Physics 140A/B
Lecture Notes on Thermodynamics and Statistical Mechanics
(A Work in Progress)

Daniel Arovas
Department of Physics
University of California, San Diego

December 15, 2025

Contents

Contents i

List of Tables xiii

List of Figures xv

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 . . . 11

1.4.1 Entropy and information theory . . . 11

1.4.2 Probability distributions from maximum entropy . . . 13

1.4.3 Continuous probability distributions . . . 17

1.5 General Aspects of Probability Distributions . . . 17
1.5.1 Discrete and continuous distributions . . . 17
1.5.2 Central limit theorem . . . 19
1.5.3 Moments and cumulants . . . 21
1.5.4 Multidimensional Gaussian integral . . . 22
1.6 Appendix : Bayesian Statistical Inference . . . 23
1.6.1 Frequentists and Bayesians . . . 23
1.6.2 Updating Bayesian priors . . . 23
1.6.3 Hyperparameters and conjugate priors . . . 25
1.6.4 The problem with priors . . . 26

2 Thermodynamics **29**
2.1 References . . . 29
2.2 What is Thermodynamics? . . . 30
2.2.1 Thermodynamic systems and state variables . . . 30
2.2.2 Mathematical interlude : exact and inexact differentials . . . 32
2.2.3 Equations of state . . . 34
2.2.4 Heat . . . 35
2.2.5 Work . . . 35
2.2.6 Pressure and temperature . . . 36
2.2.7 Standard temperature and pressure . . . 37
2.3 The Zeroth Law of Thermodynamics . . . 38
2.3.1 Gibbs phase rule . . . 39
2.4 The First Law of Thermodynamics . . . 40
2.4.1 Conservation of energy . . . 40
2.4.2 Single component systems . . . 41
2.4.3 Heat capacity . . . 42
2.4.4 Ideal gases . . . 43
2.4.5 Adiabatic transformations of ideal gases . . . 46

2.4.6 Adiabatic free expansion . . . 47
2.5 Heat Engines and the Second Law of Thermodynamics . . . 48
2.5.1 There's no free lunch so quit asking . . . 48
2.5.2 Engines and refrigerators . . . 49
2.5.3 Nothing beats a Carnot engine . . . 51
2.5.4 The Carnot cycle . . . 52
2.5.5 Carnot engine at maximum power output . . . 54
2.6 The Entropy . . . 56
2.6.1 Entropy and heat . . . 56
2.6.2 The Third Law of thermodynamics . . . 57
2.6.3 Entropy changes in cyclic processes . . . 57
2.6.4 Gibbs-Duhem relation . . . 58
2.6.5 Entropy for an ideal gas . . . 59
2.6.6 Example system . . . 61
2.6.7 Measuring the entropy of a substance . . . 63
2.7 Thermodynamic Potentials . . . 63
2.7.1 Energy E . . . 63
2.7.2 Helmholtz free energy F . . . 64
2.7.3 Enthalpy H . . . 65
2.7.4 Gibbs free energy G . . . 65
2.7.5 Grand potential Ω . . . 66
2.8 Maxwell Relations . . . 67
2.8.1 Relations deriving from $E(S,V,N)$ . . . 67
2.8.2 Relations deriving from $F(T,V,N)$ . . . 67
2.8.3 Relations deriving from $\mathsf{H}(S,p,N)$ . . . 68
2.8.4 Relations deriving from $G(T,p,N)$ . . . 68
2.8.5 Relations deriving from $\Omega(T,V,\mu)$ . . . 69
2.8.6 Generalized thermodynamic potentials . . . 70

2.9 Equilibrium and Stability . . . 70
2.9.1 Equilibrium . . . 70
2.9.2 Stability . . . 71
2.10 Applications of Thermodynamics . . . 74
2.10.1 Adiabatic free expansion revisited . . . 74
2.10.2 Energy and volume . . . 75
2.10.3 van der Waals equation of state . . . 76
2.10.4 General equation of state . . . 78
2.10.5 Thermodynamic response functions . . . 78
2.10.6 Joule effect: free expansion of a gas . . . 80
2.10.7 Throttling: the Joule-Thompson effect . . . 81
2.11 Phase Transitions and Phase Equilibria . . . 84
2.11.1 p - v - T surfaces . . . 85
2.11.2 The Clausius-Clapeyron relation . . . 86
2.11.3 Liquid-solid line in H_2O . . . 88
2.11.4 Slow melting of ice : a quasistatic but irreversible process . . . 90
2.11.5 Generalized Gibbs phase rule . . . 91
2.12 Entropy of Mixing and the Gibbs Paradox . . . 94
2.12.1 Computing the entropy of mixing . . . 94
2.12.2 Entropy and combinatorics . . . 95
2.12.3 Weak solutions and osmotic pressure . . . 97
2.12.4 Effect of impurities on boiling and freezing points . . . 99
2.12.5 Binary solutions . . . 101
2.13 Some Concepts in Thermochemistry . . . 109
2.13.1 Chemical reactions and the law of mass action . . . 109
2.13.2 Enthalpy of formation . . . 111
2.13.3 Bond enthalpies . . . 114
2.14 Appendix I : Integrating Factors . . . 116

2.15 Appendix II : Other Thermodynamic Cycles 118
2.15.1 The Stirling cycle 118
2.15.2 The Otto and Diesel cycles 119
2.15.3 The Joule-Brayton cycle 122
2.16 Appendix III : Legendre Transformations 123
2.17 Appendix IV : Useful Mathematical Relations 126

3 Ergodicity and the Approach to Equilibrium 131
3.1 References 131
3.2 Modeling the Approach to Equilibrium 132
3.2.1 Equilibrium 132
3.2.2 The master equation 132
3.2.3 Equilibrium distribution and detailed balance 133
3.2.4 Boltzmann's $\mathcal{H}$-theorem 134
3.2.5 Markov chains 135
3.3 Phase Flows in Classical Mechanics 137
3.3.1 Hamiltonian evolution 137
3.3.2 Dynamical systems and the evolution of phase space volumes 138
3.3.3 Liouville's equation and the microcanonical distribution 141
3.4 Irreversibility and Poincaré Recurrence 142
3.4.1 Poincaré recurrence theorem 143
3.4.2 Kac ring model 145
3.5 Remarks on Ergodic Theory 149
3.5.1 Definition of ergodicity 149
3.5.2 The microcanonical ensemble 150
3.5.3 Ergodicity and mixing 150
3.6 Appendix I : Normal matrices and Jordan canonical form 155
3.6.1 Contrast with singular value decomposition 156
3.6.2 Example 156

3.7 Appendix II : Formal Solution of the Master Equation 157
  3.7.1 Detailed balance 159
3.8 Appendix III : Poisson Processes and Radioactive Decay 159
  3.8.1 Poisson process 159
  3.8.2 Radioactive decay 160
3.9 Appendix IV : Transition to Ergodicity in a Simple Model 161
3.10 Appendix V : Thermalization of Quantum Systems 169
  3.10.1 Quantum dephasing 169
  3.10.2 Eigenstate thermalization hypothesis 171
  3.10.3 Entanglement and entanglement entropy 172
  3.10.4 Singular value decomposition and entanglement 173
  3.10.5 Digression on random matrices 175
  3.10.6 More precise formulation of ETH 178
  3.10.7 When is the ETH true? 179

4 Statistical Ensembles **183**
4.1 References 183
4.2 Microcanonical Ensemble (μCE) 184
  4.2.1 The microcanonical distribution function 184
  4.2.2 Density of states 185
  4.2.3 Arbitrariness in the definition of $S(E)$ 188
  4.2.4 Ultra-relativistic ideal gas 188
  4.2.5 Discrete systems 188
  4.2.6 Two systems in thermal contact 189
4.3 The Quantum Mechanical Trace 192
  4.3.1 The density matrix 192
  4.3.2 Averaging the DOS 193
  4.3.3 Coherent states 194
4.4 Ordinary Canonical Ensemble (OCE) 196

4.4.1 Canonical distribution and partition function . . . 196
4.4.2 The difference between $P(E_n)$ and P_n . . . 197
4.4.3 Additional remarks . . . 197
4.4.4 Averages within the OCE . . . 198
4.4.5 Entropy and free energy . . . 198
4.4.6 Fluctuations in the OCE . . . 199
4.4.7 Thermodynamics revisited . . . 199
4.4.8 Generalized susceptibilities . . . 201
4.5 Grand Canonical Ensemble (GCE) . . . 201
4.5.1 Grand canonical distribution and partition function . . . 201
4.5.2 Entropy and Gibbs-Duhem relation . . . 203
4.5.3 Fluctuations in the GCE . . . 204
4.5.4 Generalized susceptibilities in the GCE . . . 204
4.5.5 Gibbs ensemble . . . 205
4.6 Statistical Ensembles from Maximum Entropy . . . 206
4.6.1 Microcanonical ensemble . . . 206
4.6.2 Ordinary canonical ensemble . . . 207
4.6.3 Grand canonical ensemble . . . 207
4.6.4 Generalized Gibbs ensembles . . . 208
4.7 Ideal Gas Statistical Mechanics . . . 209
4.7.1 Maxwell velocity distribution . . . 210
4.7.2 Equipartition . . . 211
4.8 Selected Examples . . . 212
4.8.1 Spins in an external magnetic field . . . 212
4.8.2 Negative temperature (!) . . . 214
4.8.3 Adsorption . . . 215
4.8.4 Dipole gas . . . 216
4.8.5 Elasticity of wool . . . 218

4.9 Statistical Mechanics of Molecular Gases 220
4.9.1 Separation of translational and internal degrees of freedom 220
4.9.2 The internal coordinate partition function 222
4.9.3 Quantum theory of rotations 222
4.9.4 Vibrations 224
4.9.5 Two-level systems : Schottky anomaly 225
4.9.6 Electronic and nuclear excitations 226
4.10 Appendix I : Constrained Extremization of Functions 228
4.10.1 Example : volume of a cylinder 229
4.11 Appendix II : Additional Examples of Statistical Ensembles 230
4.11.1 Noninteracting spin dimers 230
4.11.2 Three state system 231
4.11.3 Spins and vacancies on a surface 232
4.11.4 Fluctuating interface 234
4.11.5 Dissociation of atomic hydrogen 236

5 Noninteracting Quantum Systems **239**
5.1 References 239
5.2 Statistical Mechanics of Noninteracting Quantum Systems 240
5.2.1 Bose and Fermi systems in the grand canonical ensemble 240
5.2.2 Quantum statistics and the Maxwell-Boltzmann limit 242
5.2.3 Single particle density of states 243
5.3 Expansion in Powers of the Density 244
5.3.1 Expansion in powers of the fugacity 244
5.3.2 Ballistic dispersion 245
5.4 Photon Statistics 246
5.4.1 Thermodynamics of the photon gas 246
5.4.2 Classical arguments for the photon gas 248
5.4.3 Surface temperature of the earth 249

5.4.4 Distribution of blackbody radiation 249
5.4.5 What if the sun emitted ferromagnetic spin waves? 251
5.5 The Ideal Bose Gas 251
5.5.1 General formulation for noninteracting systems 251
5.5.2 Ballistic dispersion 253
5.5.3 Isotherms for the ideal Bose gas 258
5.5.4 Properties of liquid ^{4}He 259
5.5.5 The Fermi distribution 264
5.5.6 $T = 0$ and the Fermi surface 264
5.5.7 The Sommerfeld expansion 267
5.5.8 Magnetic susceptibility 272

6 Classical Interacting Systems **277**
6.1 References 277
6.2 Nonideal Classical Gases 278
6.2.1 The configuration integral 278
6.2.2 One-dimensional Tonks gas 278
6.2.3 Mayer cluster expansion 280
6.2.4 Examples 286
6.3 Coulomb Systems : Plasmas and the Electron Gas 290
6.3.1 Electrostatic potential 290
6.3.2 Debye-Hückel theory 290
6.3.3 The electron gas : Thomas-Fermi screening 293

7 Mean Field Theory of Phase Transitions **297**
7.1 References 297
7.2 The van der Waals system 298
7.2.1 Equation of state 298
7.2.2 Analytic form of the coexistence curve near the critical point 302

7.2.3 History of the van der Waals equation 304

7.3 Fluids, Magnets, and the Ising Model 306

7.3.1 Lattice gas description of a fluid 306

7.3.2 Phase diagrams and critical exponents 307

7.3.3 Gibbs-Duhem relation for magnetic systems 309

7.3.4 Order-disorder transitions 310

7.4 Mean Field Theory 311

7.4.1 The mean field *Ansatz* 311

7.4.2 Zero external field 313

7.4.3 Finite external field 315

7.4.4 Magnetization dynamics 318

7.4.5 Beyond nearest neighbors 320

7.4.6 Ising model with long-ranged forces 321

7.5 Landau Theory of Phase Transitions 322

7.5.1 Quartic free energy with Ising symmetry 322

7.5.2 Cubic terms in Landau theory : first order transitions 324

7.5.3 Order parameter dynamics 326

7.5.4 Sixth order Landau theory : tricritical point 327

7.5.5 Hysteresis for the sextic potential 330

7.6 Variational Density Matrix Method 331

7.6.1 The variational principle 331

7.6.2 Variational density matrix for the Ising model 333

7.6.3 q-state Potts model 336

7.6.4 XY Model 338

7.7 Mean Field Theory of Fluctuations 341

7.7.1 Correlation and response in mean field theory 341

7.7.2 Calculation of the response functions 342

8 The Boltzmann Equation **347**

8.1 References . . . 347
8.2 Equilibrium, Nonequilibrium and Local Equilibrium . . . 348
8.3 Boltzmann Transport Theory . . . 350
  8.3.1 Derivation of the Boltzmann equation . . . 350
  8.3.2 Collisionless Boltzmann equation . . . 351
  8.3.3 Collisional invariants . . . 353
  8.3.4 Scattering processes . . . 354
  8.3.5 Detailed balance . . . 356
  8.3.6 Kinematics and cross section . . . 357
  8.3.7 $\mathcal{H}$-theorem . . . 358
8.4 Weakly Inhomogeneous Gas . . . 361
8.5 Relaxation Time Approximation . . . 363
  8.5.1 Approximation of collision integral . . . 363
  8.5.2 Computation of the scattering time . . . 363
  8.5.3 Thermal conductivity . . . 364
  8.5.4 Viscosity . . . 366
  8.5.5 Oscillating external force . . . 368
  8.5.6 Quick and dirty calculation of transport coefficiencs . . . 369
  8.5.7 Thermal diffusivity, kinematic viscosity, and Prandtl number . . . 370
8.6 The Equations of Hydrodynamics . . . 371

List of Tables

2.1 Specific heat of some common substances . . . 42
2.2 Performances of real heat engines . . . 55
2.3 Van der Waals parameters for some common gases . . . 76
2.4 Latent heats of fusion and vaporization at $p = 1$ atm . . . 100
2.5 Enthalpies of formation of some common substances . . . 112
2.6 Average bond enthalpies for some common bonds . . . 114

3.1 States and their degeneracies. . . . 136
3.2 Comparison of time and microcanonical averages (I) . . . 167
3.3 Comparison of time and microcanonical averages (II) . . . 169

4.1 Rotational and vibrational temperatures of common molecules . . . 223
4.2 Nuclear angular momentum states for homonuclear diatomic molecules . . . 228

7.1 van der Waals parameters for some common gases . . . 300
7.2 Critical exponents . . . 317

8.1 Viscosities, thermal conductivities, and Prandtl numbers for some common gases . . . 371

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 31
2.2 Gas pressure as a space-time averaged quantity 32
2.3 Two distinct paths with identical endpoints 33
2.4 The constant volume gas thermometer 37
2.5 The phase diagram of H_2O 38
2.6 Convergence of measurements in the gas thermometer 39
2.7 The first law of thermodynamics 40
2.8 C_V for one mole of H_2 gas 44
2.9 Molar heat capacities c_V for three solids 45
2.10 Adiabatic free expansion 48
2.11 A perfect engine 49
2.12 Heat engine and refrigerator 50
2.13 A wonder engine driving a Carnot refrigerator 51
2.14 The Carnot cycle 52
2.15 Decomposition of a thermodynamic loop into Carnot cycles 57
2.16 Check for instability 72
2.17 Adiabatic free expansion via a thermal path 74

2.18 Throttling . . . 82
2.19 Temperature inversion for the van der Waals gas . . . 83
2.20 Phase diagrams of single component systems . . . 84
2.21 $p(v, T)$ surface for the ideal gas equation of state . . . 85
2.22 A p-v-T surface for a substance which contracts upon freezing . . . 86
2.23 p-v-T surfaces for a substance which expands upon freezing . . . 87
2.24 Phase diagram for CO_2 in the (p, T) plane . . . 88
2.25 Projection of the p-v-T surface onto the (v, p) plane . . . 89
2.26 Surface melting data for ice . . . 90
2.27 Multicomponent system . . . 95
2.28 Mixing among three different species of particles . . . 96
2.29 Osmotic pressure . . . 99
2.30 Gibbs free energy for a binary solution . . . 102
2.31 Binary systems: chemical potential shifts and (T, μ) phase diagrams . . . 103
2.32 Phase diagram for the binary system in the (x, T) plane . . . 104
2.33 Gibbs free energy for an ideal binary solution . . . 105
2.34 Phase diagrams for azeotropes . . . 107
2.35 Binary fluid mixture in equilibrium with a vapor . . . 108
2.36 Eutectic phase diagram . . . 109
2.37 Reaction enthalpy . . . 113
2.38 Reaction enthalpy for hydrogenation of ethene . . . 115
2.39 The Stirling cycle . . . 119
2.40 The Otto cycle . . . 120
2.41 The Diesel cycle . . . 121
2.42 The Joule-Brayton cycle . . . 123
2.43 The Legendre transformation . . . 125

3.1 Time evolution of two immiscible fluids . . . 140
3.2 Projected evolution of a set $\mathcal{R}_0$ under the mapping g_τ . . . 144

3.3 Poincaré recurrence . . . 145
3.4 The Kac ring model . . . 146
3.5 Simulation of the Kac ring model . . . 147
3.6 More simulations of the Kac ring model . . . 148
3.7 An ergodic flow which is not mixing . . . 151
3.8 The baker's transformation . . . 152
3.9 Multiply iterated baker's transformation . . . 153
3.10 Arnold's cat map . . . 154
3.11 Poincaré sections for the ball and piston problem . . . 166
3.12 Long time averages of $X_{\rm av}(t) \equiv t^{-1}\int_0^t dt'\, X(t')$ . . . 167
3.13 Hierarchy of dynamical systems . . . 170

4.1 Complex integration contours $\mathcal{C}$ for inverse Laplace transform . . . 186
4.2 Two systems in thermal contact . . . 190
4.3 A system S in contact with a 'world' W . . . 193
4.4 Averaging the quantum mechanical discrete density of states . . . 194
4.5 Microscopic interpretation of the First Law of Thermodynamics . . . 200
4.6 Maxwell distribution of speeds $\varphi(v/v_0)$ . . . 211
4.7 Entropy *versus* energy for a spin system . . . 215
4.8 A crude model for the elasticity of wool . . . 218
4.9 Length *versus* temperature and force . . . 219
4.10 Heat capacity per molecule as a function of temperature . . . 226
4.11 Noninteracting spin dimers on a lattice . . . 231

5.1 Spectral density of blackbody radiation . . . 250
5.2 The Bose distribution $n^{\rm B}_\varepsilon(T,\mu)$ at fixed μ . . . 253
5.3 The polylogarithm function ${\rm Li}_s(z)$ . . . 254
5.4 Molar heat capacity of the ideal Bose gas . . . 256
5.5 Phase diagrams for the ideal Bose gas . . . 257

5.6 Phase diagram of ^{4}He 260
5.7 Specific heat of liquid ^{4}He in the vicinity of the λ-transition 261
5.8 The fountain effect 261
5.9 Strange properties of superfluid ^{4}He 263
5.10 The Fermi distribution 264
5.11 Fermi surfaces for two and three-dimensional structures 266
5.12 Deformation of the complex integration contour in eqn. 5.138 269
5.13 Fermi distributions in the presence of a magnetic field 273

6.1 The Lennard-Jones potential 280
6.2 Diagrams and the Mayer cluster expansion 282
6.3 Vertex labels in the configuration integral 283
6.4 Symmetry factors for cluster diagrams 283
6.5 Connected *versus* irreducible clusters 286
6.6 The overlap of hard sphere Mayer functions 287
6.7 Mayer function for an attractive spherical well with a repulsive core 289
6.8 The Thomas-Fermi atom 295

7.1 Pressure *versus* volume for the van der Waals gas 299
7.2 Molar free energy $f(T,v)$ of the van der Waals system 301
7.3 Pressure-volume isotherms for the van der Waals system, with Maxwell construction . . 302
7.4 Maxwell construction in the (v,p) plane 303
7.5 Universality of the liquid-gas transition 304
7.6 The lattice gas model 307
7.7 Comparison of liquid-gas and Ising ferromagnet phase diagrams 308
7.8 Order-disorder transition on the square lattice 310
7.9 Ising mean field theory at $h=0$ 313
7.10 Ising mean field theory at $h=0.1$ 315
7.11 Dissipative magnetization dynamics $\dot{m}=-f'(m)$ 319
7.12 Magnetization hysteresis 320

7.13 Phase diagram for the quartic Landau free energy 323

7.14 Behavior of the quartic free energy $f(m) = \frac{1}{2}am^2 - \frac{1}{3}ym^3 + \frac{1}{4}bm^4$ 325

7.15 Fixed points for $\varphi(u) = \frac{1}{2}ru^2 - \frac{1}{3}u^3 + \frac{1}{4}u^4$ and flow $\dot{u} = -\varphi'(u)$ 327

7.16 Behavior of the sextic free energy $f(m) = \frac{1}{2}am^2 + \frac{1}{4}bm^4 + \frac{1}{6}cm^6$ 328

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 329

7.18 Fixed points $\varphi'(u^*) = 0$ and flow $\dot{u} = -\varphi'(u)$ for the sextic potential 331

7.19 Variational field free energy *versus* magnetization 335

8.1 Level sets for a sample $f(\bar{x},\bar{p},\bar{t}) = C\,e^{-\frac{1}{2}(\bar{x}-\bar{p}\bar{t})^2}\,e^{-\frac{1}{2}\bar{p}^2}$ 353

8.2 One and two particle scattering processes 355

8.3 Graphic representation of the equation $n\,\sigma\,\bar{v}_{\rm rel}\,\tau = 1$ 364

8.4 *Gedankenexperiment* to measure shear viscosity η in a fluid 366

8.5 Experimental data on thermal conductivity and shear viscosity 368

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* (3rd 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, 3rd 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^{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

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

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

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

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:

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

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

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

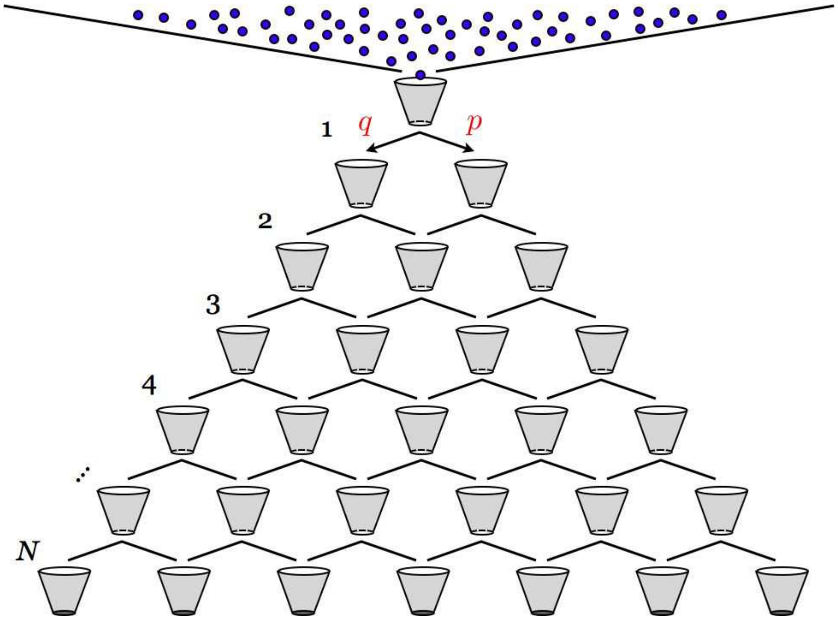


Figure 1.1: The falling ball system, which mimics a one-dimensional random walk.

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 $\mathrm{Var}(X) = 4Np\,q$. The *root mean square* deviation, ΔX_{rms}, is the square root of the variance: $\Delta X_{\mathrm{rms}} = \sqrt{\mathrm{Var}(X)}$. 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 ΔX_{rms} are proportional to $N^{1/2}$. In the limit $N \to \infty$ then, the ratio $\Delta X_{\mathrm{rms}}/\langle X\rangle$ 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

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

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

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

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

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

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 $\ln P_{N,X}$ using Stirling's asymptotic expansion

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

We then have

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

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

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

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

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

where $\mathcal{C}$ is a normalization constant[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

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

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 $\mathcal{O}(\ln N)$ and $\mathcal{O}(N^0)$ terms in the asymptotic expansion of $\ln N!$. We have ignored these terms here. Accounting for them carefully reproduces the correct value of $\mathcal{C}$ in eqn. 1.20.

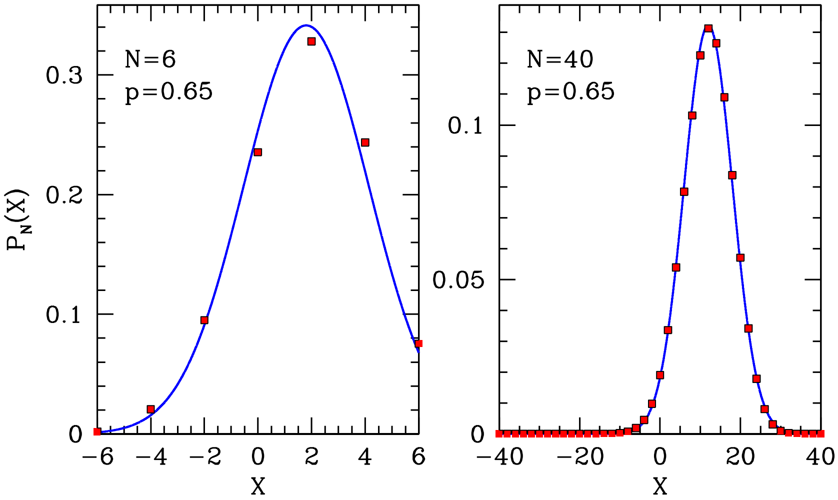


Figure 1.2: Comparison of exact distribution of eqn. 1.10 (red squares) with the Gaussian distribution of eqn. 1.19 (blue line).

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

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

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

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

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_{B}, which is the entropy in units of Boltzmann's constant. And what we have called E here is really $E/k_{\mathrm{B}}T$, 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 ω 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 elements (a, b) where $a \in A$ and $b \in B$.

Some details: If ω 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 ω such that $\omega \in A$ and $\omega \notin B$.

In probability theory, each object ω is identified as an *event*. We denote by Ω 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 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 meaning of $P(A)$ is that if events ω are chosen from Ω *at random*, 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 of a *probability space* Ω as well as a *probability measure* P. For example, in the microcanonical ensemble of classical statistical physics, the space Ω 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 $A \in \Omega$ the probability of A is $P(A) = \int d\mu\, \chi_A(\varphi)$, where $\chi_A(\varphi) = 1$ if $\varphi \in A$ and $\chi_A(\varphi) = 0$ if $\varphi \notin A$ 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, $P(A \cap B) = \text{Prob}[\omega \in A \text{ and } \omega \in B]$. 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 with an IQ greater than 60. Then $A \cap B$ would be the set of all politicians who are also American citizens, *etc. Exercise: estimate* $P(A \cap B \cap C)$.

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 $P(A\,|\,B)$. Applying eqn. 1.27 with

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

we find $P(A\,|\,B)=\frac{1}{2}$. 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 $P(B\,|\,A)+P(B\,|\,\neg A)=1$, but this is not generally the case. What is true instead is $P(B\,|\,A)+P(\neg B\,|\,A)=1$. 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, $\textsf{se}=P(B\,|\,A)$ is the sensitivity and $\textsf{sp}=P(\neg B\,|\,\neg A)$ is the specificity. We then have $P(B\,|\,\neg A)=1-P(\neg B\,|\,\neg A)$. Therefore,

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

In our previous example, $\textsf{se}=\textsf{sp}=0.9999$, 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

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

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

$$P(y)=\int\!dx\;P(y\,|\,x)\,P(x) \tag{1.30}$$

and

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

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 $P(A_i\,|\,B)$ 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. The average of any function $f(x)$ is denoted as $\mathbb{E}f$ or $\langle f\rangle$, and is defined for discrete sets as

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

where $P(x)$ is the probability of x. For continuous sets, we have

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

Typically for continuous sets we have $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{R}_{\geq 0}$. Gardiner and other authors introduce an extra symbol, X, to denote a *random variable*, with $X(x) = x$ being its value. This is formally useful but notationally confusing, so we'll avoid it here and speak loosely of x as a random variable.

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

$$\mathbb{E} f(x,y) = \langle f(x,y) \rangle = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} f(x,y)\, P(x,y) \quad , \tag{1.34}$$

with the obvious generalization to continuous sets. This generalizes to higher rank products, *i.e.* $x_i \in \mathcal{X}_i$ with $i \in \{1, \ldots, N\}$. The *covariance* of x_i and x_j is defined as

$$C_{ij} \equiv \big\langle \big(x_i - \langle x_i \rangle\big)\big(x_j - \langle x_j \rangle\big) \big\rangle = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \quad . \tag{1.35}$$

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

$$\mathbb{E} f(x) \geq f(\mathbb{E} x) \quad . \tag{1.36}$$

For continuous functions, $f(x)$ is convex if $f''(x) \geq 0$ everywhere[3]. If $f(x)$ is convex on some interval $[a,b]$ then for $x_{1,2} \in [a,b]$ we must have

$$f\big(\lambda x_1 + (1-\lambda) x_2\big) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \quad , \tag{1.37}$$

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

$$f\Big(\sum_n p_n x_n\Big) \leq \sum_n p_n f(x_n) \quad , \tag{1.38}$$

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

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

[4] 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`.

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

$$I(p) = -A\,\log p \quad , \tag{1.39}$$

where A is an arbitrary constant that can be absorbed into the base of the logarithm, since $\log_b x = \ln x / \ln b$. In statistical physics we will take $A = 1$ and use e as the base, so $I(p) = -\ln p$. 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[5].

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

$$S = \frac{\langle I_N \rangle}{N} = -\sum_n p_n \log p_n \quad , \tag{1.40}$$

which is known as the entropy of the distribution. Thus, maximizing S is equivalent to maximizing the *information* content per observation. When the logarithm is taken base two, S is known as the *Shannon entropy* of the distribution $\{p_n\}$.

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 $S = \log_2 \varGamma$, where $\varGamma$ is the total size of our discrete space of states, since $p_n = 1/\varGamma$. 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

$$g(N) = \frac{N!}{\prod_{n=1}^K N_n!} \quad . \tag{1.41}$$

---

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

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

$$\log_2 g(N) = \log_2(N!) - \sum_{n=1}^{K} \log_2(N_n!) \approx -N \sum_{n=1}^{K} p_n \log_2 p_n \quad , \tag{1.42}$$

up to terms of order unity. 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 $\log_2 g(N) = N \log_2 K$.

1.4.2 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.

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

$$\sum_{\{\sigma_j\}} P(\sigma_1, \ldots, \sigma_N) = 1 \quad , \tag{1.43}$$

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

$$P(\sigma_1, \ldots, \sigma_N) = \prod_{j=1}^{N} \big(p\, \delta_{\sigma_j,1} + q\, \delta_{\sigma_j,-1} \big) = \prod_{j=1}^{N} p^{(1+\sigma_j)/2}\, q^{(1-\sigma_j)/2} \quad , \tag{1.44}$$

corresponding to N independent steps?

The principle of maximum entropy

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

$$S = -\sum_n p_n \ln p_n \quad , \tag{1.45}$$

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 $\big\{p^{\mathsf{A}}_a\big\}$ and $\big\{p^{\mathsf{B}}_b\big\}$. The joint probability for events a and b is then $P_{a,b} = p^{\mathsf{A}}_a\, p^{\mathsf{B}}_b$. The entropy of the joint distribution is then

$$\begin{aligned} S &= -\sum_a \sum_b P_{a,b} \ln P_{a,b} = -\sum_a \sum_b p^{\mathsf{A}}_a\, p^{\mathsf{B}}_b \ln \big(p^{\mathsf{A}}_a\, p^{\mathsf{B}}_b\big) = -\sum_a \sum_b p^{\mathsf{A}}_a\, p^{\mathsf{B}}_b \big(\ln p^{\mathsf{A}}_a + \ln p^{\mathsf{B}}_b \big) \\ &= -\sum_a p^{\mathsf{A}}_a \ln p^{\mathsf{A}}_a \cdot \sum_b p^{\mathsf{B}}_b - \sum_b p^{\mathsf{B}}_b \ln p^{\mathsf{B}}_b \cdot \sum_a p^{\mathsf{A}}_a = -\sum_a p^{\mathsf{A}}_a \ln p^{\mathsf{A}}_a - \sum_b p^{\mathsf{B}}_b \ln p^{\mathsf{B}}_b = S^{\mathsf{A}} + S^{\mathsf{B}} \quad . \end{aligned}$$

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

$$S^*\big(\{p_n\},\lambda\big) = -\sum_n p_n \ln p_n - \lambda\Big(\sum_n p_n - 1\Big) \tag{1.46}$$

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

$$\begin{aligned} 0 &= \frac{\partial S^*}{\partial p_n} = -\big(\ln p_n + 1 + \lambda\big) \\ 0 &= \frac{\partial S^*}{\partial \lambda} = \sum_n p_n - 1 \quad . \end{aligned} \tag{1.47}$$

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

$$\sum_n p_n = e^{-(1+\lambda)} \cdot \sum_n 1 = \mathit{\Gamma}\, e^{-(1+\lambda)} \quad , \tag{1.48}$$

where $\mathit{\Gamma} \equiv \sum_n 1$ is the total number of possible events. Thus, $p_n = \mathit{\Gamma}^{-1}$, 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

$$S^*\big(\{p_n\},\lambda_0,\lambda_1\big) = -\sum_n p_n \ln p_n - \lambda_0\Big(\sum_n p_n - 1\Big) - \lambda_1\Big(\sum_n X_n\, p_n - X\Big) \quad . \tag{1.49}$$

We then have

$$\frac{\partial S^*}{\partial p_n} = -\big(\ln p_n + 1 + \lambda_0 + \lambda_1\, X_n\big) = 0 \quad , \tag{1.50}$$

which yields the two-parameter distribution

$$p_n = e^{-(1+\lambda_0)}\, e^{-\lambda_1 X_n} \quad . \tag{1.51}$$

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 λ_0 and λ_1, respectively:

$$\begin{aligned} 1 &= e^{-(1+\lambda_0)} \sum_n e^{-\lambda_1 X_n} \\ X &= e^{-(1+\lambda_0)} \sum_n X_n\, e^{-\lambda_1 X_n} \quad . \end{aligned} \tag{1.52}$$

General formulation

The generalization to K extra pieces of information (plus normalization) is immediately apparent. We have

$$X^a = \sum_n X^a_n\, p_n \quad , \tag{1.53}$$

and therefore we define

$$S^*\big(\{p_n\},\{\lambda_a\}\big) = -\sum_n p_n \ln p_n - \sum_{a=0}^{K} \lambda_a \Big(\sum_n X^a_n\, p_n - X^a\Big) \quad , \tag{1.54}$$

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

$$\begin{aligned} p_n &= \exp\bigg\{ -1 - \sum_{a=0}^{K} \lambda_a\, X^a_n \bigg\} \\ &= \frac{1}{Z} \exp\bigg\{ -\sum_{a=1}^{K} \lambda_a\, X^a_n \bigg\} \quad , \end{aligned} \tag{1.55}$$

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. 1.53.

Example

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

$$\begin{aligned} \sum_{\sigma_1} \cdots \sum_{\sigma_N} P(\sigma_1, \ldots, \sigma_N) &= 1 \\ \sum_{\sigma_1} \cdots \sum_{\sigma_N} P(\sigma_1, \ldots, \sigma_N) \sum_{j=1}^{N} \sigma_j &= X \quad . \end{aligned} \tag{1.56}$$

Here the discrete label n from §1.4.2 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

$$P(\sigma_1, \ldots, \sigma_N) = \mathcal{C} \, \exp\bigg\{ -\lambda \sum_j \sigma_j \bigg\} = \mathcal{C} \prod_{j=1}^{N} e^{-\lambda\, \sigma_j} \quad , \tag{1.57}$$

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

$$\mathsf{Tr}\; P \equiv \sum_{\{\sigma_j\}} P(\sigma_1, \ldots, \sigma_N) = \mathcal{C}\, \big(e^{\lambda} + e^{-\lambda}\big)^N \quad , \tag{1.58}$$

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

$$P(\sigma_1, \ldots, \sigma_N) = \prod_{j=1}^{N} \frac{e^{-\lambda\sigma_j}}{e^{\lambda} + e^{-\lambda}} = \prod_{j=1}^{N} \big(p\, \delta_{\sigma_j,1} + q\, \delta_{\sigma_j,-1} \big) \quad , \tag{1.59}$$

where

$$p = \frac{e^{-\lambda}}{e^{\lambda} + e^{-\lambda}} \quad , \quad q = 1 - p = \frac{e^{\lambda}}{e^{\lambda} + e^{-\lambda}} \quad . \tag{1.60}$$

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

Of course there are no miracles[6], 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

$$P(\sigma_1, \ldots, \sigma_N) = \frac{1}{Z} \exp\left\{ -\lambda_1 \sum_{j=1}^{N} \sigma_j - \lambda_2 \sum_{j=1}^{N-1} \sigma_j \, \sigma_{j+1} \right\} \quad , \tag{1.61}$$

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 $R_{ss'} = e^{-\lambda_1 (s+s')/2} \, e^{-\lambda_2 s s'}$ with $s, s' = \pm 1$,

$$\begin{aligned} R &= \begin{pmatrix} e^{-\lambda_1 - \lambda_2} & e^{\lambda_2} \\ e^{\lambda_2} & e^{\lambda_1 - \lambda_2} \end{pmatrix} \\ &= e^{-\lambda_2} \cosh(\lambda_1) \, \mathbb{I} - e^{-\lambda_2} \sinh(\lambda_1) \, \tau^z + e^{\lambda_2} \, \tau^x \quad , \end{aligned} \tag{1.62}$$

where τ^x and τ^z are Pauli matrices, we have that

$$Z_{\text{ring}} = \mathsf{Tr}\left(R^N\right) \quad , \quad Z_{\text{chain}} = \mathsf{Tr}\left(R^{N-1} S\right) \quad , \tag{1.63}$$

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

$$S = \begin{pmatrix} e^{-\lambda_1} & 1 \\ 1 & e^{\lambda_1} \end{pmatrix} = \cosh(\lambda_1) \, \mathbb{I} - \sinh(\lambda_1) \, \tau^z + \tau^x \quad . \tag{1.64}$$

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 $Z_{\text{ring}} = \zeta_+^N + \zeta_-^N$, where $\zeta_\pm$ are the eigenvalues of R:

$$\zeta_\pm = e^{-\lambda_2} \cosh\lambda_1 \pm \sqrt{e^{-2\lambda_2} \sinh^2\!\lambda_1 + e^{2\lambda_2}} \quad . \tag{1.65}$$

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

$$X = \left\langle \sum_{j=1}^{N} \sigma_j \right\rangle = -\frac{\partial \ln Z}{\partial \lambda_1} = -\frac{N \sinh\lambda_1}{\sqrt{\sinh^2\!\lambda_1 + e^{4\lambda_2}}} \quad . \tag{1.66}$$

We also have $E = -\partial \ln Z / \partial \lambda_2$. 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 $\varepsilon = E/N$, 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. 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).

---

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

1.4.3 Continuous probability distributions

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

$$X^a = \int\limits_{\Omega} d\mu\, X^a(\boldsymbol{\varphi})\, P(\boldsymbol{\varphi}) \quad , \tag{1.67}$$

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 Ω. Then we extremize the functional

$$S^*\big[P(\boldsymbol{\varphi}), \{\lambda_a\}\big] = -\int\limits_{\Omega} d\mu\, P(\boldsymbol{\varphi}) \ln P(\boldsymbol{\varphi}) - \sum_{a=0}^{K} \lambda_a \Bigg(\int\limits_{\Omega} d\mu\, P(\boldsymbol{\varphi})\, X^a(\boldsymbol{\varphi}) - X^a \Bigg) \tag{1.68}$$

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:

$$\ln P(\boldsymbol{\varphi}) = -1 - \sum_{a=0}^{K} \lambda_a\, X^a(\boldsymbol{\varphi}) \quad . \tag{1.69}$$

The $(K+1)$ Lagrange multipliers $\{\lambda_a\}$ are then determined from the $(K+1)$ constraint equations in eqn. 1.67.

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

$$\int\limits_{-\infty}^{\infty} dx\; P(x) = 1 \quad , \quad \int\limits_{-\infty}^{\infty} dx\; x\, P(x) = \mu \quad , \quad \int\limits_{-\infty}^{\infty} dx\; x^2\, P(x) = \mu^2 + \sigma^2 \quad . \tag{1.70}$$

Extremizing the entropy, we then obtain

$$P(x) = \mathcal{C}\, e^{-\lambda_1 x - \lambda_2 x^2} \quad , \tag{1.71}$$

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

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2} \quad . \tag{1.72}$$

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

$$\sum_{n\in\mathcal{C}} P_n = 1 \quad . \tag{1.73}$$

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

$$\langle A\rangle = \sum_{n\in\mathcal{C}} P_n\, A_n \quad . \tag{1.74}$$

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. 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 Ω is the phase space (or configuration space) of the system. Let $d\mu$ be a *measure* on this space. In general, we can write

$$d\mu = W(\varphi_1, \ldots, \varphi_n)\, d\varphi_1\, d\varphi_2 \cdots d\varphi_n \quad . \tag{1.75}$$

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

$$d\mu = \mathcal{C} \prod_{\sigma=1}^{r} dq_\sigma\, dp_\sigma \quad , \tag{1.76}$$

where $\mathcal{C}$ is a constant we shall discuss later on below[7].

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

$$\int_\Omega d\mu\, P(\boldsymbol{\varphi}) = 1 \quad . \tag{1.77}$$

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

$$\langle A\rangle = \int_\Omega d\mu\, P(\boldsymbol{\varphi})\, A(\boldsymbol{\varphi}) \quad . \tag{1.78}$$

For example, consider the Gaussian distribution

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2} \quad . \tag{1.79}$$

---

[7] 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*.

From the result[8]

$$\int\limits_{-\infty}^{\infty} dx\, e^{-\alpha x^2}\, e^{-\beta x} = \sqrt{\frac{\pi}{\alpha}}\, e^{\beta^2/4\alpha} \quad , \tag{1.80}$$

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

$$\begin{aligned} \langle x \rangle &= \mu \\ \langle x^2 \rangle - \langle x \rangle^2 &= \sigma^2 \quad . \end{aligned} \tag{1.81}$$

We call μ the *mean* and σ the *standard deviation* of the distribution, eqn. 1.79.

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 infinitesimal – 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 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 ϕ. We then have

$$\begin{aligned} P(\phi) &= \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P_N(x_1, \ldots, x_N)\, \delta\Big(\phi(x_1, \ldots, x_N) - \phi\Big) \\ &= \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P_1(x_1) \cdots P_1(x_N)\, \delta\Big(\phi(x_1, \ldots, x_N) - \phi\Big) \quad , \end{aligned} \tag{1.82}$$

where the second line is appropriate if the $\{x_j\}$ are themselves distributed independently. Note that

$$\int\limits_{-\infty}^{\infty} d\phi\, P(\phi) = 1 \quad , \tag{1.83}$$

so $P(\phi)$ is itself normalized.

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

$$P_N(X) = \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P_1(x_1) \cdots P_1(x_N)\, \delta\big(x_1 + x_2 + \ldots + x_N - X\big) \quad . \tag{1.84}$$

---

[8] Memorize this!

It is convenient to compute the Fourier transform[9] of $P(X)$:

$$\begin{aligned}\hat{P}_N(k) &= \int\limits_{-\infty}^{\infty} dX\, P_N(X)\, e^{-ikX} \\ &= \int\limits_{-\infty}^{\infty} dX \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P_1(x_1) \cdots P_1(x_N)\, \delta\big(x_1 + \ldots + x_N - X\big)\, e^{-ikX} = \big[\hat{P}_1(k)\big]^N \quad ,\end{aligned} \tag{1.85}$$

where

$$\hat{P}_1(k) = \int\limits_{-\infty}^{\infty} dx\, P_1(x)\, e^{-ikx} \tag{1.86}$$

is the Fourier transform 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 $\hat{P}_1(k)$

$$\begin{aligned}\hat{P}_1(k) &= \int\limits_{-\infty}^{\infty} dx\, P_1(x)\, \big(1 - ikx - \tfrac{1}{2}\, k^2 x^2 + \tfrac{1}{6}\, i\, k^3\, x^3 + \ldots\big) \\ &= 1 - ik\langle x\rangle - \tfrac{1}{2}\, k^2 \langle x^2\rangle + \tfrac{1}{6}\, i\, k^3 \langle x^3\rangle + \ldots \quad .\end{aligned} \tag{1.87}$$

Thus,

$$\ln \hat{P}_1(k) = -i\mu k - \tfrac{1}{2}\sigma^2 k^2 + \tfrac{1}{6}\, i\, \gamma^3\, k^3 + \ldots \quad , \tag{1.88}$$

where

$$\mu = \langle x\rangle \qquad , \qquad \sigma^2 = \langle x^2\rangle - \langle x\rangle^2 \qquad , \qquad \gamma^3 = \langle x^3\rangle - 3\,\langle x^2\rangle\,\langle x\rangle + 2\,\langle x\rangle^3 \tag{1.89}$$

We can now write

$$\big[\hat{P}_1(k)\big]^N = e^{-iN\mu k}\, e^{-N\sigma^2 k^2/2}\, e^{iN\gamma^3 k^3/6} \cdots \tag{1.90}$$

---

[9] 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 Orleans 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.

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

$$\begin{aligned} P_N(X) &= \int\limits_{-\infty}^{\infty} \frac{dk}{2\pi}\, e^{ik(X-N\mu)}\, e^{-N\sigma^2k^2/2} \Big\{ 1 + \tfrac{1}{6}\, i\, N\gamma^3 k^3 + \ldots \Big\} \\ &= \bigg(1 - \frac{\gamma^3}{6} N \frac{\partial^3}{\partial X^3} + \ldots \bigg) \frac{1}{\sqrt{2\pi N\sigma^2}}\, e^{-(X-N\mu)^2/2N\sigma^2} \\ &= \bigg(1 - \frac{\gamma^3}{6} N^{-1/2} \frac{\partial^3}{\partial \xi^3} + \ldots \bigg) \frac{1}{\sqrt{2\pi N\sigma^2}}\, e^{-\xi^2/2\sigma^2} \quad . \end{aligned} \tag{1.91}$$

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*: in the limit $N \to \infty$, the distribution of a sum of N independent random variables x_i is a Gaussian with mean $N\mu$ and standard deviation $\sqrt{N}\,\sigma$. Our only assumptions are that the mean μ and standard deviation σ 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 moment exist, where the k^{th} moment is simply $\langle x^k \rangle$, the distribution of the sum $X = \sum_{i=1}^N x_i$ is a Gaussian.

1.5.3 Moments and cumulants

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

$$\hat{P}(k_1, \ldots, k_N) = \int\limits_{-\infty}^{\infty} dx_1 \cdots \int\limits_{-\infty}^{\infty} dx_N\, P(x_1, \ldots, x_N)\, \exp\bigg(-\, i \sum_{j=1}^N k_j x_j \bigg) \quad . \tag{1.92}$$

The inverse relation is

$$P(x_1, \ldots, x_N) = \int\limits_{-\infty}^{\infty} \frac{dk_1}{2\pi} \cdots \int\limits_{-\infty}^{\infty} \frac{dk_N}{2\pi}\, \hat{P}(k_1, \ldots, k_N)\, \exp\bigg(+\, i \sum_{j=1}^N k_j x_j \bigg) \quad . \tag{1.93}$$

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,

$$\Bigg[\bigg(i\, \frac{\partial}{\partial k_1} \bigg)^{m_1} \cdots \bigg(i\, \frac{\partial}{\partial k_N} \bigg)^{m_N} \hat{P}(\boldsymbol{k}) \Bigg]_{\boldsymbol{k}=0} = \big\langle x_1^{m_1} \cdots x_N^{m_N} \big\rangle \quad . \tag{1.94}$$

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

$$\hat{P}(\boldsymbol{k}) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_N=0}^{\infty} \frac{(-ik_1)^{m_1}}{m_1!} \cdots \frac{(-ik_N)^{m_N}}{m_N!} \big\langle x_1^{m_1} \cdots x_N^{m_N} \big\rangle \quad . \tag{1.95}$$

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

$$\ln \hat{P}(\boldsymbol{k}) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_N=0}^{\infty} \frac{(-ik_1)^{m_1}}{m_1!} \cdots \frac{(-ik_N)^{m_N}}{m_N!} \, \langle\langle x_1^{m_1} \cdots x_N^{m_N} \rangle\rangle \quad . \tag{1.96}$$

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

$$\begin{aligned} \langle\langle x_i \rangle\rangle &= \langle x_i \rangle \\ \langle\langle x_i x_j \rangle\rangle &= \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \\ \langle\langle x_i x_j x_k \rangle\rangle &= \langle x_i x_j x_k \rangle - \langle x_i x_j \rangle \langle x_k \rangle - \langle x_j x_k \rangle \langle x_i \rangle - \langle x_k x_i \rangle \langle x_j \rangle + 2 \langle x_i \rangle \langle x_j \rangle \langle x_k \rangle \quad . \end{aligned} \tag{1.97}$$

1.5.4 Multidimensional Gaussian integral

Consider the multivariable Gaussian distribution,

$$P(\boldsymbol{x}) \equiv \left(\frac{\det A}{(2\pi)^n} \right)^{1/2} \exp\Big(-\tfrac{1}{2} \, x_i \, A_{ij} \, x_j \Big) \quad , \tag{1.98}$$

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

$$Z(\boldsymbol{b}) = \left(\frac{\det A}{(2\pi)^n} \right)^{1/2} \int\limits_{-\infty}^{\infty} \! dx_1 \cdots \int\limits_{-\infty}^{\infty} \! dx_n \, \exp\Big(-\tfrac{1}{2} \, x_i \, A_{ij} \, x_j + b_i \, x_i \Big) = \exp\Big(\tfrac{1}{2} \, b_i \, A^{-1}_{ij} \, b_j \Big) \quad . \tag{1.99}$$

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

$$\begin{aligned} \langle \, x_{j_1} \cdots \, x_{j_{2k}} \, \rangle &= \int \! d^n\!x \; P(\boldsymbol{x}) \, x_{j_1} \cdots \, x_{j_{2k}} = \frac{\partial^n Z(\boldsymbol{b})}{\partial b_{j_1} \cdots \, \partial b_{j_{2k}}} \bigg|_{\boldsymbol{b}=0} \\ &= \sum_{\text{contractions}} A^{-1}_{j_{\sigma(1)} j_{\sigma(2)}} \cdots A^{-1}_{j_{\sigma(2k-1)} j_{\sigma(2k)}} \quad . \end{aligned} \tag{1.100}$$

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_k, 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

$$C_{2k} = (2k-1) \cdot (2k-3) \cdots 3 \cdot 1 = \frac{(2k)!}{2^k k!} \quad . \tag{1.101}$$

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

$$\langle \, x_{j_1} x_{j_2} x_{j_3} x_{j_4} \, \rangle = A^{-1}_{j_1 j_2} A^{-1}_{j_3 j_4} + A^{-1}_{j_1 j_3} A^{-1}_{j_2 j_4} + A^{-1}_{j_1 j_4} A^{-1}_{j_2 j_3} \quad . \tag{1.102}$$

If we define $b_i = ik_i$, we have

$$\hat{P}(\boldsymbol{k}) = \exp\Big(-\tfrac{1}{2}\,k_i\,A^{-1}_{ij}\,k_j\Big) \quad , \tag{1.103}$$

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 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*[10]. 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

$$P(H\,|D) = \frac{P(D\,|\,H)\,P(H)}{P(D)} \quad . \tag{1.104}$$

---

[10] "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.

Typically the data is 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 π, rather than using the symbol P everywhere. We then have

$$\pi(\boldsymbol{\theta} \,|\, \boldsymbol{x}) = \frac{f(\boldsymbol{x} \,|\, \boldsymbol{\theta})\,\pi(\boldsymbol{\theta})}{\int\limits_{\Theta} d\boldsymbol{\theta}'\, f(\boldsymbol{x} \,|\, \boldsymbol{\theta}')\,\pi(\boldsymbol{\theta}')} \quad , \tag{1.105}$$

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 $\boldsymbol{\theta}$, $f(\boldsymbol{x} \,|\, \boldsymbol{\theta})$ 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*,

$$f(\boldsymbol{y} \,|\, \boldsymbol{x}) = \int\limits_{\Theta} d\boldsymbol{\theta}\, f(\boldsymbol{y} \,|\, \boldsymbol{\theta})\,\pi(\boldsymbol{\theta} \,|\, \boldsymbol{x}) \quad . \tag{1.106}$$

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

$$f(\boldsymbol{x}) = \int\limits_{\Theta} d\boldsymbol{\theta}\, f(\boldsymbol{x} \,|\, \boldsymbol{\theta})\,\pi(\boldsymbol{\theta}) \quad . \tag{1.107}$$

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,

$$\begin{aligned} f(x_1, \ldots, x_N \,|\, \theta) &= \prod_{j=1}^{N} \Big[(1-\theta)\,\delta_{x_j,0} + \theta\,\delta_{x_j,1} \Big] \\ &= \theta^X (1-\theta)^{N-X} \quad , \end{aligned} \tag{1.108}$$

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,

$$\pi(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\mathsf{B}(\alpha,\beta)} \quad , \tag{1.109}$$

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 α, β. Even if we limit ourselves to this form of the prior, different Bayesians might bring different assumptions about the values of α and β. Note that if we choose $\alpha = \beta = 1$, the prior distribution for θ is flat, with $\pi(\theta) = 1$.

We now compute the posterior distribution for θ:

$$\pi(\theta\,|\,x_1,\ldots,x_N)=\frac{f(x_1,\ldots,x_N\,|\,\theta)\,\pi(\theta)}{\int_0^1\!d\theta'\,f(x_1,\ldots,x_N\,|\,\theta')\,\pi(\theta')}=\frac{\theta^{X+\alpha-1}(1-\theta)^{N-X+\beta-1}}{\mathsf{B}(X+\alpha,N-X+\beta)}\quad. \tag{1.110}$$

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

$$\begin{aligned}\alpha'&=X+\alpha\\ \beta'&=N-X+\beta\quad.\end{aligned} \tag{1.111}$$

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,

$$\begin{aligned}f(x_1,\ldots,x_N)&=\int\limits_0^1\!d\theta\,f(x_1,\ldots,x_N\,|\,\theta)\,\pi(\theta)\\ &=\frac{1}{\mathsf{B}(\alpha,\beta)}\int\limits_0^1\!d\theta\,\theta^{X+\alpha-1}(1-\theta)^{N-X+\beta-1}=\frac{\mathsf{B}(X+\alpha,N-X+\beta)}{\mathsf{B}(\alpha,\beta)}\quad.\end{aligned} \tag{1.112}$$

The posterior predictive is then

$$\begin{aligned}f(y_1,\ldots,y_M\,|\,x_1,\ldots,x_N)&=\int\limits_0^1\!d\theta\,f(y_1,\ldots,y_M\,|\,\theta)\,\pi(\theta\,|\,x_1,\ldots,x_N)\\ &=\frac{1}{\mathsf{B}(X+\alpha,N-X+\beta)}\int\limits_0^1\!d\theta\,\theta^{X+Y+\alpha-1}(1-\theta)^{N-X+M-Y+\beta-1}\\ &=\frac{\mathsf{B}(X+Y+\alpha,N-X+M-Y+\beta)}{\mathsf{B}(X+\alpha,N-X+\beta)}\quad.\end{aligned} \tag{1.113}$$

1.6.3 Hyperparameters and conjugate priors

In the above example, θ is a *parameter* of the Bernoulli distribution, *i.e.* the likelihood, while quantities α and β 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

$$\pi(\boldsymbol{\theta}\,|\,\boldsymbol{x},\boldsymbol{\alpha})=\frac{f(\boldsymbol{x}\,|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta}\,|\,\boldsymbol{\alpha})}{\int\limits_{\Theta}\!d\boldsymbol{\theta}'\,f(\boldsymbol{x}\,|\,\boldsymbol{\theta}')\,\pi(\boldsymbol{\theta}'\,|\,\boldsymbol{\alpha})}\quad, \tag{1.114}$$

replacing eqn. 1.105, *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 nauseum*.

