonzero only when $i$ and $j$ are from different sublattices (they needn't be nearest neighbors only), then we can simply redefine the spin variables such that

$$\sigma'_j = \begin{cases} +\sigma_j & \text{if } j \in \text{A} \\ -\sigma_j & \text{if } j \in \text{B} \end{cases} \quad . \tag{7.343}$$

Then $\sigma_i'\sigma_j' = -\sigma_i\,\sigma_j$, and in terms of the new spin variables the exchange constant has reversed. The thermodynamic properties are invariant under such a redefinition of the spin variables.

We can see why this trick doesn't work in the presence of a magnetic field, because the field $H$ would have to be reversed on the B sublattice. In other words, the thermodynamics of an Ising ferromagnet on a bipartite lattice in a uniform applied field is identical to that of the Ising antiferromagnet, with the same exchange constant (in magnitude), in the presence of a *staggered field* $H_{\rm A} = +H$ and $H_{\rm B} = -H$.

We treat this problem using the variational density matrix method, using two independent variational parameters $m_{\rm A}$ and $m_{\rm B}$ for the two sublattices:

$$\begin{aligned}\varrho_{\rm A}(\sigma) &= \frac{1+m_{\rm A}}{2}\,\delta_{\sigma,1} + \frac{1-m_{\rm A}}{2}\,\delta_{\sigma,-1} \\ \varrho_{\rm B}(\sigma) &= \frac{1+m_{\rm B}}{2}\,\delta_{\sigma,1} + \frac{1-m_{\rm B}}{2}\,\delta_{\sigma,-1} \quad .\end{aligned} \tag{7.344}$$

With the usual adimensionalization, $f = F/NzJ$, $\theta = k_{\rm B}T/zJ$, and $h = H/zJ$, we have the free energy

$$f(m_{\rm A},m_{\rm B}) = \tfrac{1}{2}m_{\rm A}m_{\rm B} - \tfrac{1}{2}\,h\,(m_{\rm A}+m_{\rm B}) - \tfrac{1}{2}\,\theta\,s(m_{\rm A}) - \tfrac{1}{2}\,\theta\,s(m_{\rm B}) \quad , \tag{7.345}$$

where the entropy function is

$$s(m) = -\left[\frac{1+m}{2}\,\log\!\left(\frac{1+m}{2}\right) + \frac{1-m}{2}\,\log\!\left(\frac{1-m}{2}\right)\right] \quad . \tag{7.346}$$

Note that

$$\frac{ds}{dm} = -\tfrac{1}{2}\,\log\!\left(\frac{1+m}{1-m}\right) \qquad , \qquad \frac{d^2s}{dm^2} = -\frac{1}{1-m^2} \quad . \tag{7.347}$$

Differentiating $f(m_{\rm A},m_{\rm B})$ with respect to the variational parameters, we obtain two coupled mean field equations:

$$\begin{aligned}\frac{\partial f}{\partial m_{\rm A}} = 0 \quad &\Longrightarrow \quad m_{\rm B} = h - \frac{\theta}{2}\log\!\left(\frac{1+m_{\rm A}}{1-m_{\rm A}}\right) \\ \frac{\partial f}{\partial m_{\rm B}} = 0 \quad &\Longrightarrow \quad m_{\rm A} = h - \frac{\theta}{2}\log\!\left(\frac{1+m_{\rm B}}{1-m_{\rm B}}\right) \quad .\end{aligned} \tag{7.348}$$

Recognizing $\tanh^{-1}(x) = \frac{1}{2}\log\Big[(1+x)/(1-x)\Big]$, we may write these equations in an equivalent but perhaps more suggestive form:

$$m_{\rm A} = \tanh\left(\frac{h-m_{\rm B}}{\theta}\right) \qquad , \qquad m_{\rm B} = \tanh\left(\frac{h-m_{\rm A}}{\theta}\right) \quad . \tag{7.349}$$

In other words, the A sublattice sites see an internal field $H_{\rm A,int} = -zJm_{\rm B}$ from their B neighbors, and the B sublattice sites see an internal field $H_{\rm B,int} = -zJm_{\rm A}$ from their A neighbors.

We can solve these equations graphically, as in fig. 7.32. Note that there is always a paramagnetic solution with $m_{\rm A} = m_{\rm B} = m$, where

$$m = h - \frac{\theta}{2}\,\log\!\left(\frac{1+m}{1-m}\right) \qquad \Longleftrightarrow \qquad m = \tanh\left(\frac{h-m}{\theta}\right) \quad . \tag{7.350}$$

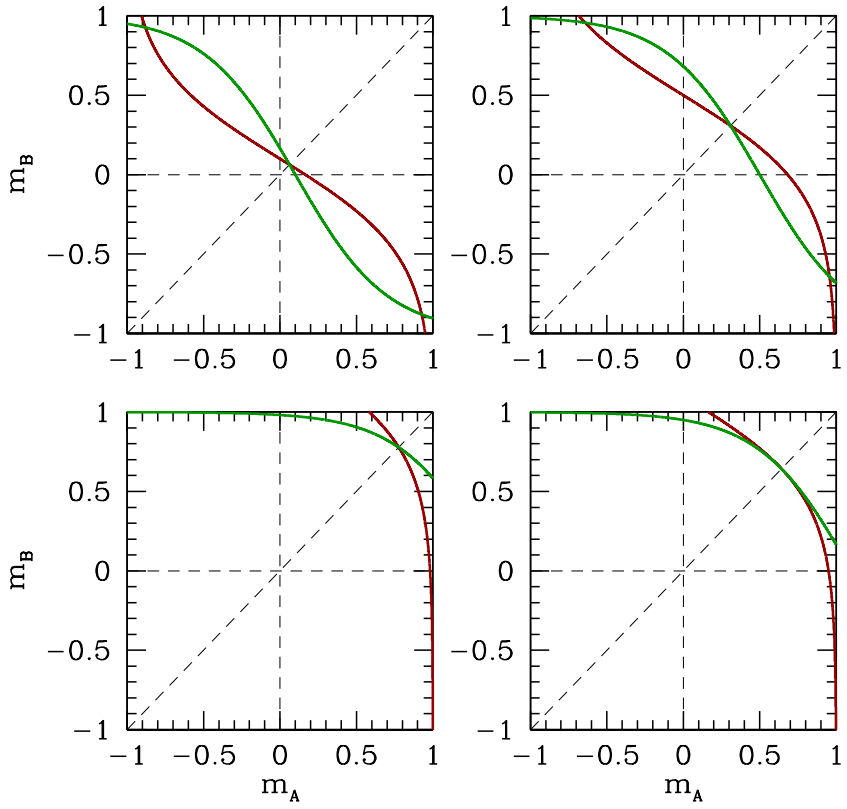



Figure 7.32: Graphical solution to the mean field equations for the Ising antiferromagnet in an external field, here for $\theta = 0.6$. Clockwise from upper left: (a) $h = 0.1$, (b) $h = 0.5$, (c) $h = 1.1$, (d) $h = 1.4$.

However, we can see from the figure that there will be *three* solutions to the mean field equations provided that $\frac{\partial m_{\rm A}}{\partial m_{\rm B}} < -1$ at the point of the solution where $m_{\rm A} = m_{\rm B} = m$. This gives us two equations with which to eliminate $m_{\rm A}$ and $m_{\rm B}$, resulting in the curve

$$h^*(\theta) = m + \frac{\theta}{2}\,\log\left(\frac{1+m}{1-m}\right) \qquad \text{with } \; m = \sqrt{1-\theta} \quad . \tag{7.351}$$

Thus, for $\theta < 1$ and $|h| < h^*(\theta)$ there are three solutions to the mean field equations. It is *usually* the case, the broken symmetry solutions, which mean those for which $m_{\rm A} \neq m_{\rm B}$ in our case, are of lower energy than the symmetric solution(s). We show the curve $h^*(\theta)$ in fig. 7.33.

We can make additional progress by defining the average and staggered magnetizations $m$ and $m_{\rm s}$,

$$m \equiv \tfrac{1}{2}(m_{\rm A} + m_{\rm B}) \qquad , \quad m_{\rm s} \equiv \tfrac{1}{2}(m_{\rm A} - m_{\rm B}) \quad . \tag{7.352}$$

We expand the free energy in terms of $m_{\rm s}$:

$$\begin{aligned} f(m, m_{\rm s}) &= \tfrac{1}{2}m^2 - \tfrac{1}{2}m_{\rm s}^2 - h\,m - \tfrac{1}{2}\,\theta\,s(m+m_{\rm s}) - \tfrac{1}{2}\,\theta\,s(m-m_{\rm s}) \\ &= \tfrac{1}{2}m^2 - h\,m - \theta\,s(m) - \tfrac{1}{2}\Big(1 + \theta\,s''(m)\Big)\,m_{\rm s}^2 - \tfrac{1}{24}\,\theta\,s''''(m)\,m_{\rm s}^4 + \ldots \quad . \end{aligned} \tag{7.353}$$

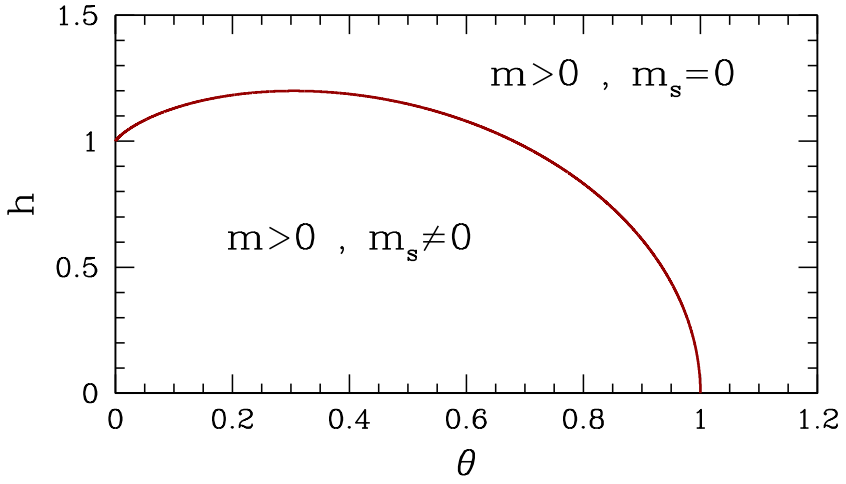



Figure 7.33: Mean field phase diagram for the Ising antiferromagnet in an external field. The phase diagram is symmetric under reflection in the $h = 0$ axis.

The term quadratic in $m_{\rm s}$ vanishes when $\theta\, s''(m) = -1$, *i.e.* when $m = \sqrt{1-\theta}$. It is easy to obtain

$$\frac{d^3s}{dm^3} = -\frac{2m}{(1-m^2)^2} \quad , \quad \frac{d^4s}{dm^4} = -\frac{2\,(1+3m^2)}{(1-m^2)^3} \quad , \tag{7.354}$$

from which we learn that the coefficient of the quartic term, $-\frac{1}{24}\,\theta\, s''''(m)$, never vanishes. Therefore the transition remains second order down to $\theta = 0$, where it finally becomes first order.

We can confirm the $\theta \to 0$ limit directly. The two competing states are the ferromagnet, with $m_{\rm A} = m_{\rm B} = \pm 1$, and the antiferromagnet, with $m_{\rm A} = -m_{\rm B} = \pm 1$. The free energies of these states are

$$f^{\rm FM} = \tfrac{1}{2} - h \quad , \quad f^{\rm AFM} = -\tfrac{1}{2} \quad . \tag{7.355}$$

There is a first order transition when $f^{\rm FM} = f^{\rm AFM}$, which yields $h = 1$.

### 7.12.3 Canted quantum antiferromagnet

Consider the following model for *quantum* $S = \frac{1}{2}$ spins:

$$\hat{H} = \sum_{\langle ij \rangle} \Big[ -J\big(\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y\big) + \Delta\, \sigma_i^z \sigma_j^z \Big] + \tfrac{1}{4} K \sum_{\langle ijkl \rangle} \sigma_i^z \sigma_j^z \sigma_k^z \sigma_l^z \quad , \tag{7.356}$$

where $\boldsymbol{\sigma}_i$ is the vector of Pauli matrices on site $i$. The spins live on a square lattice. The second sum is over all square plaquettes. All the constants $J$, $\Delta$, and $K$ are positive.

Let's take a look at the Hamiltonian for a moment. The $J$ term clearly wants the spins to align ferromagnetically in the $(x, y)$ plane (in internal spin space). The $\Delta$ term prefers antiferromagnetic alignment along the $\hat{\boldsymbol{z}}$ axis. The $K$ term discourages any kind of moment along $\hat{\boldsymbol{z}}$ and works against the $\Delta$ term. We'd like our mean field theory to capture the physics behind this competition.

Accordingly, we break up the square lattice into two interpenetrating $\sqrt{2} \times \sqrt{2}$ square sublattices (each rotated by $45^\circ$ with respect to the original), in order to be able to describe an antiferromagnetic state. In addition, we include a parameter $\alpha$ which describes the *canting angle* that the spins on these sublattices make with respect to the $\hat{\boldsymbol{x}}$-axis. That is, we write

$$\begin{aligned} \varrho_{\rm A} &= \tfrac{1}{2} + \tfrac{1}{2} m \left( \sin\alpha\, \sigma^x + \cos\alpha\, \sigma^z \right) \\ \varrho_{\rm B} &= \tfrac{1}{2} + \tfrac{1}{2} m \left( \sin\alpha\, \sigma^x - \cos\alpha\, \sigma^z \right) \quad . \end{aligned} \tag{7.357}$$

Note that $\mathsf{Tr}\ \varrho_{\rm A} = \mathsf{Tr}\ \varrho_{\rm B} = 1$ so these density matrices are normalized. Note also that the mean direction for a spin on the A and B sublattices is given by

$$\boldsymbol{m}_{\rm A,B} = \mathsf{Tr}\,(\varrho_{\rm A,B}\,\boldsymbol{\sigma}) = \pm\, m \cos\alpha\, \hat{\boldsymbol{z}} + m \sin\alpha\, \hat{\boldsymbol{x}} \quad . \tag{7.358}$$

Thus, when $\alpha = 0$, the system is an antiferromagnet with its staggered moment lying along the $\hat{\boldsymbol{z}}$ axis. When $\alpha = \frac{1}{2}\pi$, the system is a ferromagnet with its moment lying along the $\hat{\boldsymbol{x}}$ axis.

Finally, the eigenvalues of $\varrho_{\rm A,B}$ are still $\lambda_\pm = \frac{1}{2}(1 \pm m)$, hence

$$\begin{aligned} s(m) &\equiv -\,\mathsf{Tr}\,(\varrho_{\rm A} \log \varrho_{\rm A}) = -\,\mathsf{Tr}\,(\varrho_{\rm B} \log \varrho_{\rm B}) \\ &= -\left[ \frac{1+m}{2}\, \log\!\left( \frac{1+m}{2} \right) + \frac{1-m}{2}\, \log\!\left( \frac{1-m}{2} \right) \right] \quad . \end{aligned} \tag{7.359}$$

Note that we have taken $m_{\rm A} = m_{\rm B} = m$, unlike the case of the antiferromagnet in a uniform field. The reason is that there remains in our model a symmetry between A and B sublattices.

The free energy is now easily calculated:

$$\begin{aligned} F &= \mathsf{Tr}\,(\varrho \hat{H}) + k_{\rm B} T\, \mathsf{Tr}\,(\varrho \log \varrho) \\ &= -2N \Big( J\, \sin^2\!\alpha + \Delta\, \cos^2\!\alpha \Big)\, m^2 + \tfrac{1}{4} N K m^4 \cos^4\!\alpha - N k_{\rm B} T\, s(m) \end{aligned} \tag{7.360}$$

We can adimensionalize by defining $\delta \equiv \Delta/J$, $\kappa \equiv K/4J$, and $\theta \equiv k_{\rm B}T/4J$. Then the free energy per site is $f \equiv F/4NJ$ is

$$f(m,\alpha) = -\tfrac{1}{2} m^2 + \tfrac{1}{2}\big(1 - \delta\big)\, m^2 \cos^2\!\alpha + \tfrac{1}{4} \kappa\, m^4 \cos^4\!\alpha - \theta\, s(m) \quad . \tag{7.361}$$

There are two variational parameters: $m$ and $\alpha$. We thus obtain two coupled mean field equations,

$$\begin{aligned} \frac{\partial f}{\partial m} &= 0 = -m + \big(1 - \delta\big)\, m\, \cos^2\!\alpha + \kappa\, m^3 \cos^4\!\alpha + \tfrac{1}{2}\theta\, \log\!\left( \frac{1+m}{1-m} \right) \\ \frac{\partial f}{\partial \alpha} &= 0 = \Big( 1 - \delta + \kappa\, m^2 \cos^2\!\alpha \Big)\, m^2 \sin\alpha\, \cos\alpha \quad . \end{aligned} \tag{7.362}$$

Let's start with the second of the mean field equations. Assuming $m \neq 0$, it is clear from eqn. 7.361 that

$$\cos^2\!\alpha = \begin{cases} 0 & \text{if } \ \delta < 1 \\ (\delta - 1)/\kappa m^2 & \text{if } \ 1 \le \delta \le 1 + \kappa\, m^2 \\ 1 & \text{if } \ \delta \ge 1 + \kappa\, m^2 \quad . \end{cases} \tag{7.363}$$

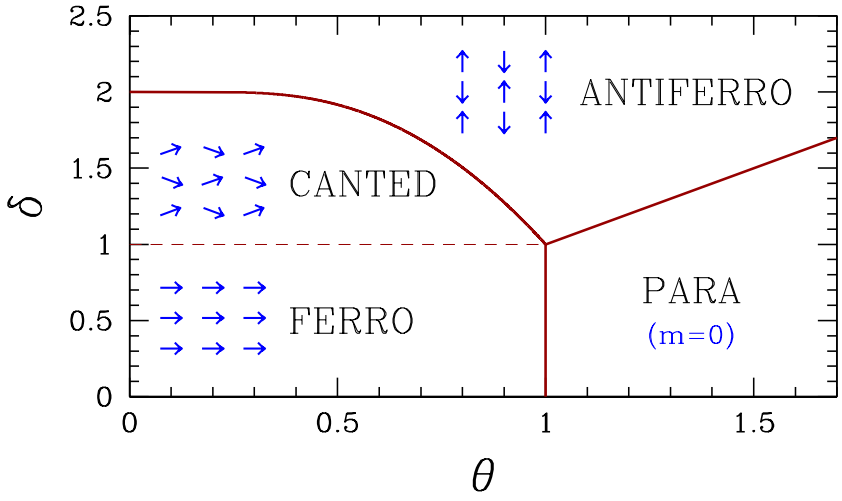



Figure 7.34: Mean field phase diagram for the model of eqn. 7.356 for the case $\kappa = 1$.

Suppose $\delta < 1$. Then we have $\cos\alpha = 0$ and the first mean field equation yields the familiar result

$$m = \tanh\left(m/\theta\right) \quad . \tag{7.364}$$

Along the $\theta$ axis, then, we have the usual ferromagnet-paramagnet transition at $\theta_{\rm c} = 1$.

For $1 < \delta < 1 + \kappa m^2$ we have canting with an angle

$$\alpha = \alpha^*(m) = \cos^{-1}\sqrt{\frac{\delta - 1}{\kappa\, m^2}} \quad . \tag{7.365}$$

Substituting this into the first mean field equation, we once again obtain the relation $m = \tanh\left(m/\theta\right)$. However, eventually, as $\theta$ is increased, the magnetization will dip below the value $m_0 \equiv \sqrt{(\delta - 1)/\kappa}$. This occurs at a dimensionless temperature

$$\theta_0 = \frac{m_0}{\tanh^{-1}(m_0)} < 1 \qquad , \qquad m_0 = \sqrt{\frac{\delta - 1}{\kappa}} \quad . \tag{7.366}$$

For $\theta > \theta_0$, we have $\delta > 1 + \kappa m^2$, and we must take $\cos^2\alpha = 1$. The first mean field equation then becomes

$$\delta m - \kappa m^3 = \frac{\theta}{2}\,\log\left(\frac{1+m}{1-m}\right) \quad , \tag{7.367}$$

or, equivalently, $m = \tanh\left((\delta m - \kappa m^3)/\theta\right)$. A simple graphical analysis shows that a nontrivial solution exists provided $\theta < \delta$. Since $\cos\alpha = \pm 1$, this solution describes an antiferromagnet, with $\boldsymbol{m}_{\rm A} = \pm m\hat{\boldsymbol{z}}$ and $\boldsymbol{m}_{\rm B} = \mp m\hat{\boldsymbol{z}}$. The resulting mean field phase diagram is then as depicted in fig. 7.34.

### 7.12.4 Coupled order parameters

Consider the Landau free energy

$$f(m,\phi) = \tfrac{1}{2}\,a_m\,m^2 + \tfrac{1}{4}\,b_m\,m^4 + \tfrac{1}{2}\,a_\phi\,\phi^2 + \tfrac{1}{4}\,b_\phi\,\phi^4 + \tfrac{1}{2}\Lambda\,m^2\,\phi^2 \quad . \tag{7.368}$$

We write

$$a_m \equiv \alpha_m \, \theta_m \qquad , \qquad a_\phi = \alpha_\phi \, \theta_\phi \quad , \tag{7.369}$$

where

$$\theta_m = \frac{T - T_{\rm c,m}}{T_0} \qquad , \qquad \theta_\phi = \frac{T - T_{\rm c,\phi}}{T_0} \quad , \tag{7.370}$$

where $T_0$ is some temperature scale. We assume without loss of generality that $T_{\rm c,m} > T_{\rm c,\phi}$. We begin by rescaling:

$$m \equiv \left(\frac{\alpha_m}{b_m}\right)^{1/2} \widetilde{m} \qquad , \qquad \phi \equiv \left(\frac{\alpha_m}{b_m}\right)^{1/2} \widetilde{\phi} \quad . \tag{7.371}$$

We then have

$$f = \varepsilon_0 \left\{ r \left( \tfrac{1}{2} \theta_m \, \widetilde{m}^2 + \tfrac{1}{4} \, \widetilde{m}^4 \right) + r^{-1} \left( \tfrac{1}{2} \, \theta_\phi \, \widetilde{\phi}^2 + \tfrac{1}{4} \, \widetilde{\phi}^4 \right) + \tfrac{1}{2} \, \lambda \, \widetilde{m}^2 \widetilde{\phi}^2 \right\} \quad , \tag{7.372}$$

where

$$\varepsilon_0 = \frac{\alpha_m \, \alpha_\phi}{(b_m \, b_\phi)^{1/2}} \qquad , \qquad r = \frac{\alpha_m}{\alpha_\phi} \left( \frac{b_\phi}{b_m} \right)^{1/2} \qquad , \qquad \lambda = \frac{\Lambda}{(b_m \, b_\phi)^{1/2}} \quad . \tag{7.373}$$

It proves convenient to perform one last rescaling, writing

$$\tilde{m} \equiv r^{-1/4} \, \mathsf{m} \qquad , \qquad \tilde{\phi} \equiv r^{1/4} \, \varphi \quad . \tag{7.374}$$

Then

$$f = \varepsilon_0 \left\{ \tfrac{1}{2} q \, \theta_m \, \mathsf{m}^2 + \tfrac{1}{4} \, \mathsf{m}^4 + \tfrac{1}{2} q^{-1} \, \theta_\phi \, \varphi^2 + \tfrac{1}{4} \, \varphi^4 + \tfrac{1}{2} \, \lambda \, \mathsf{m}^2 \, \varphi^2 \right\} \quad , \tag{7.375}$$

where

$$q = \sqrt{r} = \left( \frac{\alpha_m}{\alpha_\phi} \right)^{1/2} \left( \frac{b_\phi}{b_m} \right)^{1/4} \quad . \tag{7.376}$$

Note that we may write

$$f(\mathsf{m}, \varphi) = \frac{\varepsilon_0}{4} \begin{pmatrix} \mathsf{m}^2 & \varphi^2 \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} \mathsf{m}^2 \\ \varphi^2 \end{pmatrix} + \frac{\varepsilon_0}{2} \begin{pmatrix} \mathsf{m}^2 & \varphi^2 \end{pmatrix} \begin{pmatrix} q \, \theta_m \\ q^{-1} \, \theta_\phi \end{pmatrix} \quad . \tag{7.377}$$

The eigenvalues of the above $2 \times 2$ matrix are $1 \pm \lambda$, with corresponding eigenvectors $\left( \begin{smallmatrix} 1 \\ \pm 1 \end{smallmatrix} \right)$. Since $\varphi^2 > 0$, we are only interested in the first eigenvector $\left( \begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right)$, corresponding to the eigenvalue $1 + \lambda$. Clearly when $\lambda < -1$ the free energy is unbounded from below, which is unphysical.

We now set

$$\frac{\partial f}{\partial \mathsf{m}} = 0 \qquad , \qquad \frac{\partial f}{\partial \varphi} = 0 \quad , \tag{7.378}$$

and identify four possible phases:

- Phase I : $\mathsf{m} = 0$, $\varphi = 0$. The free energy is $f_{\rm I} = 0$.

- Phase II : $\mathsf{m} \neq 0$ with $\varphi = 0$. The free energy is

$$f = \frac{\varepsilon_0}{2}\left(q\,\theta_m\,\mathsf{m}^2 + \tfrac{1}{2}\,\mathsf{m}^4\right) \quad , \tag{7.379}$$

  hence we require $\theta_m < 0$ in this phase, in which case

$$\mathsf{m}_{\rm II} = \sqrt{-q\,\theta_m} \qquad , \qquad f_{\rm II} = -\frac{\varepsilon_0}{4}\,q^2\,\theta_m^2 \quad . \tag{7.380}$$

- Phase III : $m = 0$ with $\varphi \neq 0$. The free energy is

$$f = \frac{\varepsilon_0}{2}\left(q^{-1}\,\theta_\phi\,\varphi^2 + \tfrac{1}{2}\,\varphi^4\right) \quad , \tag{7.381}$$

  hence we require $\theta_\phi < 0$ in this phase, in which case

$$\varphi_{\rm III} = \sqrt{-q^{-1}\,\theta_\phi} \qquad , \qquad f_{\rm III} = -\frac{\varepsilon_0}{4}\,q^{-2}\,\theta_\phi^2 \quad . \tag{7.382}$$

- Phase IV : $\mathsf{m} \neq 0$ and $\varphi \neq 0$. Varying $f$ yields

$$\begin{pmatrix} 1 & \lambda \\ \lambda & 1 \end{pmatrix}\begin{pmatrix} \mathsf{m}^2 \\ \varphi^2 \end{pmatrix} = -\begin{pmatrix} q\,\theta_m \\ q^{-1}\,\theta_\phi \end{pmatrix} \quad , \tag{7.383}$$

  with solution

$$\begin{aligned} \mathsf{m}^2 &= \frac{q\,\theta_m - q^{-1}\,\theta_\phi\,\lambda}{\lambda^2 - 1} \\ \varphi^2 &= \frac{q^{-1}\,\theta_\phi - q\,\theta_m\,\lambda}{\lambda^2 - 1} \quad . \end{aligned} \tag{7.384}$$

  Since $\mathsf{m}^2$ and $\varphi^2$ must each be nonnegative, phase IV exists only over a yet-to-be-determined subset of the entire parameter space. The free energy is

$$f_{\rm IV} = \frac{q^2\,\theta_m^2 + q^{-2}\,\theta_\phi^2 - 2\lambda\,\theta_m\,\theta_\phi}{4(\lambda^2 - 1)} \quad . \tag{7.385}$$

We now define $\theta \equiv \theta_m$ and $\tau \equiv \theta_\phi - \theta_m = (T_{{\rm c},m} - T_{{\rm c},\phi})/T_0$. Note that $\tau > 0$. There are three possible temperature ranges to consider.

(1) $\theta_\phi > \theta_m > 0$. The only possible phases are I and IV. For phase IV, we must impose the conditions $m^2 > 0$ and $\phi^2 > 0$. If $\lambda^2 > 1$, then the numerators in eqns. 7.384 must each be positive:

$$\lambda < \frac{q^2\,\theta_m}{\theta_\phi} \qquad , \qquad \lambda < \frac{\theta_\phi}{q^2\,\theta_m} \qquad \Rightarrow \qquad \lambda < \min\left(\frac{q^2\,\theta_m}{\theta_\phi}\,,\,\frac{\theta_\phi}{q^2\theta_m}\right) \quad . \tag{7.386}$$

But since either $q^2\theta_m/\theta_\phi$ or its inverse must be less than or equal to unity, this requires $\lambda < -1$, which is unphysical.

If on the other hand we assume $\lambda^2 < 1$, the non-negativeness of $\mathsf{m}^2$ and $\varphi^2$ requires

$$\lambda > \frac{q^2\,\theta_m}{\theta_\phi} \quad , \quad \lambda > \frac{\theta_\phi}{q^2\,\theta_m} \quad \Rightarrow \quad \lambda > \max\left(\frac{q^2\,\theta_m}{\theta_\phi}\,,\,\frac{\theta_\phi}{q^2\theta_m}\right) > 1 \quad . \tag{7.387}$$

Thus, $\lambda > 1$ and we have a contradiction.

Therefore, the only allowed phase for $\theta > 0$ is phase I.

(2) $\theta_\phi > 0 > \theta_m$. Now the possible phases are I, II, and IV. We can immediately rule out phase I because $f_{\rm II} < f_{\rm I}$. To compare phases II and IV, we compute

$$\Delta f = f_{\rm IV} - f_{\rm II} = \frac{(q\,\lambda\,\theta_m - q^{-1}\,\theta_\phi)^2}{4(\lambda^2-1)} \quad . \tag{7.388}$$

Thus, phase II has the lower energy if $\lambda^2 > 1$. For $\lambda^2 < 1$, phase IV has the lower energy, but the conditions $\mathsf{m}^2 > 0$ and $\varphi^2 > 0$ then entail

$$\frac{q^2\theta_m}{\theta_\phi} < \lambda < \frac{\theta_\phi}{q^2\theta_m} \quad \Rightarrow \quad q^2|\theta_m| > \theta_\phi > 0 \quad . \tag{7.389}$$

Thus, $\lambda$ is restricted to the range

$$\lambda \in \left[-1\,,\,-\frac{\theta_\phi}{q^2|\theta_m|}\right] \quad . \tag{7.390}$$

With $\theta_m \equiv \theta < 0$ and $\theta_\phi \equiv \theta + \tau > 0$, the condition $q^2|\theta_m| > \theta_\phi$ is found to be

$$-\tau < \theta < -\frac{\tau}{q^2+1} \quad . \tag{7.391}$$

Thus, phase IV exists and has lower energy when

$$-\tau < \theta < -\frac{\tau}{r+1} \quad \text{and} \quad -1 < \lambda < -\frac{\theta+\tau}{r\theta} \quad , \tag{7.392}$$

where $r = q^2$.

(3) $0 > \theta_\phi > \theta_m$. In this regime, any phase is possible, however once again phase I can be ruled out since phases II and III are of lower free energy. The condition that phase II have lower free energy than phase III is

$$f_{\rm II} - f_{\rm III} = \frac{\varepsilon_0}{4}\big(q^{-2}\theta_\phi^2 - q^2\theta_m^2\big) < 0 \quad , \tag{7.393}$$

*i.e.* $|\theta_\phi| < r|\theta_m|$, which means $r|\theta| > |\theta| - \tau$. If $r > 1$ this is true for all $\theta < 0$, while if $r < 1$ phase II is lower in energy only for $|\theta| < \tau/(1-r)$.

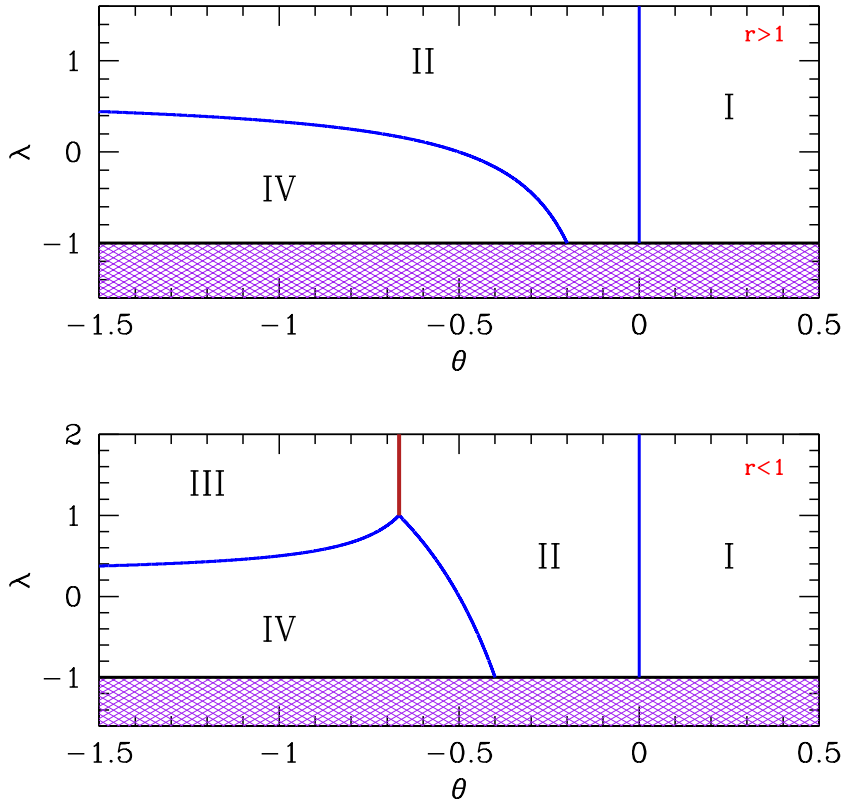



Figure 7.35: Phase diagram for $\tau = 0.5$, $r = 1.5$ (top) and $\tau = 0.5$, $r = 0.25$ (bottom). The hatched purple region is unphysical, with a free energy unbounded from below. The blue lines denote second order transitions. The thick red line separating phases II and III is a first order line.

We next need to test whether phase IV has an even lower energy than the lower of phases II and III. We have

$$f_{\rm IV} - f_{\rm II} = \frac{(q\,\lambda\,\theta_m - q^{-1}\,\theta_\phi)^2}{4(\lambda^2-1)}$$
$$f_{\rm IV} - f_{\rm III} = \frac{(q\,\theta_m - q^{-1}\,\lambda\,\theta_\phi)^2}{4(\lambda^2-1)} \quad . \tag{7.394}$$

In both cases, phase IV can only be the true thermodynamic phase if $\lambda^2 < 1$. We then require $\mathsf{m}^2 > 0$ and $\varphi^2 > 0$, which fixes

$$\lambda \in \left[ -1\,,\, \min\left( \frac{q^2\,\theta_m}{\theta_\phi}\,,\, \frac{\theta_\phi}{q^2\theta_m} \right) \right] \quad . \tag{7.395}$$

The upper limit will be the first term inside the rounded brackets if $q^2|\theta_m| < \theta_\phi$, *i.e.* if $r|\theta| < |\theta| - \tau$. This is impossible if $r > 1$, hence the upper limit is given by the second term in the rounded brackets:

$$r > 1 \;:\; \lambda \in \left[ -1\,,\, \frac{\theta+\tau}{r\,\theta} \right] \qquad \text{(condition for phase IV)} \quad . \tag{7.396}$$

If $r < 1$, then the upper limit will be $q^2\theta_m/\theta_\phi = r\theta/(\theta+\tau)$ if $|\theta| > \tau/(1-r)$, and will be $\theta_\phi/q^2\theta_m = (\theta+\tau)/r\theta$ if $|\theta| < \tau/(1-r)$.

$$\begin{aligned} r<1 \quad , -\frac{\tau}{1-r} < \theta < -\tau \;:\; \lambda \in \left[-1\,,\,\frac{\theta+\tau}{r\theta}\right] \qquad &\text{(phase IV)} \\ r<1 \quad , \theta < -\frac{\tau}{1-r} \;:\; \lambda \in \left[-1\,,\,\frac{r\theta}{\theta+\tau}\right] \qquad &\text{(phase IV)} \quad . \end{aligned} \tag{7.397}$$

Representative phase diagrams for the cases $r > 1$ and $r < 1$ are shown in fig. 7.35.

## 7.13 Appendix IV: Derivation of Ginzburg-Landau Free Energy

We can make some progress in systematically deriving the Ginzburg-Landau free energy. Consider the Ising model,

$$\frac{\hat{H}}{k_{\rm B}T} = -\tfrac{1}{2}\sum_{i,j=1}^{N} K_{ij}\,\sigma_i\,\sigma_j - \sum_{i=1}^{N} h_i\,\sigma_i + \tfrac{1}{2}\sum_{i=1}^{N} K_{ii} \quad , \tag{7.398}$$

where now $K_{ij} = J_{ij}/k_{\rm B}T$ and $h_i = H_i/k_{\rm B}T$ are the interaction energies and local magnetic fields in units of $k_{\rm B}T$. The last term on the RHS above cancels out any contribution from diagonal elements of $K_{ij}$. Our derivation makes use of a generalization of the Gaussian integral,

$$\int\limits_{-\infty}^{\infty} dx\; e^{-\frac{1}{2}ax^2 - bx} = \left(\frac{2\pi}{a}\right)^{1/2} e^{b^2/2a} \quad . \tag{7.399}$$

The generalization is

$$\int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\; e^{-\frac{1}{2}A_{ij}x_ix_j - b_ix_i} = \frac{(2\pi)^{N/2}}{\sqrt{\det A}}\, e^{\frac{1}{2}A^{-1}_{ij}\,b_ib_j} \quad , \tag{7.400}$$

where we use the Einstein convention of summing over repeated indices, and where we assume that the matrix $A$ is positive definite (else the integral diverges). This allows us to write

$$\begin{aligned} Z &= e^{-\frac{1}{2}K_{ii}}\;\mathsf{Tr}\left[e^{\frac{1}{2}K_{ij}\sigma_i\,\sigma_j}\;e^{h_i\,\sigma_i}\right] \\ &= \det{}^{-1/2}(2\pi K)\;e^{-\frac{1}{2}K_{ii}} \int\limits_{-\infty}^{\infty} d\phi_1 \cdots \int\limits_{-\infty}^{\infty} d\phi_N\; e^{-\frac{1}{2}K^{-1}_{ij}\phi_i\phi_j}\;\mathsf{Tr}\;e^{(\phi_i+h_i)\sigma_i} \\ &= \det{}^{-1/2}(2\pi K)\;e^{-\frac{1}{2}K_{ii}} \int\limits_{-\infty}^{\infty} d\phi_1 \cdots \int\limits_{-\infty}^{\infty} d\phi_N\; e^{-\frac{1}{2}K^{-1}_{ij}\phi_i\phi_j}\;e^{\sum_{i=1}^{N}\log\left[2\cosh(\phi_i+h_i)\right]} \\ &\equiv \int\limits_{-\infty}^{\infty} d\phi_1 \cdots \int\limits_{-\infty}^{\infty} d\phi_N\; e^{-\Phi(\phi_1\,,\,\ldots\,,\,\phi_N)} \quad , \end{aligned} \tag{7.401}$$

where

$$\Phi = \tfrac{1}{2}\sum_{i,j=1}^{N} K_{ij}^{-1}\,\phi_i\,\phi_j - \sum_{i=1}^{N}\log\cosh(\phi_i + h_i) + \tfrac{1}{2}\log\det(2\pi K) + \tfrac{1}{2}\,\mathsf{Tr}\;K - N\log 2 \quad . \tag{7.402}$$

We assume the model is defined on a Bravais lattice, in which case we can write $\phi_i = \phi_{\boldsymbol{R}_i}$. We can then define the Fourier transforms,

$$\phi_{\boldsymbol{R}} = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{q}} \hat{\phi}_{\boldsymbol{q}}\, e^{i\boldsymbol{q}\cdot\boldsymbol{R}} \qquad , \qquad \hat{\phi}_{\boldsymbol{q}} = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{R}} \phi_{\boldsymbol{R}}\, e^{-i\boldsymbol{q}\cdot\boldsymbol{R}} \tag{7.403}$$

and

$$\hat{K}(\boldsymbol{q}) = \sum_{\boldsymbol{R}} K(\boldsymbol{R})\, e^{-i\boldsymbol{q}\cdot\boldsymbol{R}} \quad . \tag{7.404}$$

A few remarks about the lattice structure and periodic boundary conditions are in order. For a Bravais lattice, we can write each direct lattice vector $\boldsymbol{R}$ as a sum over $d$ basis vectors with integer coefficients, *viz.* $\boldsymbol{R} = \sum_{\mu=1}^{d} n_\mu\,\boldsymbol{a}_\mu$, where $d$ is the dimension of space. The reciprocal lattice vectors $\boldsymbol{b}_\mu$ have dimensions inverse to those of $\boldsymbol{R}$ and satisfy $\boldsymbol{a}_\mu\cdot\boldsymbol{b}_\nu = 2\pi\,\delta_{\mu\nu}$. The set $\{\boldsymbol{b}_\mu\}$ is complete, hence any wavevector $\boldsymbol{q}$ may be expressed as

$$\boldsymbol{q} = \frac{1}{2\pi}\sum_{\mu=1}^{d}\theta_\mu\,\boldsymbol{b}_\mu \quad . \tag{7.405}$$

We can impose periodic boundary conditions on a system of size $L_1\times L_2\times\cdots\times L_d$ by requiring

$$\phi_{\boldsymbol{R}+\sum_{\mu=1}^{d} l_\mu L_\mu \boldsymbol{a}_\mu} = \phi_{\boldsymbol{R}} \quad . \tag{7.406}$$

This leads to the *quantization* of the wavevectors, which must then satisfy

$$e^{iL_\mu\,\boldsymbol{q}\cdot\boldsymbol{a}_\mu} = e^{iL_\mu\theta_\mu} = 1 \quad , \tag{7.407}$$

and therefore $\theta_\mu = 2\pi m_\mu/L_\mu$, where $m_\mu$ is an integer. There are then $L_1L_2\cdots L_d = N$ independent values of $\boldsymbol{q}$, which can be taken to be those corresponding to $m_\mu\in\{1,\ldots,L_\mu\}$.

Let's now expand the function $\Phi(\boldsymbol{\phi})$ in powers of the $\phi_i$, and to first order in the external fields $h_i$. We obtain

$$\Phi = \tfrac{1}{2}\sum_{\boldsymbol{q}}\Big(\hat{K}^{-1}(\boldsymbol{q}) - 1\Big)\,|\hat{\phi}_{\boldsymbol{q}}|^2 + \tfrac{1}{12}\sum_{\boldsymbol{R}}\phi_{\boldsymbol{R}}^4 - \sum_{\boldsymbol{R}} h_{\boldsymbol{R}}\,\phi_{\boldsymbol{R}} + \mathcal{O}\big(\phi^6, h^2\big) \tag{7.408}$$
$$+ \tfrac{1}{2}\,\mathsf{Tr}\;K + \tfrac{1}{2}\,\mathsf{Tr}\,\log(2\pi K) - N\log 2 \quad .$$

On a $d$-dimensional lattice, for a model with nearest neighbor interactions $K_1$ only, we have $\hat{K}(\boldsymbol{q}) = K_1\sum_{\boldsymbol{\delta}} e^{i\boldsymbol{q}\cdot\boldsymbol{\delta}}$, where $\boldsymbol{\delta}$ is a nearest neighbor separation vector. These are the eigenvalues of the matrix $K_{ij}$. We note that $K_{ij}$ is then not positive definite, since there are negative eigenvalues[38]. To fix this, we can add a term $K_0$ everywhere along the diagonal. We then have

$$\hat{K}(\boldsymbol{q}) = K_0 + K_1\sum_{\boldsymbol{\delta}}\cos(\boldsymbol{q}\cdot\boldsymbol{\delta}) \quad . \tag{7.409}$$

[38]To evoke a negative eigenvalue on a $d$-dimensional cubic lattice, set $q_\mu = \pi/a$ for all $\mu$. The eigenvalue is then $-2dK_1$.

Here we have used the inversion symmetry of the Bravais lattice to eliminate the imaginary term. The eigenvalues are all positive so long as $K_0 > zK_1$, where $z$ is the lattice coordination number. We can therefore write $\hat{K}(\boldsymbol{q}) = \hat{K}(0) - \alpha\,\boldsymbol{q}^2$ for small $\boldsymbol{q}$, with $\alpha > 0$. Thus, we can write

$$\hat{K}^{-1}(\boldsymbol{q}) - 1 = a + \kappa\,\boldsymbol{q}^2 + \dots \quad . \tag{7.410}$$

To lowest order in $\boldsymbol{q}$ the RHS is isotropic if the lattice has cubic symmetry, but anisotropy will enter in higher order terms. We'll assume isotropy at this level. This is not necessary but it makes the discussion somewhat less involved. We can now write down our Ginzburg-Landau free energy density:

$$\mathcal{F} = a\,\phi^2 + \tfrac{1}{2}\kappa\,|\boldsymbol{\nabla}\phi|^2 + \tfrac{1}{12}\,\phi^4 - h\,\phi \quad , \tag{7.411}$$

valid to lowest nontrivial order in derivatives, and to fourth order in $\phi$.

One might wonder what we have gained over the inhomogeneous variational density matrix treatment, where we found

$$\begin{aligned} F = &-\tfrac{1}{2}\sum_{\boldsymbol{q}} \hat{J}(\boldsymbol{q})\,|\hat{m}(\boldsymbol{q})|^2 - \sum_{\boldsymbol{q}} \hat{H}(-\boldsymbol{q})\,\hat{m}(\boldsymbol{q}) \\ &+ k_{\rm B}T \sum_{i=1}^{N} \left\{ \left(\frac{1+m_i}{2}\right)\log\left(\frac{1+m_i}{2}\right) + \left(\frac{1-m_i}{2}\right)\log\left(\frac{1-m_i}{2}\right) \right\} \quad . \end{aligned} \tag{7.412}$$

Surely we could expand $\hat{J}(\boldsymbol{q}) = \hat{J}(0) - \frac{1}{2}a\boldsymbol{q}^2 + \dots$ and obtain a similar expression for $\mathcal{F}$. However, such a derivation using the variational density matrix is only approximate. The method outlined in this section is in principle exact.

Let's return to our complete expression for $\Phi$:

$$\Phi(\boldsymbol{\phi}) = \Phi_0(\boldsymbol{\phi}) + \sum_{\boldsymbol{R}} v(\phi_{\boldsymbol{R}}) \quad , \tag{7.413}$$

where

$$\Phi_0(\boldsymbol{\phi}) = \tfrac{1}{2}\sum_{\boldsymbol{q}} G^{-1}(\boldsymbol{q})\,\big|\hat{\phi}(\boldsymbol{q})\big|^2 + \tfrac{1}{2}\,\mathsf{Tr}\left(\frac{1}{1+G^{-1}}\right) + \tfrac{1}{2}\,\mathsf{Tr}\,\log\left(\frac{2\pi}{1+G^{-1}}\right) - N\log 2 \quad . \tag{7.414}$$

Here we have defined

$$v(\phi) = \tfrac{1}{2}\phi^2 - \log\cosh\phi = \tfrac{1}{12}\,\phi^4 - \tfrac{1}{45}\,\phi^6 + \tfrac{17}{2520}\,\phi^8 + \dots \tag{7.415}$$

and

$$G(\boldsymbol{q}) = \frac{\hat{K}(\boldsymbol{q})}{1 - \hat{K}(\boldsymbol{q})} \quad . \tag{7.416}$$

We now want to compute

$$Z = \int D\boldsymbol{\phi}\; e^{-\Phi_0(\boldsymbol{\phi})}\; e^{-\sum_{\boldsymbol{R}} v(\phi_{\boldsymbol{R}})} \tag{7.417}$$

where $D\boldsymbol{\phi} \equiv d\phi_1\,d\phi_2\cdots d\phi_N$. We expand the second exponential factor in a Taylor series, allowing us to write

$$Z = Z_0\left\{1 - \sum_{\boldsymbol{R}} \big\langle v(\phi_{\boldsymbol{R}})\big\rangle + \tfrac{1}{2}\sum_{\boldsymbol{R}}\sum_{\boldsymbol{R}'} \big\langle v(\phi_{\boldsymbol{R}})\,v(\phi_{\boldsymbol{R}'})\big\rangle + \dots\right\} \quad , \tag{7.418}$$

where

$$\begin{aligned} Z_0 &= \int D\boldsymbol{\phi}\, e^{-\Phi_0(\boldsymbol{\phi})} \\ \log Z_0 &= \tfrac{1}{2}\,\mathsf{Tr}\left[\log(1+G) - \frac{G}{1+G}\right] + N\log 2 \end{aligned} \tag{7.419}$$

and

$$\big\langle F(\boldsymbol{\phi})\big\rangle = \frac{\int D\boldsymbol{\phi}\, F\, e^{-\Phi_0}}{\int D\boldsymbol{\phi}\, e^{-\Phi_0}} \quad . \tag{7.420}$$

To evaluate the various terms in the expansion of eqn. 7.418, we invoke Wick's theorem, which says

$$\begin{aligned} \big\langle x_{i_1}\, x_{i_2}\cdots x_{i_{2L}}\big\rangle &= \int\limits_{-\infty}^{\infty} dx_1\cdots\int\limits_{-\infty}^{\infty} dx_N\, e^{-\frac{1}{2}\mathcal{G}^{-1}_{ij}x_ix_j}\, x_{i_1}\, x_{i_2}\cdots x_{i_{2L}} \Bigg/ \int\limits_{-\infty}^{\infty} dx_1\cdots\int\limits_{-\infty}^{\infty} dx_N\, e^{-\frac{1}{2}\mathcal{G}^{-1}_{ij}x_ix_j} \\ &= \sum_{\substack{\text{all distinct}\\ \text{pairings}}} \mathcal{G}_{j_1j_2}\mathcal{G}_{j_3j_4}\cdots\mathcal{G}_{j_{2L-1}j_{2L}} \quad , \end{aligned} \tag{7.421}$$

where the sets $\{j_1,\ldots,j_{2L}\}$ are all permutations of the set $\{i_1,\ldots,i_{2L}\}$. In particular, $\big\langle x_i^4\big\rangle = 3\big(\mathcal{G}_{ii}\big)^2$, so

$$\big\langle \phi^4_{\boldsymbol{R}}\big\rangle = 3\left(\frac{1}{N}\sum_{\boldsymbol{q}} G(\boldsymbol{q})\right)^2 \quad . \tag{7.422}$$

Thus, if we write $v(\phi) \approx \frac{1}{12}\,\phi^4$ and retain only the quartic term in $v(\phi)$, we obtain

$$\begin{aligned} \frac{F}{k_{\rm B}T} = -\log Z_0 &= \tfrac{1}{2}\,\mathsf{Tr}\left[\frac{G}{1+G} - \log(1+G)\right] + \frac{1}{4N}\big(\mathsf{Tr}\;G\big)^2 - N\log 2 \\ &= -N\log 2 + \frac{1}{4N}\big(\mathsf{Tr}\;G\big)^2 - \frac{1}{4}\,\mathsf{Tr}\;\big(G^2\big) + \mathcal{O}\big(G^3\big) \quad . \end{aligned} \tag{7.423}$$

Note that if we set $K_{ij}$ to be diagonal, then $\hat{K}(\boldsymbol{q})$ and $G(\boldsymbol{q})$ are constant functions of $\boldsymbol{q}$. The $\mathcal{O}\big(G^2\big)$ term then vanishes, which is required since the free energy cannot depend on the diagonal elements of $K_{ij}$.

# Chapter 8

# The Boltzmann Equation

## 8.1 References

- H. Smith and H. H. Jensen, *Transport Phenomena* (Oxford, 1989)
  An outstanding, thorough, and pellucid presentation of the theory of Boltzmann transport in classical and quantum systems.
- P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics* (Cambridge, 2010)
  Superb, modern discussion of a broad variety of issues and models in nonequilibrium statistical physics.
- E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Pergamon, 1981)
  Volume 10 in the famous Landau and Lifshitz *Course of Theoretical Physics*. Surprisingly readable, and with many applications (some advanced).
- M. Kardar, *Statistical Physics of Particles* (Cambridge, 2007)
  A superb modern text, with many insightful presentations of key concepts. Includes a very instructive derivation of the Boltzmann equation starting from the BBGKY hierarchy.
- J. A. McLennan, *Introduction to Non-equilibrium Statistical Mechanics* (Prentice-Hall, 1989)
  Though narrow in scope, this book is a good resource on the Boltzmann equation.
- F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, 1987)
  This has been perhaps the most popular undergraduate text since it first appeared in 1967, and with good reason. The later chapters discuss transport phenomena at an undergraduate level.
- N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* ($3^{\rm rd}$ edition, North-Holland, 2007)
  This is a very readable and useful text. A relaxed but meaty presentation.

## 8.2 Equilibrium, Nonequilibrium and Local Equilibrium

Classical equilibrium statistical mechanics is described by the full $N$-body distribution,

$$\varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N;\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N) = \frac{1}{N!}\times\begin{cases} Z_N^{-1}\,e^{-\beta\hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} & \text{OCE} \\ \Xi^{-1}\,e^{\beta\mu N}\,e^{-\beta\hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} & \text{GCE} \end{cases}, \tag{8.1}$$

We assume a Hamiltonian of the form

$$\hat{H}_N = \sum_{i=1}^N \frac{\boldsymbol{p}_i^2}{2m} + \sum_{i=1}^N v(\boldsymbol{x}_i) + \sum_{i<j}^N u(\boldsymbol{x}_i-\boldsymbol{x}_j). \tag{8.2}$$

Here $v(\boldsymbol{x}) = U_{\rm ext}(\boldsymbol{x})$ is due to external forces. In the context of transport theory, $v(\boldsymbol{x})$ typically will denote the effect of an applied external field, *e.g.* $v(\boldsymbol{x}) = -q\boldsymbol{E}\cdot\boldsymbol{x}$ for a particle of charge $q$ in the presence of a uniform electric field $\boldsymbol{E}$. We write $\boldsymbol{x}_i$ and $\boldsymbol{p}_i$ for the corresponding phase space variables, the position and momentum vectors for the $i^{\rm th}$ particle, respectively. The quantity

$$\varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N)\prod_{j=1}^N d\mu_j \quad , \tag{8.3}$$

with $d\mu_j \equiv d^dx_j\,d^dp_j/h^d$, is the probability, under equilibrium conditions, of finding $N$ particles in the system, with particle #1 lying within $d^dx_1$ of $\boldsymbol{x}_1$ and having momentum within $d^dp_1$ of $\boldsymbol{p}_1$, *etc.* The temperature $T$ and chemical potential $\mu$ are constants, independent of position. Note that $\varrho_N(\{\boldsymbol{x}_i\},\{\boldsymbol{p}_i\})$ has units of $h^{-dN}$.

Nonequilibrium statistical mechanics seeks to describe thermodynamic systems which are out of equilibrium, meaning that the distribution function is not given by the Boltzmann distribution above. Rather, it is a time-dependent quantity, $\varrho_N$. In a general nonequilibrium setting, it is hopeless to make progress – we'd have to integrate the equations of motion for all the constituent particles. However, typically we are concerned with situations where external forces or constraints are imposed over some macroscopic scale. Examples would include the imposition of a voltage drop across a metal, or a temperature differential across any thermodynamic sample. In such cases, scattering at microscopic length and time scales described by the *mean free path* $\ell$ and the *collision time* $\tau$ work to establish *local equilibrium* throughout the system. A local equilibrium is a state described by a space and time varying temperature $T(\boldsymbol{r},t)$ and chemical potential $\mu(\boldsymbol{r},t)$. As we will see, the Boltzmann distribution with $T = T(\boldsymbol{r},t)$ and $\mu = \mu(\boldsymbol{r},t)$ will not be a solution to the evolution equation governing the distribution function. Rather, the distribution for systems slightly out of equilibrium will be of the form $f = f^0 + \delta f$, where $f^0$ describes a state of local equilibrium.

We will mainly be interested in the one-body distribution

$$\begin{aligned} f(\mathbf{x},\mathbf{p};t) &= \sum_{i=1}^N \big\langle\,\delta\big(\mathbf{x}-\boldsymbol{x}_i(t)\big)\,\delta\big(\mathbf{p}-\boldsymbol{p}_i(t)\big)\,\big\rangle \\ &= Nh^{-d}\int\prod_{i=2}^N d\mu_i\;\varrho_N(\mathbf{x},\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N;\mathbf{p},\boldsymbol{p}_2,\ldots,\boldsymbol{p}_N,t) \quad . \end{aligned} \tag{8.4}$$

Note that $f(\mathbf{x},\mathbf{p},t)$ has dimensions of $h^{-d}$, and $f(\mathbf{x},\mathbf{p},t)\,d^d\mathrm{x}\,d^d\mathrm{p}$ is the average number of particles found within $d^d\mathrm{x}$ of $\mathbf{x}$ and $d^d\mathrm{p}$ of $\mathbf{p}$ at time $t$. The full $k$-particle distributions, are given by

$$\begin{aligned} f_k\big(\{\mathbf{x}_i\},\{\mathbf{p}_i\},t\big) &= {\sum_{\{j_1,\ldots,j_k\}}}' \big\langle \delta\big(\mathbf{x}_1 - \boldsymbol{x}_{j_1}(t)\big)\cdots\delta\big(\mathbf{x}_k - \boldsymbol{x}_{j_k}(t)\big)\,\delta\big(\mathbf{p}_1 - \boldsymbol{p}_{j_1}(t)\big)\cdots\delta\big(\mathbf{p}_k - \boldsymbol{p}_{j_k}(t)\big)\big\rangle \\ &= \frac{N!\,h^{-kd}}{(N-k)!}\int\prod_{i=k+1}^{N} d\mu_i\;\varrho_N\big(\mathbf{x}_1,\ldots,\mathbf{x}_k,\boldsymbol{x}_{k+1},\ldots,\boldsymbol{x}_N;\mathbf{p}_1,\ldots,\mathbf{p}_k,\boldsymbol{p}_{k+1},\ldots,\boldsymbol{p}_N\,,\,t\big) \quad , \end{aligned} \tag{8.5}$$

where $\{\mathbf{x}_j\}$ as an argument of the $k$-particle distribution $f_k$ denotes the ordered set $\{\mathbf{x}_1,\ldots,\mathbf{x}_k\}$, where $k\in\{1,\ldots,N\}$ (similarly for $\{\mathbf{p}_j\}$). The prime on the sum over the indices $\{j_1,\ldots,j_k\}$ indicates that no two indices take the same value. Note that the normalization of $f_k$ is

$$\int\prod_{i=1}^{k} d^d\mathrm{x}_i\,d^d\mathrm{p}_i\;f_k\big(\{\mathbf{x}_j\},\{\mathbf{p}_j\},t\big) = \frac{N!}{(N-k)!} \quad , \tag{8.6}$$

for all $t$. We write the one-body distribution $f_1(\mathbf{x}_1,\mathbf{p}_1,t)\equiv f(\mathbf{x},\mathbf{p},t)$, where $\mathbf{x}=\mathbf{x}_1$ and $\mathbf{p}=\mathbf{p}_1$. Henceforth we will drop the distinction between boldface Roman symbols $\{\mathbf{x},\mathbf{p}\}$ and boldface math italic $\{\boldsymbol{x},\boldsymbol{p}\}$, as it will no longer be needed.

In the GCE, we sum over different particle numbers $N$. Assuming $v=0$ so that there is no one-body potential to break translational symmetry, the equilibrium distribution is time-independent and space-independent:

$$f^0(\boldsymbol{r},\boldsymbol{p}) = n\,(2\pi m k_{\mathrm{B}}T)^{-3/2}\,e^{-\boldsymbol{p}^2/2mk_{\mathrm{B}}T} = h^{-d}\,z\,e^{-\boldsymbol{p}^2/2mk_{\mathrm{B}}T} \quad , \tag{8.7}$$

where $n=N/V$ or $n=n(T,\mu)$ is the particle density in the OCE or GCE, with $z=\exp(\mu/k_{\mathrm{B}}T)=n\lambda_T^d$ the fugacity. From the one-body distribution we can compute things like the particle current density $\boldsymbol{j}$ and the energy current density $\boldsymbol{j}_\varepsilon$:

$$\boldsymbol{j}(\boldsymbol{r},t) = \int d^dp\;f(\boldsymbol{r},\boldsymbol{p};t)\,\frac{\boldsymbol{p}}{m} \tag{8.8}$$

$$\boldsymbol{j}_\varepsilon(\boldsymbol{r},t) = \int d^dp\;f(\boldsymbol{r},\boldsymbol{p};t)\;\varepsilon(\boldsymbol{p})\,\frac{\boldsymbol{p}}{m} \quad , \tag{8.9}$$

where $\varepsilon(\boldsymbol{p})=\boldsymbol{p}^2/2m$. Both quantities vanish in equilibrium, when $f=f^0$, since $f^0(\boldsymbol{r},\boldsymbol{p})$ depends only on $\boldsymbol{p}^2$ and not on the direction of $\boldsymbol{p}$. In a *steady state* nonequilibrium situation, the above quantities are time-independent.

Thermodynamics says that

$$T\,dS = dE + p\,dV - \mu\,dN \quad , \tag{8.10}$$

which may be cast as one among the corresponding current densities. Suppressing the explicit $\boldsymbol{r}$ and $t$ dependence,

$$T\,\boldsymbol{j}_s = \boldsymbol{j}_\varepsilon - \mu\,\boldsymbol{j} \quad , \tag{8.11}$$

where $\boldsymbol{j}_s$ is the entropy current density. Some authors define the *conductive heat current density* $\boldsymbol{j}_q$ (sometimes written as $\boldsymbol{q}$) as $\boldsymbol{j}_q=\boldsymbol{j}_\varepsilon - h\boldsymbol{j}$, where $h$ is the enthalpy per particle. Since $h=\mu+Ts$, where $s$ is the

entropy per particle, we have the relation $\boldsymbol{j}_q = T\boldsymbol{j}_s - Ts\boldsymbol{j}$. To make matters even more confusing, some sources (particularly in condensed matter physics) define a "thermodynamic heat current density" as $\boldsymbol{j}_{\rm th} \equiv \boldsymbol{j}_\varepsilon - \mu\,\boldsymbol{j} = T\boldsymbol{j}_s$. When there is no particle current and $\boldsymbol{j} = 0$, we have $\boldsymbol{j}_\varepsilon = \boldsymbol{j}_q = \boldsymbol{j}_{\rm th} = T\boldsymbol{j}_s$.

When the individual particles are not point particles, they possess angular momentum as well as linear momentum. Following Lifshitz and Pitaevskii, we abbreviate $\Gamma = (\boldsymbol{p}, \boldsymbol{L})$ for these two variables for the case of diatomic molecules, and $\Gamma = (\boldsymbol{p}, \boldsymbol{L}, \hat{\boldsymbol{n}} \cdot \boldsymbol{L})$ in the case of spherical top molecules, where $\hat{\boldsymbol{n}}$ is the symmetry axis of the top. We then have, in $d = 3$ dimensions,

$$d\Gamma = \begin{cases} d^3p & \text{point particles} \\ d^3p\, L\, dL\, d\Omega_{\hat{\boldsymbol{L}}} & \text{diatomic molecules} \\ d^3p\, L^2\, dL\, d\Omega_{\hat{\boldsymbol{L}}}\, d\cos\vartheta & \text{symmetric tops} \quad , \end{cases} \tag{8.12}$$

where $\vartheta = \cos^{-1}(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{L}})$. We will call the set $\Gamma$ the 'kinematic variables'. The instantaneous number density at $\boldsymbol{r}$ is then

$$n(\boldsymbol{r}, t) = \int d\Gamma\, f(\boldsymbol{r}, \Gamma; t) \quad . \tag{8.13}$$

One might ask why we do not also keep track of the angular orientation of the individual molecules. There are two reasons. First, the rotations of the molecules are generally extremely rapid, so we are justified in averaging over these motions. Second, the orientation of, say, a rotor does not enter into its energy. While the same can be said of the spatial position in the absence of external fields, (i) in the presence of external fields one must keep track of the position coordinate $\boldsymbol{r}$ since there is physical transport of particles from one region of space to another, and (ii) the collision process, which as we shall see enters the dynamics of the distribution function, takes place in real space.

## 8.3 Boltzmann Transport Theory

### 8.3.1 Derivation of the Boltzmann equation

For simplicity of presentation, we assume point particles in three space dimensions. Recall that

$$f(\boldsymbol{r}, \boldsymbol{p}, t)\, d^3r\, d^3p \equiv \begin{cases} \text{\# of particles with positions within } d^3r \text{ of} \\ \boldsymbol{r} \text{ and momenta within } d^3p \text{ of } \boldsymbol{p} \text{ at time } t. \end{cases} \tag{8.14}$$

Thus, the units of $f(\boldsymbol{r}, \boldsymbol{p}, t)$ are those of inverse action cubed, *i.e.* $(ML^2/T)^{-3}$. We now ask how the distribution function $f(\boldsymbol{r}, \boldsymbol{p}, t)$ evolves in time. It is clear that in the absence of collisions, the distribution function must satisfy the continuity equation,

$$\frac{\partial f}{\partial t} + \boldsymbol{\nabla} \cdot (\boldsymbol{u} f) = 0 \quad . \tag{8.15}$$

This is just the condition of number conservation for particles. Take care to note that $\boldsymbol{\nabla}$ and $\boldsymbol{u}$ are *six*-dimensional *phase space* vectors:

$$\begin{aligned} \boldsymbol{u} &= (\ \dot{x} \quad , \ \dot{y} \quad , \ \dot{z} \quad , \ \dot{p}_x \quad , \ \dot{p}_y \quad , \ \dot{p}_z\ ) \\ \boldsymbol{\nabla} &= \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial p_x}, \frac{\partial}{\partial p_y}, \frac{\partial}{\partial p_z} \right) \quad . \end{aligned} \tag{8.16}$$

The continuity equation describes a distribution in which each constituent particle evolves according to a prescribed dynamics, which for a mechanical system is specified by

$$\frac{d\boldsymbol{r}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}} = \boldsymbol{v}(\boldsymbol{p}) \qquad , \qquad \frac{d\boldsymbol{p}}{dt} = -\frac{\partial H}{\partial \boldsymbol{r}} = \boldsymbol{F} \quad , \tag{8.17}$$

where $\boldsymbol{F}$ is an external applied force. Here,

$$H(\boldsymbol{p}, \boldsymbol{r}) = \varepsilon(\boldsymbol{p}) + U(\boldsymbol{r}) \quad . \tag{8.18}$$

For example, under the influence of gravity, $U(\boldsymbol{r}) = m\boldsymbol{g} \cdot \boldsymbol{r}$ and $\boldsymbol{F} = -\boldsymbol{\nabla} U = -m\boldsymbol{g}$.

Note that as a consequence of the dynamics, we have $\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$, *i.e.* phase space flow is *incompressible*, provided that $\varepsilon(\boldsymbol{p})$ is a function of $\boldsymbol{p}$ alone, and not of $\boldsymbol{r}$. Thus, in the absence of collisions, we have

$$\frac{\partial f}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} f = 0 \quad . \tag{8.19}$$

The differential operator $D_t \equiv \partial_t + \boldsymbol{u} \cdot \boldsymbol{\nabla}$ is sometimes called the 'convective derivative', because $D_t f$ is the time derivative of $f$ in a comoving frame of reference. To see this, evaluate the time derivative of the time-dependent distribution $f(\boldsymbol{r}, \boldsymbol{p}, t)$ along a particle trajectory, *i.e.* substituting $\boldsymbol{r} \to \boldsymbol{r}(t)$ and $\boldsymbol{p} = \boldsymbol{p}(t)$. We have

$$\frac{d}{dt} f\big(\boldsymbol{r}(t), \boldsymbol{p}(t), t\big) = \frac{\partial f}{\partial \boldsymbol{r}} \cdot \frac{d\boldsymbol{r}}{dt}\bigg|_{\{\boldsymbol{r}(t), \boldsymbol{p}(t), t\}} + \frac{\partial f}{\partial \boldsymbol{p}} \cdot \frac{d\boldsymbol{p}}{dt}\bigg|_{\{\boldsymbol{r}(t), \boldsymbol{p}(t), t\}} + \frac{\partial f}{\partial t}\bigg|_{\{\boldsymbol{r}(t), \boldsymbol{p}(t), t\}} = \frac{Df}{Dt}\bigg|_{\{\boldsymbol{r}(t), \boldsymbol{p}(t), t\}} \quad . \tag{8.20}$$

Next we must consider the effect of collisions, which are not accounted for by the semiclassical dynamics. In a collision process, a particle with momentum $\boldsymbol{p}$ and one with momentum $\tilde{\boldsymbol{p}}$ can instantaneously convert into a pair with momenta $\boldsymbol{p}'$ and $\tilde{\boldsymbol{p}}'$, provided total momentum is conserved: $\boldsymbol{p} + \tilde{\boldsymbol{p}} = \boldsymbol{p}' + \tilde{\boldsymbol{p}}'$. This means that $D_t f \neq 0$. Rather, we should write

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \dot{\boldsymbol{r}} \cdot \frac{\partial f}{\partial \boldsymbol{r}} + \dot{\boldsymbol{p}} \cdot \frac{\partial f}{\partial \boldsymbol{p}} = \left(\frac{df}{dt}\right)_{\rm coll} \tag{8.21}$$

where the right side is known as the *collision integral*. The collision integral is in general a *function* of $\boldsymbol{r}$, $\boldsymbol{p}$, and $t$ and a *functional* of the distribution $f$. *Thus, in the absence of collisions, the convective derivative of the distribution $f(\boldsymbol{r}, \boldsymbol{p}, t)$ vanishes, meaning that the one-body distribution does not vary in time along a particle trajectory.*

We can write the Boltzmann equation as

$$\frac{\partial f}{\partial t} + \left(\frac{df}{dt}\right)_{\rm str} = \left(\frac{df}{dt}\right)_{\rm coll} \quad , \tag{8.22}$$

where

$$\left(\frac{df}{dt}\right)_{\rm str} \equiv \dot{\boldsymbol{r}} \cdot \frac{\partial f}{\partial \boldsymbol{r}} + \dot{\boldsymbol{p}} \cdot \frac{\partial f}{\partial \boldsymbol{p}} \tag{8.23}$$

is known as the *streaming term*. Thus, there are two contributions to $\partial f / \partial t$ : streaming and collisions.

### 8.3.2 Collisionless Boltzmann equation

In the absence of collisions, the Boltzmann equation is given by

$$\frac{\partial f}{\partial t} + \frac{\partial \varepsilon}{\partial \boldsymbol{p}} \cdot \frac{\partial f}{\partial \boldsymbol{r}} - \boldsymbol{\nabla} U \cdot \frac{\partial f}{\partial \boldsymbol{p}} = 0 \quad . \tag{8.24}$$

In order to gain some intuition about how the streaming term affects the evolution of the distribution $f(\boldsymbol{r}, \boldsymbol{p}, t)$, consider a case where $\boldsymbol{F} = 0$. We then have

$$\frac{\partial f}{\partial t} + \boldsymbol{v}(\boldsymbol{p}) \cdot \frac{\partial f}{\partial \boldsymbol{r}} = 0 \quad . \tag{8.25}$$

Clearly, then, *any* function of the form

$$f(\boldsymbol{r}, \boldsymbol{p}, t) = \varphi\big(\boldsymbol{r} - \boldsymbol{v}(\boldsymbol{p})\, t\,,\, \boldsymbol{p}\big) \tag{8.26}$$

will be a solution to the collisionless Boltzmann equation, where $\boldsymbol{v}(\boldsymbol{p}) = \partial\varepsilon/\partial\boldsymbol{p}$. One possible solution would be the Boltzmann distribution,

$$f(\boldsymbol{r}, \boldsymbol{p}, t) = C\, e^{\mu/k_{\rm B}T} e^{-\varepsilon(\boldsymbol{p})/k_{\rm B}T} \quad , \tag{8.27}$$

which is time-independent[1]. Here, $C$ is a constant with units of $(T/ML^2)^d$. For a ballistic dispersion, $\varepsilon(\boldsymbol{p}) = \boldsymbol{p}^2/2m$ and $C = h^{-3}$, in which case $\int d^dr\, d^dp\, f(\boldsymbol{r}, \boldsymbol{p}) = V \exp(\mu/k_{\rm B}T) = \langle \hat{N} \rangle$.

For a slightly less trivial example, let the initial distribution be $\varphi(\boldsymbol{r}, \boldsymbol{p}) = C\, e^{-\boldsymbol{r}^2/2\sigma^2} e^{-\boldsymbol{p}^2/2\kappa^2}$, so that

$$f(\boldsymbol{r}, \boldsymbol{p}, t) = C\, e^{-(\boldsymbol{r} - \frac{\boldsymbol{p}t}{m})^2/2\sigma^2}\, e^{-\boldsymbol{p}^2/2\kappa^2} \quad . \tag{8.28}$$

Consider the one-dimensional version, and rescale position, momentum, and time so that

$$f(x, p, t) = C\, e^{-(\bar{x} - \bar{p}\,\bar{t})^2/2}\, e^{-\bar{p}^2/2} \quad . \tag{8.29}$$

Now consider the level sets of $f$, where $f(x, p, t) = C\, e^{-\alpha^2/2}$. The equation for these sets is

$$\bar{x} = \bar{p}\,\bar{t} \pm \sqrt{\alpha^2 - \bar{p}^2} \quad . \tag{8.30}$$

For fixed $\bar{t}$, these level sets describe the loci in phase space of equal probability densities, with the probability density decreasing exponentially in the parameter $\alpha^2$. For $\bar{t} = 0$, the initial distribution describes a Gaussian cloud of particles with a Gaussian momentum distribution. As $\bar{t}$ increases, the distribution widens in $\bar{x}$ but not in $\bar{p}$ – each particle moves with a constant momentum, so the set of momentum values never changes. However, the level sets in the $(\bar{x}\,,\,\bar{p})$ plane become elliptical, with a semimajor axis oriented at an angle $\theta = \frac{1}{2}\tan^{-1}(2/\bar{t})$ with respect to the $\bar{x}$ axis. For $\bar{t} > 0$, the particles at the outer edges of the cloud are more likely to be moving away from the center. See fig. 8.1.

Suppose we add in a constant external force $\boldsymbol{F}$. Then it is easy to show (and left as an exercise to the reader to prove) that any function of the form

$$f(\boldsymbol{r}, \boldsymbol{p}, t) = \varphi\left(\boldsymbol{r} - \frac{\boldsymbol{p}\,t}{m} - \frac{\boldsymbol{F}t^2}{2m}\,,\, \boldsymbol{p} - \boldsymbol{F}t\right) \tag{8.31}$$

satisfies the collisionless Boltzmann equation (ballistic dispersion assumed).

[1] Indeed, any arbitrary function of $\boldsymbol{p}$ alone would be a solution. Ultimately, we require some energy exchanging processes, such as collisions, in order for any initial nonequilibrium distribution to converge to the Boltzmann distribution.

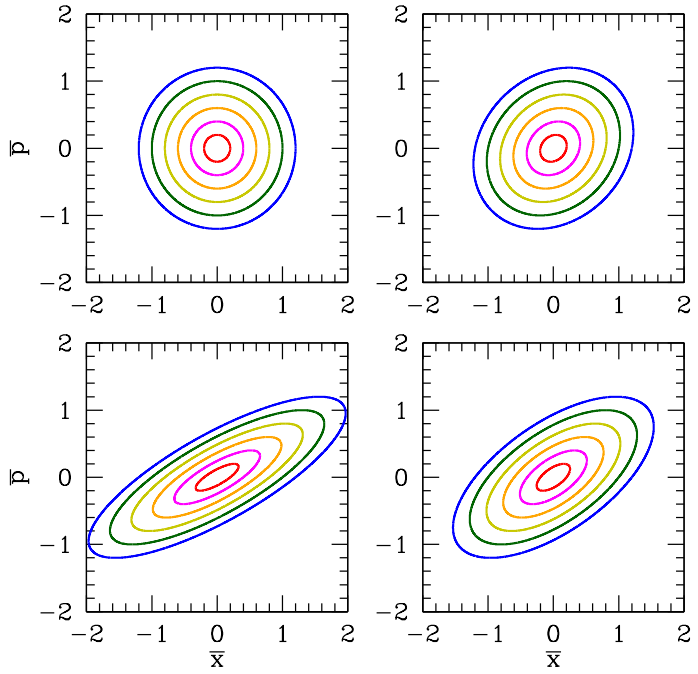



Figure 8.1: Level sets for a sample $f(\bar{x},\bar{p},\bar{t}) = C\,e^{-(\bar{x}-\bar{p}\bar{t})^2/2}e^{-\bar{p}^2/2}$, for values $f = C\,e^{-\alpha^2/2}$ with $\alpha$ in equally spaced intervals from $\alpha = 0.2$ (red) to $\alpha = 1.2$ (blue). The time variable $\bar{t}$ is taken to be $\bar{t} = 0.0$ (upper left), $0.2$ (upper right), $0.8$ (lower right), and $1.3$ (lower left).

### 8.3.3 Scattering processes

What sort of processes contribute to the collision integral? There are two broad classes to consider. The first involves potential scattering, where a particle in state $|\Gamma\rangle$ scatters, in the presence of an external potential, to a state $|\Gamma'\rangle$. Recall that $\Gamma$ is an abbreviation for the set of kinematic variables, *e.g.* $\Gamma = (\boldsymbol{p}, \boldsymbol{L})$ in the case of a diatomic molecule. For point particles, $\Gamma = (p_x, p_y, p_z)$ and $d\Gamma = d^3p$.

#### Single particle scattering

We now define the function $w\big(\Gamma' \,|\, \Gamma\big)$ such that

$$w\big(\Gamma' \,|\, \Gamma\big)\, f(\boldsymbol{r}, \Gamma; t)\, d\Gamma\, d\Gamma' = \text{rate per unit volume to scatter } \big\{|\,\Gamma + d\Gamma\,\rangle\big\} \to \big\{|\,\Gamma' + d\Gamma'\,\rangle\big\} \tag{8.32}$$

at time $t$. Here we have suppressed the time dependence in the distribution $f(\Gamma, t)$, and typically we will presume $w\big(\Gamma' \,|\, \Gamma\big)$ to be independent of both $\boldsymbol{r}$ and $t$. By $\big\{|\,\Gamma + d\Gamma\,\rangle\big\}$ we mean states with momenta within $d^3p$ of $\boldsymbol{p}$ (more generally, within $d\Gamma = d^3p\, d^3L$ of $(\boldsymbol{p}, \boldsymbol{L})$) and at the same position coordinate $\boldsymbol{r}$. We assume the rate is independent of the position $\boldsymbol{r}$ and the time $t$. The units of $w\, d\Gamma$ are therefore $1/T$. The *differential scattering cross section* for single particle scattering is then

$$d\sigma = \frac{w\big(\Gamma' \,|\, \Gamma\big)}{n_{\rm s}\,|\boldsymbol{v}|}\, d\Gamma' \quad , \tag{8.33}$$

where $\boldsymbol{v} = \boldsymbol{p}/m$ is the particle's velocity and $n_{\rm s}$ the number density of scatterers.

In computing the collision integral for the state $|\boldsymbol{r}, \Gamma\rangle$, we must take care to sum over contributions from transitions *out of* this state, *i.e.* $|\Gamma\rangle \to |\Gamma'\rangle$, which reduce $f(\boldsymbol{r}, \Gamma)$, and transitions *into* this state, *i.e.* $|\Gamma'\rangle \to |\Gamma\rangle$, which increase $f(\boldsymbol{r}, \Gamma)$. Thus, for one-body scattering, we have

$$\frac{D}{Dt} f(\boldsymbol{r}, \Gamma; t) = \int d\Gamma' \left\{ w(\Gamma \,|\, \Gamma')\, f(\boldsymbol{r}, \Gamma'; t) - w(\Gamma' \,|\, \Gamma)\, f(\boldsymbol{r}, \Gamma; t) \right\} = \left( \frac{df}{dt} \right)_{\rm coll} \quad . \tag{8.34}$$

Mathematically, in the case of point particles, where $\Gamma = \boldsymbol{p}$, this is a linear, first-order integro-differential equation in seven variables $\{x, y, z, p_x, p_y, p_z, t\}$. It can be expressed as $\partial_t f = \mathcal{L} f$, where $\mathcal{L}$ is a linear integro-differential operator. With initial conditions $f(\boldsymbol{r}, \boldsymbol{p}, t = 0)$ this can in principle be integrated forward in time, and in the autonomous case where all the rates and external forces are time-independent, we have $f(t) = \exp(\mathcal{L}t)\, f(0)$.

### Two particle scattering

The second class is that of two-particle scattering processes, *i.e.* $|\,\{\boldsymbol{r}, \Gamma\}, \{\boldsymbol{r}, \Gamma_1\}\,\rangle \to |\,\{\boldsymbol{r}, \Gamma'\}, \{\boldsymbol{r}, \Gamma_1'\}\,\rangle$. We define the scattering function $w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big)$ by

$$w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big)\, f_2(\boldsymbol{r}, \Gamma; \boldsymbol{r}, \Gamma_1; t)\, d\Gamma\, d\Gamma_1\, d\Gamma'\, d\Gamma_1' = \begin{cases} \text{rate per unit volume to scatter two particles} \\ \big\{ |\,\Gamma + d\Gamma, \Gamma_1 + d\Gamma_1\,\rangle \big\} \to \big\{ |\,\Gamma' + d\Gamma', \Gamma_1' + d\Gamma_1'\,\rangle \big\} \end{cases} \tag{8.35}$$

at time $t$. We assume that $w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big)$ is independent of $\boldsymbol{r}$ and $t$. Thus the units of $w\, d\Gamma\, d\Gamma_1$ are $L^d/T$, and the differential scattering cross section is

$$d\sigma = \frac{w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big)}{|\boldsymbol{v} - \boldsymbol{v}_1|}\, d\Gamma'\, d\Gamma_1' \quad . \tag{8.36}$$

For two-body scattering, we therefore have

$$\begin{aligned} \frac{D}{Dt} f(\boldsymbol{r}, \Gamma; t) = \int d\Gamma_1 \int d\Gamma' \int d\Gamma_1' \Big\{ & w\big(\Gamma, \Gamma_1 \,\big|\, \Gamma', \Gamma_1'\big)\, f_2(\boldsymbol{r}, \Gamma'; \boldsymbol{r}, \Gamma_1'; t) \\ & - w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big)\, f_2(\boldsymbol{r}, \Gamma; \boldsymbol{r}, \Gamma_1; t) \Big\} = \left( \frac{df}{dt} \right)_{\rm coll} \quad . \end{aligned} \tag{8.37}$$

Unlike the one-body scattering case, the kinetic equation for two-body scattering does not close, since the LHS involves the one-body distribution $f \equiv f_1$ and the RHS involves the two-body distribution $f_2$. One can then write an equation for $f_2$ which also does not close – it involves $f_3$ – and so on. The result is known as the *BBGKY hierarchy*, and is discussed in some detail in the appendix §8.10.

In practice, an approximation is made in order to close the equations. The simplest and most common such approximation, known as the *Stosszahlansatz*, or *molecular chaos assumption*, is due to Boltzmann himself. We replace the two-body distribution by a product of one-body distributions, *viz.*

$$f_2(\boldsymbol{r}, \Gamma; \boldsymbol{r}_1, \Gamma_1; t) \approx f(\boldsymbol{r}, \Gamma; t)\, f(\boldsymbol{r}_1, \Gamma_1; t) \quad . \tag{8.38}$$

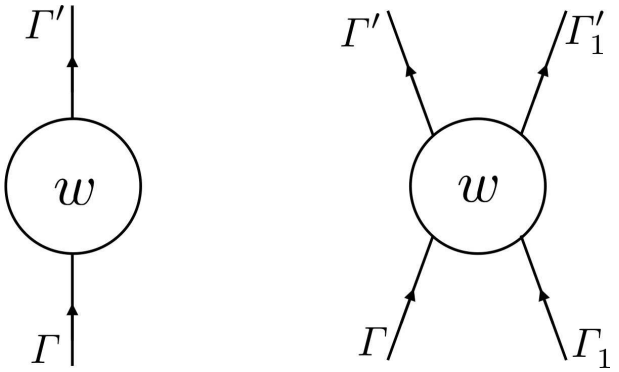



Figure 8.2: Left: single particle scattering process $|\Gamma\rangle \to |\Gamma'\rangle$. Right: two-particle scattering process $|\Gamma, \Gamma_1\rangle \to |\Gamma', \Gamma_1'\rangle$.

We then have

$$\frac{D}{Dt} f(\boldsymbol{r}, \Gamma; t) = \int d\Gamma_1 \int d\Gamma' \int d\Gamma_1' \Big\{ w\big(\Gamma, \Gamma_1 \,\big|\, \Gamma', \Gamma_1'\big)\, f(\boldsymbol{r}, \Gamma'; t)\, f(\boldsymbol{r}, \Gamma_1'; t) - w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big)\, f(\boldsymbol{r}, \Gamma; t)\, f(\boldsymbol{r}, \Gamma_1; t) \Big\} = \left(\frac{df}{dt}\right)_{\rm coll} \quad . \tag{8.39}$$

Now we have a closed but *nonlinear* first-order integro-PDE. The nonlinearity adds a new wrinkle, and as we shall see below a further approximation involves *linearizing* the collision integral.

We stress that in both cases (*i.e.* single-particle and two-particle scattering) we assume that all scattering occurs *locally*, *i.e.* the particles attain their asymptotic kinematic states on distance scales small compared to the mean interparticle separation. In this case we can treat each scattering process independently. This assumption is particular to rarefied systems, *i.e.* gases, and is not appropriate for dense liquids. The two types of scattering processes are depicted in fig. 8.2.

*Nota bene:* All $w(\Gamma'|\Gamma)$ and all $w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big)$ are non-negative!

### 8.3.4 Collisional invariants

Consider a function $A(\boldsymbol{r}, \boldsymbol{p})$ of position and momentum. Its average value at time $t$ is

$$\langle A(t) \rangle = \int d^3r\, d^3p\, A(\boldsymbol{r}, \boldsymbol{p})\, f(\boldsymbol{r}, \boldsymbol{p}, t) \quad . \tag{8.40}$$

Taking the time derivative,

$$\begin{aligned} \frac{d\langle A(t)\rangle}{dt} &= \int d^3r \int d^3p\, A(\boldsymbol{r}, \boldsymbol{p})\, \frac{\partial f}{\partial t} \\ &= \int d^3r \int d^3p\, A(\boldsymbol{r}, \boldsymbol{p}) \left\{ -\frac{\partial}{\partial \boldsymbol{r}} \cdot (\dot{\boldsymbol{r}} f) - \frac{\partial}{\partial \boldsymbol{p}} \cdot (\dot{\boldsymbol{p}} f) + \left(\frac{df}{dt}\right)_{\rm coll} \right\} \\ &= \int d^3r \int d^3p \left\{ \left( \frac{\partial A}{\partial \boldsymbol{r}} \cdot \frac{d\boldsymbol{r}}{dt} + \frac{\partial A}{\partial \boldsymbol{p}} \cdot \frac{d\boldsymbol{p}}{dt} \right) f + A(\boldsymbol{r}, \boldsymbol{p}) \left(\frac{df}{dt}\right)_{\rm coll} \right\} \quad . \end{aligned} \tag{8.41}$$

Hence, if $A$ is preserved by the dynamics between collisions, then[2]

$$\frac{dA(\boldsymbol{r}(t),\boldsymbol{p}(t))}{dt} = \frac{\partial A}{\partial \boldsymbol{r}}\cdot\frac{d\boldsymbol{r}}{dt} + \frac{\partial A}{\partial \boldsymbol{p}}\cdot\frac{d\boldsymbol{p}}{dt} = 0 \quad . \tag{8.42}$$

We therefore have that the rate of change of $\langle A(t)\rangle$ is determined wholly by the collision integral

$$\frac{d\langle A(t)\rangle}{dt} = \int d^3r \int d^3p\, A(\boldsymbol{r},\boldsymbol{p}) \left(\frac{df}{dt}\right)_{\rm coll} \quad . \tag{8.43}$$

Quantities which are then conserved in the collisions satisfy $\dot{A} = 0$. Such quantities are called *collisional invariants*. Examples of collisional invariants include the particle number ($A = 1$), the components of the total momentum ($A = p_\mu$) (in the absence of broken translational invariance, due *e.g.* to the presence of walls), and the total energy ($A = \varepsilon(\boldsymbol{p})$).

### 8.3.5 Detailed balance

Classical mechanics places some restrictions on the form of the kernel $w\big(\Gamma,\Gamma_1\,\big|\,\Gamma',\Gamma_1'\big)$. In particular, if $\Gamma^{\mathcal{T}} = (-\boldsymbol{p},-\boldsymbol{L})$ denotes the kinematic variables under time reversal ($\mathcal{T}$), then

$$w\big(\Gamma',\Gamma_1'\,\big|\,\Gamma,\Gamma_1\big) = w\big(\Gamma^{\mathcal{T}},\Gamma_1^{\mathcal{T}}\,\big|\,\Gamma'^{\mathcal{T}},\Gamma_1'^{\mathcal{T}}\big) \quad . \tag{8.44}$$

This is because the time reverse of the process $|\Gamma,\Gamma_1\rangle \to |\Gamma',\Gamma_1'\rangle$ is $|\Gamma'^{\mathcal{T}},\Gamma_1'^{\mathcal{T}}\rangle \to |\Gamma^{\mathcal{T}},\Gamma_1^{\mathcal{T}}\rangle$.

In equilibrium, we must have

$$w\big(\Gamma',\Gamma_1'\,\big|\,\Gamma,\Gamma_1\big)\; f^0(\Gamma)\, f^0(\Gamma_1)\, d^4\Gamma = w\big(\Gamma^{\mathcal{T}},\Gamma_1^{\mathcal{T}}\,\big|\,\Gamma'^{\mathcal{T}},\Gamma_1'^{\mathcal{T}}\big)\, f^0(\Gamma'^{\mathcal{T}})\, f^0(\Gamma_1'^{\mathcal{T}})\, d^4\Gamma^{\mathcal{T}} \tag{8.45}$$

where

$$d^4\Gamma \equiv d\Gamma\, d\Gamma_1\, d\Gamma' d\Gamma_1' \qquad , \qquad d^4\Gamma^{\mathcal{T}} \equiv d\Gamma^{\mathcal{T}}\, d\Gamma_1^{\mathcal{T}}\, d\Gamma'^{\mathcal{T}} d\Gamma_1'^{\mathcal{T}} \quad . \tag{8.46}$$

Since $d\Gamma = d\Gamma^{\mathcal{T}}$ *etc.*, we may cancel the differentials above, and after invoking eqn. 8.44 and suppressing the common $\boldsymbol{r}$ label, we find

$$f^0(\Gamma)\, f^0(\Gamma_1) = f^0(\Gamma'^{\mathcal{T}})\, f^0(\Gamma_1'^{\mathcal{T}}) \quad . \tag{8.47}$$

This is the condition of *detailed balance*. For the Boltzmann distribution, we have $f^0(\Gamma) = A\, e^{-\varepsilon/k_{\rm B}T}$, where $A$ is a constant and where $\varepsilon = \varepsilon(\Gamma)$ is the kinetic energy, *e.g.* $\varepsilon(\Gamma) = \boldsymbol{p}^2/2m$ in the case of point particles. Note that $\varepsilon(\Gamma^{\mathcal{T}}) = \varepsilon(\Gamma)$. Detailed balance is satisfied because the kinematics of the collision requires energy conservation:

$$\varepsilon + \varepsilon_1 = \varepsilon' + \varepsilon_1' \quad . \tag{8.48}$$

Since momentum is also kinematically conserved, *i.e.*

$$\boldsymbol{p} + \boldsymbol{p}_1 = \boldsymbol{p}' + \boldsymbol{p}_1' \quad , \tag{8.49}$$

[2]Recall from classical mechanics the definition of the *Poisson bracket*, $\{A,B\} = \partial_{\boldsymbol{r}} A\cdot\partial_{\boldsymbol{p}} B - \partial_{\boldsymbol{r}} B\cdot\partial_{\boldsymbol{p}} A$. Then from Hamilton's equations $\dot{\boldsymbol{r}} = \partial H/\partial\boldsymbol{p}$ and $\dot{\boldsymbol{p}} = -\partial H/\partial\boldsymbol{r}$, where $H(\boldsymbol{p},\boldsymbol{r},t)$ is the Hamiltonian, and thus we have $dA/dt = \{A,H\}$. Invariants have zero Poisson bracket with the Hamiltonian.

any distribution of the form

$$f^0(\Gamma) = A\,\exp\!\left(-\frac{\varepsilon(\Gamma) - \boldsymbol{V}\cdot\boldsymbol{p}}{k_{\rm B}T}\right) \tag{8.50}$$

also satisfies detailed balance, for any velocity parameter $\boldsymbol{V}$. This distribution is appropriate for gases which are flowing with average particle velocity $\boldsymbol{V}$. Note that

$$\frac{\partial}{\partial \boldsymbol{p}}\big(\varepsilon - \boldsymbol{V}\cdot\boldsymbol{p}\big) = \boldsymbol{v} - \boldsymbol{V} \quad , \tag{8.51}$$

which is the velocity relative to the local flow.

In addition to time-reversal, parity is also a symmetry of the microscopic mechanical laws. Under the parity operation $\mathcal{P}$, we have $\boldsymbol{r} \to -\boldsymbol{r}$ and $\boldsymbol{p} \to -\boldsymbol{p}$. Note that a pseudovector such as $\boldsymbol{L} = \boldsymbol{r}\times\boldsymbol{p}$ is unchanged under $\mathcal{P}$. Thus, $\Gamma^{\mathcal{P}} = (-\boldsymbol{p}, \boldsymbol{L})$. Under the combined operation of $\mathcal{C} = \mathcal{P}\mathcal{T}$, we have $\Gamma^{\mathcal{C}} = (\boldsymbol{p}, -\boldsymbol{L})$. If the microscopic Hamiltonian is invariant under $C$, then we must have

$$w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big) = w\big(\Gamma^{\mathcal{C}}, \Gamma_1^{\mathcal{C}} \,\big|\, \Gamma'^{\mathcal{C}}, \Gamma_1'^{\mathcal{C}}\big) \quad . \tag{8.52}$$

For point particles, invariance under $T$ and $P$ then means

$$w(\boldsymbol{p}', \boldsymbol{p}_1' \,|\, \boldsymbol{p}, \boldsymbol{p}_1) = w(\boldsymbol{p}, \boldsymbol{p}_1 \,|\, \boldsymbol{p}', \boldsymbol{p}_1') \quad , \tag{8.53}$$

and therefore the collision integral takes the simplified form,

$$\frac{Df(\boldsymbol{r},\boldsymbol{p},t)}{Dt} = \int d^3p_1 \int d^3p' \int d^3p_1'\, w(\boldsymbol{p}', \boldsymbol{p}_1' \,|\, \boldsymbol{p}, \boldsymbol{p}_1)\Big\{f(\boldsymbol{r},\boldsymbol{p}',t)\, f(\boldsymbol{r},\boldsymbol{p}_1',t) - f(\boldsymbol{r},\boldsymbol{p},t)\, f(\boldsymbol{r},\boldsymbol{p}_1,t)\Big\} = \left(\frac{df}{dt}\right)_{\rm coll} \quad . \tag{8.54}$$

The most general statement of detailed balance is

$$\frac{f^0(\Gamma')\, f^0(\Gamma_1')}{f^0(\Gamma)\, f^0(\Gamma_1)} = \frac{w\big(\Gamma', \Gamma_1' \,\big|\, \Gamma, \Gamma_1\big)}{w\big(\Gamma, \Gamma_1 \,\big|\, \Gamma', \Gamma_1'\big)} \quad . \tag{8.55}$$

Under this condition, the collision term vanishes for $f = f^0$, which is the equilibrium distribution.

**Detailed balance and single particle scattering**

For single particle scattering, similar considerations lead to

$$f^0(\Gamma) = f^0(\Gamma'^{\mathcal{T}}) \qquad , \qquad \frac{f^0(\Gamma')}{f^0(\Gamma)} = \frac{w(\Gamma'\,|\,\Gamma)}{w(\Gamma\,|\,\Gamma')} \quad . \tag{8.56}$$

In the first of these equations, the time reversal operation also acts on the external potential off which the particles scatter. In general, detailed balance pertains when the static equilibrium distribution $f^0(\Gamma)$ is such that the integrand of the collision term vanishes for all $\Gamma$, hence $(df^0(\Gamma)/dt)_{\rm coll} = 0$.

### 8.3.6 Kinematics and cross section

We can rewrite eqn. 8.54 in the form

$$\frac{Df(\boldsymbol{r},\boldsymbol{p},t)}{Dt} = \int d^3p_1 \int d\Omega\, |\boldsymbol{v}-\boldsymbol{v}_1|\, \frac{d\sigma}{d\Omega}\Big\{f(\boldsymbol{r},\boldsymbol{p}',t)\, f(\boldsymbol{r},\boldsymbol{p}_1',t) - f(\boldsymbol{r},\boldsymbol{p},t)\, f(\boldsymbol{r},\boldsymbol{p}_1,t)\Big\} \quad , \tag{8.57}$$

where $d\sigma/d\Omega$ is the *differential scattering cross section*. Recall the definition from scattering theory in classical mechanics,

$$d\sigma = \frac{\#\text{ of particles scattered into solid angle } d\Omega \text{ per unit time}}{\text{incident flux}} \quad . \tag{8.58}$$

If we recast the scattering problem in terms of center-of-mass and relative coordinates, we conclude that the total momentum is conserved by the collision, and furthermore that the energy in the CM frame is conserved, which means that the magnitude of the *relative* momentum is conserved. Thus, we may write $\boldsymbol{p}'-\boldsymbol{p}_1' = |\boldsymbol{p}-\boldsymbol{p}_1|\,\hat{\boldsymbol{\Omega}}$, where $\hat{\boldsymbol{\Omega}}$ is a unit vector. Then $\boldsymbol{p}'$ and $\boldsymbol{p}_1'$ are determined to be

$$\begin{aligned} \boldsymbol{p}' &= \tfrac{1}{2}\big(\boldsymbol{p}+\boldsymbol{p}_1 + |\boldsymbol{p}-\boldsymbol{p}_1|\,\hat{\boldsymbol{\Omega}}\big) \\ \boldsymbol{p}_1' &= \tfrac{1}{2}\big(\boldsymbol{p}+\boldsymbol{p}_1 - |\boldsymbol{p}-\boldsymbol{p}_1|\,\hat{\boldsymbol{\Omega}}\big) \quad . \end{aligned} \tag{8.59}$$

Recall that for the scattering of classical hard spheres of radius $a$, the differential scattering cross section is $\frac{d\sigma}{d\Omega} = a^2$. Thus, the total scattering cross section is $\sigma_{\text{tot}} = 4\pi a^2 = \pi d^2$, where $d = 2a$ is the sphere diameter. For Coulomb scattering of two point particles of charge $q$, one has

$$\frac{d\sigma}{d\Omega} = \left(\frac{me^2}{|\boldsymbol{p}_1-\boldsymbol{p}_2|^2\,\sin^2(\frac{1}{2}\vartheta)}\right)^2 \quad , \tag{8.60}$$

where $\hat{\boldsymbol{p}}_1\cdot\hat{\boldsymbol{p}}_2 = \cos\vartheta$. The total cross section for Coulomb scattering diverges since the differential cross section behaves as $\vartheta^{-4}$ as $\vartheta\to 0$.

### 8.3.7 $\mathcal{H}$-theorem

To peek ahead, we are about to prove the following. Let

$$\begin{aligned} h(\boldsymbol{r},t) &= \int d^3p\, f(\boldsymbol{r},\boldsymbol{p},t)\, \log\big[f(\boldsymbol{r},\boldsymbol{p},t)\mathsf{a}^3\big] \\ \boldsymbol{j}(\boldsymbol{r},t) &= \int d^3p\, f(\boldsymbol{r},\boldsymbol{p},t)\, \log\big[f(\boldsymbol{r},\boldsymbol{p},t)\mathsf{a}^3\big]\, \frac{d\boldsymbol{r}}{dt} \quad . \end{aligned} \tag{8.61}$$

Here $\mathsf{a}$ can be any constant which has the appropriate dimensions of action. Then if $f(\boldsymbol{r},\boldsymbol{p},t)$ evolves according to the Boltzmann equation, it is necessarily the case that

$$\frac{\partial h(\boldsymbol{r},t)}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j}(\boldsymbol{r},t) \le 0 \quad , \tag{8.62}$$

where $\boldsymbol{\nabla} \equiv \partial/\partial \boldsymbol{r}$. If we integrate over all space, and we adopt boundary conditions where $\boldsymbol{j} \to 0$ at spatial infinity,

$$\mathcal{H}(t) = \int d^3r\, h(\boldsymbol{r}, t) \quad \Rightarrow \quad \frac{d\mathcal{H}}{dt} \leq 0 \quad . \tag{8.63}$$

*Thus, Boltzmann dynamics recognizes an arrow of time.* Time increases in the direction that $\mathcal{H}$ decreases.

Let's consider the Boltzmann equation with two particle collisions. We define the local (*i.e.* $\boldsymbol{r}$-dependent) quantity

$$\rho_\varphi(\boldsymbol{r}, t) \equiv \int d^3p\, f(\boldsymbol{r}, \boldsymbol{p}, t)\, \varphi\big(f(\boldsymbol{r}, \boldsymbol{p}, t)\big) \quad , \tag{8.64}$$

where $f = f(\boldsymbol{r}, \boldsymbol{p}, t)$ and $\varphi(f)$ is arbitrary. We now compute

$$\begin{aligned}
\frac{\partial \rho_\varphi}{\partial t} &= \int d^3p\, \frac{\partial (f\varphi)}{\partial t} = \int d\Gamma\, \frac{\partial (f\varphi)}{\partial f}\, \frac{\partial f}{\partial t} \\
&= \int d^3p\, \frac{\partial (f\varphi)}{\partial f} \left\{ -\dot{\boldsymbol{r}} \cdot \frac{\partial f}{\partial \boldsymbol{r}} - \dot{\boldsymbol{p}} \cdot \frac{\partial f}{\partial \boldsymbol{p}} + \left(\frac{df}{dt}\right)_{\rm coll} \right\} \\
&= \int d^3p \left\{ -\dot{\boldsymbol{r}} \cdot \frac{\partial (f\varphi)}{\partial \boldsymbol{r}} - \dot{\boldsymbol{p}} \cdot \frac{\partial (f\varphi)}{\partial \boldsymbol{p}} + \frac{\partial (f\varphi)}{\partial f} \left(\frac{df}{dt}\right)_{\rm coll} \right\}
\end{aligned} \tag{8.65}$$

We may integrate the second term in the brackets by parts on $\boldsymbol{p}$. Assuming $f = 0$ for infinite values of the kinematic variables, which is the only physical possibility, we then have

$$\begin{aligned}
\frac{\partial \rho_\varphi}{\partial t} &= \int d^3p \left\{ -\dot{\boldsymbol{r}} \cdot \frac{\partial (f\varphi)}{\partial \boldsymbol{r}} - \frac{\partial \dot{\boldsymbol{r}}}{\partial \boldsymbol{r}}\, (f\varphi) + \frac{\partial (f\varphi)}{\partial f} \left(\frac{df}{dt}\right)_{\rm coll} \right\} \\
&= -\frac{\partial}{\partial \boldsymbol{r}} \int d^3p\, f\varphi\, \dot{\boldsymbol{r}} + \int d^3p\, \frac{\partial (f\varphi)}{\partial f} \left(\frac{df}{dt}\right)_{\rm coll} \quad .
\end{aligned} \tag{8.66}$$

Thus,

$$\frac{\partial \rho_\varphi(\boldsymbol{r}, t)}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{j}_\varphi(\boldsymbol{r}, t) = \sigma_\varphi(\boldsymbol{r}, t) \quad , \tag{8.67}$$

where

$$\begin{aligned}
\boldsymbol{j}_\varphi(\boldsymbol{r}, t) &= \int d^3p\, f(\boldsymbol{r}, \boldsymbol{p}, t)\, \varphi\big(f(\boldsymbol{r}, \boldsymbol{p}, t)\big)\, \boldsymbol{v}(\boldsymbol{p}) \\
\sigma_\varphi(\boldsymbol{r}, t) &= \int d^3p\, \frac{\partial (f\varphi)}{\partial f}\bigg|_{f(\boldsymbol{r}, \boldsymbol{p}, t)} \left(\frac{df}{dt}\right)_{\rm coll}
\end{aligned} \tag{8.68}$$

and $\dot{\boldsymbol{r}} = \boldsymbol{v}(\boldsymbol{p}) = \partial H_0/\partial \boldsymbol{p}$ is the velocity.