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 $\boldsymbol{\alpha} \in A$,

$$\pi(\boldsymbol{\theta} \,|\, \boldsymbol{x}, \boldsymbol{\alpha}) \equiv \frac{f(\boldsymbol{x} \,|\, \boldsymbol{\theta})\, \pi(\boldsymbol{\theta} \,|\, \boldsymbol{\alpha})}{\int\limits_{\Theta} d\boldsymbol{\theta}'\, f(\boldsymbol{x} \,|\, \boldsymbol{\theta}')\, \pi(\boldsymbol{\theta}' \,|\, \boldsymbol{\alpha})} \in \mathcal{P} \quad . \tag{1.115}$$

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

$$f(\boldsymbol{x} \,|\, u, s) = (2\pi s^2)^{-N/2} \exp\Bigg\{ -\frac{1}{2s^2} \sum_{j=1}^{N} (x_j - u)^2 \Bigg\} \quad . \tag{1.116}$$

The parameters here are $\boldsymbol{\theta} = \{u, s\}$. Now consider the prior distribution

$$\pi(u, s \,|\, \mu_0, \sigma_0) = (2\pi\sigma_0^2)^{-1/2} \exp\Bigg\{ -\frac{(u - \mu_0)^2}{2\sigma_0^2} \Bigg\} \quad . \tag{1.117}$$

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:

$$\begin{aligned} \pi(u, s \,|\, \boldsymbol{x}, \mu_0, \sigma_0) &\propto f(\boldsymbol{x} \,|\, u, s)\, \pi(u, s \,|\, \mu_0, \sigma_0) \\ &= \exp\Bigg\{ -\bigg(\frac{1}{2\sigma_0^2} + \frac{N}{2s^2} \bigg) u^2 + \bigg(\frac{\mu_0}{\sigma_0^2} + \frac{N\langle x \rangle}{s^2} \bigg) u - \bigg(\frac{\mu_0^2}{2\sigma_0^2} + \frac{N\langle x^2 \rangle}{2s^2} \bigg) \Bigg\} \quad , \end{aligned} \tag{1.118}$$

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

$$\pi(u, s \,|\, \boldsymbol{x}, \mu_0, \sigma_0) = (2\pi\sigma_1^2)^{-1/2} \exp\Bigg\{ -\frac{(u - \mu_1)^2}{2\sigma_1^2} \Bigg\} \quad , \tag{1.119}$$

with

$$\mu_1 = \mu_0 + \frac{N\big(\langle x \rangle - \mu_0\big)\sigma_0^2}{s^2 + N\sigma_0^2} \qquad , \qquad \sigma_1^2 = \frac{s^2\sigma_0^2}{s^2 + N\sigma_0^2} \quad . \tag{1.120}$$

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 μ_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 the for the coin flipping problem, the flat prior $\pi(\theta) = 1$ is an appropriate initial one, since it does not privilege any value of θ. 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 $\zeta = \ln\big[\theta/(1-\theta)\big]$. In terms of the new parameter ζ, 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 $\tilde{\pi}(\zeta) = \frac{1}{4}\text{sech}^2(\zeta/2)$, which is not flat. Thus what was uninformative in terms of θ has become very informative in terms of the new parameter ζ. 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: $S = -\int d\mu(\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,\ln\pi(\boldsymbol{\theta})$.

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*,

$$I_{ij}(\boldsymbol{\theta}) = -\mathbb{E}_\theta\bigg(\frac{\partial^2 \ln f(\boldsymbol{x}\,|\,\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\bigg) = -\int d\boldsymbol{x}\; f(\boldsymbol{x}\,|\,\boldsymbol{\theta})\,\frac{\partial^2 \ln f(\boldsymbol{x}\,|\,\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j} \quad . \tag{1.121}$$

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

$$\pi_{\text{J}}(\boldsymbol{\theta}) \propto \sqrt{\det I(\boldsymbol{\theta})} \quad . \tag{1.122}$$

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

$$-\frac{d^2\ln p(\boldsymbol{x}\,|\,\theta)}{d\theta^2} = \frac{X}{\theta^2} + \frac{N-X}{(1-\theta)^2} \quad , \tag{1.123}$$

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

$$I(\theta) = \frac{N}{\theta(1-\theta)} \qquad \Rightarrow \qquad \pi_{\text{J}}(\theta) = \frac{1}{\pi}\,\frac{1}{\sqrt{\theta(1-\theta)}} \quad , \tag{1.124}$$

which felicitously corresponds to a Beta distribution with $\alpha = \beta = \frac{1}{2}$. In this example the Jeffries 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 μ and standard deviation σ. Let the parameters be $(\theta_1,\theta_2) = (\mu,\sigma)$. Then

$$-\ln f(\boldsymbol{x}\,|\,\boldsymbol{\theta}) = N\ln\sqrt{2\pi} + N\ln\sigma + \frac{1}{2\sigma^2}\sum_{j=1}^N (x_j-\mu)^2 \quad , \tag{1.125}$$

and the Fisher information matrix is

$$I(\boldsymbol{\theta}) = -\frac{\partial^2 \ln f(\boldsymbol{x}\,|\,\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j} = \begin{pmatrix} N\sigma^{-2} & \sigma^{-3}\sum_j(x_j-\mu) \\ \sigma^{-3}\sum_j(x_j-\mu) & -N\sigma^{-2} + 3\sigma^{-4}\sum_j(x_j-\mu)^2 \end{pmatrix} \quad . \tag{1.126}$$

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

$$\mathbb{E}\, I(\boldsymbol{\theta}) = \begin{pmatrix} N\sigma^{-2} & 0 \\ 0 & 2N\sigma^{-2} \end{pmatrix} \tag{1.127}$$

and the Jeffries prior is $\pi_{\text{J}}(\mu,\sigma) \propto \sigma^{-2}$. This is problematic because if we choose a flat metric on the (μ,σ) upper half plane, the Jeffries 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^{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 μ, 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 ΔN inside a volume ΔV 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 ΔE 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 $N_{\!\scriptscriptstyle\mathrm{A}}$ of X, and $N_{\!\scriptscriptstyle\mathrm{A}} = 6.0221415 \times 10^{23}$ 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 = 6.0221415×10^{23} guacas.

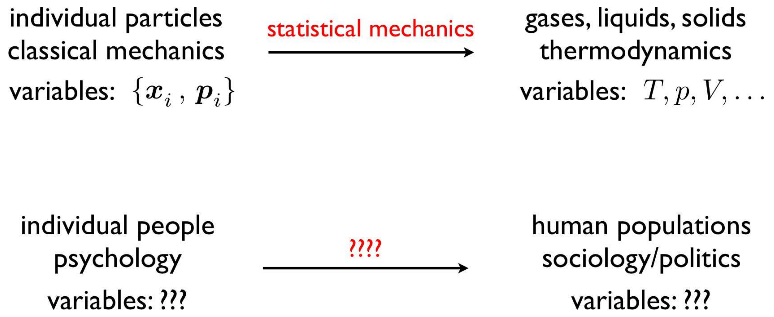


Figure 2.1: From microscale to macroscale : physical *versus* social sciences.

solids and liquids, and $10^{-5} - 10^{-3}\,\text{mol/cm}^3$ 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 ℓ, 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 τ is the *mean free time*. When we study the kinetic theory of gases, we will derive formulas for ℓ and $\bar{v}$ (and hence τ). 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 ($T = 0°\,\text{C}$, $p = 1\,\text{atm}$), the mean free path is $\ell \approx 1.1 \times 10^{-5}$ cm, the average speed is $\bar{v} \approx 480\,\text{m/s}$, and the mean free time is $\tau \approx 2.5 \times 10^{-10}$ s. Thus, particles in the gas undergo collisions at a rate $\tau^{-1} \approx 4.0 \times 10^{9}\,\text{s}^{-1}$. A measuring device, such as our spring, or a thermometer, effectively performs time and space averages. If there are N_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 $N_\text{c}^{-1/2}$ times the average. Since N_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

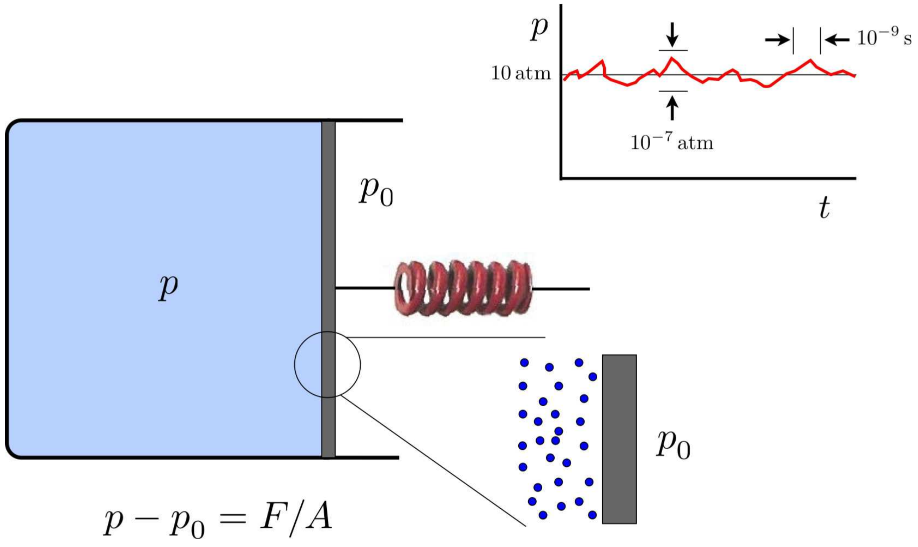


Figure 2.2: The pressure p of a gas is due to an average over space and time of the impulses due to the constituent particles.

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:

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

If ϕ is a state variable or a function of state variables, its infinitesimal change is expressed as $d\phi$, which is an *exact differential*. If ϕ is not a state variable, then its differential is in general *inexact*, and we express its infinitesimal change as $\text{đ}\phi$. 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

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

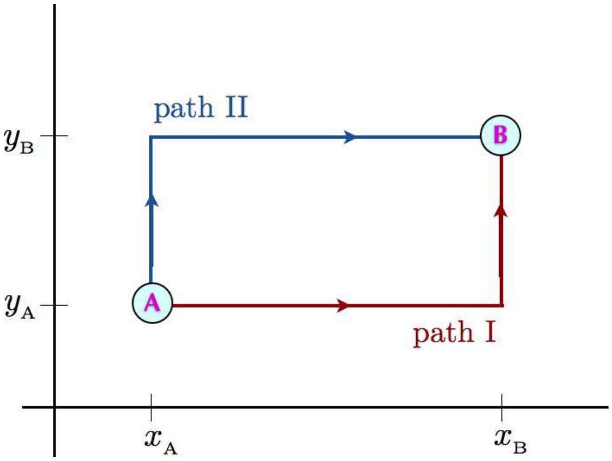


Figure 2.3: Two distinct paths with identical endpoints.

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. 2.148. In this case, we have

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

For exact differentials, the integral between fixed endpoints is path-independent:

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

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:

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

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

---

[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 continuous, which entails eqn. 2.2.

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

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

where $\mathcal{C} = \mathcal{C}_{\mathrm{II}}^{-1} \circ \mathcal{C}_{\mathrm{I}}$ is the closed path formed by appending the reverse of path II (*i.e.* C_{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 $\text{đ}F$ 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 constituent molecule, 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 g different species, we must specify $g + 2$ state variables, which may be $\{T, p, N_1, \ldots, N_g\}$, where N_a is the number of particles of species a. Another possibility is the set $(T, p, V, x_1, \ldots, x_{g-1}\}$, where the *concentration* of species a is $x_a = N_a/N$. Here, $N = \sum_{a=1}^{g} N_a$ is the total number of particles. Note that $\sum_{a=1}^{g} x_a = 1$.

It then follows that if we specify more than $g+2$ 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,

$$pV = N k_{\mathrm{B}} T \quad , \tag{2.7}$$

relating the four state variables T, p, V, and N. Here $k_{\mathrm{B}} = 1.3806503 \times 10^{-16}\,\mathrm{erg/K}$ is Boltzmann's constant. Another example is the van der Waals equation,

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

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 thermodynamics 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]

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

We will use the symbol Q to denote the amount of heat energy absorbed by a system during some given thermodynamic process, and $\text{đ}Q$ 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 $\text{đ}W$ done *by* the system as

$$\text{đ}W = \sum_i F_i\,dX_i \quad , \tag{2.11}$$

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 $\text{đ}W = p\,dV$. More generally, we write

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

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 σ, 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, $\text{đ}W = -d\boldsymbol{y}\cdot d\boldsymbol{X} - d\boldsymbol{\mu}\cdot d\boldsymbol{N}$, with $\boldsymbol{X} = \{X_1,\ldots,X_{\mathsf{d}}\}$ and $N = \{N_1,\ldots,N_{\mathsf{c}}\}$. In this general setting there are d generalized displacements and c component species. When $\delta = 1$, the volume is the only displacement.

As with heat, $\text{đ}W$ 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 $\text{đ}W$ to indicate that this is not an exact differential. More on this in section 2.2.2 below.

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:

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

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 ϱ 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:

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

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

  $$\frac{T}{T_{\mathrm{ref}}} = \frac{V}{V_{\mathrm{ref}}} \quad , \tag{2.15}$$

  where T_{ref} and V_{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

  $$\frac{T}{T_{\mathrm{ref}}} = \frac{p}{p_{\mathrm{ref}}} \quad , \tag{2.16}$$

  where T_{ref} and p_{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

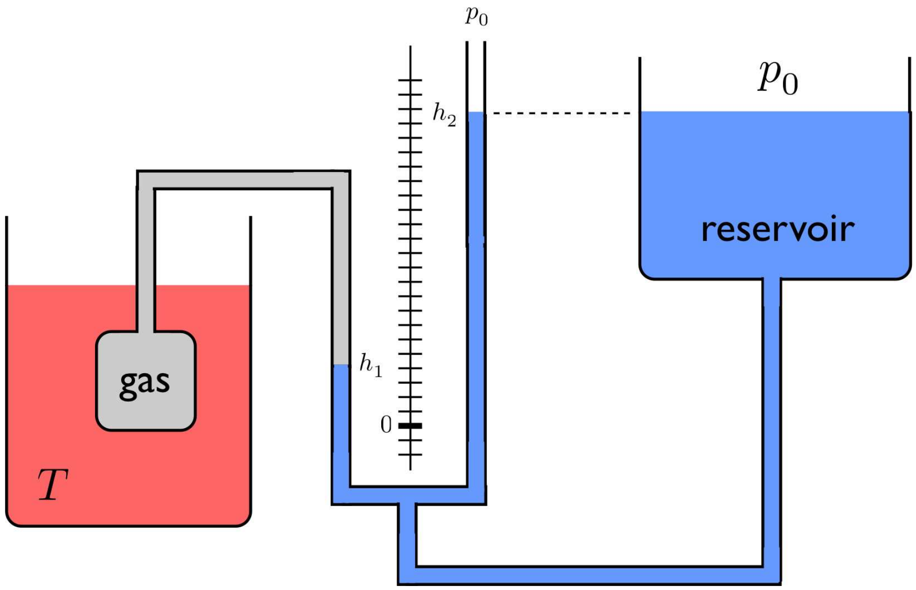


Figure 2.4: The constant volume gas thermometer. The gas is placed in thermal contact with an object of temperature T. An incompressible fluid of density ϱ is used to measure the pressure difference $\Delta p = p_{\rm gas} - p_0$.

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 $T_{\rm t}$ at thus *unique* and is *by definition* $T_{\rm t} = 273.16\,{\rm K}$. The pressure at the triple point is $611.7\,{\rm Pa} = 6.056 \times 10^{-3}\,{\rm atm}$.

A question remains: are the two modes of the thermometer compatible? *E.g.* it 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 mean by STP will keep you on your toes.

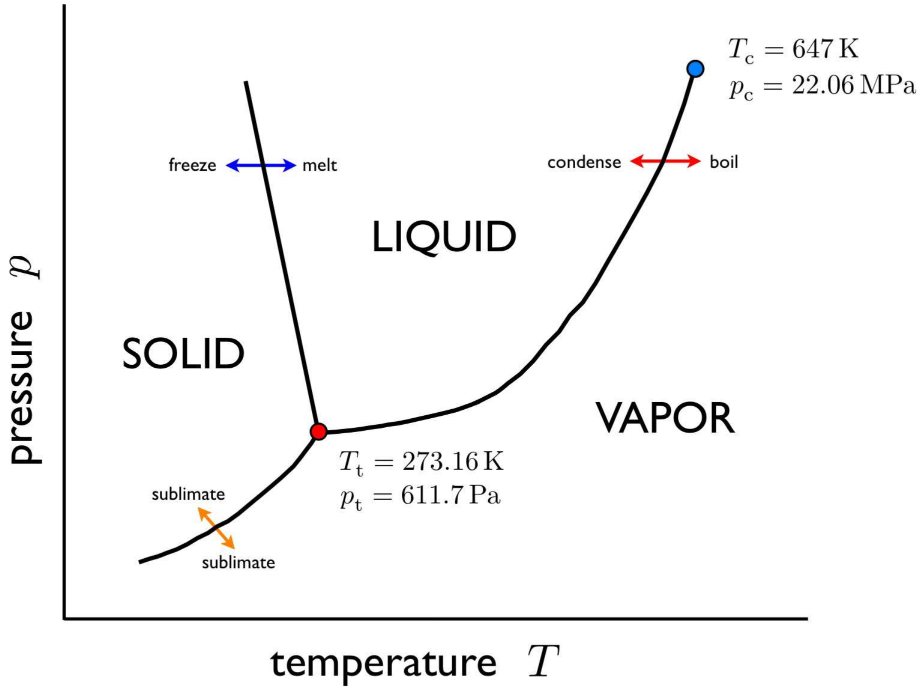


Figure 2.5: A sketch of the phase diagram of H_2O (water). Two special points are identified: the triple point (T_t, p_t) at which there is three phase coexistence, and the critical point (T_c, p_c), where the latent heat of transformation from liquid to gas vanishes. Not shown are transitions between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

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

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

where $1\,\ell = 10^6\,\mathrm{cm}^3 = 10^{-3}\,\mathrm{m}^3$ 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^{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:

$$\begin{array}{rcccl} \text{energy exchange} & \Longrightarrow & T = \text{constant} & \Longrightarrow & \text{thermal equilibrium} \\[2ex] \text{volume exchange} & \Longrightarrow & \dfrac{p}{T} = \text{constant} & \Longrightarrow & \text{mechanical equilibrium} \\[2ex] \text{particle exchange} & \Longrightarrow & \dfrac{\mu}{T} = \text{constant} & \Longrightarrow & \text{chemical equilibrium} \end{array}$$

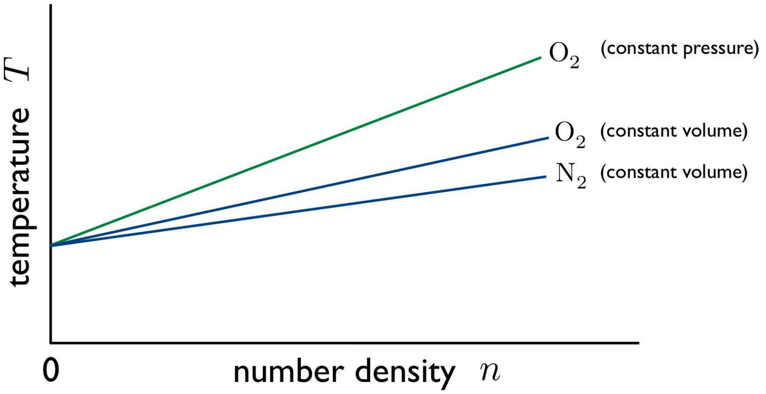


Figure 2.6: As the gas density tends to zero, the readings of the constant volume gas thermometer converge.

Equilibrium is transitive, so

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

This 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 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 §2.111) then guarantees that *any* intensive quantity ϕ may be expressed as a state function $\phi = \phi(T, p, \mu_1, \ldots, \mu_{\mathsf{c}-1})$. In particular, the g^{th} chemical potential μ_{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° C regardless of the amount of substance present.

As we shall derive in §2.11.5 below, a number φ 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 $(g+1)$-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 ($g = 1$) 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

---

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

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

phase coexistence regions, full equilibrium entails that p, T, and μ are constant throughout the system. Furthermore, *all* intensive quantities, such as the specific volume $v = n^{-1}$ or energy density ε, are also 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

$$dE = \text{đ}Q - \text{đ}W \quad . \tag{2.19}$$

We use the symbol đ in the differentials $\text{đ}Q$ and $\text{đ}W$ 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 $Q_{\text{A}} > 0$ but no work is done, so $W_{\text{A}} = 0$. 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.* $Q_{\text{B}} = 0$. However, the stirrer does work $-W_{\text{B}} > 0$ 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

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

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

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

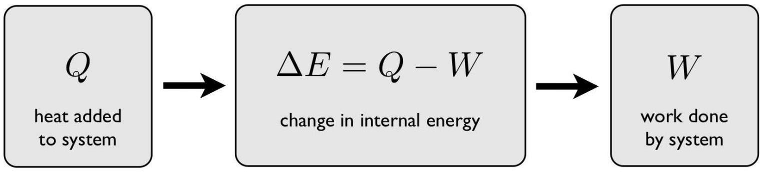


Figure 2.7: The first law of thermodynamics is a statement of energy conservation.

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 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 μ 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

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

Then

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

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°_+ 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	5.19	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.*)

2.4.3 Heat capacity

The *heat capacity* of a body, C, is by definition the ratio $\text{đ}Q/dT$ 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,

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

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

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

Unless stated as otherwise, 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

$$Q = \int\limits_{T_\text{A}}^{T_\text{B}} dT\, C(T) \quad , \tag{2.34}$$

where $C = C_V$ or $C = C_p$ depending on whether volume or pressure is held constant. For ideal gases, as we shall discuss below, $C(T)$ is constant, and thus

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

In metals at very low temperatures one finds $C = \gamma T$, where γ is a constant[8]. We then have

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

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 $T_\text{f} \in (T_1, T_2)$. 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 $\tilde{c}_\text{s,r}(T)$ are constant throughout the regime $T \in [T_1, T_2]$, we then have

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

Thus, a measurement of the final temperature T_f can be used to determine the unknown $\tilde{c}_\text{s}$. For H_2O, it requires 4.1855 J per gram of energy to raise the temperature from $T = 14.5°\text{C}$ to $T' = 15.5°\text{C}$, which is to say $\tilde{c}_{\text{H}_2\text{O}}(T = 15°\text{C}) = 4.1855\,\text{J/g}\,°\text{C}$.[9] 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 $\tilde{c}_\text{s}$.

2.4.4 Ideal gases

The ideal gas equation of state is $pV = Nk_\text{B}T$. 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[10]. Essentially, a

---

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

[9] The specific heat $\tilde{c}_{\text{H}_2\text{O}}(T)$ is weakly temperature dependent throughout much of the liquid phase. It turns out that $\tilde{c}_{\text{H}_2\text{O}}(T = 20°\text{C}) = 4.1819\,\text{J/g}\,°\text{C}$. The conversion factor between Joules and *thermochemical calories* is $1\,\text{J} = 4.184\,\text{cal}$. The so-called "15° calorie" is equivalent to 4.1855 Joules.

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

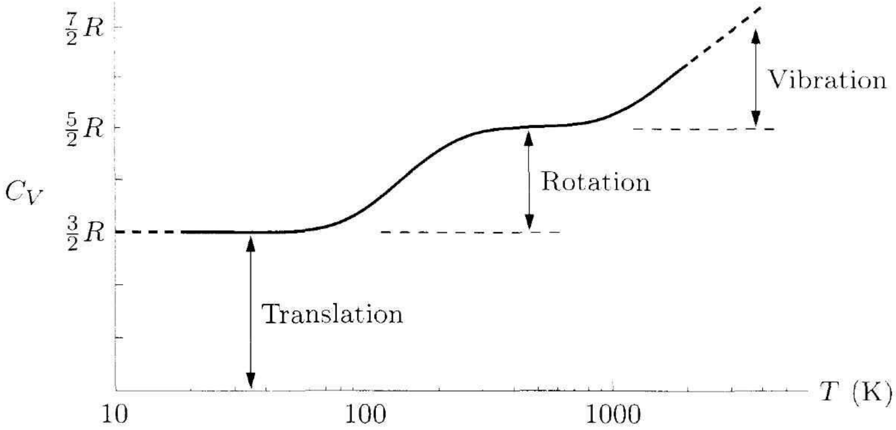


Figure 2.8: Heat capacity C_V for one mole of hydrogen (H_2) gas. At the lowest temperatures, only translational degrees of freedom are relevant, and $f = 3$. At around $200\,\mathrm{K}$, two rotational modes are excitable and $f = 5$. Above $1000\,\mathrm{K}$, the vibrational excitations begin to contribute.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

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 $\nu = N/N_{\mathrm{A}}$ is the number of moles of substance. Note that ν 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 $\nu \equiv N/N_{\mathrm{A}}$ is the number of moles, and where c_V is the molar heat capacity. We then have $c_p = c_V + R$. where $R = N_{\mathrm{A}} k_{\mathrm{B}} = 8.31457\,\mathrm{J/mol\,K}$ is the *gas constant*. We denote by $\gamma = c_p/c_V$ the ratio of specific heat at constant pressure and at constant volume.

From the kinetic theory of gases, one can show that

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

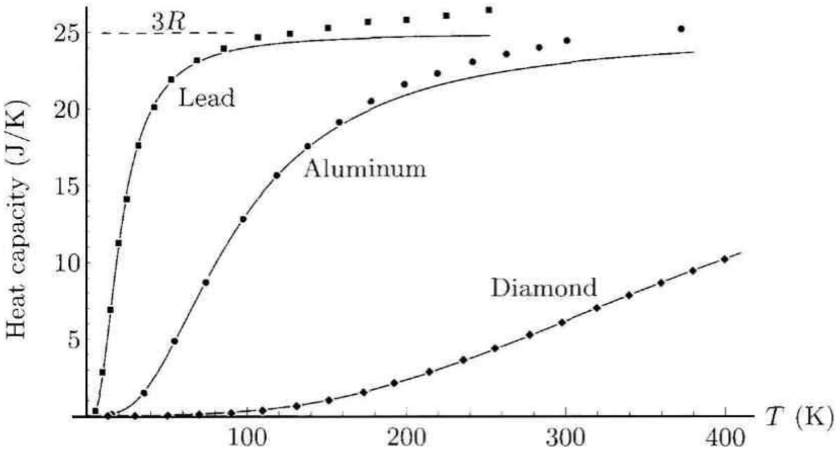


Figure 2.9: Molar heat capacities c_V for three solids. The solid curves correspond to the predictions of the Debye model, which we shall discuss later.

Digression : kinetic theory of gases

We will conclude in general from noninteracting classical statistical mechanics that the specific heat of a substance is $c_v = \frac{1}{2} f R$, 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_α 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 or greater than ε_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 $3R = 24.944\,{\rm J/mol\,K}$. 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 $T = 25^\circ\,{\rm C}$ (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, $c_p = \frac{5}{2}R = 20.7864\,\text{J/mol K}$. 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

$$\text{đ}Q = C_V\,dT + p\,dV \quad . \tag{2.40}$$

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

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

We can immediately integrate to obtain

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

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 ϱ 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 $M = 28.88\,\text{g}$ for air and $T = 293\,\text{K}$, we have $\lambda = 8.6\,\text{km}$.

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

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

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 $T_0 = T(0) = 293\,\text{K}$ 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 $dT/dz = -(1-\gamma^{-1})\,T_0/\lambda = -9.7\,\text{K/km}$. Note that in this model the atmosphere ends at a height $z_{\text{max}} = \gamma\lambda/(\gamma-1) = 30\,\text{km}$.

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 (ν 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 $E(T,V,N) = Nc_V T$, then $\Delta E = 0$ 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 $p = Nk_\text{B}T/V_1$ and the final pressure is $p' = Nk_\text{B}T/(V_1+V_2)$.

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 α, 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 $\Delta E = 0$,

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

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, although we know on general grounds that it must decrease because, as we shall see, thermodynamic stability entails a positive isothermal compressibility: $\kappa_T = -V^{-1}(\partial V/\partial p)_{T,N} > 0$.

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

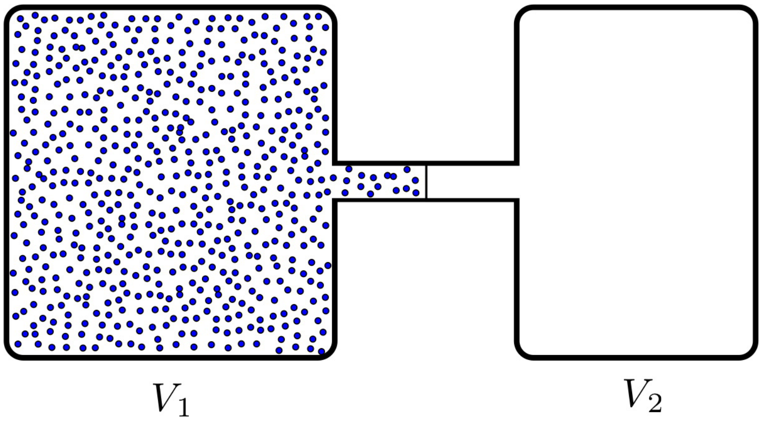


Figure 2.10: In the adiabatic free expansion of a gas, there is volume expansion with no work or heat exchange with the environment: $\Delta E = Q = W = 0$.

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

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

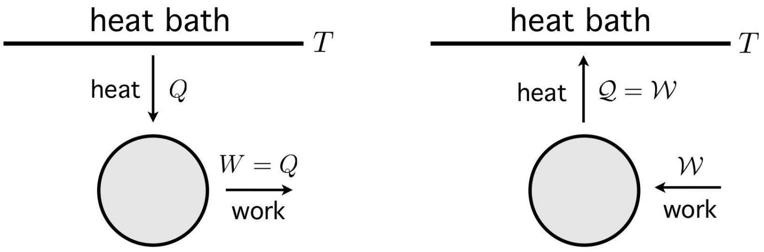
The work done *by the gas* is $W = \int p\, dV = -\mathcal{W}$. During the compression, heat energy $Q = W < 0$ is transferred to the gas from the reservoir. Thus, $\mathcal{Q} = \mathcal{W} > 0$ is given off by the gas to its environment.

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: $\text{A} \to \text{B} \to \text{C} \cdots \to \text{A}$. 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 $\frac{1}{2}k_{\rm B}T$ 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) 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.

> *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.

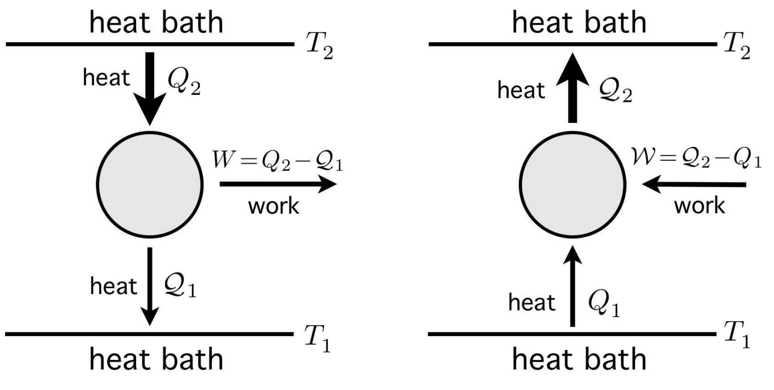
2.5.2 Engines and refrigerators

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 useful work. We define the *efficiency*, η, 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



(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.

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*, κ, defined as

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

Thus, κ 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[11] (1796 – 1832) realized that a *reversible* cyclic engine operating between 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 η_{C} can only be a function of the temperatures T_1 and T_2: $\eta_{\mathrm{C}} = \eta_{\mathrm{C}}(T_1, T_2)$. We can then define

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

Below, in §2.5.4, we will see that 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.

---

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

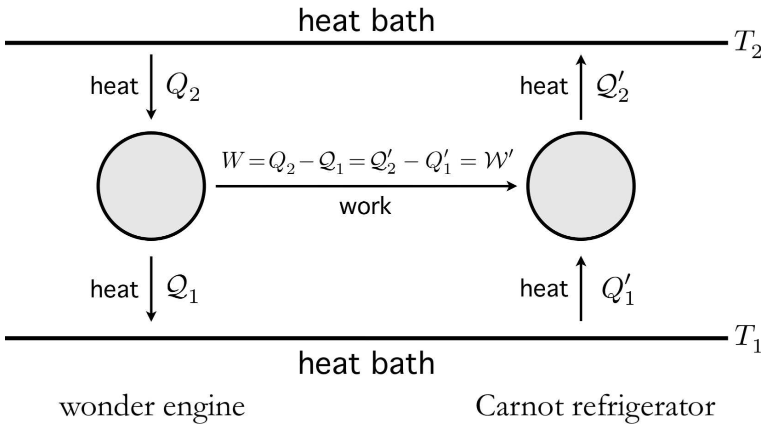


Figure 2.13: A wonder engine driving a Carnot refrigerator.

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

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

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

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

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.

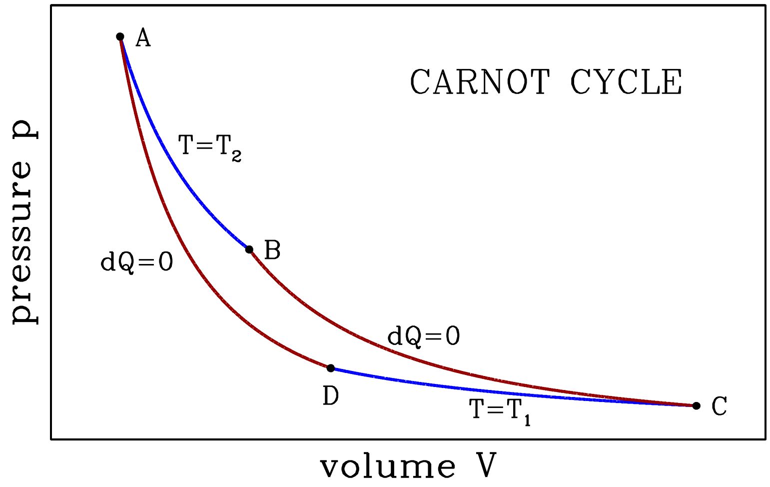


Figure 2.14: The Carnot cycle consists of two adiabats (dark red) and two isotherms (blue).

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 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,

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

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

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

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

$$W_{\rm AB} = \int\limits_{V_{\rm A}}^{V_{\rm B}} dV\,\frac{\nu R T_2}{V} = \nu R T_2\,\log\left(\frac{V_{\rm B}}{V_{\rm A}}\right) \tag{2.63}$$

$$E_{\rm A} = E_{\rm B} = \frac{\nu R T_2}{\gamma - 1} \quad , \tag{2.64}$$

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*:

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

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

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:

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

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

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 $\Delta E = 0$, we must have $Q = W$, 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[12] is as follows. During the isothermal expansion stage, the working material is maintained at a temperature $T_{\rm 2w} < T_2$. The temperature difference between the working material and the hot reservoir drives a thermal current,

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

---

[12] 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)	η_{Carnot}	η (theor.)	η (obs.)
West Thurrock (UK) Coal Fired Steam Plant	~ 25	565	0.641	0.40	0.36
CANDU (Canada) PHW Nuclear Reactor	~ 25	300	0.480	0.28	0.30
Larderello (Italy) Geothermal Steam Plant	~ 80	250	0.323	0.175	0.16

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

Here, κ_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 $T_{1\text{w}} > T_1$, which drives a thermal current to the cold reservoir,

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

Now let us assume that the upper isothermal stage requires a duration Δt_2 and the lower isotherm a duration Δt_1. Then

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

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

Since the engine is reversible, we must have

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

which says

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

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

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

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

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

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

The efficiency at maximum power is then

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

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

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

Table 2.2, taken from the article of Curzon and Albhorn (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, as shown in fig. 2.15. Each little Carnot cycle consists of two adiabats and two isotherms. We then conclude

$$\sum_i \frac{Q_i}{T_i} \longrightarrow \oint\limits_{\mathcal{C}} \frac{\text{đ}Q}{T} \le 0 \quad , \tag{2.96}$$

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:

$$dS = \frac{\text{đ}Q}{T} \quad \Longrightarrow \quad S_{\text{B}} - S_{\text{A}} = \int\limits_{\text{A}}^{\text{B}} \frac{\text{đ}Q}{T} \quad . \tag{2.97}$$

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

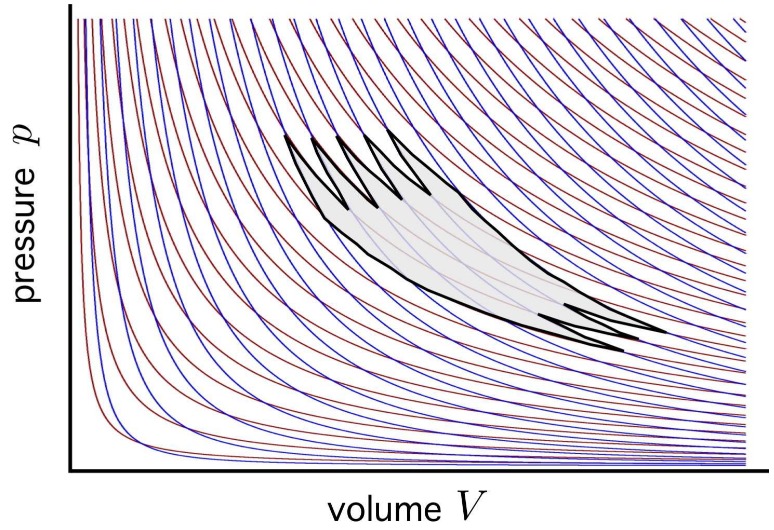


Figure 2.15: An arbitrarily shaped cycle in the $p-V$ plane can be decomposed into a number of smaller Carnot cycles. Red curves indicate isotherms and blue curves adiabats, with $\gamma = \frac{5}{3}$.

2.6.2 The Third Law of thermodynamics

Eqn. 2.97 determines the entropy up to a constant. By choosing a standard state Υ, we can define $S_\Upsilon = 0$, and then by taking $\text{A} = \Upsilon$ in the above equation, we can define the absolute entropy S for any state. However, it turns out that this seemingly arbitrary constant S_Υ 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 $S(T=0) = k_\text{B} \log g$, where g is the ground state degeneracy. Nernst's law holds when $g = 1$.

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: $\Delta S_\text{CYC} = 0$. 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

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

written as a sum over entropy changes of the engine itself, the hot reservoir, and the cold reservoir[13]. Clearly $(\Delta S)_{\text{ENGINE}} = 0$. The changes in the reservoir entropies are

$$(\Delta S)_{\text{HOT}} = \int\limits_{T=T_2} \frac{\text{đ}Q_{\text{HOT}}}{T} = -\frac{Q_2}{T_2} \; < \; 0 \tag{2.100}$$

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

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,

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

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.

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 λ. 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 λ 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,

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

---

[13] 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.

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

$$dE = \mathchar'26\mkern-12mu d Q - \mathchar'26\mkern-12mu d W = T\,dS + \sum_{j=1}^{\textsf{d}} y_j\,dX_j + \sum_{a=1}^{\textsf{c}} \mu_a\,dN_a \quad , \tag{2.106}$$

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

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

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, according to eqn. 2.105, 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

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

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

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

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

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

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

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

Integrating, we obtain

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

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

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

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

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

where $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T}$ 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

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

- *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 $dN = 0$,

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

  Integrating, we find

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

  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

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

where a is a constant. We have

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

and therefore

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

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

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

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

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

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:

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

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

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

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

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

Now we can write

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

Adding up all the work, we obtain

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

Since

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

we find once again

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

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

$$S(p,T) = S(p,T=0) + \int\limits_0^T dT'\,\frac{C_p(T')}{T'} \quad . \tag{2.141}$$

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 all but highly unusual cases, $g=1$ and $S(p,T=0)=0$.

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 to 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 and particle number. 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 $đW = -d\boldsymbol{y}\cdot d\boldsymbol{X} - d\boldsymbol{\mu}\cdot d\boldsymbol{X}$, with $\boldsymbol{X} = \{X_1,\ldots,X_{\sf d}\}$ and $\boldsymbol{N} = \{N_1,\ldots,N_{\sf c}\}$. For what we may call (p,V) systems, there is only ${\sf d}=1$ 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 that when 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 II, §2.16.

Again invoking the Second Law, we have

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

Rearranging terms, we have $đW \le -S\,dT - dF$, 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

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

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

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, $dF \le -S\,dT - p\,dV + \mu\,dN$ 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.[14] 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

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

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

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

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

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

$$dG \le -S\,dT + V\,dp + p\,dV - đW \quad , \tag{2.157}$$

---

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

which we may rearrange as $đW_{\text{free}} \le -S\,dT + V\,dp - dG$. Accordingly, we conclude, for systems at fixed (T, p), that $đW_{\text{free}} \le -dG$.

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 μ 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 Ω

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

$$d\Omega \le -S\,dT - p\,dV - \sum_{a=1}^{\mathsf{C}} N_a\, d\mu_a - đ\widetilde{W}_{\text{free}} \quad , \tag{2.161}$$

where we formally define

$$đ\widetilde{W}_{\text{free}} \equiv đW - p\,dV + \sum_{a=1}^{\mathsf{C}} \mu_a\, dN_a \quad . \tag{2.162}$$

We then have $đW_{\text{free}} \le -S\,dT - p\,dV - \boldsymbol{N} \cdot d\boldsymbol{\mu} - d\Omega$. We conclude that for systems at fixed $(T, V, \boldsymbol{\mu})$ that $đW_{\text{free}} \le -d\Omega$.

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, $d\Omega \le -S\,dT - p\,dV - N d\mu$, hence *spontaneous processes drive the grand potential* Ω *to a minimum in systems at fixed* (T, V, μ).

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 Φ itself 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 $u \neq j$ constant. Then the equality of the mixed second partial derivatives is equivalent to the conditions

$$\left(\frac{\partial R_i}{\partial u_j}\right)_{u_{k(\neq j)}} = \left(\frac{\partial R_j}{\partial u_i}\right)_{u_{k(\neq i)}} \quad .$$

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

$$T = \left(\frac{\partial E}{\partial S}\right)_{V,N} \quad , \quad -p = \left(\frac{\partial E}{\partial V}\right)_{S,N} \quad , \quad \mu = \left(\frac{\partial E}{\partial N}\right)_{S,V} \quad . \tag{2.166}$$

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

$$-S = \left(\frac{\partial F}{\partial T}\right)_{V,N} \quad , \quad -p = \left(\frac{\partial F}{\partial V}\right)_{T,N} \quad , \quad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \quad . \tag{2.171}$$

Taking the mixed second derivatives, we find

$$\frac{\partial^2 F}{\partial T\, \partial V} = -\left(\frac{\partial S}{\partial V}\right)_{T,N} = -\left(\frac{\partial p}{\partial T}\right)_{V,N} \tag{2.172}$$

$$\frac{\partial^2 F}{\partial T\, \partial N} = -\left(\frac{\partial S}{\partial N}\right)_{T,V} = \left(\frac{\partial \mu}{\partial T}\right)_{V,N} \tag{2.173}$$

$$\frac{\partial^2 F}{\partial V\, \partial N} = -\left(\frac{\partial p}{\partial N}\right)_{T,V} = \left(\frac{\partial \mu}{\partial V}\right)_{T,N} \quad . \tag{2.174}$$

2.8.3 Relations deriving from $\mathsf{H}(S,p,N)$

The enthalpy $\mathsf{H}(S,p,N)$ satisfies

$$d\mathsf{H} = T\, dS + V\, dp + \mu\, dN \quad , \tag{2.175}$$

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

$$T = \left(\frac{\partial \mathsf{H}}{\partial S}\right)_{p,N} \quad , \quad V = \left(\frac{\partial \mathsf{H}}{\partial p}\right)_{S,N} \quad , \quad \mu = \left(\frac{\partial \mathsf{H}}{\partial N}\right)_{S,p} \quad . \tag{2.176}$$

Taking the mixed second derivatives, we find

$$\frac{\partial^2 \mathsf{H}}{\partial S\, \partial p} = \left(\frac{\partial T}{\partial p}\right)_{S,N} = \left(\frac{\partial V}{\partial S}\right)_{p,N} \tag{2.177}$$

$$\frac{\partial^2 \mathsf{H}}{\partial S\, \partial N} = \left(\frac{\partial T}{\partial N}\right)_{S,p} = \left(\frac{\partial \mu}{\partial S}\right)_{p,N} \tag{2.178}$$

$$\frac{\partial^2 \mathsf{H}}{\partial p\, \partial N} = \left(\frac{\partial V}{\partial N}\right)_{S,p} = \left(\frac{\partial \mu}{\partial p}\right)_{S,N} \quad . \tag{2.179}$$

2.8.4 Relations deriving from $G(T,p,N)$

The Gibbs free energy $G(T,p,N)$ satisfies

$$dG = -S\, dT + V\, dp + \mu\, dN \quad , \tag{2.180}$$

therefore $G = G(T,p,N)$, with

$$-S = \left(\frac{\partial G}{\partial T}\right)_{p,N} \quad , \quad V = \left(\frac{\partial G}{\partial p}\right)_{T,N} \quad , \quad \mu = \left(\frac{\partial G}{\partial N}\right)_{T,p} \quad . \tag{2.181}$$

Taking the mixed second derivatives, we find

$$\frac{\partial^2 G}{\partial T\,\partial p} = -\left(\frac{\partial S}{\partial p}\right)_{T,N} = \left(\frac{\partial V}{\partial T}\right)_{p,N} \tag{2.182}$$

$$\frac{\partial^2 G}{\partial T\,\partial N} = -\left(\frac{\partial S}{\partial N}\right)_{T,p} = \left(\frac{\partial \mu}{\partial T}\right)_{p,N} \tag{2.183}$$

$$\frac{\partial^2 G}{\partial p\,\partial N} = \left(\frac{\partial V}{\partial N}\right)_{T,p} = \left(\frac{\partial \mu}{\partial p}\right)_{T,N} \quad . \tag{2.184}$$

2.8.5 Relations deriving from $\Omega(T, V, \mu)$

The grand potential $\Omega(T, V, \mu)$ satisfied

$$d\Omega = -S\,dT - p\,dV - N\,d\mu \quad , \tag{2.185}$$

hence

$$-S = \left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} \quad , \quad -p = \left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} \quad , \quad -N = \left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} \quad . \tag{2.186}$$

Taking the mixed second derivatives, we find

$$\frac{\partial^2 \Omega}{\partial T\,\partial V} = -\left(\frac{\partial S}{\partial V}\right)_{T,\mu} = -\left(\frac{\partial p}{\partial T}\right)_{V,\mu} \tag{2.187}$$

$$\frac{\partial^2 \Omega}{\partial T\,\partial \mu} = -\left(\frac{\partial S}{\partial \mu}\right)_{T,V} = -\left(\frac{\partial N}{\partial T}\right)_{V,\mu} \tag{2.188}$$

$$\frac{\partial^2 \Omega}{\partial V\,\partial \mu} = -\left(\frac{\partial p}{\partial \mu}\right)_{T,V} = -\left(\frac{\partial N}{\partial V}\right)_{T,\mu} \quad . \tag{2.189}$$

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

$$dS = \frac{1}{T}\,dE + \frac{p}{T}\,dV - \frac{\mu}{T}\,dN \quad . \tag{2.190}$$

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

$$\frac{\partial^2 S}{\partial E\,\partial V} = \left(\frac{\partial (T^{-1})}{\partial V}\right)_{E,N} = \left(\frac{\partial (pT^{-1})}{\partial E}\right)_{V,N} \quad , \tag{2.191}$$

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

$$\boxed{\texttt{THIS SPACE AVAILABLE}} \qquad dE = T\,dS + y\,dX + \mu\,dN \tag{2.192}$$

$$F = E - TS \qquad dF = -S\,dT + y\,dX + \mu\,dN \tag{2.193}$$

$$\mathsf{H} = E - yX \qquad d\mathsf{H} = T\,dS - X\,dy + \mu\,dN \tag{2.194}$$

$$G = E - TS - yX \qquad dG = -S\,dT - X\,dy + \mu\,dN \tag{2.195}$$

$$\varOmega = E - TS - \mu N \qquad d\varOmega = -S\,dT + y\,dX - N\,d\mu \quad . \tag{2.196}$$

Generalizing $(-p, V) \to (y, X)$, we also obtain, *mutatis mutandis*, the following Maxwell relations:

$$\left(\frac{\partial T}{\partial X}\right)_{S,N} = \left(\frac{\partial y}{\partial S}\right)_{X,N} \qquad \left(\frac{\partial T}{\partial N}\right)_{S,X} = \left(\frac{\partial \mu}{\partial S}\right)_{X,N} \qquad \left(\frac{\partial y}{\partial N}\right)_{S,X} = \left(\frac{\partial \mu}{\partial X}\right)_{S,N} \tag{2.197}$$

$$\left(\frac{\partial T}{\partial y}\right)_{S,N} = -\left(\frac{\partial X}{\partial S}\right)_{y,N} \qquad \left(\frac{\partial T}{\partial N}\right)_{S,y} = \left(\frac{\partial \mu}{\partial S}\right)_{y,N} \qquad \left(\frac{\partial X}{\partial N}\right)_{S,y} = -\left(\frac{\partial \mu}{\partial y}\right)_{S,N} \tag{2.198}$$

$$\left(\frac{\partial S}{\partial X}\right)_{T,N} = -\left(\frac{\partial y}{\partial T}\right)_{X,N} \qquad \left(\frac{\partial S}{\partial N}\right)_{T,X} = -\left(\frac{\partial \mu}{\partial T}\right)_{X,N} \qquad \left(\frac{\partial y}{\partial N}\right)_{T,X} = \left(\frac{\partial \mu}{\partial X}\right)_{T,N} \tag{2.199}$$

$$\left(\frac{\partial S}{\partial y}\right)_{T,N} = \left(\frac{\partial X}{\partial T}\right)_{y,N} \qquad \left(\frac{\partial S}{\partial N}\right)_{T,y} = -\left(\frac{\partial \mu}{\partial T}\right)_{y,N} \qquad \left(\frac{\partial X}{\partial N}\right)_{T,y} = -\left(\frac{\partial \mu}{\partial y}\right)_{T,N} \tag{2.200}$$

$$\left(\frac{\partial S}{\partial X}\right)_{T,\mu} = -\left(\frac{\partial y}{\partial T}\right)_{X,\mu} \qquad \left(\frac{\partial S}{\partial \mu}\right)_{T,X} = \left(\frac{\partial N}{\partial T}\right)_{X,\mu} \qquad \left(\frac{\partial y}{\partial \mu}\right)_{T,X} = -\left(\frac{\partial N}{\partial X}\right)_{T,\mu} \quad . \tag{2.201}$$

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

$$E_{\mathrm{A}} + E_{\mathrm{B}} = E \quad , \quad V_{\mathrm{A}} + V_{\mathrm{B}} = V \quad , \quad N_{\mathrm{A}} + N_{\mathrm{B}} = N \quad , \tag{2.202}$$

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.*

$$
\begin{aligned}
dS = \Bigg[\bigg(\frac{\partial S_{\rm A}}{\partial E_{\rm A}}\bigg)_{V_{\rm A},N_{\rm A}} - \bigg(\frac{\partial S_{\rm B}}{\partial E_{\rm B}}\bigg)_{V_{\rm B},N_{\rm B}}\Bigg]\,dE_{\rm A} &+ \Bigg[\bigg(\frac{\partial S_{\rm A}}{\partial V_{\rm A}}\bigg)_{E_{\rm A},N_{\rm A}} - \bigg(\frac{\partial S_{\rm B}}{\partial V_{\rm B}}\bigg)_{E_{\rm B},N_{\rm B}}\Bigg]\,dV_{\rm A}\\
&+ \Bigg[\bigg(\frac{\partial S_{\rm A}}{\partial N_{\rm A}}\bigg)_{E_{\rm A},V_{\rm A}} - \bigg(\frac{\partial S_{\rm B}}{\partial N_{\rm B}}\bigg)_{E_{\rm B},V_{\rm B}}\Bigg]\,dN_{\rm A} \quad . \qquad (2.203)
\end{aligned}
$$

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

$$
\frac{1}{T} = \bigg(\frac{\partial S}{\partial E}\bigg)_{V,N} \quad , \quad \frac{p}{T} = \bigg(\frac{\partial S}{\partial V}\bigg)_{E,N} \quad , \quad \frac{\mu}{T} = -\bigg(\frac{\partial S}{\partial N}\bigg)_{E,V} \quad . \qquad (2.204)
$$

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

$$
T_{\rm A} = T_{\rm B} \qquad \text{(thermal equilibrium)} \qquad (2.205)
$$

$$
\frac{p_{\rm A}}{T_{\rm A}} = \frac{p_{\rm B}}{T_{\rm B}} \qquad \text{(mechanical equilibrium)} \qquad (2.206)
$$

$$
\frac{\mu_{\rm A}}{T_{\rm A}} = \frac{\mu_{\rm B}}{T_{\rm B}} \qquad \text{(chemical equilibrium)} \qquad (2.207)
$$

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 $\widetilde{N} = 2N$. 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 slightly different energy, volume, and particle number than the right, as depicted in fig. 2.16. Does the entropy increase or decrease? We have

$$
\begin{aligned}
\Delta S &= S(E+\Delta E, V+\Delta V, N+\Delta N) + S(E-\Delta E, V-\Delta V, N-\Delta N) - S(2E, 2V, 2N)\\
&= \frac{\partial^2 S}{\partial E^2}\,(\Delta E)^2 + \frac{\partial^2 S}{\partial V^2}\,(\Delta V)^2 + \frac{\partial^2 S}{\partial N^2}\,(\Delta N)^2 \qquad (2.208)\\
&\qquad + 2\,\frac{\partial^2 S}{\partial E\,\partial V}\,\Delta E\,\Delta V + 2\,\frac{\partial^2 S}{\partial E\,\partial N}\,\Delta E\,\Delta N + 2\,\frac{\partial^2 S}{\partial V\,\partial N}\,\Delta V\,\Delta N \quad .
\end{aligned}
$$

Thus, we can write

$$
\Delta S = \sum_{i,j} Q_{ij}\,\Psi_i\,\Psi_j \quad , \qquad (2.209)
$$

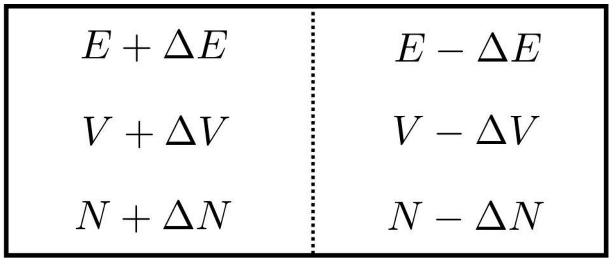


Figure 2.16: To check for an instability, we compare the energy of a system to its total energy when we reapportion its energy, volume, and particle number slightly unequally.

where

$$Q = \begin{pmatrix} \frac{\partial^2 S}{\partial E^2} & \frac{\partial^2 S}{\partial E\,\partial V} & \frac{\partial^2 S}{\partial E\,\partial N} \\ \\ \frac{\partial^2 S}{\partial E\,\partial V} & \frac{\partial^2 S}{\partial V^2} & \frac{\partial^2 S}{\partial V\,\partial N} \\ \\ \frac{\partial^2 S}{\partial E\,\partial N} & \frac{\partial^2 S}{\partial V\,\partial N} & \frac{\partial^2 S}{\partial N^2} \end{pmatrix} \tag{2.210}$$

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 η 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 2×2 submatrix of Q. A general symmetric 2×2 matrix may be written

$$Q = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \tag{2.211}$$

It is easy to solve for the eigenvalues of Q. One finds

$$\lambda_\pm = \left(\frac{a + c}{2}\right) \pm \sqrt{\left(\frac{a - c}{2}\right)^2 + b^2} \quad . \tag{2.212}$$

In order for Q to be negative definite, we require $\lambda_+ < 0$ and $\lambda_- < 0$. Thus, $\mathsf{Tr}\, Q = a + c = \lambda_+ + \lambda_- < 0$ and $\mathsf{det}\, Q = ac - b^2 = \lambda_+ \lambda_- > 0$. Taken together, these conditions require

$$a < 0 \quad , \quad c < 0 \quad , \quad ac > b^2 \quad . \tag{2.213}$$

Going back to thermodynamic variables, this requires

$$\frac{\partial^2 S}{\partial E^2} < 0 \qquad , \qquad \frac{\partial^2 S}{\partial V^2} < 0 \qquad , \qquad \frac{\partial^2 S}{\partial E^2} \cdot \frac{\partial^2 S}{\partial V^2} > \left(\frac{\partial^2 S}{\partial E\,\partial V}\right)^2 \quad . \tag{2.214}$$

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 2×2 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 $\partial^2 S/\partial E^2 = -T^{-2}(\partial T/\partial E)_V = -C_V/T^2 < 0$, 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

$$G(T_0, p_0) = E - T_0\, S + p_0\, V \quad . \tag{2.215}$$

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 ΔS and ΔV, we have

$$\begin{aligned} \Delta G = \left[\left(\frac{\partial E}{\partial S}\right)_V - T_0\right] \Delta S + \left[\left(\frac{\partial E}{\partial V}\right)_S + p_0\right] \Delta V \\ + \frac{1}{2}\left[\frac{\partial^2 E}{\partial S^2}\,(\Delta S)^2 + 2\,\frac{\partial^2 E}{\partial S\,\partial V}\,\Delta S\,\Delta V + \frac{\partial^2 E}{\partial V^2}\,(\Delta V)^2\right] + \ldots \quad . \end{aligned} \tag{2.216}$$

Equilibrium requires that the coefficients of ΔS and ΔV 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

$$Q = \begin{pmatrix} \frac{\partial^2 E}{\partial S^2} & \frac{\partial^2 E}{\partial S\,\partial V} \\ \\ \frac{\partial^2 E}{\partial S\,\partial V} & \frac{\partial^2 E}{\partial V^2} \end{pmatrix} \quad . \tag{2.217}$$

Thus, we have the following three conditions:

$$\frac{\partial^2 E}{\partial S^2} = \left(\frac{\partial T}{\partial S}\right)_V = \frac{T}{C_V} > 0 \tag{2.218}$$

$$\frac{\partial^2 E}{\partial V^2} = -\left(\frac{\partial p}{\partial V}\right)_S = \frac{1}{V\kappa_S} > 0 \tag{2.219}$$

$$\frac{\partial^2 E}{\partial S^2} \cdot \frac{\partial^2 E}{\partial V^2} - \left(\frac{\partial^2 E}{\partial S\,\partial V}\right)^2 = \frac{T}{V\kappa_S\, C_V} - \left(\frac{\partial T}{\partial V}\right)_S^2 > 0 \quad . \tag{2.220}$$

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

$$\frac{C_V}{T} > 0 \qquad , \qquad \kappa_S > 0 \qquad , \qquad \alpha_S > \sqrt{\frac{\kappa_S\, C_V}{VT}} \quad . \tag{2.221}$$

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

$$\Delta S = S_{\rm f} - S_{\rm i} = \int\limits_{V_{\rm i}}^{V_{\rm f}} \! dV \left(\frac{\partial S}{\partial V}\right)_{T,N} \quad . \tag{2.222}$$

But from a Maxwell equation deriving from F, we have

$$\left(\frac{\partial S}{\partial V}\right)_{T,N} = \left(\frac{\partial p}{\partial T}\right)_{V,N} \quad , \tag{2.223}$$

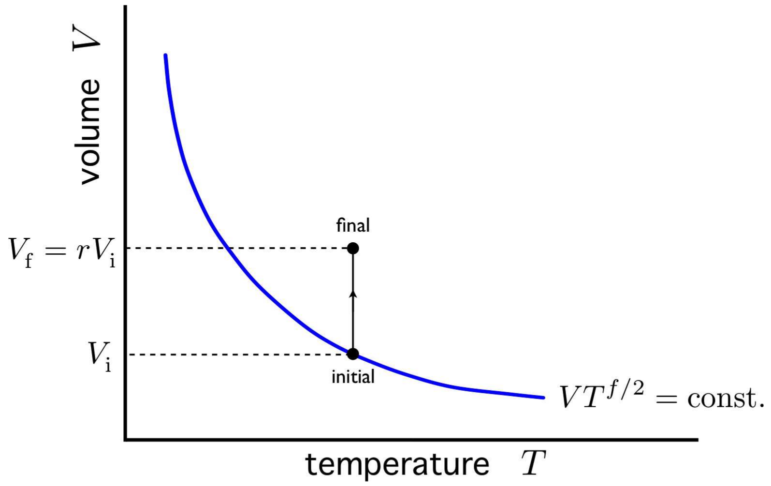


Figure 2.17: Adiabatic free expansion via a thermal path. The initial and final states do not lie along an adabat! Rather, for an ideal gas, the initial and final states lie along an isotherm.

hence

$$\Delta S = \int\limits_{V_\mathrm{i}}^{V_\mathrm{f}} dV \left(\frac{\partial p}{\partial T}\right)_{V,N} \quad . \tag{2.224}$$

For an ideal gas, we can use the equation of state $pV = Nk_\mathrm{B}T$ to obtain

$$\left(\frac{\partial p}{\partial T}\right)_{V,N} = \frac{Nk_\mathrm{B}}{V} \quad . \tag{2.225}$$

The integral can now be computed:

$$\Delta S = \int\limits_{V_\mathrm{i}}^{rV_\mathrm{i}} dV \, \frac{Nk_\mathrm{B}}{V} = Nk_\mathrm{B} \log r \quad , \tag{2.226}$$

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 $E(T, V, N) = \frac{1}{2} f N k_\mathrm{B} T$ for an ideal gas, independent of the volume. In general we should have

$$E(T, V, N) = N\,\phi\big(T, \tfrac{V}{N}\big) \quad . \tag{2.227}$$

For the ideal gas, $\phi\big(T, V/N\big) = \frac{1}{2} f k_\mathrm{B} T$ 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 $E = F + TS$,

$$\left(\frac{\partial E}{\partial V}\right)_{T,N} = \left(\frac{\partial F}{\partial V}\right)_{T,N} + T\left(\frac{\partial S}{\partial V}\right)_{T,N} = -p + T\left(\frac{\partial p}{\partial T}\right)_{V,N} \quad , \tag{2.228}$$

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

$$dE = T\left(\frac{\partial S}{\partial T}\right)_{V,N} dT + \left[T\left(\frac{\partial S}{\partial V}\right)_{T,N} - p\right] dV - \left[T\left(\frac{\partial \mu}{\partial T}\right)_{V,N} + \mu\right] dN \quad . \tag{2.229}$$

Now read off $(\partial E/\partial V)_{T,N}$ and use the same Maxwell relation as before to recover eqn. 2.228. Applying this result to the ideal gas law $pV = Nk_\mathrm{B}T$ results in the vanishing of the RHS, hence for any substance obeying the ideal gas law we must have

$$E(T, V, N) = \nu\,\varepsilon(T) = N\varepsilon(T)/N_\mathrm{A} \quad . \tag{2.230}$$

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	7404	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.*)

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[15]. This is not true, however, for the van der Waals equation of state,

$$\left(p+\frac{a}{v^2}\right)(v-b) = RT \quad , \tag{2.231}$$

where $v = N_{\mathrm{A}}V/N$ is the molar volume. We then find (always assuming constant N),

$$\left(\frac{\partial E}{\partial V}\right)_T = \left(\frac{\partial \varepsilon}{\partial v}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_V - p = \frac{a}{v^2} \quad , \tag{2.232}$$

where $E(T,V,N) \equiv \nu\,\varepsilon(T,v)$. We can integrate this to obtain

$$\varepsilon(T,v) = \omega(T) - \frac{a}{v} \quad , \tag{2.233}$$

where $\omega(T)$ is arbitrary. From eqn. 2.31, we immediately have

$$c_V = \left(\frac{\partial \varepsilon}{\partial T}\right)_v = \omega'(T) \quad . \tag{2.234}$$

What about c_p? This requires a bit of work. We start with eqn. 2.32,

$$c_p = \left(\frac{\partial \varepsilon}{\partial T}\right)_p + p\left(\frac{\partial v}{\partial T}\right)_p = c_V + \left(p+\frac{a}{v^2}\right)\left(\frac{\partial v}{\partial T}\right)_p \quad . \tag{2.235}$$

---

[15] Note $V/N = v/N_{\mathrm{A}}$.

We may rewrite the equation of state as

$$p = \frac{RT}{v-b} - \frac{a}{v^2} \quad . \tag{2.236}$$

Taking the differential of both sides,

$$\begin{aligned} dp &= \left(\frac{\partial p}{\partial T}\right)_{\!v} dT + \left(\frac{\partial p}{\partial v}\right)_{\!T} dv \\ &= \frac{R}{v-b}\, dT + \left[-\frac{RT}{(v-b)^2} + \frac{2a}{v^3}\right] dv \quad . \end{aligned} \tag{2.237}$$

Setting $dp = 0$, we may read off

$$\left(\frac{\partial v}{\partial T}\right)_{\!p} = -\left(\frac{\partial p}{\partial T}\right)_{\!v} \bigg/ \left(\frac{\partial p}{\partial v}\right)_{\!T} \quad , \tag{2.238}$$

which follows from eqn. 2.483 in the mathematical appendix §2.17 below. Putting this all together, we obtain

$$\left(\frac{\partial v}{\partial T}\right)_{\!p} = \frac{Rv^3(v-b)}{RTv^3 - 2a(v-b)^2} \quad . \tag{2.239}$$

One immediate result is the following expression for the isobaric thermal expansion coefficient,

$$\alpha_p = \frac{1}{v}\left(\frac{\partial v}{\partial T}\right)_{\!p} = \frac{Rv^2(v-b)}{RTv^3 - 2a(v-b)^2} \quad . \tag{2.240}$$

Another result is the difference $c_p - c_V$ from eqn. 2.235,

$$c_p - c_V = \left(p + \frac{a}{v^2}\right)\left(\frac{\partial v}{\partial T}\right)_{\!p} = \frac{R^2Tv^3}{RTv^3 - 2a(v-b)^2} \quad . \tag{2.241}$$

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. 2.233 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. 2.241, and

$$\begin{aligned} c_V^{\text{VDW}} &= \tfrac{1}{2} f R \\ c_p^{\text{VDW}} &= \tfrac{1}{2} f R + \frac{R^2Tv^3}{RTv^3 - 2a(v-b)^2} \quad . \end{aligned} \tag{2.242}$$

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

$$dZ = \left(\frac{\partial Z}{\partial T}\right)_{p,v} dT + \left(\frac{\partial Z}{\partial p}\right)_{T,v} dp + \left(\frac{\partial Z}{\partial v}\right)_{T,p} dv = 0 \quad . \tag{2.243}$$

We then have

$$\begin{aligned}
\left(\frac{\partial T}{\partial p}\right)_{v} &= -\left(\frac{\partial Z}{\partial p}\right)_{T,v} \Bigg/ \left(\frac{\partial Z}{\partial T}\right)_{p,v} \\
\left(\frac{\partial p}{\partial v}\right)_{T} &= -\left(\frac{\partial Z}{\partial v}\right)_{T,p} \Bigg/ \left(\frac{\partial Z}{\partial p}\right)_{T,v} \\
\left(\frac{\partial v}{\partial T}\right)_{p} &= -\left(\frac{\partial Z}{\partial T}\right)_{p,v} \Bigg/ \left(\frac{\partial Z}{\partial v}\right)_{T,p} \quad .
\end{aligned} \tag{2.244}$$

2.10.5 Thermodynamic response functions

Consider the entropy S expressed as a function of T, V, and N:

$$dS = \left(\frac{\partial S}{\partial T}\right)_{V,N} dT + \left(\frac{\partial S}{\partial V}\right)_{T,N} dV + \left(\frac{\partial S}{\partial N}\right)_{T,V} dN \quad . \tag{2.245}$$

Dividing by dT, multiplying by T, and assuming $dN = 0$ throughout, we have

$$C_p - C_V = T \left(\frac{\partial S}{\partial V}\right)_{T} \left(\frac{\partial V}{\partial T}\right)_{p} \quad . \tag{2.246}$$

Appealing to a Maxwell relation derived from $F(T, V, N)$, and then appealing to eqn. 2.482, we have

$$\left(\frac{\partial S}{\partial V}\right)_{T} = \left(\frac{\partial p}{\partial T}\right)_{V} = -\left(\frac{\partial p}{\partial V}\right)_{T} \left(\frac{\partial V}{\partial T}\right)_{p} . \tag{2.247}$$

This allows us to write

$$C_p - C_V = -T \left(\frac{\partial p}{\partial V}\right)_{T} \left(\frac{\partial V}{\partial T}\right)_{p}^{2} \quad . \tag{2.248}$$

We define the *response functions*,

$$\text{isothermal compressibility:} \quad \kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{T} = -\frac{1}{V}\frac{\partial^2 G}{\partial p^2} \tag{2.249}$$

$$\text{adiabatic compressibility:} \quad \kappa_S = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{S} = -\frac{1}{V}\frac{\partial^2 H}{\partial p^2} \tag{2.250}$$

$$\text{thermal expansivity:} \quad \alpha_p = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{p} \quad . \tag{2.251}$$

Thus,

$$C_p - C_V = V \, \frac{T\alpha_p^2}{\kappa_T} \quad , \tag{2.252}$$

or, in terms of intensive quantities,

$$c_p - c_V = \frac{v \, T\alpha_p^2}{\kappa_T} \quad , \tag{2.253}$$

where, as always, $v = VN_{\rm A}/N$ is the molar volume.

This above relation generalizes to any conjugate force-displacement pair $(-p, V) \to (y, X)$:

$$C_y - C_X = -T \left(\frac{\partial y}{\partial T}\right)_{\!X} \left(\frac{\partial X}{\partial T}\right)_{\!y} = T \left(\frac{\partial y}{\partial X}\right)_{\!T} \left(\frac{\partial X}{\partial T}\right)_{\!y}^{\!2} \quad . \tag{2.254}$$

For example, we could have $(y, X) = (H^\alpha, M^\alpha)$.

A similar relationship can be derived between the compressibilities κ_T and κ_S. We then clearly must start with the volume, writing

$$dV = \left(\frac{\partial V}{\partial p}\right)_{\!S,N} dp + \left(\frac{\partial V}{\partial S}\right)_{\!p,N} dS + \left(\frac{\partial V}{\partial p}\right)_{\!S,p} dN \quad . \tag{2.255}$$

Dividing by dp, multiplying by $-V^{-1}$, and keeping N constant, we have

$$\kappa_T - \kappa_S = -\frac{1}{V} \left(\frac{\partial V}{\partial S}\right)_{\!p} \left(\frac{\partial S}{\partial p}\right)_{\!T} \quad . \tag{2.256}$$

Again we appeal to a Maxwell relation, writing

$$\left(\frac{\partial S}{\partial p}\right)_{\!T} = -\left(\frac{\partial V}{\partial T}\right)_{\!p} \quad , \tag{2.257}$$

and after invoking the chain rule,

$$\left(\frac{\partial V}{\partial S}\right)_{\!p} = \left(\frac{\partial V}{\partial T}\right)_{\!p} \left(\frac{\partial T}{\partial S}\right)_{\!p} = \frac{T}{C_p} \left(\frac{\partial V}{\partial T}\right)_{\!p} \quad , \tag{2.258}$$

we obtain

$$\kappa_T - \kappa_S = \frac{v \, T\alpha_p^2}{c_p} \quad . \tag{2.259}$$

Comparing eqns. 2.253 and 2.259, we find

$$(c_p - c_V)\, \kappa_T = (\kappa_T - \kappa_S)\, c_p = v \, T\alpha_p^2 \quad . \tag{2.260}$$

This result entails

$$\frac{c_p}{c_V} = \frac{\kappa_T}{\kappa_S} \quad . \tag{2.261}$$

The corresponding result for magnetic systems is[16]

$$(c_H - c_M)\,\chi_T = (\chi_T - \chi_S)\,c_H = T\bigg(\frac{\partial m}{\partial T}\bigg)^2_H \quad , \tag{2.262}$$

where $m = M/\nu$ is the magnetization per mole of substance, and

$$\text{isothermal susceptibility:} \quad \chi_T = \bigg(\frac{\partial M}{\partial H}\bigg)_T = -\frac{1}{\nu}\,\frac{\partial^2 G}{\partial H^2} \tag{2.263}$$

$$\text{adiabatic susceptibility:} \quad \chi_S = \bigg(\frac{\partial M}{\partial H}\bigg)_S = -\frac{1}{\nu}\,\frac{\partial^2 \mathsf{H}}{\partial H^2} \quad . \tag{2.264}$$

Here the enthalpy and Gibbs free energy are

$$\mathsf{H} = E - HM \qquad\qquad d\mathsf{H} = T\,dS - M\,dH \tag{2.265}$$

$$G = E - TS - HM \qquad\qquad dG = -S\,dT - M\,dH \quad . \tag{2.266}$$

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

$$\chi_T^{\alpha\beta} = \bigg(\frac{\partial M^\alpha}{\partial H^\beta}\bigg)_T = -\frac{1}{\nu}\,\frac{\partial^2 G}{\partial H^\alpha\,\partial H^\beta} \tag{2.267}$$

$$\chi_S^{\alpha\beta} = \bigg(\frac{\partial M^\alpha}{\partial H^\beta}\bigg)_S = -\frac{1}{\nu}\,\frac{\partial^2 H}{\partial H^\alpha\,\partial H^\beta} \quad . \tag{2.268}$$

In this case, the enthalpy and Gibbs free energy are

$$\mathsf{H} = E - \boldsymbol{H}\cdot\boldsymbol{M} \qquad\qquad d\mathsf{H} = T\,dS - \boldsymbol{M}\cdot d\boldsymbol{H} \tag{2.269}$$

$$G = E - TS - \boldsymbol{H}\cdot\boldsymbol{M} \qquad\qquad dG = -S\,dT - \boldsymbol{M}\cdot d\boldsymbol{H} \quad . \tag{2.270}$$

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,

$$S_{\rm f} = S_{\rm i} + Nk_{\rm B}\log\bigg(\frac{V_{\rm f}}{V_{\rm i}}\bigg) \quad . \tag{2.271}$$

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

$$0 = dE = \bigg(\frac{\partial E}{\partial V}\bigg)_T dV + \bigg(\frac{\partial E}{\partial T}\bigg)_V dT \tag{2.272}$$

---

[16] Recall the subtle font difference in notation: H is enthalpy while H is magnetic field.

hence

$$\left(\frac{\partial T}{\partial V}\right)_{E} = -\frac{(\partial E/\partial V)_T}{(\partial E/\partial T)_V} = -\frac{1}{C_V}\left(\frac{\partial E}{\partial V}\right)_T \quad . \tag{2.273}$$

We also have

$$\left(\frac{\partial E}{\partial V}\right)_{T} = T\left(\frac{\partial S}{\partial V}\right)_T - p = T\left(\frac{\partial p}{\partial T}\right)_V - p \quad . \tag{2.274}$$

Thus,

$$\left(\frac{\partial T}{\partial V}\right)_{E} = \frac{1}{C_V}\left[p - T\left(\frac{\partial p}{\partial T}\right)_V\right] \quad . \tag{2.275}$$

Note that the term on the RHS vanishes for any system obeying the ideal gas law. For a nonideal gas,

$$\Delta T = \int\limits_{V_\mathrm{i}}^{V_\mathrm{f}} dV \left(\frac{\partial T}{\partial V}\right)_{E} \quad , \tag{2.276}$$

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

$$p - T\left(\frac{\partial p}{\partial T}\right)_V = -\frac{a}{v^2} = -\frac{a\nu^2}{V^2} \quad . \tag{2.277}$$

In §2.10.3 we concluded that $C_V = \frac{1}{2} f\nu R$ for the van der Waals gas, hence

$$\Delta T = -\frac{2a\nu}{fR}\int\limits_{V_\mathrm{i}}^{V_\mathrm{f}} \frac{dV}{V^2} = \frac{2a}{fR}\left(\frac{1}{v_\mathrm{f}} - \frac{1}{v_\mathrm{i}}\right) \quad . \tag{2.278}$$

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-Thompson 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

$$W = \int\limits_0^{V_\mathrm{f}} dV\, p_\mathrm{f} - \int\limits_0^{V_\mathrm{i}} dV\, p_\mathrm{i} = p_\mathrm{f} V_\mathrm{f} - p_\mathrm{i} V_\mathrm{i} \quad . \tag{2.279}$$

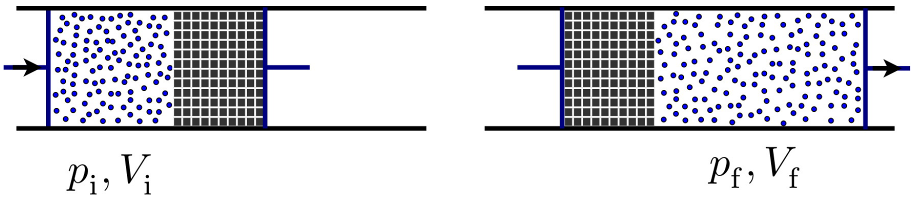


Figure 2.18: In a throttle, a gas is pushed through a porous plug separating regions of different pressure. The change in energy is the work done, hence enthalpy is conserved during the throttling process.

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

$$\begin{gathered} E_{\rm i} + p_{\rm i} V_{\rm i} = E_{\rm f} + p_{\rm f} V_{\rm f} \\ \mathsf{H}_{\rm i} = \mathsf{H}_{\rm f} \quad , \end{gathered} \tag{2.280}$$

where H is enthalpy. Thus, the throttling process is *isenthalpic*. We can therefore study it by defining a fictitious thermodynamic path along which $d\mathsf{H} = 0$. The, choosing T and p as state variables,

$$0 = d\mathsf{H} = \left(\frac{\partial \mathsf{H}}{\partial T}\right)_{\!p} dT \, + \left(\frac{\partial \mathsf{H}}{\partial p}\right)_{\!T} dp \tag{2.281}$$

hence

$$\left(\frac{\partial T}{\partial p}\right)_{\!\mathsf{H}} = -\frac{(\partial \mathsf{H}/\partial p)_T}{(\partial \mathsf{H}/\partial T)_p} \quad . \tag{2.282}$$

The numerator on the RHS is computed by writing $d\mathsf{H} = T\,dS + V\,dp$ and then dividing by dp, to obtain

$$\left(\frac{\partial \mathsf{H}}{\partial p}\right)_{\!T} = V + T\left(\frac{\partial S}{\partial p}\right)_{\!T} = V - T\left(\frac{\partial V}{\partial T}\right)_{\!p} \quad . \tag{2.283}$$

The denominator is

$$\left(\frac{\partial \mathsf{H}}{\partial T}\right)_{\!p} = \left(\frac{\partial \mathsf{H}}{\partial S}\right)_{\!p}\left(\frac{\partial S}{\partial T}\right)_{\!p} = T\left(\frac{\partial S}{\partial T}\right)_{\!p} = C_p \quad . \tag{2.284}$$

Thus,

$$\left(\frac{\partial T}{\partial p}\right)_{\!\mathsf{H}} = \frac{1}{c_p}\left[T\left(\frac{\partial v}{\partial T}\right)_{\!p} - v\right] = \frac{v}{c_p}\left(T\alpha_p - 1\right) \quad , \tag{2.285}$$

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. 2.240,

$$T\alpha_p = \frac{T}{v}\left(\frac{\partial v}{\partial T}\right)_{\!p} = \frac{RT/v}{p - \frac{a}{v^2} + \frac{2ab}{v^3}} = \frac{v-b}{v - \frac{2a}{RT}\left(\frac{v-b}{v}\right)^2} \quad . \tag{2.286}$$

Assuming $v \gg a/RT$ and $v \gg b$, we have

$$\left(\frac{\partial T}{\partial p}\right)_{\!H} = \frac{1}{c_p}\left(\frac{2a}{RT} - b\right) \quad . \tag{2.287}$$

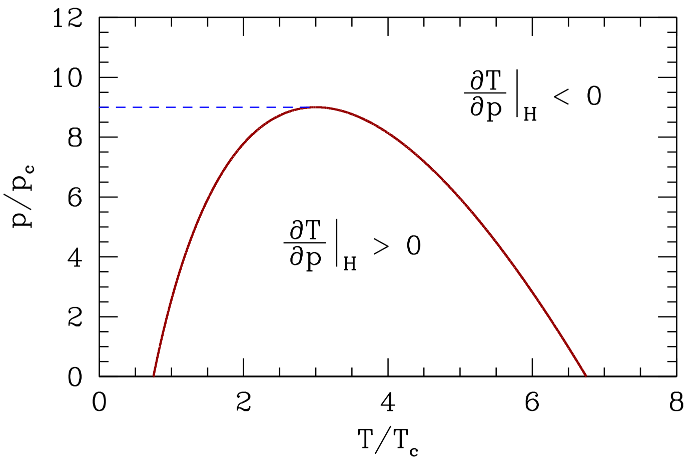


Figure 2.19: Inversion temperature $T^*(p)$ for the van der Waals gas. Pressure and temperature are given in terms of $p_c = a/27b^2$ and $T_c = 8a/27bR$, respectively.

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. 2.286 it is easy to derive

$$\frac{b}{v} = 1 - \sqrt{\frac{bRT}{2a}} \quad . \tag{2.288}$$

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

$$p = -\frac{3RT}{2b} + \sqrt{\frac{8aRT}{b^3}} - \frac{a}{b^2} \quad . \tag{2.289}$$

This is a quadratic equation for T, the solution of which is

$$T^*(p) = \frac{2a}{9\,bR}\left(2 \pm \sqrt{1 - \frac{3b^2p}{a}}\,\right)^2 \quad . \tag{2.290}$$

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

$$v_c = 3b \quad , \quad p_c = \frac{a}{27\,b^2} \quad , \quad T_c = \frac{8a}{27\,bR} \quad . \tag{2.291}$$

Values for p_c, T_c, and v_c are provided in table 2.3. If we define $v = v/v_c$, $p = p/p_c$, and $T = T/T_c$, then the van der Waals equation of state may be written in dimensionless form:

$$\left(p + \frac{3}{v^2}\right)(3v - 1) = 8T \quad . \tag{2.292}$$

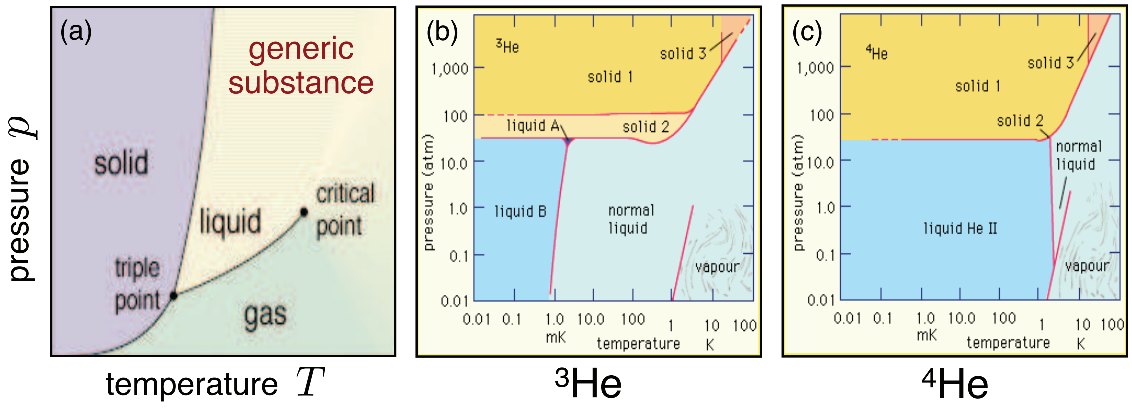


Figure 2.20: (a) Typical thermodynamic phase diagram of a single component p - v - T system, showing triple point (three phase coexistence) and critical point. (*Credit: Univ. of Helsinki.*) Also shown: phase diagrams for ^{3}He (b) and ^{4}He (c). What a difference a neutron makes! (*Credit: Brittanica.*)

In terms of the scaled parameters, the equation for the inversion curve $(\partial T/\partial p)_H = 0$ becomes

$$p = 9 - 36\Big(1 - \sqrt{\tfrac{1}{3}T}\,\Big)^2 \qquad \Longleftrightarrow \qquad T = 3\Big(1 \pm \sqrt{1 - \tfrac{1}{9}\,p}\,\Big)^2 \quad . \tag{2.293}$$

Thus, there is no inversion for $p > 9\,p_c$. We are usually interested in the upper inversion temperature, T_2^*, corresponding to the upper sign in eqn. 2.290. The maximum inversion temperature occurs for $p = 0$, where $T^*_{\rm max} = 2a/bR = 27T_c/4$. For H_2, from the data in table 2.3, we find $T^*_{\rm max}(H_2) = 224\,K$, which is within 10% of the experimentally measured value of 205 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 $2\,H_2 + O_2 \longrightarrow 2\,H_2O$, which releases energy, and we have a nice explosion.

2.11 Phase Transitions and Phase Equilibria

A typical phase diagram of a p - v - T system is shown in the 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 the thermodynamic potentials are singular. The order of the singularity is often taken as a classification of the phase transition. *I.e.* if the thermodynamic potentials E, F, G, and H have discontinuous or divergent $m^{\rm th}$ derivatives, the transition between the respective phases is said to be $m^{\rm th}$ *order*. Modern theories of phase transitions generally only recognize two possibilities: *first order transitions*, where the *order parameter* changes *discontinuously* through the transition, and *second order transitions*, where the order parameter vanishes continuously at the boundary from ordered to disordered phases[17]. We'll discuss order parameters when we get to chapter 7.

---

[17] Some exotic phase transitions in quantum matter, which do not quite fit the usual classification schemes, have recently been proposed.

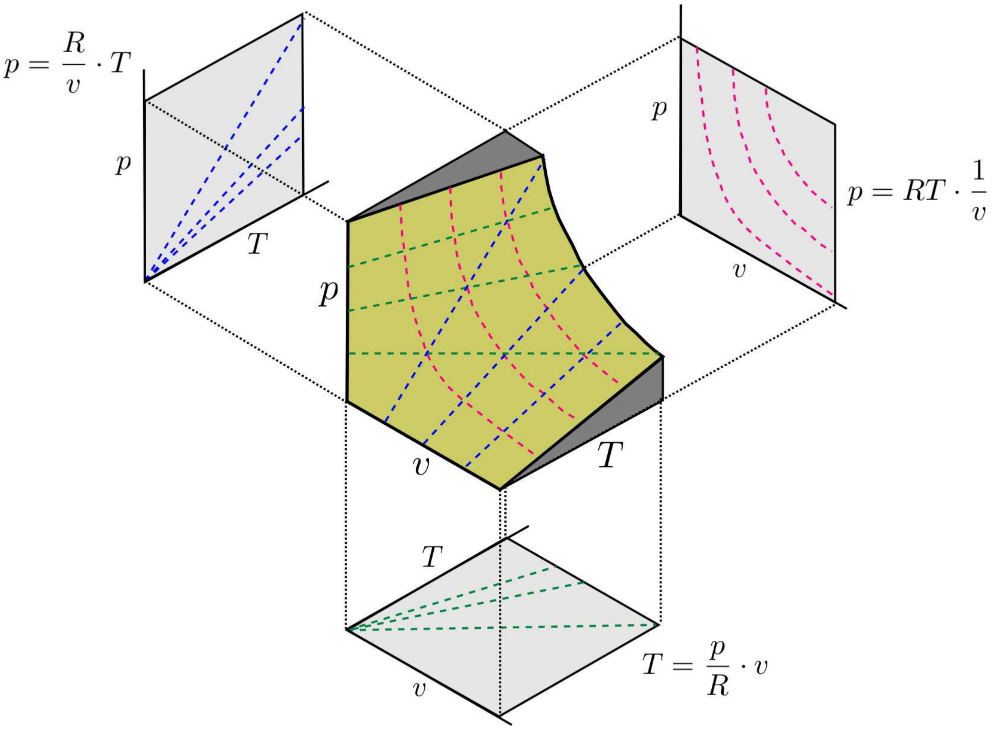


Figure 2.21: The surface $p(v,T) = RT/v$ corresponding to the ideal gas equation of state, and its projections onto the (p,T), (p,v), and (T,v) planes.

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.

2.11.1 p-v-T surfaces

The equation of state for a single component system may be written as

$$f(p,v,T) = 0 \quad . \tag{2.294}$$

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_c, the isotherms become singular. Precisely at $T = T_c$, the isotherm $p = p(v,T_c)$ becomes perfectly horizontal at $v = v_c$, which is the *critical molar volume*. This means that the isothermal compressibility, $\kappa_T = -v^{-1}(\partial v/\partial p)_T$ *diverges* at $T = T_c$. Below

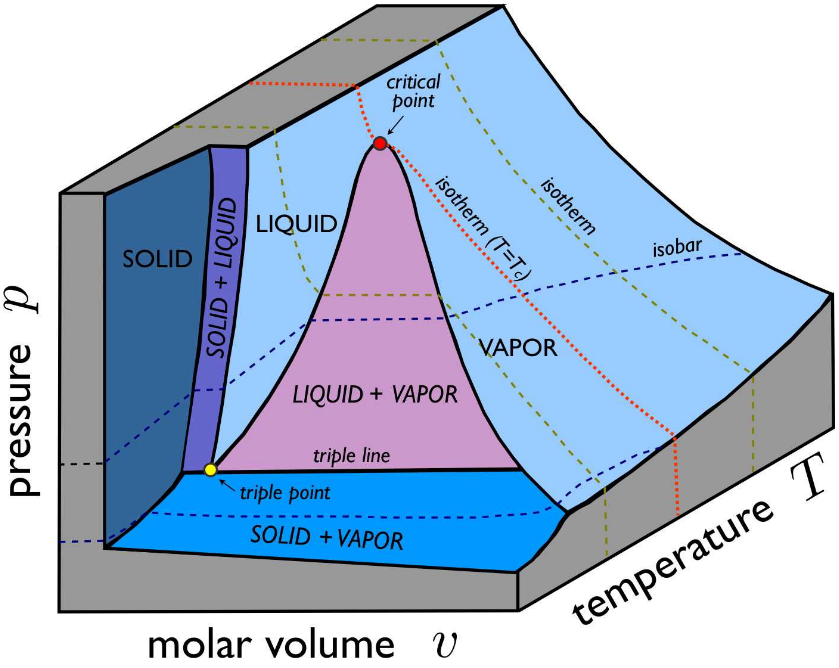


Figure 2.22: A p - v - T surface for a substance which contracts upon freezing. The red dot is the *critical point* and the red dashed line is the *critical isotherm*. The yellow dot is the *triple point* at which there is three phase coexistence of solid, liquid, and vapor.

T_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* $(\mathrm{T}_{\mathrm{t}}\,,\,p_{\mathrm{t}})$ lies at the confluence of these three coexistence regions. For H_2O, the location of the triple point and critical point are given by

$$\begin{aligned} \mathrm{T}_{\mathrm{t}} &= 273.16\,\mathrm{K} & T_{\mathrm{c}} &= 647\,\mathrm{K} \\ \mathrm{p}_{\mathrm{t}} &= 611.7\,\mathrm{Pa} = 6.037\times 10^{-3}\,\mathrm{atm} & \mathrm{p}_{\mathrm{c}} &= 22.06\,\mathrm{MPa} = 217.7\,\mathrm{atm} \end{aligned}$$

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 μ, according to eqn. 2.112. Let us define $g \equiv G/\nu = N_{\!\scriptscriptstyle\mathrm{A}}\mu$, the Gibbs free energy per mole. Then

$$dg = -s\,dT + v\,dp \quad , \tag{2.295}$$

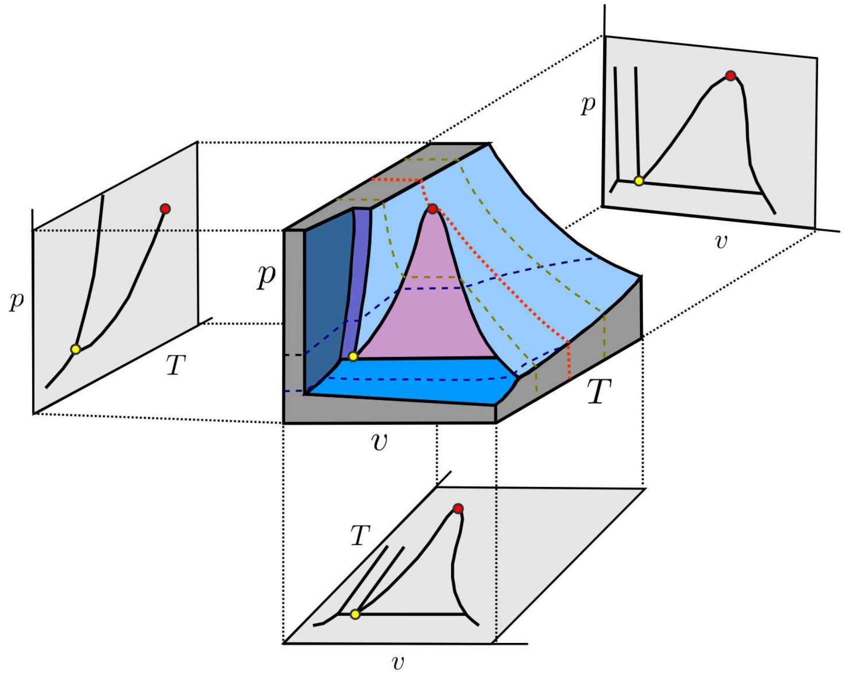


Figure 2.23: Equation of state for a substance which expands upon freezing, projected to the (v, T) and (v, p) and (T, p) planes.

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

$$dg_1 = -s_1\,dT + v_1\,dp = -s_2\,dT + v_2\,dp = dg_2 \quad . \tag{2.296}$$

Therefore, along the coexistence curve we must have

$$\bigg(\frac{dp}{dT}\bigg)_{\rm coex} = \frac{s_2 - s_1}{v_2 - v_1} = \frac{\ell}{T\,\Delta v} \quad , \tag{2.297}$$

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 ℓ per mole must be supplied in order to change from phase #1 to phase #2, even without changing p or T. If ℓ 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*,

$$\bigg(\frac{dp}{dT}\bigg)_{\rm liq-gas} = \frac{\ell}{T\,\Delta v} \approx \frac{p\,\ell}{RT^2} \quad . \tag{2.298}$$

If ℓ remains constant throughout a section of the liquid-gas coexistence curve, we may integrate the above equation to get

$$\frac{dp}{p} = \frac{\ell}{R}\,\frac{dT}{T^2} \qquad \Rightarrow \qquad p(T) = p(T_0)\,e^{\ell/RT_0}\,e^{-\ell/RT} \quad . \tag{2.299}$$

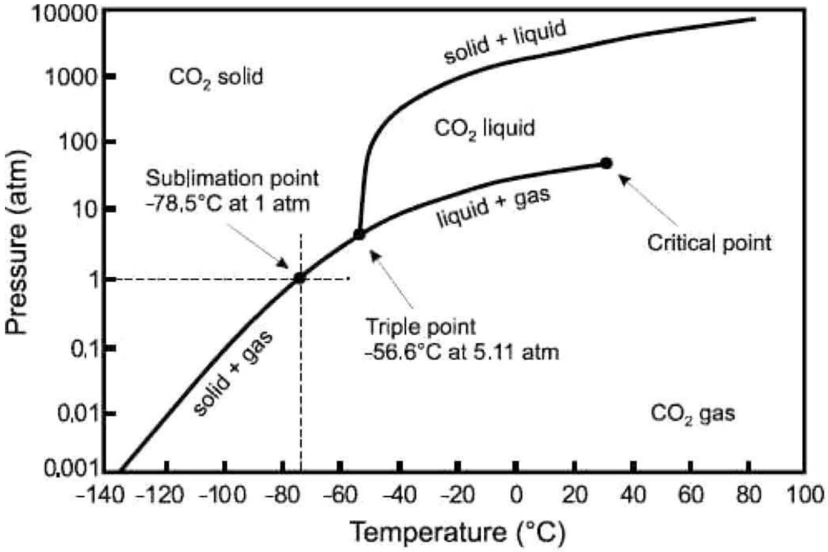


Figure 2.24: Phase diagram for $\mathrm{CO_2}$ in the (p,T) plane. (*Source: www.scifun.org.*)

2.11.3 Liquid-solid line in $\mathrm{H_2O}$

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[18]. For example at $T = 273.1\,\mathrm{K}$ and $p = 1\,\mathrm{atm}$,

$$\tilde{v}_{\mathrm{water}} = 1.00013\,\mathrm{cm}^3/\mathrm{g} \qquad , \qquad \tilde{v}_{\mathrm{ice}} = 1.0907\,\mathrm{cm}^3/\mathrm{g} \quad . \tag{2.300}$$

The latent heat of the transition is $\tilde{\ell} = 333\,\mathrm{J/g} = 79.5\,\mathrm{cal/g}$. Thus,

$$\begin{aligned}\bigg(\frac{dp}{dT}\bigg)_{\mathrm{liq-sol}} &= \frac{\tilde{\ell}}{T\,\Delta\tilde{v}} = \frac{333\,\mathrm{J/g}}{(273.1\,\mathrm{K})\,(-9.05\times10^{-2}\,\mathrm{cm}^3/\mathrm{g})} \\ &= -1.35\times10^{8}\,\frac{\mathrm{dyn}}{\mathrm{cm}^2\,\mathrm{K}} = -134\,\frac{\mathrm{atm}}{{}^\circ\mathrm{C}} \quad .\end{aligned} \tag{2.301}$$

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[19] 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 $\Delta T_{\mathrm{melt}} = -0.75^\circ\,\mathrm{C}$.

---

[18] Were the solid more dense, ponds and lakes would freeze from the bottom, with unfortunate consequences for life on earth.

[19] 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.

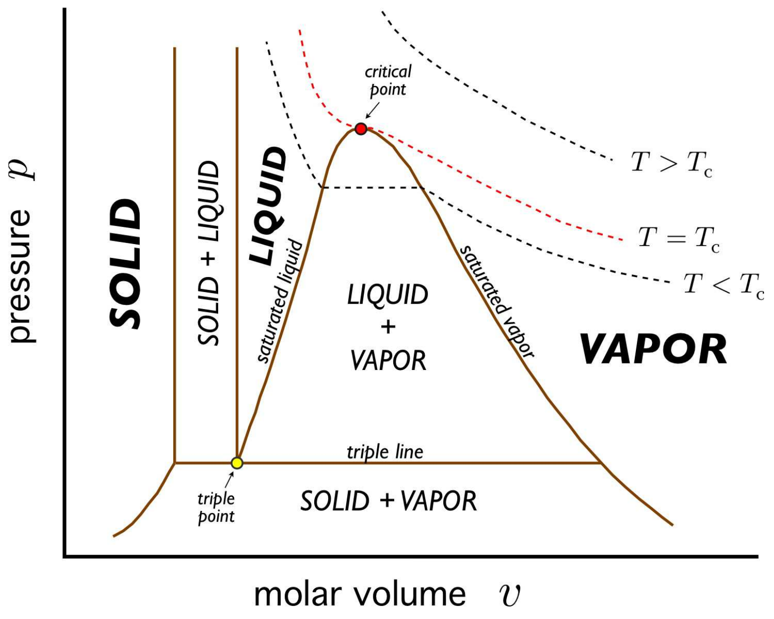


Figure 2.25: Projection of the p-v-T surface of fig. 2.22 onto the (v, p) plane.

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

$$\Delta p = \frac{Mg}{A} \approx \frac{20\,\mathrm{kg} \times 9.8\,\mathrm{m/s^2}}{(5 \times 10^{-3}\,\mathrm{m}) \times (10^{-1}\,\mathrm{m})} = 3.9 \times 10^5\,\mathrm{Pa} = 3.9\,\mathrm{atm} \quad . \tag{2.302}$$

The corresponding change in the melting temperature is thus minuscule: $\Delta T_{\mathrm{melt}} \approx -0.03^\circ\,\mathrm{C}$.

So why could my daughter skate so nicely? The answer isn't so clear![20] 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[21], theory predicts a surface melt thickness $d \propto (T_{\mathrm{m}} - T)^{-1/3}$. In fig. 2.26 we show measurements by Gilpin (1980) of the surface melt on ice, down to about -50° C. Near 0° C the melt layer thickness is about $40\,\mathrm{nm}$, but this decreases to $\sim 1\,\mathrm{nm}$ at $T = -35^\circ$ 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

---

[20] For a recent discussion, see R. Rosenberg, *Physics Today* **58**, 50 (2005).

[21] For example, they could be of the van der Waals form, due to virtual dipole fluctuations, with an attractive $1/r^6$ tail.

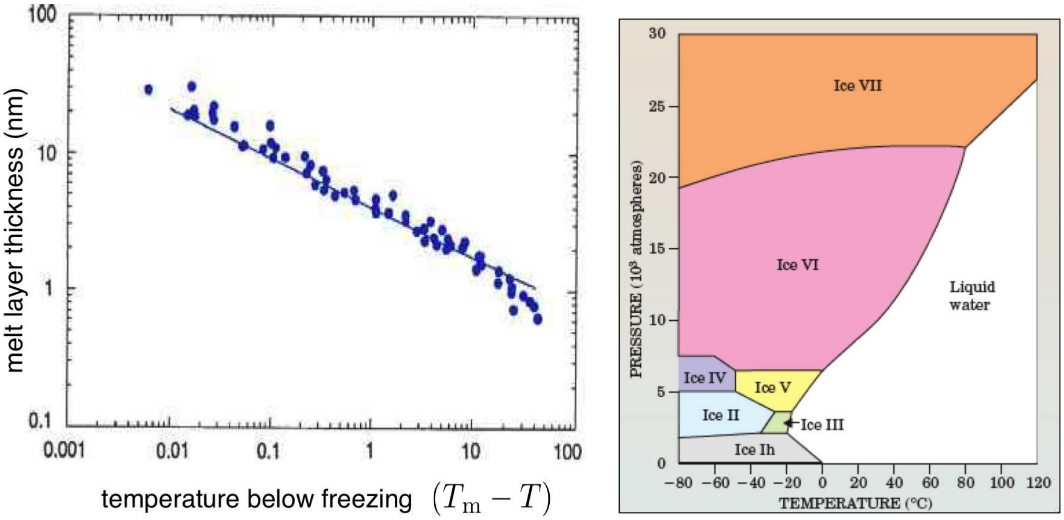


Figure 2.26: Left panel: data from R. R. Gilpin, *J. Colloid Interface Sci.* **77**, 435 (1980) showing measured thickness of the surface melt on ice at temperatures below $0°\mathrm{C}$. The straight line has slope $-\frac{1}{3}$, as predicted by theory. Right panel: phase diagram of H_2O, showing various high pressure solid phases. (*Source* : *Physics Today*, December 2005).

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.* $\Theta = 0°\,\mathrm{C}$) 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

$$Q = M\tilde{c}_{\mathrm{S}}(\Theta - T_0) + M\tilde{\ell} + M\tilde{c}_{\mathrm{L}}(T_1 - \Theta) \quad , \tag{2.303}$$

where $\tilde{c}_{\mathrm{S}}$ and $\tilde{c}_{\mathrm{L}}$ are the specific heats of ice (solid) and water (liquid), respectively[22], 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:

$$\Delta S_{\mathrm{pond}} = -\frac{Q}{T_1} \quad . \tag{2.304}$$

---

[22] We assume $\tilde{c}_{\mathrm{S}}(T)$ and $\tilde{c}_{\mathrm{L}}(T)$ have no appreciable temperature dependence, and we regard them both as constants.

Now we ask what is the entropy change of the H_2O in the ice. We have

$$\begin{aligned}\Delta S_{\rm ice} &= \int \frac{\dbar Q}{T} = \int\limits_{T_0}^{\varTheta} dT\,\frac{M\tilde c_{\rm S}}{T} + \frac{M\tilde\ell}{\varTheta} + \int\limits_{\varTheta}^{T_1} dT\,\frac{M\tilde c_{\rm L}}{T} \\ &= M\tilde c_{\rm S}\log\!\bigg(\frac{\varTheta}{T_0}\bigg) + \frac{M\tilde\ell}{\varTheta} + M\tilde c_{\rm L}\log\!\bigg(\frac{T_1}{\varTheta}\bigg) \quad .\end{aligned} \tag{2.305}$$

The total entropy change of the system is then

$$\begin{aligned}\Delta S_{\rm total} &= \Delta S_{\rm pond} + \Delta S_{\rm ice} \qquad (2.306)\\ &= M\tilde c_{\rm S}\log\!\bigg(\frac{\varTheta}{T_0}\bigg) - M\tilde c_{\rm S}\bigg(\frac{\varTheta - T_0}{T_1}\bigg) + M\tilde\ell\,\bigg(\frac{1}{\varTheta} - \frac{1}{T_1}\bigg) + M\tilde c_{\rm L}\log\!\bigg(\frac{T_1}{\varTheta}\bigg) - M\tilde c_{\rm L}\bigg(\frac{T_1 - \varTheta}{T_1}\bigg) \quad .\end{aligned}$$

Now since $T_0 < \varTheta < T_1$, we have

$$M\tilde c_{\rm S}\bigg(\frac{\varTheta - T_0}{T_1}\bigg) < M\tilde c_{\rm S}\bigg(\frac{\varTheta - T_0}{\varTheta}\bigg) \quad . \tag{2.307}$$

Therefore,

$$\Delta S > M\tilde\ell\,\bigg(\frac{1}{\varTheta} - \frac{1}{T_1}\bigg) + M\tilde c_{\rm S}\,f\big(T_0/\varTheta\big) + M\tilde c_{\rm L}\,f\big(\varTheta/T_1\big) \quad , \tag{2.308}$$

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 $T_0 < \varTheta < T_1$, 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 $\varTheta$. When $T_0 = \varTheta^-$ and $T_1 = \varTheta^+$ are infinitesimally below and above the melting temperature, the entropy change in eqn. 2.307 is a positive infinitesimal, and when $T_0 = \varTheta = T_1$ 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

$$\mu_1(p,T) = \mu_2(p,T) \quad , \tag{2.309}$$

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

$$\mu_1(p,T) = \mu_2(p,T) = \mu_3(p,T) \quad , \tag{2.310}$$

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:

$$\text{critical point} \quad \Longrightarrow \quad v_1(p,T) = v_2(p,T) \quad , \quad \mu_1(p,T) = \mu_2(p,T) \quad . \tag{2.311}$$

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 $\textsf{c}$ species, with particle numbers N_a, where $a = 1, \ldots, \textsf{c}$. It is useful to briefly recapitulate the derivation of the Gibbs-Duhem relation. The internal energy $E(S, V, N_1, \ldots, N_{\textsf{c}})$ is a homogeneous function of degree one:

$$E(\lambda S, \lambda V, \lambda N_1, \ldots, \lambda N_{\textsf{c}}) = \lambda E(S, V, N_1, \ldots, N_{\textsf{c}}) \quad . \tag{2.312}$$

From Euler's theorem for homogeneous functions (just differentiate with respect to λ and then set $\lambda = 1$), we have

$$E = TS - p\,V + \sum_{a=1}^{\textsf{c}} \mu_a\, N_a \quad . \tag{2.313}$$

Taking the differential, and invoking the First Law,

$$dE = T\,dS - p\,dV + \sum_{a=1}^{\textsf{c}} \mu_a\, dN_a \quad , \tag{2.314}$$

we arrive at the relation

$$S\,dT - V dp + \sum_{a=1}^{\textsf{c}} N_a\, d\mu_a = 0 \quad , \tag{2.315}$$

of which eqn. 2.111 is a generalization to additional internal 'work' variables. This says that the $\textsf{c} + 2$ quantities $(T, p, \mu_1, \ldots, \mu_{\textsf{c}})$ are not all independent. We can therefore write

$$\mu_{\textsf{c}} = \mu_{\textsf{c}}\big(T, p, \mu_1, \ldots, \mu_{\textsf{c}-1}\big) \quad . \tag{2.316}$$

If there are φ 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

$$\mu_{\textsf{c}}^{(j)} = \mu_{\textsf{c}}^{(j)}\Big(T, p, \mu_1^{(j)}, \ldots, \mu_{\textsf{c}-1}^{(j)}\Big) \quad . \tag{2.317}$$

Here $\mu_a^{(j)}$ is the chemical potential of the $a^{\rm th}$ species in the $j^{\rm th}$ phase. Thus, there are φ such equations relating the $2 + \textsf{c}\,\varphi$ variables $\big(T, p, \{\mu_a^{(j)}\}\big)$, meaning that only $2 + (\textsf{c} - 1)\varphi$ of them may be chosen as independent. This, then, is the dimension of 'thermodynamic space' containing a maximal number of intensive variables:

$$d(\textsf{c}, \varphi) = 2 + \varphi\,(\textsf{c} - 1) \quad . \tag{2.318}$$

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 $N = \sum_{a=1}^{\textsf{c}} N_a$ may not be conserved in the presence of chemical reactions!

Now suppose we have equilibrium among φ 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

$$\mu_a^{(j)} = \mu_a^{(j')} \tag{2.319}$$

where $j, j' \in \{1, \ldots, \varphi\}$. This gives $\mathsf{c}(\varphi - 1)$ independent equations equations[23]. Thus, we can have *phase equilibrium* among the φ phases of c species over a region of dimension

$$d_{\text{PE}}(\mathsf{c}, \varphi) = 2 + \varphi\,(\mathsf{c} - 1) - \mathsf{c}\,(\varphi - 1) = 2 + \mathsf{c} - \varphi \quad . \tag{2.320}$$

Since $d_{\text{PE}} \geq 0$, 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 r distinct chemical reactions of the form

$$\zeta_1^{(l)}\,\text{A}_1 + \zeta_2^{(l)}\,\text{A}_2 + \cdots + \zeta_{\mathsf{c}}^{(l)}\,\text{A}_{\mathsf{c}} = 0 \quad , \tag{2.321}$$

where A_a is the chemical formula for species a, and $\zeta_a^{(l)}$ is the stoichiometric coefficient for the a^{th} species in the l^{th} reaction, with $l \in \{1, \ldots, \mathsf{r}\}$, then we have an additional r constraints of the form

$$\sum_{a=1}^{\mathsf{c}} \zeta_a^{(l)}\,\mu_a^{(j)} = 0 \quad . \tag{2.322}$$

Therefore,

$$d_{\text{PE}}(\mathsf{c}, \varphi, \mathsf{r}) = 2 + \mathsf{c} - \varphi - \mathsf{r} \quad . \tag{2.323}$$

One might ask what value of j are we to use in eqn. 2.322, or do we in fact have φ such equations for each l? The answer is that eqn. 2.319 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. 2.322.

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_{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 d independent generalized displacements, the dimensions of the thermodynamic space and regions of phase equilibria are

$$d(\mathsf{c}, \varphi) = 1 + \mathsf{d} + \varphi\,(\mathsf{c} - 1) \qquad , \qquad d_{\text{PE}}(\mathsf{c}, \varphi, \mathsf{r}) = 1 + \mathsf{d} + \mathsf{c} - \varphi - \mathsf{r} \quad , \tag{2.324}$$

whence $\varphi \leq 1 + \mathsf{d} + \mathsf{c} - \mathsf{r}$.

Why do phase diagrams in the (p, v) and (T, v) plane look different than those in the (p, T) plane?[24] 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 $v = N_{\text{A}} V/N$ or chemical concentration $x_a = N_a/N$.

---

[23] Set $j = 1$ and let j' range over the $\varphi - 1$ values $2, \ldots, \varphi$.

[24] 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, μ) plane at constant p.

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 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

$$S(T,V,N) = Nk_{\rm B}\log(V/N) + N\phi(T) \quad , \tag{2.325}$$

since $(\partial S/\partial V)_{T,N} = (\partial p/\partial T)_{V,N} = Nk_{\rm B}/V$. Note that in eqn. 2.325 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 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

$$S_{\rm unmixed} = \sum_a S_a = \sum_a \Big[N_a\, k_{\rm B}\log(V_a/N_a) + N_a\,\phi_a(T)\Big] \quad . \tag{2.326}$$

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

$$S_{\rm mixed} = \sum_a S_a = \sum_a \Big[N_a\, k_{\rm B}\log(V/N_a) + N_a\,\phi_a(T)\Big] \quad . \tag{2.327}$$

Thus, the *entropy of mixing* is

$$\begin{aligned}\Delta S_{\rm mix} &= S_{\rm mixed} - S_{\rm unmixed}\\ &= \sum_a N_a\, k_{\rm B}\, \log(V/V_a) = -Nk_{\rm B}\sum_a f_a \log f_a \quad ,\end{aligned} \tag{2.328}$$

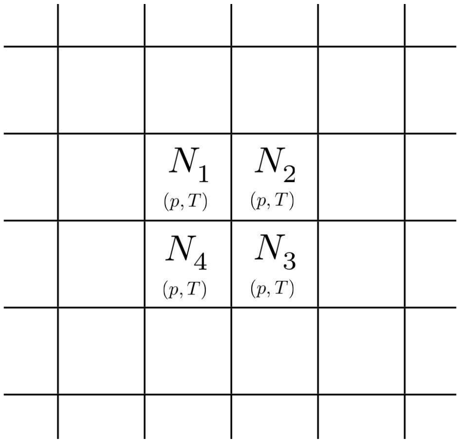


Figure 2.27: A multicomponent system consisting of isolated gases, each at temperature T and pressure p. Then system entropy increases when all the walls between the different subsystems are removed.