Thus, we arrive at eqn. 8.67, which is a continuity equation with a source term $\sigma_\varphi(\boldsymbol{r}, t)$. The source term is nonzero only in the presence of collisions. We now evaluate $\sigma_\varphi$ under the assumption that $f$ satisfies the Boltzmann equation with two particle scattering. Thus,

$$\begin{aligned}
\sigma_\varphi(\boldsymbol{r}, t) &= \int d^3p \int d^3p_1 \int d^3p' \int d^3p'_1 \left\{ w(\boldsymbol{p}', \boldsymbol{p}'_1 \,|\, \boldsymbol{p}, \boldsymbol{p}_1)\, f(\boldsymbol{p}) f(\boldsymbol{p}_1)\, \chi(\boldsymbol{p}) - w(\boldsymbol{p}, \boldsymbol{p}_1 \,|\, \boldsymbol{p}', \boldsymbol{p}'_1)\, f(\boldsymbol{p}') f(\boldsymbol{p}'_1)\, \chi(\boldsymbol{p}') \right\} \\
&= \int d^3p \int d^3p_1 \int d^3p' \int d^3p'_1\, w(\boldsymbol{p}', \boldsymbol{p}'_1 \,|\, \boldsymbol{p}, \boldsymbol{p}_1)\, f(\boldsymbol{p}) f(\boldsymbol{p}_1)\, \big(\chi(\boldsymbol{p}) - \chi'(\boldsymbol{p})\big) \quad ,
\end{aligned} \tag{8.69}$$

where

$$\chi = \frac{\partial(f\varphi)}{\partial f} = \varphi + f\,\frac{\partial\varphi}{\partial f} \quad , \tag{8.70}$$

and where we have suppressed the $\boldsymbol{r}$ and $t$ dependences. We now invoke the symmetry

$$w(\boldsymbol{p}',\boldsymbol{p}_1' \,|\, \boldsymbol{p},\boldsymbol{p}_1) = w(\boldsymbol{p}_1',\boldsymbol{p}' \,|\, \boldsymbol{p}_1,\boldsymbol{p}) \quad , \tag{8.71}$$

which allows us to write

$$\sigma = \tfrac{1}{2}\int d^3p \int d^3p_1 \int d^3p' \int d^3p_1'\; w(\boldsymbol{p}',\boldsymbol{p}_1' \,|\, \boldsymbol{p},\boldsymbol{p}_1)\, f(\boldsymbol{p})\, f(\boldsymbol{p}_1)\, \big(\chi(\boldsymbol{p}) + \chi(\boldsymbol{p}_1) - \chi(\boldsymbol{p}') - \chi(\boldsymbol{p}_1')\big) \quad , \tag{8.72}$$

This shows that $\sigma_\varphi = 0$ if $\chi(\boldsymbol{p})$ is a collisional invariant, *i.e.* if $\chi(\boldsymbol{p}) + \chi(\boldsymbol{p}_1) = \chi(\boldsymbol{p}') + \chi(\boldsymbol{p}_1')$ for every allowed collision.

Now let us fix $\varphi(f) = \log(f\mathsf{a}^3)$ and evaluate the source term $\sigma \equiv \sigma_{\varphi=\log(f\mathsf{a}^3)}$. We have

$$\sigma = -\tfrac{1}{2}\int d^3p \int d^3p_1 \int d^3p' \int d^3p_1'\; w(\boldsymbol{p}',\boldsymbol{p}_1' \,|\, \boldsymbol{p},\boldsymbol{p}_1)\, f(\boldsymbol{p}') f(\boldsymbol{p}_1') \cdot x(\boldsymbol{p},\boldsymbol{p}_1 \,|\, \boldsymbol{p}',\boldsymbol{p}_1') \log x(\boldsymbol{p},\boldsymbol{p}_1 \,|\, \boldsymbol{p}',\boldsymbol{p}_1') \quad , \tag{8.73}$$

where $x(\boldsymbol{p},\boldsymbol{p}_1,\boldsymbol{p}',\boldsymbol{p}_1') \equiv f(\boldsymbol{p})f(\boldsymbol{p}_1)/f(\boldsymbol{p}')f(\boldsymbol{p}_1')$. We next invoke the result

$$\int d^3p' \int d^3p_1'\; w(\boldsymbol{p}',\boldsymbol{p}_1' \,|\, \boldsymbol{p},\boldsymbol{p}_1) = \int d^3p' \int d^3p_1'\; w(\boldsymbol{p},\boldsymbol{p}_1 \,|\, \boldsymbol{p}',\boldsymbol{p}_1') \tag{8.74}$$

which is a statement of unitarity of the scattering matrix[3]. Multiplying both sides by $f(\boldsymbol{p})\, f(\boldsymbol{p}_1)$, then integrating over $\boldsymbol{p}$ and $\boldsymbol{p}_1$, and finally changing variables $(\boldsymbol{p},\boldsymbol{p}_1) \leftrightarrow (\boldsymbol{p}',\boldsymbol{p}_1')$, we find

$$\begin{aligned} 0 &= \int d^3p \int d^3p_1 \int d^3p' \int d^3p_1'\; w(\boldsymbol{p}',\boldsymbol{p}_1' \,|\, \boldsymbol{p},\boldsymbol{p}_1)\, \Big(f(\boldsymbol{p})f(\boldsymbol{p}_1) - f(\boldsymbol{p}')f(\boldsymbol{p}_1')\Big) \\ &= \int d^3p \int d^3p_1 \int d^3p' \int d^3p_1'\; w(\boldsymbol{p}',\boldsymbol{p}_1' \,|\, \boldsymbol{p},\boldsymbol{p}_1)\, f(\boldsymbol{p}')f(\boldsymbol{p}_1')\, \Big\{x(\boldsymbol{p},\boldsymbol{p}_1 \,|\, \boldsymbol{p}',\boldsymbol{p}_1') - 1\Big\} \quad . \end{aligned} \tag{8.75}$$

Multiplying this result by $\frac{1}{2}$ and adding it to the previous equation for $\dot{h}$, we arrive at our final result,

$$\sigma = -\tfrac{1}{2}\int d^3p \int d^3p_1 \int d^3p' \int d^3p_1'\; w(\boldsymbol{p}',\boldsymbol{p}_1' \,|\, \boldsymbol{p},\boldsymbol{p}_1)\, f(\boldsymbol{p}')f(\boldsymbol{p}_1')\; (x\log x - x + 1) \quad , \tag{8.76}$$

where $x \equiv x(\boldsymbol{p},\boldsymbol{p}_1,\boldsymbol{p}',\boldsymbol{p}_1') = f(\boldsymbol{p})f(\boldsymbol{p}_1)/f(\boldsymbol{p}')f(\boldsymbol{p}_1')$. It is now easy to prove that the function $g(x) = x\log x - x + 1$ is nonnegative for all positive $x$ values[4], which therefore entails the important result

$$\frac{\partial h(\boldsymbol{r},t)}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j}(\boldsymbol{r},t) = \sigma(\boldsymbol{r},t) \le 0 \quad . \tag{8.77}$$

Boltzmann's $\mathcal{H}$ function is the space integral of the local density $h(\boldsymbol{r})$: $\mathcal{H} = \int d^3r\; h(\boldsymbol{r})$.

[3] See Lifshitz and Pitaevskii, *Physical Kinetics*, §2.

[4] The function $g(x) = x\log x - x + 1$ satisfies $g'(x) = \log x$, hence $g'(x) < 0$ on the interval $x \in [0,1)$ and $g'(x) > 0$ on $x \in (1,\infty]$. Thus, $g(x)$ monotonically decreases from $g(0) = 1$ to $g(1) = 0$, and then monotonically increases to $g(\infty) = \infty$, never becoming negative.

Thus, everywhere in space, the source term $\sigma(\boldsymbol{r},t)$ is nonpositive. In equilibrium, $\dot{h}=0$ everywhere, which requires $x=1$, *i.e.*

$$f^0(\boldsymbol{p})\,f^0(\boldsymbol{p}_1)=f^0(\boldsymbol{p}')\,f^0(\boldsymbol{p}'_1) \quad , \tag{8.78}$$

or, taking the logarithm,

$$\log f^0(\boldsymbol{p})+\log f^0(\boldsymbol{p}_1)=\log f^0(\boldsymbol{p}')+\log f^0(\boldsymbol{p}'_1) \quad . \tag{8.79}$$

But this means that $\log f^0$ is itself a collisional invariant, and if $1$, $\boldsymbol{p}$, and $\varepsilon$ are the only collisional invariants, then $\log f^0$ must be expressible in terms of them. Thus, we may write

$$-\log f^0=\frac{\varepsilon(\boldsymbol{p})-\boldsymbol{V}\cdot\boldsymbol{p}-\mu}{k_{\rm B}T} \quad , \tag{8.80}$$

where $\mu$, $\boldsymbol{V}$, and $T$ are constants which parameterize the equilibrium distribution $f^0(\boldsymbol{p})$, corresponding to the chemical potential, flow velocity, and temperature, respectively.

## 8.4 Weakly Inhomogeneous Gas

Consider a gas which is only weakly out of equilibrium. We follow the treatment in Lifshitz and Pitaevskii, §6. As the gas is only slightly out of equilibrium, we seek a solution to the Boltzmann equation of the form $f=f^0+\delta f$, where $f^0$ describes a local equilibrium. Recall that such a distribution function is annihilated by the collision term in the Boltzmann equation but not by the streaming term, hence a correction $\delta f$ must be added in order to obtain a solution. The Boltzmann equation is written

$$\left(\frac{\partial}{\partial t}+\frac{\boldsymbol{p}}{m}\cdot\frac{\partial}{\partial\boldsymbol{r}}+\boldsymbol{F}\cdot\frac{\partial}{\partial\boldsymbol{p}}\right)\left(f^0+\delta f\right)=\left(\frac{df}{dt}\right)_{\rm coll} \quad . \tag{8.81}$$

The RHS of this equation must be of order $\delta f$ because the local equilibrium distribution $f^0$ is annihilated by the collision integral.

The most general form of local equilibrium is described by the distribution

$$f^0(\boldsymbol{r},\Gamma)=C\exp\left(\frac{\mu-\varepsilon(\Gamma)+\boldsymbol{V}\cdot\boldsymbol{p}}{k_{\rm B}T}\right) \quad , \tag{8.82}$$

where $\mu=\mu(\boldsymbol{r},t)$, $T=T(\boldsymbol{r},t)$, and $\boldsymbol{V}=\boldsymbol{V}(\boldsymbol{r},t)$ may all vary in both space and time. Note that

$$\begin{aligned} df^0&=\left(d\mu+\boldsymbol{p}\cdot d\boldsymbol{V}+(\varepsilon-\boldsymbol{V}\cdot\boldsymbol{p}-\mu)\,\frac{dT}{T}-d\varepsilon\right)\left(-\frac{\partial f^0}{\partial\varepsilon}\right)\\ &=\left(\frac{1}{n}\,dp+\boldsymbol{p}\cdot d\boldsymbol{V}+(\varepsilon-h)\,\frac{dT}{T}-d\varepsilon\right)\left(-\frac{\partial f^0}{\partial\varepsilon}\right) \quad , \end{aligned} \tag{8.83}$$

where $h=\mu+Ts$ is the enthalpy per particle, and where we drop the term proportional to $\boldsymbol{V}\cdot\boldsymbol{p}\,dT$, which on the assumption that $\boldsymbol{V}=0$ on average is second order in smallness. We have also invoked local thermodynamics in the form

$$d\mu=\left(\frac{\partial\mu}{\partial T}\right)_p dT+\left(\frac{\partial\mu}{\partial p}\right)_T dp=-s\,dT+\frac{1}{n}\,dp \quad , \tag{8.84}$$

where $s$ is the entropy per particle and $n$ is the number density. Note as well that

$$-\frac{\partial f^0}{\partial \varepsilon} = \frac{f^0}{k_{\rm B}T} \quad , \tag{8.85}$$

which we shall invoke further on below.

We wish to evaluate one of the contributions to the LHS of eqn. 8.81:

$$\frac{\partial f^0}{\partial t} + \frac{\boldsymbol{p}}{m}\cdot\frac{\partial f^0}{\partial \boldsymbol{r}} + \boldsymbol{F}\cdot\frac{\partial f^0}{\partial \boldsymbol{p}} = \left\{ \frac{\varepsilon - h}{T}\,\frac{\partial T}{\partial t} + m\boldsymbol{v}\cdot\Big[(\boldsymbol{v}\cdot\boldsymbol{\nabla})\,\boldsymbol{V}\Big] + \boldsymbol{v}\cdot\left(m\,\frac{\partial \boldsymbol{V}}{\partial t} + \frac{1}{n}\,\boldsymbol{\nabla} p\right) \right. \\ \left. + \frac{1}{n}\,\frac{\partial p}{\partial t} + \frac{\varepsilon - h}{T}\,\boldsymbol{v}\cdot\boldsymbol{\nabla} T - \boldsymbol{F}\cdot\boldsymbol{v}\right\}\left(-\frac{\partial f^0}{\partial \varepsilon}\right) \quad . \tag{8.86}$$

To simplify this, first note that Newton's laws applied to an ideal fluid give $\rho\dot{\boldsymbol{V}} = -\boldsymbol{\nabla} p$, where $\rho = mn$ is the mass density. Corrections to this result, *e.g.* viscosity and nonlinearity in $\boldsymbol{V}$, are of higher order.

Next, continuity for particle number means $\dot{n} + \boldsymbol{\nabla}\cdot(n\boldsymbol{V}) = 0$. We assume $\boldsymbol{V}$ is zero on average and that all derivatives are small, hence $\boldsymbol{\nabla}\cdot(n\boldsymbol{V}) = \boldsymbol{V}\cdot\boldsymbol{\nabla} n + n\,\boldsymbol{\nabla}\cdot\boldsymbol{V} \approx n\,\boldsymbol{\nabla}\cdot\boldsymbol{V}$. Thus,

$$\frac{\partial \log n}{\partial t} = \frac{\partial \log p}{\partial t} - \frac{\partial \log T}{\partial t} = -\boldsymbol{\nabla}\cdot\boldsymbol{V} \quad , \tag{8.87}$$

where we have invoked the ideal gas law $n = p/k_{\rm B}T$ above.

Next, we invoke conservation of entropy. If $s$ is the entropy per particle, then $ns$ is the entropy per unit volume, in which case we have the continuity equation

$$\frac{\partial (ns)}{\partial t} + \boldsymbol{\nabla}\cdot(ns\boldsymbol{V}) = n\left(\frac{\partial s}{\partial t} + \boldsymbol{V}\cdot\boldsymbol{\nabla} s\right) + s\left(\frac{\partial n}{\partial t} + \boldsymbol{\nabla}\cdot(n\boldsymbol{V})\right) = 0 \quad . \tag{8.88}$$

The second bracketed term on the RHS vanishes because of particle continuity, leaving us with the combination $\dot{s} + \boldsymbol{V}\cdot\boldsymbol{\nabla} s \approx \dot{s} = 0$ (since $\boldsymbol{V} = 0$ on average, and any gradient is first order in smallness). Now thermodynamics says

$$ds = \left(\frac{\partial s}{\partial T}\right)_p dT + \left(\frac{\partial s}{\partial p}\right)_T dp = \frac{c_p}{T}\,dT - \frac{k_{\rm B}}{p}\,dp \quad , \tag{8.89}$$

since $c_p = T\big(\partial s/\partial T\big)_p$ and $\big(\partial s/\partial p\big)_T = \big(\partial v/\partial T\big)_p$, where $v = V/N$. Here, $c_p$ is the heat capacity per particle at constant pressure[5]. Thus,

$$\frac{c_p}{k_{\rm B}}\,\frac{\partial \log T}{\partial t} - \frac{\partial \log p}{\partial t} = 0 \quad . \tag{8.90}$$

We now have in eqns. 8.87 and 8.90 two equations in the two unknowns $\partial\log T/\partial t$ and $\partial \log p/\partial t$, yielding

$$\frac{\partial \log T}{\partial t} = -\frac{k_{\rm B}}{c_V}\,\boldsymbol{\nabla}\cdot\boldsymbol{V} \qquad , \qquad \frac{\partial \log p}{\partial t} = -\frac{c_p}{c_V}\,\boldsymbol{\nabla}\cdot\boldsymbol{V} \quad . \tag{8.91}$$

[5]In the chapter on thermodynamics, we adopted a slightly different definition of $c_p$ as the heat capacity per mole. In this chapter $c_p$ is the heat capacity per particle.

Thus eqn. 8.86 becomes

$$\frac{\partial f^0}{\partial t}+\frac{\boldsymbol{p}}{m}\cdot\frac{\partial f^0}{\partial \boldsymbol{r}}+\boldsymbol{F}\cdot\frac{\partial f^0}{\partial \boldsymbol{p}}=\left(-\frac{\partial f^0}{\partial\varepsilon}\right)\left\{\frac{\varepsilon-h}{T}\,\boldsymbol{v}\cdot\boldsymbol{\nabla}T+m\,v_\alpha v_\beta\,\mathcal{Q}_{\alpha\beta}+\frac{h-Tc_p-\varepsilon}{c_V/k_{\rm B}}\,\boldsymbol{\nabla}\cdot\boldsymbol{V}-\boldsymbol{F}\cdot\boldsymbol{v}\right\}\quad,\tag{8.92}$$

where $\varepsilon=\varepsilon(\Gamma)$ and

$$\mathcal{Q}_{\alpha\beta}=\frac{1}{2}\left(\frac{\partial V_\alpha}{\partial x_\beta}+\frac{\partial V_\beta}{\partial x_\alpha}\right)\quad.\tag{8.93}$$

Therefore, the Boltzmann equation takes the form

$$\frac{\partial\,\delta f}{\partial t}+\left\{\frac{\varepsilon(\Gamma)-h}{T}\,\boldsymbol{v}\cdot\boldsymbol{\nabla}T+m\,v_\alpha v_\beta\,\mathcal{Q}_{\alpha\beta}-\frac{\varepsilon(\Gamma)-h+Tc_p}{c_V/k_{\rm B}}\,\boldsymbol{\nabla}\cdot\boldsymbol{V}-\boldsymbol{F}\cdot\boldsymbol{v}\right\}\frac{f^0}{k_{\rm B}T}=\left(\frac{df}{dt}\right)_{\rm coll}\quad.\tag{8.94}$$

Notice we have dropped the terms $\boldsymbol{v}\cdot\partial\,\delta f/\partial\boldsymbol{r}$ and $\boldsymbol{F}\cdot\partial\,\delta f/\partial\boldsymbol{p}$, since $\delta f$ must already be first order in smallness, and both the $\partial/\partial\boldsymbol{r}$ operator as well as $\boldsymbol{F}$ add a second order of smallness, which is negligible. Typically $\partial\,\delta f/\partial t$ is nonzero if the applied force $\boldsymbol{F}(t)$ is time-dependent. We use the convention of summing over repeated indices. Note that $\delta_{\alpha\beta}\,\mathcal{Q}_{\alpha\beta}=\mathcal{Q}_{\alpha\alpha}=\boldsymbol{\nabla}\cdot\boldsymbol{V}$. For ideal gases in which only translational and rotational degrees of freedom are excited, $h=c_pT$.

## 8.5 Relaxation Time Approximation

### 8.5.1 Approximation of collision integral

We now consider a very simple model of the collision integral,

$$\left(\frac{df}{dt}\right)_{\rm coll}=-\frac{f-f^0}{\tau}=-\frac{\delta f}{\tau}\quad.\tag{8.95}$$

This model is known as the *relaxation time approximation*. Here, $f^0=f^0(\boldsymbol{r},\boldsymbol{p},t)$ is a distribution function which describes a *local equilibrium* at each position $\boldsymbol{r}$ and time $t$. The quantity $\tau$ is the *relaxation time*, which can in principle be momentum-dependent, but which we shall first consider to be constant. In the absence of streaming terms, we have

$$\frac{\partial\,\delta f}{\partial t}=-\frac{\delta f}{\tau}\qquad\Longrightarrow\qquad\delta f(\boldsymbol{r},\boldsymbol{p},t)=\delta f(\boldsymbol{r},\boldsymbol{p},0)\,e^{-t/\tau}\quad.\tag{8.96}$$

The distribution $f$ then relaxes to the equilibrium distribution $f^0$ on a time scale $\tau$. We note that this approximation is obviously flawed in that all quantities – even the collisional invariants – relax to their equilibrium values on the scale $\tau$. In the Appendix II, we consider a model for the collision integral in which the collisional invariants are all preserved, but everything else relaxes to local equilibrium at a single rate.

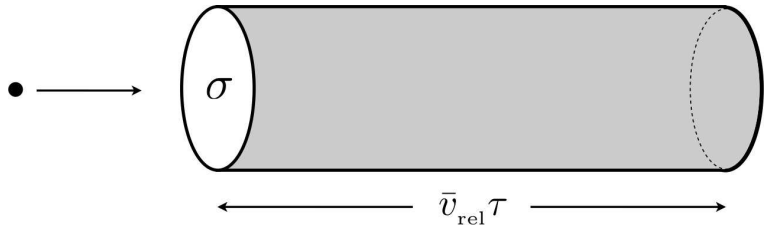



Figure 8.3: Graphic representation of the equation $n\,\sigma\,\bar{v}_{\rm rel}\,\tau = 1$, which yields the scattering time $\tau$ in terms of the number density $n$, average particle pair relative velocity $\bar{v}_{\rm rel}$, and two-particle total scattering cross section $\sigma$. The equation says that on average there must be one particle within the tube.

### 8.5.2 Computation of the scattering time

Consider two particles with velocities $\boldsymbol{v}$ and $\boldsymbol{v}'$. The average of their relative *speed* is

$$\big\langle\,|\boldsymbol{v}-\boldsymbol{v}'|\,\big\rangle = \int d^3v \int d^3v'\;P(\boldsymbol{v})\,P(\boldsymbol{v}')\,|\boldsymbol{v}-\boldsymbol{v}'| \quad , \tag{8.97}$$

where $P(\boldsymbol{v})$ is the Maxwell velocity distribution,

$$P(\boldsymbol{v}) = \left(\frac{m}{2\pi k_{\rm B}T}\right)^{3/2} \exp\left(-\frac{m\boldsymbol{v}^2}{2k_{\rm B}T}\right) \quad , \tag{8.98}$$

which follows from the Boltzmann form of the equilibrium distribution $f^0(\boldsymbol{p})$. It is left as an exercise for the student to verify that

$$\bar{v}_{\rm rel} \equiv \big\langle\,|\boldsymbol{v}-\boldsymbol{v}'|\,\big\rangle = \frac{4}{\sqrt{\pi}}\left(\frac{k_{\rm B}T}{m}\right)^{1/2} \quad . \tag{8.99}$$

Note that $\bar{v}_{\rm rel} = \sqrt{2}\,\bar{v}$, where $\bar{v}$ is the average particle speed. Let $\sigma$ be the total scattering cross section, which for hard spheres is $\sigma = \pi d^2$, where $d$ is the hard sphere diameter. Then the rate at which particles scatter is

$$\frac{1}{\tau} = n\,\bar{v}_{\rm rel}\,\sigma \quad . \tag{8.100}$$

The particle *mean free path* is then

$$\ell = \bar{v}\,\tau = \frac{1}{\sqrt{2}\,n\,\sigma} \quad . \tag{8.101}$$

While the scattering length is not temperature-dependent within this formalism, the scattering time *is* $T$-dependent, with

$$\tau(T) = \frac{1}{n\,\bar{v}_{\rm rel}\,\sigma} = \frac{\sqrt{\pi}}{4n\sigma}\left(\frac{m}{k_{\rm B}T}\right)^{1/2} \quad . \tag{8.102}$$

As $T \to 0$, the collision time diverges as $\tau \propto T^{-1/2}$, because the particles on average move more slowly at lower temperatures. The mean free path, however, is independent of $T$, and is given by $\ell = 1/\sqrt{2}n\sigma$.

### 8.5.3 Thermal conductivity

We consider a system with a temperature gradient $\boldsymbol{\nabla} T$ and seek a steady state (*i.e.* time-independent) solution to the Boltzmann equation. We assume $F_\alpha = \mathcal{Q}_{\alpha\beta} = 0$. Appealing to eqn. 8.94, and using the relaxation time approximation for the collision integral, we have

$$\delta f = -\frac{\tau(\varepsilon - c_p\, T)}{k_{\rm B} T^2}\, (\boldsymbol{v} \cdot \boldsymbol{\nabla} T)\, f^0 \quad . \tag{8.103}$$

We are now ready to compute the energy and particle currents. In order to compute the *local density* of any quantity $A(\boldsymbol{r}, \boldsymbol{p})$, we multiply by the distribution $f(\boldsymbol{r}, \boldsymbol{p})$ and integrate over momentum:

$$\rho_A(\boldsymbol{r}, t) = \int d^3p\, A(\boldsymbol{r}, \boldsymbol{p})\, f(\boldsymbol{r}, \boldsymbol{p}, t) \quad , \tag{8.104}$$

For the energy (thermal) current, we let $A = \varepsilon\, v_\alpha = \varepsilon\, p_\alpha / m$, in which case $\rho_A = j_\alpha$. Note $\int d^3p\, \boldsymbol{p}\, f^0 = 0$ since $f^0$ is isotropic in $\boldsymbol{p}$ even when $\mu$ and $T$ depend on $\boldsymbol{r}$. Thus, only $\delta f$ enters into the calculation of the various currents. Thus, the energy (thermal) current is

$$j_\varepsilon^\alpha(\boldsymbol{r}) = \int d^3p\, \varepsilon\, v^\alpha\, \delta f = -\frac{n\tau}{k_{\rm B} T^2} \left\langle\, v^\alpha v^\beta\, \varepsilon\, (\varepsilon - c_p\, T) \,\right\rangle \frac{\partial T}{\partial x^\beta} \quad , \tag{8.105}$$

where the repeated index $\beta$ is summed over, and where momentum averages are defined relative to the equilibrium distribution, *i.e.*

$$\left\langle\, \phi(\boldsymbol{p}) \,\right\rangle = \int d^3p\, \phi(\boldsymbol{p})\, f^0(\boldsymbol{p}) \Big/ \int d^3p\, f^0(\boldsymbol{p}) = \int d^3v\, P(\boldsymbol{v})\, \phi(m\boldsymbol{v}) \quad . \tag{8.106}$$

In this context, it is useful to invoke the identity $d^3p\, f^0(\boldsymbol{p}) = n\, d^3v\, P(\boldsymbol{v})$, where

$$P(\boldsymbol{v}) = \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3/2} e^{-m(\boldsymbol{v}-\boldsymbol{V})^2/2k_{\rm B}T} \tag{8.107}$$

is the Maxwell velocity distribution.

Note that if $\phi = \phi(\varepsilon)$ is a function of the energy, and if $\boldsymbol{V} = 0$, then

$$d^3p\, f^0(\boldsymbol{p}) = n\, d^3v\, P(\boldsymbol{v}) = n\, \mathcal{P}(\varepsilon)\, d\varepsilon \quad , \tag{8.108}$$

where

$$\mathcal{P}(\varepsilon) = \frac{2}{\sqrt{\pi}}\, (k_{\rm B} T)^{-3/2}\, \varepsilon^{1/2}\, e^{-\varepsilon/k_{\rm B}T} \quad , \tag{8.109}$$

is the Maxwellian distribution expressed in terms of $\varepsilon$, and is normalized according to $\int_0^\infty d\varepsilon\, \mathcal{P}(\varepsilon) = 1$.[6] Averages with respect to this distribution are given by

$$\left\langle\, \phi(\varepsilon) \,\right\rangle = \int_0^\infty d\varepsilon\, \phi(\varepsilon)\, \mathcal{P}(\varepsilon) = \frac{2}{\sqrt{\pi}}\, (k_{\rm B} T)^{-3/2} \int_0^\infty d\varepsilon\, \varepsilon^{1/2}\, \phi(\varepsilon)\, e^{-\varepsilon/k_{\rm B}T} \quad . \tag{8.110}$$

[6] $\mathcal{P}(\varepsilon)$ can also be recognized as the normalized product of the Boltzmann energy distribution $\exp(-\varepsilon/k_{\rm B}T)$ and the density of states $g(\varepsilon)$. In $d = 3$ we have $g(\varepsilon) \propto \varepsilon^{1/2}$.

For homogeneous $\phi(s) = \varepsilon^s$ we have

$$\langle\, \varepsilon^s \,\rangle = \frac{2}{\sqrt{\pi}}\,\Gamma\big(s+\tfrac{3}{2}\big)(k_{\rm B}T)^s \quad . \tag{8.111}$$

Due to spatial isotropy, it is clear that we can replace $v^\alpha\, v^\beta$ by $\frac{1}{3}\boldsymbol{v}^2\,\delta_{\alpha\beta}$ and then $\varepsilon = \frac{1}{2}m\boldsymbol{v}^2$ in eqn. 8.105. We then have $\boldsymbol{j}_\varepsilon = -\kappa\,\boldsymbol{\nabla}T$, with

$$\kappa = \frac{2n\tau}{3mk_{\rm B}T^2}\,\langle\, \varepsilon^2\big(\varepsilon - c_p\,T\big)\,\rangle = \frac{5n\tau k_{\rm B}^2 T}{2m} = \frac{\pi}{8}n\ell\bar{v}\,c_p \quad , \tag{8.112}$$

where $c_p = \frac{5}{2}k_{\rm B}$ and $\bar{v}^2 = 8k_{\rm B}T/\pi m$. The quantity $\kappa$ is called the *thermal conductivity*. Note that $\kappa \propto T^{1/2}$.

### 8.5.4 Viscosity

Consider the situation depicted in fig. 8.4. A fluid filling the space between two large flat plates at $z = 0$ and $z = d$ is set in motion by a force $\boldsymbol{F} = F\boldsymbol{x}$ applied to the upper plate; the lower plate is fixed. It is assumed that the fluid's velocity locally matches that of the plates. Fluid particles at the top have an average $x$-component of their momentum $\langle p_x\rangle = mV$. As these particles move downward toward lower $z$ values, they bring their $x$-momenta with them. Therefore there is a downward ($-\hat{\boldsymbol{z}}$-directed) flow of $\langle p_x\rangle$. Since $x$-momentum is constantly being drawn away from $z = d$ plane, this means that there is a $-\boldsymbol{x}$-directed *viscous drag* on the upper plate. The viscous drag force per unit area is given by $F_{\rm drag}/A = -\eta V/d$, where $V/d = \partial V_x/\partial z$ is the velocity gradient and $\eta$ is the *shear viscosity*. In steady state, the applied force balances the drag force, *i.e.* $F + F_{\rm drag} = 0$. Clearly in the steady state the net momentum density of the fluid does not change, and is given by $\frac{1}{2}\rho V\boldsymbol{x}$, where $\rho$ is the fluid mass density. The momentum per unit time injected into the fluid by the upper plate at $z = d$ is then extracted by the lower plate at $z = 0$. The *momentum flux density* $\Pi_{xz} = n\,\langle\, p_x\, v_z\,\rangle$ is the drag force on the upper surface per unit area: $\Pi_{xz} = -\eta\,\partial V_x/\partial z$. The units of viscosity are $[\eta] = M/LT$.

We now provide some formal definitions of viscosity. As we shall see presently, there is in fact a second type of viscosity, called second viscosity or *bulk viscosity*, which is measurable although not by the type of experiment depicted in fig. 8.4.

The momentum flux tensor $\Pi_{\alpha\beta} = n\,\langle\, p_\alpha\, v_\beta\,\rangle$ is defined to be the current of momentum component $p_\alpha$ in the direction of increasing $x_\beta$. For a gas in motion with average velocity $\boldsymbol{V}$, we have

$$\begin{aligned}\Pi_{\alpha\beta} &= nm\,\langle\,(V_\alpha + v'_\alpha)(V_\beta + v'_\beta)\,\rangle \\ &= nm\,V_\alpha V_\beta + nm\,\langle\, v'_\alpha v'_\beta\,\rangle \\ &= nm\,V_\alpha V_\beta + \tfrac{1}{3}nm\,\langle\, \boldsymbol{v}'^2\,\rangle\,\delta_{\alpha\beta} = \rho\,V_\alpha V_\beta + p\,\delta_{\alpha\beta} \quad ,\end{aligned} \tag{8.113}$$

where $\boldsymbol{v}'$ is the particle velocity in a frame moving with velocity $\boldsymbol{V}$, and where we have invoked the ideal gas law $p = nk_{\rm B}T$. The mass density is $\rho = nm$.

When $\boldsymbol{V}$ is spatially varying,

$$\Pi_{\alpha\beta} = p\,\delta_{\alpha\beta} + \rho\,V_\alpha V_\beta - \tilde{\sigma}_{\alpha\beta} \quad , \tag{8.114}$$

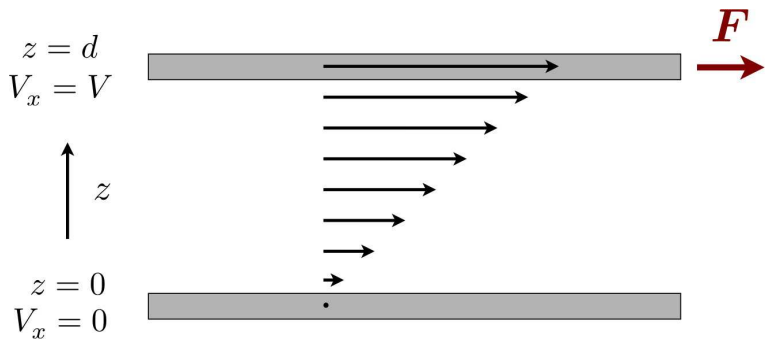



Figure 8.4: *Gedankenexperiment* to measure shear viscosity $\eta$ in a fluid. The lower plate is fixed. The viscous drag force per unit area on the upper plate is $F_{\rm drag}/A = -\eta V/d$. This must be balanced by an applied force $F$.

where $\tilde\sigma_{\alpha\beta}$ is the *viscosity stress tensor*. Any symmetric tensor, such as $\tilde\sigma_{\alpha\beta}$, can be decomposed into a sum of (i) a traceless component, and (ii) a component proportional to the identity matrix. Since $\tilde\sigma_{\alpha\beta}$ should be, to first order, linear in the spatial derivatives of the components of the velocity field $\boldsymbol{V}$, there is a unique two-parameter decomposition:

$$\begin{aligned}\tilde\sigma_{\alpha\beta} &= \eta\left(\frac{\partial V_\alpha}{\partial x_\beta} + \frac{\partial V_\beta}{\partial x_\alpha} - \tfrac{2}{3}\,\boldsymbol{\nabla}\cdot\boldsymbol{V}\,\delta_{\alpha\beta}\right) + \zeta\,\boldsymbol{\nabla}\cdot\boldsymbol{V}\,\delta_{\alpha\beta} \\ &= 2\eta\left(\mathcal{Q}_{\alpha\beta} - \tfrac{1}{3}\,\mathsf{Tr}\,(\mathcal{Q})\,\delta_{\alpha\beta}\right) + \zeta\;\mathsf{Tr}\,(\mathcal{Q})\,\delta_{\alpha\beta} \quad ,\end{aligned} \tag{8.115}$$

where $\mathcal{Q}_{\alpha\beta}$ is defined in eqn. 8.93. The coefficient of the traceless component is $\eta$, known as the *shear viscosity*. The coefficient of the component proportional to the identity is $\zeta$, known as the *bulk viscosity*. The full stress tensor $\sigma_{\alpha\beta}$ contains a contribution from the pressure:

$$\sigma_{\alpha\beta} = -p\,\delta_{\alpha\beta} + \tilde\sigma_{\alpha\beta} \quad . \tag{8.116}$$

The differential force $dF_\alpha$ that a fluid exerts on a surface element $\hat{\boldsymbol{n}}\,dA$ is $dF_\alpha = -\sigma_{\alpha\beta}\,n_\beta\,dA$, where we are using the Einstein summation convention and summing over the repeated index $\beta$. We will now compute the shear viscosity $\eta$ using the Boltzmann equation in the relaxation time approximation.

Appealing again to eqn. 8.94, with $\boldsymbol{F} = 0$ and $h = c_p T$, we find

$$\delta f = -\frac{\tau}{k_{\rm B}T}\left\{m\,v_\alpha v_\beta\,\mathcal{Q}_{\alpha\beta} + \frac{\varepsilon - c_p\,T}{T}\,\boldsymbol{v}\cdot\boldsymbol{\nabla}T - \frac{\varepsilon}{c_V/k_{\rm B}}\,\boldsymbol{\nabla}\cdot\boldsymbol{V}\right\} f^0 \quad . \tag{8.117}$$

We assume $\boldsymbol{\nabla}T = \boldsymbol{\nabla}\cdot\boldsymbol{V} = 0$, and we compute the momentum flux:

$$\begin{aligned}\Pi_{xz} &= \int d^3p\,p_x v_z\,\delta f = -\frac{nm^2\tau}{k_{\rm B}T}\,\mathcal{Q}_{\alpha\beta}\,\langle\,v_x\,v_z\,v_\alpha\,v_\beta\,\rangle \\ &= -\frac{n\tau}{k_{\rm B}T}\left(\frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x}\right)\langle\,mv_x^2\cdot mv_z^2\,\rangle = -n\tau k_{\rm B}T\left(\frac{\partial V_z}{\partial x} + \frac{\partial V_x}{\partial z}\right) \quad .\end{aligned} \tag{8.118}$$

Thus, if $V_x = V_x(z)$, we have

$$\Pi_{xz} = -n\tau k_{\rm B}T\,\frac{\partial V_x}{\partial z} \tag{8.119}$$

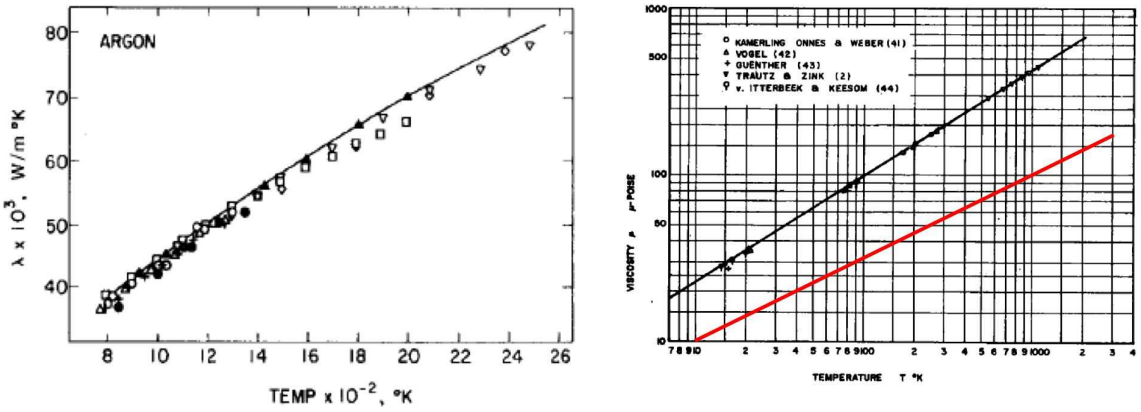



Figure 8.5: Left: thermal conductivity ($\lambda$ in figure) of Ar between $T = 800\,\mathrm{K}$ and $T = 2600\,\mathrm{K}$. The best fit to a single power law $\lambda = aT^b$ results in $b = 0.651$. Source: G. S. Springer and E. W. Wingeier, *J. Chem Phys.* **59**, 1747 (1972). Right: log-log plot of shear viscosity ($\mu$ in figure) of He between $T \approx 15\,\mathrm{K}$ and $T \approx 1000\,\mathrm{K}$. The red line has slope $\frac{1}{2}$. The slope of the data is approximately $0.633$. Source: J. Kestin and W. Leidenfrost, *Physica* **25**, 537 (1959).

from which we read off the viscosity,

$$\eta = nk_{\rm B}T\tau = \frac{\pi}{8}nm\ell\bar{v} \quad . \tag{8.120}$$

Note that $\eta(T) \propto T^{1/2}$.

How well do these predictions hold up? In fig. 8.5, we plot data for the thermal conductivity of argon and the shear viscosity of helium. Both show a clear sublinear behavior as a function of temperature, but the slope $d\log\kappa/d\log T$ is approximately $0.65$ and $d\log\eta/d\log T$ is approximately $0.63$. Clearly the simple model is not even getting the functional dependence on $T$ right, let alone its coefficient. Still, our crude theory is at least qualitatively correct.

Why do both $\kappa(T)$ as well as $\eta(T)$ decrease at low temperatures? The reason is that the heat current which flows in response to $\boldsymbol{\nabla}T$ as well as the momentum current which flows in response to $\partial V_x/\partial z$ are due to the presence of collisions, which result in momentum and energy transfer between particles. This is true even when total energy and momentum are conserved, which they are not in the relaxation time approximation. Intuitively, we might think that the viscosity should *increase* as the temperature is lowered, since common experience tells us that fluids 'gum up' as they get colder – think of honey as an extreme example. But of course honey is nothing like an ideal gas, and the physics behind the crystallization or glass transition which occurs in real fluids when they get sufficiently cold is completely absent from our approach. In our calculation, viscosity results from collisions, and with no collisions there is no momentum transfer and hence no viscosity. If, for example, the gas particles were to simply pass through each other, as though they were ghosts, then there would be no opposition to maintaining an arbitrary velocity gradient. More properly, in the dilute-gas hydrodynamic regime, collisions redistribute momentum locally and permit a Navier–Stokes description. As the collision rate tends to zero, the hydrodynamic description fails and transport becomes ballistic, so assigning a local viscosity is no longer appropriate.

### 8.5.5 Oscillating external force

Suppose a uniform oscillating external force $\boldsymbol{F}_{\rm ext}(t) = \boldsymbol{F}\, e^{-i\omega t}$ is applied. For a system of charged particles, this force would arise from an external electric field $\boldsymbol{F}_{\rm ext} = q\boldsymbol{E}\, e^{-i\omega t}$, where $q$ is the charge of each particle. We'll assume $\boldsymbol{\nabla} T = 0$. The Boltzmann equation is then written

$$\frac{\partial f}{\partial t} + \frac{\boldsymbol{p}}{m}\cdot\frac{\partial f}{\partial \boldsymbol{r}} + \boldsymbol{F}e^{-i\omega t}\cdot\frac{\partial f}{\partial \boldsymbol{p}} = -\frac{f - f^0}{\tau} \quad . \tag{8.121}$$

We again write $f = f^0 + \delta f$, and we assume $\delta f$ is spatially constant. Thus,

$$\frac{\partial\, \delta f}{\partial t} + \boldsymbol{F}e^{-i\omega t}\cdot\boldsymbol{v}\,\frac{\partial f^0}{\partial \varepsilon} = -\frac{\delta f}{\tau} \quad . \tag{8.122}$$

If we assume $\delta f(t) = \delta f(\omega)\, e^{-i\omega t}$ then the above differential equation is converted to an algebraic equation, with solution

$$\delta f(t) = -\frac{\tau\, e^{-i\omega t}}{1 - i\omega\tau}\,\frac{\partial f^0}{\partial \varepsilon}\,\boldsymbol{F}\cdot\boldsymbol{v} \quad . \tag{8.123}$$

We now compute the particle current:

$$\begin{aligned} j_\alpha(\boldsymbol{r},t) = \int d^3p\, v_\alpha\, \delta f &= \frac{\tau\, e^{-i\omega t}}{1 - i\omega\tau}\cdot\frac{F_\beta}{k_{\rm B}T}\int d^3p\, f^0(\boldsymbol{p})\, v_\alpha\, v_\beta \\ &= \frac{\tau\, e^{-i\omega t}}{1 - i\omega\tau}\cdot\frac{nF_\alpha}{3k_{\rm B}T}\int d^3v\, P(\boldsymbol{v})\,\boldsymbol{v}^2 = \frac{n\tau}{m}\cdot\frac{F_\alpha\, e^{-i\omega t}}{1 - i\omega\tau} \quad . \end{aligned} \tag{8.124}$$

If the particles are electrons, with charge $q = -e$, then the electrical current is $(-e)$ times the particle current. We then obtain

$$j_\alpha^{\rm (elec)}(t) = \frac{ne^2\tau}{m}\cdot\frac{E_\alpha\, e^{-i\omega t}}{1 - i\omega\tau} \equiv \sigma_{\alpha\beta}(\omega)\, E_\beta\, e^{-i\omega t} \quad , \tag{8.125}$$

where

$$\sigma_{\alpha\beta}(\omega) = \frac{ne^2\tau}{m}\cdot\frac{1}{1 - i\omega\tau}\,\delta_{\alpha\beta} \tag{8.126}$$

is the frequency-dependent electrical conductivity tensor. Of course for fermions such as electrons, we should be using the Fermi distribution in place of the Maxwell-Boltzmann distribution for $f^0(\boldsymbol{p})$. This affects the relation between $n$ and $\mu$ only, and the final result for the conductivity tensor $\sigma_{\alpha\beta}(\omega)$ is unchanged.

### 8.5.6 Quick and dirty calculation of transport coefficients

Suppose we have some averaged intensive quantity $\phi$ which is spatially dependent through $T(\boldsymbol{r})$ or $\mu(\boldsymbol{r})$ or $\boldsymbol{V}(\boldsymbol{r})$. For simplicity we will write $\phi = \phi(z)$. We wish to compute the current of $\phi$ across some surface whose equation is $dz = 0$. If the mean free path is $\ell$, then the value of $\phi$ for particles crossing this surface in the $+\hat{\boldsymbol{z}}$ direction is $\phi(z - \ell\cos\theta)$, where $\theta$ is the angle the particle's velocity makes with

respect to $\hat{\boldsymbol{z}}$, *i.e.* $\cos\theta = v_z/v$. We perform the same analysis for particles moving in the $-\hat{\boldsymbol{z}}$ direction, for which $\phi = \phi(z + \ell\cos\theta)$. The current of $\phi$ through this surface is then

$$\begin{aligned}\boldsymbol{j}_\phi &= n\hat{\boldsymbol{z}}\int\limits_{v_z>0} d^3v\, P(\boldsymbol{v})\, v_z\, \phi(z - \ell\cos\theta) + n\hat{\boldsymbol{z}}\int\limits_{v_z<0} d^3v\, P(\boldsymbol{v})\, v_z\, \phi(z + \ell\cos\theta) \\ &= -n\ell\,\frac{\partial\phi}{\partial z}\hat{\boldsymbol{z}}\int d^3v\, P(\boldsymbol{v})\,\frac{v_z^2}{v} = -\tfrac{1}{3}n\bar{v}\ell\,\frac{\partial\phi}{\partial z}\,\hat{\boldsymbol{z}} \quad ,\end{aligned} \tag{8.127}$$

where $\bar{v} = (8k_{\rm B}T/\pi m)^{1/2}$ is the average particle speed. If the $z$-dependence of $\phi$ comes through the dependence of $\phi$ on the local temperature $T$, then we have

$$\boldsymbol{j}_\phi = -\tfrac{1}{3}\,n\ell\bar{v}\,\frac{\partial\phi}{\partial T}\,\boldsymbol{\nabla}T \equiv -K\,\boldsymbol{\nabla}T \quad , \tag{8.128}$$

where

$$K = \tfrac{1}{3}n\ell\bar{v}\,\frac{\partial\phi}{\partial T} \tag{8.129}$$

is the transport coefficient. With $\phi = \langle\varepsilon\rangle$ we have $\partial\phi/\partial T = c_V$, where $c_V$ is the heat capacity per particle at constant volume. We then find $\boldsymbol{j}_q = -\kappa\,\boldsymbol{\nabla}T$ with thermal conductivity $\kappa = \frac{1}{3}n\ell\bar{v}\,c_V$. Our Boltzmann equation calculation (eqn. 8.112) yielded a similar result, but involving $c_p$ rather than $c_V$, and with a different prefactor.

We can make a similar argument for the viscosity. In this case $\phi = \langle p_x\rangle$ is spatially varying through its dependence on the flow velocity $\boldsymbol{V}(\boldsymbol{r})$. Clearly $\partial\phi/\partial V_x = m$, hence

$$j^z_{p_x} = \Pi_{xz} = -\tfrac{1}{3}nm\ell\bar{v}\,\frac{\partial V_x}{\partial z} \quad , \tag{8.130}$$

from which we identify the viscosity, $\eta = \frac{1}{3}nm\ell\bar{v}$. Once again, this agrees in its functional dependences with the Boltzmann equation calculation in the relaxation time approximation. Only the coefficients differ.

### 8.5.7 Thermal diffusivity, kinematic viscosity, and Prandtl number

Suppose, under conditions of constant pressure, we add heat $q$ per unit volume to an ideal gas. We know from thermodynamics that its temperature will then increase by an amount $\Delta T = q/nc_p$. If a heat current $\boldsymbol{j}_q$ flows, then the continuity equation for energy flow requires

$$nc_p\,\frac{\partial T}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j}_q = 0 \quad . \tag{8.131}$$

In a system where there is no net particle current, the heat current $\boldsymbol{j}_q$ is the same as the energy current $\boldsymbol{j}_\varepsilon$, and since $\boldsymbol{j}_\varepsilon = -\kappa\,\boldsymbol{\nabla}T$, we obtain a diffusion equation for temperature,

$$\frac{\partial T}{\partial t} = \frac{\kappa}{nc_p}\,\nabla^2 T \quad . \tag{8.132}$$

| Gas | $\eta$ ($\mu$Pa · s) | $\kappa$ (mW/m · K) | $c_p/k_{\rm B}$ | Pr |
|---|---|---|---|---|
| He | 19.5 | 149 | 2.50 | 0.682 |
| Ar | 22.3 | 17.4 | 2.50 | 0.666 |
| Xe | 22.7 | 5.46 | 2.50 | 0.659 |
| $H_2$ | 8.67 | 179 | 3.47 | 0.693 |
| $N_2$ | 17.6 | 25.5 | 3.53 | 0.721 |
| $O_2$ | 20.3 | 26.0 | 3.50 | 0.711 |
| $CH_4$ | 11.2 | 33.5 | 4.29 | 0.74 |
| $CO_2$ | 14.8 | 18.1 | 4.47 | 0.71 |
| $NH_3$ | 10.1 | 24.6 | 4.50 | 0.90 |

Table 8.1: Viscosities, thermal conductivities, and Prandtl numbers for some common gases at $T = 293\,\mathrm{K}$ and $p = 1\,\mathrm{atm}$. (Source: Table 1.1 of Smith and Jensen, with data for triatomic gases added.)

The combination $a \equiv \kappa/nc_p$ is known as the *thermal diffusivity*. Our Boltzmann equation calculation in the relaxation time approximation yielded the result $\kappa = nk_{\rm B}T\tau c_p/m$. Thus, we find $a = k_{\rm B}T\tau/m$ via this method. Note that the dimensions of $a$ are the same as for any diffusion constant $D$, namely $[a] = L^2/T$.

Another quantity with dimensions of $L^2/T$ is the *kinematic viscosity*, $\nu = \eta/\rho$, where $\rho = nm$ is the mass density. We found $\eta = nk_{\rm B}T\tau$ from the relaxation time approximation calculation, hence $\nu = k_{\rm B}T\tau/m$. The ratio $\nu/a$, called the *Prandtl number*, $\mathsf{Pr} = \eta c_p/m\kappa$, is dimensionless. According to our calculations, $\mathsf{Pr} = 1$. According to table 8.1, most monatomic gases have $\mathsf{Pr} \approx \frac{2}{3}$.

## 8.6 Diffusion and the Lorentz model

### 8.6.1 Failure of the relaxation time approximation

As we remarked above, the relaxation time approximation fails to conserve any of the collisional invariants. It is therefore unsuitable for describing hydrodynamic phenomena such as diffusion. To see this, let $f(\boldsymbol{r}, \boldsymbol{v}, t)$ be the distribution function, here written in terms of $(\boldsymbol{r}, \boldsymbol{v}, t)$ rather than $(\boldsymbol{r}, \boldsymbol{p}, t)$ as before[7]. In the absence of external forces, the Boltzmann equation in the relaxation time approximation is

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f}{\partial \boldsymbol{r}} = -\frac{f - f^0}{\tau} \quad . \tag{8.133}$$

[7]The difference is trivial, since $\boldsymbol{p} = m\boldsymbol{v}$.

We can solve this equation by first defining $\delta f(\boldsymbol{r},\boldsymbol{v},t) \equiv f(\boldsymbol{r},\boldsymbol{v},t) - f^0(\boldsymbol{v})$, and then taking the Laplace transform in time and the Fourier transform in space,

$$\delta\hat{f}(\boldsymbol{k},\boldsymbol{v},s) = \int\limits_0^\infty dt\, e^{-st} \int d^3r\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\, \delta f(\boldsymbol{r},\boldsymbol{v},t) \quad , \tag{8.134}$$

resulting in

$$\big(s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}\big)\,\delta\hat{f}(\boldsymbol{k},\boldsymbol{v},s) = \delta f(\boldsymbol{k},\boldsymbol{v},t=0) \quad \Rightarrow \quad \delta\hat{f}(\boldsymbol{k},\boldsymbol{v},s) = \frac{\delta f(\boldsymbol{k},\boldsymbol{v},t=0)}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \quad . \tag{8.135}$$

Taking the inverse transforms,

$$\begin{aligned}
\delta f(\boldsymbol{r},\boldsymbol{v},t) &= \int \frac{d^dk}{(2\pi)^d} \int\limits_{\mathcal{C}} \frac{ds}{2\pi i}\, \delta\hat{f}(\boldsymbol{k},\boldsymbol{v},s)\, e^{st} \\
&= \int \frac{d^dk}{(2\pi)^d}\, \delta f(\boldsymbol{k},\boldsymbol{v},t=0)\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \int\limits_{\mathcal{C}} \frac{ds}{2\pi i}\, \frac{e^{st}}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \\
&= \int \frac{d^dk}{(2\pi)^d}\, \delta f(\boldsymbol{k},\boldsymbol{v},t=0)\, e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{v}t)}\, e^{-t/\tau} = \delta f(\boldsymbol{r}-\boldsymbol{v}t,\boldsymbol{v},0)\, e^{-t/\tau} \quad ,
\end{aligned} \tag{8.136}$$

where $\mathcal{C}$ is an integration contour from $c - i\infty$ to $c + i\infty$ where $c$ is chosen so that $\mathcal{C}$ lies to the right of all singularities of the integrand. Thus, choosing an initial distribution that is localized at $\boldsymbol{r} = 0$ and $\boldsymbol{v} = \boldsymbol{v}_0$, we find it evolves according to

$$\delta f(\boldsymbol{r},\boldsymbol{v},0) = N\,\delta(\boldsymbol{r})\,\delta(\boldsymbol{v}-\boldsymbol{v}_0) \quad \Rightarrow \quad \delta f(\boldsymbol{r},\boldsymbol{v},t) = N\,\delta(\boldsymbol{r}-\boldsymbol{v}_0t)\,\delta(\boldsymbol{v}-\boldsymbol{v}_0)\,e^{-t/\tau} \quad . \tag{8.137}$$

This result is profoundly unphysical – it says that particle number is not conserved. Not only that, but the spatial distribution remains instantaneously localized about $\boldsymbol{r} = \boldsymbol{v}t$, whereas we expect that in a model of random impurity elastic scattering the particle velocity distribution should become isotropic.

### 8.6.2 Modified Boltzmann equation and its solution

To remedy this unphysical aspect, consider the modified Boltzmann equation,

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\cdot\frac{\partial f}{\partial \boldsymbol{r}} = \frac{1}{\tau}\left[-f + \int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, f\right] \equiv \frac{1}{\tau}\big(\mathbb{P} - 1\big) f \quad , \tag{8.138}$$

where $\mathbb{P}$ is a projector onto a space of isotropic functions of $\boldsymbol{v}$: $\mathbb{P}F = \int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, F(\boldsymbol{v}) = \int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, F(v\,\hat{\boldsymbol{v}})$ for any function $F(\boldsymbol{v})$. Note that $\mathbb{P}F$ is a function of the speed $v = |\boldsymbol{v}|$. Since the equilibrium distribution $f^0(\boldsymbol{v}) = f^0(v)$ is the Maxwell distribution, it is isotropic in velocity space, *i.e.* $f^0(\boldsymbol{v}) = f^0(v)$, we have that $(\mathbb{P}-1)f^0(\boldsymbol{v}) = 0$, and thus the above modified Boltzmann equation holds for $f$ as well as for $\delta f$. Note that the number density $n(\boldsymbol{r},t) = \int d^3v\, f(\boldsymbol{r},\boldsymbol{v},t)$ and number current density $\boldsymbol{j}(\boldsymbol{r},t) = \int d^3v\, f(\boldsymbol{r},\boldsymbol{v},t)\,\boldsymbol{v}$ satisfy the continuity equation

$$\frac{\partial n(\boldsymbol{r},t)}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j}(\boldsymbol{r},t) = 0 \quad , \tag{8.139}$$

since integrating over $\hat{\boldsymbol{v}}$ annihilates the RHS of eqn. 8.138 for all $(\boldsymbol{r}, v)$. Thus, total particle number is conserved, which is not the case in the naïve relaxation time approximation.

The model in eqn. 8.138 is known as the *Lorentz model*[8]. To solve it, we again begin with the double Laplace-Fourier transform,

$$\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) = \int_0^\infty dt\, e^{-st} \int d^3r\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\, f(\boldsymbol{r}, \boldsymbol{v}, t) \quad . \tag{8.140}$$

Applying this transform to eqn. 8.138, we obtain

$$\left(s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}\right) \hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) = \tau^{-1}\, \mathbb{P}\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) + f(\boldsymbol{k}, \boldsymbol{v}, t=0) \quad . \tag{8.141}$$

We now solve for $\mathbb{P}\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s)$:

$$\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) = \frac{\tau^{-1}}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}}\, \mathbb{P}\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) + \frac{f(\boldsymbol{k}, \boldsymbol{v}, t=0)}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \quad , \tag{8.142}$$

which entails

$$\mathbb{P}\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) = \left[\int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, \frac{\tau^{-1}}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}}\right] \mathbb{P}\hat{f}(\boldsymbol{k}, \boldsymbol{v}, \boldsymbol{s}) + \int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, \frac{f(\boldsymbol{k}, \boldsymbol{v}, t=0)}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \quad . \tag{8.143}$$

Now we have

$$\int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, \frac{\tau^{-1}}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} = \int_{-1}^{1} dx\, \frac{\tau^{-1}}{s + ivkx + \tau^{-1}} = \frac{1}{vk} \tan^{-1}\!\left(\frac{vk\tau}{1+\tau s}\right) \quad . \tag{8.144}$$

Thus,

$$\mathbb{P}f(\boldsymbol{k}, \boldsymbol{v}, s) = \left[1 - \frac{1}{vk\tau} \tan^{-1}\!\left(\frac{vk\tau}{1+\tau s}\right)\right]^{-1} \int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, \frac{f(\boldsymbol{k}, \boldsymbol{v}, t=0)}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \quad . \tag{8.145}$$

We now have the solution to Lorentz's modified Boltzmann equation:

$$\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) = \frac{\tau^{-1}}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \left[1 - \frac{1}{vk\tau} \tan^{-1}\!\left(\frac{vk\tau}{1+\tau s}\right)\right]^{-1} \int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, \frac{f(\boldsymbol{k}, \boldsymbol{v}, t=0)}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} + \frac{f(\boldsymbol{k}, \boldsymbol{v}, t=0)}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \quad . \tag{8.146}$$

Thus,

$$f(\boldsymbol{r}, \boldsymbol{v}, t) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \int_{\mathcal{C}} \frac{ds}{2\pi i}\, e^{st} \left\{ \frac{\tau^{-1}}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \left[1 - \frac{1}{vk\tau} \tan^{-1}\!\left(\frac{vk\tau}{1+\tau s}\right)\right]^{-1} \int \frac{d\hat{\boldsymbol{v}}}{4\pi}\, \frac{f(\boldsymbol{k}, \boldsymbol{v}, t=0)}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \right\}$$
$$+ f(\boldsymbol{r} - \boldsymbol{v}t\, \boldsymbol{v}, 0)\, e^{-t/\tau} \quad . \tag{8.147}$$

[8] See the excellent discussion in the book by Krapivsky, Redner, and Ben-Naim, cited in §8.1.

Note that the last term $f(\boldsymbol{r} - \boldsymbol{v}t, \boldsymbol{v}, 0)\,\exp(-t/\tau)$ vanishes as $t \to \infty$. However, we have already seen that the full expression must satisfy the continuity equation for $n(\boldsymbol{r}, t)$. We will now show that in the long time limit, when the exponentially decaying transient may be neglected, the remaining result is an expanding cloud describing isotropic particle diffusion.

Let us again consider an initial distribution which is perfectly localized in both $\boldsymbol{r}$ and $\boldsymbol{v}$:

$$f(\boldsymbol{r}, \boldsymbol{v}, t = 0) = N\,\delta(\boldsymbol{r})\,\delta(\boldsymbol{v} - \boldsymbol{v}_0) \quad . \tag{8.148}$$

For these initial conditions, we find

$$\int \frac{d\hat{\boldsymbol{v}}}{4\pi}\,\frac{f(\boldsymbol{k}, \boldsymbol{v}, t = 0)}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} = \frac{1}{s + i\boldsymbol{v}_0\cdot\boldsymbol{k} + \tau^{-1}} \cdot \frac{N\,\delta(v - v_0)}{4\pi v_0^2} \quad . \tag{8.149}$$

We are interested in the long time limit of $f(\boldsymbol{r}, v, t)$, where $t \gg \tau$. Long times are dominated in the Laplace transform by $s \sim t^{-1}$. We also assume the distribution in space becomes smooth, as we shall show, and expand in the regime where $s\tau \ll 1$ and $vk\tau \ll 1$. We then have

$$1 - \frac{1}{vk\tau}\tan^{-1}\left(\frac{vk\tau}{1 + \tau s}\right) = s\tau + \tfrac{1}{3}k^2v^2\tau^2 + \ldots \quad , \tag{8.150}$$

and therefore

$$\begin{aligned}\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) &= \frac{\tau^{-1}}{s + i\boldsymbol{v}\cdot\boldsymbol{k} + \tau^{-1}} \cdot \frac{\tau^{-1}}{s + i\boldsymbol{v}_0\cdot\boldsymbol{k} + \tau^{-1}} \cdot \frac{1}{s + \frac{1}{3}v_0^2\,k^2\,\tau + \ldots} \cdot \frac{N\,\delta(v - v_0)}{4\pi v_0^2} + \frac{N\,\delta(\boldsymbol{v} - \boldsymbol{v}_0)}{s + i\boldsymbol{v}_0\cdot\boldsymbol{k} + \tau^{-1}} \\ &\approx \frac{1}{s + \frac{1}{3}v_0^2\,k^2\,\tau} \cdot \frac{N\,\delta(v - v_0)}{4\pi v_0^2} + \frac{N\,\delta(\boldsymbol{v} - \boldsymbol{v}_0)}{s + i\boldsymbol{v}_0\cdot\boldsymbol{k} + \tau^{-1}} \quad ,\end{aligned} \tag{8.151}$$

since $\tau^{-1}$ is dominant over $s$ and $i\boldsymbol{v}\cdot\boldsymbol{k}$ in the denominators of the first two multiplicative factors on the RHS of the top equation. We then have

$$\hat{f}(\boldsymbol{k}, \boldsymbol{v}, s) \approx \frac{1}{s + \frac{1}{3}v_0^2\,k^2\,\tau} \cdot \frac{N\,\delta(v - v_0)}{4\pi v_0^2} + \frac{N\,\delta(\boldsymbol{v} - \boldsymbol{v}_0)}{s + i\boldsymbol{v}_0\cdot\boldsymbol{k} + \tau^{-1}} \quad . \tag{8.152}$$

Performing the inverse Laplace and Fourier transforms, and dropping the transient term for $t \gg \tau$, we obtain our final result,

$$f(\boldsymbol{r}, \boldsymbol{v}, t \gg \tau) = (4\pi Dt)^{-3/2}\,e^{-r^2/4Dt} \cdot \frac{N\,\delta(v - v_0)}{4\pi v_0^2} \quad , \tag{8.153}$$

where the *diffusion constant* is $D = \frac{1}{3}v_0^2\,\tau$. The units are $[D] = L^2/T$. Integrating over velocities, we have the density

$$n(\boldsymbol{r}, t \gg \tau) = \int d^3v\; f(\boldsymbol{r}, \boldsymbol{v}, t \gg \tau) = N\,(4\pi Dt)^{-3/2}\,e^{-r^2/4Dt} \quad . \tag{8.154}$$

Note that $\int d^3r\; n(\boldsymbol{r}, t \gg \tau) = N$ at all times in this limit. In addition to particle number being conserved, we see that the late time distribution $f(\boldsymbol{r}, \boldsymbol{v}, t \gg \tau)$ is isotropic in both $\boldsymbol{r}$ as well as $\boldsymbol{v}$.

## 8.7 Linearized Boltzmann Equation

### 8.7.1 Linearizing the collision integral

We now return to the classical Boltzmann equation and consider a more formal treatment of the collision term in the linear approximation. We will assume time-reversal symmetry, in which case

$$\left(\frac{df}{dt}\right)_{\rm coll} = \int d^3p_1 \int d^3p' \int d^3p_1' \; w(\boldsymbol{p}', \boldsymbol{p}_1' \,|\, \boldsymbol{p}, \boldsymbol{p}_1) \left\{ f(\boldsymbol{p}')\, f(\boldsymbol{p}_1') - f(\boldsymbol{p})\, f(\boldsymbol{p}_1) \right\} \quad . \tag{8.155}$$

The collision integral is nonlinear in the distribution $f$. We *linearize* by writing

$$f(\boldsymbol{p}) = f^0(\boldsymbol{p}) + f^0(\boldsymbol{p})\,\psi(\boldsymbol{p}) \quad , \tag{8.156}$$

where we assume $\psi(\boldsymbol{p})$ is small. We then have, to first order in $\psi$,

$$\left(\frac{df}{dt}\right)_{\rm coll} = f^0(\boldsymbol{p})\,\hat{L}\psi + \mathcal{O}(\psi^2) \quad , \tag{8.157}$$

where the action of the *linearized collision operator* is given by

$$\begin{aligned} \hat{L}\psi &= \int d^3p_1 \int d^3p' \int d^3p_1' \; w(\boldsymbol{p}', \boldsymbol{p}_1' \,|\, \boldsymbol{p}, \boldsymbol{p}_1)\, f^0(\boldsymbol{p}_1) \left\{ \psi(\boldsymbol{p}') + \psi(\boldsymbol{p}_1') - \psi(\boldsymbol{p}) - \psi(\boldsymbol{p}_1) \right\} \\ &= \int d^3p_1 \int d\Omega \, |\boldsymbol{v} - \boldsymbol{v}_1| \, \frac{d\sigma}{d\Omega}\, f^0(\boldsymbol{p}_1) \left\{ \psi(\boldsymbol{p}') + \psi(\boldsymbol{p}_1') - \psi(\boldsymbol{p}) - \psi(\boldsymbol{p}_1) \right\} \quad , \end{aligned} \tag{8.158}$$

where we have invoked eqn. 8.57 to write the RHS in terms of the differential scattering cross section. In deriving the above result, we have made use of the detailed balance relation,

$$f^0(\boldsymbol{p})\, f^0(\boldsymbol{p}_1) = f^0(\boldsymbol{p}')\, f^0(\boldsymbol{p}_1') \quad . \tag{8.159}$$

We have also suppressed the $\boldsymbol{r}$ dependence in writing $f(\boldsymbol{p})$, $f^0(\boldsymbol{p})$, and $\psi(\boldsymbol{p})$.

From eqn. 8.94, we then have the linearized equation

$$\left(\hat{L} - \frac{\partial}{\partial t}\right)\psi = Y, \tag{8.160}$$

where, for point particles,

$$Y = \frac{1}{k_{\rm B}T}\left\{ \frac{\varepsilon(\boldsymbol{p}) - c_p T}{T}\, \boldsymbol{v}\cdot\boldsymbol{\nabla} T + m\, v_\alpha v_\beta\, \mathcal{Q}_{\alpha\beta} - \frac{k_{\rm B}\,\varepsilon(\boldsymbol{p})}{c_V}\, \boldsymbol{\nabla}\cdot\boldsymbol{V} - \boldsymbol{F}\cdot\boldsymbol{v} \right\} \quad . \tag{8.161}$$

Eqn. 8.160 is an inhomogeneous linear equation, which can be solved by inverting the operator $\hat{L} - \frac{\partial}{\partial t}$.

### 8.7.2 Linear algebraic properties of $\hat{L}$

Although $\hat{L}$ is an integral operator, it shares many properties with other linear operators with which you are familiar, such as matrices and differential operators. We can define an *inner product*[9],

$$\langle\, \psi_1 \,|\, \psi_2 \,\rangle \equiv \int d^3p\, f^0(\boldsymbol{p})\, \psi_1(\boldsymbol{p})\, \psi_2(\boldsymbol{p}) \quad . \tag{8.162}$$

Note that this is not the usual Hilbert space inner product from quantum mechanics, since the factor $f^0(\boldsymbol{p})$ is included in the metric. This is necessary in order that $\hat{L}$ be *self-adjoint*: $\langle\, \psi_1 \,|\, \hat{L}\psi_2 \,\rangle = \langle\, \hat{L}\psi_1 \,|\, \psi_2 \,\rangle$.

We can now define the spectrum of *normalized eigenfunctions* of $\hat{L}$, which we write as $\phi_n(\boldsymbol{p})$. The eigenfunctions satisfy the eigenvalue equation,

$$\hat{L}\,\phi_n = -\lambda_n\,\phi_n \quad , \tag{8.163}$$

and may be chosen to be orthonormal, *i.e.* $\langle\, \phi_m \,|\, \phi_n \,\rangle = \delta_{mn}$ . Of course, in order to obtain the eigenfunctions $\phi_n$ we must have detailed knowledge of the function $w(\boldsymbol{p}', \boldsymbol{p}_1' \,|\, \boldsymbol{p}, \boldsymbol{p}_1)$.

Recall that there are five collisional invariants, which are the particle number, the three components of the total particle momentum, and the particle energy. To each collisional invariant, there is an associated eigenfunction $\phi_n$ with eigenvalue $\lambda_n = 0$. One can check that these normalized eigenfunctions are

$$\phi_n(\boldsymbol{p}) = \frac{1}{\sqrt{n}} \qquad , \qquad \phi_{p_\alpha}(\boldsymbol{p}) = \frac{p_\alpha}{\sqrt{nmk_{\rm B}T}} \qquad , \qquad \phi_\varepsilon(\boldsymbol{p}) = \sqrt{\frac{2}{3n}}\left(\frac{\varepsilon(\boldsymbol{p})}{k_{\rm B}T} - \frac{3}{2}\right) \quad . \tag{8.164}$$

If there are no temperature, chemical potential, or bulk velocity gradients, and there are no external forces, then $Y = 0$ and the only changes to the distribution are from collisions. The linearized Boltzmann equation becomes

$$\frac{\partial\psi}{\partial t} = \hat{L}\psi \quad . \tag{8.165}$$

We can therefore write the most general solution in the form

$$\psi(\boldsymbol{p}, t) = {\sum_n}' C_n\, \phi_n(\boldsymbol{p})\, e^{-\lambda_n t} \quad , \tag{8.166}$$

where the prime on the sum reminds us that collisional invariants are to be excluded. All the eigenvalues $\lambda_n$, aside from the five zero eigenvalues for the collisional invariants, must be positive. Any negative eigenvalue would cause $\psi(\boldsymbol{p}, t)$ to increase without bound, and an initial nonequilibrium distribution would not relax to the equilibrium $f^0(\boldsymbol{p})$, which we regard as unphysical. Henceforth we will drop the prime on the sum but remember that $C_n = 0$ for the five collisional invariants.

Recall also the particle, energy, and thermal (heat) currents,

$$\begin{aligned}
\boldsymbol{j} &= \int d^3p\, \boldsymbol{v}\, f(\boldsymbol{p}) = \int d^3p\, f^0(\boldsymbol{p})\, \boldsymbol{v}\, \psi(\boldsymbol{p}) = \langle\, \boldsymbol{v} \,|\, \psi \,\rangle \\
\boldsymbol{j}_\varepsilon &= \int d^3p\, \boldsymbol{v}\, \varepsilon\, f(\boldsymbol{p}) = \int d^3p\, f^0(\boldsymbol{p})\, \boldsymbol{v}\, \varepsilon\, \psi(\boldsymbol{p}) = \langle\, \boldsymbol{v}\, \varepsilon \,|\, \psi \,\rangle \\
\boldsymbol{j}_q &= \int d^3p\, \boldsymbol{v}\, (\varepsilon - \mu)\, f(\boldsymbol{p}) = \int d^3p\, f^0(\boldsymbol{p})\, \boldsymbol{v}\, (\varepsilon - \mu)\, \psi(\boldsymbol{p}) = \langle\, \boldsymbol{v}\, (\varepsilon - \mu) \,|\, \psi \,\rangle \quad .
\end{aligned} \tag{8.167}$$

[9]The requirements of an inner product $\langle f|g\rangle$ are symmetry, linearity, and non-negative definiteness.

Note $\boldsymbol{j}_q = \boldsymbol{j}_\varepsilon - \mu \boldsymbol{j}$.

### 8.7.3 Steady state solution to the linearized Boltzmann equation

Under steady state conditions, there is no time dependence, and the linearized Boltzmann equation takes the form $\hat{L}\psi = Y$. We may expand $\psi$ in the eigenfunctions $\phi_n$ and write $\psi = \sum_n C_n \, \phi_n$. Applying $\hat{L}$ and taking the inner product with $\phi_j$, we have

$$C_j = -\frac{1}{\lambda_j} \langle \, \phi_j \, | \, Y \, \rangle \quad . \tag{8.168}$$

Thus, the formal solution to the linearized Boltzmann equation is

$$\psi(\boldsymbol{p}) = -\sum_n \frac{1}{\lambda_n} \langle \, \phi_n \, | \, Y \, \rangle \, \phi_n(\boldsymbol{p}) \quad . \tag{8.169}$$

This solution is applicable provided $| \, Y \, \rangle$ is orthogonal to the five collisional invariants.

#### Thermal conductivity

For the thermal conductivity, we take $\boldsymbol{\nabla} T = \partial_x T \, \hat{\boldsymbol{x}}$, and

$$Y = \frac{1}{k_{\rm B} T^2} \frac{\partial T}{\partial x} \cdot X_\kappa \quad , \tag{8.170}$$

where $X_\kappa \equiv (\varepsilon - c_p T) \, v_x$. Under the conditions of no particle flow ($\boldsymbol{j} = 0$), we have $\boldsymbol{j}_q = -\kappa \, \partial_x T \, \hat{\boldsymbol{x}}$. Then we have

$$\langle \, X_\kappa \, | \, \psi \, \rangle = -\kappa \, \frac{\partial T}{\partial x} \quad . \tag{8.171}$$

#### Viscosity

For the viscosity, we take

$$Y = \frac{m}{k_{\rm B} T} \frac{\partial V_x}{\partial y} \cdot X_\eta \quad , \tag{8.172}$$

with $X_\eta = v_x \, v_y$. We then

$$\Pi_{xy} = \langle \, m \, v_x \, v_y \, | \, \psi \, \rangle = -\eta \, \frac{\partial V_x}{\partial y} \quad . \tag{8.173}$$

Thus,

$$\langle \, X_\eta \, | \, \psi \, \rangle = -\frac{\eta}{m} \frac{\partial V_x}{\partial y} \quad . \tag{8.174}$$

### 8.7.4 Variational approach

Following the treatment in chapter 1 of Smith and Jensen, define $\hat{H} \equiv -\hat{L}$. We have that $\hat{H}$ is a positive semidefinite operator, whose only zero eigenvalues correspond to the collisional invariants. We then have the Schwarz inequality,

$$\langle\,\psi\,|\,\hat{H}\,|\,\psi\,\rangle \cdot \langle\,\phi\,|\,\hat{H}\,|\,\phi\,\rangle \geq \langle\,\phi\,|\,\hat{H}\,|\,\psi\,\rangle^2 \quad , \tag{8.175}$$

for any two Hilbert space vectors $|\,\psi\,\rangle$ and $|\,\phi\,\rangle$. Consider now the above calculation of the thermal conductivity. We have

$$\hat{H}\psi = -\frac{1}{k_{\rm B}T^2}\frac{\partial T}{\partial x}\,X_\kappa \tag{8.176}$$

and therefore

$$\kappa = \frac{k_{\rm B}T^2}{(\partial T/\partial x)^2}\,\langle\,\psi\,|\,\hat{H}\,|\,\psi\,\rangle \geq \frac{1}{k_{\rm B}T^2}\,\frac{\langle\,\phi\,|\,X_\kappa\,\rangle^2}{\langle\,\phi\,|\,\hat{H}\,|\,\phi\,\rangle} \quad . \tag{8.177}$$

Similarly, for the viscosity, we have

$$\hat{H}\psi = -\frac{m}{k_{\rm B}T}\frac{\partial V_x}{\partial y}\,X_\eta \quad , \tag{8.178}$$

from which we derive

$$\eta = \frac{k_{\rm B}T}{(\partial V_x/\partial y)^2}\,\langle\,\psi\,|\,\hat{H}\,|\,\psi\,\rangle \geq \frac{m^2}{k_{\rm B}T}\,\frac{\langle\,\phi\,|\,X_\eta\,\rangle^2}{\langle\,\phi\,|\,\hat{H}\,|\,\phi\,\rangle} \quad . \tag{8.179}$$

In order to get a good lower bound, we want $\phi$ in each case to have a good overlap with $X_{\kappa,\eta}$. One approach then is to take $\phi = X_{\kappa,\eta}$, which guarantees that the overlap will be finite (and not zero due to symmetry, for example). We illustrate this method with the viscosity calculation. We have

$$\eta \geq \frac{m^2}{k_{\rm B}T}\,\frac{\langle\,v_xv_y\,|\,v_xv_y\,\rangle^2}{\langle\,v_xv_y\,|\,\hat{H}\,|\,v_xv_y\,\rangle} \quad . \tag{8.180}$$

With $\hat{H} = -\hat{L}$ we have

$$\langle\,\phi\,|\,\hat{H}\,|\,\psi\,\rangle = \int d^3p\,f^0(\boldsymbol{p})\,\phi(\boldsymbol{p})\int d^3p_1\int d\Omega\,\frac{d\sigma}{d\Omega}\,|\boldsymbol{v}-\boldsymbol{v}_1|\,f^0(\boldsymbol{p}_1)\left\{\psi(\boldsymbol{p})+\psi(\boldsymbol{p}_1)-\psi(\boldsymbol{p}')-\psi(\boldsymbol{p}_1')\right\} \quad . \tag{8.181}$$

Here the kinematics of the collision guarantee total energy and momentum conservation, so $\boldsymbol{p}'$ and $\boldsymbol{p}_1'$ are determined as in eqn. 8.59.

We have $d\Omega = \sin\chi\,d\chi\,d\varphi$, where $\chi$ is the scattering angle depicted in fig. 8.6 and $\varphi$ is the azimuthal angle of the scattering. The differential scattering cross section is obtained by elementary mechanics and is known to be

$$\frac{d\sigma}{d\Omega} = \left|\frac{d(b^2/2)}{d\cos\chi}\right| \quad , \tag{8.182}$$

where $b$ is the *impact parameter*. The scattering angle is

$$\chi(b,u) = \pi - 2\int\limits_{r_{\rm p}}^{\infty} dr\,\frac{b}{\sqrt{r^4 - b^2r^2 - \frac{2U(r)r^4}{\tilde{m}u^2}}} \quad , \tag{8.183}$$

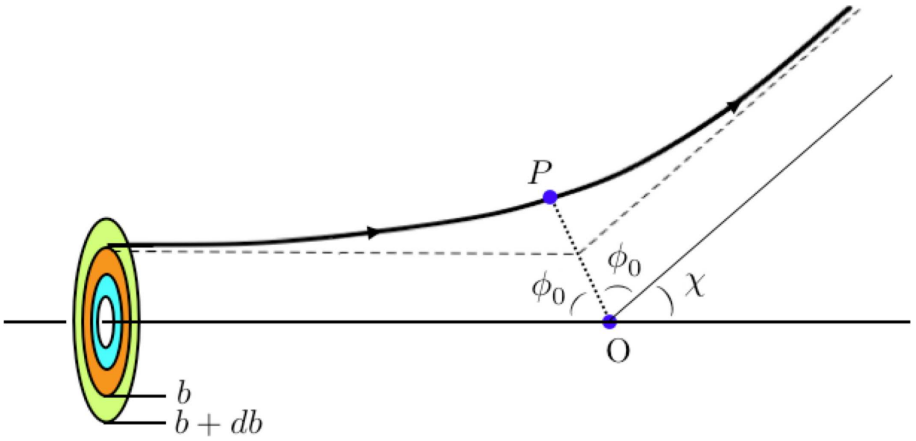



Figure 8.6: Scattering in the CM frame. O is the force center and $P$ is the point of periapsis. The impact parameter is $b$, and $\chi$ is the scattering angle. $\phi_0$ is the angle through which the relative coordinate moves between periapsis and infinity.

where $\tilde{m} = \frac{1}{2}m$ is the reduced mass, and $r_{\rm p}$ is the relative coordinate separation at periapsis, *i.e.* the distance of closest approach, which occurs when $\dot{r} = 0$, *i.e.*

$$\tfrac{1}{2}\tilde{m}u^2 = \frac{\ell^2}{2\tilde{m}r_{\rm p}^2} + U(r_{\rm p}) \quad , \tag{8.184}$$

where $\ell = \tilde{m}ub$ is the relative coordinate angular momentum.

We work in center-of-mass coordinates, so the velocities are

$$\begin{aligned} \boldsymbol{v} &= \boldsymbol{V} + \tfrac{1}{2}\boldsymbol{u} \qquad & \boldsymbol{v}' &= \boldsymbol{V} + \tfrac{1}{2}\boldsymbol{u}' \\ \boldsymbol{v}_1 &= \boldsymbol{V} - \tfrac{1}{2}\boldsymbol{u} \qquad & \boldsymbol{v}_1' &= \boldsymbol{V} - \tfrac{1}{2}\boldsymbol{u}' \end{aligned} \quad , \tag{8.185}$$

with $|\boldsymbol{u}| = |\boldsymbol{u}'|$ and $\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{u}}' = \cos\chi$. Then if $\psi(\boldsymbol{p}) = v_x v_y$, we have

$$\Delta(\psi) \equiv \psi(\boldsymbol{p}) + \psi(\boldsymbol{p}_1) - \psi(\boldsymbol{p}') - \psi(\boldsymbol{p}_1') = \tfrac{1}{2}\big(u_x u_y - u_x' u_y'\big) \quad . \tag{8.186}$$

We may write

$$\boldsymbol{u}' = u\left(\sin\chi\cos\varphi\,\hat{\mathbf{e}}_1 + \sin\chi\sin\varphi\,\hat{\mathbf{e}}_2 + \cos\chi\,\hat{\mathbf{e}}_3\right) \quad , \tag{8.187}$$

where $\hat{\mathbf{e}}_3 = \hat{\boldsymbol{u}}$. With this parameterization, we have

$$\int_0^{2\pi} d\varphi\,\tfrac{1}{2}\big(u_\alpha u_\beta - u_\alpha' u_\beta'\big) = -\pi\sin^2\chi\,\left(u^2\,\delta_{\alpha\beta} - 3u_\alpha u_\beta\right) \quad . \tag{8.188}$$

Note that we have used here the relation

$$e_{1\alpha}\,e_{1\beta} + e_{2\alpha}\,e_{2\beta} + e_{3\alpha}\,e_{3\beta} = \delta_{\alpha\beta} \quad , \tag{8.189}$$

which holds since the LHS is a projector $\sum_{i=1}^3 |\hat{\mathbf{e}}_i\rangle\langle\hat{\mathbf{e}}_i|$.

It is convenient to define the following integral:

$$R(u) \equiv \int_0^\infty db\, b\, \sin^2\chi(b,u) \quad . \tag{8.190}$$

Since the Jacobian satisfies

$$\left| \det \frac{(\partial \boldsymbol{v}, \partial \boldsymbol{v}_1)}{(\partial \boldsymbol{V}, \partial \boldsymbol{u})} \right| = 1 \quad , \tag{8.191}$$

we have

$$\langle\, v_x v_y \,|\, \hat{H} \,|\, v_x v_y \,\rangle = n^2 \left( \frac{m}{2\pi k_{\rm B} T} \right)^3 \int d^3V \int d^3u\, e^{-m\boldsymbol{V}^2/k_{\rm B}T}\, e^{-m\boldsymbol{u}^2/4k_{\rm B}T} \cdot u \cdot \tfrac{3\pi}{2} u_x u_y \cdot R(u) \cdot v_x v_y \quad . \tag{8.192}$$

This yields

$$\langle\, v_x v_y \,|\, \hat{H} \,|\, v_x v_y \,\rangle = \tfrac{\pi}{40}\, n^2 \left\langle u^5\, R(u) \right\rangle \quad , \tag{8.193}$$

where

$$\left\langle F(u) \right\rangle \equiv \int_0^\infty du\, u^2\, e^{-mu^2/4k_{\rm B}T}\, F(u) \Big/ \int_0^\infty du\, u^2\, e^{-mu^2/4k_{\rm B}T} \quad . \tag{8.194}$$

It is easy to compute the term in the numerator of eqn. 8.180:

$$\langle\, v_x v_y \,|\, v_x v_y \,\rangle = n \left( \frac{m}{2\pi k_{\rm B} T} \right)^{3/2} \int d^3v\, e^{-m\boldsymbol{v}^2/2k_{\rm B}T}\, v_x^2\, v_y^2 = n \left( \frac{k_{\rm B} T}{m} \right)^2 \quad . \tag{8.195}$$

Putting it all together, we find

$$\eta \geq \frac{40\,(k_{\rm B}T)^3}{\pi\, m^2} \Big/ \left\langle u^5\, R(u) \right\rangle \quad . \tag{8.196}$$

The computation for $\kappa$ is a bit more tedious. One has $\psi(\boldsymbol{p}) = (\varepsilon - c_p T)\, v_x$, in which case

$$\Delta(\psi) = \tfrac{1}{2} m \Big[ (\boldsymbol{V} \cdot \boldsymbol{u})\, u_x - (\boldsymbol{V} \cdot \boldsymbol{u}')\, u_x' \Big] \quad . \tag{8.197}$$

Ultimately, one obtains the lower bound

$$\kappa \geq \frac{150\, k_{\rm B}\, (k_{\rm B}T)^3}{\pi\, m^3} \Big/ \left\langle u^5\, R(u) \right\rangle \quad . \tag{8.198}$$

Thus, independent of the potential, this variational calculation yields a Prandtl number of

$$\mathsf{Pr} = \frac{\nu}{a} = \frac{\eta\, c_p}{m\, \kappa} = \tfrac{2}{3} \quad , \tag{8.199}$$

which is very close to what is observed in dilute monatomic gases (see Tab. 8.1).

While the variational expressions for $\eta$ and $\kappa$ are complicated functions of the potential, for hard sphere scattering the calculation is simple, because $b = d\sin\phi_0 = d\cos(\frac{1}{2}\chi)$, where $d$ is the hard sphere diameter. Thus, the impact parameter $b$ is independent of the relative speed $u$, and one finds $R(u) = \frac{1}{3}d^2$. Then

$$\big\langle u^5\,R(u)\big\rangle = \tfrac{1}{3}d^2\big\langle u^5\big\rangle = \frac{128}{\sqrt{\pi}}\left(\frac{k_{\rm B}T}{m}\right)^{5/2} d^2 \tag{8.200}$$

and one finds

$$\eta \geq \frac{5\,(mk_{\rm B}T)^{1/2}}{16\sqrt{\pi}\,d^2} \quad , \quad \kappa \geq \frac{75\,k_{\rm B}}{64\sqrt{\pi}\,d^2}\left(\frac{k_{\rm B}T}{m}\right)^{1/2} \quad . \tag{8.201}$$

## 8.8 The Equations of Hydrodynamics

We now derive the equations governing fluid flow. The equations of mass and momentum balance are

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot(\rho\,\boldsymbol{V}) = 0 \tag{8.202}$$

$$\frac{\partial(\rho\,V_\alpha)}{\partial t} + \frac{\partial\Pi_{\alpha\beta}}{\partial x^\beta} = 0 \quad , \tag{8.203}$$

where

$$\Pi_{\alpha\beta} = \rho\,V_\alpha V_\beta + p\,\delta_{\alpha\beta} - \overbrace{\left\{\eta\left(\frac{\partial V_\alpha}{\partial x_\beta} + \frac{\partial V_\beta}{\partial x_\alpha} - \tfrac{2}{3}\,\boldsymbol{\nabla}\cdot\boldsymbol{V}\,\delta_{\alpha\beta}\right) + \zeta\,\boldsymbol{\nabla}\cdot\boldsymbol{V}\,\delta_{\alpha\beta}\right\}}^{\tilde\sigma_{\alpha\beta}} \quad . \tag{8.204}$$

Substituting the continuity equation into the momentum balance equation, one arrives at

$$\rho\,\frac{\partial\boldsymbol{V}}{\partial t} + \rho\,(\boldsymbol{V}\cdot\boldsymbol{\nabla})\boldsymbol{V} = -\boldsymbol{\nabla}p + \eta\,\nabla^2\boldsymbol{V} + (\zeta + \tfrac{1}{3}\eta)\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{V}) \quad , \tag{8.205}$$

which, together with continuity, are known as the *Navier-Stokes equations*. These equations are supplemented by an equation describing the conservation of energy,

$$T\frac{\partial s}{\partial t} + T\,\boldsymbol{\nabla}\cdot(s\boldsymbol{V}) = \tilde\sigma_{\alpha\beta}\,\frac{\partial V_\alpha}{\partial x^\beta} + \boldsymbol{\nabla}\cdot(\kappa\boldsymbol{\nabla}T) \quad , \tag{8.206}$$

where $\tilde\sigma_{\alpha\beta}$ is the viscosity stress tensor of eqn. 8.115 (also identified in eqn. 8.204 above). Note that the LHS of eqn. 8.205 is $\rho\,D\boldsymbol{V}/Dt$, where $D/Dt$ is the convective derivative. Multiplying by a differential volume, this gives the mass times the acceleration of a differential local fluid element. The RHS, multiplied by the same differential volume, gives the differential force on this fluid element in a frame instantaneously moving with constant velocity $\boldsymbol{V}$. Thus, this is Newton's Second Law for the fluid.

If the fluid is incompressible, then $\boldsymbol{\nabla}\cdot\boldsymbol{V} = 0$, and with the kinematic viscosity $\nu \equiv \eta/\rho$ we have the incompressible Navier-Stokes equations,

$$\frac{\partial\boldsymbol{V}}{\partial t} + (\boldsymbol{V}\cdot\boldsymbol{\nabla})\boldsymbol{V} = -\frac{1}{\rho}\,\boldsymbol{\nabla}p + \nu\,\nabla^2\boldsymbol{V} \quad . \tag{8.207}$$

Since $[\eta] = M/LT$ and $[\rho] = M/L^3$, we have $[\nu] = L^2/T$. At $T = 20^\circ$C,

$$\nu_{\rm H_2O} = 10^{-2}\,{\rm cm}^2/{\rm sec} \quad , \quad \nu_{\rm air} = 0.15\,{\rm cm}^2/{\rm sec} \quad , \quad \nu_{\rm glycerine} = 6.8\,{\rm cm}^2/{\rm sec} \quad . \tag{8.208}$$

## 8.9 Nonequilibrium Quantum Transport

### 8.9.1 Boltzmann equation for quantum systems

Almost everything we have derived thus far can be applied, *mutatis mutandis*, to quantum systems. The main difference is that the distribution $f^0$ corresponding to local equilibrium is no longer of the Maxwell-Boltzmann form, but rather of the Bose-Einstein or Fermi-Dirac form,

$$f^0(\boldsymbol{r},\boldsymbol{k},t) = \left\{ \exp\left( \frac{\varepsilon(\boldsymbol{k}) - \mu(\boldsymbol{r},t)}{k_{\rm B}T(\boldsymbol{r},t)} \right) \mp 1 \right\}^{-1} \quad , \tag{8.209}$$

where the top sign applies to bosons and the bottom sign to fermions. Here we shift to the more common notation for quantum systems in which we write the distribution in terms of the wavevector $\boldsymbol{k} = \boldsymbol{p}/\hbar$ rather than the momentum $\boldsymbol{p}$. The quantum distributions satisfy detailed balance with respect to the quantum collision integral

$$\left(\frac{df}{dt}\right)_{\rm coll} = \int \frac{d^3k_1}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} \int \frac{d^3k'_1}{(2\pi)^3}\, w \left\{ f' f'_1 \left(1 \pm f\right)\left(1 \pm f_1\right) - f f_1 \left(1 \pm f'\right)\left(1 \pm f'_1\right) \right\} \tag{8.210}$$

where $w = w(\boldsymbol{k},\boldsymbol{k}_1 \,|\, \boldsymbol{k}',\boldsymbol{k}'_1)$, $f = f(\boldsymbol{k})$, $f_1 = f(\boldsymbol{k}_1)$, $f' = f(\boldsymbol{k}')$, and $f'_1 = f(\boldsymbol{k}'_1)$, and where we have assumed time-reversal and parity symmetry. Detailed balance requires

$$\frac{f}{1 \pm f} \cdot \frac{f_1}{1 \pm f_1} = \frac{f'}{1 \pm f'} \cdot \frac{f'_1}{1 \pm f'_1} \quad , \tag{8.211}$$

where $f = f^0$ is the equilibrium distribution. One can check that

$$f = \frac{1}{e^{\beta(\varepsilon-\mu)} \mp 1} \qquad \Rightarrow \qquad \frac{f}{1 \pm f} = e^{\beta(\mu-\varepsilon)} \quad , \tag{8.212}$$

which are the Boltzmann factors, which we have already shown to satisfy detailed balance. For the streaming term, we have

$$\begin{aligned} df^0 &= k_{\rm B}T \frac{\partial f^0}{\partial \varepsilon}\, d\left(\frac{\varepsilon-\mu}{k_{\rm B}T}\right) \\ &= k_{\rm B}T \frac{\partial f^0}{\partial \varepsilon} \left\{ -\frac{d\mu}{k_{\rm B}T} - \frac{(\varepsilon-\mu)\,dT}{k_{\rm B}T^2} + \frac{d\varepsilon}{k_{\rm B}T} \right\} \\ &= -\frac{\partial f^0}{\partial \varepsilon} \left\{ \frac{\partial \mu}{\partial \boldsymbol{r}} \cdot d\boldsymbol{r} + \frac{\varepsilon-\mu}{T} \frac{\partial T}{\partial \boldsymbol{r}} \cdot d\boldsymbol{r} - \frac{\partial \varepsilon}{\partial \boldsymbol{k}} \cdot d\boldsymbol{k} \right\} \quad , \end{aligned} \tag{8.213}$$

from which we read off

$$\begin{aligned} \frac{\partial f^0}{\partial \boldsymbol{r}} &= -\frac{\partial f^0}{\partial \varepsilon} \left\{ \frac{\partial \mu}{\partial \boldsymbol{r}} + \frac{\varepsilon-\mu}{T} \frac{\partial T}{\partial \boldsymbol{r}} \right\} \\ \frac{\partial f^0}{\partial \boldsymbol{k}} &= \hbar \boldsymbol{v} \frac{\partial f^0}{\partial \varepsilon} \quad . \end{aligned} \tag{8.214}$$

Here we have assumed $\boldsymbol{V} = 0$ for the unforced flow velocity of the electron fluid. The most important application is to the theory of electron transport in metals and semiconductors, in which case $f^0$ is the Fermi distribution. In this case, the quantum collision integral also receives a contribution from one-body scattering in the presence of an external potential $U(\boldsymbol{r})$, which is given by Fermi's Golden Rule:

$$\begin{aligned}\left(\frac{\partial f(\boldsymbol{k})}{\partial t}\right)'_{\rm coll} &= \frac{2\pi}{\hbar}\sum_{\boldsymbol{k}'\in\hat\Omega} |\langle\, \boldsymbol{k}'\,|\,U\,|\,\boldsymbol{k}\,\rangle|^2\,\big(f(\boldsymbol{k}')-f(\boldsymbol{k})\big)\,\delta\big(\varepsilon(\boldsymbol{k})-\varepsilon(\boldsymbol{k}')\big) \\ &= \frac{2\pi}{\hbar V}\int\limits_{\hat\Omega}\frac{d^3k}{(2\pi)^3}\,|\,\hat U(\boldsymbol{k}-\boldsymbol{k}')|^2\,\big(f(\boldsymbol{k}')-f(\boldsymbol{k})\big)\,\delta\big(\varepsilon(\boldsymbol{k})-\varepsilon(\boldsymbol{k}')\big) \quad .\end{aligned} \tag{8.215}$$

The wavevectors are now restricted to the first Brillouin zone, and the dispersion $\varepsilon(\boldsymbol{k})$ is no longer the ballistic form $\varepsilon = \hbar^2\boldsymbol{k}^2/2m$ but rather the dispersion for electrons in a particular energy band of a solid[10]. Note that $f = f^0$ satisfies detailed balance with respect to one-body collisions as well[11].

In the presence of a weak electric field $\boldsymbol{E}$ and a (not necessarily weak) magnetic field $\boldsymbol{B}$, we have, within the relaxation time approximation, $f = f^0 + \delta f$ with

$$\frac{\partial\,\delta f}{\partial t} - \frac{e}{\hbar c}\,\boldsymbol{v}\times\boldsymbol{B}\cdot\frac{\partial\,\delta f}{\partial\boldsymbol{k}} - \boldsymbol{v}\cdot\left[e\,\boldsymbol{\mathcal{E}} + \frac{\varepsilon-\mu}{T}\,\boldsymbol{\nabla}\,T\right]\frac{\partial f^0}{\partial\varepsilon} = -\frac{\delta f}{\tau} \quad , \tag{8.216}$$

where $\boldsymbol{\mathcal{E}} = -\boldsymbol{\nabla}(\phi - \mu/e) = \boldsymbol{E} - e^{-1}\boldsymbol{\nabla}\mu$ is the gradient of the 'electrochemical potential' $\phi - e^{-1}\mu$. In deriving the above equation, we have worked to lowest order in small quantities. This entails dropping terms like $\boldsymbol{v}\cdot\partial\,\delta f/\partial\boldsymbol{r}$ (higher order in spatial derivatives) and $\boldsymbol{E}\cdot\partial\,\delta f/\partial\boldsymbol{k}$ (both $\boldsymbol{E}$ and $\delta f$ are assumed small). Typically $\tau$ is energy-dependent, *i.e.* $\tau = \tau\big(\varepsilon(\boldsymbol{k})\big)$.

We can use eqn. 8.216 to compute the electrical current $\boldsymbol{j}$ and the thermal current $\boldsymbol{j}_{\rm th}$,

$$\boldsymbol{j} = -2e\int\limits_{\hat\Omega}\frac{d^3k}{(2\pi)^3}\,\boldsymbol{v}\,\delta f \qquad , \qquad \boldsymbol{j}_{\rm th} = 2\int\limits_{\hat\Omega}\frac{d^3k}{(2\pi)^3}\,(\varepsilon-\mu)\,\boldsymbol{v}\,\delta f \quad . \tag{8.217}$$

Here the factor of 2 is from spin degeneracy of the electrons (we neglect Zeeman splitting).

In the presence of a time-independent temperature gradient and electric field, and taking $\boldsymbol{B} = 0$, the linearized Boltzmann equation in the relaxation time approximation has the solution

$$\delta f = -\tau(\varepsilon)\,\boldsymbol{v}\cdot\left(e\boldsymbol{\mathcal{E}} + \frac{\varepsilon-\mu}{T}\,\boldsymbol{\nabla}\,T\right)\left(-\frac{\partial f^0}{\partial\varepsilon}\right) \quad . \tag{8.218}$$

We now consider both the electrical current[12] $\boldsymbol{j}$ as well as the thermal current density $\boldsymbol{j}_{\rm th}$. One readily

[10] We neglect interband scattering here, which can be important in practical applications, but which is beyond the scope of these notes.

[11] The transition rate from $|\boldsymbol{k}'\rangle$ to $|\boldsymbol{k}\rangle$ is proportional to the matrix element and to the product $f'(1-f)$. The reverse process is proportional to $f(1-f')$. Subtracting these factors, one obtains $f'-f$, and therefore the nonlinear terms felicitously cancel in eqn. 8.215.