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

$$S(E,V,N) = k_{\rm B} \log W(E,V,N) \quad . \tag{2.329}$$

Consider a system consisting of 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 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

$$W_a = \binom{B_a}{N_a} = \frac{B_a!}{N_a!\,(B_a - N_a)!} \quad . \tag{2.330}$$

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,

$$\log(K!) = K \log K - K + \tfrac{1}{2} \log K + \tfrac{1}{2} \log(2\pi) + \mathcal{O}(K^{-1}) \quad , \tag{2.331}$$

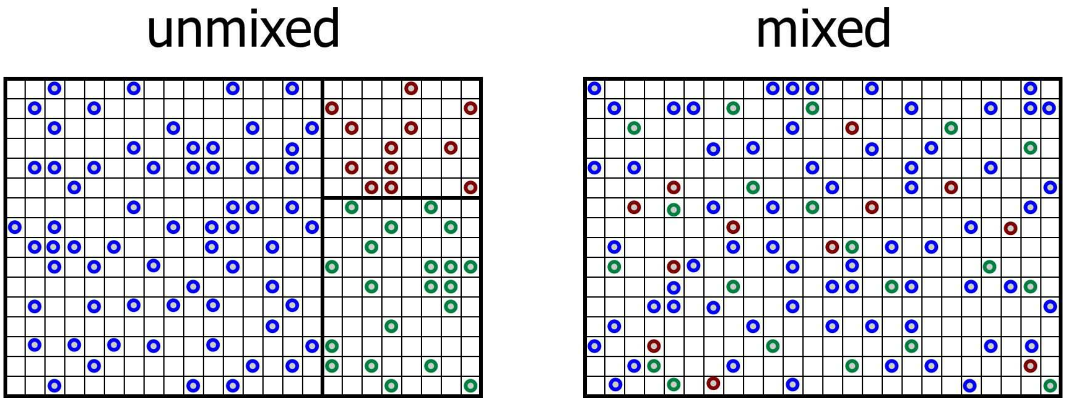


Figure 2.28: Mixing among three different species of particles. The mixed configuration has an additional entropy, ΔS_{mix}.

which is an asymptotic expansion valid for $K \gg 1$. One then finds for $B, N \gg 1$, with $x = N/B \in [0,1]$,

$$\begin{aligned}\log\binom{B}{N} &= \Big(B\log B - B\Big) - \Big(xB\log(xB) - xB\Big) - \Big((1-x)B\log\big((1-x)B\big) - (1-x)B\Big) \\ &= -B\Big[x\log x + (1-x)\log(1-x)\Big] \quad . \end{aligned} \tag{2.332}$$

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

$$S_{\text{unmixed}} = k_{\text{B}}\sum_{a=1}^{\mathsf{c}}\log W_a = -k_{\text{B}}\sum_{a=1}^{\mathsf{c}} B_a\Big[x_a\log x_a + (1-x_a)\log(1-x_a)\Big] \quad , \tag{2.333}$$

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_{c} particles of species c is

$$W_{\text{mixed}} = \frac{B!}{N_0!\,N_1!\cdots N_{\mathsf{c}}!} \quad , \tag{2.334}$$

where $N_0 = B - \sum_{a=1}^{\mathsf{c}} N_a$ is the number of vacant boxes. Again using Stirling's rule, we find

$$S_{\text{mixed}} = -k_{\text{B}}B\sum_{a=0}^{\mathsf{c}} y_a\log y_a \quad , \tag{2.335}$$

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

$$y_a = \frac{N_a}{B} = \frac{N_a}{B_a}\cdot\frac{B_a}{B} = x_a\,f_a \quad , \tag{2.336}$$

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

$$S_{\text{mixed}} = -k_{\text{B}}B\sum_{a=1}^{\mathsf{c}} \Big[x_a\, f_a \log(f_a) + f_a\, x_a \log x_a\Big] - k_{\text{B}}B\,\overbrace{\Big(1-\sum_{a=1}^{\mathsf{c}} x_a\, f_a\Big)}^{y_0} \log\Big(1-\sum_{a=1}^{\mathsf{c}} x_a f_a\Big) \tag{2.337}$$

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 c initial regions, then

$$\begin{aligned} S_{\text{unmixed}} &= k_{\text{B}}B\Big[x\log x + (1-x)\log(1-x)\Big] \\ S_{\text{mixed}} &= -k_{\text{B}}Bx\sum_{a=1}^{\mathsf{c}} f_a \log f_a - k_{\text{B}}B\Big[x\log x + (1-x)\log(1-x\Big] \end{aligned} \tag{2.338}$$

and thus, after identifying $N = xB$ is the total number of particles (*i.e.* occupied little boxes), the entropy of mixing is

$$\Delta S_{\text{mix}} = -Nk_{\text{B}}\sum_{a=1}^{\mathsf{c}} f_a \log f_a \quad , \tag{2.339}$$

where $N = \sum_{a=1}^{\mathsf{c}} N_a$ is the total number of particles among all species (excluding vacancies) and

$$f_a = \frac{N_a}{B} = \frac{N_a}{N_0 + N} \tag{2.340}$$

is the fraction of all boxes occupied by 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

$$\Delta S_{\text{mix}} = -k_{\text{B}}\Bigg[N_0\log\bigg(\frac{N_0}{N_0+\widetilde{N}}\bigg) + \sum_{a=1}^{\tilde{\mathsf{c}}} N_a \log\bigg(\frac{N_a}{N_0+\widetilde{N}}\bigg)\Bigg] \quad , \tag{2.341}$$

where $\widetilde{N} = \sum_{a=1}^{\tilde{\mathsf{c}}} N_a$ is the total number of solute molecules, summed over all solvent species. The total number of solvent 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

$$\Delta S_{\text{mix}} = -k_{\text{B}}\sum_{a=1}^{\tilde{\mathsf{c}}} \Bigg[N_a\log\bigg(\frac{N_a}{N_0}\bigg) - N_a\Bigg] + \sum_{a,b=1}^{\tilde{\mathsf{c}}} \frac{N_a N_b}{N_0} + \mathcal{O}\bigg(\frac{\widetilde{N}^3}{N_0^2}\bigg) \quad . \tag{2.342}$$

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,

$$\begin{aligned} G\big(T,p,N_0,\{N_a\}\big) = N_0\, g_0(T,p) + k_{\rm B}T\sum_{a=1}^{\tilde{\mathsf{c}}} N_a \log\bigg(\frac{N_a}{eN_0}\bigg) \\ + \sum_{a=1}^{\tilde{\mathsf{c}}} N_a\,\omega_a(T,p) + \frac{1}{2N_0}\sum_{a,b=1}^{\tilde{\mathsf{c}}} A_{ab}(T,p)\,N_a\,N_b \quad . \end{aligned} \tag{2.343}$$

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. 2.342. 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

$$\mu_0(T,p) = \frac{\partial G}{\partial N_0} = g_0(T,p) - k_{\rm B}T\sum_{a=1}^{\tilde{\mathsf{c}}} x_a - \tfrac{1}{2}\sum_{a,b=1}^{\tilde{\mathsf{c}}} A_{ab}(T,p)\,x_a\,x_b \quad , \tag{2.344}$$

and the chemical potential for species a is

$$\mu_a(T,p) = \frac{\partial G}{\partial N_a} = k_{\rm B}T\log x_a + \omega_a(T,p) + \sum_{b=1}^{\tilde{\mathsf{c}}} A_{ab}(T,p)\,x_b \quad , \tag{2.345}$$

where $x_a = N_a/N_0$ is the concentrations 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 $\widetilde{N} = \sum_{a=1}^{\tilde{\mathsf{c}}} N_a$ is the total number of solute molecules. To lowest order, then, we have

$$\mu_0(T,p) = g_0(T,p) - x\,k_{\rm B}T \tag{2.346}$$

$$\mu_a(T,p) = k_{\rm B}T\log x_a + \omega_a(T,p) \quad , \tag{2.347}$$

where $x = \sum_{a=1}^{\tilde{\mathsf{c}}} x_a$ is the total solute concentration.

If we add sugar to a solution confined by a semipermeable membrane[25], 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 μ_0 is the same on each side. This means

$$g_0(T,p_{\rm R}) - xk_{\rm B}T = g_0(T,p_{\rm L}) \quad , \tag{2.348}$$

where $p_{\rm L,R}$ is the pressure on the left and right sides of the membrane, and $x = N/N_0$ is again the total solute concentration. This equation once again tells us that the pressure p cannot be the same on

---

[25] 'Semipermeable' in this context means permeable to the solvent but not the solute(s).

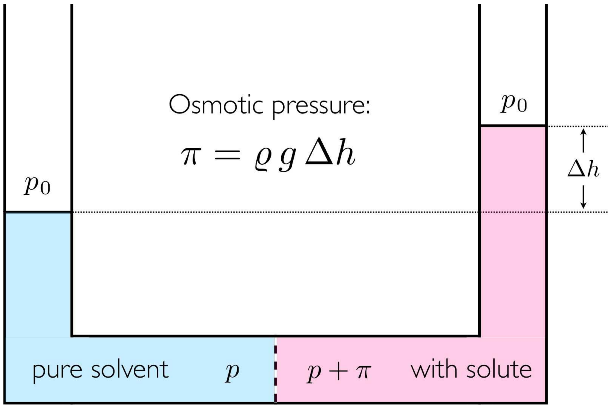


Figure 2.29: Osmotic pressure causes the column on the right side of the U-tube to rise higher than the column on the left by an amount $\Delta h = \pi/\varrho\, g$.

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

$$\pi = x\, k_{\rm B} T \bigg/ \left(\frac{\partial \mu_0}{\partial p}\right)_T \quad . \tag{2.349}$$

But a Maxwell relation (§2.8) guarantees

$$\left(\frac{\partial \mu}{\partial p}\right)_{T,N} = \left(\frac{\partial V}{\partial N}\right)_{T,p} = v(T,p)/N_{\rm A} \quad , \tag{2.350}$$

where $v(T,p)$ is the molar volume of the solvent.

$$\pi v = xRT \quad , \tag{2.351}$$

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 ν moles of sucrose ($C_{12}H_{22}O_{11}$) per kilogram ($55.52\,{\rm mol}$) of water at $30^\circ\,{\rm C}$. We find $\pi = 2.5\,{\rm atm}$ when $\nu = 0.1$.

One might worry about the expansion in powers of π when π 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 κ_T is the isothermal compressibility. For water one has $\kappa_T \approx 4.4 \times 10^{-5}\,({\rm atm})^{-1}$, 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:

$$\mu^0_{\rm L}(T,p) = \mu^0_{\rm V}(T,p) \quad . \tag{2.352}$$

Here we write μ^0 for μ 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

Substance	Latent Heat of Fusion $\tilde{\ell}_f$ J/g	Melting Point °C	Latent Heat of Vaporization $\tilde{\ell}_v$ J/g	Boiling Point °C
C_2H_5OH	108	-114	855	78.3
NH_3	339	-75	1369	-33.34
CO_2	184	-57	574	-78
He	–	–	21	-268.93
H	58	-259	455	-253
Pb	24.5	372.3	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.

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

$$\mu_L(T,p,x) = \mu_L^0(T,p) - x k_B T \quad , \tag{2.353}$$

where x is the dimensionless solute concentration, summed over all species. The condition for liquid-vapor coexistence now becomes

$$\mu_L^0(T,p) - x k_B T = \mu_V^0(T,p) \quad . \tag{2.354}$$

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

$$\mu_L^0(T_0^*,p) + \bigg(\frac{\partial \mu_L^0}{\partial T}\bigg)_{\!p} \big(T - T_0^*\big) - x k_B T = \mu_V^0(T_0^*,p) + \bigg(\frac{\partial \mu_V^0}{\partial T}\bigg)_{\!p} \big(T - T_0^*\big) \quad . \tag{2.355}$$

Note that

$$\bigg(\frac{\partial \mu}{\partial T}\bigg)_{\!p,N} = -\bigg(\frac{\partial S}{\partial N}\bigg)_{\!T,p} \tag{2.356}$$

from a Maxwell relation deriving from exactness of dG. Since S is extensive, we can write $S = (N/N_A)\, s(T,p)$, where $s(T,p)$ is the molar entropy. Solving for T, we obtain

$$T^*(p,x) = T_0^*(p) + \frac{xR\big[T_0^*(p)\big]^2}{\ell_v(p)} \quad , \tag{2.357}$$

where $\ell_v = T_0^* \cdot (s_V - s_L)$ 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*.

As an example, consider seawater, which contains approximately $35\,$g of dissolved 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

$$x = \frac{2 \cdot 35}{23.0 + 35.4} \bigg/ \frac{1000}{18} \approx 0.022 \quad . \tag{2.358}$$

The latent heat of vaporization of H_2O at atmospheric pressure is $\ell = 40.7\,\mathrm{kJ/mol}$, hence

$$\Delta T^* = \frac{(0.022)(8.3\,\mathrm{J/mol\,K})(373\,\mathrm{K})^2}{4.1 \times 10^4\,\mathrm{J/mol}} \approx 0.6\,\mathrm{K} \quad . \tag{2.359}$$

Put another way, the boiling point elevation of H_2O at atmospheric pressure is about $0.28°\mathrm{C}$ 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 $K = 0.51°\mathrm{C\,kg/mol}$.

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 $\ell_\mathrm{f} = T_\mathrm{f}^0 \cdot (s_\mathrm{LIQUID} - s_\mathrm{SOLID}) = 6.01\,\mathrm{kJ/mol}$[26] We thus predict a *freezing point depression* of $\Delta T^* = -xR\big[T_0^*\big]^2/\ell_\mathrm{f} = 1.03°\mathrm{C} \cdot x[\%]$. This can be expressed once again as $\Delta T^* = -Km$, with $K = 1.86°\mathrm{C\,kg/mol}$[27].

2.12.5 Binary solutions

Consider a binary solution, and write the Gibbs free energy $G(T, p, N_\mathsf{A}, N_\mathsf{B})$ as

$$\begin{aligned} G(T, p, N_\mathsf{A}, N_\mathsf{B}) = N_\mathsf{A}\,\mu_\mathsf{A}^0(T, p) + N_\mathsf{B}\,\mu_\mathsf{B}^0(T, p) + N_\mathsf{A} k_\mathrm{B} T \, \log\left(\frac{N_\mathsf{A}}{N_\mathsf{A} + N_\mathsf{B}}\right) \\ + N_\mathsf{B}\, k_\mathrm{B} T \, \log\left(\frac{N_\mathsf{B}}{N_\mathsf{A} + N_\mathsf{B}}\right) + \lambda\, \frac{N_\mathsf{A} N_\mathsf{B}}{N_\mathsf{A} + N_\mathsf{B}} \quad . \end{aligned} \tag{2.360}$$

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, $g(T, p, x) = G/(N_\mathsf{A} + N_\mathsf{B})$, in terms of T, p, and the concentration $x \equiv x_\mathsf{B} = N_\mathsf{B}/(N_\mathsf{A} + N_\mathsf{B})$ of species B (hence $x_\mathsf{A} = 1 - x$ is the concentration of species A). Then

$$g(T, p, x) = (1 - x)\,\mu_\mathsf{A}^0 + x\,\mu_\mathsf{B}^0 + k_\mathrm{B} T \Big[x \log x + (1 - x) \log(1 - x)\Big] + \lambda\, x\,(1 - x) \quad . \tag{2.361}$$

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

$$\frac{\partial g}{\partial x} = \mu_\mathsf{B}^0 - \mu_\mathsf{A}^0 + k_\mathrm{B} T \, \log\left(\frac{x}{1 - x}\right) + \lambda\,(1 - 2x) \tag{2.362}$$

---

[26] See table 2.4, and recall $M = 18\,\mathrm{g}$ is the molar mass of H_2O.

[27] 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 ΔT^* is positive.

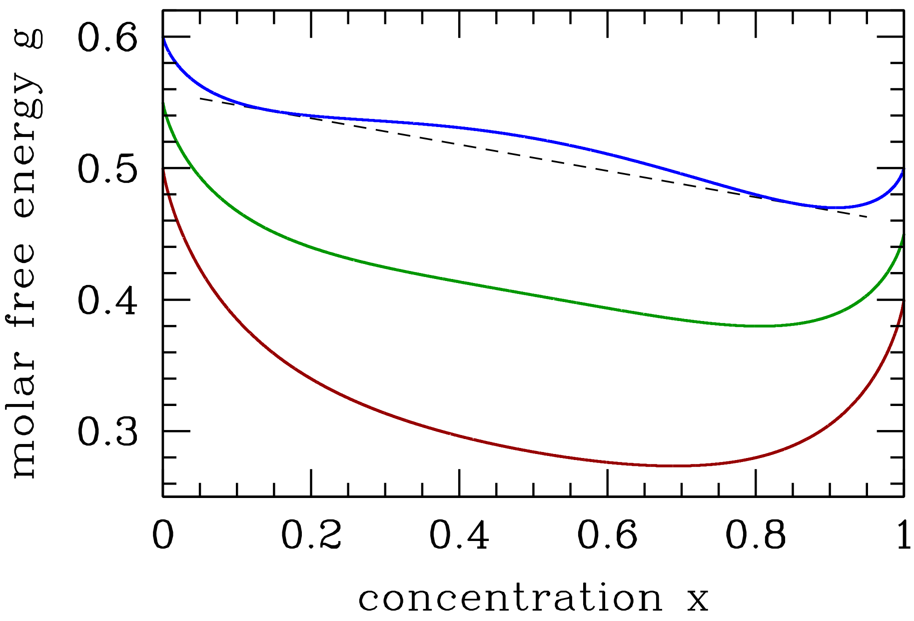


Figure 2.30: Gibbs free energy per particle for a binary solution as a function of concentration $x = x_{\mathsf{B}}$ of the B species (pure A at the left end $x = 0$; pure B at the right end $x = 1$), in units of the interaction parameter λ. Dark red curve: $T = 0.65\,\lambda/k_{\rm B} > T_{\rm c}$; 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 $\mu^0_{\mathsf{A}} = 0.60\,\lambda - 0.50\,k_{\rm B}T$ and $\mu^0_{\mathsf{B}} = 0.50\,\lambda - 0.50\,k_{\rm B}T$. 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.

and

$$\frac{\partial^2 g}{\partial x^2} = \frac{k_{\rm B}T}{x} + \frac{k_{\rm B}T}{1-x} - 2\lambda \quad . \tag{2.363}$$

The spinodal is given by the solution to the equation $\partial^2 g/\partial x^2 = 0$, which is

$$T^*(x) = \frac{2\lambda}{k_{\rm B}}\, x\,(1-x) \quad . \tag{2.364}$$

Since $x\,(1-x)$ achieves its maximum value of $\frac{1}{4}$ at $x = \frac{1}{2}$, we have $T^* \le k_{\rm B}/2\lambda$.

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 λ. When $T < \lambda/2k_{\rm B}$, there 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

$$\frac{g(x_2) - g(x_1)}{x_2 - x_1} = \frac{\partial g}{\partial x}\bigg|_{x_1} = \frac{\partial g}{\partial x}\bigg|_{x_2} \quad . \tag{2.365}$$

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

$$g(1-x) - g(x) = (1-2x)\left(\mu^0_{\mathsf{B}} - \mu^0_{\mathsf{A}}\right) \quad , \tag{2.366}$$

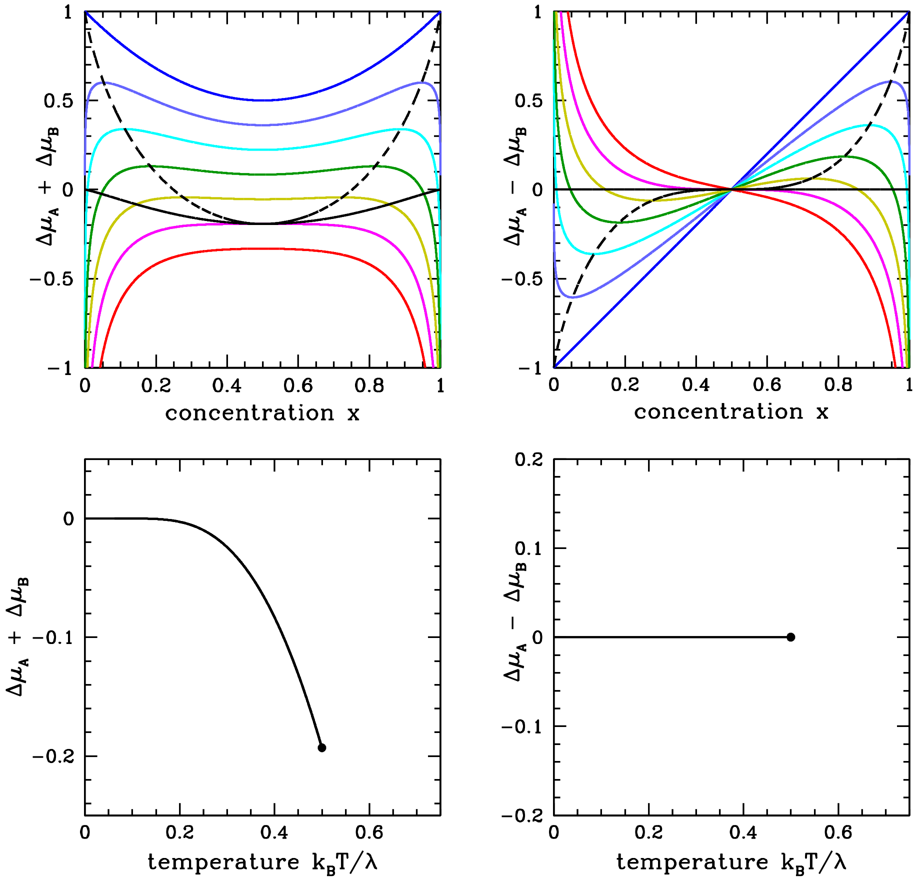


Figure 2.31: Upper panels: chemical potential shifts $\Delta\mu_{\pm} = \Delta\mu_{\text{A}} \pm \Delta\mu_{\text{B}}$ *versus* concentration $x = x_{\text{B}}$. The dashed black line is the spinodal, and the solid black line the coexistence boundary. Temperatures range from $T = 0$ (dark blue) to $T = 0.6\,\lambda/k_{\text{B}}$ (red) in units of $0.1\,\lambda/k_{\text{B}}$. Lower panels: phase diagram in the $(T, \Delta\mu_{\pm})$ planes. The black dot is the critical point.

and invoking eqns. 2.365 and 2.362 we obtain the solution

$$T_{\text{coex}}(x) = \frac{\lambda}{k_{\text{B}}} \cdot \frac{1-2x}{\log\!\big(\frac{1-x}{x}\big)} \quad . \tag{2.367}$$

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 $T^*(x) < T < T_{\text{coex}}(x)$, 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 λ itself is temperature-dependent, there can be multiple solutions to eqns. 2.364 and 2.367. For example, one could take

$$\lambda(T) = \frac{\lambda_0\, T^2}{T^2 + T_0^2} \quad . \tag{2.368}$$

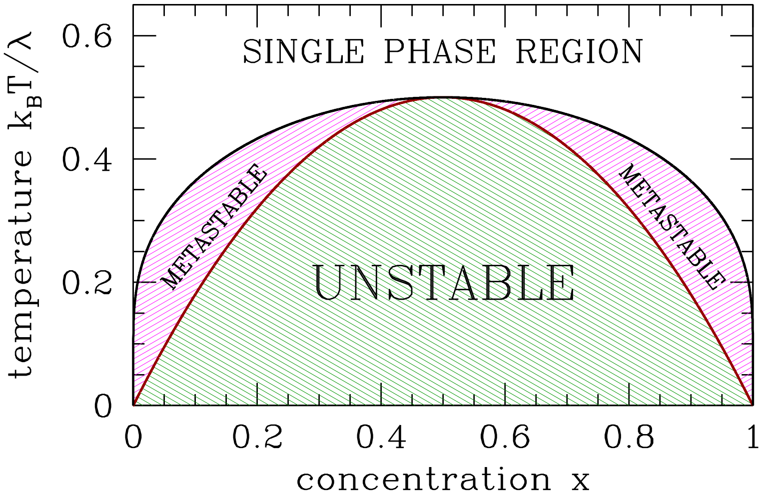


Figure 2.32: Phase diagram for the binary system. The black curve is the coexistence curve, and the dark red curve is the spinodal. A-rich material is to the left and B-rich to the right.

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.

It is instructive to consider the phase diagram in the (T,μ) plane. We define the chemical potential shifts,

$$\Delta\mu_{\sf A} \equiv \mu_{\sf A} - \mu^0_{\sf A} = k_{\rm B}T\log(1-x) + \lambda\, x^2 \tag{2.369}$$

$$\Delta\mu_{\sf B} \equiv \mu_{\sf B} - \mu^0_{\sf B} = k_{\rm B}T\log x + \lambda\,(1-x)^2 \quad , \tag{2.370}$$

and their sum and difference, $\Delta\mu_\pm \equiv \Delta\mu_{\sf A} \pm \Delta\mu_{\sf B}$. From the Gibbs-Duhem relation, we know that we can write $\mu_{\sf B}$ as a function of T, p, and $\mu_{\sf A}$. 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 $x = x_{\sf B}$ changes discontinuously across the phase boundary. The critical point lies at $(T,\Delta\mu_-) = (\lambda/2k_{\rm B}\,,\,0)$.

If we choose $N = N_{\sf A} + N_{\sf B}$ to be the extensive variable, then fixing N means $dN_{\sf A} + dN_{\sf B} = 0$. So st fixed T and p,

$$dG\big|_{T,p} = \mu_{\sf A}\,dN_{\sf A} + \mu_{\sf B}\,dN_{\sf B} \qquad \Rightarrow \qquad dg\big|_{T,p} = -\Delta\mu_-\,dx \quad . \tag{2.371}$$

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,Δ_-) 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.

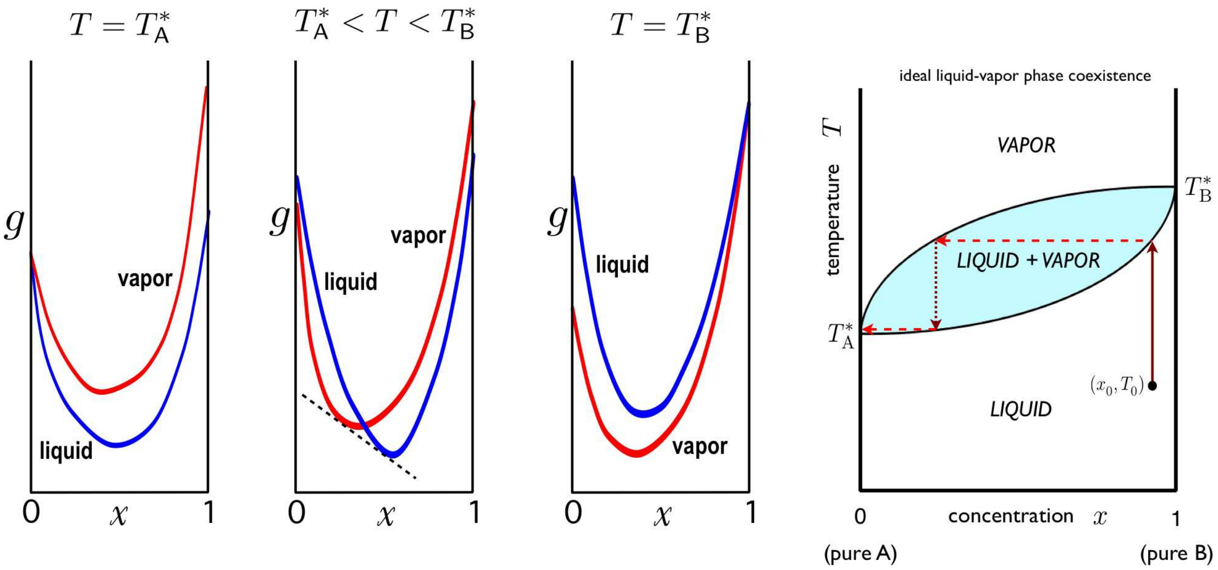


Figure 2.33: Gibbs free energy per particle g for an ideal binary solution for temperatures $T \in [T_A^*, T_B^*]$. The Maxwell construction is shown for the case $T_A^* < T < T_B^*$. Right: phase diagram, showing two-phase region and distillation sequence in (x, T) space.

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 ($x = 1$, $k_B T/\lambda = 0.4$) 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*[28]. 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 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[29]. 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

---

[28] An emulsion is a mixture of two or more immiscible liquids.

[29] We assume the boiling temperatures are not exactly equal.

the mixture? We begin by writing the free energies of the mixed liquid and mixed vapor phases as

$$g_{\rm L}(T,p,x) = (1-x)\,\mu^{\rm L}_{\sf A}(T,p) + x\,\mu^{\rm L}_{\sf B}(T,p) + k_{\rm B}T\Big[x\log x + (1-x)\log(1-x)\Big] + \lambda_{\rm L}\,x(1-x) \tag{2.372}$$

$$g_{\rm V}(T,p,x) = (1-x)\,\mu^{\rm V}_{\sf A}(T,p) + x\,\mu^{\rm V}_{\sf B}(T,p) + k_{\rm B}T\Big[x\log x + (1-x)\log(1-x)\Big] + \lambda_{\rm V}\,x(1-x) \quad . \tag{2.373}$$

Typically $\lambda_{\rm V} \approx 0$. Consider these two free energies as functions of the concentration x, at fixed T and p. If the curves never cross, and $g_{\rm L}(x) < g_{\rm V}(x)$ 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 $g_{\rm V}(x) < g_{\rm L}(x)$ 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 $\lambda_{\rm L} = \lambda_{\rm V} = 0$, and setting $g_{\rm L} = g_{\rm V}$ requires

$$(1-x)\,\Delta\mu_{\sf A}(T,p) + x\,\Delta\mu_{\sf B}(T,p) = 0 \quad , \tag{2.374}$$

where $\Delta\mu_{\sf A/B}(T,p) = \mu^{\rm L}_{\sf A/B}(T,p) - \mu^{\rm V}_{\sf A/B}(T,p)$. Expanding the chemical potential about a given temperature T^*,

$$\mu(T,p) = \mu(T^*,p) - s(T^*,p)\,(T-T^*) - \frac{c_p(T^*,p)}{2T}\,(T-T^*)^2 + \ldots \quad , \tag{2.375}$$

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 $(\partial s/\partial T)_{\mu,N} = c_p/T$. Thus, expanding $\Delta\mu_{\sf A/B}$ about $T^*_{\sf A/B}$, we have

$$\begin{aligned} \Delta\mu_{\sf A} &\equiv \mu^{\rm L}_{\sf A} - \mu^{\rm V}_{\sf A} = (s^{\rm V}_{\sf A} - s^{\rm L}_{\sf A})(T-T^*_{\sf A}) + \frac{c^{\rm V}_{p,{\sf A}} - c^{\rm L}_{p,{\sf A}}}{2T^*_{\sf A}}\,(T-T^*_{\sf A})^2 + \ldots \\ \Delta\mu_{\sf B} &\equiv \mu^{\rm L}_{\sf B} - \mu^{\rm V}_{\sf B} = (s^{\rm V}_{\sf B} - s^{\rm L}_{\sf B})(T-T^*_{\sf B}) + \frac{c^{\rm V}_{p,{\sf B}} - c^{\rm L}_{p,{\sf B}}}{2T^*_{\sf B}}\,(T-T^*_{\sf B})^2 + \ldots \end{aligned} \tag{2.376}$$

We assume $s^{\rm V}_{\sf A/B} > s^{\rm L}_{\sf A/B}$, *i.e.* the vapor phase has greater entropy per particle. Thus, $\Delta\mu_{\sf A/B}(T)$ changes sign from negative to positive as T rises through $T^*_{\sf A/B}$. If we assume that these are the *only* sign changes for $\Delta\mu_{\sf A/B}(T)$ at fixed p, then eqn. 2.374 can only be solved for $T \in [T^*_{\sf A}, T^*_{\sf B}]$. 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 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

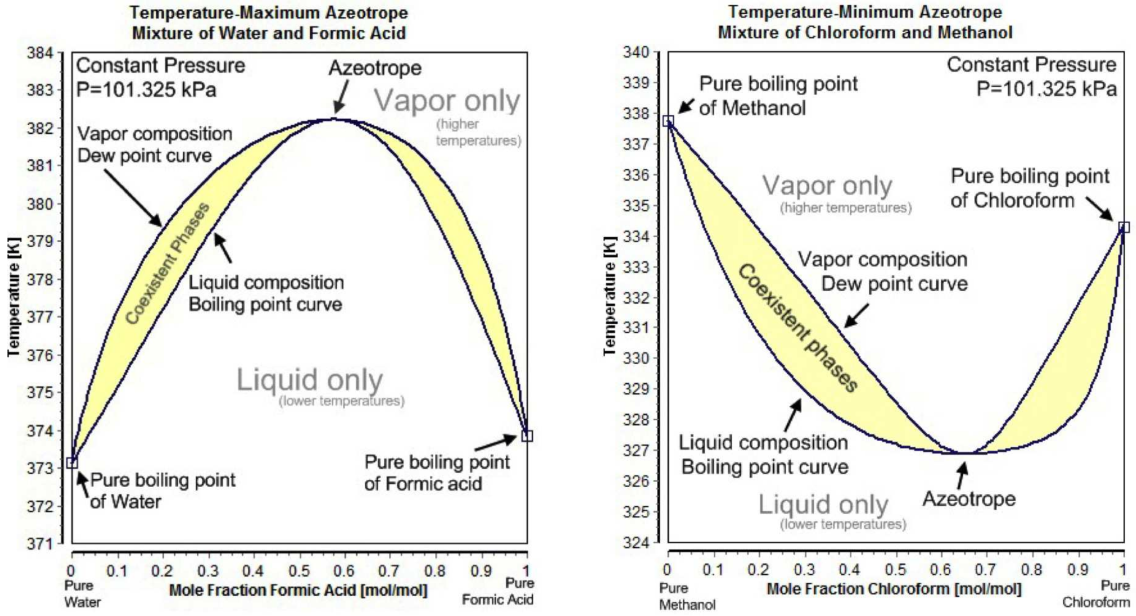


Figure 2.34: Negative (left) and positive (right) azeotrope phase diagrams. From *Wikipedia*.

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*.

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 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 $T^*_{\mathsf{A,B}}$. 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 $T^*_{\mathsf{A,B}}$, 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

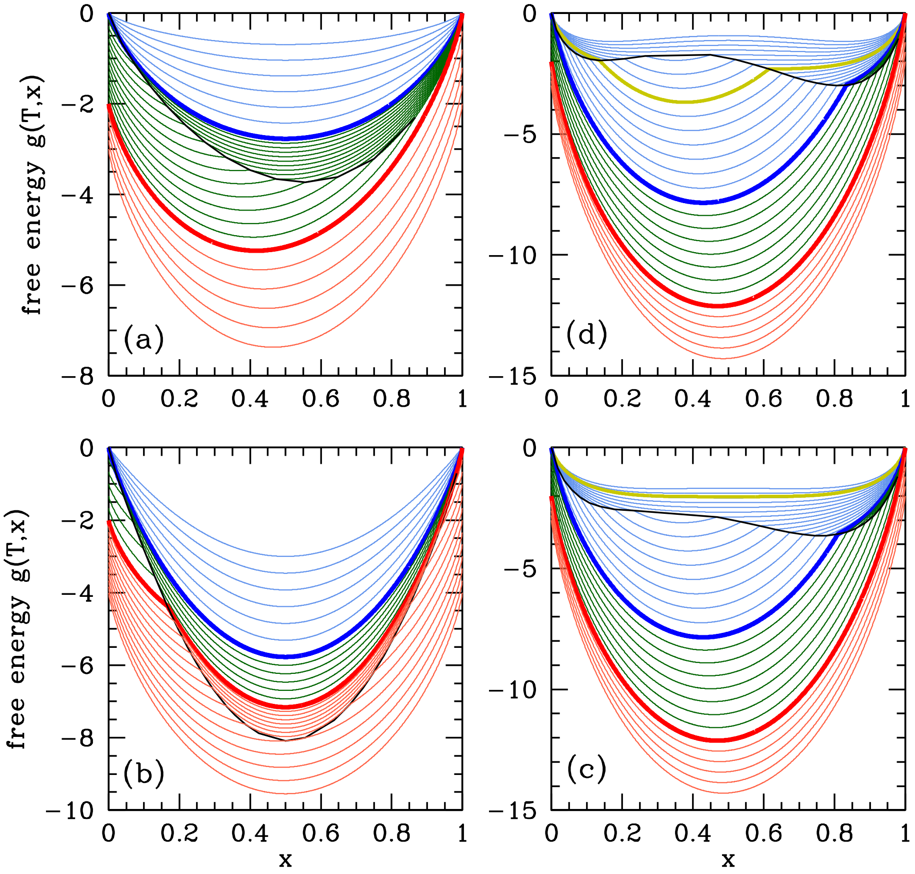


Figure 2.35: Free energies *before Maxwell constructions* for a binary fluid mixture in equilibrium with a vapor ($\lambda_V = 0$). Panels show (a) $\lambda_L = 0$ (ideal fluid), (b) $\lambda_L < 0$ (miscible fluid; negative azeotrope), (c) $\lambda^L_{AB} > 0$ (positive azeotrope), (d) $\lambda^L_{AB} > 0$ (heteroazeotrope). Thick blue and red lines correspond to temperatures T^*_A and T^*_B, respectively, with $T^*_A < T^*_B$. Thin blue and red curves are for temperatures outside the range $[T^*_A, T^*_B]$. The black curves show the locus of points where g is discontinuous, *i.e.* where the liquid and vapor free energy curves cross. The yellow curve in (d) corresponds to the coexistence temperature for the fluid mixture. In this case the azeotrope forms within the coexistence region.

(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 $T^*_{C_2H_5OH} = 78.4°C$, $T^*_{H_2O} = 100°C$, while the azeotrope boils at $T^*_{AZ} = 78.2°C$. 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 $\lambda_L \neq 0$, in which case one can find *two* solutions to the energy crossing condition $g_V(x) = g_L(x)$. With two such crossings come two Maxwell constructions, hence the phase diagrams in fig. 2.34. *Generally*, negative azeotropes are found in systems with $\lambda_L < 0$, whereas positive azeotropes are found when $\lambda_L > 0$. As we've seen, such repulsive interactions between the A and B components in general lead to a phase separation below a coexistence temperature $T_{COEX}(x)$

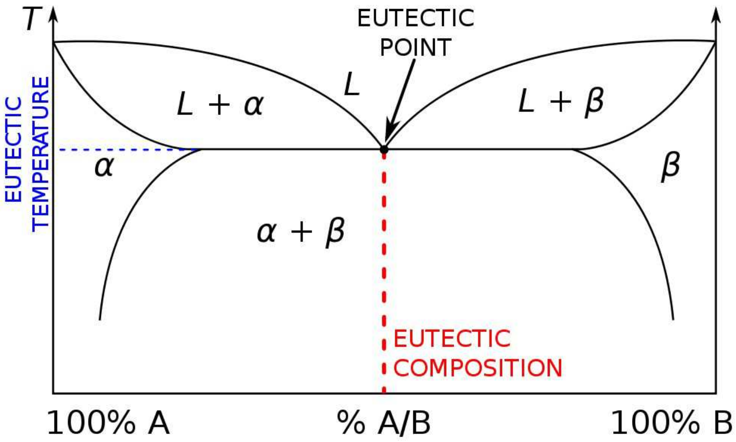


Figure 2.36: Phase diagram for a eutectic mixture in which a liquid L is in equilibrium with two solid phases α and β. The same phase diagram holds for heteroazeotropes, where a vapor is in equilibrium with two liquid phases.

given by eqn. 2.367. 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

$$\zeta_1 \, \mathrm{A}_1 + \zeta_2 \, \mathrm{A}_2 + \cdots + \zeta_{\mathsf{c}} \, \mathrm{A}_{\mathsf{c}} = 0 \quad , \tag{2.377}$$

where A_a represents a chemical formula and ζ_a the corresponding stoichiometric coefficient. For example, we could have

$$-3 \, \mathrm{H}_2 - \mathrm{N}_2 + 2 \, \mathrm{NH}_3 = 0 \qquad (3 \, \mathrm{H}_2 + \mathrm{N}_2 \rightleftharpoons 2 \, \mathrm{NH}_3) \tag{2.378}$$

for which

$$\zeta(\mathrm{H}_2) = -3 \quad , \quad \zeta(\mathrm{N}_2) = -1 \quad , \quad \zeta(\mathrm{NH}_3) = 2 \quad . \tag{2.379}$$

When $\zeta_a > 0$, the corresponding A_a is a *product*; when $\zeta_a < 0$, the corresponding A_a is a *reactant*. The bookkeeping of the coefficients ζ_a which ensures conservation of each individual species of atom in the reaction(s) is known as *stoichiometry*[30]

---

[30] 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

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

$$dG = -S\,dT + V\,dp + \sum_i \mu_a\,dN_a \quad , \tag{2.380}$$

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 $\sum_{a=1}^{\textsf{c}} \zeta_a\,\mu_a = 0$.

Let us investigate the consequences of this relation for ideal gases. From eqn. 2.327 and the relation $\mu_a = -T(\partial S/\partial N_a)$, the chemical potential of the $a^{\rm th}$ species is of the form

$$\mu_a(T,p) = k_{\rm B}T\log p_a + k_{\rm B}T\log\psi_a(T) \quad , \tag{2.381}$$

where each $\psi_a(T)$ has units of inverse pressure[31]. Here $p_a = p\,x_a$ is the partial pressure of species a, where $x_a = N_a\big/\sum_b N_b$ is its dimensionless concentration[32]. In equilibrium we must have

$$\sum_{a=1}^{\textsf{c}} \zeta_a\,\mu_a = \sum_a \zeta_a \Big[\log x_a + \log p + \log\psi_a(T)\Big] = 0 \quad . \tag{2.382}$$

Exponentiating, we obtain the *law of mass action*:

$$\prod_a x_a^{\zeta_a} = \prod_a \big(p\,\psi_a(T)\big)^{-\zeta_a} \equiv \kappa(p,T) \quad . \tag{2.383}$$

The quantity $\kappa(p,T)$ is called the *equilibrium constant*. When κ 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 REACTANTS $\rightleftharpoons$ PRODUCTS is shifted to the right, *i.e.* the products are plentiful and the reactants are scarce. When κ 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.

---

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.

[31] In terms of the functions $\phi_a(T)$ from eqn. 2.327, one has $\psi_a(T) = -\phi_a(T)/k_{\rm B} - \log(k_{\rm B}T/e)$.

[32] Chemists sometimes write $x_a = [{\rm A}_a]$ for the concentration of species a.

The rate at which the equilibrium constant changes with temperature is given by

$$\left(\frac{\partial \log \kappa}{\partial T}\right)_p = -\sum_a \zeta_a \, \frac{\partial \log \psi_a(T)}{\partial T} \quad . \tag{2.384}$$

Now from eqn. 2.381 we have that the enthalpy per particle for species i is

$$\mathsf{h}_a = \mu_a - T\left(\frac{\partial \mu_a}{\partial T}\right)_{p_a} \quad , \tag{2.385}$$

since $\mathsf{H} = G + TS$ and $S = -(\partial G/\partial T)_p$. We find

$$\mathsf{h}_a = -k_{\rm B} T^2 \, \frac{\partial \log \psi_a(T)}{\partial T} \quad , \tag{2.386}$$

and thus

$$\left(\frac{\partial \log \kappa}{\partial T}\right)_p = \frac{\sum_i \zeta_a \, \mathsf{h}_a}{k_{\rm B} T^2} = \frac{\Delta \mathsf{h}}{k_{\rm B} T^2} \quad , \tag{2.387}$$

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 reaction $\mathrm{H_2} + \mathrm{I_2} \rightleftharpoons 2\,\mathrm{HI}$. We have

$$\zeta(\mathrm{H_2}) = -1 \quad , \quad \zeta(\mathrm{I_2}) = -1 \qquad\qquad \zeta(\mathrm{HI}) = 2 \quad . \tag{2.388}$$

Suppose our initial system consists of ν_1^0 moles of $\mathrm{H_2}$, $\nu_2^0 = 0$ moles of $\mathrm{I_2}$, and ν_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

$$\alpha \equiv \frac{x_3^0 - x_3}{x_3} \quad , \tag{2.389}$$

in which case we have

$$x_1 = x_1^0 + \tfrac{1}{2}\alpha x_3^0 \quad , \quad x_2 = \tfrac{1}{2}\alpha x_3^0 \quad , \quad x_3 = (1-\alpha)\, x_3^0 \quad . \tag{2.390}$$

Then the law of mass action gives

$$\frac{4\,(1-\alpha)^2}{\alpha(\alpha + 2r)} = \kappa \quad . \tag{2.391}$$

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 enthalpy of this reaction is positive: $\Delta\mathsf{h} > 0$.

2.13.2 Enthalpy of formation

Most chemical reactions take place under constant pressure. The heat $Q_{\rm if}$ associated with a given isobaric process is

$$Q_{\rm if} = \int\limits_{\rm i}^{\rm f} dE + \int\limits_{\rm i}^{\rm f} p\, dV = (E_{\rm f} - E_{\rm i}) + p\,(V_{\rm f} - V_{\rm i}) = \mathsf{H}_{\rm f} - \mathsf{H}_{\rm i} \quad , \tag{2.392}$$

Formula	Name	State	ΔH^0_f kJ/mol	Formula	Name	State	ΔH^0_f 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	$Ca_3P_2O_8$	Calcium phosphate	gas	-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.

where H is the *enthalpy*,

$$\mathsf{H} = E + pV \quad . \tag{2.393}$$

Note that the enthalpy 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: $\Delta\mathsf{H} = \mathsf{H}_f - \mathsf{H}_i$. If $\Delta\mathsf{H} < 0$ for a given reaction, we call it *exothermic* – this is the case when $Q_{if} < 0$ 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*. Endothermic reactions require that heat be supplied in order for the reaction to proceed. Photosynthesis is an example of an endothermic reaction.

Suppose we have two reactions

$$A + B \xrightarrow{(\Delta\mathsf{H})_1} C \tag{2.394}$$

and

$$C + D \xrightarrow{(\Delta\mathsf{H})_2} E \quad . \tag{2.395}$$

Then we may write

$$A + B + D \xrightarrow{(\Delta\mathsf{H})_3} E \quad , \tag{2.396}$$

with

$$(\Delta\mathsf{H})_1 + (\Delta\mathsf{H})_2 = (\Delta\mathsf{H})_3 \quad . \tag{2.397}$$

We can use this additivity of reaction enthalpies to define a *standard molar enthalpy of formation*. We first define the *standard state* of a pure substance at a given temperature to be its state (gas, liquid, or solid) at a pressure $p = 1\,$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* $\Delta\mathsf{H}^0_f(X)$ 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

$$S + O_2 \xrightarrow{\Delta\mathsf{H}^0_f[SO_2]} SO_2 \quad . \tag{2.398}$$

The enthalpy of formation of any substance in its standard state is zero at all temperatures, by definition: $\Delta\mathsf{H}^0_f[O_2] = \Delta\mathsf{H}^0_f[He] = \Delta\mathsf{H}^0_f[K] = \Delta\mathsf{H}^0_f[Mn] = 0$, *etc.*

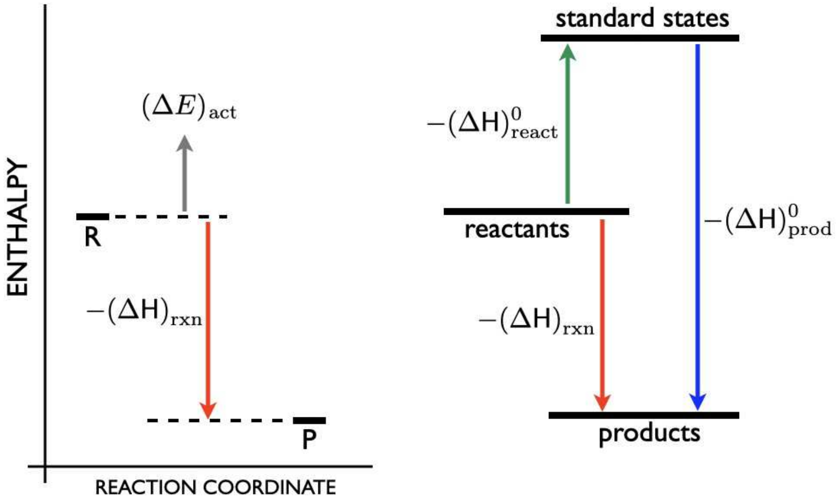


Figure 2.37: Left panel: reaction enthalpy and activation energy (exothermic case shown). Right panel: reaction enthalpy as a difference between enthalpy of formation of reactants and products.

Suppose now we have a reaction

$$a\,A + b\,B \xrightarrow{\Delta\mathsf{H}} c\,C + d\,D \quad . \tag{2.399}$$

To compute the reaction enthalpy $\Delta\mathsf{H}$, 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

$$\Delta\mathsf{H} = -a\,\Delta\mathsf{H}^0_{\rm f}(A) - b\,\Delta\mathsf{H}^0_{\rm f}(B) + c\,\Delta\mathsf{H}^0_{\rm f}(C) + d\,\Delta\mathsf{H}^0_{\rm f}(D) \quad . \tag{2.400}$$

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

$$d\mathsf{H} = dE + p\,dV + V\,dp = đQ + V\,dp \quad , \tag{2.401}$$

hence

$$C_p = \left(\frac{đQ}{dT}\right)_p = \left(\frac{\partial\mathsf{H}}{\partial T}\right)_p \quad . \tag{2.402}$$

We therefore have

$$\mathsf{H}(T,p,\nu) = \mathsf{H}(T_0,p,\nu) + \nu\int\limits_{T_0}^{T} dT'\,c_p(T') \quad . \tag{2.403}$$

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:

$$c_p(T) = \alpha + \beta\,T + \gamma\,T^2 \quad . \tag{2.404}$$

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).)*

Two examples ($300\,\mathrm{K} < T < 1500\,\mathrm{K}$, $p = 1\,\mathrm{atm}$):

$$\begin{array}{rlll} \mathrm{O}_2: & \alpha = 25.503\,\dfrac{\mathrm{J}}{\mathrm{mol\,K}} \quad , & \beta = 13.612\times 10^{-3}\,\dfrac{\mathrm{J}}{\mathrm{mol\,K^2}} \quad , & \gamma = -42.553\times 10^{-7}\,\dfrac{\mathrm{J}}{\mathrm{mol\,K^3}} \\ \mathrm{H_2O}: & \alpha = 30.206\,\dfrac{\mathrm{J}}{\mathrm{mol\,K}} \quad , & \beta = 9.936\times 10^{-3}\,\dfrac{\mathrm{J}}{\mathrm{mol\,K^2}} \quad , & \gamma = 11.14\times 10^{-7}\,\dfrac{\mathrm{J}}{\mathrm{mol\,K^3}} \end{array}$$

If all the gaseous components in a reaction can be approximated as ideal, then we may write

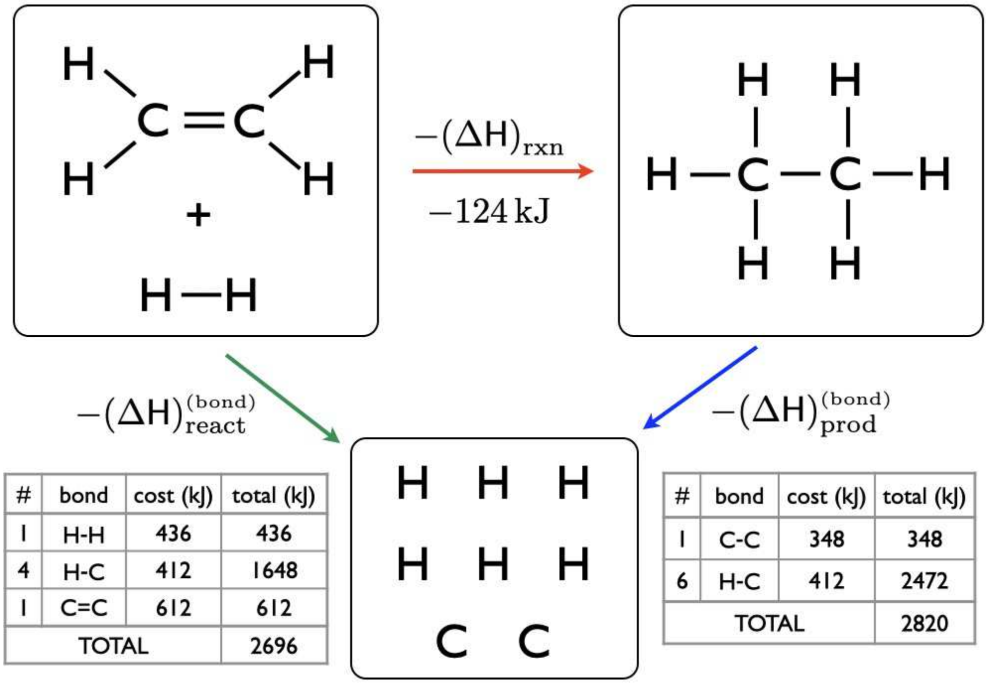
$$(\Delta\mathsf{H})_{\mathrm{rxn}} = (\Delta E)_{\mathrm{rxn}} + \sum_a \zeta_a\, RT \quad , \tag{2.405}$$

where the subscript ‘rxn’ stands for ‘reaction’. Here $(\Delta E)_{\mathrm{rxn}}$ 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



#	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 $2\,H_2(g) + O_2(g) \longrightarrow 2\,H_2O(l)$. We then have, from the table,

$$\begin{aligned}(\Delta H)_{\rm rxn} &= 2\,\mathsf{h}[H{-}H] + \mathsf{h}[O{=}O] - 4\,\mathsf{h}[H{-}O] \\ &= -483\,\mathrm{kJ/mol\ O_2} \quad .\end{aligned} \tag{2.406}$$

Thus, 483 kJ of heat would be released for every two moles of H_2O produced, if the H_2O were in the gaseous phase. Since H_2O is liquid at STP, we should also include the condensation energy of the gaseous water vapor into liquid water. At $T = 100°\mathrm{C}$ the latent heat of vaporization is $\tilde{\ell} = 2270\,\mathrm{J/g}$, but at $T = 20°\mathrm{C}$, one has $\tilde{\ell} = 2450\,\mathrm{J/g}$, hence with $M = 18$ we have $\ell = 44.1\,\mathrm{kJ/mol}$. Therefore, the heat produced by the reaction $2\,H_2(g) + O_2(g) \rightleftharpoons 2\,H_2O(l)$ is $(\Delta H)_{\rm rxn} = -571.2\,\mathrm{kJ\,/\,mol\ O_2}$. Since the reaction produces two moles of water, we conclude that the enthalpy of formation of liquid water at STP is half this value: $\Delta H^0_f[H_2O] = 285.6\,\mathrm{kJ/mol}$.