[12] In this section we use $\boldsymbol{j}$ to denote electrical current, rather than particle number current as before.

obtains

$$\begin{aligned} \boldsymbol{j} &= -2e \int\limits_{\hat{\Omega}} \frac{d^3k}{(2\pi)^3}\ \boldsymbol{v}\,\delta f \equiv L_{11}\,\boldsymbol{\mathcal{E}} - L_{12}\,\boldsymbol{\nabla} T \\ \boldsymbol{j}_{\rm th} &= 2 \int\limits_{\hat{\Omega}} \frac{d^3k}{(2\pi)^3}\ (\varepsilon-\mu)\,\boldsymbol{v}\,\delta f \equiv L_{21}\,\boldsymbol{\mathcal{E}} - L_{22}\,\boldsymbol{\nabla} T \end{aligned} \tag{8.219}$$

where the *transport coefficients* $L^{11}$ *etc.* are matrices:

$$\begin{aligned} L_{11}^{\alpha\beta} &= \frac{e^2}{4\pi^3\hbar}\int d\varepsilon\,\tau(\varepsilon)\left(-\frac{\partial f^0}{\partial\varepsilon}\right)\int dS_\varepsilon\,\frac{v^\alpha\,v^\beta}{|\boldsymbol{v}|} \\ L_{21}^{\alpha\beta} = T L_{12}^{\alpha\beta} &= -\frac{e}{4\pi^3\hbar}\int d\varepsilon\,\tau(\varepsilon)\,(\varepsilon-\mu)\left(-\frac{\partial f^0}{\partial\varepsilon}\right)\int dS_\varepsilon\,\frac{v^\alpha\,v^\beta}{|\boldsymbol{v}|} \\ L_{22}^{\alpha\beta} &= \frac{1}{4\pi^3\hbar T}\int d\varepsilon\,\tau(\varepsilon)\,(\varepsilon-\mu)^2\left(-\frac{\partial f^0}{\partial\varepsilon}\right)\int dS_\varepsilon\,\frac{v^\alpha\,v^\beta}{|\boldsymbol{v}|} \quad , \end{aligned} \tag{8.220}$$

where $dS_\varepsilon$ is a differential surface element in $\boldsymbol{k}$-space at fixed energy $\varepsilon(\boldsymbol{k}) = \varepsilon$, and

$$d^3k = \frac{d\varepsilon\,dS_\varepsilon}{|\partial\varepsilon/\partial\boldsymbol{k}|} = \frac{d\varepsilon\,dS_\varepsilon}{\hbar|\boldsymbol{v}|} \quad . \tag{8.221}$$

If we define the hierarchy of integral expressions

$$\mathcal{J}_n^{\alpha\beta} \equiv \frac{1}{4\pi^3\hbar}\int d\varepsilon\,\tau(\varepsilon)\,(\varepsilon-\mu)^n\left(-\frac{\partial f^0}{\partial\varepsilon}\right)\int dS_\varepsilon\,\frac{v^\alpha\,v^\beta}{|\boldsymbol{v}|} \tag{8.222}$$

then we may write

$$L_{11}^{\alpha\beta} = e^2\,\mathcal{J}_0^{\alpha\beta} \quad , \qquad L_{21}^{\alpha\beta} = T L_{12}^{\alpha\beta} = -e\,\mathcal{J}_1^{\alpha\beta} \quad , \qquad L_{22}^{\alpha\beta} = \frac{1}{T}\,\mathcal{J}_2^{\alpha\beta} \quad . \tag{8.223}$$

The linear relations in eqn. (8.219) may be recast in the following form:

$$\begin{aligned} \boldsymbol{\mathcal{E}} &= \rho\,\boldsymbol{j} + Q\,\boldsymbol{\nabla} T \\ \boldsymbol{j}_{\rm th} &= \sqcap\,\boldsymbol{j} - \kappa\,\boldsymbol{\nabla} T \quad , \end{aligned} \tag{8.224}$$

where the matrices $\rho$, $Q$, $\sqcap$, and $\kappa$ are given by

$$\rho = L_{11}^{-1} \quad , \quad Q = L_{11}^{-1}\,L_{12} \quad , \quad \sqcap = L_{21}\,L_{11}^{-1} \quad , \quad \kappa = L_{22} - L_{21}\,L_{11}^{-1}\,L_{12} \quad , \tag{8.225}$$

or, in terms of the $\mathcal{J}_n$,

$$\rho = \frac{1}{e^2}\,\mathcal{J}_0^{-1} \quad , \quad Q = -\frac{1}{eT}\,\mathcal{J}_0^{-1}\,\mathcal{J}_1 \quad , \quad \sqcap = -\frac{1}{e}\,\mathcal{J}_1\,\mathcal{J}_0^{-1} \quad , \quad \kappa = \frac{1}{T}\left(\mathcal{J}_2 - \mathcal{J}_1\,\mathcal{J}_0^{-1}\,\mathcal{J}_1\right) \quad , \tag{8.226}$$

These equations describe a wealth of transport phenomena:

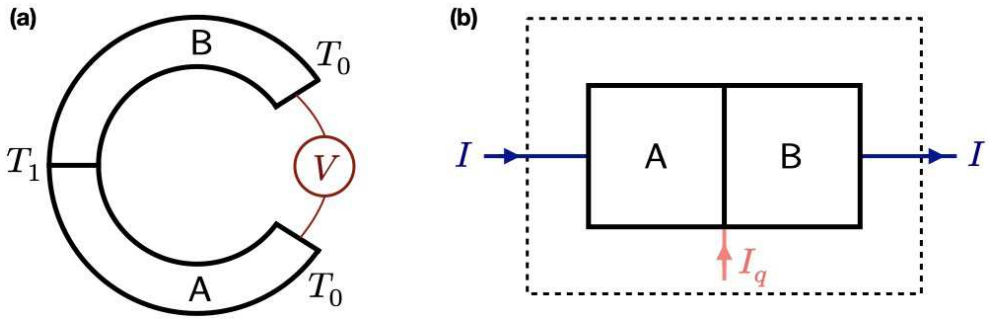



Figure 8.7: (a) A thermocouple is a junction formed of two dissimilar metals. With no electrical current passing, an electric field is generated in the presence of a temperature gradient, resulting in a voltage $V = V_{\rm A} - V_{\rm B} = (Q_{\rm B} - Q_{\rm A})(T_1 - T_0)$. (b) A sketch of a Peltier effect refrigerator. An electrical current $I$ is passed through a junction between two dissimilar metals. If the dotted line represents the boundary of a thermally well-insulated body, then the body cools when $\Pi_{\rm B} > \Pi_{\rm A}$, in order to maintain a heat current balance at the junction, according to $I_{\rm th} = (\Pi_{\rm B} - \Pi_{\rm A})I$.

- Electrical resistance ($\boldsymbol{\nabla} T = \boldsymbol{B} = 0$)
  An electrical current $\boldsymbol{j}$ will generate an electric field $\boldsymbol{\mathcal{E}} = \rho \boldsymbol{j}$, where $\rho$ is the electrical resistivity.
- Peltier effect ($\boldsymbol{\nabla} T = \boldsymbol{B} = 0$)
  An electrical current $\boldsymbol{j}$ will generate a heat current $\boldsymbol{j}_{\rm th} = \Pi \boldsymbol{j}$, where $\Pi$ is the Peltier coefficient.
- Thermal conduction ($\boldsymbol{j} = \boldsymbol{B} = 0$)
  A temperature gradient $\boldsymbol{\nabla} T$ gives rise to a heat current $\boldsymbol{j}_{\rm th} = -\kappa \boldsymbol{\nabla} T$, where $\kappa$ is the thermal conductivity.
- Seebeck effect ($\boldsymbol{j} = \boldsymbol{B} = 0$)
  A temperature gradient $\boldsymbol{\nabla} T$ gives rise to an electric field $\boldsymbol{\mathcal{E}} = Q \boldsymbol{\nabla} T$, where $Q$ is the Seebeck coefficient.

One practical way to measure the thermopower is to form a junction between two dissimilar metals, A and B. The junction is held at temperature $T_1$ and the other ends of the metals are held at temperature $T_0$. One then measures a voltage difference between the free ends of the metals – this is known as the Seebeck effect. Integrating the electric field from the free end of A to the free end of B gives

$$V_{\rm A} - V_{\rm B} = -\int_{\rm A}^{\rm B} \boldsymbol{\mathcal{E}} \cdot d\boldsymbol{l} = (Q_{\rm B} - Q_{\rm A})(T_1 - T_0) \quad . \tag{8.227}$$

What one measures here is really the difference in thermopowers of the two metals. For an absolute measurement of $Q_{\rm A}$, replace B by a superconductor ($Q = 0$ for a superconductor). A device which converts a temperature gradient into an emf is known as a *thermocouple*.

The Peltier effect has practical applications in refrigeration technology. Suppose an electrical current $I$ is passed through a junction between two dissimilar metals, A and B. Due to the difference in Peltier coefficients, there will be a net heat current into the junction of $W = (\Pi_{\rm A} - \Pi_{\rm B})\, I$. Note that this is

proportional to $I$, rather than the familiar $I^2$ result from Joule heating. The sign of $W$ depends on the direction of the current. If a second junction is added, to make an ABA configuration, then heat absorbed at the first junction will be liberated at the second. [13]

### 8.9.2 The Heat Equation

We begin with the continuity equations for charge density $\rho$ and energy density $\varepsilon$:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{j} = 0 \quad , \quad \frac{\partial \varepsilon}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{j}_\varepsilon = \boldsymbol{j} \cdot \boldsymbol{E} \quad , \tag{8.228}$$

where $\boldsymbol{E}$ is the electric field[14]. Now we invoke local thermodynamic equilibrium and write

$$\frac{\partial \varepsilon}{\partial t} = \frac{\partial \varepsilon}{\partial n}\frac{\partial n}{\partial t} + \frac{\partial \varepsilon}{\partial T}\frac{\partial T}{\partial t} = -\frac{\mu}{e}\frac{\partial \rho}{\partial t} + c_V \frac{\partial T}{\partial t} \quad , \tag{8.229}$$

where $n$ is the electron *number density* ($n = -\rho/e$) and $c_V$ is the specific heat. We may now write

$$\begin{aligned} c_V \frac{\partial T}{\partial t} &= \frac{\partial \varepsilon}{\partial t} + \frac{\mu}{e}\frac{\partial \rho}{\partial t} \\ &= \boldsymbol{j} \cdot \boldsymbol{E} - \boldsymbol{\nabla} \cdot \boldsymbol{j}_\varepsilon - \frac{\mu}{e}\boldsymbol{\nabla} \cdot \boldsymbol{j} = \boldsymbol{j} \cdot \boldsymbol{\mathcal{E}} - \boldsymbol{\nabla} \cdot \boldsymbol{j}_{\rm th} \quad . \end{aligned} \tag{8.230}$$

Invoking $\boldsymbol{j}_{\rm th} = \sqcap \boldsymbol{j} - \kappa \boldsymbol{\nabla} T$, we see that if there is no electrical current ($\boldsymbol{j} = 0$), we obtain the *heat equation*

$$c_V \frac{\partial T}{\partial t} = \kappa_{\alpha\beta} \frac{\partial^2 T}{\partial x^\alpha \, \partial x^\beta} \quad . \tag{8.231}$$

This results in a time scale $\tau_T$ for temperature diffusion $\tau_T = \mathcal{C} L^2 c_V / \kappa$, where $L$ is a typical length scale and $\mathcal{C}$ is a numerical constant. For a cube of size $L$ subjected to a sudden external temperature change, $L$ is the side length and $\mathcal{C} = 1/3\pi^2$ (solve by separation of variables).

### 8.9.3 Calculation of Transport Coefficients

We will henceforth assume that sufficient crystalline symmetry exists (*e.g.* cubic symmetry) to render all the transport coefficients multiples of the identity matrix. Under such conditions, we may write $\mathcal{J}_n^{\alpha\beta} = \mathcal{J}_n \, \delta_{\alpha\beta}$ with

$$\mathcal{J}_n = \frac{1}{12\pi^3 \hbar} \int d\varepsilon \, \tau(\varepsilon) \, (\varepsilon - \mu)^n \left( -\frac{\partial f^0}{\partial \varepsilon} \right) \int dS_\varepsilon \, |\boldsymbol{v}| \quad . \tag{8.232}$$

The low-temperature behavior is extracted using the Sommerfeld expansion,

$$\mathcal{I} \equiv \int\limits_{-\infty}^{\infty} d\varepsilon \, H(\varepsilon) \left( -\frac{\partial f^0}{\partial \varepsilon} \right) = \pi \mathcal{D} \csc(\pi \mathcal{D}) \, H(\varepsilon) \Big|_{\varepsilon = \mu} = H(\mu) + \frac{\pi^2}{6} (k_{\rm B} T)^2 \, H''(\mu) + \ldots \quad , \tag{8.233}$$

[13] To create a refrigerator, stick the cold junction inside a thermally insulated box and the hot junction outside the box.

[14] Note that it is $\boldsymbol{E} \cdot \boldsymbol{j}$ and *not* $\boldsymbol{\mathcal{E}} \cdot \boldsymbol{j}$ which is the source term in the energy continuity equation.

where $\mathcal{D} \equiv k_{\rm B}T\,\frac{\partial}{\partial\varepsilon}$ is a dimensionless differential operator[15].

Let us now perform some explicit calculations in the case of a parabolic band with an energy-independent scattering time $\tau$. In this case, one readily finds

$$\mathcal{J}_n = \frac{\sigma_0}{e^2}\,\varepsilon_{\rm F}^{-3/2}\,\pi\mathcal{D}\csc\pi\mathcal{D}\,\varepsilon^{3/2}\,(\varepsilon-\mu)^n\Big|_{\varepsilon=\mu} \quad , \tag{8.234}$$

where $\sigma_0 = ne^2\tau/m^*$. Note that

$$n = \frac{1}{3\pi^2}\left(\frac{2m^*\varepsilon_{\rm F}}{\hbar^2}\right)^{3/2} \tag{8.235}$$

and that $\varepsilon_{\rm F}$ and $\mu$ are related by

$$\varepsilon_{\rm F}^{3/2} = \pi\mathcal{D}\csc\pi\mathcal{D}\,\varepsilon^{3/2}\,\Big|_{\varepsilon=\mu} \quad . \tag{8.236}$$

Thus,

$$\mathcal{J}_0 = \frac{\sigma_0}{e^2} \quad , \quad \mathcal{J}_1 = \frac{\sigma_0}{e^2}\,\frac{\pi^2}{2}\,\frac{(k_{\rm B}T)^2}{\mu} + \ldots \quad , \quad \mathcal{J}_2 = \frac{\sigma_0}{e^2}\,\frac{\pi^2}{3}(k_{\rm B}T)^2 + \ldots \quad , \tag{8.237}$$

from which we obtain the low-$T$ results $\rho = \sigma_0^{-1}$,

$$Q = -\frac{\pi^2}{2}\,\frac{k_{\rm B}^2T}{e\,\varepsilon_{\rm F}} \quad , \quad \kappa = \frac{\pi^2}{3}\,\frac{n\tau}{m^*}\,k_{\rm B}^2T \quad , \tag{8.238}$$

and of course $\Pi = TQ$. The predicted universal ratio

$$\frac{\kappa}{\sigma T} = \frac{\pi^2}{3}\,(k_{\rm B}/e)^2 = 2.45\times10^{-8}\,{\rm V}^2\,{\rm K}^{-2} \quad , \tag{8.239}$$

is known as the *Wiedemann-Franz law*. Note also that our result for the thermopower is unambiguously negative. In actuality, several nearly free electron metals have positive low-temperature thermopowers (Cs and Li, for example). What went wrong? The constant $\tau$, parabolic-band approximation is too restrictive. Realistic band structure and energy-dependence of scattering time can alter or reverse the thermopower, while phonon drag can supply an additional contribution.

### 8.9.4 Onsager Relations

Transport phenomena are described in general by a set of linear relations,

$$J_i = L_{ik}\,F_k \quad , \tag{8.240}$$

[15]Remember that physically the fixed quantities are temperature and total carrier number density (or charge density, in the case of electron and hole bands), and *not* temperature and chemical potential. An equation of state relating $n$, $\mu$, and $T$ is then inverted to obtain $\mu(n,T)$, so that all results ultimately may be expressed in terms of $n$ and $T$.

where the $\{F_k\}$ are *generalized forces* and the $\{J_i\}$ are *generalized currents*. Moreover, to each force $F_i$ corresponds a unique conjugate current $J_i$, such that the rate of internal entropy production is

$$\dot{S} = \sum_i F_i \, J_i \quad \Longrightarrow \quad F_i = \frac{\delta \dot{S}}{\delta J_i} \quad . \tag{8.241}$$

The Onsager relations (also known as *Onsager reciprocity*) state that

$$L_{ik}(\boldsymbol{B}) = \eta_i \, \eta_k \, L_{ki}(-\boldsymbol{B}) \quad , \tag{8.242}$$

where $\eta_i$ describes the parity of $J_i$ under time reversal:

$$J_i^T = \eta_i \, J_i \quad , \tag{8.243}$$

where $J_i^T$ is the time reverse of $J_i$. To justify the Onsager relations requires a microscopic description of our nonequilibrium system.

The Onsager relations have some remarkable consequences. For example, they require, for $\boldsymbol{B} = 0$, that the thermal conductivity tensor $\kappa_{ij}$ of any crystal must be symmetric, independent of the crystal structure. In general, this result does not follow from considerations of crystalline symmetry. It also requires that for every 'off-diagonal' transport phenomenon, *e.g.* the Seebeck effect, there exists a distinct corresponding phenomenon, *e.g.* the Peltier effect.

For the transport coefficients studied, Onsager reciprocity means that in the presence of an external magnetic field,

$$\rho_{\alpha\beta}(\boldsymbol{B}) = \rho_{\beta\alpha}(-\boldsymbol{B}) \qquad , \qquad \kappa_{\alpha\beta}(\boldsymbol{B}) = \kappa_{\beta\alpha}(-\boldsymbol{B}) \qquad , \qquad \sqcap_{\alpha\beta}(\boldsymbol{B}) = T \, Q_{\beta\alpha}(-\boldsymbol{B}) \quad . \tag{8.244}$$

Let's consider an isotropic system in a weak magnetic field, and expand the transport coefficients to first order in $\boldsymbol{B}$:

$$\begin{aligned} \rho_{\alpha\beta}(\boldsymbol{B}) &= \rho \, \delta_{\alpha\beta} + \nu \, \epsilon_{\alpha\beta\gamma} \, B^\gamma \\ \kappa_{\alpha\beta}(\boldsymbol{B}) &= \kappa \, \delta_{\alpha\beta} + \varpi \, \epsilon_{\alpha\beta\gamma} \, B^\gamma \\ Q_{\alpha\beta}(\boldsymbol{B}) &= Q \, \delta_{\alpha\beta} + \zeta \, \epsilon_{\alpha\beta\gamma} \, B^\gamma \\ \sqcap_{\alpha\beta}(\boldsymbol{B}) &= \sqcap \, \delta_{\alpha\beta} + \theta \, \epsilon_{\alpha\beta\gamma} B^\gamma \quad . \end{aligned} \tag{8.245}$$

Onsager reciprocity requires $\sqcap = T \, Q$ and $\theta = T \, \zeta$. We can now write

$$\begin{aligned} \boldsymbol{\mathcal{E}} &= \rho \, \boldsymbol{j} + \nu \, \boldsymbol{j} \times \boldsymbol{B} + Q \, \boldsymbol{\nabla} T + \zeta \, \boldsymbol{\nabla} T \times \boldsymbol{B} \\ \boldsymbol{j}_{\rm th} &= \sqcap \, \boldsymbol{j} + \theta \, \boldsymbol{j} \times \boldsymbol{B} - \kappa \, \boldsymbol{\nabla} T - \varpi \, \boldsymbol{\nabla} T \times \boldsymbol{B} \quad . \end{aligned} \tag{8.246}$$

There are several new phenomena lurking:

- Hall effect ($\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = j_y = 0$)
  An electrical current $\boldsymbol{j} = j_x \, \boldsymbol{x}$ and a field $\boldsymbol{B} = B_z \, \hat{\boldsymbol{z}}$ yield an electric field $\boldsymbol{\mathcal{E}}$. The Hall coefficient is $R_{\rm H} = \mathcal{E}_y / j_x \, B_z = -\nu$.

- Ettingshausen effect ($\frac{\partial T}{\partial x} = j_y = j_{\rm th,y} = 0$)
  An electrical current $\boldsymbol{j} = j_x\,\boldsymbol{x}$ and a field $\boldsymbol{B} = B_z\,\hat{\boldsymbol{z}}$ yield a temperature gradient $\frac{\partial T}{\partial y}$. The Ettingshausen coefficient is $P = \frac{\partial T}{\partial y}\big/ j_x\,B_z = -\theta/\kappa$.

- Nernst effect ($j_x = j_y = \frac{\partial T}{\partial y} = 0$)
  A temperature gradient $\boldsymbol{\nabla} T = \frac{\partial T}{\partial x}\,\boldsymbol{x}$ and a field $\boldsymbol{B} = B_z\,\hat{\boldsymbol{z}}$ yield an electric field $\boldsymbol{\mathcal{E}}$. The Nernst coefficient is $\Lambda = \mathcal{E}_y\big/\frac{\partial T}{\partial x}\,B_z = -\zeta$.

- Righi-Leduc effect ($j_x = j_y = j_{\rm th,y} = 0$)
  A temperature gradient $\boldsymbol{\nabla} T = \frac{\partial T}{\partial x}\,\boldsymbol{x}$ and a field $\boldsymbol{B} = B_z\,\hat{\boldsymbol{z}}$ yield an orthogonal temperature gradient $\frac{\partial T}{\partial y}$. The Righi-Leduc coefficient is $\mathcal{L} = \frac{\partial T}{\partial y}\big/\frac{\partial T}{\partial x}B_z = \varpi/\kappa$.

## 8.10 Appendix I: The BBGKY Hierarchy and the Boltzmann Equation

The procedure by which the Boltzmann equation for a gas is obtained from $N$-particle Hamiltonian dynamics of its constituent particles is known as the *Boltzmann-Grad limit*. An excellent source, which we follow here, is Kardar (Particles), §3.3. The formal derivation of the Boltzmann equation from reversible hard sphere dynamics in the low-density limit is generally credited to R. Lanford, "Time Evolution of Large Classical Systems" in *Dynamical Systems, Theory and Applications* (Springer Lecture Notes in Physics, 1975).

### 8.10.1 BBGKY hierarchy

We start with the time-dependent distribution function, $\varrho_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N, t)$, on $N$-particle phase space, which is assumed normalized according to

$$\int \prod_{j=1}^{N} d\mu_j\; \varrho_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_N, t) = 1 \quad , \tag{8.247}$$

where $d\mu_j \equiv d^d x_j\, d^d p_j$ (we assume $d = 3$). We define the $s$-particle distribution function,

$$f_s\big(\{\boldsymbol{x}_i\}, \{\boldsymbol{p}_i\}, t\big) = \frac{N!}{(N-s)!} \int \prod_{i=s+1}^{N} d\mu_i\; \varrho_N\big(\{\boldsymbol{x}_j\}, \{\boldsymbol{p}_j\}, t\big) \equiv \frac{N!}{(N-s)!}\, \varrho_s\big(\{\boldsymbol{x}_i\}, \{\boldsymbol{p}_i\}, t\big) \quad . \tag{8.248}$$

We adopt the notation that $\{\boldsymbol{x}_j\} \equiv \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s\}$ when it appears as an argument of a function $f_s$ or $\varrho_s$, and similarly for $\{\boldsymbol{p}_j\}$. The dynamics of $\varrho_N$ are given by Liouville's equation, $\partial_t\, \varrho_N + \{\varrho_N, H_N\} = 0$, where

$$H_N\big(\{\boldsymbol{x}_j\}, \{\boldsymbol{p}_j\}\big) = \sum_{j=1}^{N} \left[ \frac{\boldsymbol{p}_j^2}{2m} + v(\boldsymbol{x}_j) \right] + \sum_{i<j}^{N} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \quad , \tag{8.249}$$

where $v(\boldsymbol{x})$ is an external potential. We may write, for any $k \in \{1, \ldots, N\}$,

$$H_N = H_k + \bar{H}_k + H'_k \quad , \tag{8.250}$$

where

$$\begin{aligned} H_k &= \sum_{j=1}^{k} \left[ \frac{\boldsymbol{p}_j^2}{2m} + v(\boldsymbol{x}_j) \right] + \sum_{0<i<j}^{k} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \\ \bar{H}_k &= \sum_{j=k+1}^{N} \left[ \frac{\boldsymbol{p}_j^2}{2m} + v(\boldsymbol{x}_j) \right] + \sum_{k<i<j}^{N} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \\ H'_k &= \sum_{i=1}^{k} \sum_{j=k+1}^{N} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \quad . \end{aligned} \tag{8.251}$$

We further assume that $u(\boldsymbol{x}_i - \boldsymbol{x}_j) = u(|\boldsymbol{x}_i - \boldsymbol{x}_j|)$ is a central potential. The dynamics of $\varrho_s$ is then given by

$$\frac{\partial \varrho_s}{\partial t} = \int \prod_{j=s+1}^{N} d\mu_j \, \frac{\partial \varrho_N}{\partial t} = - \int \prod_{j=s+1}^{N} d\mu_j \, \big\{ \varrho_N, H_s + \bar{H}_s + H'_s \big\} \quad . \tag{8.252}$$

We now evaluate each of the three contributions to the Poisson bracket.

The first contribution is easy:

$$\left[ \frac{\partial \varrho_s}{\partial t} \right]_{(1)} = - \int \prod_{j=s+1}^{N} d\mu_j \, \big\{ \varrho_N, H_s \big\} = - \big\{ \varrho_s, H_s \big\} \quad . \tag{8.253}$$

The second contribution is

$$\begin{aligned} \left[ \frac{\partial \varrho_s}{\partial t} \right]_{(2)} &= - \int \prod_{j=s+1}^{N} d\mu_j \big\{ \varrho_N, \bar{H}_s \big\} \\ &= \int \prod_{j=s+1}^{N} d\mu_j \sum_{k=s+1}^{N} \left[ \frac{\partial \varrho_N}{\partial \boldsymbol{p}_k} \cdot \frac{\partial \bar{H}_s}{\partial \boldsymbol{x}_k} - \frac{\partial \varrho_N}{\partial \boldsymbol{x}_k} \cdot \frac{\partial \bar{H}_s}{\partial \boldsymbol{p}_k} \right] \\ &= \int \prod_{j=s+1}^{N} d\mu_j \sum_{k=s+1}^{N} \left[ \frac{\partial \varrho_N}{\partial \boldsymbol{p}_k} \cdot \overbrace{\left( \frac{\partial v(\boldsymbol{x}_k)}{\partial \boldsymbol{x}_k} + \frac{1}{2} \sum_{\ell=s+1}^{N} \frac{\partial u(\boldsymbol{x}_k - \boldsymbol{x}_\ell)}{\partial \boldsymbol{x}_k} \right)}^{\text{independent of } \boldsymbol{p}_k} - \frac{\partial \varrho_N}{\partial \boldsymbol{x}_k} \cdot \frac{\boldsymbol{p}_k}{m} \right] = 0 \quad . \end{aligned} \tag{8.254}$$

Here we integrate by parts on both terms inside the square brackets. The term in rounded brackets is independent of $\boldsymbol{p}_k$, while $\boldsymbol{p}_k/m$ is independent of $\boldsymbol{x}_k$ . Thus, the second contribution vanishes.

The third contribution is

$$\begin{aligned}\left[\frac{\partial \varrho_s}{\partial t}\right]_{(3)} &= -\int \prod_{j=s+1}^{N} d\mu_j \, \{\varrho_N, H_s'\} \\ &= \int \prod_{j=s+1}^{N} d\mu_j \sum_{k=1}^{s} \sum_{\ell=s+1}^{N} \left( \frac{\partial \rho_N}{\partial \boldsymbol{p}_k} \cdot \frac{\partial u(\boldsymbol{x}_k - \boldsymbol{x}_\ell)}{\partial \boldsymbol{x}_k} - \frac{\partial \rho_N}{\partial \boldsymbol{p}_\ell} \cdot \frac{\partial u(\boldsymbol{x}_k - \boldsymbol{x}_\ell)}{\partial \boldsymbol{x}_\ell} \right) \\ &= (N-s) \int d\mu_{s+1} \sum_{k=1}^{s} \frac{\partial u(\boldsymbol{x}_k - \boldsymbol{x}_{s+1})}{\partial \boldsymbol{x}_k} \cdot \frac{\partial}{\partial \boldsymbol{p}_k} \left( \int \prod_{\ell=s+2}^{N} d\mu_\ell \, \varrho_N \right)\end{aligned} \tag{8.255}$$

Thus, we arrive at

$$\frac{\partial \varrho_s}{\partial t} + \{\rho_s, H_s\} = (N-s) \int d\mu_{s+1} \sum_{k=1}^{s} \frac{\partial u(\boldsymbol{x}_k - \boldsymbol{x}_{s+1})}{\partial \boldsymbol{x}_k} \cdot \frac{\partial \varrho_{s+1}}{\partial \boldsymbol{p}_k} \quad , \tag{8.256}$$

or

$$\frac{\partial f_s}{\partial t} + \{f_s, H_s\} = \int d\mu_{s+1} \sum_{k=1}^{s} \frac{\partial u(\boldsymbol{x}_k - \boldsymbol{x}_{s+1})}{\partial \boldsymbol{x}_k} \cdot \frac{\partial f_{s+1}}{\partial \boldsymbol{p}_k} \quad . \tag{8.257}$$

This is the BBGKY hierarchy.

### 8.10.2 Boltzmann equation

At each level of the hierarchy, we may identify three types of terms, each of which is associated with its own time scale. The size of each type is roughly inversely proportional to its characteristic time scale[16]. The first type of term arises from the noninteracting terms in $H_s$, *i.e.* the one-body terms, which do not involve the interaction potential $u(\boldsymbol{x}_i - \boldsymbol{x}_j)$. The time scale associated with these terms behaves as $\tau_v \sim L/V$, where $V$ is a characteristic velocity and $L$ the linear system size. This extrinsic scale becomes arbitrarily large as $L \to \infty$. The second type of term are those on the LHS of the equation for $\varrho_s$. Assuming $u(\boldsymbol{x})$ is short-ranged, such as a van der Waals or Lennard-Jones potential, the typical distance scale is $d \sim 1$ Å, which entails a collision time $\tau_{\rm c} \sim 10^{-12}$ s. Finally, the third type at each level are terms on the RHS which are proportional to the interaction potential $u$. Because the distribution function $f_{s+1}$ enters, these terms require an additional particle be present, and the time constant is longer than $\tau$ by a factor $1/nd^3 \sim 10^4$. One concludes that the terms proportional to $nd^3/\tau_{\rm c}$ that appear on the RHS at each level may be dropped, because they are dominated by the terms proportional to $1/\tau_{\rm c}$ on the LHS at any given level. This means that each level of the hierarchy provides a closed equation for $f_s$, with the sole exception of $s = 1$, where there is no interaction term on the LHS.

At level $s = 2$, the equation for $f_2$ is

$$\left\{ \frac{\partial}{\partial t} + \frac{\boldsymbol{p}_1}{m} \cdot \frac{\partial}{\partial \boldsymbol{x}_1} + \frac{\boldsymbol{p}_2}{m} \cdot \frac{\partial}{\partial \boldsymbol{x}_2} - \frac{\partial u(\boldsymbol{x}_2 - \boldsymbol{x}_1)}{\partial \boldsymbol{x}_1} \cdot \frac{\partial}{\partial \boldsymbol{p}_1} - \frac{\partial u(\boldsymbol{x}_2 - \boldsymbol{x}_1)}{\partial \boldsymbol{x}_2} \cdot \frac{\partial}{\partial \boldsymbol{p}_2} \right\} f_2(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{p}_1, \boldsymbol{p}_2, t) = 0 \quad . \tag{8.258}$$

[16] See Kardar §3.3 for a more detailed discussion.

On time scales long compared with $\tau_c$, the distribution $f_2$ reaches a steady state value. In the presence of slowly time-varying parameters such as the temperature $T(\boldsymbol{r},t)$ or chemical potential $\mu(\boldsymbol{r},t)$, or in the presence of a time-dependent external force $\boldsymbol{F}(\boldsymbol{r},t)$, the distribution $f_2$ will evolve, but on a longer time scale associated with these respective variations. Therefore we can fix the time $t$ and freeze these slow changes and ask about the steady state distribution $f_2^{\rm ss}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{p}_1,\boldsymbol{p}_2,t)$. In discussing the 'instantaneous steady state' – a seemingly nonsensical collocation which however has meaning given the aforementioned separation of time scales – we suppress for the moment the $t$ label and we define $\phi(\boldsymbol{X},\boldsymbol{x},\boldsymbol{p}_1,\boldsymbol{p}_2) \equiv f_2(\boldsymbol{x}_1,\boldsymbol{x}_1,\boldsymbol{p}_2,\boldsymbol{p}_2)$, where $\boldsymbol{X} \equiv \frac{1}{2}(\boldsymbol{x}_2+\boldsymbol{x}_2)$ is the center-of-mass coordinate and $\boldsymbol{x} \equiv \boldsymbol{x}_2 - \boldsymbol{x}_1$ is the relative coordinate. We then have

$$\left\{\left(\frac{\boldsymbol{p}_1+\boldsymbol{p}_2}{2m}\right)\cdot\frac{\partial}{\partial \boldsymbol{X}} + \left(\frac{\boldsymbol{p}_1-\boldsymbol{p}_2}{m}\right)\cdot\frac{\partial}{\partial \boldsymbol{x}} + \boldsymbol{\nabla}u(\boldsymbol{x})\cdot\left(\frac{\partial}{\partial \boldsymbol{p}_1}-\frac{\partial}{\partial \boldsymbol{p}_2}\right)\right\}\phi(\boldsymbol{X},\boldsymbol{x},\boldsymbol{p}_1,\boldsymbol{p}_2) = 0 \quad . \tag{8.259}$$

We recognize this as a linear partial differential equation in the 12 variables $\{\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{p}_1,\boldsymbol{p}_2\}$, which, in principle, may be solved by the method of characteristics. This entails introducing a scalar variable $\tau$ and solving the coupled ODEs

$$\frac{d\boldsymbol{X}}{d\tau} = \frac{\boldsymbol{p}_1+\boldsymbol{p}_2}{2m} \quad , \quad \frac{d\boldsymbol{x}}{d\tau} = \frac{\boldsymbol{p}_2-\boldsymbol{p}_1}{m} \quad , \quad \frac{d\boldsymbol{p}_1}{d\tau} = \boldsymbol{\nabla}u(\boldsymbol{x}) \quad , \quad \frac{d\boldsymbol{p}_2}{d\tau} = -\boldsymbol{\nabla}u(\boldsymbol{x}) \quad . \tag{8.260}$$

This defines a path in the 12-dimensional space. Along such a path,

$$\frac{d\phi\big(\boldsymbol{X}(\tau),\boldsymbol{x}(\tau),\boldsymbol{p}_1(\tau),\boldsymbol{p}_2(\tau)\big)}{d\tau} = \frac{\partial\phi}{\partial\boldsymbol{X}}\cdot\frac{d\boldsymbol{X}}{d\tau} + \frac{\partial\phi}{\partial\boldsymbol{x}}\cdot\frac{d\boldsymbol{x}}{d\tau} + \frac{\partial\phi}{\partial\boldsymbol{p}_1}\cdot\frac{d\boldsymbol{p}_1}{d\tau} + \frac{\partial\phi}{\partial\boldsymbol{p}_2}\cdot\frac{d\boldsymbol{p}_2}{d\tau} = 0 \quad . \tag{8.261}$$

The $N=12$ variable PDE of eqn. 8.258 thus has been converted into a system of $N+1=13$ coupled ODEs of eqns. 8.260 and 8.261. We see that $\tau$ functions as a time variable, and the coupled ODEs of eqn. 8.260 are simply Hamilton's equations of motion. Note that $\boldsymbol{P} \equiv \boldsymbol{p}_1+\boldsymbol{p}_2$ satisfies $d\boldsymbol{P}/dt = 0$ and we have the familiar CM motion $\boldsymbol{X}(\tau) = \boldsymbol{X}(0) + \boldsymbol{P}\tau/M$, with $M=2m$ the total mass. Initial conditions for the $N$ variables $\boldsymbol{w} = \{w_1,\ldots,w_N\}$ at $\tau=0$ are specified by fixing a hypersurface parameterized by $N-1$ variables $\boldsymbol{\zeta} = \{\zeta_1,\ldots,\zeta_{N-1}\}$, where

$$w_j(\tau=0) = h_j(\zeta_1,\ldots,\zeta_{N-1}) \qquad , \qquad \phi(\tau=0) = F(\zeta_1,\ldots,\zeta_{N-1}) \quad , \tag{8.262}$$

and where the $\{h_j(\boldsymbol{\zeta})\}$ define the hypersurface. A simple way to think of this is to define the hypersurface by the relation $w_N = 0$, and take $w_j(\tau=0) = \zeta_j$ for $j \in \{1,\ldots,N-1\}$, *i.e.* $h_j(\boldsymbol{\zeta}) = \zeta_j$. Each characteristic is labeled by an $(N-1)$-tuple $\boldsymbol{\zeta}$, and the solution along a characteristic is $\phi(\boldsymbol{w}) = \phi(\tau;\boldsymbol{\zeta})$[17].

Now we turn to level $s=1$ of the hierarchy, where we have

$$\frac{\partial f}{\partial t} + \frac{\boldsymbol{p}_1}{m}\cdot\frac{\partial f}{\partial \boldsymbol{x}_1} - \boldsymbol{\nabla}v(\boldsymbol{x}_1)\cdot\frac{\partial f}{\partial \boldsymbol{p}_1} = \int d\mu_2\,\frac{\partial u(\boldsymbol{x}_2-\boldsymbol{x}_1)}{\partial \boldsymbol{x}_1}\cdot\left(\frac{\partial f_2}{\partial \boldsymbol{p}_1}-\frac{\partial f_2}{\partial \boldsymbol{p}_2}\right) \tag{8.263}$$

We now invoke the relation

$$-\boldsymbol{\nabla}u(\boldsymbol{x})\cdot\left(\frac{\partial f_2}{\partial \boldsymbol{p}_1}-\frac{\partial f_2}{\partial \boldsymbol{p}_2}\right) = \left(\frac{\boldsymbol{p}_1+\boldsymbol{p}_2}{2m}\right)\cdot\frac{\partial f_2}{\partial \boldsymbol{X}} + \left(\frac{\boldsymbol{p}_1-\boldsymbol{p}_2}{m}\right)\cdot\frac{\partial f_2}{\partial \boldsymbol{x}} \quad , \tag{8.264}$$

[17] If two characteristics cross, a shock has occurred.

leading us to

$$\frac{\partial f}{\partial t} + \frac{\boldsymbol{p}_1}{m}\cdot\frac{\partial f}{\partial \boldsymbol{x}_1} - \boldsymbol{\nabla} v(\boldsymbol{x}_1)\cdot\frac{\partial f}{\partial \boldsymbol{p}_1} = \int d^3p_2 \int d^3x \left(\frac{\boldsymbol{p}_1-\boldsymbol{p}_2}{m}\right)\cdot\frac{\partial f_2(\boldsymbol{x}_1,\boldsymbol{x}_1+\boldsymbol{x},\boldsymbol{p}_1,\boldsymbol{p}_2,t)}{\partial \boldsymbol{x}} \quad , \tag{8.265}$$

where we have restored the time label $t$. In arriving at this last equation, note that we have dropped terms proportional to $\partial f_2/\partial \boldsymbol{X}$. This is justified by the presumption that the center-of-mass dynamics are relatively slow. Note also that on the RHS we are expressing the spatial coordinates in terms of $\boldsymbol{x}_1$ and $\boldsymbol{x}$, hence we have substituted $\boldsymbol{x}_2 = \boldsymbol{x}_1 + \boldsymbol{x}$ in the second vector argument of $f_2$ .

Consider the integral on the RHS of eqn. 8.265 over the relative coordinate $\boldsymbol{x}$. Following Kardar, at fixed $\boldsymbol{x}_1$, $\boldsymbol{p}_1$, and $\boldsymbol{p}_2$ , we choose a coordinate system for $\boldsymbol{x}$ where one axis is parallel to $\hat{\boldsymbol{n}} = (\boldsymbol{p}_1-\boldsymbol{p}_2)/|\boldsymbol{p}_1-\boldsymbol{p}_2|$ and write

$$\boldsymbol{x} = x_\parallel\,\hat{\boldsymbol{n}} + \boldsymbol{b} \quad , \tag{8.266}$$

where $\hat{\boldsymbol{n}}\cdot\boldsymbol{b} = 0$. The vector $\boldsymbol{b}$ is known as the *impact parameter vector*. We then have

$$\left(\frac{df}{dt}\right)_{\rm coll} = \int d^3p_2 \int d^2b \left|\frac{\boldsymbol{p}_1-\boldsymbol{p}_2}{m}\right| \left\{\tilde{f}_2^+(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1,\boldsymbol{p}_2,t) - \tilde{f}_2^-(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1,\boldsymbol{p}_2,t)\right\} \quad , \tag{8.267}$$

where

$$\tilde{f}_2^\pm(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1,\boldsymbol{p}_2,t) \equiv f_2(\boldsymbol{x}_1,\boldsymbol{x}_1+x_\perp^+\,\hat{\boldsymbol{n}}+\boldsymbol{b},\boldsymbol{p}_1,\boldsymbol{p}_2,t) \quad , \tag{8.268}$$

and where $x_\parallel^\pm$ is the value of $\hat{\boldsymbol{n}}\cdot\boldsymbol{x}$ after $(+)$ or before $(-)$ the collision. We assume the potential $u(\boldsymbol{x})$ is sufficiently short-ranged that $x_\parallel^\pm$ may be evaluated a short distance from the collision center $\boldsymbol{x} = 0$. In a scattering process, the momenta $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ prior to the scattering event evolve into $\boldsymbol{p}_1'$ and $\boldsymbol{p}_2'$ after the event. The relation between the incoming and outgoing momenta is given by eqn. 8.59, *i.e.*

$$\begin{aligned}\boldsymbol{p}_1' &= \tfrac{1}{2}\big(\boldsymbol{p}_1+\boldsymbol{p}_2+|\boldsymbol{p}_1-\boldsymbol{p}_2|\,\hat{\boldsymbol{\Omega}}\big)\\ \boldsymbol{p}_2' &= \tfrac{1}{2}\big(\boldsymbol{p}_1+\boldsymbol{p}_2-|\boldsymbol{p}_1-\boldsymbol{p}_2|\,\hat{\boldsymbol{\Omega}}\big) \quad ,\end{aligned} \tag{8.269}$$

where $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Omega}}(\boldsymbol{b})$ is a unit vector which is determined by the impact parameter vector $\boldsymbol{b}$. Assuming the system is time-reversal invariant, we may relate $\tilde{f}_2^+$ and $\tilde{f}_2^-$ by

$$\tilde{f}_2^+(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1,\boldsymbol{p}_2,t) = \tilde{f}_2^-(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1',\boldsymbol{p}_2',t) \quad . \tag{8.270}$$

This suggests that we could write the collision integral in either of two ways:

$$\begin{aligned}\left(\frac{df}{dt}\right)_{\rm coll,-} &= \int d^3p_2 \int d\Omega\,\frac{d\sigma}{d\Omega}\,|\boldsymbol{v}_1-\boldsymbol{v}_2|\left\{\tilde{f}_2^-(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1',\boldsymbol{p}_2',t) - \tilde{f}_2^-(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1,\boldsymbol{p}_2,t)\right\}\\ \left(\frac{df}{dt}\right)_{\rm coll,+} &= \int d^3p_2 \int d\Omega\,\frac{d\sigma}{d\Omega}\,|\boldsymbol{v}_1-\boldsymbol{v}_2|\left\{\tilde{f}_2^+(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1,\boldsymbol{p}_2,t) - \tilde{f}_2^+(\boldsymbol{x}_1,\boldsymbol{b},\boldsymbol{p}_1',\boldsymbol{p}_2',t)\right\} \quad .\end{aligned} \tag{8.271}$$

Here we have also written

$$d^2b = \frac{d\sigma}{d\Omega}\,d\Omega \quad , \tag{8.272}$$

where $d\Omega$ is the differential solid angle of the scattering process. The expression $d\sigma/d\Omega$, which should be familiar from both classical mechanics and quantum mechanics, is the *differential scattering cross-section*. We assume there is a unique relationship between the scattering direction $\hat{\boldsymbol{\Omega}}(\boldsymbol{b})$ and the impact

parameter vector $\boldsymbol{b}$, which is to say that the $\boldsymbol{b}$ plane can be mapped to the unit sphere $\hat{\boldsymbol{\Omega}}(\boldsymbol{b})$ in a one-to-one fashion. In general, $d\sigma/d\Omega \to 0$ for $|\boldsymbol{b}| \gg r_{\rm c}$, where $r_{\rm c}$ is the length scale over which the scattering occurs, which is presumed to be microscopic. At long last, we make the assumption,

$$\tilde{f}_2^{\pm}(\boldsymbol{x}_1, \boldsymbol{b}, \boldsymbol{p}_1, \boldsymbol{p}_2, t) \approx f_1(\boldsymbol{x}_1, \boldsymbol{p}_1, t) \times f_1(\boldsymbol{x}_2, \boldsymbol{p}_2, t) \quad . \tag{8.273}$$

This is known as the *Stosszahlansatz* or the *hypothesis of molecular chaos*, and amounts to assuming that the velocities of colliding particles are uncorrelated. But are they uncorrelated before $(-)$ or after $(+)$ the collision? The answer makes a huge difference! It determines which of the candidate collision integrals in eqn. 8.271 is correct.

Our Boltzmann equation for the one-body distribution $f(\boldsymbol{r}, \boldsymbol{p}, t)$ is now given by

$$\begin{aligned}\left(\frac{df}{dt}\right)_{\rm coll} &= \pm \int d^3p_1 \int d\Omega \, \frac{d\sigma}{d\Omega} \left| \frac{\boldsymbol{p} - \boldsymbol{p}_1}{m} \right| \Big\{ f(\boldsymbol{r}, \boldsymbol{p}, t) f(\boldsymbol{r}, \boldsymbol{p}_1, t) - f(\boldsymbol{r}, \boldsymbol{p}', t) f(\boldsymbol{r}, \boldsymbol{p}_1', t) \Big\} \\ &= \pm \int d^3p_1 \int d^3p' \int d^3p_1' \, w(\boldsymbol{p}, \boldsymbol{p}_1 \,|\, \boldsymbol{p}', \boldsymbol{p}_1') \Big\{ f(\boldsymbol{r}, \boldsymbol{p}, t) f(\boldsymbol{r}, \boldsymbol{p}_1, t) - f(\boldsymbol{r}, \boldsymbol{p}', t) f(\boldsymbol{r}, \boldsymbol{p}_1', t) \Big\}\end{aligned} \tag{8.274}$$

where we have used eqn. 8.36 in the last line. According to the results of §8.3.7, we see that choosing $\tilde{f}_2^+$ in eqn. 8.273 results in a generically *positive* source for Boltzmann's $\mathcal{H}$ density, $h$, which is to say $\partial_t h + \boldsymbol{\nabla} \cdot \boldsymbol{j} \geq 0$, so for this choice we have an "anti $\mathcal{H}$-theorem". Thermodynamically, this means that entropy of an isolated system *decreases* (since $S = -k_{\rm B}\mathcal{H}$), which is the exact opposite of what the second law of thermodynamics requires. Thus, the correct choice is to take $\tilde{f}_2^-$ in eqn. 8.273, in which case our Boltzmann equation takes the final form[18]

$$\frac{\partial f}{\partial t} + \frac{\boldsymbol{p}}{m} \cdot \frac{\partial f}{\partial \boldsymbol{r}} - \frac{\partial v}{\partial \boldsymbol{r}} \cdot \frac{\partial f}{\partial \boldsymbol{p}} = \int d^3p_1 \int d^3p' \int d^3p_1' \, w(\boldsymbol{p}, \boldsymbol{p}_1 \,|\, \boldsymbol{p}', \boldsymbol{p}_1') \Big\{ f(\boldsymbol{r}, \boldsymbol{p}', t) f(\boldsymbol{r}, \boldsymbol{p}_1', t) - f(\boldsymbol{r}, \boldsymbol{p}, t) f(\boldsymbol{r}, \boldsymbol{p}_1, t) \Big\} \quad . \tag{8.275}$$

Why is it necessary to invoke the *Stosszahlansatz* for particles which are just about to collide, but *not* for particles which have just collided? This is a deep question, and the essence of Loschmidt's objection to the Boltzmann equation and the $\mathcal{H}$-theorem it entails (also called *Loschmidt's paradox*) – how can irreversibility result from microscopic dynamics which are invariant under time-reversal? As Lanford (1975) stresses, "the BBGKY hierarchy is time-reversal invariant but the Boltzmann equation is not." The answer appears to be quite subtle and still somewhat contentious. Bodineau *et al.*[19] claim that the BBGKY dynamics for dilute hard spheres in $d$ dimensions[20] converges to the solution of the Boltzmann equation, *i.e.* one must specify pre-collisional data.

[18] Note also that if we were to impose the *Stosszahlansatz* on each term inside the curly brackets on the RHS of eqn. 8.267, the collision integral would vanish!

[19] See T. Bodineau, I. Gallagher, L. Saint-Raymond, and S. Simonella, "One-sided convergence in the Boltzmann-Grad limit," *Annales de la Faculté des Sciences de Toulouse*, vol. XXVIII, no. 5, p. 985 (2018).

[20] The Boltzmann-Grad limit is defined by $N$ hard spheres each of diameter $\delta$ in a box of volume $V$, with $n\delta^{d-1}\ell \sim 1$, where $n = N/V \to \infty$ is the number density of spheres and $\ell$ is the finite mean free path. Thus $\delta \sim n^{-1/(d-1)} \to 0$.

## 8.11 Appendix II: Boltzmann Equation and Collisional Invariants

*Problem* : The linearized Boltzmann operator $L\psi$ is a complicated functional. Suppose we replace $L$ by $\mathcal{L}$, where

$$\mathcal{L}\psi = -\gamma\,\psi(\boldsymbol{v},t) + \gamma\Big(\frac{m}{2\pi k_{\rm B}T}\Big)^{3/2}\int d^3u\,\exp\Big(-\frac{m\boldsymbol{u}^2}{2k_{\rm B}T}\Big) \times\left\{1+\frac{m}{k_{\rm B}T}\,\boldsymbol{u}\cdot\boldsymbol{v}+\frac{2}{3}\left(\frac{m\boldsymbol{u}^2}{2k_{\rm B}T}-\frac{3}{2}\right)\left(\frac{m\boldsymbol{v}^2}{2k_{\rm B}T}-\frac{3}{2}\right)\right\}\psi(\boldsymbol{u},t) \quad . \tag{8.276}$$

Show that $\mathcal{L}$ shares all the important properties of $L$. What is the meaning of $\gamma$? Expand $\psi(\boldsymbol{v},t)$ in spherical harmonics and Sonine polynomials,

$$\psi(\boldsymbol{v},t) = \sum_{r\ell m} a_{r\ell m}(t)\,S^r_{\ell+\frac{1}{2}}(x)\,x^{\ell/2}\,Y_{\ell,m}(\hat{\boldsymbol{n}}), \tag{8.277}$$

with $x = mv^2/2k_{\rm B}T$, and thus express the action of the linearized Boltzmann operator algebraically on the expansion coefficients $a_{r\ell m}(t)$.

The Sonine polynomials $S^n_\alpha(x)$ are a complete, orthogonal set which are convenient to use in the calculation of transport coefficients. They are defined as

$$S^n_\alpha(x) = \sum_{m=0}^{n}\frac{\Gamma(\alpha+n+1)\,(-x)^m}{\Gamma(\alpha+m+1)\,(n-m)!\,m!} \quad , \tag{8.278}$$

and satisfy the generalized orthogonality relation

$$\int_0^\infty dx\,e^{-x}\,x^\alpha\,S^n_\alpha(x)\,S^{n'}_\alpha(x) = \frac{\Gamma(\alpha+n+1)}{n!}\,\delta_{n,n'} \quad . \tag{8.279}$$

*Solution* : The 'important properties' of $L$ are that it annihilates the five collisional invariants, *i.e.* 1, $\boldsymbol{v}$, and $v^2$, and that all other eigenvalues are *negative*. That this is true for $\mathcal{L}$ can be verified by an explicit calculation.

Plugging the conveniently parameterized form of $\psi(\boldsymbol{v},t)$ into $\mathcal{L}$, we have

$$\mathcal{L}\psi = -\gamma\sum_{r\ell m}a_{r\ell m}(t)\,S^r_{\ell+\frac{1}{2}}(x)\,x^{\ell/2}\,Y_{\ell,m}(\hat{\boldsymbol{n}}) \;+\; \frac{\gamma}{2\pi^{3/2}}\sum_{r\ell m}a_{r\ell m}(t)\int_0^\infty dx_1\,x_1^{1/2}\,e^{-x_1} \times\int d\hat{\boldsymbol{n}}_1\Big[1+2\,x^{1/2}x_1^{1/2}\,\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}}_1+\tfrac{2}{3}\big(x-\tfrac{3}{2}\big)\big(x_1-\tfrac{3}{2}\big)\Big]\,S^r_{\ell+\frac{1}{2}}(x_1)\,x_1^{\ell/2}\,Y_{\ell,m}(\hat{\boldsymbol{n}}_1) \quad , \tag{8.280}$$

where we've used

$$u = \sqrt{\frac{2k_{\rm B}T}{m}}\,x_1^{1/2} \qquad , \qquad du = \sqrt{\frac{k_{\rm B}T}{2m}}\,x_1^{-1/2}\,dx_1 \quad . \tag{8.281}$$

Now recall $Y_{0,0}(\hat{\boldsymbol{n}}) = \frac{1}{\sqrt{4\pi}}$ and

$$Y_{1,1}(\hat{\boldsymbol{n}}) = -\sqrt{\frac{3}{8\pi}}\,\sin\theta\,e^{i\varphi} \qquad Y_{1,0}(\hat{\boldsymbol{n}}) = \sqrt{\frac{3}{4\pi}}\,\cos\theta \qquad Y_{1,-1}(\hat{\boldsymbol{n}}) = +\sqrt{\frac{3}{8\pi}}\,\sin\theta\,e^{-i\varphi} \tag{8.282}$$

$$S^0_{1/2}(x) = 1 \qquad S^0_{3/2}(x) = 1 \qquad S^1_{1/2}(x) = \tfrac{3}{2} - x \quad ,$$

which allows us to write

$$\begin{aligned} 1 &= 4\pi\,Y_{0,0}(\hat{\boldsymbol{n}})\,Y^*_{0,0}(\hat{\boldsymbol{n}}_1) \\ \hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}}_1 &= \frac{4\pi}{3}\Big[\,Y_{1,0}(\hat{\boldsymbol{n}})\,Y^*_{1,0}(\hat{\boldsymbol{n}}_1) + Y_{1,1}(\hat{\boldsymbol{n}})\,Y^*_{1,1}(\hat{\boldsymbol{n}}_1) + Y_{1,-1}(\hat{\boldsymbol{n}})\,Y^*_{1,-1}(\hat{\boldsymbol{n}}_1)\,\Big] \quad . \end{aligned} \tag{8.283}$$

We can do the integrals by appealing to the orthogonality relations for the spherical harmonics and Sonine polynomials:

$$\begin{aligned} \int d\hat{\boldsymbol{n}}\,Y_{\ell,m}(\hat{\boldsymbol{n}})\,Y^*_{\ell',m'}(\hat{\boldsymbol{n}}) &= \delta_{\ell,\ell'}\,\delta_{m,m'} \\ \int_0^\infty dx\,e^{-x}\,x^\alpha\,S^n_\alpha(x)\,S^{n'}_\alpha(x) &= \frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)}\,\delta_{n,n'} \quad . \end{aligned} \tag{8.284}$$

Integrating first over the direction vector $\hat{\boldsymbol{n}}_1$,

$$\begin{aligned} \mathcal{L}\psi = -\gamma\sum_{r\ell m} a_{r\ell m}(t)\,S^r_{\ell+\frac12}(x)\,x^{\ell/2}\,Y_{\ell,m}(\hat{\boldsymbol{n}}) \\ + \frac{2\gamma}{\sqrt{\pi}}\sum_{r\ell m} a_{r\ell m}(t)\int_0^\infty dx_1\,x_1^{1/2}\,e^{-x_1}\int d\hat{\boldsymbol{n}}_1\,\Big[Y_{0,0}(\hat{\boldsymbol{n}})\,Y^*_{0,0}(\hat{\boldsymbol{n}}_1)\,S^0_{1/2}(x)\,S^0_{1/2}(x_1) \\ + \tfrac23\,x^{1/2}x_1^{1/2}\sum_{m'=-1}^{1} Y_{1,m'}(\hat{\boldsymbol{n}})\,Y^*_{1,m'}(\hat{\boldsymbol{n}}_1)\,S^0_{3/2}(x)\,S^0_{3/2}(x_1) \\ + \tfrac23\,Y_{0,0}(\hat{\boldsymbol{n}})\,Y^*_{0,0}(\hat{\boldsymbol{n}}_1)\,S^1_{1/2}(x)\,S^1_{1/2}(x_1)\Big]\,S^r_{\ell+\frac12}(x_1)\,x_1^{\ell/2}\,Y_{\ell,m}(\hat{\boldsymbol{n}}_1) \quad , \end{aligned} \tag{8.285}$$

we obtain the intermediate result

$$\begin{aligned} \mathcal{L}\psi = -\gamma\sum_{r\ell m} a_{r\ell m}(t)\,S^r_{\ell+\frac12}(x)\,x^{\ell/2}\,Y_{\ell,m}(\hat{\boldsymbol{n}}) \\ + \frac{2\gamma}{\sqrt{\pi}}\sum_{r\ell m} a_{r\ell m}(t)\int_0^\infty dx_1\,x_1^{1/2}\,e^{-x_1}\Big[Y_{0,0}(\hat{\boldsymbol{n}})\,\delta_{l,0}\,\delta_{m,0}\,S^0_{1/2}(x)\,S^0_{1/2}(x_1) \\ + \tfrac23\,x^{1/2}x_1^{1/2}\sum_{m'=-1}^{1} Y_{1,m'}(\hat{\boldsymbol{n}})\,\delta_{l,1}\,\delta_{m,m'}\,S^0_{3/2}(x)\,S^0_{3/2}(x_1) \\ + \tfrac23\,Y_{0,0}(\hat{\boldsymbol{n}})\,\delta_{l,0}\,\delta_{m,0}\,S^1_{1/2}(x)\,S^1_{1/2}(x_1)\Big]\,S^r_{\ell+\frac12}(x_1)\,x_1^{1/2}. \end{aligned} \tag{8.286}$$

Appealing now to the orthogonality of the Sonine polynomials, and recalling that

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi} \qquad , \qquad \Gamma(1) = 1 \qquad , \qquad \Gamma(z+1) = z\,\Gamma(z) \quad , \tag{8.287}$$

we integrate over $x_1$. For the first term in brackets, we invoke the orthogonality relation with $n = 0$ and $\alpha = \frac{1}{2}$, giving $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi}$. For the second bracketed term, we have $n = 0$ but $\alpha = \frac{3}{2}$, and we obtain $\Gamma(\frac{5}{2}) = \frac{3}{2}\,\Gamma(\frac{3}{2})$, while the third bracketed term leads to $n = 1$ and $\alpha = \frac{1}{2}$, also yielding $\Gamma(\frac{5}{2}) = \frac{3}{2}\,\Gamma(\frac{3}{2})$. Thus, we obtain the simple and pleasing result

$$\mathcal{L}\psi = -\gamma {\sum_{r\ell m}}' a_{r\ell m}(t)\, S^r_{\ell+\frac{1}{2}}(x)\, x^{\ell/2}\, Y_{\ell,m}(\hat{\boldsymbol{n}}) \tag{8.288}$$

where the prime on the sum indicates that the set

$$\mathrm{CI} = \Big\{(0,0,0) \quad , \quad (1,0,0) \quad , \quad (0,1,1) \quad , \quad (0,1,0) \quad , \quad (0,1,-1)\Big\} \tag{8.289}$$

are to be excluded from the sum. But these are just the functions which correspond to the five collisional invariants! Thus, we learn that

$$\psi_{r\ell m}(\boldsymbol{v}) = \mathcal{N}_{r\ell m}\, S^r_{\ell+\frac{1}{2}}(x)\, x^{\ell/2}\, Y_{\ell,m}(\hat{\boldsymbol{n}}), \tag{8.290}$$

is an eigenfunction of $\mathcal{L}$ with eigenvalue $-\gamma$ if $(r,\ell,m)$ does not correspond to one of the five collisional invariants. In the latter case, the eigenvalue is zero. Thus, the algebraic action of $\mathcal{L}$ on the coefficients $a_{r\ell m}$ is

$$(\mathcal{L}a)_{r\ell m} = \begin{cases} -\gamma\, a_{r\ell m} & \text{if } (r,\ell,m) \notin \mathrm{CI} \\ = 0 & \text{if } (r,\ell,m) \in \mathrm{CI} \end{cases} \tag{8.291}$$

The quantity $\tau = \gamma^{-1}$ is the relaxation time.

It is pretty obvious that $\mathcal{L}$ is self-adjoint, since

$$\begin{aligned} \langle\, \phi \,|\, \mathcal{L}\psi \,\rangle = \langle\, \mathcal{L}\phi \,|\, \psi \,\rangle &\equiv \int d^3v\, f^0(\boldsymbol{v})\, \phi(\boldsymbol{v})\, \mathcal{L}[\psi(\boldsymbol{v})] \\ &= -\gamma\, n \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3/2} \int d^3v\, \exp\left(-\frac{m\boldsymbol{v}^2}{2k_{\rm B}T}\right) \phi(\boldsymbol{v})\, \psi(\boldsymbol{v}) \\ &\quad + \gamma\, n \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3} \int d^3v \int d^3u\, \exp\left(-\frac{m\boldsymbol{u}^2}{2k_{\rm B}T}\right) \exp\left(-\frac{m\boldsymbol{v}^2}{2k_{\rm B}T}\right) \\ &\qquad \times \phi(\boldsymbol{v}) \left[1 + \frac{m}{k_{\rm B}T}\, \boldsymbol{u}\cdot\boldsymbol{v} + \frac{2}{3}\left(\frac{m\boldsymbol{u}^2}{2k_{\rm B}T} - \frac{3}{2}\right)\left(\frac{m\boldsymbol{v}^2}{2k_{\rm B}T} - \frac{3}{2}\right)\right] \psi(\boldsymbol{u}) \quad , \end{aligned} \tag{8.292}$$

where $n$ is the bulk number density and $f^0(\boldsymbol{v})$ is the Maxwellian velocity distribution.

# Chapter 9

# Stochastic Processes

## 9.1 References

- C. Gardiner, *Stochastic Methods* ($4^{\text{th}}$ edition, Springer-Verlag, 2010)
  Very clear and complete text on stochastic methods, with many applications.
- N. G. Van Kampen *Stochastic Processes in Physics and Chemistry* ($3^{\text{rd}}$ edition, North-Holland, 2007)
  Another standard text. Very readable, but less comprehensive than Gardiner.
- Z. Schuss, *Theory and Applications of Stochastic Processes* (Springer-Verlag, 2010)
  In-depth discussion of continuous path stochastic processes and connections to partial differential equations.
- R. Mahnke, J. Kaupužs, and I. Lubashevsky, *Physics of Stochastic Processes* (Wiley, 2009)
  Introductory sections are sometimes overly formal, but a good selection of topics.
- A. N. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea, 1956)
  The *Urtext* of mathematical probability theory.

## 9.2 Introduction to Stochastic Processes

A *stochastic process* is one which is partially random, *i.e.* it is not wholly deterministic. Typically the randomness is due to phenomena at the microscale, such as the effect of fluid molecules on a small particle, such as a piece of dust in the air. The resulting motion (called *Brownian motion* in the case of particles moving in a fluid) can be described only in a statistical sense. That is, the full motion of the system is a *functional* of one or more independent random variables. The motion is then described by its averages with respect to the various random distributions.

### 9.2.1 Diffusion and Brownian motion

*Fick's law* (1855) is a phenomenological relationship between number current $\boldsymbol{j}$ and number density gradient $\boldsymbol{\nabla} n$, given by $\boldsymbol{j} = -D\,\boldsymbol{\nabla} n$. Combining this with the continuity equation $\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$, one arrives at the diffusion equation[1],

$$\frac{\partial n}{\partial t} = \boldsymbol{\nabla}\cdot(D\,\boldsymbol{\nabla} n) \quad . \tag{9.1}$$

Note that the diffusion constant $D$ may be position-dependent. The applicability of Fick's law was experimentally verified in many different contexts and has applicability to a wide range of transport phenomena in physics, chemistry, biology, ecology, geology, *etc.*

The eponymous Robert Brown, a botanist, reported in 1827 on the random motions of pollen grains suspended in water, which he viewed through a microscope. Apparently this phenomenon attracted little attention until the work of Einstein (1905) and Smoluchowski (1906), who showed how it is described by kinetic theory, in which the notion of randomness is essential, and also connecting it to Fick's laws of diffusion. Einstein began with the ideal gas law for osmotic pressure, $p = nk_{\rm B}T$. In steady state, the osmotic force per unit volume acting on the solute (*e.g.* pollen in water), $-\boldsymbol{\nabla} p$, must be balanced by viscous forces. Assuming the solute consists of spherical particles of radius $a$, the viscous force per unit volume is given by the hydrodynamic *Stokes drag* per particle $\boldsymbol{F} = -6\pi\eta a\boldsymbol{v}$ times the number density $n$, where $\eta$ is the dynamical viscosity of the solvent. Thus, $\boldsymbol{j} = n\boldsymbol{v} = -D\,\boldsymbol{\nabla} n$, where $D = k_{\rm B}T/6\pi a\eta$.

To connect this to kinetic theory, Einstein reasoned that the solute particles were being buffeted about randomly by the solvent, and he treated this problem statistically. While a given pollen grain is not significantly affected by any single collision with a water molecule, after some characteristic microscopic time $\tau$ the grain has effectively forgotten its initial conditions. Assuming there are no global currents, on average each grain's velocity is zero. Einstein posited that over an interval $\tau$, the number of grains which move a distance within $d^3\Delta$ of $\boldsymbol{\Delta}$ is $n\,\phi(\boldsymbol{\Delta})\,d^3\Delta$, where $\phi(\boldsymbol{\Delta}) = \phi\big(|\boldsymbol{\Delta}|\big)$ is isotropic and also normalized according to $\int d^3\Delta\,\phi(\Delta) = 1$. Then

$$n(\boldsymbol{x}, t+\tau) = \int d^3\Delta\, n(\boldsymbol{x}-\boldsymbol{\Delta}, t)\,\phi(\boldsymbol{\Delta}) \quad , \tag{9.2}$$

Taylor expanding in both space and time, to lowest order in $\tau$ one recovers the diffusion equation,

[1]The equation $\boldsymbol{j} = -D\,\boldsymbol{\nabla} n$ is sometimes called Fick's first law, and the continuity equation $\partial_t n = -\boldsymbol{\nabla}\cdot\boldsymbol{j}$ Fick's second law.

$\partial_t n = D\,\nabla^2 n$, where the diffusion constant is given by

$$D = \frac{1}{6\tau}\int d^3\Delta\;\phi(\Delta)\,\Delta^2 \quad . \tag{9.3}$$

The diffusion equation with constant $D$ is easily solved by taking the spatial Fourier transform. One then has, in $d$ spatial dimensions,

$$\frac{\partial\hat{n}(\boldsymbol{k},t)}{\partial t} = -D\boldsymbol{k}^2\hat{n}(\boldsymbol{k},t) \qquad \Rightarrow \qquad n(\boldsymbol{x},t) = \int\frac{d^dk}{(2\pi)^d}\,\hat{n}(\boldsymbol{k},t_0)\,e^{-D\boldsymbol{k}^2(t-t_0)}\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \quad . \tag{9.4}$$

If $n(\boldsymbol{x},t_0) = \delta(\boldsymbol{x}-\boldsymbol{x}_0)$, corresponding to $\hat{n}(\boldsymbol{k},t_0) = e^{-i\boldsymbol{k}\cdot\boldsymbol{x}_0}$, we have

$$n(\boldsymbol{x},t) = \big(4\pi D|t-t_0|\big)^{-d/2}\exp\left\{-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{4D|t-t_0|}\right\} \quad , \tag{9.5}$$

where $d$ is the dimension of space.

**WTF just happened?**

We're so used to diffusion processes that most of us overlook a rather striking aspect of the above solution to the diffusion equation. At $t = t_0$, the probability density is $P(\boldsymbol{x}, t=t_0) = \delta(\boldsymbol{x}-\boldsymbol{x}_0)$, which means all the particles are sitting at $\boldsymbol{x} = \boldsymbol{x}_0$. For any $t > t_0$, the solution is given by Eqn. 9.5, which is nonzero for all $\boldsymbol{x}$. If we take a value of $\boldsymbol{x}$ such that $|\boldsymbol{x}-\boldsymbol{x}_0| > ct$, where $c$ is the speed of light, we see that there is a finite probability, however small, for particles to diffuse at superluminal speeds. Clearly this is nonsense. The error lies in the diffusion equation itself, which does not recognize any limiting propagation speed. For most processes, this defect is harmless, as we are not interested in the extreme tails of the distribution. Diffusion phenomena and the applicability of the diffusion equation are well-established in virtually every branch of science. To account for a finite propagation speed, one is forced to consider various generalizations of the diffusion equation. Some examples are discussed in the appendix §9.6.

### 9.2.2 Langevin equation

Consider a particle of mass $M$ subjected to dissipative and random forcing. We'll examine this system in one dimension to gain an understanding of the essential physics. We write

$$\dot{u} + \gamma u = \frac{F}{M} + \eta(t) \quad . \tag{9.6}$$

Here, $u$ is the particle's velocity, $\gamma$ is the damping rate due to friction, $F$ is a constant external force, and $\eta(t)$ is a *stochastic random force*. This equation, known as the Langevin equation, describes a ballistic particle being buffeted by random forcing events[2]. Think of a particle of dust as it moves in the atmosphere. $F$ would then represent the external force due to gravity and $\eta(t)$ the random forcing

[2]See the appendix in §9.7 for the solution of the Langevin equation for a particle in a harmonic well.

due to interaction with the air molecules. For a sphere of radius $a$ moving in a fluid of dynamical viscosity $\eta$ – *not to be confused with the stochastic function* $\eta(t)$*!* – hydrodynamics gives $\gamma = 6\pi\eta a/M$, where $M$ is the mass of the particle. It is illustrative to compute $\gamma$ in some setting. Consider a micron sized droplet ($a = 10^{-4}\,\text{cm}$) of some liquid of density $\rho \sim 1.0\,\text{g/cm}^3$ moving in air at $T = 20^\circ\,\text{C}$. The viscosity of air is $\eta = 1.8 \times 10^{-4}\,\text{g/cm}\cdot\text{s}$ at this temperature[3]. If the droplet density is constant, then $\gamma = 9\eta/2\rho a^2 = 8.1 \times 10^4\,\text{s}^{-1}$, hence the time scale for viscous relaxation of the particle is $\tau = \gamma^{-1} = 12\,\mu\text{s}$. We should stress that the viscous damping on the particle is of course due to the fluid (*e.g.*, air) molecules, in some average 'coarse-grained' sense. The random component to the force $\eta(t)$ would then represent the fluctuations with respect to this average.

We can easily integrate this equation:

$$\begin{aligned}\frac{d}{dt}\left(u\,e^{\gamma t}\right) &= \frac{F}{M}\,e^{\gamma t} + \eta(t)\,e^{\gamma t} \\ u(t) &= u(0)\,e^{-\gamma t} + \frac{F}{\gamma M}\left(1 - e^{-\gamma t}\right) + \int_0^t ds\,\eta(s)\,e^{\gamma(s-t)}\end{aligned} \tag{9.7}$$

Note that $u(t)$ is indeed a functional of the random function $\eta(t)$. We can therefore only compute averages in order to describe the motion of the system.

The first average we will compute is that of $u$ itself. In so doing, we assume that $\eta(t)$ has zero mean: $\big\langle \eta(t) \big\rangle = 0$. Then

$$\big\langle u(t) \big\rangle = u(0)\,e^{-\gamma t} + \frac{F}{\gamma M}\left(1 - e^{-\gamma t}\right) \quad . \tag{9.8}$$

On the time scale $\gamma^{-1}$, the initial conditions $u(0)$ are effectively forgotten, and asymptotically for $t \gg \gamma^{-1}$ we have $\big\langle u(t) \big\rangle \to F/\gamma M$, which is the terminal velocity.

Next, consider

$$\big\langle u^2(t) \big\rangle = \big\langle u(t) \big\rangle^2 + \int_0^t ds_1 \int_0^t ds_2\, e^{\gamma(s_1 - t)}\, e^{\gamma(s_2 - t)}\, \big\langle \eta(s_1)\,\eta(s_2) \big\rangle \quad . \tag{9.9}$$

We now need to know the two-time correlator $\big\langle \eta(s_1)\,\eta(s_2) \big\rangle$. We assume that the correlator is a function only of the time difference $\Delta s = s_1 - s_2$, and that the random force $\eta(s)$ has zero average, $\big\langle \eta(s) \big\rangle = 0$, and autocorrelation

$$\big\langle \eta(s_1)\,\eta(s_2) \big\rangle = \phi(s_1 - s_2) \quad . \tag{9.10}$$

The function $\phi(s)$ is the *autocorrelation function* of the random force. A macroscopic object moving in a fluid is constantly buffeted by fluid particles over its entire perimeter. These different fluid particles are almost completely uncorrelated, hence $\phi(s)$ is essentially zero except on a very small time scale $\tau_\phi$, which is the time a single fluid particle spends interacting with the object. We can take $\tau_\phi \to 0$ and approximate $\phi(s) \approx \Gamma\,\delta(s)$. We shall determine the value of $\Gamma$ from equilibrium thermodynamic considerations below.

[3]The cgs unit of viscosity is the *Poise* (P). $1\,\text{P} = 1\,\text{g/cm}\cdot\text{s}$.

With this form for $\phi(s)$, we can easily calculate the equal time momentum autocorrelation:

$$\big\langle u^2(t)\big\rangle = \big\langle u(t)\big\rangle^2 + \Gamma\int\limits_0^t ds\, e^{2\gamma(s-t)} = \big\langle u(t)\big\rangle^2 + \frac{\Gamma}{2\gamma}\left(1 - e^{-2\gamma t}\right) \quad . \tag{9.11}$$

Consider the case where $F = 0$ and the limit $t \gg \gamma^{-1}$. We demand that the object thermalize at temperature $T$. Thus, we impose the condition

$$\big\langle\, \tfrac{1}{2}Mu^2(t)\,\big\rangle = \tfrac{1}{2}k_{\rm B}T \qquad \Longrightarrow \qquad \Gamma = \frac{2\gamma k_{\rm B}T}{M} \quad . \tag{9.12}$$

This fixes the value of $\Gamma$.

We can now compute the general momentum autocorrelator:

$$\big\langle v(t)\,v(t')\big\rangle - \big\langle v(t)\big\rangle\big\langle v(t')\big\rangle = \int\limits_0^t ds \int\limits_0^{t'} ds'\, e^{\gamma(s-t)}\, e^{\gamma(s'-t')}\,\big\langle \zeta(s)\,\zeta(s')\big\rangle = \frac{\Gamma}{2\gamma}\Big[e^{-\gamma|t-t'|} - e^{-\gamma|t+t'|}\Big] \quad . \tag{9.13}$$

In the limit where $t$ and $t'$ each tend to infinity with finite $|t-t'|$, the above covariance is approximately $\mathsf{Cov}\big[v(t),v(t')\big] \approx (\Gamma/2\gamma)\exp\big(-\gamma|t-t'|\big)$.

Let's now compute the position $x(t)$. We find

$$x(t) = \big\langle x(t)\big\rangle + \int\limits_0^t ds \int\limits_0^s ds_1\, \eta(s_1)\, e^{\gamma(s_1 - s)} \quad , \tag{9.14}$$

where

$$\big\langle x(t)\big\rangle = x(0) + \frac{1}{\gamma}\left(u(0) - \frac{F}{\gamma M}\right)\left(1 - e^{-\gamma t}\right) + \frac{Ft}{\gamma M} \quad . \tag{9.15}$$

Note that for $\gamma t \ll 1$ we have $\big\langle x(t)\big\rangle = x(0) + u(0)\,t + \frac{1}{2}M^{-1}Ft^2 + \mathcal{O}(t^3)$, as is appropriate for ballistic particles moving under the influence of a constant force. This long time limit of course agrees with our earlier evaluation for the terminal velocity, $\big\langle u(\infty)\big\rangle = F/\gamma M$. We next compute the position autocorrelation:

$$\begin{aligned}\big\langle x(t)\,x(t')\big\rangle - \big\langle x(t)\big\rangle\big\langle x(t')\big\rangle &= \int\limits_0^t ds \int\limits_0^{t'} ds'\, e^{-\gamma(s+s')} \int\limits_0^s du \int\limits_0^{s'} du'\, e^{\gamma(u+u')}\,\big\langle \eta(u)\,\eta(u')\big\rangle \\ &= \frac{2k_{\rm B}T}{\gamma M}\mathsf{min}(t,t') + \mathcal{O}(1) \quad .\end{aligned}$$

In particular, the equal time autocorrelator is

$$\big\langle x^2(t)\big\rangle - \big\langle x(t)\big\rangle^2 = \frac{2k_{\rm B}T\,t}{\gamma M} \equiv 2D\,t \quad , \tag{9.16}$$

at long times, up to terms of order unity. Here, $D = \Gamma/2\gamma^2 = k_{\rm B}T/\gamma M$ is the *diffusion constant*. For a liquid droplet of radius $a = 1\,\mu{\rm m}$ moving in air at $T = 293\,{\rm K}$, for which $\eta = 1.8 \times 10^{-4}\,{\rm P}$, we have

$$D = \frac{k_{\rm B}T}{6\pi\eta a} = \frac{(1.38 \times 10^{-16}\,{\rm erg/K})\,(293\,{\rm K})}{6\pi\,(1.8 \times 10^{-4}\,{\rm P})\,(10^{-4}\,{\rm cm})} = 1.19 \times 10^{-7}\,{\rm cm}^2/{\rm s} \quad . \tag{9.17}$$

This result presumes that the droplet is large enough compared to the intermolecular distances in the fluid that one can adopt a continuum approach and invoke the Navier-Stokes equations for laminar flow.

If we consider molecular diffusion, the situation is quite a bit different. The diffusion constant is then $D = \frac{1}{3}\ell\bar{v}$, where $\ell$ is the mean free path and $\bar{v}$ the average molecular speed. This estimate treats collisions as approximately randomizing the direction of motion, and a more detailed treatment results in a modified numerical prefactor. For identical hard sphere molecules of radius $a$, the total scattering cross section is $\sigma = 4\pi a^2$ and elementary kinetic theory gives

$$\ell = \frac{1}{\sqrt{2}\,n\sigma} \qquad , \qquad \bar{v} = \sqrt{\frac{8k_{\rm B}T}{\pi m}} \quad . \tag{9.18}$$

At $T = 293\,{\rm K}$ and $p = 1\,{\rm atm}$, the number density is $n = p/k_{\rm B}T = 2.51 \times 10^{19}\,{\rm cm}^{-3}$. Since air is some 79% $N_2$, we take $a = 1.9 \times 10^{-8}\,{\rm cm}$ and $m = 28.0\,{\rm amu} = 4.65 \times 10^{-23}\,{\rm g}$. We then obtain an average molecular speed of $\bar{v} = 471\,{\rm m/s}$. a mean free path of $\ell = 6.2 \times 10^{-6}\,{\rm cm}$, and a diffusion constant of $D \simeq 0.097\,{\rm cm}^2/{\rm s}$. While this is much larger than the diffusion constant for micron-sized droplets, it is nevertheless too small to explain various common experiences. For particles initially released from a point source, the mean square radial displacement is $\langle r^2\rangle = 6Dt$, thus reaching an rms radius of $d = 10\,{\rm cm}$ should take approximately $d^2/6D \simeq 170\,{\rm s}$, or nearly three minutes. Yet if someone in the next seat emits a foul odor[4], you detect the offending emission in on the order of a second. What this tells us is that diffusion isn't the only transport process involved in these and like phenomena. More important here are *convection* currents which distribute the scent much more rapidly.

## 9.3 Distributions and Functionals

### 9.3.1 Basic definitions

Let $x \in \mathbb{R}$ be a random variable, and $P(x)$ a probability distribution for $x$. The average of any function $\phi(x)$ is then

$$\big\langle \phi(x) \big\rangle = \int\limits_{-\infty}^{\infty} dx\; P(x)\,\phi(x) \Bigg/ \int\limits_{-\infty}^{\infty} dx\; P(x) \quad . \tag{9.19}$$

Let $\eta(t)$ be a random *function* of $t$, with $\eta(t) \in \mathbb{R}$, and let $P\big[\eta(t)\big]$ be the probability distribution *functional* for $\eta(t)$. Then if $\Phi\big[\eta(t)\big]$ is a functional of $\eta(t)$, the average of $\Phi$ is given by

$$\int D\eta\; P\big[\eta(t)\big]\; \Phi\big[\eta(t)\big] \Bigg/ \int D\eta\, P\big[\eta(t)\big] \tag{9.20}$$

[4]Gross!

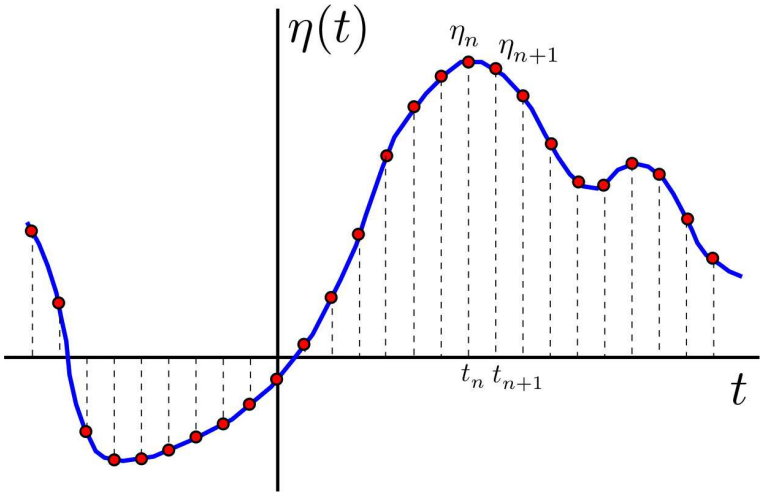



Figure 9.1: Discretization of a continuous function $\eta(t)$. Upon discretization, a functional $\Phi\big[\eta(t)\big]$ becomes an ordinary multivariable function $\Phi(\{\eta_j\})$.

The expression $\int D\eta\ P[\eta]\,\Phi[\eta]$ is a *functional integral*. A functional integral is a continuum limit of a multivariable integral. Suppose $\eta(t)$ were defined on a set of $t$ values $t_n = n\tau$. A functional of $\eta(t)$ becomes a multivariable function of the values $\eta_n \equiv \eta(t_n)$. The metric then becomes $D\eta = \prod_n d\eta_n$ .

In fact, for our purposes we will not need to know any details about the functional measure $D\eta$ ; we will finesse this delicate issue[5]. Consider the *generating functional*,

$$Z\big[J(t)\big] = \int D\eta\ P[\eta]\ \exp\left\{\int_{-\infty}^{\infty} dt\ J(t)\,\eta(t)\right\} \quad . \tag{9.21}$$

It is clear that

$$\frac{1}{Z[J]}\,\frac{\delta^n Z[J]}{\delta J(t_1)\cdots\delta J(t_n)}\bigg|_{J(t)=0} = \big\langle \eta(t_1)\cdots\eta(t_n)\big\rangle \quad . \tag{9.22}$$

The function $J(t)$ is an arbitrary *source function*. We functionally differentiate with respect to it in order to find the $\eta$-field correlators. The functional derivative $\delta Z\big[J(t)\big]/\delta J(s)$ can be computed by substituting $J(t) \to J(t) + \epsilon\,\delta(t-s)$ inside the functional $Z[J]$, and then taking the ordinary derivative with respect to $\varepsilon$, *i.e.*

$$\frac{\delta Z\big[J(t)\big]}{\delta J(s)} = \frac{dZ\big[J(t) + \varepsilon\,\delta(t-s)\big]}{d\varepsilon}\bigg|_{\varepsilon=0} \quad . \tag{9.23}$$

Thus the functional derivative $\delta Z\big[J(t)\big]/\delta J(s)$ tells us how the functional $Z[J]$ changes when the function $J(t)$ is replaced by $J(t) + \varepsilon\,\delta(t-s)$. Equivalently, one may eschew this $\varepsilon$ prescription and use the familiar chain rule from differential calculus, supplemented by the rule $\delta J(t)/\delta J(s) = \delta(t-s)$ .

[5] A discussion of measure for functional integrals is found in R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*.

Let's compute the generating functional for a class of distributions of the Gaussian form,

$$P[\eta] = \exp\left\{-\frac{1}{2\Gamma}\int_{-\infty}^{\infty} dt\,\left(\tau^2\,\dot\eta^2 + \eta^2\right)\right\} = \exp\left\{-\frac{1}{2\Gamma}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\left(1+\omega^2\tau^2\right)\left|\hat\eta(\omega)\right|^2\right\} \quad . \tag{9.24}$$

Then Fourier transforming the source function $J(t)$, it is easy to see that

$$Z[J] = Z[0]\cdot\exp\left\{\frac{\Gamma}{2}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\frac{\left|\hat J(\omega)\right|^2}{1+\omega^2\tau^2}\right\} \quad . \tag{9.25}$$

Note that with $\eta(t)\in\mathbb{R}$ and $J(t)\in\mathbb{R}$ we have $\eta^*(\omega)=\eta(-\omega)$ and $\hat J^*(\omega)=\hat J(-\omega)$. Transforming back to real time, we have

$$Z[J] = Z[0]\cdot\exp\left\{\frac{1}{2}\int_{-\infty}^{\infty}dt\int_{-\infty}^{\infty}dt'\,J(t)\,G(t-t')\,J(t')\right\} \quad , \tag{9.26}$$

where

$$G(s) = \frac{\Gamma}{2\tau}\,e^{-|s|/\tau} \qquad , \qquad \widehat G(\omega) = \frac{\Gamma}{1+\omega^2\tau^2} \tag{9.27}$$

is the *Green's function*, in real and Fourier space. Note that

$$\int_{-\infty}^{\infty} ds\,G(s) = \widehat G(0) = \Gamma \quad . \tag{9.28}$$

We can now compute

$$\begin{aligned}\left\langle\eta(t_1)\,\eta(t_2)\right\rangle &= G(t_1-t_2)\\ \left\langle\eta(t_1)\,\eta(t_2)\,\eta(t_3)\,\eta(t_4)\right\rangle &= G(t_1-t_2)\,G(t_3-t_4) + G(t_1-t_3)\,G(t_2-t_4) + G(t_1-t_4)\,G(t_2-t_3) \quad .\end{aligned} \tag{9.29}$$

The generalization is now easy to prove, and is known as *Wick's theorem*:

$$\left\langle\eta(t_1)\cdots\eta(t_{2n})\right\rangle = \sum_{\text{contractions}} G(t_{i_1}-t_{i_2})\cdots G(t_{i_{2n-1}}-t_{i_{2n}}) \quad , \tag{9.30}$$

where the sum is over all distinct *contractions* of the sequence $1\cdot 2\cdots 2n$ into products of pairs. How many terms are there? Some simple combinatorics answers this question. Choose the index $1$. There are $(2n-1)$ other time indices with which it can be contracted. Now choose another index. There are $(2n-3)$ indices with which *that* index can be contracted. And so on. We thus obtain

$$C(n) \equiv \left\{\begin{matrix}\text{\# of contractions}\\ \text{of 1-2-3}\cdots 2n\end{matrix}\right\} = (2n-1)(2n-3)\cdots 3\cdot 1 = \frac{(2n)!}{2^n\,n!} \quad . \tag{9.31}$$

### 9.3.2 Correlations for the Langevin equation

Now suppose we have the Langevin equation

$$\frac{du}{dt} + \gamma u = \eta(t) \tag{9.32}$$

with $u(0) = 0$. We wish to compute the joint probability density

$$P(u_1, t_1; \ldots; u_N, t_N) = \Big\langle \delta\big(u_1 - u(t_1)\big) \cdots \delta\big(u_N - u(t_N)\big) \Big\rangle \quad , \tag{9.33}$$

where the average is over all realizations of the random variable $\eta(t)$:

$$\big\langle F\big[\eta(t)\big] \big\rangle = \int D\eta \, P\big[\eta(t)\big] \, F\big[\eta(t)\big] \quad . \tag{9.34}$$

Using the integral representation of the Dirac $\delta$-function, we have

$$P(u_1, t_1; \ldots; u_N, t_N) = \int\limits_{-\infty}^{\infty} \frac{d\omega_1}{2\pi} \, \cdots \int\limits_{-\infty}^{\infty} \frac{d\omega_N}{2\pi} \, e^{-i(\omega_1 u_1 + \ldots + \omega_N u_N)} \left\langle e^{i\omega_1 u(t_1)} \cdots e^{i\omega_N u(t_N)} \right\rangle \quad . \tag{9.35}$$

Now integrating the Langevin equation with the initial condition $u(0) = 0$ gives

$$u(t_j) = \int\limits_0^{t_j} dt \, e^{\gamma(t - t_j)} \, \eta(t) \quad , \tag{9.36}$$

and therefore we may write

$$\sum_{j=1}^{N} \omega_j \, u(t_j) = \int\limits_{-\infty}^{\infty} dt \, f(t) \, \eta(t) \qquad , \qquad f(t) = \sum_{j=1}^{N} \omega_j \, e^{\gamma(t - t_j)} \, \Theta(t) \, \Theta(t_j - t) \quad . \tag{9.37}$$

We assume that the random variable $\eta(t)$ is distributed as a Gaussian, with $\big\langle \eta(t) \, \eta(t') \big\rangle = G(t - t')$, as described above. Using our previous results, we may perform the functional integral over $\eta(t)$ to obtain

$$\left\langle \exp \, i \int\limits_{-\infty}^{\infty} dt \, f(t) \, \eta(t) \right\rangle = \exp\left\{ -\frac{1}{2} \int\limits_{-\infty}^{\infty} dt \int\limits_{-\infty}^{\infty} dt' \, G(t - t') \, f(t) \, f(t') \right\} = \exp\left\{ -\frac{1}{2} \sum_{j,j'=1}^{N} M_{jj'} \, \omega_j \, \omega_{j'} \right\} \quad , \tag{9.38}$$

where $M_{jj'} = M(t_j, t_{j'})$ with

$$M(t, t') = \int\limits_0^t ds \int\limits_0^{t'} ds' \, G(s - s') \, e^{\gamma(s - t)} \, e^{\gamma(s' - t')} \quad . \tag{9.39}$$

We now have

$$\begin{aligned} P(u_1,t_1;\ldots;u_N,t_N) &= \int\limits_{-\infty}^{\infty}\frac{d\omega_1}{2\pi}\,\cdots\int\limits_{-\infty}^{\infty}\frac{d\omega_N}{2\pi}\,e^{-i(\omega_1u_1+\ldots+\omega_Nu_N)}\,\exp\left\{-\tfrac{1}{2}\sum_{j,j'=1}^{N}M_{jj'}\,\omega_j\,\omega_{j'}\right\} \\ &= \det{}^{-1/2}(2\pi M)\,\exp\left\{-\tfrac{1}{2}\sum_{j,j'=1}^{N}M^{-1}_{jj'}\,u_j\,u_{j'}\right\}\quad . \end{aligned} \tag{9.40}$$

In the limit $G(s)=\Gamma\,\delta(s)$, we have

$$M_{jj'}=\Gamma\int\limits_0^{\min(t_j,t_{j'})}dt\,e^{2\gamma t}\,e^{-\gamma(t_j+t_{j'})}=\frac{\Gamma}{2\gamma}\left(e^{-\gamma|t_j-t_{j'}|}-e^{-\gamma(t_j+t_{j'})}\right)\quad . \tag{9.41}$$

Setting $N=1$, one has $M_{11}(t)=\frac{\Gamma}{2\gamma}\big(1-e^{-2\gamma t}\big)$, in which case the one-point distribution is found to be

$$\begin{aligned} P(u,t) &= \int\limits_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,e^{-i\omega u}\,\exp\left(-\frac{\Gamma(1-e^{-2\gamma t})\,\omega^2}{4\gamma}\right) \\ &= \sqrt{\frac{\gamma}{\pi\Gamma\,(1-e^{-2\gamma t})}}\,\exp\left(-\frac{\gamma u^2}{\Gamma(1-e^{-2\gamma t})}\right) \end{aligned} \tag{9.42}$$

This is a Gaussian distribution which interpolates between $P(u,0)=\delta(u)$ at time $t=0$ and at late times takes the form

$$P(u,t\gg\gamma^{-1})=(\gamma/\pi\Gamma)^{1/2}\exp(-\gamma u^2/\Gamma)\quad . \tag{9.43}$$

The conditional distribution $P(u_1,t_1\,|\,u_2,t_2)=P(u_1,t_1;u_2,t_2)/P(u_2,t_2)$ is then found to be

$$P(u_1,t_1\,|\,u_2,t_2)=\sqrt{\frac{\gamma/\pi\Gamma}{1-e^{-2\gamma(t_1-t_2)}}}\,\exp\left\{-\frac{\gamma}{\Gamma}\cdot\frac{\big(u_1-e^{-\gamma(t_1-t_2)}\,u_2\big)^2}{1-e^{-2\gamma(t_1-t_2)}}\right\}\quad . \tag{9.44}$$

Here we have assumed $t_{1,2}\gg\gamma^{-1}$, but have made no assumptions about the magnitude of $\gamma|t_1-t_2|$. Note that $P(u_1,t_1\,|\,u_2,t_2)$ tends to $P(u_1,t_1)$ independent of the most recent condition, so long as we are in the limit $|t_1-t_2|\gg\gamma^{-1}$.

As we shall discuss below, a *Markov process* is one where, at any given time, the statistical properties of the subsequent evolution are fully determined by the state of the system at that time. Equivalently, every *conditional probability* depends only on the *most recent condition*. Is $u(t)$ a continuous time Markov process? It is within the white noise approximation $G(s)=\Gamma\,\delta(s)$. Note that $u(t)$ satisfies a first order differential equation, hence only the initial condition on $u$ is necessary in order to derive its probability distribution at any time in the future. Explicitly, we can compute $P(u_1t_1|\,u_2t_2,u_3t_3)$ and show that it is independent of $u_3$ and $t_3$ for $t_1>t_2>t_3$. This is true regardless of the relative sizes of $t_j-t_{j+1}$ and $\gamma^{-1}$.

While $u(t)$ defines a Markov process, its integral $x(t)$ does not. This is because more information than the initial value of $x$ is necessary in order to integrate forward to a solution at future times. Since $x(t)$

satisfies a second order ODE, its conditional probabilities should in principle depend only on its *two most recent conditions*. We could also consider the evolution of the pair $\boldsymbol{\varphi} = (x, u)$ in *phase space*, writing

$$\frac{d}{dt}\begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & -\gamma \end{pmatrix}\begin{pmatrix} x \\ u \end{pmatrix} + \begin{pmatrix} 0 \\ \eta(t) \end{pmatrix} \quad , \tag{9.45}$$

or $\dot{\boldsymbol{\varphi}} = A\boldsymbol{\varphi} + \boldsymbol{\eta}(t)$, where $A$ is the above $2 \times 2$ matrix, and the stochastic term $\boldsymbol{\eta}(t)$ has only a lower component. The paths $\boldsymbol{\varphi}(t)$ are also Markovian, because they are determined by a first order set of coupled ODEs. In the limit where $t_j - t_{j+1} \gg \gamma^{-1}$, $x(t)$ effectively becomes Markovian, because we interrogate the paths on time scales where the separations are such that the particle has 'forgotten' its initial velocity.

See the appendix §9.8 for the solution of general linear ODEs with random forcing.

## 9.4 The Fokker-Planck Equation

### 9.4.1 Basic derivation

Suppose $x(t)$ is a stochastic variable. We define the quantity

$$\delta x(t) \equiv x(t + \delta t) - x(t) \quad , \tag{9.46}$$

and we assume

$$\big\langle \delta x(t) \big\rangle = F_1\big(x(t)\big)\,\delta t \qquad , \qquad \big\langle \big[\delta x(t)\big]^2 \big\rangle = F_2\big(x(t)\big)\,\delta t \quad , \tag{9.47}$$

but also that $\big\langle \big[\delta x(t)\big]^n \big\rangle = \mathcal{O}\big((\delta t)^2\big)$ for $n > 2$. The $n = 1$ term is due to *drift* and the $n = 2$ term is due to *diffusion*. Now consider the conditional probability density, $P(x, t \,|\, x_0, t_0)$, defined to be the probability distribution for $x \equiv x(t)$ given that $x(t_0) = x_0$. The conditional probability density satisfies the composition rule,

$$P(x_2, t_2 \,|\, x_0, t_0) = \int\limits_{-\infty}^{\infty} dx_1\, P(x_2, t_2 \,|\, x_1, t_1)\, P(x_1, t_1 \,|\, x_0, t_0) \quad , \tag{9.48}$$

for any value of $t_1$. This is also known as the *Chapman-Kolmogorov equation*. In words, what it says is that the probability density for a particle being at $x_2$ at time $t_2$, given that it was at $x_0$ at time $t_0$, is given by the product of the probability density for being at $x_2$ at time $t_2$ given that it was at $x_1$ at $t_1$, multiplied by that for being at $x_1$ at $t_1$ given it was at $x_0$ at $t_0$, integrated over $x_1$. This should be intuitively obvious, since if we pick any time $t_1 \in [t_0, t_2]$, then the particle had to be *somewhere* at that time. What is perhaps not obvious is why the conditional probability $P(x_2, t_2 \,|\, x_1, t_1)$ does not also depend on $(x_0, t_0)$. This is so if the system is described by a *Markov process*, about which we shall have more to say below in §9.5.1. At any rate, a picture is worth a thousand words: see Fig. 9.2.

Proceeding, we may write

$$P(x, t + \delta t \,|\, x_0, t_0) = \int\limits_{-\infty}^{\infty} dx'\, P(x, t + \delta t \,|\, x', t)\, P(x', t \,|\, x_0, t_0) \quad . \tag{9.49}$$

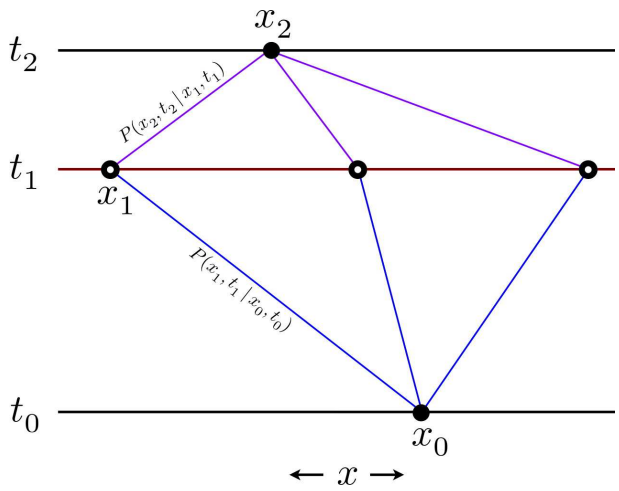



Figure 9.2: Interpretive sketch of the mathematics behind the Chapman-Kolmogorov equation.

Now

$$\begin{aligned} P(x,t+\delta t\,|\,x',t) &= \big\langle \delta\big(x-\delta x(t)-x'\big)\big\rangle \\ &= \left\{1 + \big\langle \delta x(t)\big\rangle \frac{d}{dx'} + \tfrac{1}{2}\big\langle \big[\delta x(t)\big]^2\big\rangle \frac{d^2}{dx'^2} + \ldots \right\} \delta(x-x') \\ &= \delta(x-x') + F_1(x') \frac{d\,\delta(x-x')}{dx'}\,\delta t + \tfrac{1}{2} F_2(x') \frac{d^2\delta(x-x')}{dx'^2}\,\delta t + \mathcal{O}\big((\delta t)^2\big) \quad , \end{aligned} \tag{9.50}$$

where the average is over the random variables. We now insert this result into eqn. 9.49, integrate by parts, divide by $\delta t$, and then take the limit $\delta t \to 0$. The result is the Fokker-Planck equation,

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}\big[F_1(x)\,P(x,t)\big] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\big[F_2(x)\,P(x,t)\big] \quad . \tag{9.51}$$

### 9.4.2 Brownian motion redux

Let's apply our Fokker-Planck equation to a description of Brownian motion. From our earlier results, we have $F_1(x) = F/\gamma M$ and $F_2(x) = 2D$. A formal proof of these results is left as an exercise for the reader. The Fokker-Planck equation is then

$$\frac{\partial P}{\partial t} = -u\frac{\partial P}{\partial x} + D\frac{\partial^2 P}{\partial x^2} \quad , \tag{9.52}$$

where $u = F/\gamma M$ is the average terminal velocity. If we make a Galilean transformation and define $y = x - ut$ and $s = t$, then our Fokker-Planck equation takes the form

$$\frac{\partial P}{\partial s} = D\frac{\partial^2 P}{\partial y^2} \quad . \tag{9.53}$$

This is known as the *diffusion equation*. Eqn. 9.52 is also a diffusion equation, rendered in a moving frame.

While the Galilean transformation is illuminating, we can easily solve eqn. 9.52 without it. Let's take a look at this equation after Fourier transforming from $x$ to $q$:

$$P(x,t) = \int_{-\infty}^{\infty} \frac{dq}{2\pi}\, e^{iqx}\, \hat{P}(q,t) \tag{9.54}$$

$$\hat{P}(q,t) = \int_{-\infty}^{\infty} dx\, e^{-iqx}\, P(x,t) \quad . \tag{9.55}$$

Then as should be well known to you by now, we can replace the operator $\frac{\partial}{\partial x}$ with multiplication by $iq$, resulting in

$$\frac{\partial}{\partial t}\,\hat{P}(q,t) = -(Dq^2 + iqu)\,\hat{P}(q,t) \quad , \tag{9.56}$$

with solution

$$\hat{P}(q,t) = e^{-Dq^2 t}\, e^{-iqut}\, \hat{P}(q,0) \quad . \tag{9.57}$$

We now apply the inverse transform to get back to $x$-space:

$$\begin{aligned} P(x,t) &= \int_{-\infty}^{\infty} \frac{dq}{2\pi}\, e^{iqx}\, e^{-Dq^2 t}\, e^{-iqut} \int_{-\infty}^{\infty} dx'\, e^{-iqx'}\, P(x',0) \\ &= \int_{-\infty}^{\infty} dx'\, P(x',0) \int_{-\infty}^{\infty} \frac{dq}{2\pi}\, e^{-Dq^2 t}\, e^{iq(x-ut-x')} = \int_{-\infty}^{\infty} dx'\, K(x-x',t)\, P(x',0) \quad , \end{aligned} \tag{9.58}$$

where

$$K(x,t) = \frac{1}{\sqrt{4\pi D t}}\, e^{-(x-ut)^2/4Dt} \tag{9.59}$$

is the *diffusion kernel*. We now have a recipe for obtaining $P(x,t)$ given the initial conditions $P(x,0)$. If $P(x,0) = \delta(x)$, describing a particle confined to an infinitesimal region about the origin, then $P(x,t) = K(x,t)$ is the probability distribution for finding the particle at $x$ at time $t$. There are two aspects to $K(x,t)$ which merit comment. The first is that the center of the distribution moves with velocity $u$. This is due to the presence of the external force. The second is that the standard deviation $\sigma = \sqrt{2Dt}$ is increasing in time, so the distribution is not only shifting its center but it is also getting broader as time evolves. This movement of the center and broadening are what we have called *drift* and *diffusion*, respectively.

### 9.4.3 Ornstein-Uhlenbeck process

Starting from any initial condition $P(x,0)$, the Fokker-Planck equation for Brownian motion, even with drift, inexorably evolves the distribution $P(x,t)$ toward an infinitesimal probability uniformly spread throughout all space. Consider now the Fokker-Planck equation with $F_2(x) = 2D$ as before, but with $F_1(x) = -\beta x$. Thus we have diffusion but also drift, where the local velocity is $-\beta x$. For $x > 0$, probability which diffuses to the right will also drift to the left, so there is a competition between drift and diffusion. Who wins?