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 $\Delta H^0_f[C(g)] = 718\,\mathrm{kJ/mol}$. We may now write

$$\begin{aligned} 2\,\mathrm{C(g)} + 4\,\mathrm{H(g)} &\xrightarrow{-2260\,\mathrm{kJ}} \mathrm{C_2H_4(g)} \\ 2\,\mathrm{C(s)} &\xrightarrow{1436\,\mathrm{kJ}} 2\,\mathrm{C(g)} \\ 2\,\mathrm{H_2(g)} &\xrightarrow{872\,\mathrm{kJ}} 4\,\mathrm{H(g)} \quad . \end{aligned}$$

Thus, using Hess's law, *i.e.* adding up these reaction equations, we have

$$2\,\mathrm{C(s)} + 2\,\mathrm{H_2(g)} \xrightarrow{48\,\mathrm{kJ}} \mathrm{C_2H_4(g)} \quad .$$

Thus, the formation of ethene is endothermic. For ethane,

$$\begin{aligned} 2\,\mathrm{C(g)} + 6\,\mathrm{H(g)} &\xrightarrow{-2820\,\mathrm{kJ}} \mathrm{C_2H_6(g)} \\ 2\,\mathrm{C(s)} &\xrightarrow{1436\,\mathrm{kJ}} 2\,\mathrm{C(g)} \\ 3\,\mathrm{H_2(g)} &\xrightarrow{1306\,\mathrm{kJ}} 6\,\mathrm{H(g)} \end{aligned}$$

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,

$$\mathrm{C_2H_4(g)} + \mathrm{H_2(g)} \xrightarrow{-48\,\mathrm{kJ}} 2\,\mathrm{C(s)} + 3\,\mathrm{H_2(g)} \xrightarrow{-76\,\mathrm{kJ}} \mathrm{C_2H_6(g)} \quad , \tag{2.407}$$

which says

$$\mathrm{C_2H_4(g)} + \mathrm{H_2(g)} \xrightarrow{-124\,\mathrm{kJ}} \mathrm{C_2H_6(g)} \quad . \tag{2.408}$$

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

$$đW = A_i\, dx_i \quad . \tag{2.409}$$

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* $\exp\big[L(\boldsymbol{x})\big]$ is a function which, when divided into $đF$, yields an exact differential:

$$dU = e^{-L}\, đW = \frac{\partial U}{\partial x_i}\, dx_i \quad . \tag{2.410}$$

Clearly we must have

$$\frac{\partial^2 U}{\partial x_i\, \partial x_j} = \frac{\partial}{\partial x_i}\left(e^{-L} A_j\right) = \frac{\partial}{\partial x_j}\left(e^{-L} A_i\right) \quad . \tag{2.411}$$

Applying the Leibniz rule and then multiplying by e^L yields

$$\frac{\partial A_j}{\partial x_i} - A_j \, \frac{\partial L}{\partial x_i} = \frac{\partial A_i}{\partial x_j} - A_i \, \frac{\partial L}{\partial x_j} \quad . \tag{2.412}$$

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

$$\Omega_{ijk} \, \frac{\partial L}{\partial x_k} = F_{ij} \quad , \tag{2.413}$$

where

$$\Omega_{ijk} = A_j \, \delta_{ik} - A_i \, \delta_{jk} \tag{2.414}$$

$$F_{ij} = \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j} \quad . \tag{2.415}$$

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

$$\frac{dy}{dx} = -\frac{A_x(x,y)}{A_y(x,y)} \quad . \tag{2.416}$$

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

$$\begin{aligned} 0 = \frac{dU_c}{dx} &= \frac{\partial U_x}{\partial x} + \frac{\partial U_c}{\partial y} \, \frac{dy}{dx} \\ &= \frac{\partial U_c}{\partial x} - \frac{A_x}{A_y} \, \frac{\partial U_c}{\partial y} \quad . \end{aligned} \tag{2.417}$$

Thus,

$$\frac{\partial U_c}{\partial x} \, A_y = \frac{\partial U_c}{\partial y} \, A_x \equiv e^{-L} A_x \, A_y \quad . \tag{2.418}$$

This equation defines the integrating factor L :

$$L = -\log\!\bigg(\frac{1}{A_x} \, \frac{\partial U_c}{\partial x}\bigg) = -\log\!\bigg(\frac{1}{A_y} \, \frac{\partial U_c}{\partial y}\bigg) \quad . \tag{2.419}$$

We now have that

$$A_x = e^L \, \frac{\partial U_c}{\partial x} \quad , \quad A_y = e^L \, \frac{\partial U_c}{\partial y} \quad , \tag{2.420}$$

and hence

$$e^{-L} \, \text{đ}W = \frac{\partial U_c}{\partial x} \, dx + \frac{\partial U_c}{\partial y} \, dy = dU_c \quad . \tag{2.421}$$

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

$$p_{\rm A} V_1 = p_{\rm B} V_2 \qquad , \qquad p_{\rm D} V_1 = p_{\rm C} V_2 \quad , \tag{2.422}$$

which says

$$\frac{p_{\rm B}}{p_{\rm A}} = \frac{p_{\rm C}}{p_{\rm D}} = \frac{V_1}{V_2} \quad . \tag{2.423}$$

AB: This isothermal expansion is the power stroke. Assuming ν moles of ideal gas throughout, we have $pV = \nu R T_2 = p_1 V_1$, hence

$$W_{\rm AB} = \int\limits_{V_1}^{V_2} dV \, \frac{\nu R T_2}{V} = \nu R T_2 \, \log\bigg(\frac{V_2}{V_1}\bigg) \quad . \tag{2.424}$$

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

$$\Delta E_{\rm BC} = E_{\rm C} - E_{\rm B} = \frac{\nu R (T_1 - T_2)}{\gamma - 1}, \tag{2.425}$$

which is negative. Since $W_{\rm BC} = 0$, we have $Q_{\rm BC} = \Delta E_{\rm BC}$.

CD: Isothermal compression. Clearly

$$W_{\rm CD} = \int\limits_{V_2}^{V_1} dV \, \frac{\nu R T_1}{V} = -\nu R T_1 \, \log\bigg(\frac{V_2}{V_1}\bigg) \quad . \tag{2.426}$$

Since CD is an isotherm, we have $E_{\rm C} = E_{\rm D}$, and from $\Delta E_{\rm CD} = 0$ we conclude $Q_{\rm CD} = W_{\rm CD}$.

DA: Isochoric heating. Since $dV = 0$ we have $W_{\rm DA} = 0$. The energy change is given by

$$\Delta E_{\rm DA} = E_{\rm A} - E_{\rm D} = \frac{\nu R (T_2 - T_1)}{\gamma - 1} \quad , \tag{2.427}$$

which is positive, and opposite to $\Delta E_{\rm BC}$. Since $W_{\rm DA} = 0$, we have $Q_{\rm DA} = \Delta E_{\rm DA}$.

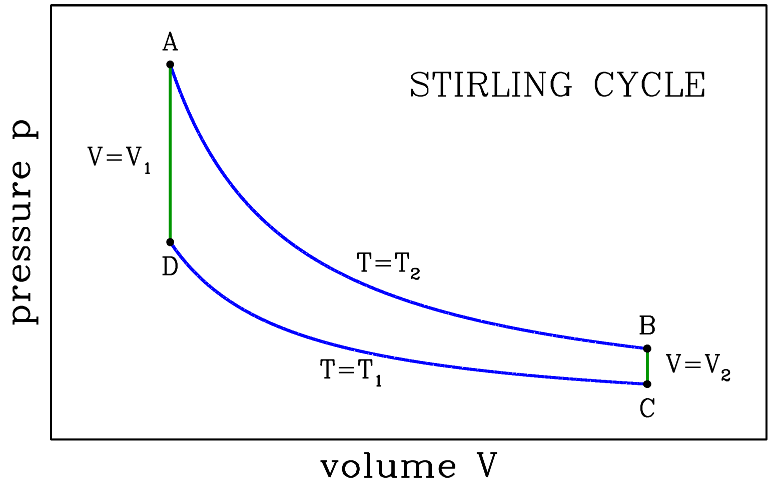


Figure 2.39: A Stirling cycle consists of two isotherms (blue) and two isochores (green).

We now add up all the work contributions to obtain

$$\begin{aligned} W &= W_{\rm AB} + W_{\rm BC} + W_{\rm CD} + W_{\rm DA} \\ &= \nu R (T_2 - T_1)\, \log\left(\frac{V_2}{V_1}\right) \quad . \end{aligned} \tag{2.428}$$

The cycle efficiency is once again

$$\eta = \frac{W}{Q_{\rm AB}} = 1 - \frac{T_1}{T_2} \quad . \tag{2.429}$$

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,

$$p_{\rm A}\, V_1^\gamma = p_{\rm B}\, V_2^\gamma \quad , \quad p_{\rm D}\, V_1^\gamma = p_{\rm C}\, V_2^\gamma \quad , \tag{2.430}$$

which says

$$\frac{p_{\rm B}}{p_{\rm A}} = \frac{p_{\rm C}}{p_{\rm D}} = \left(\frac{V_1}{V_2}\right)^\gamma \quad . \tag{2.431}$$

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

$$W_{\rm AB} = \frac{p_{\rm A} V_1 - p_{\rm B} V_2}{\gamma - 1} = \frac{p_{\rm A} V_1}{\gamma - 1}\left[1 - \left(\frac{V_1}{V_2}\right)^{\gamma - 1}\right] \quad . \tag{2.432}$$

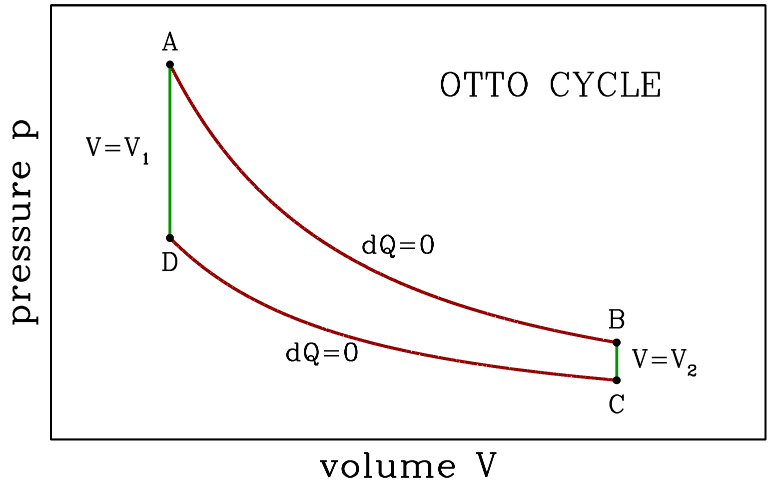


Figure 2.40: An Otto cycle consists of two adiabats (dark red) and two isochores (green).

Note that this result can also be obtained from the adiabatic equation of state $pV^\gamma = p_{\rm A} V_1^\gamma$:

$$W_{\rm AB} = \int\limits_{V_1}^{V_2} p\,dV = p_{\rm A} V_1^\gamma \int\limits_{V_1}^{V_2} dV\,V^{-\gamma} = \frac{p_{\rm A} V_1}{\gamma-1}\left[1-\left(\frac{V_1}{V_2}\right)^{\gamma-1}\right] \quad . \tag{2.433}$$

BC: Isochoric cooling (exhaust); $dV = 0$ hence $W_{\rm BC} = 0$. The heat $Q_{\rm BC}$ absorbed is then

$$Q_{\rm BC} = E_{\rm C} - E_{\rm B} = \frac{V_2}{\gamma-1}\,(p_{\rm C} - p_{\rm B}) \quad . \tag{2.434}$$

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 $W_{\rm CD} = E_{\rm C} - E_{\rm D}$:

$$W_{\rm CD} = \frac{p_{\rm C} V_2 - p_{\rm D} V_1}{\gamma-1} = -\frac{p_{\rm D} V_1}{\gamma-1}\left[1-\left(\frac{V_1}{V_2}\right)^{\gamma-1}\right] \quad . \tag{2.435}$$

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

$$Q_{\rm DA} = E_{\rm A} - E_{\rm D} = \frac{V_1}{\gamma-1}\,(p_{\rm A} - p_{\rm D}) \quad . \tag{2.436}$$

The total work done per cycle is then

$$\begin{aligned} W &= W_{\rm AB} + W_{\rm BC} + W_{\rm CD} + W_{\rm DA} \\ &= \frac{(p_{\rm A} - p_{\rm D})V_1}{\gamma-1}\left[1-\left(\frac{V_1}{V_2}\right)^{\gamma-1}\right] \quad , \end{aligned} \tag{2.437}$$

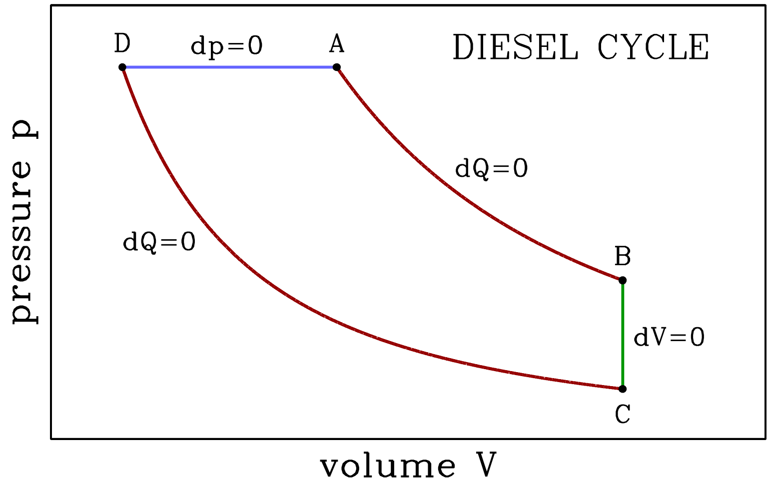


Figure 2.41: A Diesel cycle consists of two adiabats (dark red), one isobar (light blue), and one isochore (green).

and the efficiency is defined to be

$$\eta \equiv \frac{W}{Q_{\rm DA}} = 1 - \left(\frac{V_1}{V_2}\right)^{\gamma-1} \quad . \tag{2.438}$$

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 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,

$$\begin{aligned} W &= p_{\rm A}(V_{\rm A} - V_{\rm D}) + \frac{p_{\rm A} V_{\rm A} - p_{\rm B} V_{\rm B}}{\gamma - 1} + \frac{p_{\rm C} V_{\rm C} - p_{\rm D} V_{\rm D}}{\gamma - 1} \\ &= \frac{\gamma\, p_{\rm A}(V_{\rm A} - V_{\rm D})}{\gamma - 1} - \frac{(p_{\rm B} - p_{\rm C}) V_{\rm B}}{\gamma - 1} \end{aligned} \tag{2.439}$$

and

$$Q_{\rm DA} = \frac{\gamma\, p_{\rm A}(V_{\rm A} - V_{\rm D})}{\gamma - 1} \quad . \tag{2.440}$$

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,

$$p_{\rm B} = p_{\rm A} \cdot \left(\frac{V_{\rm A}}{V_{\rm B}}\right)^{\gamma} \quad , \qquad p_{\rm C} = p_{\rm A} \cdot \left(\frac{V_{\rm D}}{V_{\rm B}}\right)^{\gamma} \quad , \tag{2.441}$$

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

$$\eta = \frac{W}{Q_{\rm DA}} = 1 - \frac{1}{\gamma}\,\frac{r^{1-\gamma}(s^\gamma - 1)}{s-1} \quad . \tag{2.442}$$

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 Thus,

$$p_2\, V_{\rm A}^\gamma = p_1\, V_{\rm D}^\gamma \qquad , \qquad p_2\, V_{\rm B}^\gamma = p_1\, V_{\rm C}^\gamma \quad , \tag{2.443}$$

which says

$$\frac{V_{\rm D}}{V_{\rm A}} = \frac{V_{\rm C}}{V_{\rm B}} = \left(\frac{p_2}{p_1}\right)^{\gamma^{-1}} \quad . \tag{2.444}$$

AB: This isobaric expansion at $p = p_2$ is the power stroke. We have

$$W_{\rm AB} = \int\limits_{V_{\rm A}}^{V_{\rm B}} \! dV\, p_2 = p_2\,(V_{\rm B} - V_{\rm A}) \tag{2.445}$$

$$\Delta E_{\rm AB} = E_{\rm B} - E_{\rm A} = \frac{p_2\,(V_{\rm B} - V_{\rm A})}{\gamma - 1} \tag{2.446}$$

$$Q_{\rm AB} = \Delta E_{\rm AB} + W_{\rm AB} = \frac{\gamma\, p_2\,(V_{\rm B} - V_{\rm A})}{\gamma - 1} \quad . \tag{2.447}$$

BC: Adiabatic expansion; $Q_{\rm BC} = 0$ and $W_{\rm BC} = E_{\rm B} - E_{\rm C}$. The work done by the gas is

$$\begin{aligned} W_{\rm BC} &= \frac{p_2 V_{\rm B} - p_1 V_{\rm C}}{\gamma - 1} = \frac{p_2 V_{\rm B}}{\gamma - 1}\left(1 - \frac{p_1}{p_2}\cdot\frac{V_{\rm C}}{V_{\rm B}}\right) \\ &= \frac{p_2\, V_{\rm B}}{\gamma - 1}\left[1 - \left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}}\right] \quad . \end{aligned} \tag{2.448}$$

CD: Isobaric compression at $p = p_1$.

$$W_{\rm CD} = \int\limits_{V_{\rm C}}^{V_{\rm D}} \! dV\, p_1 = p_1\,(V_{\rm D} - V_{\rm C}) = -p_2\,(V_{\rm B} - V_{\rm A}) \left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}} \tag{2.449}$$

$$\Delta E_{\rm CD} = E_{\rm D} - E_{\rm C} = \frac{p_1\,(V_{\rm D} - V_{\rm C})}{\gamma - 1} \tag{2.450}$$

$$Q_{\rm CD} = \Delta E_{\rm CD} + W_{\rm CD} = -\frac{\gamma\, p_2}{\gamma - 1}\,(V_{\rm B} - V_{\rm A}) \left(\frac{p_1}{p_2}\right)^{1-\gamma^{-1}} \quad . \tag{2.451}$$

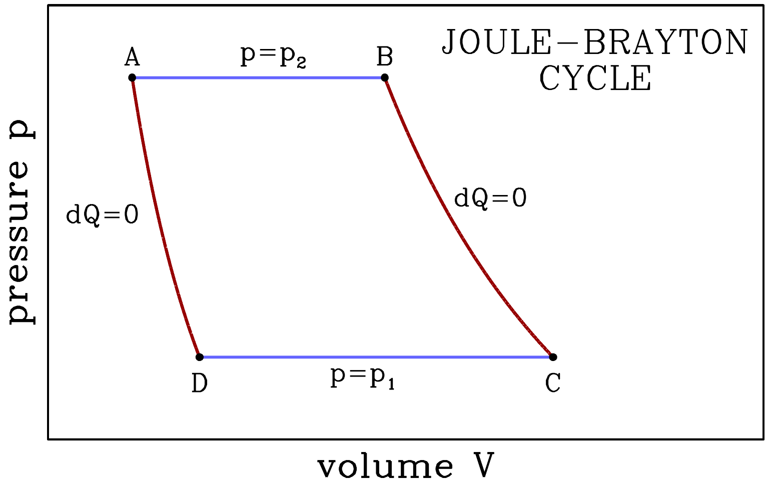


Figure 2.42: A Joule-Brayton cycle consists of two adiabats (dark red) and two isobars (light blue).

DA: Adiabatic expansion; $Q_{\rm DA}=0$ and $W_{\rm DA}=E_{\rm D}-E_{\rm A}$. The work done by the gas is

$$\begin{aligned} W_{\rm DA} &= \frac{p_1 V_{\rm D} - p_2 V_{\rm A}}{\gamma-1} = -\frac{p_2 V_{\rm A}}{\gamma-1}\bigg(1-\frac{p_1}{p_2}\cdot\frac{V_{\rm D}}{V_{\rm A}}\bigg) \\ &= -\frac{p_2\,V_{\rm A}}{\gamma-1}\left[1-\bigg(\frac{p_1}{p_2}\bigg)^{1-\gamma^{-1}}\right] \quad . \end{aligned} \tag{2.452}$$

The total work done per cycle is then

$$\begin{aligned} W &= W_{\rm AB} + W_{\rm BC} + W_{\rm CD} + W_{\rm DA} \\ &= \frac{\gamma\,p_2\,(V_{\rm B}-V_{\rm A})}{\gamma-1}\left[1-\bigg(\frac{p_1}{p_2}\bigg)^{1-\gamma^{-1}}\right] \end{aligned} \tag{2.453}$$

and the efficiency is defined to be

$$\eta \equiv \frac{W}{Q_{\rm AB}} = 1-\bigg(\frac{p_1}{p_2}\bigg)^{1-\gamma^{-1}} \quad . \tag{2.454}$$

2.16 Appendix III : Legendre Transformations

A *convex function* of a single variable $f(x)$ is one for which $f''(x)>0$ everywhere. The *Legendre transform* of a 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 $g(p)=F\big(x(p),p\big)$.[33] The value $x(p)$ is unique if $f(x)$ is convex, since $x(p)$ is determined by the equation

$$f'\big(x(p)\big)=p \quad . \tag{2.455}$$

---

[33] Note that $g(p)$ may be a negative number, if the line $y=px$ lies everywhere below $f(x)$.

Note that from $p = f'\big(x(p)\big)$ we have, according to the chain rule,

$$\frac{d}{dp}\, f'\big(x(p)\big) = f''\big(x(p)\big)\, x'(p) \qquad \Longrightarrow \qquad x'(p) = \Big[f''\big(x(p)\big)\Big]^{-1} \quad . \tag{2.456}$$

From this, we can prove that $g(p)$ is itself convex:

$$\begin{aligned} g'(p) &= \frac{d}{dp}\Big[p\, x(p) - f\big(x(p)\big)\Big] \\ &= p\, x'(p) + x(p) - f'\big(x(p)\big)\, x'(p) = x(p) \quad , \end{aligned} \tag{2.457}$$

hence

$$g''(p) = x'(p) = \Big[f''\big(x(p)\big)\Big]^{-1} > 0 \quad . \tag{2.458}$$

In higher dimensions, the generalization of the definition $f''(x) > 0$ is that a function $F(x_1, \ldots, x_n)$ is convex if the matrix of second derivatives, called the *Hessian*,

$$H_{ij}(\boldsymbol{x}) = \frac{\partial^2 F}{\partial x_i\, \partial x_j} \tag{2.459}$$

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

$$\boldsymbol{G}(\boldsymbol{p}) = \boldsymbol{p} \cdot \boldsymbol{x} - F(\boldsymbol{x}) \tag{2.460}$$

where

$$\boldsymbol{p} = \boldsymbol{\nabla} F \quad . \tag{2.461}$$

Note that

$$dG = \boldsymbol{x} \cdot d\boldsymbol{p} + \boldsymbol{p} \cdot d\boldsymbol{x} - \boldsymbol{\nabla} F \cdot d\boldsymbol{x} = \boldsymbol{x} \cdot d\boldsymbol{p} \quad , \tag{2.462}$$

which establishes that G is a function of $\boldsymbol{p}$ and that

$$\frac{\partial G}{\partial p_j} = x_j \quad . \tag{2.463}$$

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.

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

$$G(\boldsymbol{p}, \boldsymbol{y}) = \boldsymbol{p} \cdot \boldsymbol{x} - F(\boldsymbol{x}, \boldsymbol{y}) \quad , \tag{2.464}$$

where

$$p_a = \frac{\partial F}{\partial x_a} \qquad (a = 1, \ldots, m) \quad . \tag{2.465}$$

These equations are then to be inverted to yield

$$x_a = x_a(\boldsymbol{p}, \boldsymbol{y}) = \frac{\partial G}{\partial p_a} \quad . \tag{2.466}$$

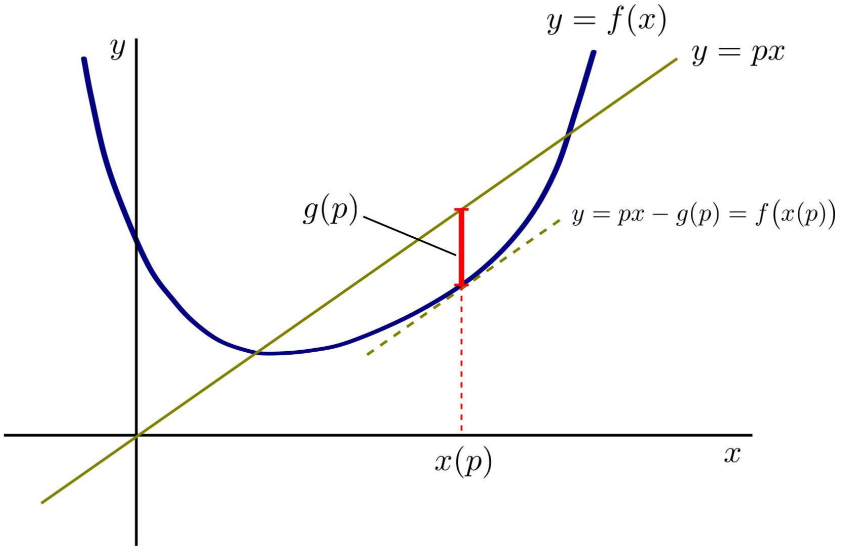


Figure 2.43: Construction for the Legendre transformation of a function $f(x)$.

Note that

$$p_a = \frac{\partial F}{\partial x_a}\big(\boldsymbol{x}(\boldsymbol{p},\boldsymbol{y}),\boldsymbol{y}\big) \quad . \tag{2.467}$$

Thus, from the chain rule,

$$\delta_{ab} = \frac{\partial p_a}{\partial p_b} = \frac{\partial^2 F}{\partial x_a\,\partial x_c}\,\frac{\partial x_c}{\partial p_b} = \frac{\partial^2 F}{\partial x_a\,\partial x_c}\,\frac{\partial^2 G}{\partial p_c\,\partial p_b} \quad , \tag{2.468}$$

which says

$$\frac{\partial^2 G}{\partial p_a\,\partial p_b} = \frac{\partial x_a}{\partial p_b} = \mathsf{K}^{-1}_{ab} \quad , \tag{2.469}$$

where the $m \times m$ partial Hessian is

$$\frac{\partial^2 F}{\partial x_a\,\partial x_b} = \frac{\partial p_a}{\partial x_b} = \mathsf{K}_{ab} \quad . \tag{2.470}$$

Note that $\mathsf{K}_{ab} = \mathsf{K}_{ba}$ is symmetric. And with respect to the $\boldsymbol{y}$ coordinates,

$$\frac{\partial^2 G}{\partial y_\mu\,\partial y_\nu} = -\frac{\partial^2 F}{\partial y_\mu\,\partial y_\nu} = -\mathsf{L}_{\mu\nu} \quad , \tag{2.471}$$

where

$$\mathsf{L}_{\mu\nu} = \frac{\partial^2 F}{\partial y_\mu\,\partial y_\nu} \tag{2.472}$$

is the partial Hessian in the $\boldsymbol{y}$ coordinates. Now it is easy to see that if the full $q \times q$ Hessian matrix H_{ij} is positive definite, then any submatrix such as K_{ab} or $\mathsf{L}_{\mu\nu}$ 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

$$\frac{\partial x_i}{\partial z_k} = \frac{\partial x_i}{\partial y_j}\,\frac{\partial y_j}{\partial z_k} \quad . \tag{2.473}$$

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

$$\mathsf{det}\bigg(\frac{\partial x_i}{\partial z_k}\bigg) \equiv \frac{\partial(x_1, \ldots, x_n)}{\partial(z_1, \ldots, z_n)} \quad . \tag{2.474}$$

Such a determinant is called a Jacobian. Now if $A = BC$, then $\mathsf{det}(A) = \mathsf{det}(B) \cdot \mathsf{det}(C)$. Thus,

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(z_1, \ldots, z_n)} = \frac{\partial(x_1, \ldots, x_n)}{\partial(y_1, \ldots, y_n)} \cdot \frac{\partial(y_1, \ldots, y_n)}{\partial(z_1, \ldots, z_n)} \quad . \tag{2.475}$$

Recall also that

$$\frac{\partial x_i}{\partial x_k} = \delta_{ik} \quad . \tag{2.476}$$

Consider the case $n = 2$. We have

$$\frac{\partial(x,y)}{\partial(u,v)} = \mathsf{det}\begin{pmatrix} \big(\frac{\partial x}{\partial u}\big)_v & \big(\frac{\partial x}{\partial v}\big)_u \\ \big(\frac{\partial y}{\partial u}\big)_v & \big(\frac{\partial y}{\partial v}\big)_u \end{pmatrix} = \bigg(\frac{\partial x}{\partial u}\bigg)_v \bigg(\frac{\partial y}{\partial v}\bigg)_u - \bigg(\frac{\partial x}{\partial v}\bigg)_u \bigg(\frac{\partial y}{\partial u}\bigg)_v \quad . \tag{2.477}$$

We also have

$$\frac{\partial(x,y)}{\partial(u,v)} \cdot \frac{\partial(u,v)}{\partial(r,s)} = \frac{\partial(x,y)}{\partial(r,s)} \quad . \tag{2.478}$$

From this simple mathematics follows several very useful results.

1) First, write

$$\frac{\partial(x,y)}{\partial(u,v)} = \bigg[\frac{\partial(u,v)}{\partial(x,y)}\bigg]^{-1} \quad . \tag{2.479}$$

Now let $v = y$:

$$\frac{\partial(x,y)}{\partial(u,y)} = \bigg(\frac{\partial x}{\partial u}\bigg)_y = \frac{1}{\big(\frac{\partial u}{\partial x}\big)_y} \quad . \tag{2.480}$$

Thus,

$$\bigg(\frac{\partial x}{\partial u}\bigg)_y = 1 \Big/ \bigg(\frac{\partial u}{\partial x}\bigg)_y \tag{2.481}$$

2) Second, we have

$$\frac{\partial(x,y)}{\partial(u,y)} = \bigg(\frac{\partial x}{\partial u}\bigg)_y = \frac{\partial(x,y)}{\partial(x,u)} \cdot \frac{\partial(x,u)}{\partial(u,y)} = -\bigg(\frac{\partial y}{\partial u}\bigg)_x \bigg(\frac{\partial x}{\partial y}\bigg)_u \quad ,$$

which is to say

$$\left(\frac{\partial x}{\partial y}\right)_{u}\left(\frac{\partial y}{\partial u}\right)_{x}=-\left(\frac{\partial x}{\partial u}\right)_{y} \quad . \tag{2.482}$$

Invoking eqn. 2.481, we conclude that

$$\left(\frac{\partial x}{\partial y}\right)_{u}\left(\frac{\partial y}{\partial u}\right)_{x}\left(\frac{\partial u}{\partial x}\right)_{y}=-1 \quad . \tag{2.483}$$

3) Third, we have

$$\frac{\partial(x,v)}{\partial(u,v)}=\frac{\partial(x,v)}{\partial(y,v)}\cdot\frac{\partial(y,v)}{\partial(u,v)} \quad , \tag{2.484}$$

which says

$$\left(\frac{\partial x}{\partial u}\right)_{v}=\left(\frac{\partial x}{\partial y}\right)_{v}\left(\frac{\partial y}{\partial u}\right)_{v} \tag{2.485}$$

This is simply the chain rule of partial differentiation.

4) Fourth, we have

$$\begin{aligned}\frac{\partial(x,y)}{\partial(u,y)}&=\frac{\partial(x,y)}{\partial(u,v)}\cdot\frac{\partial(u,v)}{\partial(u,y)}\\&=\left(\frac{\partial x}{\partial u}\right)_{v}\left(\frac{\partial y}{\partial v}\right)_{u}\left(\frac{\partial v}{\partial y}\right)_{u}-\left(\frac{\partial x}{\partial v}\right)_{u}\left(\frac{\partial y}{\partial u}\right)_{v}\left(\frac{\partial v}{\partial y}\right)_{u} \quad ,\end{aligned} \tag{2.486}$$

which says

$$\left(\frac{\partial x}{\partial u}\right)_{y}=\left(\frac{\partial x}{\partial u}\right)_{v}-\left(\frac{\partial x}{\partial y}\right)_{u}\left(\frac{\partial y}{\partial u}\right)_{v} \tag{2.487}$$

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,

$$\left(\frac{\partial S}{\partial V}\right)_{p,T}=\frac{S}{V} \quad . \tag{2.488}$$

The reason should be obvious. In the above example, $S(p,V,T)=V\phi(p,T)$, where ϕ 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

$$\left(\frac{\partial S}{\partial V}\right)_{p,T}=\left(\frac{\partial S}{\partial V}\right)_{p,\mu}=\left(\frac{\partial S}{\partial V}\right)_{n,T}=\frac{S}{V} \quad , \tag{2.489}$$

where $n=N/V$ is the particle density.

6) Sixth, suppose we have a function $\varPhi(y,v)$ and we write

$$d\varPhi=x\,dy+u\,dv \quad . \tag{2.490}$$

That is,

$$x=\left(\frac{\partial \varPhi}{\partial y}\right)_{\!v}\equiv \varPhi_y \qquad , \qquad u=\left(\frac{\partial \varPhi}{\partial v}\right)_{\!y}\equiv \varPhi_v \quad . \tag{2.491}$$

Now we may write

$$dx=\varPhi_{yy}\,dy+\varPhi_{yv}\,dv \tag{2.492}$$

$$du=\varPhi_{vy}\,dy+\varPhi_{vv}\,dv \quad . \tag{2.493}$$

If we demand $du=0$, this yields

$$\left(\frac{\partial x}{\partial u}\right)_{\!v}=\frac{\varPhi_{yy}}{\varPhi_{vy}} \quad . \tag{2.494}$$

Note that $\varPhi_{vy}=\varPhi_{yv}$. From the equation $du=0$ we also derive

$$\left(\frac{\partial y}{\partial v}\right)_{\!u}=-\frac{\varPhi_{vv}}{\varPhi_{vy}} \quad . \tag{2.495}$$

Next, we use eqn. 2.493 with $du=0$ to eliminate dy in favor of dv, and then substitute into eqn. 2.492. This yields

$$\left(\frac{\partial x}{\partial v}\right)_{\!u}=\varPhi_{yv}-\frac{\varPhi_{yy}\,\varPhi_{vv}}{\varPhi_{vy}} \quad . \tag{2.496}$$

Finally, eqn. 2.493 with $dv=0$ yields

$$\left(\frac{\partial y}{\partial u}\right)_{\!v}=\frac{1}{\varPhi_{vy}} \quad . \tag{2.497}$$

Combining the results of eqns. 2.494, 2.495, 2.496, and 2.497, we have

$$\begin{aligned}\frac{\partial(x,y)}{\partial(u,v)}&=\left(\frac{\partial x}{\partial u}\right)_{\!v}\left(\frac{\partial y}{\partial v}\right)_{\!u}-\left(\frac{\partial x}{\partial v}\right)_{\!u}\left(\frac{\partial y}{\partial u}\right)_{\!v}\\&=\left(\frac{\varPhi_{yy}}{\varPhi_{vy}}\right)\left(-\frac{\varPhi_{vv}}{\varPhi_{vy}}\right)-\left(\varPhi_{yv}-\frac{\varPhi_{yy}\,\varPhi_{vv}}{\varPhi_{vy}}\right)\left(\frac{1}{\varPhi_{vy}}\right)=-1 \quad .\end{aligned} \tag{2.498}$$

Thus, if $\varPhi=E(S,V)$, then $(x,y)=(T,S)$ and $(u,v)=(-p,V)$, we have

$$\frac{\partial(x,y)}{\partial(u,v)}=\frac{\partial(T,S)}{\partial(-p,V)}=-1 \quad . \tag{2.499}$$

Nota bene: It is important to understand what other quantities are kept constant, else problems may arise. For example, take $\varPhi=G(p,N)$, whence $(x,y)=(V,p)$ and $(u,v)=(\mu,N)$, hence eqn. 2.498 yields

$$\frac{\partial(x,y)}{\partial(u,v)}=\frac{\partial(V,p)}{\partial(\mu,N)}=\frac{\partial(-p,V)}{\partial(\mu,N)}=-1 \quad . \tag{2.500}$$

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 $\varPhi = F(T,N)$, then $(x,y) = (-S,T)$ and $(u,v) = (\mu,N)$, and again from eqn. 2.498 we have

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{\partial(-S,T)}{\partial(\mu,N)} = \frac{\partial(T,S)}{\partial(\mu,N)} = -1 \quad . \tag{2.501}$$

What has gone wrong?

The problem is that we haven't been consistent in what is being held constant. In eqn. 2.499 we are implicitly taking $dN = 0$ since $(y,v) = (S,V)$ and $E = E(S,V,N)$. In eqn. 2.500 we take $dT = 0$ since $(y,v) = (p,N)$ and $G = G(T,p,N)$. And in eqn. 2.501 we take $dp = 0$ 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

$$\frac{\partial(x,y,s)}{\partial(u,v,s)} = \left(\frac{\partial x}{\partial u}\right)_{v,s} \left(\frac{\partial y}{\partial v}\right)_{u,s} - \left(\frac{\partial x}{\partial v}\right)_{u,s} \left(\frac{\partial y}{\partial u}\right)_{v,s} = -1 \quad . \tag{2.502}$$

Thus, if

$$dE = T\,dS + y\,dX + \mu\,dN \quad , \tag{2.503}$$

where $(y,X) = (-p,V)$ or (H^α, M^α) or (E^α, P^α), the appropriate thermodynamic relations are

$$\begin{aligned} \frac{\partial(T,S,N)}{\partial(y,X,N)} &= -1 & \qquad \frac{\partial(T,S,\mu)}{\partial(y,X,\mu)} &= -1 \\ \frac{\partial(\mu,N,X)}{\partial(T,S,X)} &= -1 & \qquad \frac{\partial(\mu,N,y)}{\partial(T,S,y)} &= -1 \\ \frac{\partial(y,X,S)}{\partial(\mu,N,S)} &= -1 & \qquad \frac{\partial(y,X,T)}{\partial(\mu,N,T)} &= -1 \end{aligned} \tag{2.504}$$

For example,

$$\frac{\partial(T,S,N)}{\partial(-p,V,N)} = \frac{\partial(-p,V,S)}{\partial(\mu,N,S)} = \frac{\partial(\mu,N,V)}{\partial(T,S,V)} = -1 \tag{2.505}$$

and

$$\frac{\partial(T,S,\mu)}{\partial(-p,V,\mu)} = \frac{\partial(-p,V,T)}{\partial(\mu,N,T)} = \frac{\partial(\mu,N,-p)}{\partial(T,S,-p)} = -1 \quad . \tag{2.506}$$

If we are careful, then the results in eq. 2.504 can be quite handy, especially when used in conjunction with eqn. 2.475. For example, we have

$$\left(\frac{\partial S}{\partial V}\right)_{T,N} = \frac{\partial(T,S,N)}{\partial(T,V,N)} = \overbrace{\frac{\partial(T,S,N)}{\partial(p,V,N)}}^{=1} \cdot \frac{\partial(p,V,N)}{\partial(T,V,N)} = \left(\frac{\partial p}{\partial T}\right)_{V,N} \quad , \tag{2.507}$$

which is one of the Maxwell relations derived from the exactness of $dF(T,V,N)$. Some other examples include

$$\left(\frac{\partial V}{\partial S}\right)_{p,N} = \frac{\partial(V,p,N)}{\partial(S,p,N)} = \overbrace{\frac{\partial(V,p,N)}{\partial(S,T,N)}}^{=1} \cdot \frac{\partial(S,T,N)}{\partial(S,p,N)} = \left(\frac{\partial T}{\partial p}\right)_{S,N} \tag{2.508}$$

$$\left(\frac{\partial S}{\partial N}\right)_{T,p} = \frac{\partial(S,T,p)}{\partial(N,T,p)} = \overbrace{\frac{\partial(S,T,p)}{\partial(\mu,N,p)}}^{=1} \cdot \frac{\partial(\mu,N,p)}{\partial(N,T,p)} = -\left(\frac{\partial \mu}{\partial T}\right)_{p,N} \quad , \tag{2.509}$$

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

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = -\frac{\partial(y,x,z)}{\partial(u,v,w)} = \frac{\partial(y,x,z)}{\partial(w,v,u)} = \ldots \quad . \tag{2.510}$$

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

$$\frac{\partial(T,S,N)}{\partial(\bullet,\bullet,N)} = \frac{\partial(T,S,N)}{\partial(p,V,N)}\,\frac{\partial(p,V,N)}{\partial(\bullet,\bullet,N)} = \frac{\partial(p,V,N)}{\partial(\bullet,\bullet,N)} \quad , \tag{2.511}$$

where each $\bullet$ is a distinct arbitrary state variable other than N.

If our Jacobian involves S, V, and N, we write

$$\frac{\partial(S,V,N)}{\partial(\bullet,\bullet,N)} = \frac{\partial(S,V,N)}{\partial(T,V,N)} \cdot \frac{\partial(T,V,N)}{\partial(\bullet,\bullet,N)} = \frac{C_V}{T} \cdot \frac{\partial(T,V,N)}{\partial(\bullet,\bullet,N)} \quad . \tag{2.512}$$

If our Jacobian involves S, p, and N, we write

$$\frac{\partial(S,p,N)}{\partial(\bullet,\bullet,N)} = \frac{\partial(S,p,N)}{\partial(T,p,N)} \cdot \frac{\partial(T,p,N)}{\partial(\bullet,\bullet,N)} = \frac{C_p}{T} \cdot \frac{\partial(T,p,N)}{\partial(\bullet,\bullet,N)} \quad . \tag{2.513}$$

For example,

$$\left(\frac{\partial T}{\partial p}\right)_{S,N} = \frac{\partial(T,S,N)}{\partial(p,S,N)} = \overbrace{\frac{\partial(T,S,N)}{\partial(p,V,N)}}^{=1} \cdot \frac{\partial(p,V,N)}{\partial(p,T,N)} \cdot \frac{\partial(p,T,N)}{\partial(p,S,N)} = \frac{T}{C_p}\left(\frac{\partial V}{\partial T}\right)_{p,N} \tag{2.514}$$

$$\left(\frac{\partial V}{\partial p}\right)_{S,N} = \frac{\partial(V,S,N)}{\partial(p,S,N)} = \frac{\partial(V,S,N)}{\partial(V,T,N)} \cdot \frac{\partial(V,T,N)}{\partial(p,T,N)} \cdot \frac{\partial(p,T,N)}{\partial(p,S,N)} = \frac{C_V}{C_p}\left(\frac{\partial V}{\partial p}\right)_{T,N} \quad . \tag{2.515}$$

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. 2.261.

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 Modeling 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, $N_{\mathrm{A}} = 6.02 \times 10^{23}$. 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 τ, called the collision time, and over some microscopic distance scale, ℓ, 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 we can write this equation as

$$\frac{dP_i}{dt} = -\sum_j \varGamma_{ij}\,P_j \quad , \tag{3.2}$$

where

$$\varGamma_{ij} = \begin{cases} -W_{ij} & \text{if } i \neq j \\ \sum_k' W_{kj} & \text{if } i = j \quad , \end{cases} \tag{3.3}$$

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 may take $W_{ii} \equiv 0$ (no sum on i). Fermi's Golden Rule of quantum mechanics says that

$$W_{ij} = \frac{2\pi}{\hbar}\,\big|\langle\, i \,|\, \hat{V} \,|\, j \,\rangle\big|^2\,\varrho(E_j) \quad , \tag{3.4}$$

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 Equilibrium distribution and detailed balance

If the transition rates W_{ij} are themselves time-independent, then we may formally write

$$P_i(t) = \big(e^{-\Gamma t}\big)_{ij}\, P_j(0) \quad . \tag{3.5}$$

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:

$$\frac{d}{dt}\sum_i P_i = -\sum_{i,j} \Gamma_{ij}\, P_j = -\sum_j \bigg(P_j \sum_i \Gamma_{ij}\bigg) = 0 \quad . \tag{3.6}$$

We conclude that $\vec{L} = (1, 1, \ldots, 1)$ is a left eigenvector of Γ with eigenvalue $\lambda = 0$. The corresponding right eigenvector, which we write as $P_i^{\rm eq}$, satisfies $\sum_j \Gamma_{ij} P_j^{\rm eq} = 0$, and is a stationary (*i.e.* time independent) solution to the master equation. In order for probability to be conserved, it must be that Re $\lambda_a \geq 0$ for each eigenvalue λ_a. Generally, there is only one right/left eigenvector pair corresponding to $\lambda = 0$, in which case any initial probability distribution $P_i(0)$ converges to $P_i^{\rm eq}$ as $t \to \infty$, as shown in the appendix §3.7. Note, however, that since the matrix Γ is in general not symmetric, its eigenvectors may not span, which is to say that it may contain nontrivial Jordan blocks when it is brought to canonical form. In the generic case, we have

$$\Gamma = \sum_{a=1}^{N} \lambda_a \,|\, R^a \,\rangle\langle\, L^a \,| \qquad \Rightarrow \qquad \Gamma_{ij} = \sum_{a=1}^{N} \lambda_a\, R_i^a\, L_j^a \quad , \tag{3.7}$$

with $\langle L^a | R^b \rangle = \delta^{ab}$. See the appendix in §3.6 below for a complete discussion.

In equilibrium, the net rate of transitions into a state $|\, i \,\rangle$ is equal to the rate of transitions out of $|\, i \,\rangle$. If, for each state $|\, j \,\rangle$ the transition rate from $|\, i \,\rangle$ to $|\, j \,\rangle$ is equal to the transition rate from $|\, j \,\rangle$ to $|\, i \,\rangle$, we say that the rates satisfy the condition of *detailed balance*. In other words, $W_{ij}\, P_j^{\rm eq} = W_{ji}\, P_i^{\rm eq}$. Assuming $W_{ij} \neq 0$ and $P_j^{\rm eq} \neq 0$, we can divide to obtain

$$\frac{W_{ji}}{W_{ij}} = \frac{P_j^{\rm eq}}{P_i^{\rm eq}} \quad . \tag{3.8}$$

Note that detailed balance is a stronger condition than that required for a stationary solution to the master equation. When detailed balance holds, then $\Gamma_{ji}\, \pi_i = \Gamma_{ij}\, \pi_j \ \forall \ i, j$, where $\pi_i \equiv P_i^{\rm eq}$. But then we have

$$\Delta_{ij} \equiv \pi_i^{-1/2}\, \Gamma_{ij}\, \pi_j^{1/2} = \Delta_{ji} \quad . \tag{3.9}$$

The real symmetric matrix Δ has real eigenvalues λ_a and a complete and orthonormal set of left/right eigenvectors ϕ_i^a. 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 $\sum_i \Gamma_{ij} = 0$, we have that $\pi_i^{1/2}$ is a left eigenvector of Δ with eigenvalue zero. Therefore $R_j^{(1)} = \pi_j$, which is the equilibrium distribution.

If $\Gamma = \Gamma^{\rm T}$ 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 Γ. The system then satisfies the conditions of detailed balance. See §3.8 below for an example of this formalism applied to a model of radioactive decay.

3.2.4 Boltzmann's $\mathcal{H}$-theorem

Suppose for the moment that Γ is a symmetric matrix, *i.e.* $\Gamma_{ij} = \Gamma_{ji}$. Then construct the function

$$\mathcal{H}(t) = \sum_i P_i(t) \, \ln P_i(t) \quad . \tag{3.10}$$

Then

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= \sum_i \frac{dP_i}{dt} \left(1 + \ln P_i\right) = \sum_i \frac{dP_i}{dt} \, \ln P_i \\ &= -\sum_{i,j} \Gamma_{ij} \, P_j \ln P_i = \sum_{i,j} \Gamma_{ij} \, P_j \big(\ln P_j - \ln P_i\big) \quad , \end{aligned} \tag{3.11}$$

where we have used $\sum_i \Gamma_{ij} = 0$. Now switch $i \leftrightarrow j$ in the above sum and add the terms to get

$$\frac{d\mathcal{H}}{dt} = \frac{1}{2} \sum_{i,j} \Gamma_{ij} \big(P_i - P_j\big) \left(\ln P_i - \ln P_j\right) \quad . \tag{3.12}$$

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

$$(x - y) \left(\ln x - \ln y\right) \geq 0 \quad , \tag{3.13}$$

we conclude

$$\frac{d\mathcal{H}}{dt} \leq 0 \quad . \tag{3.14}$$

In equilibrium, P_i^{eq} is a constant, independent of i. We write

$$P_i^{\text{eq}} = \frac{1}{\Omega} \quad , \quad \Omega = \sum_i 1 \quad \Longrightarrow \quad \mathcal{H} = -\ln \Omega \quad , \tag{3.15}$$

with $\Omega = N$, the matrix dimension of Γ.

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

$$\overline{W}_{ij} \equiv W_{ij} \, P_j^{\text{eq}} = W_{ji} \, P_i^{\text{eq}} = \overline{W}_{ji} \quad , \tag{3.16}$$

and the generalized $\mathcal{H}$-function,

$$\mathcal{H}(t) \equiv \sum_i P_i(t) \, \ln\!\left(\frac{P_i(t)}{P_i^{\text{eq}}}\right) \quad . \tag{3.17}$$

Then

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= -\frac{1}{2} \sum_{i,j} \big(W_{ji} P_i - W_{ij} P_j\big) \, \ln\!\left(\frac{W_{ji} P_i}{W_{ij} P_j}\right) \\ &= -\frac{1}{2} \sum_{i,j} \overline{W}_{ij} \left(\frac{P_i}{P_i^{\text{eq}}} - \frac{P_j}{P_j^{\text{eq}}}\right) \left[\ln\!\left(\frac{P_i}{P_i^{\text{eq}}}\right) - \ln\!\left(\frac{P_j}{P_j^{\text{eq}}}\right)\right] \leq 0 \quad . \end{aligned} \tag{3.18}$$

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

$$P_i(t+1) = \sum_j Q_{ij}\, P_j(t) \quad . \tag{3.19}$$

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. Interating eqn. 3.19, we have

$$P_i(t+n) = \big[Q^n\big]_{ij}\, P_j(t) \quad . \tag{3.20}$$

The transition matrix Q is in general non-normal, meaning $Q \neq Q^{\intercal}$, which means that its 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

$$Q = \sum_{a=1}^{n} \lambda_a \,|\, R^a \,\rangle\langle\, L^a \,| \quad , \tag{3.21}$$

with $|\mathsf{Re}\,\lambda_a| \leq 1$ 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 §3.6. From the condition $\sum_i Q_{ij} = 1$, we have that $\vec{L}^{(1)} = (1,1,\ldots,1)$ is a left eigenvector corresponding to eigenvalue $\lambda = 1$, and the corresponding right eigenvector $|\, R^{(1)} \,\rangle$ is the asymptotic steady state, with $P^{\rm eq}_j = \langle\, j \,|\, R^1 \,\rangle = R^a_j$. Since $Q \in \mathsf{GL}(N,\mathbb{R})$, where N is the matrix dimension (which we take here to be finite), if $\lambda \in \mathbb{C}$ is an eigenvalue, then λ^* 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 4×4 matrix.

Solution: There are four states available: Now 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 $Q_{21} = 1$

$\lvert j\,\rangle$	room A	room B	$g^{\rm A}_j$	$g^{\rm B}_j$	$g^{\rm TOT}_j$
$\lvert 1\,\rangle$	TTT	EEEE	1	1	1
$\lvert 2\,\rangle$	TTE	EEET	3	4	12
$\lvert 3\,\rangle$	TEE	EETT	3	6	18
$\lvert 4\,\rangle$	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 $\lvert 2\,\rangle$, the nonzero elements are

$$Q_{12} = \tfrac{1}{4} \times \tfrac{1}{3} = \tfrac{1}{12} \quad , \quad Q_{22} = \tfrac{3}{4} \times \tfrac{1}{3} + \tfrac{1}{4} \times \tfrac{2}{3} = \tfrac{5}{12} \quad , \quad Q_{32} = \tfrac{1}{2} \quad . \tag{3.22}$$

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 $\lvert 3\,\rangle$, 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 $\lvert 4\,\rangle$, the nonzero elements are $Q_{34} = \frac{3}{4}$ and $Q_{44} = \frac{1}{4}$. The full transition matrix is then

$$Q = \begin{pmatrix} 0 & \frac{1}{12} & 0 & 0 \\ 1 & \frac{5}{12} & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{3}{4} \\ 0 & 0 & \frac{1}{6} & \frac{1}{4} \end{pmatrix} \quad . \tag{3.23}$$

Note that $\sum_i Q_{ij} = 1$ for all $j = 1,2,3,4$. This guarantees that $L^{(1)} = (1\,,\,1\,,\,1\,,\,1)$ is a left eigenvector of Q with eigenvalue 1. The corresponding right eigenvector is obtained by setting $Q_{ij}\,\psi^{(1)}_j = \psi^{(1)}_i$. Simultaneously solving these four equations and normalizing so that $\sum_j \psi^{(1)}_j = 1$, we easily obtain

$$\psi^{(1)} = \frac{1}{35} \begin{pmatrix} 1 \\ 12 \\ 18 \\ 4 \end{pmatrix} \quad . \tag{3.24}$$

This is the state we converge to after repeated application of the transition matrix Q. If we decompose $Q = \sum_{\alpha=1}^{4} \lambda_\alpha \,\lvert\, \psi^{(\alpha)}\,\rangle\langle\, L^{(\alpha)}\,\rvert$, then in the limit $t \to \infty$ we have $Q^t \approx \lvert\, \psi^{(1)}\,\rangle\langle\, L^{(1)}\,\rvert$, where $\lambda_1 = 1$, since the remaining eigenvalues are all less than 1 in magnitude[2]. Thus, Q^t acts as a *projector* onto the state $\lvert\, \psi^{(1)}\,\rangle$. 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, $\lim_{t\to\infty} P_j(t) = \psi^{(1)}_j$, and we find $P_3(\infty) = \frac{18}{35}$. Note that the equilibrium distribution satisfies detailed balance:

$$\psi^{(1)}_j = \frac{g^{\rm TOT}_j}{\sum_l g^{\rm TOT}_l} \quad . \tag{3.25}$$

---

[2] 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

$$\dot{p}_\sigma = \frac{d}{dt}\bigg(\frac{\partial L}{\partial \dot{q}_\sigma}\bigg) = \frac{\partial L}{\partial q_\sigma} \quad , \tag{3.26}$$

where p_σ is the canonical momentum conjugate to the generalized coordinate q_σ, *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,

$$H(q,p) = \sum_{\sigma=1}^N p_\sigma\,\dot{q}_\sigma - L \quad . \tag{3.27}$$

Note that

$$\begin{aligned} dH &= \sum_{\sigma=1}^N \bigg(p_\sigma\,d\dot{q}_\sigma + \dot{q}_\sigma\,dp_\sigma - \frac{\partial L}{\partial q_\sigma}\,dq_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma}\,d\dot{q}_\sigma\bigg) - \frac{\partial L}{\partial t}\,dt \\ &= \sum_{\sigma=1}^N \bigg(\dot{q}_\sigma\,dp_\sigma - \frac{\partial L}{\partial q_\sigma}\,dq_\sigma\bigg) - \frac{\partial L}{\partial t}\,dt \quad . \end{aligned} \tag{3.28}$$

Thus, we obtain Hamilton's equations of motion,

$$\frac{\partial H}{\partial p_\sigma} = \dot{q}_\sigma \quad , \quad \frac{\partial H}{\partial q_\sigma} = -\frac{\partial L}{\partial q_\sigma} = -\dot{p}_\sigma \tag{3.29}$$

and

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \quad . \tag{3.30}$$

Define the rank $2N$ vector $\boldsymbol{\varphi}$ by its components,

$$\varphi_i = \begin{cases} q_i & \text{if } 1 \le i \le N \\ p_{i-N} & \text{if } N \le i \le 2N \end{cases} \quad . \tag{3.31}$$

Then we may write Hamilton's equations compactly as

$$\dot{\varphi}_i = J_{ij}\,\frac{\partial H}{\partial \varphi_j} \quad , \tag{3.32}$$

where

$$J = \begin{pmatrix} 0_{N\times N} & 1_{N\times N} \\ -1_{N\times N} & 0_{N\times N} \end{pmatrix} \tag{3.33}$$

is a rank $2N$ 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

$$\begin{aligned} \frac{dF}{dt} &= \frac{\partial F}{\partial t} + \sum_{\sigma=1}^{N} \bigg(\frac{\partial F}{\partial q_\sigma} \frac{dq_\sigma}{dt} + \frac{\partial F}{\partial p_\sigma} \frac{dp_\sigma}{dt} \bigg) \\ &= \frac{\partial F}{\partial t} + \big\{ F, H \big\} \quad , \end{aligned} \tag{3.34}$$

where $\big\{ \bullet, \bullet \big\}$ is the *Poisson bracket*,

$$\big\{ A, B \big\} = \sum_{\sigma=1}^{N} \bigg(\frac{\partial A}{\partial q_\sigma} \frac{\partial B}{\partial p_\sigma} - \frac{\partial A}{\partial p_\sigma} \frac{\partial B}{\partial q_\sigma} \bigg) \quad . \tag{3.35}$$

3.3.2 Dynamical systems and the evolution of phase space volumes

Consider a general dynamical system,

$$\frac{d\boldsymbol{\varphi}}{dt} = \boldsymbol{V}(\boldsymbol{\varphi}) \quad , \tag{3.36}$$

where $\boldsymbol{\varphi}(t)$ is a point in an n-dimensional phase space. Consider now a compact[3] 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 $t = 0$:

$$\mathcal{R}(t) = \big\{ \boldsymbol{\varphi}(t) \,\big|\, \boldsymbol{\varphi}(0) \in \mathcal{R}_0 \big\} \quad . \tag{3.37}$$

Now consider the volume $\Omega(t)$ of the set $\mathcal{R}(t)$. We have

$$\Omega(t) = \int\limits_{\mathcal{R}(t)} \! d\mu \tag{3.38}$$

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

$$\Omega(t + dt) = \int\limits_{\mathcal{R}(t+dt)} \! d\mu' = \int\limits_{\mathcal{R}(t)} \! d\mu \, \left| \frac{\partial \varphi_i(t+dt)}{\partial \varphi_j(t)} \right| \quad , \tag{3.39}$$

where

$$\left| \frac{\partial \varphi_i(t+dt)}{\partial \varphi_j(t)} \right| \equiv \frac{\partial(\varphi'_1, \ldots, \varphi'_n)}{\partial(\varphi_1, \ldots, \varphi_n)} \tag{3.40}$$

---

[3] 'Compact' in the parlance of mathematical analysis means 'closed and bounded'.

is a determinant, which is the Jacobean 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

$$\varphi_i(t + dt) = \varphi_i(t) + V_i\big(\boldsymbol{\varphi}(t)\big)\, dt + \mathcal{O}(dt^2) \tag{3.41}$$

and therefore

$$\frac{\partial\varphi_i(t + dt)}{\partial\varphi_j(t)} = \delta_{ij} + \frac{\partial V_i}{\partial\varphi_j}\, dt + \mathcal{O}(dt^2) \quad . \tag{3.42}$$

We now make use of the matrix equality $\ln \mathsf{det}\, M = \mathsf{Tr}\, \ln M$, which gives us[4], for small ε,

$$\mathsf{det}\big(1 + \varepsilon A\big) = \exp \mathsf{Tr}\, \ln\big(1 + \varepsilon A\big) = 1 + \varepsilon\ \mathsf{Tr}\, A + \tfrac{1}{2}\,\varepsilon^2\Big(\big(\mathsf{Tr}\, A\big)^2 - \mathsf{Tr}\,(A^2)\Big) + \ldots \tag{3.43}$$

Thus,

$$\Omega(t + dt) = \Omega(t) + \int\limits_{\mathcal{R}(t)} d\mu\ \boldsymbol{\nabla}\!\cdot\! \boldsymbol{V}\ dt + \mathcal{O}(dt^2) \quad , \tag{3.44}$$

which says

$$\frac{d\Omega}{dt} = \int\limits_{\mathcal{R}(t)} d\mu\ \boldsymbol{\nabla}\!\cdot\! \boldsymbol{V} = \int\limits_{\partial\mathcal{R}(t)} dS\ \hat{\boldsymbol{n}} \cdot \boldsymbol{V} \tag{3.45}$$

Here, the divergence is the *phase space divergence*,

$$\boldsymbol{\nabla}\!\cdot\! \boldsymbol{V} = \sum_{i=1}^{n} \frac{\partial V_i}{\partial\varphi_i} \quad , \tag{3.46}$$

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,

$$\frac{\partial\varrho}{\partial t} + \boldsymbol{\nabla}\!\cdot\!(\varrho \boldsymbol{V}) = 0 \quad . \tag{3.47}$$

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

$$\frac{d}{dt}\int\limits_{\mathcal{R}} d\mu\, \varrho = -\int\limits_{\mathcal{R}} d\mu\ \boldsymbol{\nabla}\!\cdot\!(\varrho \boldsymbol{V}) = -\int\limits_{\partial\mathcal{R}} dS\ \hat{\boldsymbol{n}} \cdot (\varrho \boldsymbol{V}) \quad . \tag{3.48}$$

It is perhaps helpful to think of ϱ as a charge density, in which case $\boldsymbol{J} = \varrho \boldsymbol{V}$ is the current density. The above equation then says

$$\frac{dQ_{\mathcal{R}}}{dt} = -\int\limits_{\partial\mathcal{R}} dS\ \hat{\boldsymbol{n}} \cdot \boldsymbol{J} \quad , \tag{3.49}$$

---

[4] The equality $\ln \mathsf{det}\, M = \mathsf{Tr}\, \ln M$ is most easily proven by bringing the matrix to diagonal form via a similarity transformation, and proving the equality for diagonal matrices.

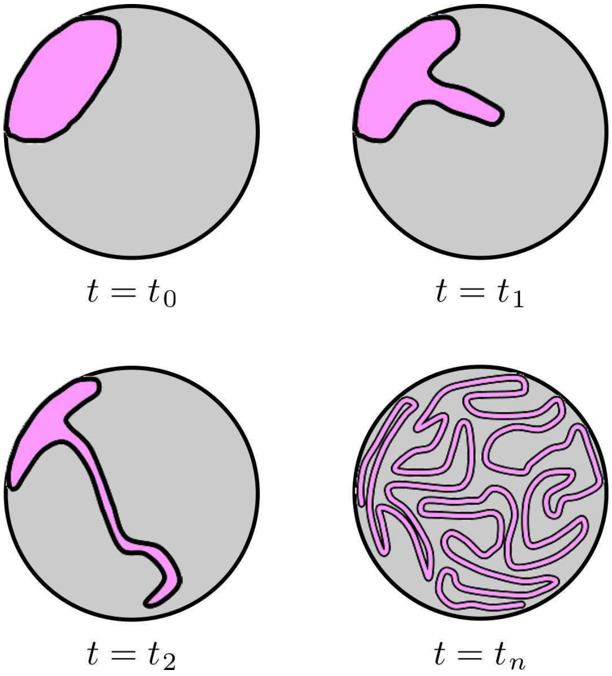


Figure 3.1: Time evolution of two immiscible fluids. The local density remains constant.

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

$$\frac{\partial \varrho}{\partial t} + \boldsymbol{V} \cdot \boldsymbol{\nabla} \varrho + \varrho \, \boldsymbol{\nabla} \cdot \boldsymbol{V} = 0 \quad . \tag{3.50}$$

But now suppose that the phase flow is divergenceless, *i.e.* $\boldsymbol{\nabla} \cdot \boldsymbol{V} = 0$. Then we have

$$\frac{D\varrho}{Dt} \equiv \left(\frac{\partial}{\partial t} + \boldsymbol{V} \cdot \boldsymbol{\nabla} \right) \varrho = 0 \quad . \tag{3.51}$$

The combination inside the brackets above is known as the *convective derivative*. It tells us the total rate of change of ϱ for an observer *moving with the phase flow*. That is

$$\begin{aligned} \frac{d}{dt} \, \varrho\big(\boldsymbol{\varphi}(t), t\big) &= \frac{\partial \varrho}{\partial \varphi_i} \, \frac{d\varphi_i}{dt} + \frac{\partial \varrho}{\partial t} \\ &= \sum_{i=1}^{n} V_i \, \frac{\partial \varrho}{\partial \varphi_i} + \frac{\partial \varrho}{\partial t} = \frac{D\varrho}{Dt} \quad . \end{aligned} \tag{3.52}$$

If $D\varrho/Dt = 0$, the local density remains the same during the evolution of the system. If we consider the 'characteristic function'

$$\varrho(\boldsymbol{\varphi}, t = 0) = \begin{cases} 1 & \text{if } \boldsymbol{\varphi} \in \mathcal{R}_0 \\ 0 & \text{otherwise} \end{cases} \tag{3.53}$$

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

$$\dot{q}_\sigma = +\frac{\partial H}{\partial p_\sigma} \quad , \qquad \dot{p}_\sigma = -\frac{\partial H}{\partial q_\sigma} \tag{3.54}$$

A point in phase space is specified by N positions q_σ and N momenta p_σ, hence the dimension of phase space is $n = 2N$:

$$\boldsymbol{\varphi} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} \quad , \qquad \boldsymbol{V} = \begin{pmatrix} \dot{\boldsymbol{q}} \\ \dot{\boldsymbol{p}} \end{pmatrix} = \begin{pmatrix} +\partial H/\partial \boldsymbol{p} \\ -\partial H/\partial \boldsymbol{q} \end{pmatrix} \quad . \tag{3.55}$$

Hamilton's equations of motion guarantee that the phase space flow is divergenceless:

$$\begin{aligned} \nabla\!\cdot\!\boldsymbol{V} &= \sum_{\sigma=1}^{N} \left\{ \frac{\partial \dot{q}_\sigma}{\partial q_\sigma} + \frac{\partial \dot{p}\sigma}{\partial p_\sigma} \right\} \\ &= \sum_{\sigma=1}^{N} \left\{ \frac{\partial}{\partial q_\sigma}\left(\frac{\partial H}{\partial p_\sigma}\right) + \frac{\partial}{\partial p_\sigma}\left(-\frac{\partial H}{\partial q_\sigma}\right) \right\} = 0 \quad . \end{aligned} \tag{3.56}$$

Thus, we have that the convective derivative vanishes, *viz.*

$$\frac{D\varrho}{Dt} \equiv \frac{\partial \varrho}{\partial t} + \boldsymbol{V}\!\cdot\!\nabla\varrho = 0 \quad , \tag{3.57}$$

for any distribution $\varrho(\boldsymbol{\varphi}, t)$ on phase space. 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

$$\frac{\partial \varrho}{\partial t} = -\dot{\boldsymbol{\varphi}} \cdot \boldsymbol{\nabla}\varrho = -\big\{\varrho, H\big\} \quad . \tag{3.58}$$

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*:

$$i\hat{L} = \sum_{\sigma=1}^{N} \left\{ \frac{\partial H}{\partial p_\sigma}\frac{\partial}{\partial q_\sigma} - \frac{\partial H}{\partial q_\sigma}\frac{\partial}{\partial p_\sigma} \right\} \quad . \tag{3.59}$$

Eqn. 3.58, 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

$$\frac{\partial \hat{\varrho}}{\partial t} = \frac{i}{\hbar}\left[\,\hat{\varrho}, \hat{H}\,\right] \quad , \tag{3.60}$$

whence we infer the correspondence

$$\{\varrho, H\} \to \frac{1}{i\hbar}\,\big[\,\hat{\varrho}, \hat{H}\,\big] \quad . \tag{3.61}$$

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 Λ_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

$$\dot{\boldsymbol{\varphi}} \cdot \boldsymbol{\nabla}\varrho = \sum_{a=1}^{k} \frac{\partial\varrho}{\partial\Lambda_a}\,\dot{\boldsymbol{\varphi}} \cdot \boldsymbol{\nabla}\Lambda_a = 0 \quad . \tag{3.62}$$

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,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{D(E)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{\int\! d\mu\,\delta\big(E - H(\boldsymbol{\varphi})\big)} \quad , \tag{3.63}$$

is a fixed point solution of Liouville's equation. If there were a second conserved quantity, $\Lambda(\boldsymbol{\varphi})$, the generalized microcanonical distribution,

$$\varrho_{E,Q}(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)}{D(E,Q)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)}{\int\! d\mu\,\delta\big(E - H(\boldsymbol{\varphi})\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)} \quad , \tag{3.64}$$

would be a solution to Liouville's equation for arbitrary E and Q.

Similarly, the Gibbs distribution,

$$\varrho_\beta(\boldsymbol{\varphi}) = \frac{1}{Z(\beta)}\,e^{-\beta H(\boldsymbol{\varphi})} \quad , \tag{3.65}$$

where $Z(\beta) = \mathsf{Tr}\ e^{-\beta H(\boldsymbol{\varphi})}$ is the partition function, satisfies $\big\{\varrho_\beta, H\big\} = 0$. In the presence of multiple conserved quantities, one defines the *generalized Gibbs distribution*,

$$\varrho_{\boldsymbol{\beta}}(\boldsymbol{\varphi}) = \frac{1}{Z(\boldsymbol{\beta})}\,\exp\!\bigg(-\sum_{a=1}^{k} \beta_a\,\Lambda_a(\boldsymbol{\varphi})\bigg) \quad , \tag{3.66}$$

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: $dH/dt \le 0$, where H is Boltzmann's H-function. However, the microscopic laws of physics are (almost) time-reversal invariant[5], so how can we understand the emergence of irreversibility? Furthermore, any dynamics which are deterministic and volume-preserving in a finite phase space exhibits 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_τ be the 'τ-advance mapping' which evolves points in phase space according to Hamilton's equations. Assume that g_τ 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.*

$$\int d\mu\, \delta\big(E - H(\boldsymbol{q},\boldsymbol{p})\big) < \infty \quad , \tag{3.67}$$

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_τ, where m is finite.

Proof: Assume the theorem fails; we will show this assumption results in a contradiction. Consider the set Υ formed from the union of all sets $g_\tau^k\, \mathcal{R}$ for all m:

$$\Upsilon = \bigcup_{k=0}^{\infty} g_\tau^k\, \mathcal{R}_0 \tag{3.68}$$

We assume that the set $\{g_\tau^k\, \mathcal{R}_0 \,|\, k \in \mathbb{N}_0\}$ is disjoint[6]. The volume of a union of disjoint sets is the sum of the individual volumes. Thus,

$$\text{vol}(\Upsilon) = \sum_{k=0}^{\infty} \text{vol}\big(g_\tau^k\, \mathcal{R}_0\big) = \text{vol}(\mathcal{R}_0) \cdot \sum_{k=0}^{\infty} 1 = \infty \quad , \tag{3.69}$$

since $\text{vol}\big(g_\tau^k\, \mathcal{R}_0\big) = \text{vol}\big(\mathcal{R}_0\big)$ from volume preservation. But clearly Υ 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 $k < l$. Apply the inverse g_τ^{-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

---

[5] Actually, the microscopic laws of physics are *not* time-reversal invariant, but rather are invariant under the product PCT, where P is parity, C is charge conjugation, and T is time reversal.

[6] The natural numbers $\mathbb{N}_0$ is the set of non-negative integers $\{0, 1, 2, \ldots\}$.

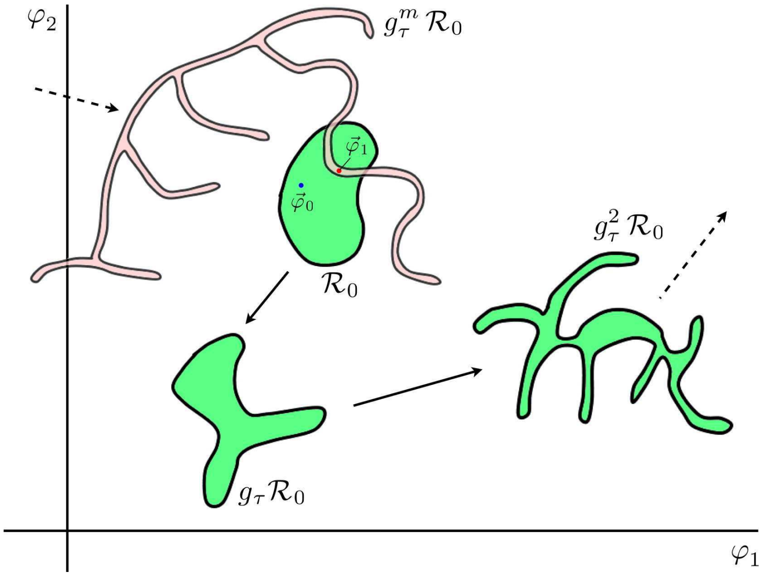


Figure 3.2: Successive images of a set $\mathcal{R}_0$ under the τ-advance mapping g_τ, projected onto a two-dimensional phase plane. The Poincaré recurrence theorem guarantees that if phase space has finite volume, and g_τ is invertible and volume preserving, then for any set $\mathcal{R}_0$ there exists an integer m such that $\mathcal{R}_0 \cap g_\tau^m\,\mathcal{R}_0 \neq \emptyset$.

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 $T(\boldsymbol{p}) > E_{\text{TOT}}$, where $T(\boldsymbol{p})$ is the single particle kinetic energy function[7]. 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![8] In this case, we could define the region $\mathcal{R}_0$ as

$$\mathcal{R}_0 = \big\{(q_1,\ldots,q_r,p_1,\ldots,p_r) \;\big|\; |q_i - q_i^0| \le \Delta q \text{ and } |p_j - p_j^0| \le \Delta p \;\forall\, i,j\big\} \quad , \tag{3.70}$$

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 $f : \mathbb{R} \to \mathbb{R}$ 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 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.

---

[7] In the nonrelativistic limit, $T = p^2/2m$. For relativistic particles, we have $T = (p^2c^2 + m^2c^4)^{1/2} - mc^2$.

[8] 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.

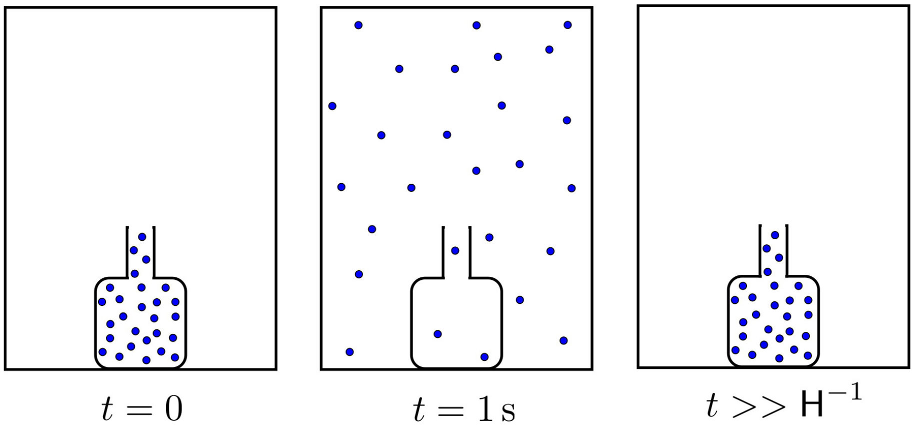


Figure 3.3: Poincaré recurrence guarantees that if we remove the cap from a bottle of perfume in an otherwise evacuated room, all the perfume molecules will eventually return to the bottle! (Here H is the Hubble constant.)

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: $\Omega(t) = e^{-2\beta 2}\,\Omega(0)$, and are not preserved by the dynamics. The proof of recurrence therefore fails. In this case, it is possible for the set Υ 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 θ_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 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.

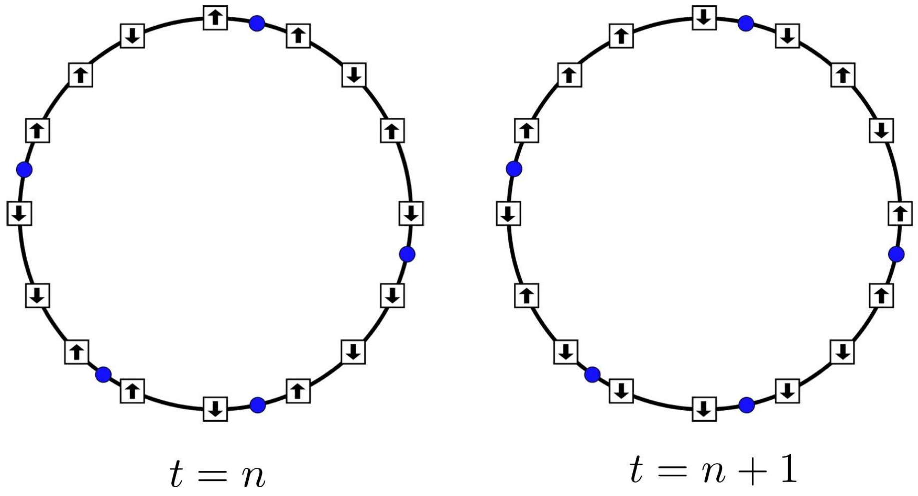


Figure 3.4: Left: A configuration of the Kac ring with $N = 16$ sites and $F = 4$ flippers. The flippers, which live on the links, are represented by blue dots. Right: The ring system after one time step. Evolution proceeds by clockwise rotation. Spins passing through flippers are flipped.

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

$$p_{n+1} = (1-x)\,p_n + x\,(1-p_n) \quad , \tag{3.71}$$

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 3.71 can be solved immediately:

$$p_n = \tfrac{1}{2} + (1-2x)^n\,(p_0 - \tfrac{1}{2}) \quad , \tag{3.72}$$

which decays exponentially to the equilibrium value of $p_{\text{eq}} = \frac{1}{2}$ with time scale

$$\tau(x) = -\frac{1}{\ln|1-2x|} \quad . \tag{3.73}$$

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 $m_{\text{eq}} = 0$. Note that for $\frac{1}{2} < x < 1$ that the magnetization reverses sign each time step, as well as decreasing exponentially in magnitude.

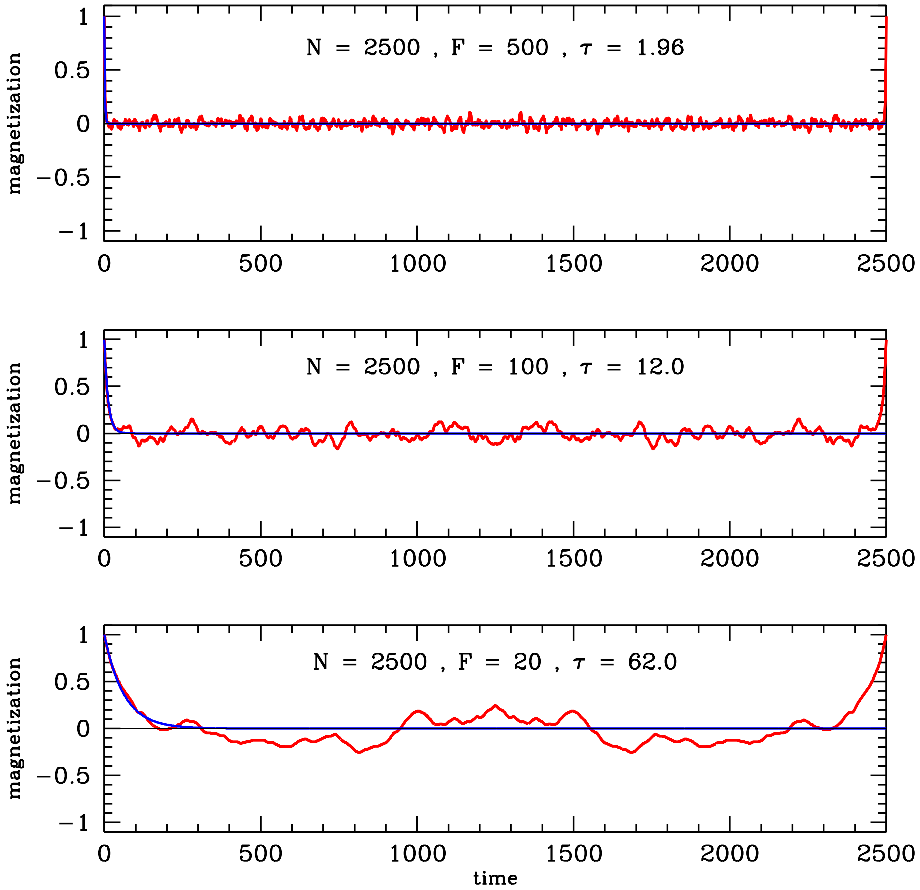


Figure 3.5: Three simulations of the Kac ring model with $N = 2500$ sites and three different concentrations of flippers. The red line shows the magnetization as a function of time, starting from an initial configuration in which 100% of the spins are up. The blue line shows the prediction of the *Stosszahlansatz*, which yields an exponentially decaying magnetization with time constant τ.

The assumption that leads to equation 3.71 is called the *Stosszahlansatz*[9], 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 $N = 1000$ sites, with $F = 100$ and $F = 24$ flippers. Note how the magnetization decays and fluctuates about the equilibrium value $m_{\text{eq}} = 0$, but that after N iterations m recovers its initial value: $m_N = m_0$.

---

[9] Unfortunately, many important physicists were German and we have to put up with a legacy of long German words like *Gedankenexperiment*, *Zitterbewegung*, *Brehmsstrahlung*, *Stosszahlansatz*, *Kartoffelsalat*, *etc.*

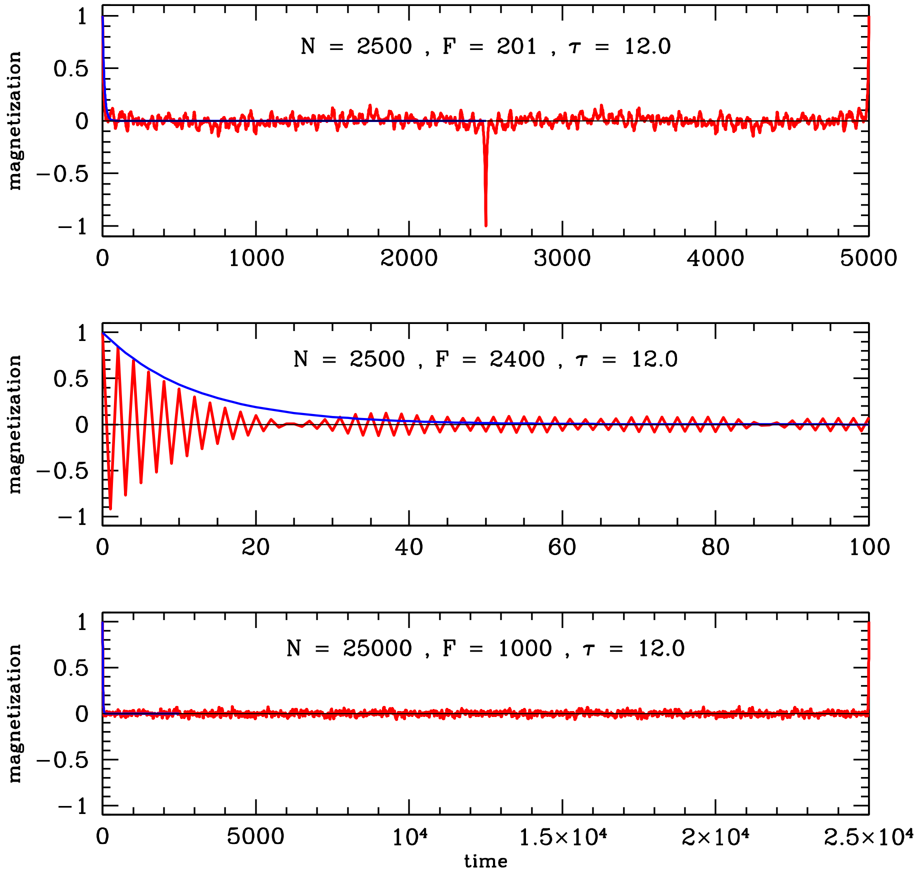


Figure 3.6: Simulations of the Kac ring model. Top: $N = 2500$ sites with $F = 201$ flippers. After 2500 iterations, each spin has flipped an odd number of times, so the recurrence time is $2N$. Middle: $N = 2500$ with $F = 2400$, resulting in a near-complete reversal of the population with every iteration. Bottom: $N = 25000$ with $N = 1000$, showing long time equilibration and dramatic resurgence of the spin population.

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.

In figure 3.6 we plot two other simulations. The top 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 $f(\boldsymbol{\varphi})$ is defined as

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_t = \lim_{T\to\infty} \frac{1}{T}\int\limits_0^T \! dt\, f\big(\boldsymbol{\varphi}(t)\big) \quad . \tag{3.74}$$

For a Hamiltonian system, the *phase space average* of the same function is defined by

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_E = \int\! d\mu\, f(\boldsymbol{\varphi})\, \delta\big(E - H(\boldsymbol{\varphi})\big) \bigg/ \int\! d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\big) \quad , \tag{3.75}$$

where $H(\boldsymbol{\varphi}) = H(\boldsymbol{q},\boldsymbol{p})$ is the Hamiltonian, and where $\delta(x)$ is the Dirac δ-function. The energy is fixed to be $E = H\big(\boldsymbol{\varphi}(t=0)\big)$. Thus,

$$\text{ergodicity} \quad \Longleftrightarrow \quad \big\langle f(\boldsymbol{\varphi})\big\rangle_t = \big\langle f(\boldsymbol{\varphi})\big\rangle_E \quad , \tag{3.76}$$

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 $\mathcal{R}$:

$$\chi_{\mathcal{R}}(\boldsymbol{\varphi}) = \begin{cases} 1 & \text{if } \boldsymbol{\varphi} \in \mathcal{R} \\ 0 & \text{otherwise,} \end{cases} \tag{3.77}$$

where $H(\boldsymbol{\varphi}) = E$ for all $\boldsymbol{\varphi} \in \mathcal{R}$, so $\mathsf{dim}\,\mathcal{R} = 2n-1$, where $n = 2dN$ is the total dimension of phase space for N point particles in d space dimensions. Then

$$\big\langle \chi_{\mathcal{R}}(\boldsymbol{\varphi})\big\rangle_t = \lim_{T\to\infty} \left(\frac{\text{time spent in } \mathcal{R}}{T}\right) \quad . \tag{3.78}$$

If the system is ergodic, then

$$\big\langle \chi_{\mathcal{R}}(\boldsymbol{\varphi})\big\rangle_t = P(\mathcal{R}) = \frac{D_{\mathcal{R}}(E)}{D(E)} \quad , \tag{3.79}$$

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,

$$D(E) = \int\! d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\big) = \int\limits_{\mathcal{S}_E}\! d\Sigma_E \quad , \tag{3.80}$$

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. 3.83. 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 ζ_E is a phase space coordinate locally perpendicular to $\mathcal{S}_E$. We then define[10]

$$d\Sigma_E \equiv \frac{dS_E}{|\boldsymbol{\nabla} H|}\bigg|_{H(\boldsymbol{\varphi})=E} \quad , \tag{3.81}$$

and we may now write $d\mu = dE \, d\Sigma_E$. Note that we may also express $D(E)$ as

$$D(E) = \frac{d}{dE}\int d\mu \, \Theta\big(E - H(\boldsymbol{\varphi})\big) \equiv \frac{d\Omega(E)}{dE} \quad , \tag{3.82}$$

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

$$D_{\mathcal{R}}(E) = \int\limits_{\mathcal{R}} d\Sigma_E \quad . \tag{3.83}$$

Note that $\mathcal{R} \subset \mathcal{S}_E$.

3.5.2 The microcanonical ensemble

The distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{D(E)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{\int d\mu \, \delta\big(E - H(\boldsymbol{\varphi})\big)} \quad , \tag{3.84}$$

defines the *microcanonical ensemble* (μCE) of Gibbs. We could also write

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_E = \frac{1}{D(E)}\int\limits_{\mathcal{S}_E} d\Sigma_E \, f(\boldsymbol{\varphi}) \quad , \tag{3.85}$$

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,

---

[10] 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 A^N, where A stands for action, *i.e.* $[d\mu] = \mathrm{ML}^2/\mathrm{T}$. Thus while the product $d\mu = dS_E \, d\zeta_E$ has units of 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 A^N/E. To resolve any confusion, one may choose to rescale so that all phase space coordinates are dimensionless.

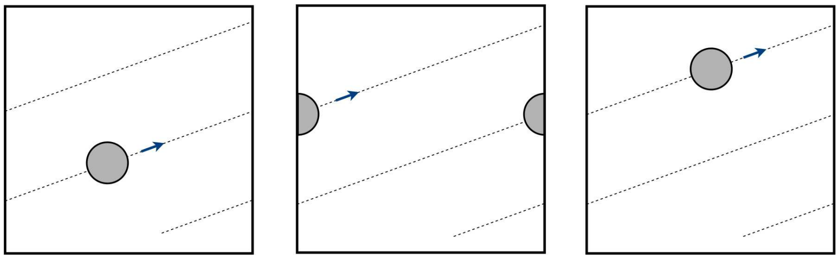


Figure 3.7: Constant phase space velocity at an irrational angle over a toroidal phase space is ergodic, but not mixing. A circle remains a circle, and a blob remains a blob.

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 $p = p_0 + \alpha(q - q_0)$.

Now consider the average of some function $f(q,p)$. We can write $f(q,p)$ in terms of its Fourier transform,

$$f(q,p) = \sum_{m,n} \hat{f}_{m,n}\, e^{2\pi i(mq+np)} \quad . \tag{3.86}$$

We have, then,

$$f\big(q(t),p(t)\big) = \sum_{m,n} \hat{f}_{m,n}\, e^{2\pi i(mq_0+np_0)}\, e^{2\pi i(m+\alpha n)t} \quad . \tag{3.87}$$

We can now perform the time average of f:

$$\begin{aligned} \big\langle f(q,p)\big\rangle_t &= \hat{f}_{0,0} + \lim_{T\to\infty} \frac{1}{T} \sum_{m,n}{}' \hat{f}_{m,n}\, e^{2\pi i(mq_0+np_0)}\, \frac{e^{2\pi i(m+\alpha n)T} - 1}{2\pi i(m+\alpha n)} \\ &= \hat{f}_{0,0} \quad \text{if } \alpha \notin \mathbb{Q} \quad . \end{aligned} \tag{3.88}$$

Clearly,

$$\big\langle f(q,p)\big\rangle_E = \int\limits_0^1 \! dq \int\limits_0^1 \! dp\, f(q,p) = \hat{f}_{0,0} = \big\langle f(q,p)\big\rangle_t \quad , \tag{3.89}$$

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 $\exp\{2\pi i(m+\alpha n)t\} = 1$ whenever $(m,n) = (kr,-ks)$ for any $k \in \mathbb{Z}$. Thus,

$$\big\langle f(q,p)\big\rangle_t = \sum_{k=-\infty}^{\infty} \hat{f}_{kr,-ks}\, e^{2\pi i k(rq_0 - sp_0)} \quad , \tag{3.90}$$

which is not the same as $\big\langle f(q,p)\big\rangle_{\mu\mathrm{CE}} = \hat{f}_{0,0}$.

The situation is depicted in fig. 3.7. If we start with the characteristic function of a disc,

$$\varrho(q,p,t=0) = \Theta\big(a^2 - (q-q_0)^2 - (p-p_0)^2\big) \quad , \tag{3.91}$$

then it remains the characteristic function of a disc:

$$\varrho(q,p,t) = \Theta\big(a^2 - (q-q_0-t)^2 - (p-p_0-\alpha t)^2\big) \quad , \tag{3.92}$$

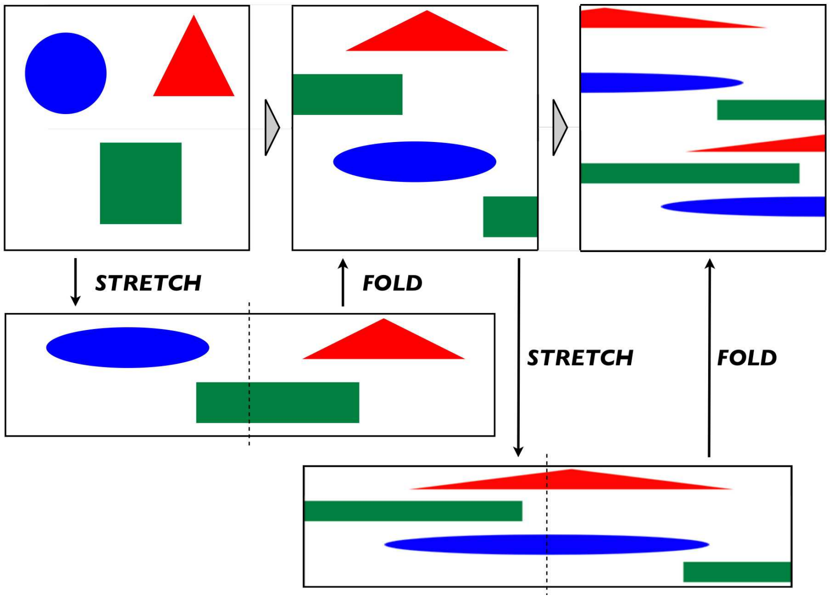


Figure 3.8: The baker's transformation is a successive stretching, cutting, and restacking.

For an example of a transition to ergodicity in a simple dynamical Hamiltonian model, see §3.9.

A stronger condition one could impose is the following. Let A and B be subsets of $\mathcal{S}_E$. Define the *measure*

$$\nu(A) = \frac{\int d\Sigma_E \, \chi_A(\varphi)}{\int d\Sigma_E} = \frac{D_A(E)}{D(E)} \quad , \tag{3.93}$$

where $\chi_A(\varphi)$ 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

$$\text{mixing} \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} \nu\Big(g^n A \cap B\Big) = \nu(A)\,\nu(B) \quad . \tag{3.94}$$

In other words, the fraction of B covered by the n^{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

$$\begin{aligned} \big\langle f(\varphi)\big\rangle &= \int \! d\mu \, \varrho(\varphi,t)\, f(\varphi) \xrightarrow[t\to\infty]{} \int \! d\mu \, f(\varphi)\, \delta\big(E - H(\varphi)\big) \bigg/ \! \int \! d\mu \, \delta\big(E - H(\varphi)\big) \\ &\equiv \mathsf{Tr}\Big[f(\varphi)\,\delta\big(E - H(\varphi)\big)\Big] \Big/ \, \mathsf{Tr}\Big[\delta\big(E - H(\varphi)\big)\Big] \quad . \end{aligned} \tag{3.95}$$

Physically, we can imagine regions of phase space being successively stretched and folded. During the

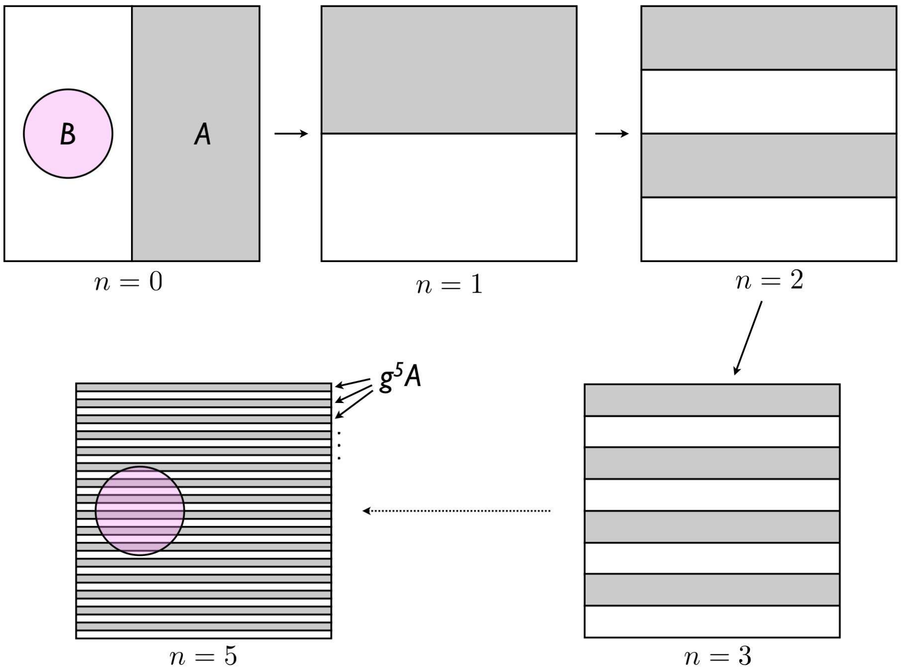


Figure 3.9: The multiply iterated baker's transformation. The set A covers half the phase space and its area is preserved under the map. Initially, the fraction of B covered by A is zero. After many iterations, the fraction of B covered by g^nA approaches $\frac{1}{2}$.

stretching process, the volume is preserved, so the successive stretch and fold operations map phase space back onto itself.

An example of a mixing system is the *baker's transformation*, depicted in fig. 3.8, and defined by

$$g(q,p) = \begin{cases} \big(2q\,,\,\frac{1}{2}p\big) & \text{if} \quad 0 \le q < \frac{1}{2} \\ \big(2q-1\,,\,\frac{1}{2}p+\frac{1}{2}\big) & \text{if} \quad \frac{1}{2} \le q < 1 \quad . \end{cases} \tag{3.96}$$

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,

$$g^{-1}(q,p) = \begin{cases} \big(\frac{1}{2}q\,,\,2p\big) & \text{if} \quad 0 \le p < \frac{1}{2} \\ \big(\frac{1}{2}q+\frac{1}{2}\,,\,2p-1\big) & \text{if} \quad \frac{1}{2} \le p < 1 \quad . \end{cases} \tag{3.97}$$

Another example of a mixing system is Arnold's 'cat map'[11]

$$g(q,p) = \big(\,[q+p]\,,\,[q+2p]\,\big) \quad , \tag{3.98}$$

---

[11] The cat map gets its name from its initial application, by Arnold, to the image of a cat's face.

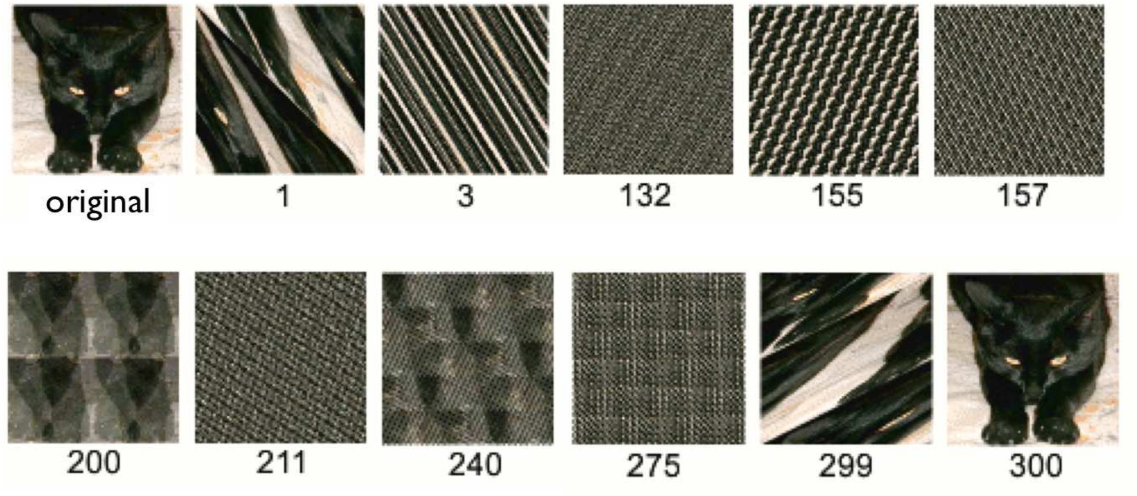


Figure 3.10: The Arnold cat map applied to an image of 150×150 pixels. After 300 iterations, the image repeats itself. (*Source: Wikipedia*)

where $[x]$ denotes the fractional part of x. One can write this in matrix form as

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}}^{M} \begin{pmatrix} q \\ p \end{pmatrix} \bmod \mathbb{Z}^2 \quad . \tag{3.99}$$

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

$$\begin{pmatrix} q \\ p \end{pmatrix} = \overbrace{\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}}^{M^{-1}} \begin{pmatrix} q' \\ p' \end{pmatrix} \bmod \mathbb{Z}^2 \quad . \tag{3.100}$$

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. 3.13, understanding the characteristic features of each successive refinement[12].

3.6 Appendix I : Normal matrices and Jordan canonical form

If a matrix A is normal, meaning $\big[A, A^\dagger\big] = 0$, 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 $\langle\!\langle\, L^\alpha \,\|\, R^\beta \,\rangle\!\rangle = \delta^{\alpha\beta}$, 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 $\mathsf{dim}(A) = n$, 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 $\tilde{A} = Q^{-1} A\, Q$, where Q is invertible and

$$\tilde{A} = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_b \end{pmatrix} \quad . \tag{3.101}$$

Here b is the number of Jordan blocks, where each block J_α is of the form

$$J_\alpha = \begin{pmatrix} \lambda_\alpha & 1 & & \\ & \lambda_\alpha & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_\alpha \end{pmatrix} \quad . \tag{3.102}$$

Thus each J_α is tridiagonal, with diagonal elements all given by λ_α and each element directly above the diagonal equal to one. We denote the right and left eigenvectors of $\tilde{A}$ as $\|\, \widetilde{R}^\alpha \,\rangle\!\rangle = Q^{-1} \,\|\, R^\alpha \,\rangle\!\rangle$ and $\langle\!\langle\, \widetilde{L}^\alpha \,\| = \langle\!\langle\, L^\alpha \,\|\, Q$, respectively. Each J_α has only one right eigenvector, $\psi^\alpha_j = \delta_{j,1}$, whose corresponding left eigenvector is $\chi^\alpha_j = \delta_{j,n_\alpha}$, where $n_\alpha = \mathsf{rank}(J_\alpha)$. Note $n = \mathsf{rank}(A)$ 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 1×1 matrix λ_α . For a non-normal matrix A, its eigenvalues λ_α 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 $\langle\!\langle\, L^\alpha \,\|\, R^\beta \,\rangle\!\rangle = \langle\!\langle\, \widetilde{L}^\alpha \,\|\, \widetilde{R}^\beta \,\rangle\!\rangle = 0$. For eigenspaces with $n_\alpha = 1$, we may choose $\langle\!\langle\, L^\alpha \,\|\, R^\beta \,\rangle\!\rangle = \langle\!\langle\, \widetilde{L}^\alpha \,\|\, \widetilde{R}^\beta \,\rangle\!\rangle = \delta^{\alpha\beta}$, but for the nontrivial Jordan blocks with $n_\alpha > 1$ we have $\langle\!\langle\, \widetilde{L}^\alpha \,\|\, \widetilde{R}^\alpha \,\rangle\!\rangle = 0$, as we have seen in the previous paragraph, and therefore $\langle\!\langle\, L^\alpha \,\|\, R^\alpha \,\rangle\!\rangle = 0$. Real symmetric matrices are all normal, with no Jordan blocks. For complex symmetric matrices, we may have nontrivial Jordan blocks.

---

[12] 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.

Since $\langle\!\langle\, L \,\|\, R \,\rangle\!\rangle = \langle\!\langle\, \widetilde{L} \,\|\, \widetilde{R} \,\rangle\!\rangle = 0$, we may write

$$A = \bigoplus_{n_\alpha=1} \lambda_\alpha \,\|\, R^\alpha \,\rangle\!\rangle\langle\!\langle\, L^\alpha \,\| \bigoplus_{n_\beta>1} Q J_\beta \, Q^{-1} \tag{3.103}$$

and raising A to the k power yields

$$A^k = \bigoplus_{n_\alpha=1} \lambda^k_\alpha \,\|\, R^\alpha \,\rangle\!\rangle\langle\!\langle\, L^\alpha \,\| \bigoplus_{n_\beta>1} Q J^k_\beta \, Q^{-1} \quad . \tag{3.104}$$

Note that J^k_β is upper triangular with all diagonal elements given by λ^k_β. Note that for complex symmetric matrices, the left and right eigenvectors are identical and we may write

$$L^\alpha_j = R^\alpha_j \equiv \varPsi^\alpha_j \qquad \Rightarrow \qquad \langle\!\langle\, L^\alpha \,\| = \langle\!\langle\, \varPsi^\alpha \,\| \quad , \quad \|\, R^\alpha \,\rangle\!\rangle = \|\, \varPsi^\alpha \,\rangle\!\rangle \quad , \tag{3.105}$$

with no complex conjugation, *i.e.* $\langle\!\langle\, L^\alpha \,\|\, j \,\rangle\!\rangle = \langle\!\langle\, j \,\|\, R^\alpha \,\rangle\!\rangle \equiv \varPsi^\alpha_j$.

3.6.1 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 $A = UD\,V^\dagger$, 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 rows of U and V are the singular vectors. Note that $A^\dagger A = VD^2\,V^\dagger$ is $n \times n$ and $AA^\dagger = UD^2\,U^\dagger$ is $m \times m$. If we define

$$R(\lambda) = \prod_{j=1}^{k} \big(\lambda - d^2_j\big) \quad , \tag{3.106}$$

Then

$$P(\lambda) \equiv \mathsf{det}(\lambda - A^\dagger A) = \lambda^{n-k} R(\lambda) \qquad , \qquad Q(\lambda) \equiv \mathsf{det}(\lambda - AA^\dagger) = \lambda^{m-k} R(\lambda) \quad . \tag{3.107}$$

For any square $n \times n$ complex matrix A we therefore have two decompositions, via JCF and SVD, *viz.*

$$A = QJ\,Q^{-1} = UD\,V^\dagger \quad , \tag{3.108}$$

where J is the Jordan canonical form of A. When A is normal, $k = n$ and $U = V = Q$, *i.e.* the two decompositions are equivalent.

3.6.2 Example

Consider the real asymmetric matrix

$$A = \begin{pmatrix} 2a & -4a^2 \\ \frac{1}{4} & 0 \end{pmatrix} \quad , \tag{3.109}$$

where $a \in \mathbb{Z}$ is any real number. The characteristic polynomial is $F(\lambda) = \mathsf{det}(\lambda - A) = (\lambda - a)^2$ and there is a single eigenvalue, $\lambda = a$. The right and left eigenvectors are found to be

$$\| \, R \, \rangle\!\rangle = \begin{pmatrix} 4a \\ 1 \end{pmatrix} \quad , \quad \langle\!\langle \, L \, \| = \begin{pmatrix} \frac{1}{4} & -a \end{pmatrix} \quad , \tag{3.110}$$

where the normalization is arbitrary. Note $\langle\!\langle \, L \, \| \, R \, \rangle\!\rangle = 0$. The matrix A is brought to JCF by the similarity transformation $\tilde{A} = Q^{-1} A \, Q$ with

$$Q = \begin{pmatrix} 4a & 4 \\ 1 & 0 \end{pmatrix} \quad , \quad Q^{-1} = \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & -a \end{pmatrix} \quad , \quad \tilde{A} = Q^{-1} A \, Q = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \quad . \tag{3.111}$$

Note that

$$\| \, \widetilde{R} \, \rangle\!\rangle = Q^{-1} \, \| \, R \, \rangle\!\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \langle\!\langle \, \widetilde{L} \, \| = \langle\!\langle \, L \, \| \, Q = \begin{pmatrix} 0 & 1 \end{pmatrix} \quad , \tag{3.112}$$

and that $\langle\!\langle \, \widetilde{L} \, \| \, \widetilde{R} \, \rangle\!\rangle = 0$.

Adding another row and column to our matrix A, consider the matrix

$$B = \begin{pmatrix} 2a & -4a^2 & 0 \\ \frac{1}{4} & 0 & 0 \\ 0 & 0 & b \end{pmatrix} \quad , \tag{3.113}$$

where both a and b are arbitrary real numbers. Since $\mathsf{det}(\lambda - B) = (\lambda - a)^2 \, (\lambda - b)$, B thus has two eigenvalues: $\lambda_1 = a$ and $\lambda_2 = b$. The decomposition of B is then

$$B = Q J_1 Q^{-1} \bigoplus \lambda_2 \, \| \, R_2 \, \rangle\!\rangle \langle\!\langle \, L_2 \, \| \quad , \tag{3.114}$$

where $\langle\!\langle \, L_2 \, \| = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ and $\| \, R_2 \, \rangle\!\rangle = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^{\mathsf{T}}$. 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 Appendix II : Formal Solution of the Master Equation

Recall the master equation $\dot{P}_i = -\varGamma_{ij} \, P_j$. The matrix $\varGamma_{ij}$ is real but not necessarily symmetric. For such a matrix, the left eigenvectors L^α_i and the right eigenvectors R^β_j are in general not related by a simple transpose operation:

$$\begin{aligned} L^\alpha_i \, \varGamma_{ij} &= \lambda_\alpha \, L^\alpha_j \\ \varGamma_{ij} \, R^\beta_j &= \lambda_\beta \, R^\beta_i \quad . \end{aligned} \tag{3.115}$$

We denote the right and eigenvectors by $\| \, R^\alpha \, \rangle\!\rangle$ and $\langle\!\langle \, L^\alpha \, \|$, which are column vectors and row vectors, respectively. Thus $\varGamma \, \| \, R^\alpha \, \rangle\!\rangle = \lambda_\alpha \| \, R^\alpha \, \rangle\!\rangle$ and $\langle\!\langle \, L^\alpha \, \| \, \varGamma = \langle\!\langle \, L^\alpha \, \| \, \lambda_\alpha$, the second of which may be written as $\varGamma^{\mathsf{T}} \, \| \, L^\alpha \, \rangle\!\rangle = \lambda_\alpha \, \| \, L^\alpha \, \rangle\!\rangle$, where the column vector $\| \, L^\alpha \, \rangle\!\rangle$ is the transpose of the row vector $\langle\!\langle \, L^\alpha \, \|$ and $\varGamma^{\mathsf{T}}$ is the matrix transpose of $\varGamma$. The characteristic polynomial is the same in both cases:

$$F(\lambda) \equiv \mathsf{det} \, (\lambda - \varGamma) = \mathsf{det} \, (\lambda - \varGamma^{\mathsf{T}}) \quad , \tag{3.116}$$

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^α on the right by R^β_j and summing over j, and multiplying the eigenvector equation for R^β on the left by L^α_i and summing over i, and subtracting the two results yields $(\lambda_\alpha - \lambda_\beta)\,\langle\!\langle\, L^\alpha \,\|\, R^\beta \,\rangle\!\rangle = 0\,$, where the inner product is now $\langle\!\langle\, L^\alpha \,\|\, R^\beta \,\rangle\!\rangle = \sum_i L^\alpha_i\, R^\beta_i$ with no complex conjugation on the bra vector. We may now demand $\langle\!\langle\, L^\alpha \,\|\, R^\beta \,\rangle\!\rangle = \delta_{\alpha\beta}\,$, which is our eigenvector normalization condition. As discussed above in §3.6, in the event that $\varGamma$ contains nontrivial Jordan blocks, its eigenvectors do not span. However, this is a nongeneric state of affairs, and here we assume that $\varGamma$ contains no nontrivial Jordan blocks.

We have seen that $\langle\!\langle\, L \,\| = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$ is a left eigenvector of the matrix $\varGamma$ with eigenvalue $\lambda = 0$, since $\sum_i \varGamma_{ij} = 0$. We do not know *a priori* the corresponding right eigenvector, which depends on other details of $\varGamma_{ij}$. 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 $\varGamma$ span, then, let's expand the probability distribution $P_i(t)$ in the right eigenvectors of $\varGamma$, writing

$$P_i(t) = \sum_\alpha C_\alpha(t)\, R^\alpha_i \quad , \tag{3.117}$$

where $\alpha \in \{0, 1, \ldots, n-1\}$, where n is the rank of $\varGamma$. Then

$$\frac{dP_i}{dt} = \sum_\alpha \frac{dC_\alpha}{dt}\, R^\alpha_i = -\varGamma_{ij}\, P_j = -\sum_\alpha C_\alpha\, \varGamma_{ij}\, R^\alpha_j = -\sum_\alpha \lambda_\alpha\, C_\alpha\, R^\alpha_i \quad . \tag{3.118}$$

This allows us to write

$$\frac{dC_\alpha}{dt} = -\lambda_\alpha\, C_\alpha \qquad \Longrightarrow \qquad C_\alpha(t) = C_\alpha(0)\, e^{-\lambda_\alpha t} \quad . \tag{3.119}$$

Hence, we can write

$$P_i(t) = \sum_\alpha C_\alpha(0)\, e^{-\lambda_\alpha t}\, R^\alpha_i \quad . \tag{3.120}$$

Let $\alpha = 1$ correspond to the left eigenvector $\langle\!\langle\, L^1 \,\| = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$. The corresponding eigenvalue is $\lambda_1 = 1\,$. It is now easy to see that $\mathrm{Re}\,(\lambda_\alpha) > 0$ for all $\alpha > 1$, or else the probabilities will become negative[13]. For suppose $\mathrm{Re}\,(\lambda_\alpha) < 0$ for some α. Then as $t \to \infty$, the sum in eqn. 3.120 will be dominated by the term for which λ_α has the largest negative real part; all other contributions will be subleading. But we must have $\sum_i R^\alpha_i = 0$ since $\|\, R^\alpha \,\rangle\!\rangle$ must be orthogonal to the left eigenvector $\langle\!\langle\, L^0 \,\|$. Therefore, at least one component of R^α_i (*i.e.* for some value of i) must have a negative real part, which means a negative probability![14] 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 $P_i(t) \to P^{\mathrm{eq}}_i$ as $t \to \infty$, relaxing to the $\lambda = 0$ right eigenvector, with $\mathrm{Re}\,(\lambda_\alpha) \geq 0$ for all α.

---

[13] We presume that the eigenvalue $\lambda = 0$ is nondegenerate.

[14] 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 $P_i(t)$.

3.7.1 Detailed balance

Consider an arbitrary nonnegative real upper triangular matrix T with $T_{ij} \geq 0$ for all $1 \leq i < j \leq n$. Let π_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

$$W_{ij} = \begin{cases} \pi_i \, T_{ij} & \text{if } i < j \\ \pi_j \, T_{ji} & \text{if } i > j \quad . \end{cases} \tag{3.121}$$

and take this to be the matrix of transition rates so that the master equation is as in Eqn. 3.1

$$\frac{dP_i}{dt} = \sum_j \big(W_{ij} \, P_j - W_{ji} \, P_i \big) \quad . \tag{3.122}$$

Since $W_{ij}/W_{ji} = \pi_i/\pi_j$, the matrix W satisfies detailed balance relative to the distribution π. With $\varGamma_{ij} \equiv W_{ij}$ for $i \neq j$ and $\varGamma_{ii} \equiv \sum'_k W_{ki}$ (with $k = i$ excluded from the sum) as before, we recover the form of the master equation $\dot{P}_i = -\sum_j \varGamma_{ij} \, P_j$.

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 π, we conclude that $\frac{1}{2}(n-1)(n+2)$ parameters are required to specify W_{ij} . But if we drop the constraint of detailed balance, then all the elements of W_{ij} not lying on the diagonal are independent, corresponding to $n(n-1)$ parameters. Note that we may set $W_{ii} = 0$ for all i.