We can solve this model exactly. Starting with the FPE

$$\partial_t P = \partial_x(\beta x P) + D\,\partial_x^2 P \quad , \tag{9.60}$$

we first Fourier transform

$$\hat{P}(k,t) = \int\limits_{-\infty}^{\infty} dx\; P(x,t)\, e^{-ikx} \quad . \tag{9.61}$$

Expressed in terms of independent variables $k$ and $t$, one finds that the FPE becomes

$$\partial_t \hat{P} + \beta k\, \partial_k \hat{P} = -Dk^2 \hat{P} \quad . \tag{9.62}$$

This is known as a *quasilinear partial differential equation*, and a general method of solution for such equations is the *method of characteristics*, which is briefly reviewed in appendix §9.9. A quasilinear PDE in $N$ independent variables can be transformed into $N+1$ coupled ODEs. Applying the method to Eqn. 9.62, one finds

$$\hat{P}(k,t) = \hat{P}\big(k\,e^{-\beta t}, t=0\big)\;\exp\left\{-\frac{D}{2\beta}\left(1-e^{-2\beta t}\right)k^2\right\} \quad . \tag{9.63}$$

Suppose $P(x,0) = \delta(x-x_0)$, in which case $\hat{P}(k,0) = e^{-ikx_0}$. We may now apply the inverse Fourier transform to obtain

$$P(x,t) = \sqrt{\frac{\beta}{2\pi D}\cdot\frac{1}{1-e^{-2\beta t}}}\;\exp\left\{-\frac{\beta}{2D}\,\frac{\left(x-x_0\,e^{-\beta t}\right)^2}{1-e^{-2\beta t}}\right\} \quad . \tag{9.64}$$

Taking the limit $t\to\infty$, we obtain the asymptotic distribution

$$P(x,t\to\infty) = \sqrt{\frac{\beta}{2\pi D}}\;e^{-\beta x^2/2D} \quad , \tag{9.65}$$

which is a Gaussian centered at $x=0$, with standard deviation $\sigma = \sqrt{D/\beta}$.

Physically, the drift term $F_1(x) = -\beta x$ arises when the particle is confined to a harmonic well. The equation of motion is then $\ddot{x} + \gamma\dot{x} + \omega_0^2 x = \eta$, which is discussed in the appendix, §9.7. If we average over the random forcing, then setting the acceleration to zero yields the local drift velocity $v_{\rm drift} = -\omega_0^2\,x/\gamma$, hence $\beta = \omega_0^2/\gamma$. Solving by Laplace transform, one has $L(z) = z^2 + \gamma z + \omega_0^2$, with roots $z_\pm = -\frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} - \omega_0^2}$, and

$$K(s) = \frac{e^{z_+ s} - e^{z_- s}}{z_+ - z_-}\,\Theta(s) \quad . \tag{9.66}$$

Note that $\mathrm{Re}\,(z_\pm) < 0$. Plugging this result into Eqn. 9.185 and integrating, we find

$$\lim_{t\to\infty} M(t,t) = \frac{\Gamma}{2\gamma\omega_0^2} \quad , \tag{9.67}$$

hence the asymptotic distribution is

$$P(x,t\to\infty) = \sqrt{\frac{\gamma\omega_0^2}{\pi\Gamma}}\;e^{-\gamma\omega_0^2 x^2/\Gamma} \quad . \tag{9.68}$$

Comparing with Eqn. 9.65, we once again find $D = \Gamma/2\gamma^2$. Does the Langevin particle in a harmonic well describe an Ornstein-Uhlenbeck process for finite $t$? It does in the limit $\gamma \to \infty$, $\omega_0 \to \infty$, $\Gamma \to \infty$, with $\beta = \omega_0^2/\gamma$ and $D = \Gamma/2\gamma^2$ finite. In this limit, one has $M(t,t) = \beta^{-1}D\left(1 - e^{-2\beta t}\right)$. For $\gamma < \infty$, the velocity relaxation time is finite, and on time scales shorter than $\gamma^{-1}$ the path $x(t)$ is not Markovian.

In the Ornstein-Uhlenbeck model, drift would like to collapse the distribution to a delta-function at $x = 0$, whereas diffusion would like to spread the distribution infinitely thinly over all space. In that sense, both terms represent extremist inclinations. Yet in the limit $t \to \infty$, drift and diffusion gracefully arrive at a grand compromise, with neither achieving its ultimate goal. The asymptotic distribution is centered about $x = 0$, but has a finite width. There is a lesson here for the United States Congress, if only they understood math.

### 9.4.4 Equilibrium distribution

If we set $\partial P/\partial t = 0$ in eqn. 9.51, we obtain an equation for the steady-state distribution $P_\infty(x)$,

$$\frac{d}{dx}\big[F_1(x)\,P_\infty(x)\big] = \frac{1}{2}\,\frac{d^2}{dx^2}\big[F_2(x)\,P_\infty(x)\big] \quad . \tag{9.69}$$

We assume that $P_\infty(\pm\infty) = 0$ in order to ensure normalizability. We may then integrate once, multiply and divide the LHS by $F_2(x)$, and thereby obtain

$$\frac{d}{dx}\log\big[F_2(x)\,P_\infty(x)\big] = \frac{2\,F_1(x)}{F_2(x)} \quad . \tag{9.70}$$

The solution is then

$$P_\infty(x) = \frac{\mathcal{N}}{F_2(x)}\,\exp\left\{\int_a^x dx'\,\frac{2\,F_1(x')}{F_2(x')}\right\} \quad , \tag{9.71}$$

where the constant $\mathcal{N}$ is chosen to guarantee normalization, *i.e.* $\int_{-\infty}^{\infty} dx\,P_\infty(x) \equiv 1$.

## 9.5 Formal Theory of Stochastic Processes

Here we follow the presentation in chapter 3 in the book by C. Gardiner. Given a time-dependent random variable $\boldsymbol{X}(t)$, we define the probability distribution

$$P(\boldsymbol{x},t) = \big\langle \delta\big(\boldsymbol{x} - \boldsymbol{X}(t)\big)\big\rangle \quad , \tag{9.72}$$

where the average is over different realizations of the random process. $P(\boldsymbol{x},t)$ is a density with units $L^{-d}$. This distribution is normalized according to $\int d\boldsymbol{x}\,P(\boldsymbol{x},t) = 1$, where $d\boldsymbol{x} = d^dx$ is the differential for the spatial volume, and does not involve time. If we integrate over some region $A$, we obtain

$$P_A(t) = \int_A d\boldsymbol{x}\,P(\boldsymbol{x},t) = \text{probability that } \boldsymbol{X}(t) \in A \quad . \tag{9.73}$$

We define the *joint probability distributions* as follows:

$$P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2\,;\,\ldots\,;\,\boldsymbol{x}_N, t_N) = \Big\langle \delta\big(\boldsymbol{x}_1 - \boldsymbol{X}(t_1)\big)\cdots\delta\big(\boldsymbol{x}_N - \boldsymbol{X}(t_N)\big)\Big\rangle \quad . \tag{9.74}$$

From the joint probabilities we may form *conditional probability distributions*

$$P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2\,;\,\ldots\,;\,\boldsymbol{x}_N, t_N\,|\,\boldsymbol{y}_1, \tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M, \tau_M) = \frac{P(\boldsymbol{x}_1, t_1\,;\,\ldots\,;\,\boldsymbol{x}_N, t_N\,;\,\boldsymbol{y}_1, \tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M, \tau_M)}{P(\boldsymbol{y}_1, \tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M, \tau_M)} \quad . \tag{9.75}$$

Although the times can be in any order, by convention we order them so they decrease from left to right:

$$t_1 > \cdots > t_N > \tau_1 > \cdots \tau_M \quad . \tag{9.76}$$

### 9.5.1 Markov processes

In a *Markov process*, any conditional probability is determined by its *most recent condition*. Thus,

$$P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2\,;\,\ldots\,;\,\boldsymbol{x}_N, t_N\,|\,\boldsymbol{y}_1, \tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M, \tau_M) = P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2\,;\,\ldots\,;\,\boldsymbol{x}_N, t_N\,|\,\boldsymbol{y}_1, \tau_1) \quad , \tag{9.77}$$

where the ordering of the times is as in Eqn. 9.76. This definition entails that *all* probabilities may be constructed from $P(\boldsymbol{x}, t)$ and from the conditional distribution $P(\boldsymbol{x}, t\,|\,\boldsymbol{y}, \tau)$. Clearly we must have that $P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2) = P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_2, t_2)\,P(\boldsymbol{x}_2, t_2)$. At the next level, we have

$$\begin{aligned} P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2\,;\,\boldsymbol{x}_3, t_3) &= P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_2, t_2\,;\,\boldsymbol{x}_3, t_3)\,P(\boldsymbol{x}_2, t_2\,;\,\boldsymbol{x}_3, t_3) \\ &= P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_2, t_2)\,P(\boldsymbol{x}_2, t_2\,|\,\boldsymbol{x}_3, t_3)\,P(\boldsymbol{x}_3, t_3) \quad . \end{aligned}$$

Proceeding thusly, we have

$$P(\boldsymbol{x}_1, t_1\,;\,\ldots\,;\,\boldsymbol{x}_N, t_N) = P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_2, t_2)\,P(\boldsymbol{x}_2, t_2\,|\,\boldsymbol{x}_3, t_3)\cdots P(\boldsymbol{x}_{N-1}, t_{N-1}\,|\,\boldsymbol{x}_N, t_N)\,P(\boldsymbol{x}_N, t_N) \quad , \tag{9.78}$$

so long as $t_1 > t_2 > \ldots > t_N$.

#### Chapman-Kolmogorov equation

The probability density $P(\boldsymbol{x}_1, t_1)$ can be obtained from the joint probability density $P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2)$ by integrating over $\boldsymbol{x}_2$:

$$P(\boldsymbol{x}_1, t_1) = \int\! d\boldsymbol{x}_2\,P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2) = \int\! d\boldsymbol{x}_2\,P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_2, t_2)\,P(\boldsymbol{x}_2, t_2) \quad . \tag{9.79}$$

Similarly[6],

$$P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_3, t_3) = \int\! d\boldsymbol{x}_2\,P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_2, t_2\,;\,\boldsymbol{x}_3, t_3)\,P(\boldsymbol{x}_2, t_2\,|\,\boldsymbol{x}_3, t_3) \quad . \tag{9.80}$$

For Markov processes, then,

$$P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_3, t_3) = \int\! d\boldsymbol{x}_2\,P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_2, t_2)\,P(\boldsymbol{x}_2, t_2\,|\,\boldsymbol{x}_3, t_3) \quad . \tag{9.81}$$

For discrete spaces, we have $\int d\boldsymbol{x} \to \sum_{\boldsymbol{x}}$, and $\sum_{\boldsymbol{x}_2} P(\boldsymbol{x}_1, t_1\,|\,\boldsymbol{x}_2, t_2)\,P(\boldsymbol{x}_2, t_2\,|\,\boldsymbol{x}_3, t_3)$ is a matrix multiplication.

[6]Because $P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2\,|\,\boldsymbol{x}_3, t_3) = \big[P(\boldsymbol{x}_1, t_1\,;\,\boldsymbol{x}_2, t_2\,;\,\boldsymbol{x}_3, t_3)/P(\boldsymbol{x}_2, t_2\,;\,\boldsymbol{x}_3, t_3)\big]\cdot\big[P(\boldsymbol{x}_2, t_2\,;\,\boldsymbol{x}_3, t_3)/P(\boldsymbol{x}_3, t_3)\big]$.

**Do Markov processes exist in nature and are they continuous?**

A random walk in which each step is independently and identically distributed is a Markov process. Consider now the following arrangement. You are given a bag of marbles, an initial fraction $p_0$ of which are red, $q_0$ of which are green, and $r_0$ of which are blue, with $p_0 + q_0 + r_0 = 1$. Let $\sigma_j = +1$, $0$, or $-1$ according to whether the $j^{\rm th}$ marble selected is red, green, or blue, respectively, and define $X_n = \sum_{j=1}^n \sigma_j$, which would correspond to the position of a random walker who steps either to the right ($\sigma_j = +1$), remain stationary ($\sigma_j = 0$), or steps left ($\sigma_j = -1$) during each discrete time interval. If the bag is infinite, then $\{X_1, X_2\,, \ldots\}$ is a Markov process. The probability for $\sigma_j = +1$ remains at $p = p_0$ and is unaffected by the withdrawal of any finite number of marbles from the bag. But if the contents of the bag are *finite*, then the probability $p$ changes with discrete time, and in such a way that cannot be determined from the instantaneous value of $X_n$ alone. Note that if there were only two colors of marbles, and $\sigma_j \in \{+1\,,\,-1\}$, then given $X_0 = 0$ and knowledge of the initial number of marbles in the bag, specifying $X_n$ tells us everything we need to know about the composition of the bag at time $n$. But with three possibilities $\sigma_j \in \{+1\,,\,0\,,\,-1\}$ we need to know the entire history in order to determine the current values of $p$, $q$, and $r$. The reason is that the sequences $0000$, $001\overline{1}$, $\overline{1}11\overline{1}$ (with $\overline{1} \equiv -1$) all have the same effect on the displacement $X$, but result in a different composition remaining in the bag.

In physical systems, processes we might model as random have a finite correlation time. We saw above that the correlator of the random force $\eta(t)$ in the Langevin equation is written $\big\langle \eta(t)\,\eta(t+s)\big\rangle = \phi(s)$, where $\phi(s)$ decays to zero on a time scale $\tau_\phi$. For time differences $|s| < \tau_\phi$, the system is not Markovian. In addition, the system itself may exhibit some memory. For example, in the Langevin equation $\dot u + \gamma u = \eta(t)$, there is a time scale $\gamma^{-1}$ over which the variable $u(t)$ forgets its previous history. Still, if $\tau_\phi = 0$, $u(t)$ is a Markov process, because the equation is first order and therefore only the most recent condition is necessary in order to integrate forward from some past time $t = t_0$ to construct the statistical ensemble of functions $u(t)$ for $t > t_0$. For second order equations, such as $\ddot x + \gamma \dot x = \eta(t)$, two initial conditions are required, hence diffusion paths $X(t)$ are only Markovian on time scales beyond $\gamma^{-1}$, over which the memory of the initial velocity is lost. More generally, if $\boldsymbol{\varphi}$ is an $N$-component vector in phase space, and

$$\frac{d\varphi_i}{dt} = A_i(\boldsymbol{\varphi}, t) + B_{ij}(\boldsymbol{\varphi}, t)\,\eta_j(t) \quad , \tag{9.82}$$

where we may choose $\big\langle \eta_i(t)\,\eta_j(t')\big\rangle = \delta_{ij}\,\delta(t-t')$, then the path $\boldsymbol{\varphi}(t)$ is a Markov process.

While a random variable $\boldsymbol{X}(t)$ may take values in a continuum, as a function of time it may still exhibit discontinuous jumps. That is to say, even though time $t$ may evolve continuously, the sample paths $\boldsymbol{X}(t)$ may be discontinuous. As an example, consider the Brownian motion of a particle moving in a gas or fluid. On the scale of the autocorrelation time, the velocity changes discontinuously, while the position $\boldsymbol{X}(t)$ evolves continuously (although not smoothly). The condition that sample paths $\boldsymbol{X}(t)$ evolve continuously is known as the *Lindeberg condition*,

$$\lim_{\tau\to 0}\,\frac{1}{\tau}\int\limits_{|\boldsymbol{x}-\boldsymbol{y}|>\varepsilon} d\boldsymbol{y}\;P(\boldsymbol{y}, t+\tau\,|\,\boldsymbol{x}, t) = 0 \quad , \tag{9.83}$$

for all $\varepsilon > 0$. If this condition is satisfied, then the sample paths $\boldsymbol{X}(t)$ are continuous with probability one. Two examples:

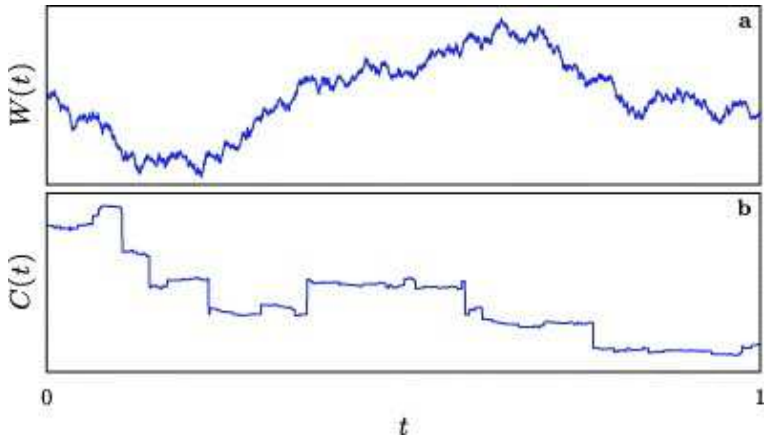



Figure 9.3: (a) Wiener process sample path $W(t)$. (b) Cauchy process sample path $C(t)$. From K. Jacobs and D. A. Steck, *New J. Phys.* **13**, 013016 (2011).

(1) **Wiener process**: As we shall discuss below, this is a pure diffusion process with no drift or jumps, with

$$P(x,t\,|\,x',t') = \frac{1}{\sqrt{4\pi D|t-t'|}}\,\exp\left(-\frac{(x-x')^2}{4D|t-t'|}\right) \tag{9.84}$$

in one space dimension. The Lindeberg condition is satisfied, and the sample paths $\boldsymbol{X}(t)$ are continuous.

(2) **Cauchy process**: This is a process in which sample paths exhibit finite jumps, and hence are not continuous. In one space dimension,

$$P(x,t\,|\,x',t') = \frac{c|t-t'|}{\pi\big[(x-x')^2+c^2(t-t')^2\big]} \quad , \tag{9.85}$$

where $c$ has units of velocity. Note that in both this case and the Wiener process described above, we have $\lim_{t-t'\to 0} P(xt\,|\,x't') = \delta(x-x')$. However in this example the Lindeberg condition is not satisfied.

To simulate, given $\boldsymbol{x}_n = \boldsymbol{X}(t=n\tau)$, choose $\boldsymbol{y} \in \mathcal{D}_b(\boldsymbol{x}_n)$, where $\mathcal{D}_b(\boldsymbol{x}_n)$ is a ball of radius $b > \varepsilon$ centered at $\boldsymbol{x}_n$. Then evaluate the probability $p \equiv P(\boldsymbol{y},(n+1)\tau\,|\,\boldsymbol{x},n\tau)/M(\boldsymbol{y})$ where $M(\boldsymbol{y})$ is a proposal density such that $p \le 1$.. If $p$ exceeds a random number drawn from a uniform distribution on $[0,1]$, accept and set $\boldsymbol{x}_{n+1} = \boldsymbol{X}\big((n+1)\tau\big) = \boldsymbol{y}$. Else reject and choose a new $\boldsymbol{y}$ and proceed as before.

### 9.5.2 Martingales

A *Martingale* is a stochastic process for which the *conditional average* of the random variable $\boldsymbol{X}(t)$ does not change from its most recent condition. That is,

$$\big\langle \boldsymbol{x}(t)\,\big|\,\{\boldsymbol{y}_1\,\tau_1\,;\,\boldsymbol{y}_2,\tau_2\,;\,\ldots\,;\,\boldsymbol{y}_M,\tau_M\}\big\rangle = \int d\boldsymbol{x}\,P(\boldsymbol{x},t\,|\,\boldsymbol{y}_1,\tau_1\,;\,\ldots\,;\,\boldsymbol{y}_M,\tau_M)\,\boldsymbol{x}=\boldsymbol{y}_1 \quad . \tag{9.86}$$

In this sense, a Martingale is a stochastic process which represents a 'fair game'. Not every Martingale is a Markov process, and not every Markov process is a Martingale. The Wiener process is a Martingale.

We now come to a very important fact about Martingales, which we will here derive in $d = 1$ dimension. For $t_1 > t_2$,

$$\begin{aligned}\big\langle x(t_1)\,x(t_2)\big\rangle &= \int\!dx_1 \int\!dx_2\; P(x_1,t_1\,;\,x_2,t_2)\,x\,x_2 = \int\!dx_1 \int\!dx_2\; P(x_1,t_1\,|\,x_2,t_2)\,P(x_2,t_2)\,x_1\,x_2 \\ &= \int\!dx_2\; P(x_2,t_2)\,x_2 \int\!dx_1\; P(x_1,t_1\,|\,x_2,t_2)\,x_1 = \int\!dx_2\; P(x_2,t_2)\,x_2^2 = \big\langle x^2(t_2)\big\rangle \quad .\end{aligned} \tag{9.87}$$

One can further show that, for $t_1 > t_2 > t_3$ ,

$$\Big\langle \big[x(t_1) - x(t_2)\big]\big[x(t_2) - x(t_3)\big]\Big\rangle = 0 \quad , \tag{9.88}$$

which says that *at the level of pair correlations*, past performance provides no prediction of future results.

### 9.5.3 Differential Chapman-Kolmogorov equations

Suppose the following conditions apply:

$$|\boldsymbol{y}-\boldsymbol{x}| > \varepsilon \quad \Longrightarrow \quad \lim_{\tau\to 0}\frac{1}{\tau}\,P(\boldsymbol{y},t+\tau\,|\,\boldsymbol{x},t) = W(\boldsymbol{y}\,|\,\boldsymbol{x},t) \tag{9.89}$$

$$\lim_{\tau\to 0}\frac{1}{\tau}\int\limits_{|\boldsymbol{y}-\boldsymbol{x}|<\varepsilon}\!\!d\boldsymbol{y}\;(y_\mu - x_\mu)\,P(\boldsymbol{y},t+\tau\,|\,\boldsymbol{x},t) = A_\mu(\boldsymbol{x},t) + \mathcal{O}(\varepsilon) \tag{9.90}$$

$$\lim_{\tau\to 0}\frac{1}{\tau}\int\limits_{|\boldsymbol{y}-\boldsymbol{x}|<\varepsilon}\!\!d\boldsymbol{y}\;(y_\mu - x_\mu)\,(y_\nu - x_\nu)\,P(\boldsymbol{y},t+\tau\,|\,\boldsymbol{x},t) = B_{\mu\nu}(\boldsymbol{x},t) + \mathcal{O}(\varepsilon) \quad , \tag{9.91}$$

where the last two conditions hold uniformly in $\boldsymbol{x}$, $t$, and $\varepsilon$. Then following §3.4.1 and §3.6 of Gardiner, one obtains the *forward differential Chapman-Kolmogorov equation* (DCK+),

$$\begin{aligned}\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t} &= -\sum_\mu \frac{\partial}{\partial x_\mu}\Big[A_\mu(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big] + \frac{1}{2}\sum_{\mu,\nu}\frac{\partial^2}{\partial x_\mu\,\partial x_\nu}\Big[B_{\mu\nu}(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big] \\ &\qquad + \int\!d\boldsymbol{y}\,\Big[W(\boldsymbol{x}\,|\,\boldsymbol{y},t)\,P(\boldsymbol{y},t\,|\,\boldsymbol{x}',t') - W(\boldsymbol{y}\,|\,\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big] \quad ,\end{aligned} \tag{9.92}$$

and the *backward differential Chapman-Kolmogorov equation* (DCK−),

$$\begin{aligned}\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t'} &= -\sum_\mu A_\mu(\boldsymbol{x}',t')\,\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial x'_\mu} - \frac{1}{2}\sum_{\mu,\nu} B_{\mu\nu}(\boldsymbol{x}',t')\,\frac{\partial^2 P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial x'_\mu\,\partial x'_\nu} \\ &\qquad + \int\!d\boldsymbol{y}\,W(\boldsymbol{y}\,|\,\boldsymbol{x}',t')\Big[P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t') - P(\boldsymbol{x},t\,|\,\boldsymbol{y},t')\Big] \quad .\end{aligned} \tag{9.93}$$

Note that the Lindeberg condition requires that

$$\lim_{\tau\to 0}\frac{1}{\tau}\int\limits_{|\boldsymbol{x}-\boldsymbol{y}|>\varepsilon}\!\!d\boldsymbol{y}\;P(\boldsymbol{y},t+\tau\,|\,\boldsymbol{x},t) = \int\limits_{|\boldsymbol{x}-\boldsymbol{y}|>\varepsilon}\!\!d\boldsymbol{y}\;W(\boldsymbol{y}\,|\,\boldsymbol{x},t) = 0 \quad , \tag{9.94}$$

which must hold for any $\varepsilon > 0$. Taking the limit $\varepsilon \to 0$, we conclude[7] $W(\boldsymbol{y} \,|\, \boldsymbol{x}, t) = 0$ *if* the Lindeberg condition is satisfied. If there are any jump processes, *i.e.* if $W(\boldsymbol{y} \,|\, \boldsymbol{x}, t)$ does not identically vanish for all values of its arguments, then Lindeberg is violated, and the paths are discontinuous.

Some applications:

(1) **Master equation**: If $A_\mu(\boldsymbol{x}, t) = 0$ and $B_{\mu\nu}(\boldsymbol{x}, t) = 0$, then we have from DCK+,

$$\frac{\partial P(\boldsymbol{x}, t \,|\, \boldsymbol{x}', t')}{\partial t} = \int d\boldsymbol{y} \, \Big[ W(\boldsymbol{x} \,|\, \boldsymbol{y}, t) \, P(\boldsymbol{y}, t \,|\, \boldsymbol{x}', t') - W(\boldsymbol{y} \,|\, \boldsymbol{x}, t) \, P(\boldsymbol{x}, t \,|\, \boldsymbol{x}', t') \Big] \quad . \tag{9.95}$$

Let's integrate this equation over a time interval $\Delta t$. Assuming $P(\boldsymbol{x}, t \,|\, \boldsymbol{x}', t) = \delta(\boldsymbol{x} - \boldsymbol{x}')$, we have

$$P(\boldsymbol{x}, t + \Delta t \,|\, \boldsymbol{x}', t) = \Big[ 1 - \Delta t \int d\boldsymbol{y} \, W(\boldsymbol{y} \,|\, \boldsymbol{x}', t) \Big] \delta(\boldsymbol{x} - \boldsymbol{x}') + W(\boldsymbol{x} \,|\, \boldsymbol{x}', t) \, \Delta t \quad . \tag{9.96}$$

Thus,

$$Q(\boldsymbol{x}', t + \Delta t, t) = 1 - \Delta t \int d\boldsymbol{y} \, W(\boldsymbol{y} \,|\, \boldsymbol{x}', t) \tag{9.97}$$

is the probability for a particle to remain at $\boldsymbol{x}'$ over the interval $\big[t, t + \Delta t\big]$ given that it was at $\boldsymbol{x}'$ at time $t$. Iterating this relation, we find

$$\begin{aligned} Q(\boldsymbol{x}, t, t_0) &= \Big( 1 - \Lambda(\boldsymbol{x}, t - \Delta t) \, \Delta t \Big) \Big( 1 - \Lambda(\boldsymbol{x}, t - 2\Delta t) \, \Delta t \Big) \cdots \Big( 1 - \Lambda(\boldsymbol{x}, t_0) \, \Delta t \Big) \overbrace{Q(\boldsymbol{x}, t_0, t_0)}^{1} \\ &= \mathcal{P} \exp \left\{ - \int_{t_0}^{t} dt' \, \Lambda(\boldsymbol{x}, t') \right\} \quad , \end{aligned} \tag{9.98}$$

where $\Lambda(\boldsymbol{x}, t) = \int d\boldsymbol{y} \, W(\boldsymbol{y} \,|\, \boldsymbol{x}, t)$ and $\mathcal{P}$ is the path ordering operator which places earlier times to the right.

The interpretation of the function $W(\boldsymbol{y} \,|\, \boldsymbol{x}, t)$ is that it is the probability density *rate* for the random variable $\boldsymbol{X}$ to jump from $\boldsymbol{x}$ to $\boldsymbol{y}$ at time $\boldsymbol{t}$. Thus, the dimensions of $W(\boldsymbol{y} \,|\, \boldsymbol{x}, t)$ are $L^{-d} \, T^{-1}$. Such processes are called *jump processes*. For discrete state spaces, the master equation takes the form

$$\frac{\partial P(\boldsymbol{n}, t \,|\, \boldsymbol{n}', t')}{\partial t} = \sum_{\boldsymbol{m}} \Big[ W(\boldsymbol{n} \,|\, \boldsymbol{m}, t) \, P(\boldsymbol{m}, t \,|\, \boldsymbol{n}', t') - W(\boldsymbol{m} \,|\, \boldsymbol{n}, t) \, P(\boldsymbol{n}, t \,|\, \boldsymbol{n}', t') \Big] \quad . \tag{9.99}$$

Here $W(\boldsymbol{n} \,|\, \boldsymbol{m}, t)$ has units $T^{-1}$, and corresponds to the rate of transitions from state $\boldsymbol{m}$ to state $\boldsymbol{n}$ at time $t$.

[7] What about the case $\boldsymbol{y} = \boldsymbol{x}$, which occurs for $\varepsilon = 0$, which is never actually reached throughout the limiting procedure? The quantity $W(\boldsymbol{x} \,|\, \boldsymbol{x}, t)$ corresponds to the rate at which the system jumps from $\boldsymbol{x}$ to $\boldsymbol{x}$ at time $t$, which is not a jump process at all. Note that the contribution from $\boldsymbol{y} = \boldsymbol{x}$ cancels from the DCK$\pm$ equations. In other words, we can set $W(\boldsymbol{x} \,|\, \boldsymbol{x}, t) \equiv 0$.

(2) **Fokker-Planck equation**: If $W(\boldsymbol{x}\,|\,\boldsymbol{y},t)=0$, DCK+ gives

$$\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t}=-\sum_\mu\frac{\partial}{\partial x_\mu}\Big[A_\mu(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big]+\tfrac{1}{2}\sum_{\mu,\nu}\frac{\partial^2}{\partial x_\mu\,\partial x_\nu}\Big[B_{\mu\nu}(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big]\quad,\tag{9.100}$$

which is a more general form of the Fokker-Planck equation we studied in §9.4 above. Defining the average $\big\langle F(\boldsymbol{x},t)\big\rangle=\int d^dx\,F(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')$, via integration by parts we derive

$$\begin{aligned}\frac{d}{dt}\big\langle x_\mu\big\rangle&=\big\langle A_\mu\big\rangle\\ \frac{d}{dt}\big\langle x_\mu\,x_\nu\big\rangle&=\big\langle x_\mu\,A_\nu\big\rangle+\big\langle A_\mu\,x_\nu\big\rangle+\tfrac{1}{2}\big\langle B_{\mu\nu}+B_{\nu\mu}\big\rangle\quad.\end{aligned}\tag{9.101}$$

For the case where $A_\mu(\boldsymbol{x},t)$ and $B_{\mu\nu}(\boldsymbol{x},t)$ are constants independent of $\boldsymbol{x}$ and $t$, we have the solution

$$P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')=\det{}^{-1/2}\big[2\pi B\,\Delta t\big]\exp\left\{-\frac{1}{2\,\Delta t}\,\big(\Delta x_\mu-A_\mu\,\Delta t\big)\,B^{-1}_{\mu\nu}\,\big(\Delta x_\nu-A_\nu\,\Delta t\big)\right\}\quad,\tag{9.102}$$

where $\Delta\boldsymbol{x}\equiv\boldsymbol{x}-\boldsymbol{x}'$ and $\Delta t\equiv t-t'$. This is normalized so that the integral over $\boldsymbol{x}$ is unity. If we subtract out the *drift* $\boldsymbol{A}\,\Delta t$, then clearly

$$\Big\langle\big(\Delta x_\nu-A_\nu\,\Delta t\big)\,\big(\Delta x_\mu-A_\mu\,\Delta t\big)\Big\rangle=B_{\mu\nu}\,\Delta t\quad,\tag{9.103}$$

which is diffusive.

(3) **Liouville equation**: If $W(\boldsymbol{x}\,|\,\boldsymbol{y},t)=0$ and $B_{\mu\nu}(\boldsymbol{x},t)=0$, then DCK+ gives

$$\frac{\partial P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')}{\partial t}=-\sum_\mu\frac{\partial}{\partial x_\mu}\Big[A_\mu(\boldsymbol{x},t)\,P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')\Big]\quad.\tag{9.104}$$

This is Liouville's equation from classical mechanics, also known as the continuity equation. Suppressing the $(\boldsymbol{x}',t')$ variables, the above equation is equivalent to

$$\frac{\partial\varrho}{\partial t}+\boldsymbol{\nabla}\cdot(\varrho\,\boldsymbol{v})=0\quad,\tag{9.105}$$

where $\varrho(\boldsymbol{x},t)=P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')$ and $\boldsymbol{v}(\boldsymbol{x},t)=\boldsymbol{A}(\boldsymbol{x},t)$. The product of $\boldsymbol{A}$ and $P$ is the current $\boldsymbol{j}=\varrho\boldsymbol{v}$. To find the general solution, we assume the initial conditions are $P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t)=\delta(\boldsymbol{x}-\boldsymbol{x}')$. Then if $\boldsymbol{x}(t;\boldsymbol{x}')$ is the solution to the ODE

$$\frac{d\boldsymbol{x}(t)}{dt}=\boldsymbol{A}\big(\boldsymbol{x}(t),t\big)\tag{9.106}$$

with boundary condition $\boldsymbol{x}(t')=\boldsymbol{x}'$, then by applying the chain rule, we see that

$$P(\boldsymbol{x},t\,|\,\boldsymbol{x}',t')=\delta\big(\boldsymbol{x}-\boldsymbol{x}(t;\boldsymbol{x}')\big)\tag{9.107}$$

solves the Liouville equation. Thus, the probability density remains a $\delta$-function for all time.

### 9.5.4 Stationary Markov processes and ergodic properties

Stationary Markov processes satisfy a time translation invariance:

$$P(\boldsymbol{x}_1, t_1\,;\,\ldots\,;\,\boldsymbol{x}_N, t_N) = P(\boldsymbol{x}_1, t_1+\tau\,;\,\ldots\,;\,\boldsymbol{x}_N, t_N+\tau) \quad . \tag{9.108}$$

This entails $P(\boldsymbol{x}, t) = P(\boldsymbol{x})$ and $P(\boldsymbol{x}_1, t_1 \,|\, \boldsymbol{x}_2, t_2) = P(\boldsymbol{x}_1, t_1 - t_2 \,|\, \boldsymbol{x}_2, 0)$. Consider the case of one space dimension and define the time average

$$\overline{X}_T \equiv \frac{1}{T}\int\limits_{T/2}^{T/2} dt\; x(t) \quad . \tag{9.109}$$

We use a bar to denote time averages and angular brackets $\langle\cdots\rangle$ to denote averages over the randomness. Thus, $\langle \overline{X}_T \rangle = \langle x \rangle$, which is time-independent for a stationary Markov process. The variance of $\overline{X}_T$ is

$$\mathsf{Var}\big(\overline{X}_T\big) = \frac{1}{T^2}\int\limits_{T/2}^{T/2} dt \int\limits_{T/2}^{T/2} dt'\; \big\langle x(t)\, x(t')\big\rangle_{\rm c} \quad , \tag{9.110}$$

where the connected average is $\langle AB\rangle_{\rm c} = \langle AB\rangle - \langle A\rangle\langle B\rangle$. We define the correlation function

$$C(t_1 - t_2) \equiv \big\langle x(t_1)\, x(t_2)\big\rangle = \int\limits_{-\infty}^{\infty} dx_1 \int\limits_{-\infty}^{\infty} dx_2\; x_1\, x_2\; P(x_1, t_1\,;\, x_2, t_2) \tag{9.111}$$

and the connected correlator

$$C_{\rm c}(t_1 - t_2) \equiv \big\langle x(t)\, x(t')\big\rangle_{\rm c} = C(t_1 - t_2) - \langle x(t_1)\rangle\langle x(t_2)\rangle \quad . \tag{9.112}$$

If $C_{\rm c}(\tau)$ decays to zero sufficiently rapidly with $\tau$, for example as an exponential $e^{-\gamma\tau}$, then $\mathsf{Var}\big(\overline{X}_T\big) \to 0$ as $T \to \infty$, which means that $X_{T\to\infty} = \langle x\rangle$. Thus the time average is the ensemble average, which means the process is *ergodic*. Typically one assumes $\langle x(t)\rangle = 0$ in which case $C_{\rm c}(\tau) = C(\tau)$.

#### Wiener-Khinchin theorem

Define the quantity

$$\hat{x}_T(\omega) = \int\limits_{-T/2}^{T/2} dt\; x(t)\, e^{i\omega t} \quad . \tag{9.113}$$

The *spectral function* $S_T(\omega)$ is given by

$$S_T(\omega) = \left\langle \frac{1}{T}\,\big|\hat{x}_T(\omega)\big|^2 \right\rangle \quad . \tag{9.114}$$

We are interested in the limit $T \to \infty$. Does $S(\omega) \equiv S_{T\to\infty}(\omega)$ exist?

Observe that

$$\Big\langle \big|\hat{x}_T(\omega)\big|^2 \Big\rangle = \int\limits_{-T/2}^{T/2} dt_1 \int\limits_{-T/2}^{T/2} dt_2 \; e^{i\omega(t_2-t_1)} \; \overbrace{\big\langle x(t_1)\, x(t_2)\big\rangle}^{C(t_1-t_2)} = \int\limits_{-T}^{T} d\tau \, e^{-i\omega\tau} \, C(\tau) \, \big(T-|\tau|\big) \quad . \tag{9.115}$$

Thus,

$$S(\omega) = \lim_{T\to\infty} \int\limits_{-\infty}^{\infty} d\tau \, e^{i\omega\tau} \, C(\tau) \left(1 - \frac{|\tau|}{T}\right) \Theta\big(T-|\tau|\big) = \int\limits_{-\infty}^{\infty} d\tau \, e^{i\omega\tau} \, C(\tau) \quad . \tag{9.116}$$

The second equality above follows from Lebesgue's dominated convergence theorem, which you can look up on Wikipedia[8]. We therefore conclude the limit exists and is given by the Fourier transform of the correlation function $C(\tau) = \big\langle x(t)\, x(t+\tau)\big\rangle$.

Another way to derive this WK theorem is to write $C(\tau) = \int\limits_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \, \hat{C}(\omega') \, e^{-i\omega'\tau}$, in which case

$$\begin{aligned} \Big\langle \big|\hat{x}_T(\omega)\big|^2 \Big\rangle &= \int\limits_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \int\limits_{-T/2}^{T/2} dt_1 \int\limits_{-T/2}^{T/2} dt_2 \; e^{i(\omega-\omega')(t_2-t_1)} \, \hat{C}(\omega') \\ S_T(\omega) = \Big\langle \frac{1}{T} \big|\hat{x}_T(\omega)\big|^2 \Big\rangle &= \int\limits_{-\infty}^{\infty} \frac{d\omega'}{2\pi T} \, \hat{C}(\omega') \left( \frac{2\sin\big((\omega-\omega')\,T/2\big)}{(\omega-\omega')} \right)^2 \quad . \end{aligned} \tag{9.117}$$

But since

$$\lim_{T\to\infty} \frac{2\sin^2\big((\omega-\omega')\,T/2\big)}{\pi T\,(\omega-\omega')^2} = \delta(\omega-\omega') \quad , \tag{9.118}$$

we have $S(\omega) = \lim_{T\to\infty} S_T(\omega) = \hat{C}(\omega)$.

### 9.5.5 Approach to stationary solution

We have seen, for example, how in general an arbitrary initial state of the master equation will converge exponentially to an equilibrium distribution. For stationary Markov processes, the conditional distribution $P(\boldsymbol{x}, t \,|\, \boldsymbol{x}', t')$ converges to an equilibrium distribution $P_{\rm eq}(\boldsymbol{x})$ as $t - t' \to \infty$. How can we understand this convergence in terms of the differential Chapman-Kolmogorov equation? We summarize here the results in §3.7.3 of Gardiner.

Suppose $P_1(\boldsymbol{x}, t)$ and $P_2(\boldsymbol{x}, t)$ are each solutions to the DCK+ equation, and furthermore that $W(\boldsymbol{x} \,|\, \boldsymbol{x}', t)$, $A_\mu(\boldsymbol{x}, t)$, and $B_{\mu\nu}(\boldsymbol{x}, t)$ are all independent of $t$. Define the *Lyapunov functional*

$$K[P_1, P_2, t] = \int d\boldsymbol{x} \, \Big( P_1 \log(P_1/P_2) + P_2 - P_1 \Big) \quad . \tag{9.119}$$

[8] If we define the one parameter family of functions $C_T(\tau) = C(\tau)\left(1 - \frac{|\tau|}{T}\right) \Theta(T - |\tau|)$, then as $T \to \infty$ the function $C_T(\tau)\, e^{i\omega\tau}$ converges pointwise to $C(\tau)\, e^{i\omega\tau}$, and if $|C(\tau)|$ is integrable on $\mathbb{R}$, the theorem guarantees the second equality in Eqn. 9.116.

Since $P_{1,2}(\boldsymbol{x},t)$ are both normalized, the integrals of the last two terms inside the big round brackets cancel. Nevertheless, it is helpful to express $K$ in this way since, factoring out $P_1$ from the terms inside the brackets, we may use $f(z) = z - \log z - 1 \geq 0$ for $z \in \mathbb{R}_+$, where $z = P_2/P_1$. Thus, $K \geq 0$, and the minimum value is obtained for $P_1(x,t) = P_2(x,t)$.

Next, evaluate the time derivative $\dot{K}$:

$$\frac{dK}{dt} = \int d\boldsymbol{x} \left\{ \frac{\partial P_1}{\partial t} \cdot \Big[ \log P_1 - \log P_2 + 1 \Big] - \frac{\partial P_2}{\partial t} \cdot \frac{P_1}{P_2} \right\} \quad . \tag{9.120}$$

We now use DCK+ to obtain $\partial_t P_{1,2}$ and evaluate the contributions due to drift, diffusion, and jump processes. One finds

$$\left( \frac{dK}{dt} \right)_{\rm drift} = -\sum_\mu \int d\boldsymbol{x} \, \frac{\partial}{\partial x_\mu} \Big[ A_\mu \, P_1 \, \log \left( P_1/P_2 \right) \Big] \tag{9.121}$$

$$\left( \frac{dK}{dt} \right)_{\rm diff} = -\frac{1}{2} \sum_{\mu,\nu} \int d\boldsymbol{x} \, B_{\mu\nu} \, \frac{\partial \log(P_1/P_2)}{\partial x_\mu} \, \frac{\partial \log(P_1/P_2)}{\partial x_\nu} + \frac{1}{2} \int d\boldsymbol{x} \, \frac{\partial^2}{\partial x_\mu \, \partial x_\nu} \Big[ B_{\mu\nu} \, P_1 \, \log(P_1/P_2) \Big] \tag{9.122}$$

$$\left( \frac{dK}{dt} \right)_{\rm jump} = \int d\boldsymbol{x} \int d\boldsymbol{x}' \, W(\boldsymbol{x} \,|\, \boldsymbol{x}') \, P_2(\boldsymbol{x}', t) \Big[ \phi' \log(\phi/\phi') - \phi + \phi' \Big] \quad , \tag{9.123}$$

where $\phi(\boldsymbol{x},t) \equiv P_1(\boldsymbol{x},t)/P_2(\boldsymbol{x},t)$ in the last line. Dropping the total derivative terms, which we may set to zero at spatial infinity, we see that $\dot{K}_{\rm drift} = 0$, $\dot{K}_{\rm diff} \leq 0$, and $\dot{K}_{\rm jump} \leq 0$. Barring pathological cases[9], one has that $K(t)$ is a nonnegative decreasing function. Since $K = 0$ when $P_1(\boldsymbol{x},t) = P_2(\boldsymbol{x},t) = P_{\rm eq}(\boldsymbol{x})$, we see that the Lyapunov analysis confirms that $K$ is nonincreasing. For cases where $K$ is decreasing to zero, if we set $P_2(\boldsymbol{x},t) = P_{\rm eq}(\boldsymbol{x})$, we conclude that $P_1(\boldsymbol{x},t)$ converges to $P_{\rm eq}(\boldsymbol{x})$ as $t \to \infty$.

One can easily construct examples where $K$ is constant. For example, consider the deterministic drift equation on the circle $\theta \in [0, 2\pi)$,

$$\frac{\partial P}{\partial t} = -\omega \frac{\partial P}{\partial \theta} \quad , \tag{9.124}$$

which is a Fokker-Planck equation with constant $F_1(\theta) = \omega$ and $F_2(\theta) = 0$. It is then easy to see that

$$P(\theta,t) = P(\theta - \omega t, 0) \equiv P_0(\theta - \omega t) \quad . \tag{9.125}$$

Let us ccmpute $K[P, P_{\rm uni}, t]$ where $P_{\rm uni}(\theta) = (2\pi)^{-1}$ is the uniform distribution. Then

$$K[P, P_{\rm uni}, t] = \int_0^{2\pi} d\theta \, P_0(\theta - \omega t) \log \left( \frac{P_0(\theta - \omega t)}{P_{\rm uni}(\theta)} \right) = \int_0^{2\pi} d\theta \, P_0(\theta) \log \left( \frac{P_0(\theta)}{P_{\rm uni}(\theta)} \right) = K[P, P_{\rm uni}, 0] \quad . \tag{9.126}$$

More interesting examples may be concocted. For example the dynamics may respect a conservation law, in which case two distributions with different values of their conserved 'charges' can never converge to the same distribution.

[9] See Gardiner, §3.7.3.

## 9.6 Appendix I: Nonlinear diffusion

### 9.6.1 PDEs with infinite propagation speed

Starting from an initial probability density $P(\boldsymbol{x}, t=0) = \delta(\boldsymbol{x})$, we saw how Fickian diffusion, described by the equation $\partial_t P = \boldsymbol{\nabla}\cdot(D\,\boldsymbol{\nabla} P)$, gives rise to the solution

$$P(\boldsymbol{x}, t) = (4\pi D t)^{-d/2}\, e^{-\boldsymbol{x}^2/4Dt} \quad , \tag{9.127}$$

for all $t > 0$, assuming $D$ is a constant. As remarked in §9.2.1, this violates any physical limits on the speed of particle propagation, including that set by special relativity, because $P(\boldsymbol{x}, t) > 0$ for all $\boldsymbol{x}$ at any finite value of $t$.

It's perhaps good to step back at this point and recall the solution to the one-dimensional discrete random walk, where after each time increment the walker moves to the right ($\Delta X = 1$) with probability $p$ and to the left ($\Delta X = -1$) with probability $1-p$. To make things even simpler we'll consider the case with no drift, *i.e.* $p = \frac{1}{2}$. The distribution for $X$ after $N$ time steps is of the binomial form:

$$P_N(X) = 2^{-N} \binom{N}{\frac{1}{2}(N-X)} \quad . \tag{9.128}$$

Invoking Stirling's asymptotic result $\log K! = K\log K - K + \mathcal{O}(\log K)$ for $K \gg 1$, one has[10]

$$P_N(X) \simeq \sqrt{\frac{2}{\pi N}}\, e^{-X^2/2N} \quad . \tag{9.129}$$

We note that the distribution in Eqn. 9.128 is cut off at $|X| = N$, so that $P_N(X) = 0$ for $|X| > N$. This reflects the fact that the walker travels at a fixed speed of one step per time interval. This feature is lost in Eqn. 9.129, because the approximation which led to this result is not valid in the tails of the distribution. One might wonder about the results of §9.3 in this context, since we ultimately obtained a diffusion form for $P(x,t)$ using an exact functional averaging method. However, since we assumed a Gaussian probability functional for the random forcing $\eta(t)$, there is a finite probability for arbitrarily large values of the forcing. For example, consider the distribution of the integrated force $\phi = \int_{t_1}^{t_2} dt\, \eta(t)$:

$$P(\phi, \Delta t) = \left\langle \delta\left(\phi - \int_{t_1}^{t_2} dt\, \eta(t)\right)\right\rangle = \frac{1}{\sqrt{2\pi\Gamma\,\Delta t}}\, e^{-\phi^2/2\Gamma\Delta t} \quad , \tag{9.130}$$

where $\Delta t = t_2 - t_1$. This distribution is nonzero for arbitrarily large values of $\phi$.

Mathematically, the diffusion equation is an example of what is known as a *parabolic* partial differential equation. The Navier-Stokes equations of hydrodynamics are also parabolic PDEs. The other two classes are called *elliptic* and *hyperbolic*. Paradigmatic examples of these classes include the Laplace and

[10] The prefactor in this equation seems to be twice the expected $(2\pi N)^{-1/2}$, but since each step results in $\Delta X = \pm 1$, if we start from $X_0 = 0$ then after $N$ steps $X$ will be even if $N$ is even and odd if $N$ is odd. Therefore the continuum limit for the normalization condition on $P_N(X)$ is $\sum_X P_N(X) \approx \frac{1}{2}\int_{-\infty}^{\infty} dX\, P_N(X) = 1$.

Helmholtz equations (elliptic) and the wave equation and Maxwell's equations (hyperbolic). Hyperbolic equations propagate information at finite propagation speed. For second order PDEs of the form

$$A_{ij}\,\frac{\partial^2\Psi}{\partial x_i\,\partial x_j} + B_i\,\frac{\partial\Psi}{\partial x_i} + C\Psi = S \quad , \tag{9.131}$$

the PDE is elliptic if the matrix $A$ is positive definite or negative definite, parabolic if $A$ has one zero eigenvalue, and hyperbolic if $A$ is nondegenerate and indefinite (*e.g.* one positive and one negative eigenvalue, in the case of two independent variables). Accordingly, one way to remedy the unphysical propagation speed in the diffusion equation is to deform it to a hyperbolic PDE such as the *telegrapher's equation*,

$$\tau\,\frac{\partial^2\Psi}{\partial t^2} + \frac{\partial\Psi}{\partial t} + \gamma\Psi = D\,\frac{\partial^2\Psi}{\partial x^2} \quad . \tag{9.132}$$

When $\gamma = 0$, the solution for the initial condition $\Psi(x,0) = 0$ and $\partial_t\Psi(x,0) = \tau^{-1}\delta(x)$ is

$$\Psi(x,t) = \frac{1}{\sqrt{4D\tau}}\,e^{-t/2\tau}\,I_0\left(\sqrt{\left(\frac{t}{2\tau}\right)^2 - \frac{x^2}{4D\tau}}\right)\Theta\Big(\sqrt{D/\tau}\,t - |x|\Big) \quad . \tag{9.133}$$

Note that $\Psi(x,t)$ vanishes for $|x| > ct$, where $c = \sqrt{D/\tau}$ is the maximum propagation speed. One can check that in the limit $\tau \to 0$ one recovers the familiar diffusion kernel.

When there are more than two independent variables, the classification depends in more detail on the spectrum of the symmetric principal matrix $A$ in eqn. 9.131 and is as follows:

- elliptic: all eigenvalues of the symmetric principal matrix have the same sign;
- parabolic: the matrix is singular, with the nonzero eigenvalues having the same sign in the standard simple case;
- hyperbolic: one eigenvalue has the opposite sign from all the others;
- more general mixed signatures are sometimes called *ultrahyperbolic.*

**The telegrapher's equation**

To derive the telegrapher's equation, consider the section of a transmission line shown in Fig. 9.4. Let $V(x,t)$ be the electrical potential on the top line, with $V = 0$ on the bottom (*i.e.* ground). Per unit length $a$, the potential drop along the top line is $\Delta V = a\,\partial_x V = -IR - L\,\partial_t I$, and the current drop is $\Delta I = a\,\partial_x I = -GV - C\,\partial_t V$. Differentiating the first equation with respect to $x$ and using the second for $\partial_x I$, one arrives at Eqn. 9.132 with $\tau = LC/(RC+GL)$, $\gamma = RG/(RC+GL)$, and $D = a^2/(RC+GL)$.

### 9.6.2 The porous medium and $p$-Laplacian equations

Another way to remedy this problem with the diffusion equation is to consider some nonlinear extensions thereof[11]. Two such examples have been popular in the mathematical literature, the *porous medium*

[11] See J. L. Vazquez, *The Porous Medium Equation* (Oxford, 2006).

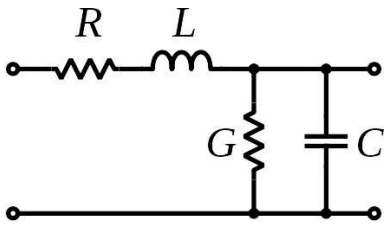



Figure 9.4: Repeating unit of a transmission line. *Credit: Wikipedia*

*equation* (PME),

$$\frac{\partial u}{\partial t} = \nabla^2\big(u^m\big) \quad , \tag{9.134}$$

and the *p-Laplacian equation*,

$$\frac{\partial u}{\partial t} = \boldsymbol{\nabla}\cdot\Big(|\boldsymbol{\nabla} u|^{p-2}\,\boldsymbol{\nabla} u\Big) \quad . \tag{9.135}$$

Both these equations introduce a nonlinearity whereby the diffusion constant $D$ depends on the field $u$. For example, the PME can be rewritten $\partial_t u = \boldsymbol{\nabla}\cdot\big(m\,u^{m-1}\boldsymbol{\nabla} u\big)$, whence $D = m\,u^{m-1}$. For the $p$-Laplacian equation, $D = |\boldsymbol{\nabla} u|^{p-2}$. These nonlinearities strangle the diffusion when $u$ or $|\boldsymbol{\nabla} u|$ gets small, preventing the solution from advancing infinitely fast.

As its name betokens, the PME describes fluid flow in a porous medium. A fluid moving through a porous medium is described by three fundamental equations:

(i) *Continuity:* In a medium with porosity $\varepsilon$, the continuity equation becomes $\varepsilon\,\partial_t\varrho + \nabla\cdot(\varrho\boldsymbol{v}) = 0$, where $\varrho$ is the fluid density. This is because in a volume $\Omega$ where the fluid density is changing at a rate $\partial_t\varrho$, the rate of change of fluid mass is $\varepsilon\Omega\,\partial_t\varrho$.

(ii) *Darcy's law:* First articulated in 1856 by the French hydrologist Henry Darcy, this says that the flow velocity is directly proportional to the pressure gradient according to the relation $\boldsymbol{v} = -(K/\mu)\boldsymbol{\nabla} p$, where the *permeability* $K$ depends on the medium but not on the fluid, and $\mu$ is the shear viscosity of the fluid.

(iii) *Fluid equation of state:* This is a relation between the pressure $p$ and the density $\varrho$ of the fluid. For isentropic ideal gases, $p = A\,\varrho^\gamma$ where $A$ is a constant and $\gamma = c_p/c_V$ is the specific heat ratio.

Putting these three equations together, we obtain

$$\frac{\partial\varrho}{\partial t} = C\,\nabla^2\big(\varrho^m\big) \quad , \tag{9.136}$$

where $C = A\gamma k/(\gamma+1)\varepsilon\mu$ and $m = 1+\gamma$.

### 9.6.3 Illustrative solutions

A class of solution to the PME was discussed in the Russian literature in the early 1950s in a series of papers by Zeldovich, Kompaneets, and Barenblatt. The ZKB solution, which is isotropic in $d$ space

dimensions, is of the scaling form,

$$U(r,t) = t^{-\alpha}\,F\big(r\,t^{-\alpha/d}\big) \qquad ; \qquad F(\xi) = \Big(C - k\,\xi^2\Big)_+^{\frac{1}{m-1}} \quad , \tag{9.137}$$

where $r = |\boldsymbol{x}|$,

$$\alpha = \frac{d}{(m-1)d+2} \qquad , \qquad k = \frac{m-1}{2m}\cdot\frac{1}{(m-1)d+2} \quad , \tag{9.138}$$

and the $+$ subscript in the definition of $F(\xi)$ in Eqn. 9.137 indicates that the function is cut off and vanishes when the quantity inside the round brackets becomes negative. We also take $m > 1$, which means that $\alpha < \frac{1}{2}d$. The quantity $C$ is determined by initial conditions. The scaling form is motivated by the fact that the PME conserves the integral of $u(\boldsymbol{x},t)$ over all space, provided the current $\boldsymbol{j} = -mu^{m-1}\boldsymbol{\nabla}u$ vanishes at spatial infinity. Explicitly, we have

$$\int d^dx\,U(\boldsymbol{x},t) = \Omega_d\int_0^\infty dr\,r^{d-1}\,t^{-\alpha}\,F\big(r\,t^{-\alpha/d}\big) = \Omega_d\int_0^\infty ds\,s^{d-1}\,F(s) \quad , \tag{9.139}$$

where $\Omega_d$ is the total solid angle in $d$ space dimensions. The above integral is therefore independent of $t$, which means that the integral of $U$ is conserved. Therefore as $t\to 0$, we must have $U(\boldsymbol{x},t=0) = A\,\delta(\boldsymbol{x})$, where $A$ is a constant which can be expressed in terms of $C$, $m$, and $d$. We plot the behavior of this solution for the case $m = 2$ and $d = 1$ in Fig. 9.5, and compare and contrast it to the solution of the diffusion equation. Note that the solutions to the PME have compact support, *i.e.* they vanish identically for $r > \sqrt{C/k}\;t^{\alpha/d}$, which is consistent with a finite maximum speed of propagation. A similar point source solution to the $p$-Laplacian equation in $d = 1$ was obtained by Barenblatt:

$$U(x,t) = t^{-m}\Big(C - k\,|\xi|^{1+m^{-1}}\Big)^{\frac{m}{m-1}} \quad , \tag{9.140}$$

for arbitrary $C > 0$, with $\xi = x\,t^{-1/2m}$, and $k = (m-1)(2m)^{-(m+1)/m}$.

To derive the ZKB solution of the porous medium equation, it is useful to write the PME in terms of the 'pressure' variable $v = \frac{m}{m-1}\,u^{m-1}$. The PME then takes the form

$$\frac{\partial v}{\partial t} = (m-1)\,v\,\nabla^2 v + (\boldsymbol{\nabla}v)^2 \quad . \tag{9.141}$$

We seek an isotropic solution in $d$ space dimensions, and posit the scaling form

$$V(\boldsymbol{x},t) = t^{-\lambda}\,G\big(r\,t^{-\mu}\big) \quad , \tag{9.142}$$

where $r = |\boldsymbol{x}|$. Acting on isotropic functions, the Laplacian is given by $\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{d-1}{r}\frac{\partial}{\partial r}$. Defining $\xi = r\,t^{-\mu}$, we have

$$\frac{\partial V}{\partial t} = -t^{-1}\Big[\lambda\,G + \mu\,\xi\,G'\Big] \qquad , \qquad \frac{\partial V}{\partial r} = t^{-(\lambda+\mu)}\,G' \qquad , \qquad \frac{\partial^2 V}{\partial r^2} = t^{-(\lambda+2\mu)}\,G'' \quad , \tag{9.143}$$

whence

$$-\Big[\lambda\,G + \mu\,\xi\,G'\Big]\,t^{-1} = \Big[(m-1)\,G\,G'' + (m-1)\,(d-1)\,\xi^{-1}\,G\,G' + (G')^2\Big]\,t^{-2(\lambda+\mu)} \quad . \tag{9.144}$$

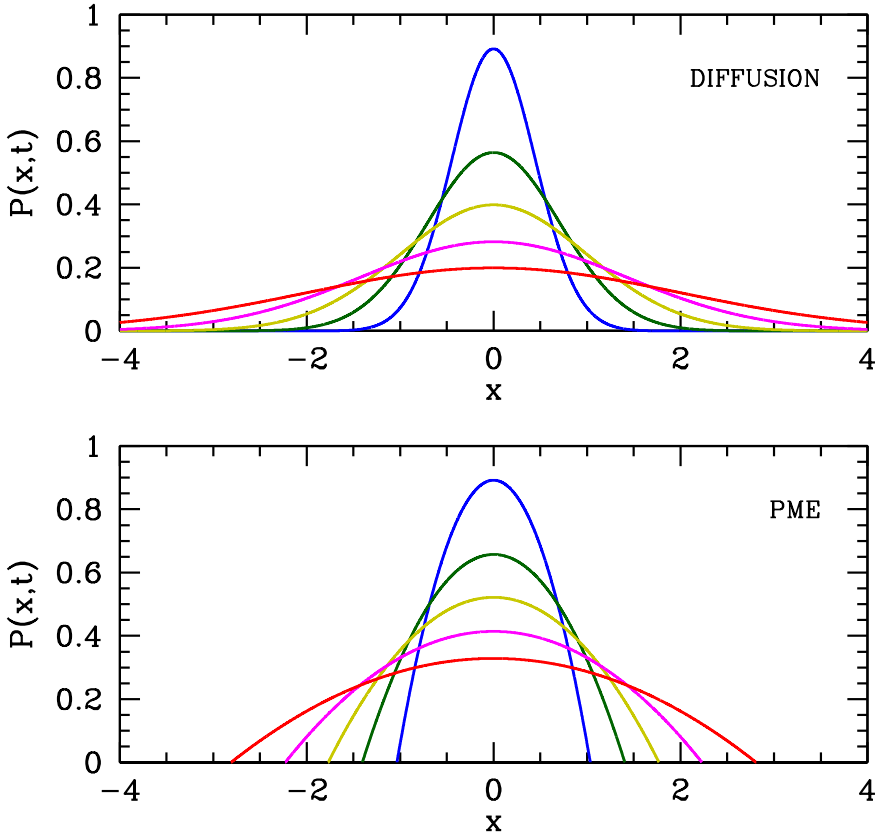



Figure 9.5: Top panel: evolution of the diffusion equation with $D = 1$ and $\sigma = 1$ for times $t = 0.1$, $0.25$, $0.5$, $1.0$, and $2.0$. Bottom panel: evolution of the porous medium equation with $m = 2$ and $d = 1$ and $C$ chosen so that $P(x = 0, t = 0.1)$ is equal to the corresponding value in the top panel (*i.e.* the peak of the blue curve).

At this point we can read off the result $\lambda + \mu = \frac{1}{2}$ and eliminate the $t$ variable, which validates our initial scaling form hypothesis. What remains is

$$\lambda\, G + \mu\, \xi G' + (m-1)\, G\, G'' + (m-1)(d-1)\, \xi^{-1}\, G\, G' + (G')^2 = 0 \quad . \tag{9.145}$$

Inspection now shows that this equation has a solution of the form $G(\xi) = A - b\,\xi^2$. Plugging this in, we find

$$\lambda = (m-1)\,\alpha \quad , \quad \mu = \frac{\alpha}{d} \quad , \quad b = \frac{\alpha}{2d} \quad , \quad \alpha \equiv \frac{d}{(m-1)\,d + 2} \quad . \tag{9.146}$$

The quadratic function $G(\xi) = A - b\,\xi^2$ goes negative for $\xi^2 > A/b$, which is clearly unphysical in the context of diffusion. To remedy this, Zeldovich *et al.* proposed to take the maximum value of $G(\xi)$ and zero. Clearly $G = 0$ is a solution, hence $G(\xi) = \big(A - b\,\xi^2\big)_+$ is a solution for $|\xi| < \sqrt{A/b}$ and for $|\xi| > \sqrt{A/b}$, but what about the points $\xi = \pm\sqrt{A/b}$? The concern is that the second derivative $G''(\xi)$ has a delta function singularity at those points, owing to the discontinuity of $G'(\xi)$. However, an examination of Eqn. 9.145 shows that $G''$ is multiplied by $G$, and we know that $\lim_{x\to 0} x\,\delta(x) = 0$. The remaining nonzero terms in this equation are then $\big[\mu\,\xi + G'(\xi)\big]G'(\xi)$, which agreeably vanishes. So we

have a solution of the form[12]

$$V(\boldsymbol{x},t) = \frac{1}{t}\left(A'\,t^{2\alpha/d} - \alpha\,\boldsymbol{x}^2\right)_+ \quad , \tag{9.147}$$

where $A' = 2dA$.

## 9.7 Appendix II: Langevin equation for a particle in a harmonic well

Consider next the equation

$$\ddot{X} + \gamma\dot{X} + \omega_0^2 X = \frac{F}{M} + \eta(t) \quad , \tag{9.148}$$

where $F$ is a constant force. We write $X = x_0 + x$ and measure $x$ relative to the potential minimum $x_0 = F/M\omega_0^2$, yielding

$$\ddot{x} + \gamma\,\dot{x} + \omega_0^2\,x = \eta(t) \quad . \tag{9.149}$$

We solve via Laplace transform. Recall

$$\check{x}(z) = \int_0^\infty dt\; e^{-zt}\,x(t) \qquad , \qquad x(t) = \int_{\mathcal{C}} \frac{dz}{2\pi i}\; e^{+zt}\,\check{x}(z) \quad , \tag{9.150}$$

where the contour $\mathcal{C}$ extends from $c - i\infty$ to $c + i\infty$ such that all poles of the $\check{x}(z)$ lie to the left of $\mathcal{C}$. Then

$$\begin{aligned}\int_0^\infty dt\; e^{-zt}\left(\ddot{x} + \gamma\,\dot{x} + \omega_0^2\,x\right) &= -(z+\gamma)\,x(0) - \dot{x}(0) + \left(z^2 + \gamma z + \omega_0^2\right)\check{x}(z) \\ &= \int_0^\infty dt\; e^{-zt}\,\eta(t) = \check{\eta}(z) \quad .\end{aligned} \tag{9.151}$$

Thus, we have

$$\check{x}(z) = \frac{(z+\gamma)\,x(0) + \dot{x}(0)}{z^2 + \gamma z + \omega_0^2} + \frac{1}{z^2 + \gamma z + \omega_0^2}\int_0^\infty dt\; e^{-zt}\,\eta(t) \quad . \tag{9.152}$$

Now we may write

$$z^2 + \gamma z + \omega_0^2 = (z - z_+)(z - z_-) \quad , \tag{9.153}$$

where $z_\pm = -\frac{1}{2}\gamma \pm \sqrt{\frac{1}{4}\gamma^2 - \omega_0^2}$. Note that $\mathrm{Re}\,(z_\pm) \le 0$ and that $z_\mp = -\gamma - z_\pm$.

Performing the inverse Laplace transform, we obtain

$$x(t) = \frac{x(0)}{z_+ - z_-}\left(z_+\,e^{z_- t} - z_-\,e^{z_+ t}\right) + \frac{\dot{x}(0)}{z_+ - z_-}\left(e^{z_+ t} - e^{z_- t}\right) + \int_0^\infty ds\; K(t-s)\,\eta(s) \quad , \tag{9.154}$$

[12] Actually the result $\lim_{x\to 0} x\,\delta(x) = 0$ is valid in the distribution sense, *i.e.* underneath an integral, provided $x\,\delta(x)$ is multiplied by a nonsingular function of $x$. Thus, Eqn. 9.147 constitutes a *weak solution* to the pressure form of the porous medium equation 9.141. Zeldovich *et al.* found numerically that cutting off the negative part of $A - b\,\xi^2$ is appropriate. Mathematically, Vazquez has shown that when the initial data are taken within a suitable class of integrable functions, the weak solution exists and is unique.

where

$$K(t-s) = \frac{\Theta(t-s)}{(z_+ - z_-)}\left(e^{z_+(t-s)} - e^{z_-(t-s)}\right) \tag{9.155}$$

is the *response kernel* and $\Theta(t-s)$ is the step function which is unity for $t > s$ and zero otherwise. The response is *causal*, *i.e.* $x(t)$ depends on $\eta(s)$ for all previous times $s < t$, but not for future times $s > t$. Note that $K(\tau)$ decays exponentially for $\tau \to \infty$, if $\mathsf{Re}(z_\pm) < 0$. The marginal case where $\omega_0 = 0$ and $z_+ = 0$ corresponds to the diffusion calculation we performed in the previous section.

It is now easy to compute

$$\left\langle x^2(t)\right\rangle_{\rm c} = \Gamma\int\limits_0^t ds\; K^2(s) = \frac{\Gamma}{2\omega_0^2\gamma} \quad (t\to\infty) \tag{9.156}$$

$$\left\langle \dot{x}^2(t)\right\rangle_{\rm c} = \Gamma\int\limits_0^t ds\; \dot{K}^2(s) = \frac{\Gamma}{2\gamma} \qquad (t\to\infty) \quad , \tag{9.157}$$

where the connected average is defined by $\langle AB\rangle_{\rm c} = \langle AB\rangle - \langle A\rangle\langle B\rangle$. Therefore,

$$\left\langle \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}M\omega_0^2 x^2\right\rangle_{t\to\infty} = \frac{M\Gamma}{2\gamma} \quad . \tag{9.158}$$

Setting this equal to $2\times\frac{1}{2}k_{\rm B}T$ by equipartition again yields $\Gamma = 2\gamma k_{\rm B}T/M$.

## 9.8 Appendix III: General Linear Autonomous Inhomogeneous ODEs

### 9.8.1 Solution by Fourier transform

We can solve general autonomous linear inhomogeneous ODEs of the form

$$\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1\frac{dx}{dt} + a_0\, x = \eta(t) \quad . \tag{9.159}$$

We can write this as

$$\mathcal{L}_t\, x(t) = \eta(t) \quad , \tag{9.160}$$

where $\mathcal{L}_t$ is the $n^{\rm th}$ order differential operator

$$\mathcal{L}_t = \frac{d^n}{dt^n} + a_{n-1}\frac{d^{n-1}}{dt^{n-1}} + \ldots + a_1\frac{d}{dt} + a_0 \quad . \tag{9.161}$$

The general solution to the inhomogeneous equation is given by

$$x(t) = x_{\rm h}(t) + \int\limits_{-\infty}^{\infty} dt'\; G(t,t')\,\eta(t') \quad , \tag{9.162}$$

where $G(t,t')$ is the Green's function. Note that $\mathcal{L}_t\, x_{\rm h}(t) = 0$. Thus, in order for eqns. 9.160 and 9.162 to be true, we must have

$$\mathcal{L}_t\, x(t) = \overbrace{\mathcal{L}_t\, x_{\rm h}(t)}^{\text{this vanishes}} + \int\limits_{-\infty}^{\infty} dt'\, \mathcal{L}_t\, G(t,t')\, \eta(t') = \eta(t) \quad , \tag{9.163}$$

which means that $\mathcal{L}_t\, G(t,t') = \delta(t-t')$, where $\delta(t-t')$ is the Dirac $\delta$-function.

If the differential equation $\mathcal{L}_t\, x(t) = \eta(t)$ is defined over some finite or semi-infinite $t$ interval with prescribed boundary conditions on $x(t)$ at the endpoints, then $G(t,t')$ will depend on $t$ and $t'$ separately. For the case we are now considering, let the interval be the entire real line $t \in (-\infty, \infty)$. Then $G(t,t') = G(t-t')$ is a function of the single variable $t-t'$.

Note that $\mathcal{L}_t = \mathcal{L}\big(\frac{d}{dt}\big)$ may be considered a function of the differential operator $\frac{d}{dt}$. If we now Fourier transform the equation $\mathcal{L}_t\, x(t) = \eta(t)$, we obtain

$$\begin{aligned}\int\limits_{-\infty}^{\infty} dt\, e^{i\omega t}\, \eta(t) &= \int\limits_{-\infty}^{\infty} dt\, e^{i\omega t} \left\{ \frac{d^n}{dt^n} + a_{n-1}\, \frac{d^{n-1}}{dt^{n-1}} + \ldots + a_1\, \frac{d}{dt} + a_0 \right\} x(t) \\ &= \int\limits_{-\infty}^{\infty} dt\, e^{i\omega t} \left\{ (-i\omega)^n + a_{n-1}\, (-i\omega)^{n-1} + \ldots + a_1\, (-i\omega) + a_0 \right\} x(t) \quad .\end{aligned} \tag{9.164}$$

Thus, if we define

$$\hat{\mathcal{L}}(\omega) = \sum_{k=0}^{n} a_k\, (-i\omega)^k \quad , \tag{9.165}$$

then we have $\hat{\mathcal{L}}(\omega)\, \hat{x}(\omega) = \hat{\eta}(\omega)$ , where $a_n \equiv 1$. According to the Fundamental Theorem of Algebra, the $n^{\rm th}$ degree polynomial $\hat{\mathcal{L}}(\omega)$ may be uniquely factored over the complex $\omega$ plane into a product over $n$ roots:

$$\hat{\mathcal{L}}(\omega) = (-i)^n\, (\omega - \omega_1)(\omega - \omega_2) \cdots (\omega - \omega_n) \quad . \tag{9.166}$$

If the $\{a_k\}$ are all real, then $\big[\hat{\mathcal{L}}(\omega)\big]^* = \hat{\mathcal{L}}(-\omega^*)$, hence if $\Omega$ is a root then so is $-\Omega^*$. Thus, the roots appear in pairs which are symmetric about the imaginary axis. *I.e.* if $\Omega = a + ib$ is a root, then so is $-\Omega^* = -a + ib$.

The general solution to the homogeneous equation is

$$x_{\rm h}(t) = \sum_{\sigma=1}^{n} A_\sigma\, e^{-i\omega_\sigma t} \quad , \tag{9.167}$$

which involves $n$ arbitrary complex constants $A_i$. The susceptibility, or Green's function in Fourier space, $\hat{G}(\omega)$ is then

$$\hat{G}(\omega) = \frac{1}{\hat{\mathcal{L}}(\omega)} = \frac{i^n}{(\omega - \omega_1)(\omega - \omega_2) \cdots (\omega - \omega_n)} \quad , \tag{9.168}$$

Note that $\big[\hat{G}(\omega)\big]^* = \hat{G}(-\omega)$, which is equivalent to the statement that $G(t-t')$ is a real function of its argument. The general solution to the inhomogeneous equation is then

$$x(t) = x_{\rm h}(t) + \int\limits_{-\infty}^{\infty} dt'\, G(t-t')\,\eta(t') \quad , \tag{9.169}$$

where $x_{\rm h}(t)$ is the solution to the homogeneous equation, *i.e.* with zero forcing, and where

$$\begin{aligned} G(t-t') &= \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, e^{-i\omega(t-t')}\,\hat{G}(\omega) \\ &= i^n \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, \frac{e^{-i\omega(t-t')}}{(\omega-\omega_1)(\omega-\omega_2)\cdots(\omega-\omega_n)} = \sum_{\sigma=1}^{n} \frac{e^{-i\omega_\sigma(t-t')}}{i\,\hat{\mathcal{L}}'(\omega_\sigma)}\,\Theta(t-t') \quad , \end{aligned} \tag{9.170}$$

where we assume that $\mathrm{Im}\,\omega_\sigma < 0$ for all $\sigma$. This guarantees *causality* – the *response* $x(t)$ to the *influence* $\eta(t')$ is nonzero only for $t > t'$.

As an example, consider the familiar case $\hat{\mathcal{L}}(\omega) = -\omega^2 - i\gamma\omega + \omega_0^2$, which can be written as $\hat{\mathcal{L}}(\omega) = -(\omega-\omega_+)\,(\omega-\omega_-)$ with $\omega_\pm = -\frac{i}{2}\gamma \pm \beta$, and $\beta = \left(\omega_0^2 - \frac{1}{4}\gamma^2\right)^{1/2}$. This yields $\hat{\mathcal{L}}'(\omega_\pm) = \mp(\omega_+ - \omega_-) = \mp 2\beta$, hence according to equation 9.170,

$$\begin{aligned} G(s) &= \left\{ \frac{e^{-i\omega_+ s}}{i\mathcal{L}'(\omega_+)} + \frac{e^{-i\omega_- s}}{i\mathcal{L}'(\omega_-)} \right\} \Theta(s) \\ &= \left\{ \frac{e^{-\gamma s/2}\,e^{-i\beta s}}{-2i\beta} + \frac{e^{-\gamma s/2}\,e^{i\beta s}}{2i\beta} \right\} \Theta(s) = \beta^{-1}\,e^{-\gamma s/2}\,\sin(\beta s)\,\Theta(s) \quad . \end{aligned} \tag{9.171}$$

Now let us evaluate the two-point correlation function $\big\langle x(t)\,x(t')\big\rangle$, assuming the noise is correlated according to $\big\langle \eta(s)\,\eta(s')\big\rangle = \phi(s-s')$. We assume $t, t' \to \infty$ so the transient contribution $x_{\rm h}$ is negligible. We then have

$$\big\langle x(t)\,x(t')\big\rangle = \int\limits_{-\infty}^{\infty} ds \int\limits_{-\infty}^{\infty} ds'\, G(t-s)\,G(t'-s')\,\big\langle \eta(s)\,\eta(s')\big\rangle = \int\limits_{-\infty}^{\infty} \frac{d\omega}{2\pi}\,\hat{\phi}(\omega)\,\big|\hat{G}(\omega)\big|^2\,e^{i\omega(t-t')} \quad . \tag{9.172}$$

### 9.8.2 Solution by Laplace transform

Laplace transforms are useful when dealing with functions which vanish for sufficiently negative values of their argument. This is the case, for example, when an oscillator in a quiescent state is suddenly forced by an external drive which turns on at some time $t_0$ (typically $t_0 = 0$). The Laplace transform $\check{x}(z)$ is obtained from a function $x(t)$ via the integral transform

$$\check{x}(z) = \int\limits_{0}^{\infty} dt\, e^{-zt}\,x(t) \quad . \tag{9.173}$$

The inverse transform is given by

$$x(t) = \int\limits_{c-i\infty}^{c+i\infty} \frac{dz}{2\pi i}\, e^{zt}\, \check{x}(z) \quad , \tag{9.174}$$

where the integration contour is a straight line which lies to the right of any singularities of $\check{x}(z)$ in the complex $z$ plane. Note the similarity to the Fourier transform.

Now let's take the Laplace transform of Eqn. 9.160. Note that integration by parts yields

$$\int\limits_0^\infty dt\, e^{-zt}\, \frac{df}{dt} = z\check{f}(z) - f(0) \tag{9.175}$$

for any function $f(t)$. Applying this result iteratively, we find the Laplace transform of Eqn. 9.160 is

$$L(z)\, \check{x}(z) = \check{\eta}(z) + R_0(z) \quad , \tag{9.176}$$

where

$$L(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0 \tag{9.177}$$

is an $n^{\rm th}$ order polynomial in $z$ with coefficients $a_j$ for $j \in \{0, \ldots, n\}$, and

$$R_0(z) = a_n\, x^{(n-1)}(0) + \big(z a_n + a_{n-1}\big)\, x^{(n-2)}(0) + \cdots + \big(z^{n-1} a_n + \ldots + a_1\big)\, x(0) \tag{9.178}$$

and $x^{(k)}(t) = d^k x/dt^k$. We now have

$$\check{x}(z) = \frac{1}{L(z)} \Big\{ \check{\eta}(z) + R_0(z) \Big\} \quad . \tag{9.179}$$

The formal solution to Eqn. 9.160 is then given by the inverse Laplace transform. One finds

$$x(t) = \int\limits_0^t dt'\, K(t-t')\, \eta(t') + x_{\rm h}(t) \quad , \tag{9.180}$$

where $x_{\rm h}(t)$ is a solution to the homogeneous equation $\mathcal{L}_t\, x(t) = 0$, and

$$K(s) = \int\limits_{c-i\infty}^{c+i\infty} \frac{dz}{2\pi i}\, \frac{e^{zs}}{L(z)} = \sum_{l=1}^n \frac{e^{z_l s}}{L'(z_l)} \quad . \tag{9.181}$$

Note that $K(s)$ vanishes for $s < 0$ because then we can close the contour in the far right half plane. The RHS of the above equation follows from the fundamental theorem of algebra, which allows us to factor $L(z)$ as

$$L(z) = a_n (z - z_1) \cdots (z - z_n) \quad , \tag{9.182}$$

with all the roots $z_l$ lying to the left of the contour. In deriving the RHS of Eqn. 9.181, we assume that all roots are distinct[13]. The general solution to the homogeneous equation is

$$x_{\rm h}(t) = \sum_{l=1}^n A_l\, e^{z_l t} \quad , \tag{9.183}$$

[13] If two or more roots are degenerate, one can still use this result by first inserting a small spacing $\varepsilon$ between the degenerate roots and then taking $\varepsilon \to 0$.

again assuming the roots are nondegenerate[14]. In order that the homogeneous solution not grow with time, we must have $\mathrm{Re}\,(z_l) \le 0$ for all $l$. For example, if $\mathcal{L}_t = \frac{d}{dt} + \gamma$, then $L(z) = z + \gamma$ and $K(s) = e^{-\gamma s}$. If $\mathcal{L}_t = \frac{d^2}{dt^2} + \gamma \frac{d}{dt}$, then $L(z) = z^2 + \gamma z$ and $K(s) = (1 - e^{-\gamma s})/\gamma$.

Let us assume that all the initial derivatives $d^k x(t)/dt^k$ vanish at $t = 0$, hence $x_{\rm h}(t) = 0$. Now let us compute the generalization of Eqn. 9.35,

$$\begin{aligned} P(x_1, t_1; \ldots; x_N, t_N) &= \int_{-\infty}^{\infty} \frac{d\omega_1}{2\pi} \cdots \int_{-\infty}^{\infty} \frac{d\omega_N}{2\pi}\, e^{-i(\omega_1 x_1 + \ldots + \omega_N x_N)} \left\langle e^{i\omega_1 x(t_1)} \ldots e^{i\omega_N x(t_N)} \right\rangle \\ &= \det{}^{-1/2}(2\pi M)\, \exp\left\{ -\tfrac{1}{2} \sum_{j,j'=1}^{N} M^{-1}_{jj'}\, x_j\, x_{j'} \right\} \quad , \end{aligned} \tag{9.184}$$

where

$$M(t, t') = \int_0^t ds \int_0^{t'} ds'\, G(s - s')\, K(t - s)\, K(t' - s') \quad , \tag{9.185}$$

with $G(s - s') = \langle \eta(s)\, \eta(s') \rangle$ as before. For $t \gg \gamma^{-1}$, we have $K(s) = \gamma^{-1}$, and if we take the correlator to be $G(s - s') = \Gamma\, \delta(s - s')$ we obtain $M(t, t') = \Gamma\, \min(t, t')/\gamma^2 = 2D\, \min(t, t')$. We then have the expected diffusion probability, $P(x, t) = \exp\left(-x^2/4Dt\right)/\sqrt{4\pi D t}$.

### 9.8.3 Higher order ODEs

Note that any $n^{\rm th}$ order ODE, of the general form

$$\frac{d^n x}{dt^n} = F\left( x\,,\, \frac{dx}{dt}\,,\, \ldots\,,\, \frac{d^{n-1}x}{dt^{n-1}} \right) \quad , \tag{9.186}$$

may be represented by the first order system $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$. To see this, define $\varphi_k = d^{k-1}x/dt^{k-1}$, with $k = 1, \ldots, n$. Thus, for $k < n$ we have $\dot{\varphi}_k = \varphi_{k+1}$, and $\dot{\varphi}_n = F$. In other words,

$$\overbrace{\frac{d}{dt}\begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_{n-1} \\ \varphi_n \end{pmatrix}}^{\dot{\boldsymbol{\varphi}}} = \overbrace{\begin{pmatrix} \varphi_2 \\ \vdots \\ \varphi_n \\ F(\varphi_1, \ldots, \varphi_n) \end{pmatrix}}^{\boldsymbol{V}(\boldsymbol{\varphi})} \quad . \tag{9.187}$$

An inhomogeneous linear $n^{\rm th}$ order ODE,

$$\frac{d^n x}{dt^n} + a_{n-1} \frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_1 \frac{dx}{dt} + a_0\, x = \eta(t) \tag{9.188}$$

[14] If a particular root $z_j$ appears $k$ times, then one has solutions of the form $e^{z_j t}$, $t\, e^{z_j t}$, $\ldots$ $t^{k-1}\, e^{z_j t}$.

may be written in matrix form, as

$$\frac{d}{dt}\begin{pmatrix}\varphi_1\\ \varphi_2\\ \vdots\\ \varphi_n\end{pmatrix} = \overbrace{\begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1}\end{pmatrix}}^{Q}\begin{pmatrix}\varphi_1\\ \varphi_2\\ \vdots\\ \varphi_n\end{pmatrix} + \overbrace{\begin{pmatrix}0\\ 0\\ \vdots\\ \eta(t)\end{pmatrix}}^{\boldsymbol{\xi}} \quad . \tag{9.189}$$

Thus,

$$\dot{\boldsymbol{\varphi}} = Q\,\boldsymbol{\varphi} + \boldsymbol{\xi} \quad , \tag{9.190}$$

and if the coefficients $a_k$ are time-independent the ODE is also said to be *autonomous*.

For the homogeneous case where $\eta(t) = 0$, and hence $\boldsymbol{\xi}(t) = 0$, the solution is obtained by exponentiating the constant matrix $Qt$:

$$\boldsymbol{\varphi}(t) = \exp(Qt)\,\boldsymbol{\varphi}(0) \quad . \tag{9.191}$$

Recall that the exponential of a matrix may be given meaning by its Taylor series expansion. If the ODE is not autonomous, then $Q = Q(t)$ is time-dependent, and the solution is given by the path-ordered exponential,

$$\boldsymbol{\varphi}(t) = \mathcal{P}\exp\left\{\int_0^t dt'\,Q(t')\right\}\boldsymbol{\varphi}(0) \quad , \tag{9.192}$$

where $\mathcal{P}$ is the *path ordering operator* which places earlier times to the right. As defined, the equation $\dot{\boldsymbol{\varphi}} = \boldsymbol{V}(\boldsymbol{\varphi})$ is autonomous, since the $t$-advance mapping $g_t$ depends only on $t$ and on no other time variable. However, by extending the phase space $\mathbb{M} \ni \boldsymbol{\varphi}$ from $\mathbb{M} \to \mathbb{M} \times \mathbb{R}$, which is of dimension $n+1$, one can describe arbitrary time-dependent ODEs.

In general, path ordered exponentials are difficult to compute analytically. We will henceforth consider the autonomous case where $Q$ is a constant matrix in time. We will assume the matrix $Q$ is real, but other than that it has no helpful symmetries. Assuming it has no Jordan blocks, which is the generic state of affairs, we may decompose $Q$ into left and right eigenvectors, viz

$$Q_{ij} = \sum_{\sigma=1}^{n} \nu_\sigma\, R_i^\sigma\, L_j^\sigma \quad . \tag{9.193}$$

Or, in bra-ket notation, $Q = \sum_\sigma \nu_\sigma\,|\,R_\sigma\,\rangle\langle\,L_\sigma\,|$. We adopt the normalization convention $\langle\,L^\sigma\,|\,R^{\sigma'}\,\rangle = \delta^{\sigma\sigma'}$, where $\{\nu_\sigma\}$ are the eigenvalues of $Q$:

$$Q\,|\,R^\sigma\,\rangle = \nu_\sigma\,|\,R^\sigma\,\rangle \qquad , \qquad \langle\,L^\sigma\,|\,Q = \nu_\sigma\,\langle\,L^\sigma\,| \quad . \tag{9.194}$$

The eigenvalues may be real or complex. Since the characteristic polynomial $P(\nu) = \det(\nu\mathbb{1} - Q)$ has real coefficients, we know that the eigenvalues of $Q$ are either real or come in complex conjugate pairs.

Consider, for example, the $n = 2$ system

$$Q = \begin{pmatrix}0 & 1\\ -\omega_0^2 & -\gamma\end{pmatrix} \quad . \tag{9.195}$$

The eigenvalues are as before: $\nu_\pm = -\frac{1}{2}\gamma \pm \left(\frac{1}{4}\gamma^2 - \omega_0^2\right)^{1/2}$. The left and right eigenvectors are

$$\boldsymbol{L}^\pm = \frac{\pm 1}{\nu_+ - \nu_-}\begin{pmatrix} -\nu_\mp & 1 \end{pmatrix} \qquad , \qquad \boldsymbol{R}^\pm = \begin{pmatrix} 1 \\ \nu_\pm \end{pmatrix} \quad . \tag{9.196}$$

The utility of working in a left-right eigenbasis is apparent once we reflect upon the result

$$f(Q) = \sum_{\sigma=1}^{n} f(\nu_\sigma)\, |\, R^\sigma \,\rangle\, \langle\, L^\sigma \,| \tag{9.197}$$

for any function $f$. Thus, the solution to the general autonomous homogeneous case is

$$\begin{aligned} |\, \varphi(t)\,\rangle &= \sum_{\sigma=1}^{n} e^{\nu_\sigma t}\, |\, R^\sigma \,\rangle\, \langle\, L^\sigma \,|\, \varphi(0)\,\rangle \\ \varphi_i(t) &= \sum_{\sigma=1}^{n} e^{\nu_\sigma t}\, R^\sigma_i \sum_{j=1}^{n} L^\sigma_j\, \varphi_j(0) \quad . \end{aligned} \tag{9.198}$$

If $\mathrm{Re}\,(\nu_\sigma) \le 0$ for all $\sigma$, then the initial conditions $\boldsymbol{\varphi}(0)$ are forgotten on time scales $\tau_\sigma = -(\mathrm{Re}\,\nu_\sigma)^{-1}$. Physicality demands that this is the case.

Now let's consider the inhomogeneous case where $\eta(t) \neq 0$. We begin by recasting eqn. 9.190 in the form

$$\frac{d}{dt}\big(e^{-Qt}\,\boldsymbol{\varphi}\big) = e^{-Qt}\,\boldsymbol{\xi}(t) \quad . \tag{9.199}$$

We can integrate this directly:

$$\boldsymbol{\varphi}(t) = e^{Qt}\,\boldsymbol{\varphi}(0) + \int_0^t ds\, e^{Q(t-s)}\,\boldsymbol{\xi}(s) \quad . \tag{9.200}$$

In component notation,

$$\varphi_i(t) = \sum_{\sigma=1}^{n} e^{\nu_\sigma t}\, R^\sigma_i\, \langle\, L^\sigma \,|\, \varphi(0)\,\rangle + \sum_{\sigma=1}^{n} R^\sigma_i \int_0^t ds\, e^{\nu_\sigma (t-s)}\, \langle\, L^\sigma \,|\, \xi(s)\,\rangle. \tag{9.201}$$

Note that the first term on the RHS is the solution to the homogeneous equation, as must be the case when $\boldsymbol{\xi}(s) = 0$.

The solution in eqn. 9.201 holds for general $Q$ and $\boldsymbol{\xi}(s)$. For the particular form of $Q$ and $\eta(s)$ in eqn. 9.189, we can proceed further. For starters, $\langle\, L^\sigma \,|\, \xi(s)\,\rangle = L^\sigma_n\, \eta(s)$. We can further exploit a special feature of the $Q$ matrix to analytically determine all its left and right eigenvectors. Applying $Q$ to the right eigenvector $|R^\sigma\rangle$, we find $R^\sigma_j = \nu_\sigma\, R^\sigma_{j-1}$ for $j > 1$. We are free to choose $R^\sigma_1 = 1$ for all $\sigma$ and defer the issue of normalization to the derivation of the left eigenvectors. Thus, we obtain the pleasingly simple result, $R^\sigma_k = \nu_\sigma^{k-1}$. Applying $Q$ to the left eigenvector $\langle\, L^\sigma \,|$, we obtain

$$\begin{aligned} -a_0\, L^\sigma_n &= \nu_\sigma\, L^\sigma_1 \\ L^\sigma_{j-1} - a_{j-1}\, L^\sigma_n &= \nu_\sigma\, L^\sigma_j \qquad (j > 1) \quad . \end{aligned} \tag{9.202}$$

From these equations we may derive

$$L_k^\sigma = -\frac{L_n^\sigma}{\nu_\sigma}\sum_{j=0}^{k-1} a_j\,\nu_\sigma^{j-k-1} = +\frac{L_n^\sigma}{\nu_\sigma}\sum_{j=k}^{n} a_j\,\nu_\sigma^{j-k-1} \quad . \tag{9.203}$$

The equality in the above equation is derived using the result $P(\nu_\sigma) = \sum_{j=0}^n a_j\,\nu_\sigma^j = 0$. Recall also that $a_n \equiv 1$. We now impose the normalization condition,

$$\sum_{j=1}^{n} L_j^\sigma\,R_j^\sigma = 1 \quad . \tag{9.204}$$

This condition determines our last remaining unknown quantity (for a given $\sigma$), $L_n^\sigma$ :

$$\langle\, L^\sigma \,|\, R^\sigma \,\rangle = L_n^\sigma \sum_{k=1}^{n} k\,a_k\,\nu_\sigma^{k-1} = P'(\nu_\sigma)\,L_n^\sigma \quad , \tag{9.205}$$

where $P'(\nu)$ is the first derivative of the characteristic polynomial. Thus, we find $L_n^\sigma = 1/P'(\nu_\sigma)$ .

Now let us evaluate the general two-point correlation function,

$$C_{jj'}(t,t') \equiv \big\langle \varphi_j(t)\,\varphi_{j'}(t')\big\rangle - \big\langle\varphi_j(t)\big\rangle\,\big\langle\varphi_{j'}(t')\big\rangle \quad . \tag{9.206}$$

We write

$$\big\langle\eta(s)\,\eta(s')\big\rangle = \phi(s-s') = \int\limits_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\hat\phi(\omega)\,e^{-i\omega(s-s')} \quad . \tag{9.207}$$

When $\hat\phi(\omega) = \Gamma$ is constant, we have $\big\langle\eta(s)\,\eta(s')\big\rangle = \Gamma\,\delta(s-s')$. This is the case of so-called *white noise*, when all frequencies contribute equally. The more general case when $\hat\phi(\omega)$ is frequency-dependent is known as *colored noise*. Appealing to eqn. 9.201, we have

$$C_{jj'}(t,t') = \sum_{\sigma,\sigma'}\frac{\nu_\sigma^{j-1}}{P'(\nu_\sigma)}\,\frac{\nu_{\sigma'}^{j'-1}}{P'(\nu_{\sigma'})}\int\limits_0^t ds\,e^{\nu_\sigma(t-s)}\int\limits_0^{t'} ds'\,e^{\nu_{\sigma'}(t'-s')}\,\phi(s-s') \tag{9.208}$$

$$= \sum_{\sigma,\sigma'}\frac{\nu_\sigma^{j-1}}{P'(\nu_\sigma)}\,\frac{\nu_{\sigma'}^{j'-1}}{P'(\nu_{\sigma'})}\int\limits_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\frac{\hat\phi(\omega)\,(e^{-i\omega t}-e^{\nu_\sigma t})(e^{i\omega t'}-e^{\nu_{\sigma'}t'})}{(\omega - i\nu_\sigma)(\omega+i\nu_{\sigma'})} \quad . \tag{9.209}$$

In the limit $t,t'\to\infty$, assuming $\mathrm{Re}\,(\nu_\sigma) < 0$ for all $\sigma$ (*i.e.* no diffusion), the exponentials $e^{\nu_\sigma t}$ and $e^{\nu_{\sigma'}t'}$ may be neglected, and we then have

$$C_{jj'}(t,t') = \sum_{\sigma,\sigma'}\frac{\nu_\sigma^{j-1}}{P'(\nu_\sigma)}\,\frac{\nu_{\sigma'}^{j'-1}}{P'(\nu_{\sigma'})}\int\limits_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\frac{\hat\phi(\omega)\,e^{-i\omega(t-t')}}{(\omega-i\nu_\sigma)(\omega+i\nu_{\sigma'})} \quad . \tag{9.210}$$

### 9.8.4 Kramers-Krönig relations

Suppose $\hat{\chi}(\omega) \equiv \hat{G}(\omega)$ is analytic in the UHP[15]. Then for all $\nu$, we must have

$$\int_{-\infty}^{\infty} \frac{d\nu}{2\pi} \, \frac{\hat{\chi}(\nu)}{\nu - \omega + i\epsilon} = 0 \quad , \tag{9.211}$$

where $\epsilon$ is a positive infinitesimal. The reason is simple: just close the contour in the UHP, assuming $\hat{\chi}(\omega)$ vanishes sufficiently rapidly that Jordan's lemma can be applied. Clearly this is an extremely weak restriction on $\hat{\chi}(\omega)$, given the fact that the denominator already causes the integrand to vanish as $|\omega|^{-1}$.

Let us examine the function

$$\frac{1}{\nu - \omega + i\epsilon} = \frac{\nu - \omega}{(\nu - \omega)^2 + \epsilon^2} - \frac{i\epsilon}{(\nu - \omega)^2 + \epsilon^2} \quad . \tag{9.212}$$

which we have separated into real and imaginary parts. Under an integral sign, the first term, in the limit $\epsilon \to 0$, is equivalent to taking a *principal part* of the integral. That is, for any function $F(\nu)$ which is regular at $\nu = \omega$,

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \frac{d\nu}{2\pi} \, \frac{\nu - \omega}{(\nu - \omega)^2 + \epsilon^2} \, F(\nu) \equiv \wp \int_{-\infty}^{\infty} \frac{d\nu}{2\pi} \, \frac{F(\nu)}{\nu - \omega} \quad . \tag{9.213}$$

The *principal part* symbol $\wp$ means that the singularity at $\nu = \omega$ is elided, either by regularizing the function $1/(\nu - \epsilon)$ as above, or by simply cutting out a region of integration of width $\epsilon$ on either side of $\nu = \omega$.

The imaginary part is more interesting. Let us write

$$h(u) \equiv \frac{\epsilon}{u^2 + \epsilon^2} \quad . \tag{9.214}$$

For $|u| \gg \epsilon$, $h(u) \simeq \epsilon/u^2$, which vanishes as $\epsilon \to 0$. For $u = 0$, $h(0) = 1/\epsilon$ which diverges as $\epsilon \to 0$. Thus, $h(u)$ has a huge peak at $u = 0$ and rapidly decays to 0 as one moves off the peak in either direction a distance greater than $\epsilon$. Finally, note that

$$\int_{-\infty}^{\infty} du \, h(u) = \pi \quad , \tag{9.215}$$

a result which itself is easy to show using contour integration. Putting it all together, this tells us that

$$\lim_{\epsilon \to 0} \frac{\epsilon}{u^2 + \epsilon^2} = \pi \delta(u) \quad . \tag{9.216}$$

Thus, for positive infinitesimal $\epsilon$, when part of an integrand, multiplied by any smooth function of $u$,

$$\frac{1}{u \pm i\epsilon} = \frac{\wp}{u} \mp i\pi\delta(u) \quad , \tag{9.217}$$

[15] In this section, we use the notation $\hat{\chi}(\omega)$ for the susceptibility, rather than $\hat{G}(\omega)$

a most useful result.

We now return to our initial result 9.211, and we separate $\hat{\chi}(\omega)$ into real and imaginary parts:

$$\hat{\chi}(\omega) = \hat{\chi}'(\omega) + i\hat{\chi}''(\omega) \quad . \tag{9.218}$$

(In this equation, the primes do not indicate differentiation with respect to argument.) We therefore have, for every real value of $\omega$,

$$0 = \int\limits_{-\infty}^{\infty} \frac{d\nu}{2\pi} \left[\chi'(\nu) + i\chi''(\nu)\right] \left[\frac{\wp}{\nu-\omega} - i\pi\delta(\nu-\omega)\right] \quad . \tag{9.219}$$

Taking the real and imaginary parts of this equation, we derive the *Kramers-Krönig relations*:

$$\chi'(\omega) = +\wp \int\limits_{-\infty}^{\infty} \frac{d\nu}{\pi}\, \frac{\hat{\chi}''(\nu)}{\nu-\omega} \qquad , \qquad \chi''(\omega) = -\wp \int\limits_{-\infty}^{\infty} \frac{d\nu}{\pi}\, \frac{\hat{\chi}'(\nu)}{\nu-\omega} \quad . \tag{9.220}$$

## 9.9 Appendix IV: Method of Characteristics

### 9.9.1 Quasilinear partial differential equations

Consider the quasilinear PDE

$$a_1(\boldsymbol{x},\phi)\,\frac{\partial\phi}{\partial x_1} + a_2(\boldsymbol{x},\phi)\,\frac{\partial\phi}{\partial x_2} + \ldots + a_N(\boldsymbol{x},\phi)\,\frac{\partial\phi}{\partial x_N} = b(\boldsymbol{x},\phi) \quad . \tag{9.221}$$

This PDE is called 'quasilinear' because it is linear in the derivatives $\partial\phi/\partial x_j$. The $N$ independent variables are the elements of the vector $\boldsymbol{x} = (x_1, \ldots, x_N)$. A solution is a function $\phi(\boldsymbol{x})$ which satisfies the PDE.

Now consider a curve $\boldsymbol{x}(s)$ parameterized by a single real variable $s$ satisfying

$$\frac{dx_j}{ds} = a_j\big(\boldsymbol{x},\phi(\boldsymbol{x})\big) \quad , \tag{9.222}$$

where $\phi(\boldsymbol{x})$ is a solution of eqn. 9.221. Along such a curve, which is called a *characteristic*, the variation of $\phi$ is

$$\frac{d\phi}{ds} = \sum_{j=1}^{N} \frac{\partial\phi}{\partial x_j}\,\frac{dx_j}{ds} = b\big(\boldsymbol{x}(s),\phi\big) \quad . \tag{9.223}$$

Thus, we have converted our PDE into a set of $N+1$ ODEs. To integrate, we must supply some initial conditions of the form

$$g\big(\boldsymbol{x},\phi\big)\Big|_{s=0} = 0 \quad . \tag{9.224}$$

This defines an $(N-1)$-dimensional hypersurface, parameterized by $\{\zeta_1, \ldots, \zeta_{N-1}\}$:

$$\begin{aligned} x_j(s=0) &= h_j(\zeta_1, \ldots, \zeta_{N-1}) \qquad , \qquad j \in \{1, \ldots, N\} \\ \phi(s=0) &= f(\zeta_1, \ldots, \zeta_{N-1}) \quad . \end{aligned} \tag{9.225}$$

If we can solve for all the characteristic curves, then the solution of the PDE follows. For every $\boldsymbol{x}$, we identify the characteristic curve upon which $\boldsymbol{x}$ lies. The characteristics are identified by their parameters $(\zeta_1, \ldots, \zeta_{N-1})$. The solution is then $\phi(\boldsymbol{x}) = \phi(s; \zeta_1, \ldots, \zeta_{N-1})$. If two or more characteristics cross, the solution is multi-valued, or a shock has occurred.

### 9.9.2 Example

Consider the PDE

$$\phi_t + t^2\,\phi_x = -x\,\phi \quad . \tag{9.226}$$

We identify $a_1(t, x, \phi) = 1$ and $a_2(t, x, \phi) = t^2$, as well as $b(t, x, \phi) = -x\,\phi$. The characteristics are curves $\big(t(s), x(s)\big)$ satisfying

$$\frac{dt}{ds} = 1 \qquad , \qquad \frac{dx}{ds} = t^2 \quad . \tag{9.227}$$

The variation of $\phi$ along each of the characteristics is given by

$$\frac{d\phi}{ds} = -x\,\phi \quad . \tag{9.228}$$

The initial data are expressed parametrically as

$$t(s=0) = 0 \quad , \quad x(s=0) = \zeta \quad , \quad \phi(s=0) = f(\zeta) \quad . \tag{9.229}$$

We now solve for the characteristics. We have

$$\frac{dt}{ds} = 1 \quad \Rightarrow \quad t(s, \zeta) = s \quad . \tag{9.230}$$

It then follows that

$$\frac{dx}{ds} = t^2 = s^2 \quad \Rightarrow \quad x(s, \zeta) = \zeta + \tfrac{1}{3}s^3 \quad . \tag{9.231}$$

Finally, we have

$$\frac{d\phi}{ds} = -x\,\phi = -\big(\zeta + \tfrac{1}{3}s^3\big)\,\phi \quad \Rightarrow \quad \phi(s, \zeta) = f(\zeta)\,\exp\big(-\tfrac{1}{12}s^4 - s\zeta\big) \quad . \tag{9.232}$$

We may now eliminate $(\zeta, s)$ in favor of $(x, t)$, writing $s = t$ and $\zeta = x - \frac{1}{3}t^3$, yielding the solution

$$\phi(x, t) = \phi\big(x - \tfrac{1}{3}t^3, t = 0\big)\,\exp\big(\tfrac{1}{4}t^4 - xt\big) \quad . \tag{9.233}$$