3.8 Appendix III : 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

$$\frac{dP_n}{dt} = \sum_m \big(W_{nm} \, P_m - W_{mn} \, P_n \big) \quad . \tag{3.123}$$

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

$$W_{nm} = \begin{cases} \lambda & \text{if } n = m + 1 \\ 0 & \text{if } n \neq m + 1 \, . \end{cases} \tag{3.124}$$

We then have

$$\frac{dP_n}{dt} = \lambda \big(P_{n-1} - P_n \big) \, . \tag{3.125}$$

The *generating function* $P(z,t)=\sum_{n=0}^{\infty} z^n P_n(t)$ then satisfies

$$\frac{\partial P}{\partial t}=\lambda(z-1)\,P \quad\Rightarrow\quad P(z,t)=e^{(z-1)\lambda t}\,P(z,0)\ . \tag{3.126}$$

If the initial distribution is $P_n(0)=\delta_{n,0}$, then

$$P_n(t)=\frac{(\lambda t)^n}{n!}\,e^{-\lambda t}\ , \tag{3.127}$$

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

$$\langle n^k\rangle = e^{-\alpha}\left(\alpha\frac{\partial}{\partial\alpha}\right)^{\!k} e^{\alpha}\ . \tag{3.128}$$

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

$$W_{nm}=\begin{cases}0 & \text{if } n\ge m\\ m\gamma & \text{if } n=m-1\\ 0 & \text{if } n<m-1\end{cases}\quad . \tag{3.129}$$

Here, γ is the decay rate of an individual atom, which can be determined from quantum mechanics. The master equation then tells us

$$\frac{dP_n}{dt}=(n+1)\,\gamma\,P_{n+1}-n\,\gamma\,P_n \quad . \tag{3.130}$$

The interpretation here is as follows: let $|\,n\,\rangle$ denote a state in which n atoms are excited. Then $P_n(t)=\big|\langle\,n\,|\,P(t)\,\rangle\big|^2$. 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

$$\begin{aligned}\frac{dN}{dt}&=\sum_{n=0}^{\infty} n\Big[(n+1)\,\gamma\,P_{n+1}-n\,\gamma\,P_n\Big]\\ &=\gamma\sum_{n=0}^{\infty}\Big[n(n-1)\,P_n-n^2P_n\Big]=-\gamma\sum_{n=0}^{\infty} n\,P_n=-\gamma\,N \quad .\end{aligned} \tag{3.131}$$

Thus, $N(t)=N(0)\,e^{-\gamma t}$. The relaxation time is $\tau=\gamma^{-1}$, and the equilibrium distribution is $P_n^{\text{eq}}=\delta_{n,0}$. Note that this satisfies detailed balance.

We can go a bit farther here. Let us define

$$P(z,t)\equiv\sum_{n=0}^{\infty} z^n\,P_n(t) \quad . \tag{3.132}$$

This is sometimes called a *generating function*. Then

$$\begin{aligned} \frac{\partial P}{\partial t} &= \gamma \sum_{n=0}^{\infty} z^n \Big[(n+1)\, P_{n+1} - n\, P_n \Big] \\ &= \gamma\, \frac{\partial P}{\partial z} - \gamma z\, \frac{\partial P}{\partial z} \quad . \end{aligned} \tag{3.133}$$

Thus,

$$\frac{1}{\gamma}\,\frac{\partial P}{\partial t} - (1-z)\,\frac{\partial P}{\partial z} = 0 \quad . \tag{3.134}$$

We now see that any function $f(\xi)$ satisfies the above equation, where $\xi = \gamma t - \ln(1-z)$. Thus, we can write

$$P(z,t) = f\big(\gamma t - \ln(1-z)\big) \quad . \tag{3.135}$$

Setting $t=0$ we have $P(z,0) = f\big(-\ln(1-z)\big)$, and inverting this result we obtain $f(u) = P(1-e^{-u},0)$, which entails

$$P(z,t) = P\big(1 + (z-1)\, e^{-\gamma t}\, ,\, 0\big) \quad . \tag{3.136}$$

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

$$N(t) = \sum_{n=0}^{\infty} n\, P_n(t) = \frac{\partial P}{\partial z}\bigg|_{z=1} = \, e^{-\gamma t}\, P(1,0) = N(0)\, e^{-\gamma t} \quad . \tag{3.137}$$

3.9 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 $dX/dt = \partial\bar{H}/\partial\bar{P}$, *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

$$\langle X\rangle_t = \lim_{T\to\infty} \frac{1}{T}\int\limits_0^T \! dt\, X(t) \quad .$$

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

$$\langle \gamma\rangle_t = \lim_{T\to\infty} \frac{1}{T}\int\limits_0^T \! dt\, \sum_i \delta(t-t_i) \quad ,$$

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. 3.11.

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:

$$\big\langle F(\boldsymbol{\varphi})\big\rangle_{\mu\text{ce}} \equiv \widetilde{\mathsf{Tr}}\left[F(\boldsymbol{\varphi})\,\delta\big(E - H(\boldsymbol{\varphi})\big)\right]\Big/\widetilde{\mathsf{Tr}}\,\delta\big(E - H(\boldsymbol{\varphi})\big)$$

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 ∞, 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 β-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,

$$\begin{aligned} D(E) &= 2\pi\sqrt{r}\int\limits_0^\infty dw \int\limits_0^\infty dX \int\limits_0^X dx\; \delta\big(w + X + rx - E\big) \\ &= \frac{2\pi}{\sqrt{r}}\int\limits_0^\infty dw \int\limits_0^\infty dX \int\limits_0^X dx\;\; \Theta(E - w - X)\, \Theta(X + rX - E + w) \\ &= \frac{2\pi}{\sqrt{r}}\int\limits_0^E dw \int\limits_{\frac{E-w}{1+r}}^{E-w} dX = \frac{2\pi\sqrt{r}}{1+r}\int\limits_0^E dq\; q = \frac{\sqrt{r}}{1+r}\cdot \pi E^2 \qquad , \end{aligned}$$

with $q = E - w$. Similarly,

$$\begin{aligned} N(E) &= 2\pi\sqrt{r}\int\limits_0^\infty dw \int\limits_0^\infty dX\; X \int\limits_0^X dx\; \delta\big(w + X + rx - E\big) \\ &= \frac{2\pi}{\sqrt{r}}\int\limits_0^\infty dw \int\limits_0^\infty dX\; X \int\limits_0^X dx\;\; \Theta(E - w - X)\, \Theta(X + rX - E + w) \\ &= \frac{2\pi}{\sqrt{r}}\int\limits_0^E dw \int\limits_{\frac{E-w}{1+r}}^{E-w} dX\; X = \frac{2\pi}{\sqrt{r}}\int\limits_0^E dq \left(1 - \frac{1}{(1+r)^2}\right)\cdot \tfrac{1}{2}q^2 = \left(\frac{2+r}{1+r}\right)\cdot \frac{\sqrt{r}}{1+r}\cdot \tfrac{1}{3}\pi E^3 \qquad . \end{aligned}$$

(c) Using the general result

$$\delta\big(F(x) - A\big) = \sum_i \frac{\delta(x - x_i)}{\big|F'(x_i)\big|} \qquad ,$$

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

$$N(E) = \widetilde{\mathsf{Tr}}\,\big[\Theta(p)\, r^{-1} p\, \delta(x - 0^+)\, \delta(E - H)\big] \quad .$$

The Laplace transform is

$$\begin{aligned} \widehat{N}(\beta) &= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_0^\infty dp\; r^{-1}\, p\, e^{-\beta p^2/2r} \int\limits_0^\infty dX\; e^{-\beta X} \\ &= \sqrt{\frac{2\pi}{\beta}}\cdot \frac{1}{\beta}\cdot \frac{1}{\beta} = \sqrt{2\pi}\,\beta^{-5/2} \qquad . \end{aligned}$$

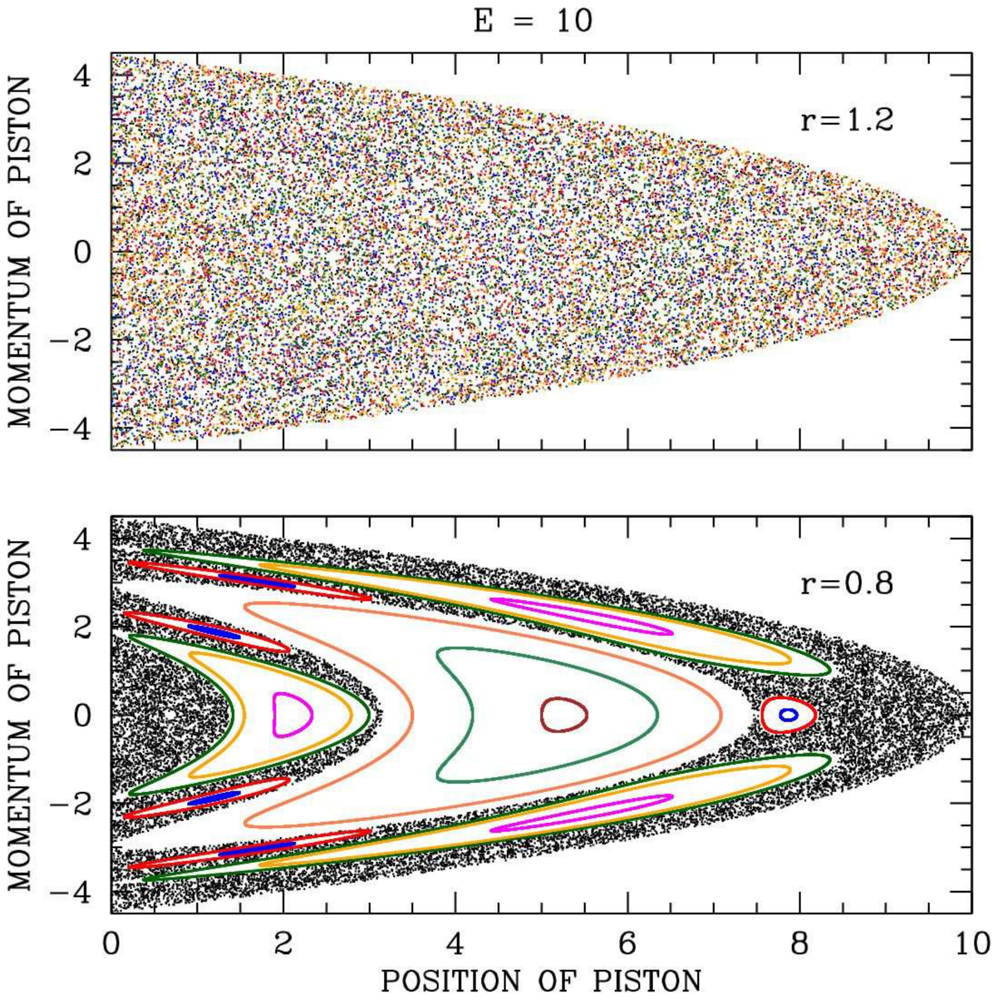


Figure 3.11: Poincaré sections for the ball and piston head problem. Each color corresponds to a different initial condition. When the mass ratio $r = m/M$ exceeds unity, the system apparently becomes ergodic.

Thus, $N(E) = \frac{4\sqrt{2}}{3}\, E^{3/2}$ and

$$\langle\gamma\rangle = \frac{N(E)}{D(E)} = \tfrac{4\sqrt{2}}{3\pi}\bigg(\frac{1+r}{\sqrt{r}}\bigg) E^{-1/2} \quad .$$

(d) When the constraint $X \ge x$ is removed, we integrate over all phase space. We then have

$$\begin{aligned}\widehat{D}(\beta) &= \mathrm{Tr}\; e^{-\beta H}\\ &= \int\limits_{-\infty}^{\infty}\! dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty}\! dp\, e^{-\beta p^2/2r} \int\limits_{0}^{\infty}\! dX\, e^{-\beta X} \int\limits_{0}^{\infty}\! dx\, e^{-\beta r x} = \frac{2\pi\sqrt{r}}{\beta^3} \quad .\end{aligned}$$

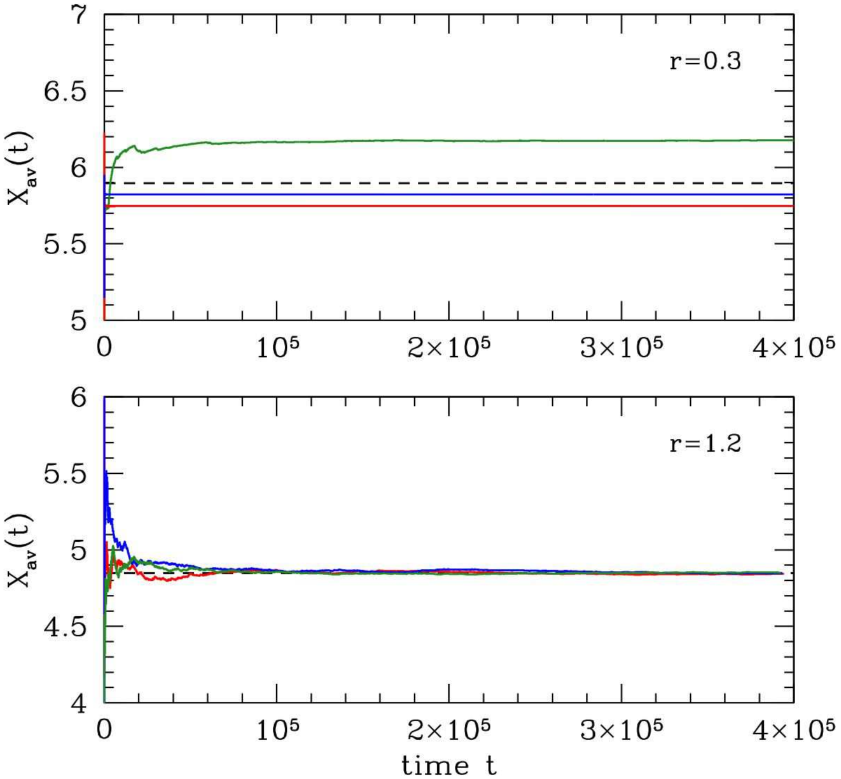


Figure 3.12: Long time running numerical averages $X_{\rm av}(t) \equiv t^{-1}\int_0^t dt'\,X(t')$ for $r = 0.3$ (top) and $r = 1.2$ (bottom), each for three different initial conditions, with $E = 10$ in all cases. Note how in the $r = 0.3$ case the long time average is dependent on the initial condition, 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 $r = 0.9$. 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 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 $\big\{(\mathrm{x},\mathrm{y})\,\big|\,\mathrm{x}\ge 0\,,\ \mathrm{y}\ge \mathrm{x}\,\mathrm{ctn}\,\phi\big\}$ for some fixed angle $\phi = \tan^{-1}\sqrt{\frac{m}{M}}$.

r	$X(0)$	N_{b}	$\langle X(t)\rangle$	$\langle X\rangle_{\mu\mathrm{ce}}$	$\langle\gamma(t)\rangle$	$\langle\gamma\rangle_{\mu\mathrm{ce}}$
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_{b} ranging from 10^4 to 10^9.

To see this, pass to relative ($\mathcal{X}$) and center-of-mass ($\mathcal{Y}$) coordinates,

$$\mathcal{X} = X - x \qquad \mathcal{P}_x = \frac{mP - Mp}{M+m}$$
$$\mathcal{Y} = \frac{MX + mx}{M+m} \qquad \mathcal{P}_y = P + p \quad .$$

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 $\mathsf{y} \geq \sqrt{\frac{M}{m}}\,\mathsf{x}$.

3.10 Appendix V : Thermalization of Quantum Systems

3.10.1 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

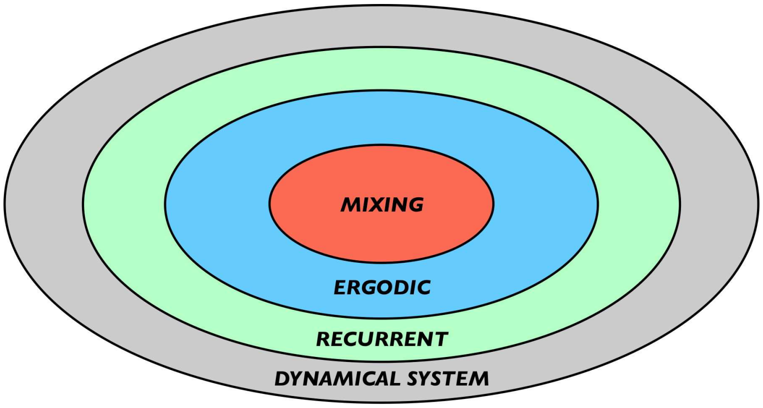


Figure 3.13: The hierarchy of dynamical systems.

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 $\big\{ |\,\Psi_\alpha\,\rangle \big\}$ 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

$$|\Psi(t)\rangle = \sum_\alpha C_\alpha \, e^{-iE_\alpha t/\hbar} \, |\Psi_\alpha\rangle \quad . \tag{3.138}$$

The energy is distributed according to the time-independent function

$$P(E) = \langle\, \Psi(t)\,|\,\delta(E-\hat{H})\,|\,\Psi(t)\,\rangle = \sum_\alpha |C_\alpha|^2 \, \delta(E - E_\alpha) \quad . \tag{3.139}$$

Thus, the average energy is time-independent and is given by

$$\langle E\rangle = \langle\, \Psi(t)\,|\,\hat{H}\,|\,\Psi(t)\,\rangle = \int\limits_{-\infty}^{\infty} \! dE \; P(E) \, E = \sum_\alpha |C_\alpha|^2 \, E_\alpha \quad . \tag{3.140}$$

The root mean square fluctuations of the energy are given by

$$(\Delta E)_{\rm rms} = \Big\langle \big(E - \langle E\rangle\big)^2 \Big\rangle^{1/2} = \sqrt{\sum_\alpha |C_\alpha|^2 \, E_\alpha^2 - \Big(\sum_\alpha |C_\alpha|^2 \, E_\alpha\Big)^2} \quad . \tag{3.141}$$

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

$$\langle \mathcal{A}(t)\rangle = \langle\, \Psi(t)\,|\,\mathcal{A}\,|\,\Psi(t)\,\rangle = \sum_{\alpha,\beta} C_\alpha^* \, C_\beta \, e^{i(E_\alpha - E_\beta)t/\hbar} \, \mathcal{A}_{\alpha\beta} \quad , \tag{3.142}$$

where $\mathcal{A}_{\alpha\beta} = \langle\, \Psi_\alpha \,|\, \mathcal{A} \,|\, \Psi_\beta \,\rangle$. In the limit of large times, we have

$$\langle \mathcal{A} \rangle_t \equiv \lim_{T\to\infty} \frac{1}{T} \int\limits_0^T \! dt \, \langle \mathcal{A}(t) \rangle = \sum_\alpha |C_\alpha|^2 \, \mathcal{A}_{\alpha\alpha} \quad . \tag{3.143}$$

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.

3.10.2 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,

$$\langle \mathcal{A} \rangle_t = \sum_\alpha |C_\alpha|^2 \, \mathcal{A}_{\alpha\alpha} = \frac{\sum_\alpha \mathcal{A}_{\alpha\alpha} \, \Theta(E_\alpha \in I)}{\sum_\alpha \Theta(E_\alpha \in I)} \equiv \langle \mathcal{A} \rangle_E , \tag{3.144}$$

where $I = \big[E, E + \Delta E\big]$ is an energy interval of width ΔE. So once again, time averages are micro canonical averages.

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[15]. Now consider a partition of $U = W \cup S$, where $W \gg S$. 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.*

$$|\Psi_\alpha\rangle = \sum_{i=1}^{N_W} \sum_{j=1}^{N_S} \Phi^\alpha_{ij} \, |\psi^W_i\rangle \otimes |\psi^S_j\rangle \quad . \tag{3.145}$$

Here $N_{W/S}$ is the size of the basis for W/S. The *reduced density matrix* for S is defined as

$$\varrho_S = \mathop{\mathsf{Tr}}_W P_\alpha = \sum_{j,j'=1}^{N_S} \Bigg(\sum_{i=1}^{N_W} \Phi^\alpha_{ij} \, \Phi^{\alpha}_{ij'} \Bigg) \, |\psi^S_j\rangle\langle\psi^S_{j'}| \quad . \tag{3.146}$$

---

[15] 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.

The claim is that ϱ_S approximates a *thermal density matrix* on S, *i.e.*

$$\varrho_S \approx \frac{1}{Z_S}\, e^{-\beta \hat{H}_S} \quad , \tag{3.147}$$

where $\hat{H}_S$ is some local Hamiltonian on S and $Z_S = \mathsf{Tr}\; e^{-\beta \hat{H}_S}$, so that ϱ_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_α/V_U to the thermal energy density, *i.e.*

$$\frac{E_\alpha}{V} = \frac{1}{V}\, \frac{\mathsf{Tr}\; \hat{H}\, e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} \quad . \tag{3.148}$$

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 ϱ_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 ϱ_S by its Gibbs form in eqn. 3.147 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 ξ_T from ∂S, where ξ_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 ξ_T. How do we define ξ_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 $(d_\partial(\boldsymbol{r})/\xi_T)^{-p}$, where p is a critical exponent. 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 is discussed in the appendix §3.10.4.

3.10.3 Entanglement and entanglement entropy

The *von Neumann entanglement entropy* between S and W is defined to be

$$S_{\mathrm{vN}} = -\,\mathsf{Tr}\;\varrho_S \log \varrho_S = -\sum_j \lambda_j \log \lambda_j \quad , \tag{3.149}$$

where $\{\lambda_j\}$ are the eigenvalues of the reduced density matrix ϱ_S. Each eigenvalue λ_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 $\varrho_S = \mathsf{Tr}_W\,\varrho$ does not correspond to a pure state. *I.e.*, as stated above, ϱ_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 $\varrho_S = \mathsf{Tr}_W |\Psi\rangle\langle\Psi| = |\chi_S\rangle\langle\chi_S|$ is pure, and $S_{\text{vN}} = 0$. One then has that W and S are *unentangled*.

For a simple example, consider a system consisting of two spin-$\frac{1}{2}$ objects, are the wavefunction

$$|\,\Psi\,\rangle = \alpha\,|\!\uparrow\downarrow\,\rangle + \beta\,|\!\downarrow\uparrow\,\rangle \tag{3.150}$$

with $|\alpha|^2 + |\beta|^2 = 1$. Then the two spin density matrix is

$$\varrho = |\alpha|^2\,|\!\uparrow\downarrow\,\rangle\langle\,\uparrow\downarrow\!| + |\beta|^2\,|\!\downarrow\uparrow\,\rangle\langle\,\downarrow\uparrow\!| + \alpha\beta^*\,|\!\uparrow\downarrow\,\rangle\langle\,\downarrow\uparrow\!| + \alpha^*\beta\,|\!\downarrow\uparrow\,\rangle\langle\,\uparrow\downarrow\!| \quad . \tag{3.151}$$

Taking the trace over the first spin, the reduced density matrix $\tilde{\varrho}$ is given by

$$\tilde{\varrho} = \mathop{\mathsf{Tr}}_{1}\,\varrho = |\alpha|^2\,|\!\downarrow\,\rangle\langle\,\downarrow\!| + |\beta|^2\,|\!\uparrow\,\rangle\langle\,\uparrow\!| = \begin{pmatrix} |\beta|^2 & 0 \\ 0 & |\alpha|^2 \end{pmatrix} \quad . \tag{3.152}$$

The entanglement entropy is then

$$S_{\text{vN}} = -|\alpha|^2 \log|\alpha|^2 - |\beta|^2 \log|\beta|^2 \quad , \tag{3.153}$$

and if our initial state is a singlet with $\alpha = \frac{1}{\sqrt{2}}$ and $\beta = -\frac{1}{\sqrt{2}}$, we obtain $S_{\text{vN}} = \log 2$.[16]

3.10.4 Singular value decomposition and entanglement

Consider a quantum system whose xwavefunction 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

$$|\,\Psi\,\rangle = \sum_{a=1}^{N_A}\sum_{b=1}^{N_B} \Psi_{ab}\,|\,\psi^A_a\,\rangle \otimes |\,\psi^B_b\,\rangle \quad , \tag{3.154}$$

where Ψ_{ab} is a complex $N_A \times N_B$ matrix. Any such matrix has a *singular value decomposition* [17] given by $\Psi = UDV^{\mathsf{T}}$, where (i) U is a complex rectangular matrix of dimensions $N_A \times N_{\text{s}}$ satisfying $U^\dagger U = \mathbb{I}_{N_{\text{s}} \times N_{\text{s}}}$,

---

[16] 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 δ.

[17] See §3.6.1 below as well.

(ii) D is a square diagonal matrix of dimensions $N_\text{s} \times N_\text{s}$ with positive real entries, (iii) V is a complex rectangular matrix of dimensions $N_B \times N_\text{s}$ satisfying $V^\dagger V = \mathbb{I}_{N_\text{s} \times N_\text{s}}$, and (iv) $N_\text{s} \le \text{min}(N_A, N_B)$. Writing $D = \text{diag}\{d_1, \ldots, d_{N_\text{s}}\}$ with $p \in \{1, \ldots, N_\text{s}\}$ the *singular values*, we obtain

$$| \, \Psi \, \rangle = \sum_{p=1}^{N_\text{s}} d_p \, | \, \chi_p^A \, \rangle \otimes | \, \chi_p^B \, \rangle \quad , \tag{3.155}$$

where

$$| \, \chi_p^A \, \rangle = \sum_{a=1}^{N_A} U_{ap} \, | \, \psi_a^A \, \rangle \qquad , \qquad | \, \chi_p^B \, \rangle = \sum_{b=1}^{N_B} V_{bp} \, | \, \psi_b^B \, \rangle \quad . \tag{3.156}$$

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

$$\begin{aligned} \varrho_A &= \underset{B}{\text{Tr}} \, | \, \Psi \, \rangle \langle \, \Psi \, | = \sum_{p=1}^{N_\text{s}} d_p^2 \, | \, \chi_p^A \, \rangle \langle \, \chi_p^A \, | \\ \varrho_B &= \underset{A}{\text{Tr}} \, | \, \Psi \, \rangle \langle \, \Psi \, | = \sum_{p=1}^{N_\text{s}} d_p^2 \, | \, \chi_p^B \, \rangle \langle \, \chi_p^B \, | \end{aligned} \tag{3.157}$$

Since $\text{Tr} \, \varrho_A = \text{Tr} \, \varrho_B = 1$, we must have that the squares of the singular values sum to unity: $\sum_{p=1}^{N_\text{s}} d_p^2 = 1$. The entanglement entropy is then

$$S_\text{vN} \equiv S_A = S_B = -\sum_{p=1}^{N_\text{s}} d_p^2 \log(d_p^2) \quad . \tag{3.158}$$

Writing $d_p \equiv \exp(-\xi_p/2)$, we have

$$\varrho_A = \sum_{p=1}^{N_\text{s}} e^{-\xi_p} \, | \, \chi_p^A \, \rangle \langle \, \chi_p^A \, | \equiv e^{-\hat{K}} \quad , \tag{3.159}$$

and we identify the *entanglement Hamiltonian* as

$$\hat{K} \equiv \sum_{p=1}^{N_\text{s}} \xi_p \, | \, \chi_p^A \, \rangle \langle \, \chi_p^A \, | \quad , \tag{3.160}$$

with $\sum_{p=1}^{N_\text{s}} \exp(-\xi_p) = 1$. Thus, $\text{spec}(\hat{K}) = \{\xi_1, \ldots, \xi_{N_\text{s}}\}$ is the spectrum of eigenvalues of the entanglement Hamiltonian. In the language of eqn. 3.147, we have $\hat{K} = \beta \hat{H}_S + k_\text{B} T \log Z_S$, with $A = S$ and $B = W$. In general there is no guarantee that the entanglement Hamiltonian $\hat{K}$ is local, but ETH assumes it to be so within the bulk of A.

3.10.5 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

$$P_\beta(H) = C_\beta \, e^{-\beta \, \mathsf{Tr}\,(H^2)/2} \quad , \tag{3.161}$$

with C_β 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 matrix rank[18]. These invariances require the Gaussian form of the probability distribution $P_\beta(H)$. 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

$$H_{mn} = \begin{cases} H^{(0)}_{mn} & \text{(GOE)} \\ H^{(0)}_{mn} + iH^{(1)}_{mn} & \text{(GUE)} \\ H^{(0)}_{mn}\,\mathbf{1} + H^{(1)}_{mn}\,\boldsymbol{i} + H^{(2)}_{mn}\,\boldsymbol{j} + H^{(3)}_{mn}\,\boldsymbol{k} & \text{(GSE)} \quad , \end{cases} \tag{3.162}$$

with $m \le n$, where $H^{(\mu)}_{mn} \in \mathbb{R}$ for $\mu \in \{0,1,2,3\}$, $H^{(0)}_{nm} = H^{(0)}_{mn}$, and $H^{(r)}_{nm} = -H^{(r)}_{mn}$ for $r \in \{1,2,3\}$. Here, $\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$ are the quaternions (see below). The measure dH in these three ensembles is defined as

$$dH = \begin{cases} \prod_{j\le k} dH^{(0)}_{jk} & \text{(GOE)} \\ \prod_{j\le k} dH^{(0)}_{jk} \prod_{m<n} dH^{(1)}_{mn} & \text{(GUE)} \\ \prod_{j\le k} dH^{(0)}_{jk} \prod_{m<n} \prod_{r=1}^{3} dH^{(r)}_{mn} & \text{(GUE)} \quad . \end{cases} \tag{3.163}$$

The standard reference on the subject is M. L. Mehta, *Random Matrices*, 3rd ed. (Academic Press, 2004)[19]. One can then compute

$$\langle H_{ij} H_{mn} \rangle = \begin{cases} \frac{1}{2}\,\delta_{in}\delta_{jm} + \frac{1}{2}\,\delta_{im}\delta_{jn} & \text{(GOE)} \\ \frac{1}{2}\,\delta_{in}\delta_{jm} & \text{(GUE)} \\ \frac{1}{2}\,\delta_{in}\delta_{jm} - \frac{1}{4}\,\delta_{im}\delta_{jn} & \text{(GSE)} \quad . \end{cases} \tag{3.164}$$

---

[18] For quaternionic matrices, expressing each quaternionic number as a 2 × 2 matrix (see below), the matrix rank becomes $2N$.

[19] 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.

For each ensemble, the representative matrices have real eigenvalues λ_j whose joint probability distribution is given by (see Mehta)[20]

$$\mathcal{P}_\beta(\lambda_1,\ldots,\lambda_N) = \mathcal{C}_\beta \prod_{j<k}^N |\lambda_j - \lambda_k|^\beta \prod_{l=1}^N e^{-\beta\lambda_l^2/2} \quad . \tag{3.165}$$

Remarkably, the density of states $\sigma(\lambda)$ of the eigenvalues for each ensemble follows the Wigner semicircle law in the large-N limit,

$$\sigma(\lambda) = \frac{1}{\pi}\sqrt{2N-\lambda^2}\,\Theta(2N-\lambda^2) \quad . \tag{3.166}$$

Note that $\int\limits_{-\infty}^{\infty} d\lambda\, \sigma(\lambda) = N$.

A few words about quaternionic numbers

A *complex quaternionic number* q is written as $q = a\boldsymbol{1} + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$, where $a,b,c,d \in \mathbb{C}$ and $\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$ are the quaternions: $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -\boldsymbol{1}$ and $\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}$, $\boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}$, and $\boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}$. In a *real quaternionic number* we have $a,b,c,d \in \mathbb{R}$.. The quaternionic algebra is satisfied if we take $\boldsymbol{1}$ as the 2×2 unit matrix and $(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) = (iZ, iY, iX)$, where $\{X,Y,Z\}$ are the Pauli matrices. We may then express each quaternionic number q as a 2×2 matrix, *viz.*

$$q = a\boldsymbol{1} + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} \quad . \tag{3.167}$$

There are three conjugacy operations defined for quaternions: complex conjugation q^*, quaternion conjugation $\bar{q}$, and Hermitian conjugation $q^\dagger$:

$$\begin{aligned} q^* &= a^*\boldsymbol{1} + b^*\boldsymbol{i} + c^*\boldsymbol{j} + d^*\boldsymbol{k} = \begin{pmatrix} a^*+ib^* & c^*+id^* \\ -c^*+id^* & a^*-ib^* \end{pmatrix} \\ \bar{q} &= a\boldsymbol{1} - b\boldsymbol{i} - c\boldsymbol{j} - d\boldsymbol{k} = \begin{pmatrix} a-ib & -c-id \\ c-id & a+ib \end{pmatrix} \\ q^\dagger &= a^*\boldsymbol{1} - b^*\boldsymbol{i} - c^*\boldsymbol{j} - d^*\boldsymbol{k} = \begin{pmatrix} a^*-ib^* & -c^*-id^* \\ c^*-id^* & a^*+ib^* \end{pmatrix} \quad ; \end{aligned} \tag{3.168}$$

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 2×2 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 $q \in \mathbb{H}$,

$$|q|^2 \equiv q\bar{q} = \bar{q}q = (a^2+b^2+c^2+d^2)\,\boldsymbol{1} = \begin{pmatrix} a^2+b^2+c^2+d^2 & 0 \\ 0 & a^2+b^2+c^2+d^2 \end{pmatrix} \quad . \tag{3.169}$$

---

[20] For the GSE, each eigenvalue is doubly degenerate due to Kramers' theorem, a consequence of time-reversal invariance and half-odd-integer spin.

Let a quaternionic matrix Q have elements $q_{mn} = a_{mn}\,\mathbf{1} + b_{mn}\,\boldsymbol{i} + c_{mn}\,\boldsymbol{j} + d_{mn}\,\boldsymbol{k}$. 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 $\mathsf{Tr}\,\mathbf{1} \equiv 1$ even though the $\mathbf{1}$ on the LHS may be regarded as a 2×2 matrix.

Eigenvector statistics

In the limit $N \to \infty$, the eigenvectors of random matrices are independent random vectors in any basis. Let ψ^α_j be the α component of the j^{th} eigenvector of a given random matrix. The probability distribution for a given eigenvector is given by

$$P\big(\{\psi^\alpha_j , \bar{\psi}^\alpha_j\}\big) \propto \prod_{j=1}^N \delta\bigg(\sum_{\alpha=1}^N |\psi^\alpha_j|^2 - 1\bigg) \quad , \tag{3.170}$$

where $\bar{\psi}^\alpha_j = \psi^\alpha_j$ for GOE, $\bar{\psi}^\alpha_j = \psi^{\alpha *}_j$ for GUE, and $\bar{\psi}^\alpha_j$ is the quaternion conjugate of ψ^α_j for GSE. It then follows that

$$\begin{aligned}
\text{GOE}: &\quad \langle \psi^\alpha_i \psi^\beta_j \rangle = \frac{1}{N}\,\delta_{ij}\,\delta^{\alpha\beta} \\
\text{GUE}: &\quad \langle \psi^\alpha_i \bar{\psi}^\beta_j \rangle = \frac{1}{N}\,\delta_{ij}\,\delta^{\alpha\beta} \quad , \quad \langle \psi^\alpha_i \psi^\beta_j \rangle = \langle \bar{\psi}^\alpha_i \bar{\psi}^\beta_j \rangle = 0 \\
\text{GSE}: &\quad \langle \psi^\alpha_i \bar{\psi}^\beta_j \rangle = \frac{1}{N}\,\delta_{ij}\,\delta^{\alpha\beta} \quad , \quad \langle \psi^\alpha_i \psi^\beta_j \rangle = \langle \bar{\psi}^\alpha_i \bar{\psi}^\beta_j \rangle = -\frac{1}{2N}\,\delta_{ij}\,\delta^{\alpha\beta} \quad .
\end{aligned} \tag{3.171}$$

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 $|\, k \,\rangle$:

$$A_{jk} = \langle\, j \,|\, \hat{A} \,|\, k \,\rangle = \sum_\alpha A_\alpha \,\langle\, j \,|\, \alpha \,\rangle\langle\, \alpha \,|\, k \,\rangle = \sum_\alpha A_\alpha\, \bar{\psi}^\alpha_j\, \psi^\alpha_k \quad . \tag{3.172}$$

Then

$$\langle A_{jk} \rangle = \sum_\alpha A_\alpha\, \langle \bar{\psi}^\alpha_j\, \psi^\alpha_k \rangle = \frac{1}{N} \sum_\alpha A_\alpha\, \delta_{jk} = \frac{\mathsf{Tr}\,(\hat{A})}{N}\, \delta_{jk} \quad . \tag{3.173}$$

We may also compute

$$\langle A_{jk}\, \bar{A}_{mn} \rangle = \sum_{\alpha,\beta} A_\alpha\, A_\beta\, \langle \bar{\psi}^\alpha_j\, \psi^\alpha_k\, \psi^\beta_m\, \bar{\psi}^\beta_n \rangle \tag{3.174}$$

using Wick's theorem, *viz.*

$$\langle \bar{\psi}^\alpha_j\, \psi^\alpha_k\, \psi^\beta_m\, \bar{\psi}^\beta_n \rangle = \langle \bar{\psi}^\alpha_j\, \psi^\alpha_k \rangle\langle \psi^\beta_m\, \bar{\psi}^\beta_n \rangle + \langle \bar{\psi}^\alpha_j\, \psi^\beta_m \rangle\langle \psi^\alpha_k\, \bar{\psi}^\beta_n \rangle + \langle \bar{\psi}^\alpha_j\, \bar{\psi}^\beta_n \rangle\langle \psi^\beta_m\, \psi^\alpha_k \rangle \quad . \tag{3.175}$$

Thus we obtain

$$\begin{aligned}
\text{GOE}: &\quad \langle \bar{\psi}^\alpha_j\, \psi^\alpha_k\, \psi^\beta_m\, \bar{\psi}^\beta_n \rangle = \frac{1}{N^2}\big(\delta_{jk}\,\delta_{mn} + \delta_{jm}\,\delta_{kn}\,\delta^{\alpha\beta} + \delta_{jn}\,\delta_{km}\,\delta^{\alpha\beta}\big) \\
\text{GUE}: &\quad \langle \bar{\psi}^\alpha_j\, \psi^\alpha_k\, \psi^\beta_m\, \bar{\psi}^\beta_n \rangle = \frac{1}{N^2}\big(\delta_{jk}\,\delta_{mn} + \delta_{jm}\,\delta_{kn}\,\delta^{\alpha\beta}\big) \\
\text{GSE}: &\quad \langle \bar{\psi}^\alpha_j\, \psi^\alpha_k\, \psi^\beta_m\, \bar{\psi}^\beta_n \rangle = \frac{1}{N^2}\big(\delta_{jk}\,\delta_{mn} + \delta_{jm}\,\delta_{kn}\,\delta^{\alpha\beta} + \tfrac{1}{4}\,\delta_{jn}\,\delta_{km}\,\delta^{\alpha\beta}\big)
\end{aligned} \tag{3.176}$$

We use these results to obtain the following:

$$\langle |A_{mn}|^2\rangle - |\langle A_{mn}\rangle|^2 = \frac{\mathsf{Tr}\,(\hat{A}^2)}{N^2}\cdot\begin{cases} 1+\delta_{mn} & \text{(GOE)} \\ 1 & \text{(GUE)} \\ 1+\frac{1}{4}\,\delta_{mn} & \text{(GSE)} \end{cases} . \tag{3.177}$$

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 $2:1:\frac{5}{4}$.

3.10.6 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.*[21]

$$A_{mn} = \langle\, m \,|\, \hat{A} \,|\, n\,\rangle = A(E)\,\delta_{mn} + e^{-S(E)/2}\, f_A(E,\omega)\, R_{mn} \tag{3.178}$$

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 rank of $\hat{A}$. We further assume that $A(E)$ and $f(E,\omega)$ are smooth functions of their arguments. Additionally, one has

$$\begin{aligned} A_{mn} \in \mathbb{R} \quad &: \quad R_{nm} = R_{mn} \quad , \quad f_A(E,-\omega) = f_A(E,\omega) \\ A_{mn} \in \mathbb{C} \quad &: \quad R_{nm} = \bar{R}_{mn} \quad , \quad f_A(E,-\omega) = f^*_A(E,\omega) \quad , \end{aligned} \tag{3.179}$$

for systems with and without time-reversal symmetry, respectively.

ETH *versus* random matrix theory

ETH reduces to random matrix theory within a given small energy window. Eqn. 3.178 is to be contrasted with the corresponding result for random matrix theory,

$$A_{mn} = \langle A\rangle\,\delta_{mn} + \sqrt{\frac{\langle A^2\rangle - \langle A\rangle^2}{\mathcal{D}}}\, R_{mn} \quad , \tag{3.180}$$

where $\langle A\rangle \equiv \mathcal{D}^{-1}\,\mathsf{Tr}\,\hat{A}$ and $\langle A^2\rangle \equiv \mathcal{D}^{-1}\,\mathsf{Tr}\,(\hat{A}^2)$ 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. 3.177. ETH thus reduces to RMT within any small window which contains $\mathcal{O}(\mathcal{D})$ states.

---

[21] See the review article by L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *Adv. Phys.*, **65**, 239 (2016) and references therein.

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*[22]. To see this, note

$$\mathsf{Tr}\,(\varrho\hat{A}) \approx \frac{1}{Z}\int dE\, D(E)\, e^{-\beta E} A(E) \approx A(E^*) \quad . \tag{3.181}$$

Here it is assumed that the function $A(E)$ in Eqn. 3.178 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

$$D(E) \sim \exp\left(-\frac{(E-E_{\rm c})^2}{2\sigma^2}\right) \quad , \tag{3.182}$$

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

$$\varepsilon^*(\beta) \equiv \frac{E^*}{N} \approx \varepsilon_{\rm c} - \beta w \quad , \tag{3.183}$$

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

$$\varepsilon^*(\beta) = \frac{1}{NZ}\,\mathsf{Tr}\,(\hat{H}e^{-\beta\hat{H}}) \quad . \tag{3.184}$$

3.10.7 When is the ETH true?

There is no rigorous 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 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 ψ_j whose energies ε_j the localized portion of the eigenspectrum decay exponentially, as $|\psi_j(\boldsymbol{r})|^2 \sim \exp\big(-|\boldsymbol{r}-\boldsymbol{r}_j|/\xi(\varepsilon_j)\big)$, where $\boldsymbol{r}_j$ is some position in space associated with ψ_j and $\xi(\varepsilon_j)$ is the *localization length*. Within the localized portion of the spectrum, $\xi(\varepsilon)$ is finite. As ε 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[23]. Exponentially localized states are unable to thermalize with other distantly removed localized states. Of course, all noninteracting systems will violate

---

[22] See V. Khemani, A. Chandran, H. Kim, and S. L. Sondhi, *Phys. Rev. E* **90**, 052133 (2014).

[23] Recall that in systems with no disorder, eigenstates exhibit Bloch periodicity in space.

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.

Quaternionic numbers

A *complex quaternionic number* q is written as $q = a\,1 + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$, where $a, b, c, d \in \mathbb{C}$ and $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ are the quaternions: $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1$ and $\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}$, $\boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}$, and $\boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}$. In a *real quaternionic number* we have $a, b, c, d \in \mathbb{R}$.. The quaternionic algebra is satisfied if we take 1 as the 2×2 unit matrix and $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}) = (iZ, iY, iX)$, where $\{X, Y, Z\}$ are the Pauli matrices. We may then express each quaternionic number q as a 2×2 matrix, *viz.*

$$q = a\,1 + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} = \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix} \quad . \tag{3.185}$$

There are three conjugacy operations defined for quaternions: complex conjugation q^*, quaternion conjugation $\bar{q}$, and Hermitian conjugation $q^\dagger$:

$$\begin{aligned} q^* &= a^*\,1 + b^*\,\boldsymbol{i} + c^*\,\boldsymbol{j} + d^*\,\boldsymbol{k} = \begin{pmatrix} a^* + ib^* & c^* + id^* \\ -c^* + id^* & a^* - ib^* \end{pmatrix} \\ \bar{q} &= a\,1 - b\,\boldsymbol{i} - c\,\boldsymbol{j} - d\,\boldsymbol{k} = \begin{pmatrix} a - ib & -c - id \\ c - id & a + ib \end{pmatrix} \\ q^\dagger &= a^*\,1 - b^*\,\boldsymbol{i} - c^*\,\boldsymbol{j} - d^*\,\boldsymbol{k} = \begin{pmatrix} a^* - ib^* & -c^* - id^* \\ c^* - id^* & a^* + ib^* \end{pmatrix} \quad ; \end{aligned} \tag{3.186}$$

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 2×2 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 $q \in \mathbb{H}$,

$$|q|^2 \equiv q\bar{q} = \bar{q}q = (a^2 + b^2 + c^2 + d^2)1 = \begin{pmatrix} a^2 + b^2 + c^2 + d^2 & 0 \\ 0 & a^2 + b^2 + c^2 + d^2 \end{pmatrix} \quad . \tag{3.187}$$

Let a quaternionic matrix Q have elements $q_{mn} = a_{mn}\,1 + b_{mn}\,\boldsymbol{i} + c_{mn}\,\boldsymbol{j} + d_{mn}\,\boldsymbol{k}$. 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 $\mathsf{Tr}\ 1 \equiv 1$ even though the 1 on the LHS may be regarded as a 2×2 matrix.

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 (μCE)

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:

$$\text{ergodicity} \quad \Longleftrightarrow \quad \big\langle f(\boldsymbol{\varphi})\big\rangle_t = \big\langle f(\boldsymbol{\varphi})\big\rangle_E \quad , \tag{4.1}$$

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 δ-function[1]. 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 quantitied $\Lambda_a(\boldsymbol{\varphi})$ 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 Tr means 'trace', which entails an integration over all phase space:

$$\mathsf{Tr}\; A(q,p) \equiv \frac{1}{N!}\prod_{i=1}^{N} \int \! \frac{d^d\!p_i\, d^d\!q_i}{(2\pi\hbar)^d}\, A(q,p) \quad . \tag{4.6}$$

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*[2]. The normalization factor $(2\pi\hbar)^{-Nd}$ renders the trace dimensionless. Again,

---

[1] We write the Hamiltonian as $\hat{H}$ (classical or quantum) in order to distinguish it from magnetic field, H.

[2] 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:

$$d\mu = \frac{1}{N!} \prod_{i=1}^{N} \frac{d^d p_i \, d^d q_i}{(2\pi\hbar)^d} \quad . \tag{4.7}$$

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

$$\begin{aligned} D(E)\,\Delta E &= \int\limits_{E}^{E+\Delta E} \! dE' \int d\mu\, \delta(E' - \hat{H}) = \int\limits_{E<\hat{H}<E+\Delta E} \! d\mu \\ &= \text{\# of states with energies between } E \text{ and } E + \Delta E \quad . \end{aligned} \tag{4.9}$$

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

$$D(E) = \frac{1}{N!} \int \prod_{i=1}^{N} \frac{d^d p_i \, d^d q_i}{(2\pi\hbar)^d} \, \delta\bigg(E - \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m}\bigg) \quad . \tag{4.11}$$

We shall calculate $D(E)$ in two ways. The first method utilizes the *Laplace transform*, $Z(\beta)$:

$$Z(\beta) = \mathcal{L}\big[D(E)\big] \equiv \int\limits_{0}^{\infty} \! dE\, e^{-\beta E}\, D(E) = \mathsf{Tr}\ e^{-\beta\hat{H}} \quad . \tag{4.12}$$

The inverse Laplace transform is then

$$D(E) = \mathcal{L}^{-1}\big[Z(\beta)\big] \equiv \int\limits_{c-i\infty}^{c+i\infty} \! \frac{d\beta}{2\pi i}\, e^{\beta E}\, Z(\beta) \quad , \tag{4.13}$$

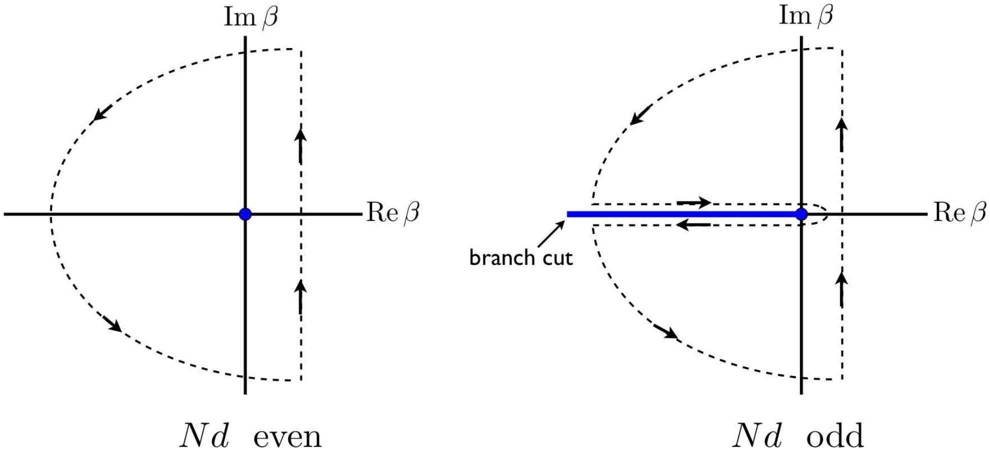


Figure 4.1: Complex integration contours $\mathcal{C}$ for inverse Laplace transform $\mathcal{L}^{-1}\big[Z(\beta)\big] = D(E)$. When the product dN is odd, there is a branch cut along the negative $\text{Re}\,\beta$ axis.

where c is such that the integration contour is to the right of any singularities of $Z(\beta)$ in the complex β-plane. We then have

$$\begin{aligned} Z(\beta) &= \frac{1}{N!}\prod_{i=1}^{N}\int\!\frac{d^dx_i\,d^dp_i}{(2\pi\hbar)^d}\;e^{-\beta\boldsymbol{p}_i^2/2m} \\ &= \frac{V^N}{N!}\left(\int\limits_{-\infty}^{\infty}\!\frac{dp}{2\pi\hbar}\;e^{-\beta p^2/2m}\right)^{Nd} = \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\beta^{-Nd/2} \quad . \end{aligned} \tag{4.14}$$

The inverse Laplace transform is then

$$\begin{aligned} D(E) &= \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\oint\limits_{\mathcal{C}}\!\frac{d\beta}{2\pi i}\;e^{\beta E}\,\beta^{-Nd/2} \\ &= \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\frac{E^{\frac{1}{2}Nd-1}}{\Gamma(Nd/2)} \quad , \end{aligned} \tag{4.15}$$

exactly as before. The integration contour for the inverse Laplace transform is extended in an infinite semicircle in the left half β-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 $\text{Re}\,\beta$ 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,

$$\ln N! = N\ln N - N + \tfrac{1}{2}\ln N + \tfrac{1}{2}\ln(2\pi) + \mathcal{O}\big(N^{-1}\big) \quad , \tag{4.17}$$

we may define the *statistical entropy*,

$$S(E,V,N) \equiv k_{\rm B} \ln D(E,V,N) = N k_{\rm B}\, \phi\bigg(\frac{E}{N}\,,\,\frac{V}{N}\bigg) + \mathcal{O}(\ln N) \quad , \tag{4.18}$$

where

$$\phi\bigg(\frac{E}{N}\,,\,\frac{V}{N}\bigg) = \frac{d}{2}\,\ln\bigg(\frac{E}{N}\bigg) + \ln\bigg(\frac{V}{N}\bigg) + \frac{d}{2}\,\ln\bigg(\frac{m}{d\pi\hbar^2}\bigg) + \big(1+\tfrac{1}{2}d\big) \quad . \tag{4.19}$$

Recall $k_{\rm B} = 1.3806503 \times 10^{-16}\,{\rm erg/K}$ is Boltzmann's constant.

Second method

The second method invokes a mathematical trick. First, let's rescale $p^\alpha_i \equiv \sqrt{2mE}\, u^\alpha_i$. We then have

$$D(E) = \frac{V^N}{N!}\left(\frac{\sqrt{2mE}}{h}\right)^{\!Nd} \frac{1}{E}\int\! d^M\!u\; \delta\big(u_1^2 + u_2^2 + \ldots + u_M^2 - 1\big) \quad . \tag{4.20}$$

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 δ-functions. We can now write

$$d^M\!u = u^{M-1}\, du\; d\Omega_M \quad , \tag{4.21}$$

where $d\Omega_M$ is the M-dimensional differential solid angle. We now have our answer:[3]

$$D(E) = \frac{V^N}{N!}\left(\frac{\sqrt{2m}}{h}\right)^{\!Nd} E^{\frac{1}{2}Nd-1} \cdot \tfrac{1}{2}\,\Omega_{Nd} \quad . \tag{4.22}$$

What remains is for us to compute Ω_M, the total solid angle in M dimensions. We do this by a nifty mathematical trick. Consider the integral

$$\begin{aligned} \mathcal{I}_M &= \int\! d^M\!u\; e^{-u^2} = \Omega_M \int_0^\infty\! du\; u^{M-1}\, e^{-\boldsymbol{u}^2} \\ &= \tfrac{1}{2}\Omega_M \int_0^\infty\! ds\; s^{\frac{1}{2}M-1}\, e^{-s} = \tfrac{1}{2}\Omega_M\, \Gamma\big(\tfrac{1}{2}M\big) \quad , \end{aligned} \tag{4.23}$$

where $s = u^2$, and where

$$\Gamma(z) = \int_0^\infty\! dt\; t^{z-1}\, e^{-t} \tag{4.24}$$

---

[3] The factor of $\frac{1}{2}$ preceding Ω_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 $z\,\Gamma(z) = \Gamma(z+1)$.[4] On the other hand, we can compute $\mathcal{I}_M$ in Cartesian coordinates, writing

$$\mathcal{I}_M = \left(\int\limits_{-\infty}^{\infty} du_1\, e^{-u_1^2} \right)^{M} = \big(\sqrt{\pi}\big)^M \quad . \tag{4.25}$$

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 ΔE in eqn. 4.9, and define $\tilde{D}(E;\Delta E) = D(E)\,\Delta E$ and $S(E;\Delta E) \equiv k_{\rm B} \ln \tilde{D}(E;\Delta E)$. The definition of statistical entropy then involves the arbitrary parameter ΔE, however this only affects $S(E)$ in an additive way. That is,

$$S(E,V,N;\Delta E_1) = S(E,V,N;\Delta E_2) + k_{\rm B} \ln\!\left(\frac{\Delta E_1}{\Delta E_2}\right) \quad . \tag{4.27}$$

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

$$Z(\beta) = \frac{V^N}{N!}\,\frac{\Omega_d^N}{h^N d} \left(\int\limits_0^{\infty} dp\, p^{d-1}\, e^{-\beta c p} \right)^{N} = \frac{V^N}{N!} \left(\frac{\Gamma(d)\,\Omega_d}{c^d\, h^d\, \beta^d} \right)^{N} \quad . \tag{4.28}$$

The statistical entropy is $S(E,V,N) = k_{\rm B} \ln D(E,V,N) = N k_{\rm B}\, \phi\big(\frac{E}{N}, \frac{V}{N}\big)$, with

$$\phi\!\left(\frac{E}{N}, \frac{V}{N}\right) = d\, \ln\!\left(\frac{E}{N}\right) + \ln\!\left(\frac{V}{N}\right) + \ln\!\left(\frac{\Omega_d\, \Gamma(d)}{(dhc)^d}\right) + (d+1) \tag{4.29}$$

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 σ_i takes discrete values. The number of ways

---

[4] 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,[5] 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

$$S(E,N,B) = k_{\rm B} \ln \Omega(E,N,B) = -N k_{\rm B} \left[\left(\frac{1+m}{2}\right) \ln\left(\frac{1+m}{2}\right) + \left(\frac{1-m}{2}\right) \ln\left(\frac{1-m}{2}\right) \right] \quad . \tag{4.32}$$

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

$$\begin{aligned} 0 &= \frac{1}{P_1(E_1)} \frac{dP_1(E_1)}{dE_1} = \frac{\partial}{\partial E_1} \ln P_1(E_1) \\ &= \frac{\partial}{\partial E_1} \ln D_1(E_1) + \frac{\partial}{\partial E_1} \ln D_2(E - E_1) \quad . \end{aligned} \tag{4.35}$$

---

[5] 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.

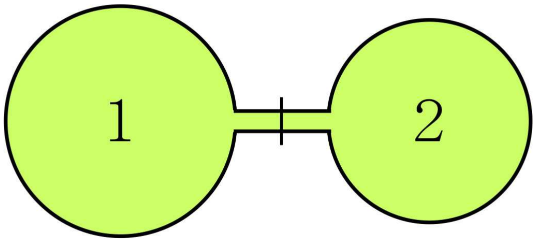


Figure 4.2: Two systems in thermal contact.

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

$$\frac{1}{T} = \bigg(\frac{\partial S}{\partial E}\bigg)_{V,N} \quad , \tag{4.38}$$

a result familiar from thermodynamics. The difference is now we have a more rigorous definition of the entropy. When the total entropy S is maximized, we have that $T_1 = T_2$. Once again, two systems in thermal contact and can exchange energy will in equilibrium 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

$$\begin{aligned} S_{\text{NR}} &= \tfrac{1}{2}Nd\,k_{\text{B}} \ln\bigg(\frac{E}{N}\bigg) + Nk_{\text{B}} \ln\bigg(\frac{V}{N}\bigg) + \text{const.} \\ S_{\text{UR}} &= Nd\,k_{\text{B}} \ln\bigg(\frac{E}{N}\bigg) + Nk_{\text{B}} \ln\bigg(\frac{V}{N}\bigg) + \text{const.} \quad . \end{aligned} \tag{4.39}$$

Invoking eqn. 4.38, we then have $E_{\text{NR}} = \frac{1}{2}Nd\,k_{\text{B}}T$ and $E_{\text{UR}} = Nd\,k_{\text{B}}T$.

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

$$\ln P_1(E_1^* + \Delta E_1) = \ln P_1(E_1^*) + \frac{1}{2k_{\text{B}}} \frac{\partial^2 S_1}{\partial E_1^2}\bigg|_{E_1^*} (\Delta E_1)^2 + \frac{1}{2k_{\text{B}}} \frac{\partial^2 S_2}{\partial E_2^2}\bigg|_{E_2^*} (\Delta E_1)^2 + \ldots \quad , \tag{4.40}$$

where $E_2^* = E - E_1^*$. We must now evaluate

$$\frac{\partial^2 S}{\partial E^2} = \frac{\partial}{\partial E}\bigg(\frac{1}{T}\bigg) = -\frac{1}{T^2}\bigg(\frac{\partial T}{\partial E}\bigg)_{V,N} = -\frac{1}{T^2\,C_V} \quad , \tag{4.41}$$

where $C_V = \big(\partial E/\partial T\big)_{V,N}$ is the heat capacity. Thus,

$$P_1 = P_1^* \, e^{-(\Delta E_1)^2/2k_B T^2 \bar{C}_V} \quad , \tag{4.42}$$

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 ΔE_1 can now be computed:

$$\big\langle (\Delta E_1)^2 \big\rangle = k_B T^2 \, \bar{C}_V \qquad \Longrightarrow \qquad (\Delta E_1)_{\text{RMS}} = T \sqrt{k_B \bar{C}_V} \quad . \tag{4.44}$$

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 ΔE_1 are proportional to the *square root* of the system size, whereas E_1 itself is extensive. Thus, the ratio $(\Delta E_1)_{\text{RMS}}/E_1 \propto V^{-1/2}$ scales as the inverse square root of the volume. The distribution $P_1(E_1)$ is thus *extremely sharp*.

The full distribution function for the energy within the microcanonical ensemble is

$$P(\mathcal{E}) = \big\langle \delta(\mathcal{E} - \hat{H}) \big\rangle = \frac{\mathsf{Tr}\; \delta(\mathcal{E} - \hat{H}) \, e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} = \frac{1}{Z} \, D(\mathcal{E}) \, e^{-\beta \mathcal{E}} \quad . \tag{4.45}$$

Thus,

$$P(\mathcal{E}) = \frac{e^{-\beta[\mathcal{E} - TS(\mathcal{E})]}}{\int \! d\mathcal{E}' \, e^{-\beta[\mathcal{E}' - TS(\mathcal{E}')]}} \quad , \tag{4.46}$$

where $S(\mathcal{E}) = k_B \ln D(\mathcal{E})$ 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

$$S(E + \delta\mathcal{E}) = S(E) + \frac{\delta\mathcal{E}}{T} - \frac{\big(\delta\mathcal{E}\big)^2}{2T^2 \, C_V} + \ldots \tag{4.47}$$

Recall that $S''(E) = \partial(T^{-1})/\partial E = -1/T^2 C_V$. Thus,

$$\mathcal{E} - T\,S(\mathcal{E}) = E - T\,S(E) + \frac{(\delta\mathcal{E})^2}{2T\,C_V} + \mathcal{O}\big((\delta\mathcal{E})^3\big) \quad . \tag{4.48}$$

Applying this to both numerator and denominator of eqn. 4.46, we obtain[6]

$$P(\mathcal{E}) = \mathcal{N} \exp\!\left[- \frac{(\delta\mathcal{E})^2}{2k_B T^2 \, C_V} \right] \quad , \tag{4.49}$$

where $\mathcal{N} = (2\pi k_B T^2 C_V)^{-1/2}$ 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 $(\Delta\mathcal{E})_{\text{RMS}} = \sqrt{k_B T^2 \, C_V}$. This is a consequence of the Central Limit Theorem.

---

[6] In applying eqn. 4.48 to the denominator of eqn. 4.46, 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.

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

First, let us consider that our system S will in general be in contact with a world W. We call the union of S and W the universe, $U = W \cup S$. Let $|\, w \,\rangle$ denote a quantum mechanical state of W, and let $|\, s \,\rangle$ denote a quantum mechanical state of S. Then the most general wavefunction we can write is of the form

$$|\, \Psi \,\rangle = \sum_{w,s} \Psi_{w,s} \, |\, s \,\rangle \otimes |\, s \,\rangle \quad . \tag{4.50}$$

Now let us compute the expectation value of some operator $\hat{\mathcal{A}}$ which acts as the identity within W, meaning $\langle\, w \,|\, \hat{\mathcal{A}} \,|\, w' \,\rangle = \hat{A}\, \delta_{ww'}$, where $\hat{A}$ is the 'reduced' operator acting within S alone. We then have

$$\langle\, \Psi \,|\, \hat{\mathcal{A}} \,|\, \Psi \,\rangle = \sum_{w,w'} \sum_{s,s'} \Psi^*_{w,s} \, \Psi_{w',s'} \, \delta_{ww'} \, \langle\, s \,|\, \hat{A} \,|\, s' \,\rangle = \mathsf{Tr}\,\big(\hat{\varrho}\, \hat{A}\big) \quad , \tag{4.51}$$

where

$$\hat{\varrho} = \sum_{w} \sum_{s,s'} \Psi^*_{w,s} \, \Psi_{w,s'} \, |\, s' \,\rangle \langle\, s \,| \tag{4.52}$$

is the *density matrix* for S. The time-dependence of $\hat{\varrho}$ is easily found:

$$\hat{\varrho}(t) = \sum_{w} \sum_{s,s'} \Psi^*_{w,s} \, \Psi_{w,s'} \, |\, s'(t) \,\rangle \langle\, s(t) \,| = e^{-i\hat{H}t/\hbar} \, \hat{\varrho} \, e^{+i\hat{H}t/\hbar} \quad , \tag{4.53}$$

where $\hat{H}$ is the Hamiltonian for the system S. Thus, we find

$$i\hbar \, \frac{\partial \hat{\varrho}}{\partial t} = \big[\hat{H}, \hat{\varrho}\big] \quad . \tag{4.54}$$

Note that the density matrix evolves according to a slightly different equation than an operator in the Heisenberg picture, for which

$$\hat{A}(t) = e^{+iHt/\hbar} \, A \, e^{-i\hat{H}t/\hbar} \qquad \Longrightarrow \qquad i\hbar \, \frac{\partial \hat{A}}{\partial t} = \big[\hat{A}, \hat{H}\big] = -\big[\hat{H}, \hat{A}\big] \quad . \tag{4.55}$$

For Hamiltonian systems, we found that the phase space distribution $\varrho(q,p,t)$ evolved according to the Liouville equation, $i\, \partial\varrho/\partial t = L\, \varrho$, where the Liouvillian L is the differential operator

$$L = -i \sum_{j=1}^{Nd} \left\{ \frac{\partial \hat{H}}{\partial p_j} \, \frac{\partial}{\partial q_j} - \frac{\partial \hat{H}}{\partial q_j} \, \frac{\partial}{\partial p_j} \right\} \quad . \tag{4.56}$$

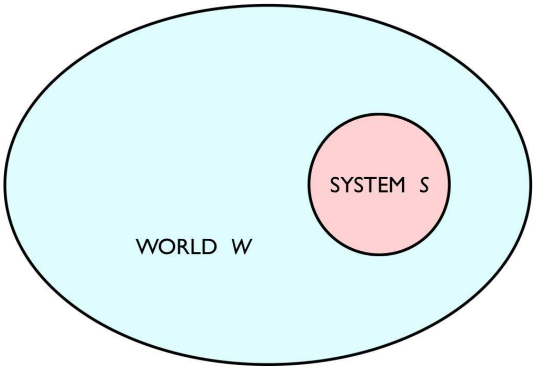


Figure 4.3: A system S in contact with a 'world' W. The union of the two, universe $U = W \cup S$, is said to be the 'universe'.

Accordingly, any distribution $\varrho(\Lambda_1, \ldots, \Lambda_k)$ which is a function of constants of the motion $\Lambda_a(q,p)$ is a stationary solution to the Liouville equation: $\partial_t \, \varrho(\Lambda_1, \ldots, \Lambda_k) = 0$. Similarly, any quantum mechanical density matrix which commutes with the Hamiltonian is a stationary solution to eqn. 4.54. The corresponding microcanonical distribution is $\hat{\varrho}_E = \delta\big(E - \hat{H}\big)$.

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

$$D(E) = \mathsf{Tr}\;\delta\big(E - \hat{H}\big) = \sum_l \delta(E - E_l) \quad , \tag{4.57}$$

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 ΔE:

$$\overline{D(E)} = \frac{1}{\Delta E} \int\limits_E^{E+\Delta E} \! dE' \, D(E') \quad . \tag{4.58}$$

If we take the limit $\Delta E \to 0$ but with $\Delta E \gg \delta E$, where δE 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 $\delta E \sim 1/D(E) = \exp\big[-N\phi(\varepsilon, v)\big]$ 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 ΔE.

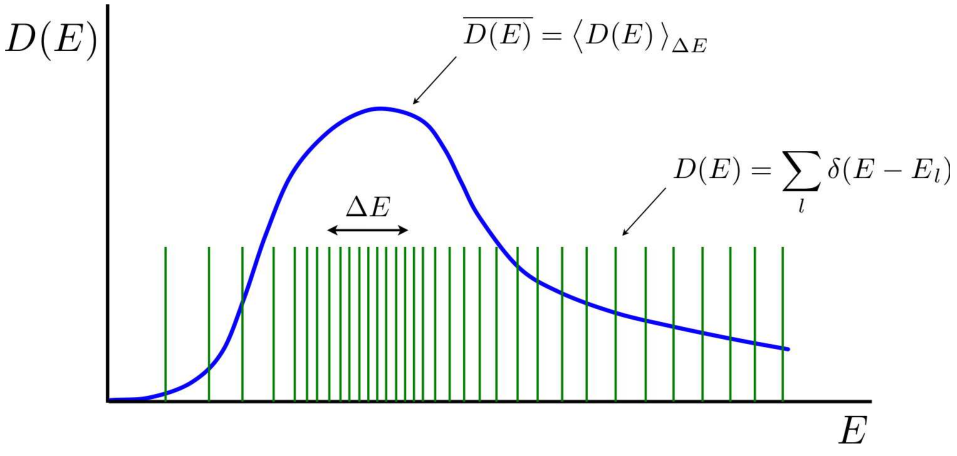


Figure 4.4: Averaging the quantum mechanical discrete density of states yields a continuous curve.

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

$$\hat{H}_0 = \frac{p^2}{2m} + \tfrac{1}{2} m\,\omega_0^2\, q^2 = \hbar\omega_0\left(a^\dagger a + \tfrac{1}{2}\right) \quad , \tag{4.59}$$

where a and $a^\dagger$ are *ladder operators* satisfying $\left[a, a^\dagger\right] = 1$, which can be taken to be

$$a = \ell\,\frac{\partial}{\partial q} + \frac{q}{2\ell} \qquad , \qquad a^\dagger = -\ell\,\frac{\partial}{\partial q} + \frac{q}{2\ell} \quad , \tag{4.60}$$

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

$$\psi_0(q) = (2\pi\ell^2)^{-1/4}\, e^{-q^2/4\ell^2} \quad . \tag{4.61}$$

The normalized *coherent state* $|\,z\,\rangle$ is defined as

$$|\,z\,\rangle = e^{-\frac{1}{2}|z|^2}\, e^{z a^\dagger}\,|\,0\,\rangle = e^{-\frac{1}{2}|z|^2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}}\,|\,n\,\rangle \quad . \tag{4.62}$$

The overlap of coherent states is given by

$$\langle\, z_1 \,|\, z_2 \,\rangle = e^{-\frac{1}{2}|z_1|^2}\, e^{-\frac{1}{2}|z_2|^2}\, e^{\bar{z}_1 z_2} \quad , \tag{4.63}$$

hence different coherent states are not orthogonal. Despite this nonorthogonality, the coherent states allow a simple resolution of the identity,

$$1 = \int\!\frac{d^2\!z}{2\pi i}\,|\,z\,\rangle\langle\,z\,| \quad ; \quad \frac{d^2\!z}{2\pi i} \equiv \frac{d\,\text{Re}z\; d\,\text{Im}z}{\pi} \tag{4.64}$$

which is straightforward to establish.

To gain some physical intuition about the coherent states, define

$$z \equiv \frac{Q}{2\ell} + \frac{i\ell P}{\hbar} \tag{4.65}$$

and write $|\, z \,\rangle \equiv |\, Q, P \,\rangle$. One finds (*exercise!*)

$$\psi_{Q,P}(q) = \langle\, q \,|\, z \,\rangle = (2\pi\ell^2)^{-1/4} \, e^{-iPQ/2\hbar} \, e^{iPq/\hbar} \, e^{-(q-Q)^2/4\ell^2} \quad , \tag{4.66}$$

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

$$\langle\, Q, P \,|\, q \,|\, Q, P \,\rangle = \langle\, z \,|\, \ell\,(a + a^\dagger) \,|\, z \,\rangle = 2\ell \, \mathrm{Re}\, z = Q \tag{4.67}$$

$$\langle\, Q, P \,|\, p \,|\, Q, P \,\rangle = \langle\, z \,|\, \frac{\hbar}{2i\ell}\,(a - a^\dagger) \,|\, z \,\rangle = \frac{\hbar}{\ell} \, \mathrm{Im}\, z = P \tag{4.68}$$

as well as

$$\langle\, Q, P \,|\, q^2 \,|\, Q, P \,\rangle = \langle\, z \,|\, \ell^2\,(a + a^\dagger)^2 \,|\, z \,\rangle = Q^2 + \ell^2 \tag{4.69}$$

$$\langle\, Q, P \,|\, p^2 \,|\, Q, P \,\rangle = -\langle\, z \,|\, \frac{\hbar^2}{4\ell^2}(a - a^\dagger)^2 \,|\, z \,\rangle = P^2 + \frac{\hbar^2}{4\ell^2} \quad . \tag{4.70}$$

Thus, the root mean square fluctuations in the coherent state $|\, Q, P \,\rangle$ are

$$\Delta q = \ell = \sqrt{\frac{\hbar}{2m\omega_0}} \qquad , \qquad \Delta p = \frac{\hbar}{2\ell} = \sqrt{\frac{m\hbar\omega_0}{2}} \quad , \tag{4.71}$$

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

$$\langle\, Q, P \,|\, \hat{A}(q,p) \,|\, Q, P \,\rangle = A(Q,P) + \mathcal{O}(\hbar) \quad , \tag{4.72}$$

where $A(Q,P)$ is formed from $\hat{A}(q,p)$ by replacing $q \to Q$ and $p \to P$.

Since

$$\frac{d^2 z}{2\pi i} \equiv \frac{d\,\mathrm{Re} z \; d\,\mathrm{Im} z}{\pi} = \frac{dQ\, dP}{2\pi\hbar} \quad , \tag{4.73}$$

we can write the trace using coherent states as

$$\mathsf{Tr}\, \hat{A} = \frac{1}{2\pi\hbar} \int\limits_{-\infty}^{\infty} \! dQ \int\limits_{-\infty}^{\infty} \! dP \, \langle\, Q, P \,|\, \hat{A} \,|\, Q, P \,\rangle \quad . \tag{4.74}$$

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 ω_0 is arbitrary in our discussion. By increasing ω_0, the states become more localized in q and more plane wave like in p. However, so long as ω_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

$$\begin{aligned} P_n &= \lim_{\Delta E \to 0} \frac{D_{\rm W}(E_{\rm U} - E_n)\,\Delta E}{D_{\rm U}(E_{\rm U})\,\Delta E} \\ &= \frac{\text{\# of states accessible to } W \text{ given that } E_{\rm S} = E_n}{\text{total \# of states in } U} \quad . \end{aligned} \tag{4.75}$$

Then

$$\begin{aligned} \ln P_n &= \ln D_{\rm W}(E_{\rm U} - E_n) - \ln D_{\rm U}(E_{\rm U}) \\ &= \ln D_{\rm W}(E_{\rm U}) - \ln D_{\rm U}(E_{\rm U}) - E_n\, \frac{\partial \ln D_{\rm W}(E)}{\partial E}\bigg|_{E=E_{\rm U}} + \ldots \quad . \end{aligned} \tag{4.76}$$

The higher order terms are negligible if $\mathsf{vol}(S) \ll \mathsf{vol}(W)$. In this case we have $\ln P_n = -\alpha - \beta E_n$, with

$$\beta = \frac{\partial \ln D_{\rm W}(E)}{\partial E}\bigg|_{E=E_{\rm U}} \equiv \frac{1}{k_{\rm B}T} \quad . \tag{4.77}$$

The constant α is fixed by the requirement that $\sum_n P_n = 1$, and thus we obtain

$$P_n = \frac{1}{Z}\, e^{-E_n/k_{\rm B}T} \quad , \qquad Z(T,V,N) = \mathsf{Tr}\; e^{-\beta \hat{H}} = \sum_n e^{-E_n/k_{\rm B}T} \quad . \tag{4.78}$$

We define the Helmholtz free energy $F(T,V,N)$ as

$$F(T,V,N) = -k_{\rm B}T \ln Z(T,V,N) \quad . \tag{4.79}$$

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

$$\hat{\varrho} = \frac{e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} \quad , \tag{4.80}$$

which is known as the *Gibbs distribution*. Note that $\big[\hat{\varrho}, \hat{H}\big] = 0$, 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

$$P(E_{\rm S}, E_{\rm W}) = D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm W})\, \delta(E_{\rm U} - E_{\rm S} - E_{\rm W})/D_{\rm U}(E_{\rm U}) \quad , \tag{4.81}$$

where

$$D_{\rm U}(E_{\rm U}) = \int\limits_{-\infty}^{\infty} \!dE_{\rm S}\; D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm U} - E_{\rm S}) \quad , \tag{4.82}$$

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

$$P(E_{\rm S}) = \frac{D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm U} - E_{\rm S})}{D_{\rm U}(E_{\rm U})} \quad , \tag{4.83}$$

as we have in eqn. 4.75. Suppose we wish to know the probability P_n that the system is in a *particular* state $|\, n\,\rangle$ with energy E_n. Clearly

$$P_n = \lim_{\Delta E \to 0} \frac{\text{probability that } E_{\rm S} \in [E_n, E_n + \Delta E]}{\#\text{ of S states with } E_{\rm S} \in [E_n, E_n + \Delta E]} = \frac{P(E_n)\, \Delta E}{D_{\rm S}(E_n)\, \Delta E} = \frac{D_{\rm W}(E_{\rm U} - E_n)}{D_{\rm U}(E_{\rm U})} \quad . \tag{4.84}$$

4.4.3 Additional remarks

The formula of eqn. 4.75 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

$$\begin{aligned} D_{\rm W}(E_{\rm U} - E_n)\, \Delta E &\to e^{S_{\rm W}(E_{\rm U} - E_n\, , \Delta E)} \equiv \mathop{\mathsf{Tr}}_{\rm W} \int\limits_{E_{\rm U} - E_n}^{E_{\rm U} - E_n + \Delta E} \!\! dE\; \delta(E - \hat{H}_{\rm W}) \\ D_{\rm U}(E_{\rm U})\, \Delta E &\to e^{S_{\rm U}(E_{\rm U}\, , \Delta E)} \equiv \mathop{\mathsf{Tr}}_{\rm U} \int\limits_{E_{\rm U}}^{E_{\rm U} + \Delta E} \!\! dE\; \delta(E - \hat{H}_{\rm U}) \quad . \end{aligned} \tag{4.85}$$

The off-diagonal matrix elements of $\varrho_{\rm S}$ are negligible in the thermodynamic limit.

---

[7] See T.-C. Lu and T. Grover, arXiv 1709.08784.

4.4.4 Averages within the OCE

To compute averages within the OCE,

$$\langle \hat{A} \rangle = \mathsf{Tr}\,\big(\hat{\varrho}\,\hat{A}\big) = \frac{\sum_n \langle n|\hat{A}|n\rangle \, e^{-\beta E_n}}{\sum_j e^{-\beta E_j}} \quad , \tag{4.86}$$

where we have conveniently taken the trace in a basis of energy eigenstates. In the classical limit, we have

$$\varrho(\boldsymbol{\varphi}) = \frac{1}{Z}\, e^{-\beta \hat{H}(\boldsymbol{\varphi})} \quad , \quad Z = \mathsf{Tr}\; e^{-\beta \hat{H}} = \int \! d\mu \, e^{-\beta \hat{H}(\boldsymbol{\varphi})} \quad , \tag{4.87}$$

with $d\mu = \frac{1}{N!}\prod_{j=1}^N (d^dq_j \, d^dp_j/h^d)$ for identical particles ('Maxwell-Boltzmann statistics'). Thus,

$$\langle A \rangle = \mathsf{Tr}\,(\varrho A) = \frac{\int \! d\mu \; A(\boldsymbol{\varphi}) \, e^{-\beta \hat{H}(\boldsymbol{\varphi})}}{\int \! d\mu \, e^{-\beta \hat{H}(\boldsymbol{\varphi})}} \quad . \tag{4.88}$$

4.4.5 Entropy and free energy

The *Boltzmann entropy* is defined by

$$S = -k_{\rm B}\, \mathsf{Tr}\,\big(\hat{\varrho}\ln\hat{\varrho}\big) = -k_{\rm B} \sum_n P_n \, \ln P_n \quad . \tag{4.89}$$

The Boltzmann entropy and the statistical entropy $S = k_{\rm B} \ln D(E)$ are identical in the thermodynamic limit. Since $\ln P_n = \beta(F - E_n)$, we have

$$S = -k_{\rm B} \sum_n P_n \, \big(\beta F - \beta E_n\big) = -\frac{F}{T} + \frac{\langle \, \hat{H} \, \rangle}{T} \quad , \tag{4.90}$$

which is to say $F = E - TS$, where

$$E = \sum_n P_n \, E_n = \frac{\mathsf{Tr}\; \hat{H}\, e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} \tag{4.91}$$

is the average energy. We also see that

$$Z = \mathsf{Tr}\; e^{-\beta \hat{H}} = \sum_n e^{-\beta E_n} \quad \Longrightarrow \quad E = \frac{\sum_n E_n \, e^{-\beta E_n}}{\sum_j e^{-\beta E_j}} = -\frac{\partial}{\partial\beta}\,\ln Z = \frac{\partial}{\partial\beta}\big(\beta F\big) \quad . \tag{4.92}$$

Thus, $F(T,V,N)$ is a Legendre transform of $E(S,V,N)$, with

$$dF = -S\,dT - p\,dV + \mu\,dN \quad , \tag{4.93}$$

which means

$$S = -\bigg(\frac{\partial F}{\partial T}\bigg)_{V,N} \quad , \quad p = -\bigg(\frac{\partial F}{\partial V}\bigg)_{T,N} \quad , \quad \mu = +\bigg(\frac{\partial F}{\partial N}\bigg)_{T,V} \quad . \tag{4.94}$$

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

$$\begin{aligned} -\frac{\partial E}{\partial \beta} = k_{\rm B} T^2 \, \frac{\partial E}{\partial T} &= \frac{\partial^2 \ln Z}{\partial \beta^2} \\ &= \frac{\textsf{Tr} \; \hat{H}^2 \, e^{-\beta \hat{H}}}{\textsf{Tr} \; e^{-\beta \hat{H}}} - \left(\frac{\textsf{Tr} \; \hat{H} \, e^{-\beta \hat{H}}}{\textsf{Tr} \; e^{-\beta \hat{H}}} \right)^{\!2} \\ &= \big\langle \hat{H}^2 \big\rangle - \big\langle \hat{H} \big\rangle^2 \quad . \end{aligned} \tag{4.95}$$

Thus, the heat capacity is related to the fluctuations in the energy, just as we saw in §4.2.6:

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{V,N} = \frac{1}{k_{\rm B} T^2} \Big(\big\langle \hat{H}^2 \big\rangle - \big\langle \hat{H} \big\rangle^2 \Big) \qquad \Rightarrow \qquad \frac{(\Delta E)_{\rm RMS}}{E} = \frac{\sqrt{k_{\rm B} T^2 \, C_V}}{E} \quad . \tag{4.96}$$

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

$$\left(\frac{(\Delta E)_{\rm RMS}}{E} \right)_{\rm NRIG} = \sqrt{\frac{2}{Nd}} \quad , \tag{4.97}$$

which scales as $N^{-1/2}$ and thus vanishes in the thermodynamic limit.

4.4.7 Thermodynamics revisited

The average energy within the OCE is

$$E = \sum_n E_n P_n \quad , \tag{4.98}$$

and therefore

$$dE = \sum_n E_n \, dP_n + \sum_n P_n \, dE_n = đQ - đW \quad , \tag{4.99}$$

where

$$đQ = \sum_n E_n \, dP_n \qquad , \qquad đW = -\sum_n P_n \, dE_n \quad . \tag{4.100}$$

Finally, from $P_n = Z^{-1} \, e^{-E_n / k_{\rm B} T}$, we can write

$$E_n = -k_{\rm B} T \ln Z - k_{\rm B} T \ln P_n \quad , \tag{4.101}$$

with which we obtain

$$\begin{aligned} đQ = \sum_n E_n \, dP_n &= -k_{\rm B} T \ln Z \sum_n dP_n - k_{\rm B} T \sum_n \ln P_n \, dP_n \\ &= T \, d\Big(- k_{\rm B} \sum_n P_n \ln P_n \Big) = T \, dS \quad . \end{aligned} \tag{4.102}$$

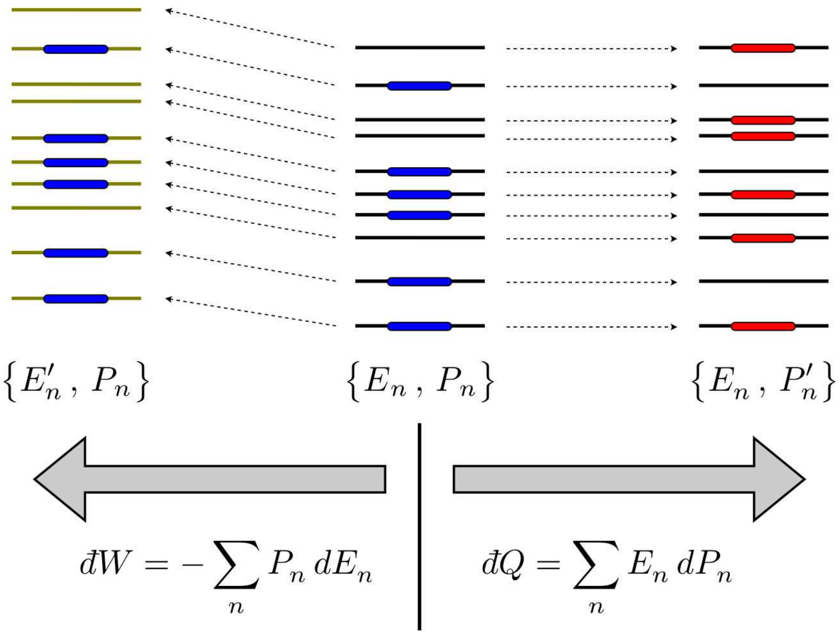


Figure 4.5: Microscopic, statistical interpretation of the First Law of Thermodynamics.

Note also that

$$\begin{aligned}\text{đ}W = -\sum_n P_n\,dE_n &= -\sum_n P_n\bigg(\sum_i \frac{\partial E_n}{\partial X_i}\,dX_i\bigg)\\ &= -\sum_{n,i} P_n\,\langle\, n\,|\,\frac{\partial \hat{H}}{\partial X_i}\,|\,n\,\rangle\,dX_i \equiv \sum_i F_i\,dX_i \quad ,\end{aligned} \tag{4.103}$$

so the generalized force F_i conjugate to the generalized displacement dX_i is

$$F_i = -\sum_n P_n\,\frac{\partial E_n}{\partial X_i} = -\bigg\langle \frac{\partial \hat{H}}{\partial X_i}\bigg\rangle \quad . \tag{4.104}$$

This is the force acting *on* the system[8]. 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. 4.99 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 $(\Delta E)_{\text{I}} = -W$ 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: $(\Delta E)_{\text{II}} = Q$. This provides us with a statistical and microscopic interpretation of the First Law of Thermodynamics.

---

[8] In deriving eqn. 4.104, 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.

4.4.8 Generalized susceptibilities

Suppose our Hamiltonian is of the form

$$\hat{H} = \hat{H}(\lambda) = \hat{H}_0 - \lambda\,\hat{Q} \quad , \tag{4.105}$$

where λ is an intensive parameter, such as magnetic field. Then $Z(\lambda) = \mathsf{Tr}\; e^{-\beta(\hat{H}_0 - \lambda\hat{Q})}$ and

$$\frac{1}{Z}\,\frac{\partial Z}{\partial\lambda} = \beta\cdot\frac{1}{Z}\,\mathsf{Tr}\,\Big(\hat{Q}\,e^{-\beta\hat{H}(\lambda)}\Big) = \beta\,\langle\hat{Q}\rangle \quad . \tag{4.106}$$

But then from $Z = e^{-\beta F}$ we have

$$Q(\lambda,T) = \langle\,\hat{Q}\,\rangle = -\bigg(\frac{\partial F}{\partial\lambda}\bigg)_T \quad . \tag{4.107}$$

Typically we will take Q to be an extensive quantity. We can now define the *susceptibility* χ as

$$\chi = \frac{1}{V}\,\frac{\partial Q}{\partial\lambda} = -\frac{1}{V}\,\frac{\partial^2 F}{\partial\lambda^2} \quad . \tag{4.108}$$

The volume factor in the denominator ensures that χ is intensive.

It is important to realize that we have assumed here that $\big[\hat{H}_0\,,\,\hat{Q}\big] = 0$, *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 $\big\{\hat{Q}_k\big\}$ satisfying $\big[\hat{Q}_k\,,\,\hat{Q}_{k'}\big] = 0$ and $\big[\hat{H}_0\,,\,\hat{Q}_k\big] = 0$, for all k and k'. Then for the Hamiltonian

$$\hat{H}(\boldsymbol{\lambda}) = \hat{H}_0 - \sum_k \lambda_k\,\hat{Q}_k \quad , \tag{4.109}$$

we have that

$$Q_k(\boldsymbol{\lambda},T) = \langle\,\hat{Q}_k\,\rangle = -\bigg(\frac{\partial F}{\partial\lambda_k}\bigg)_{T,\,\{N_a\},\,\{\lambda_{j\,(\neq k)}\}} \tag{4.110}$$

and we may define an entire matrix of susceptibilities,

$$\chi_{kl} \equiv \frac{1}{V}\,\frac{\partial Q_k}{\partial\lambda_l} = -\frac{1}{V}\,\frac{\partial^2 F}{\partial\lambda_k\,\partial\lambda_l} \quad . \tag{4.111}$$

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

$$\begin{aligned} P_n &= \lim_{\Delta E\to 0}\,\lim_{\Delta N\to 0}\,\frac{D_{\rm W}(E_{\rm U}-E_n\,,\,N_{\rm U}-N_n)\,\Delta E\,\Delta N}{D_{\rm U}(E_{\rm U},N_{\rm U})\,\Delta E\,\Delta N}\\ &= \frac{\text{\# of states accessible to } W \text{ given that } E_{\rm S}=E_n \text{ and } N_{\rm S}=N_n}{\text{total \# of states in } U}\qquad . \end{aligned} \tag{4.112}$$

Then

$$\begin{aligned} \ln P_n &= \ln D_{\rm W}(E_{\rm U}-E_n\,,\,N_{\rm U}-N_n) - \ln D_{\rm U}(E_{\rm U},N_{\rm U})\\ &= \ln D_{\rm W}(E_{\rm U},N_{\rm U}) - \ln D_{\rm U}(E_{\rm U},N_{\rm U})\\ &\qquad\qquad - E_n\,\frac{\partial \ln D_{\rm W}(E,N)}{\partial E}\bigg|_{\substack{E=E_{\rm U}\\ N=N_{\rm U}}} - N_n\,\frac{\partial \ln D_{\rm W}(E,N)}{\partial N}\bigg|_{\substack{E=E_{\rm U}\\ N=N_{\rm U}}} + \,\ldots \end{aligned} \tag{4.113}$$

and thus $\ln P_n = -\alpha - \beta E_n + \beta\mu N_n$, with

$$\beta = \frac{\partial \ln D_{\rm W}(E,N)}{\partial E}\bigg|_{\substack{E=E_{\rm U}\\ N=N_{\rm U}}} = \frac{1}{k_{\rm B}T}\qquad,\qquad \mu = -k_{\rm B}T\,\frac{\partial \ln D_{\rm W}(E,N)}{\partial N}\bigg|_{\substack{E=E_{\rm U}\\ N=N_{\rm U}}}\qquad . \tag{4.114}$$

The quantity μ 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 α is fixed by the requirement that $\sum_n P_n = 1$:

$$\begin{aligned} P_n &= \frac{1}{\Xi}\,e^{-\beta(E_n-\mu N_n)}\\ \Xi(\beta,V,\mu) &= \sum_n e^{-\beta(E_n-\mu N_n)} = \mathsf{Tr}\;e^{-\beta\hat K}\qquad . \end{aligned} \tag{4.115}$$

Thus, the quantum mechanical *grand canonical density matrix* is given by

$$\hat\varrho = \frac{e^{-\beta\hat K}}{\mathsf{Tr}\;e^{-\beta\hat K}}\qquad , \tag{4.116}$$

with $\hat K = \hat H - \mu\hat N$. Note that $\big[\hat\varrho,\hat K\big]=0$. 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:

$$\Omega(T,V,\mu) = -k_{\rm B}T\ln\Xi(T,V,\mu)\qquad\Longleftrightarrow\qquad \Xi = \exp(-\Omega/k_{\rm B}T)\quad , \tag{4.117}$$

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, $\ln P_n = (\Omega - E_n + \mu N_n)/k_{\rm B}T$.

If $\big[\hat{H},\hat{N}\big]=0$, the grand potential may be expressed as a sum over contributions from each N sector, *viz.*

$$\varXi(T,V,\mu)=\sum_N e^{\beta\mu N}\,Z(T,V,N) \quad . \tag{4.118}$$

When there is more than one species, we have several chemical potentials $\{\mu_a\}$, and accordingly we define

$$\hat{K}=\hat{H}-\sum_a \mu_a\,\hat{N}_a \quad , \tag{4.119}$$

with $\varXi=\mathsf{Tr}\;e^{-\beta\hat{K}}$ as before. To compute averages within the GCE, we use the grand canonical density matrix $\hat{\varrho}$:

$$\big\langle\hat{A}\big\rangle=\mathsf{Tr}\,\big(\hat{\varrho}\,\hat{A}\big)=\frac{\sum_n \langle n|\hat{A}|n\rangle\;e^{-\beta(E_n-\mu N_n)}}{\sum_j e^{-\beta(E_j-\mu N_j)}} \quad , \tag{4.120}$$

4.5.2 Entropy and Gibbs-Duhem relation

In the GCE, the Boltzmann entropy is

$$\begin{aligned}S=-k_{\rm B}\sum_n P_n\ln P_n&=-k_{\rm B}\sum_n P_n\Big(\beta\varOmega-\beta E_n+\beta\mu N_n\Big)\\ &=-\frac{\varOmega}{T}+\frac{\langle\hat{H}\rangle}{T}-\frac{\mu\,\langle\hat{N}\rangle}{T} \quad ,\end{aligned} \tag{4.121}$$

which says

$$\varOmega=E-TS-\mu N \quad , \tag{4.122}$$

where

$$E=\sum_n E_n\,P_n=\mathsf{Tr}\big(\hat{\varrho}\,\hat{H}\big) \qquad , \qquad N=\sum_n N_n\,P_n=\mathsf{Tr}\big(\hat{\varrho}\,\hat{N}\big) \quad . \tag{4.123}$$

Therefore, $\varOmega(T,V,\mu)$ is a double Legendre transform of $E(S,V,N)$, with

$$d\varOmega=-S\,dT-p\,dV-N d\mu \quad , \tag{4.124}$$

which entails

$$S=-\left(\frac{\partial\varOmega}{\partial T}\right)_{V,\mu} \qquad , \qquad p=-\left(\frac{\partial\varOmega}{\partial V}\right)_{T,\mu} \qquad , \qquad N=-\left(\frac{\partial\varOmega}{\partial\mu}\right)_{T,V} \quad . \tag{4.125}$$

Since $\varOmega(T,V,\mu)$ is an extensive quantity, we must be able to write $\varOmega=V\omega(T,\mu)$. We identify the function $\omega(T,\mu)$ as the negative of the pressure:

$$\frac{\partial\varOmega}{\partial V}=-\frac{k_{\rm B}T}{\varXi}\left(\frac{\partial\varXi}{\partial V}\right)_{T,\mu}=\frac{1}{\varXi}\sum_n\frac{\partial E_n}{\partial V}\,e^{-\beta(E_n-\mu N_n)}=\left(\frac{\partial E}{\partial V}\right)_{T,\mu}=-p(T,\mu) \quad . \tag{4.126}$$

Therefore, $\varOmega=-pV$, 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,

$$d\varOmega=-S\,dT-p\,dV-N d\mu=-p\,dV-V dp \quad\Rightarrow\quad S\,dT-V dp+N d\mu=0 \quad . \tag{4.127}$$

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

$$N = \langle \hat{N} \rangle = \frac{\mathsf{Tr}\ \hat{N}\, e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H}-\mu\hat{N})}} = \frac{1}{\beta}\frac{\partial}{\partial\mu}\ln \varXi = -\frac{\partial\varOmega}{\partial\mu} \quad . \tag{4.128}$$

Therefore,

$$\frac{1}{\beta}\frac{\partial N}{\partial\mu} = -\frac{1}{\beta}\frac{\partial^2\varOmega}{\partial\mu^2} = \frac{\mathsf{Tr}\ \hat{N}^2\, e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H}-\mu\hat{N})}} - \left(\frac{\mathsf{Tr}\ \hat{N}\, e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H}-\mu\hat{N})}}\right)^2 = \big\langle \hat{N}^2 \big\rangle - \big\langle \hat{N} \big\rangle^2 \quad , \tag{4.129}$$

and thus

$$\frac{\big\langle \hat{N}^2 \big\rangle - \big\langle \hat{N} \big\rangle^2}{\big\langle \hat{N} \big\rangle^2} = \frac{k_{\rm B}T}{N^2}\left(\frac{\partial N}{\partial\mu}\right)_{T,V} = \frac{k_{\rm B}T}{V}\,\kappa_T \quad , \tag{4.130}$$

where κ_T is the isothermal compressibility. Note:

$$\begin{aligned}\left(\frac{\partial N}{\partial\mu}\right)_{T,V} &= \frac{\partial(N,T,V)}{\partial(\mu,T,V)} = -\frac{\partial(N,T,V)}{\partial(V,T,\mu)}\\ &= -\frac{\partial(N,T,V)}{\partial(N,T,p)}\cdot \overbrace{\frac{\partial(N,T,p)}{\partial(V,T,p)}}^{N/V} \cdot \overbrace{\frac{\partial(V,T,p)}{\partial(N,T,\mu)}}^{1} \cdot \overbrace{\frac{\partial(N,T,\mu)}{\partial(V,T,\mu)}}^{N/V}\\ &= -\frac{N^2}{V^2}\left(\frac{\partial V}{\partial p}\right)_{T,N} = \frac{N^2}{V}\,\kappa_T \quad .\end{aligned} \tag{4.131}$$

We thus arrive at the result

$$\frac{(\Delta N)_{\rm RMS}}{N} = \sqrt{\frac{k_{\rm B}T\,\kappa_T}{V}} \quad , \tag{4.132}$$

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. 4.96.

4.5.4 Generalized susceptibilities in the GCE

We can appropriate the results from §4.4.8 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,

$$\hat{K} = \hat{H}_0 - \sum_a \mu_a\,\hat{N}_a - \sum_k \lambda_k\,\hat{Q}_k \quad . \tag{4.133}$$

We then have

$$\varOmega\big(T,V,\{\mu_a\},\{\lambda_k\}\big) = -k_{\rm B}T\,\mathsf{Tr}\ e^{-\hat{K}/k_{\rm B}T} \quad , \tag{4.134}$$

whence

$$d\Omega = -S\,dT - p\,dV + \sum_a N_a\,d\mu_a - \sum_k Q_k\,d\lambda_k \quad , \tag{4.135}$$

where the un-hatted quantities $\{N_a, Q_k\}$ are statistical averages within the GCE, *viz.*

$$N_a = \langle \hat{N}_a \rangle = -\frac{\partial \Omega}{\partial \mu_a} \quad , \quad Q_k = \langle \hat{Q}_k \rangle = -\frac{\partial \Omega}{\partial \lambda_k} \quad . \tag{4.136}$$

This leads to various generalized susceptibilities and cross-susceptibilities,

$$\frac{\partial N_a}{\partial \mu_b} = -\frac{\partial^2 \Omega}{\partial \mu_a \partial \mu_b} \quad , \quad \frac{\partial Q_k}{\partial \lambda_l} = -\frac{\partial^2 \Omega}{\partial \lambda_k \partial \lambda_l} \quad , \quad \frac{\partial N_a}{\partial \lambda_k} = \frac{\partial Q_k}{\partial \mu_a} = -\frac{\partial^2 \Omega}{\partial \mu_a \partial \lambda_k} \quad . \tag{4.137}$$

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

$$P_n = \lim_{\Delta E \to 0}\, \lim_{\Delta V \to 0} \frac{D_{\rm W}(E_{\rm U} - E_n\,,\, V_{\rm U} - V_n)\,\Delta E\,\Delta V}{D_{\rm U}(E_{\rm U}, V_{\rm U})\,\Delta E\,\Delta V} . \tag{4.138}$$

Then

$$\begin{aligned} \ln P_n &= \ln D_{\rm W}(E_{\rm U} - E_n\,,\, V_{\rm U} - V_n) - \ln D_{\rm U}(E_{\rm U}, V_{\rm U}) \qquad (4.139)\\ &= \ln D_{\rm W}(E_{\rm U}, V_{\rm U}) - \ln D_{\rm U}(E_{\rm U}, V_{\rm U}) - E_n\, \frac{\partial \ln D_{\rm W}(E,V)}{\partial E}\bigg|_{\substack{E=E_{\rm U}\\ V=V_{\rm U}}} - V_n\, \frac{\partial \ln D_{\rm W}(E,V)}{\partial V}\bigg|_{\substack{E=E_{\rm U}\\ V=V_{\rm U}}} + \ldots , \end{aligned}$$

and thus $\ln P_n = -\alpha - \beta E_n - \beta p\, V_n$, where the constants β and p are given by

$$\beta = \frac{\partial \ln D_{\rm W}(E,V)}{\partial E}\bigg|_{\substack{E=E_{\rm U}\\ V=V_{\rm U}}} = \frac{1}{k_{\rm B} T} \quad , \quad p = k_{\rm B} T\, \frac{\partial \ln D_{\rm W}(E,V)}{\partial V}\bigg|_{\substack{E=E_{\rm U}\\ V=V_{\rm U}}} \quad . \tag{4.140}$$

The corresponding partition function is

$$Y(T,p,N) = \mathsf{Tr}\; e^{-\beta(\hat{H} + pV)} = \frac{1}{V_0} \int\limits_0^\infty \! dV\, e^{-\beta p V}\, Z(T,V,N) \equiv e^{-\beta G(T,p,N)} \quad , \tag{4.141}$$

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 $G(V_0') = G(V_0) + k_{\rm B} T \ln(V_0'/V_0)$, 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.*

$$\begin{aligned} G &= E - TS + pV \\ dG &= -S\,dT + V\,dp + \mu\,dN \quad , \end{aligned} \tag{4.142}$$

which entails

$$S = -\bigg(\frac{\partial G}{\partial T}\bigg)_{p,N} \quad , \quad V = +\bigg(\frac{\partial G}{\partial p}\bigg)_{T,N} \quad , \quad \mu = +\bigg(\frac{\partial G}{\partial N}\bigg)_{T,p} \quad . \tag{4.143}$$

4.6 Statistical Ensembles from Maximum Entropy

The basic principle: maximize the Boltzmann entropy,

$$S = -k_{\rm B} \sum_n P_n \ln P_n \quad , \tag{4.144}$$

subject to a set of *constraints*. Constrained extremization using Lagrange's method of undetermined multipliers is reviews in the appendix, §4.10.

4.6.1 Microcanonical ensemble

We maximize S subject to the single constraint

$$C = \sum_n P_n - 1 = 0 \quad . \tag{4.145}$$

We implement the constraint $C = 0$ with a Lagrange multiplier, $\bar{\lambda} \equiv k_{\rm B}\,\lambda$, writing

$$S^* = S - k_{\rm B}\,\lambda\,C \quad , \tag{4.146}$$

and freely extremizing over the distribution $\{P_n\}$ and the Lagrange multiplier λ. Thus,

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B}\lambda\,\delta C - k_{\rm B}\,C\,\delta\lambda \\ &= -k_{\rm B} \sum_n \Big[\ln P_n + 1 + \lambda\Big]\delta P_n - k_{\rm B}\,C\,\delta\lambda \equiv 0 \quad . \end{aligned} \tag{4.147}$$

We conclude that $C = 0$ and that

$$\ln P_n = -\big(1 + \lambda\big) \quad , \tag{4.148}$$

and we fix λ by the normalization condition $\sum_n P_n = 1$. This gives $P_n = 1/\Omega$, with

$$\Omega = \sum_n \Theta(E + \Delta E - E_n)\,\Theta(E_n - E) \quad , \tag{4.149}$$

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

$$C_1 = \sum_n P_n - 1 = 0 \quad , \qquad C_2 = \sum_n E_n \, P_n - E = 0 \quad . \tag{4.150}$$

We now have two Lagrange multipliers. We write

$$S^* = S - k_\mathrm{B} \sum_{j=1}^{2} \lambda_j \, C_j \quad , \tag{4.151}$$

and we freely extremize over $\{P_n\}$ and $\{C_j\}$. We therefore have

$$\begin{aligned} \delta S^* &= \delta S - k_\mathrm{B} \sum_n \big(\lambda_1 + \lambda_2 \, E_n\big) \, \delta P_n - k_\mathrm{B} \sum_{j=1}^{2} C_j \, \delta\lambda_j \\ &= -k_\mathrm{B} \sum_n \Big[\ln P_n + 1 + \lambda_1 + \lambda_2 \, E_n\Big] \delta P_n - k_\mathrm{B} \sum_{j=1}^{2} C_j \, \delta\lambda_j \equiv 0 \quad . \end{aligned} \tag{4.152}$$

Thus, $C_1 = C_2 = 0$ and

$$\ln P_n = -\big(1 + \lambda_1 + \lambda_2 \, E_n\big) \quad . \tag{4.153}$$

We define $\lambda_2 \equiv \beta$ and we fix λ_1 by normalization. This yields

$$P_n = \frac{1}{Z} \, e^{-\beta E_n} \quad , \qquad Z = \sum_n e^{-\beta E_n} = \mathsf{Tr} \, e^{-\beta \hat{H}} \quad . \tag{4.154}$$

4.6.3 Grand canonical ensemble

We maximize S subject to the three constraints

$$C_1 = \sum_n P_n - 1 = 0 \quad , \quad C_2 = \sum_n E_n \, P_n - E = 0 \quad , \quad C_3 = \sum_n N_n \, P_n - N = 0 \quad . \tag{4.155}$$

We now have three Lagrange multipliers. We write

$$S^* = S - k_\mathrm{B} \sum_{j=1}^{3} \lambda_j \, C_j \quad , \tag{4.156}$$

and hence

$$\begin{aligned} \delta S^* &= \delta S - k_\mathrm{B} \sum_n \big(\lambda_1 + \lambda_2 \, E_n + \lambda_3 \, N_n\big) \, \delta P_n - k_\mathrm{B} \sum_{j=1}^{3} C_j \, \delta\lambda_j \\ &= -k_\mathrm{B} \sum_n \Big[\ln P_n + 1 + \lambda_1 + \lambda_2 \, E_n + \lambda_3 \, N_n\Big] \delta P_n - k_\mathrm{B} \sum_{j=1}^{3} C_j \, \delta\lambda_j \equiv 0 \quad . \end{aligned} \tag{4.157}$$

Thus, $C_1 = C_2 = C_3 = 0$ and

$$\ln P_n = -\big(1 + \lambda_1 + \lambda_2\, E_n + \lambda_3\, N_n\big) \quad . \tag{4.158}$$

We define $\lambda_2 \equiv \beta$ and $\lambda_3 \equiv -\beta\mu$, and we fix λ_1 by normalization. This yields

$$P_n = \frac{1}{\Xi}\, e^{-\beta(E_n - \mu N_n)} \qquad , \qquad \Xi = \sum_n e^{-\beta(E_n - \mu N_n)} = \mathsf{Tr}\; e^{-\beta(\hat{H} - \mu\hat{N})} \quad . \tag{4.159}$$

4.6.4 Generalized Gibbs ensembles

Suppose we have p constraints of the form $\langle \hat{Q}_k \rangle = Q_k$ with $k \in \{1, \ldots, \mathsf{p}\}$ and 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

$$\begin{aligned}\frac{1}{k_{\rm B}}\, S^*(P_1, \ldots, P_\Omega\,,\, \lambda\,,\, \eta_1\,,\, \ldots, \eta_{\mathsf{c}}\,,\, \theta_1, \ldots, \theta_{\mathsf{p}}) &= -\sum_n P_n \ln P_n - \lambda\bigg(\sum_n P_n - 1\bigg) \\ &\quad - \sum_{k=1}^{\mathsf{p}} \theta_k \bigg(\sum_n P_n\, Q_{k,n} - Q_k\bigg) - \sum_{a=1}^{\mathsf{c}} \eta_a \bigg(\sum_n P_n\, N_{a,n} - N_a\bigg) \quad ,\end{aligned} \tag{4.160}$$

where λ, $\{\theta_k\}$ and $\{\eta_a\}$ comprise $1+\mathsf{p}+\mathsf{c}$ undetermined Lagrange multipliers. The total number of states is taken to be Ω, *i.e.* $n \in \{1, \ldots, \Omega\}$. Setting the variation $\delta S^* = 0$, we obtain the following $\Omega + 1 + \mathsf{p} + \mathsf{c}$ equations:

$$\begin{aligned} 0 &= \frac{1}{k_{\rm B}} \frac{\partial S^*}{\partial P_n} = -\ln P_n - (1+\lambda) - \sum_{k=1}^{\mathsf{p}} \theta_k\, Q_{k,n} - \sum_{a=1}^{\mathsf{c}} \eta_a\, N_{a,n} \\ 0 &= \frac{1}{k_{\rm B}} \frac{\partial S^*}{\partial \lambda} = 1 - \sum_n P_n \\ 0 &= \frac{1}{k_{\rm B}} \frac{\partial S^*}{\partial \theta_k} = Q_k - \sum_n P_n\, Q_{k,n} \\ 0 &= \frac{1}{k_{\rm B}} \frac{\partial S^*}{\partial \eta_a} = N_a - \sum_n P_n\, N_{a,n} \quad . \end{aligned} \tag{4.161}$$

Thus, the probability distribution is given by

$$P_n(\boldsymbol{\theta}, \boldsymbol{\eta}) = \frac{1}{\mathcal{Z}(\boldsymbol{\theta}, \boldsymbol{\eta})}\, \exp\bigg\{ -\sum_{k=1}^{\mathsf{p}} \theta_k\, Q_{k,n} - \sum_{a=1}^{\mathsf{c}} \eta_a\, N_{a,n} \bigg\} \tag{4.162}$$

with

$$\mathcal{Z}(\boldsymbol{\theta}, \boldsymbol{\eta}) = \sum_n \exp\bigg\{ -\sum_{k=1}^{\mathsf{p}} \theta_k\, Q_{k,n} - \sum_{a=1}^{\mathsf{c}} \eta_a\, N_{a,n} \bigg\} \tag{4.163}$$

as well as

$$Q_k = \sum_n P_n(\boldsymbol{\theta}, \boldsymbol{\eta})\, Q_{k,n} \qquad , \qquad N_a = \sum_n P_n(\boldsymbol{\theta}, \boldsymbol{\eta})\, N_{a,n} \quad , \tag{4.164}$$

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

$$\hat{K}(\boldsymbol{\lambda},\boldsymbol{\mu}) \equiv \hat{H} - \sum_{k=2}^{\mathsf{p}} \lambda_k\, \hat{Q}_k - \sum_{a=1}^{\mathsf{c}} \mu_a\, \hat{N}_a \quad , \tag{4.165}$$

we have

$$\mathcal{Z}(T,\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathsf{Tr}\; e^{-\beta\hat{K}} = \sum_n e^{-K_n/k_{\rm B}T} \tag{4.166}$$

and

$$P_n(T,\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathcal{Z}^{-1}\, e^{-K_n/k_{\rm B}T} \qquad , \qquad K_n = \langle\, n \,|\, \hat{K} \,|\, n \,\rangle \quad . \tag{4.167}$$

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

$$\begin{aligned} Z(T,V,N) &= \frac{1}{N!}\prod_{i=1}^{N}\int \frac{d^d\!x_i\, d^d\!p_i}{(2\pi\hbar)^d}\, e^{-\beta \boldsymbol{p}_i^2/2m} \\ &= \frac{V^N}{N!}\left(\int\limits_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}\, e^{-\beta p^2/2m}\right)^{Nd} = \frac{1}{N!}\left(\frac{V}{\lambda_T^d}\right)^{N} \quad , \end{aligned} \tag{4.168}$$

where λ_T is the *thermal wavelength*,

$$\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T} \quad . \tag{4.169}$$

The physical interpretation of λ_T is that it is 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

$$\begin{aligned} \Xi(T,V,\mu) &= \sum_{N=0}^{\infty} e^{\beta\mu N}\, Z(T,V,N) \\ &= \sum_{N=1}^{\infty}\frac{1}{N!}\left(\frac{V e^{\mu/k_{\rm B}T}}{\lambda_T^d}\right)^{N} = \exp\left(\frac{V e^{\mu/k_{\rm B}T}}{\lambda_T^d}\right) \quad . \end{aligned} \tag{4.170}$$

From $\Xi = e^{-\Omega/k_{\rm B}T}$, we have the grand potential is

$$\Omega(T,V,\mu) = -Vk_{\rm B}T\, e^{\mu/k_{\rm B}T}\big/\lambda_T^d \quad . \tag{4.171}$$

Since $\Omega = -pV$ (see §4.5.2), we have

$$p(T,\mu) = k_{\rm B}T\, \lambda_T^{-d}\, e^{\mu/k_{\rm B}T} \quad . \tag{4.172}$$

The number density can also be calculated:

$$n = \frac{N}{V} = -\frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = \lambda_T^{-d}\, e^{\mu/k_\mathrm{B}T} \quad . \tag{4.173}$$

Combined, the last two equations recapitulate the ideal gas law, $pV = Nk_\mathrm{B}T$.

4.7.1 Maxwell velocity distribution

The distribution function for momenta is given by

$$g(\boldsymbol{p}) = \Big\langle \frac{1}{N}\sum_{i=1}^{N} \delta(\boldsymbol{p}_i - \boldsymbol{p})\Big\rangle \quad . \tag{4.174}$$

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 $\langle A\rangle = \mathsf{Tr}\,\big(Ae^{-\beta\hat{H}}\big)/\,\mathsf{Tr}\; e^{-\beta\hat{H}}$. Setting $i=1$, all the integrals other than that over $\boldsymbol{p}_1$ divide out between numerator and denominator. We then have

$$g(\boldsymbol{p}) = \frac{\int d^3\!p_1\, \delta(\boldsymbol{p}_1 - \boldsymbol{p})\, e^{-\beta\boldsymbol{p}_1^2/2m}}{\int d^3\!p_1\; e^{-\beta\boldsymbol{p}_1^2/2m}} = (2\pi m k_\mathrm{B}T)^{-3/2}\, e^{-\beta\boldsymbol{p}^2/2m} \quad . \tag{4.175}$$

Textbooks commonly refer to the *velocity distribution* $f(\boldsymbol{v})$, which is related to $g(\boldsymbol{p})$ by

$$f(\boldsymbol{v})\, d^3\!v = g(\boldsymbol{p})\, d^3\!p \quad . \tag{4.176}$$

Hence,

$$f(\boldsymbol{v}) = \left(\frac{m}{2\pi k_\mathrm{B}T}\right)^{3/2} e^{-m\boldsymbol{v}^2/2k_\mathrm{B}T} \quad . \tag{4.177}$$

This is known as the *Maxwell velocity distribution*. Note that the distributions are normalized, *viz.*

$$\int d^3\!p\; g(\boldsymbol{p}) = \int d^3\!v\; f(\boldsymbol{v}) = 1 \quad . \tag{4.178}$$

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

$$\tilde{f}(v) = 4\pi\, v^2 f(\boldsymbol{v}) = 4\pi\left(\frac{m}{2\pi k_\mathrm{B}T}\right)^{3/2} v^2\, e^{-mv^2/2k_\mathrm{B}T} \quad . \tag{4.179}$$

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 $v_0 = \sqrt{k_\mathrm{B}T/m}$, in which case

$$\tilde{f}(v) = \frac{1}{v_0}\,\varphi(v/v_0) \qquad , \qquad \varphi(s) = \sqrt{\tfrac{2}{\pi}}\, s^2\, e^{-s^2/2} \quad . \tag{4.180}$$

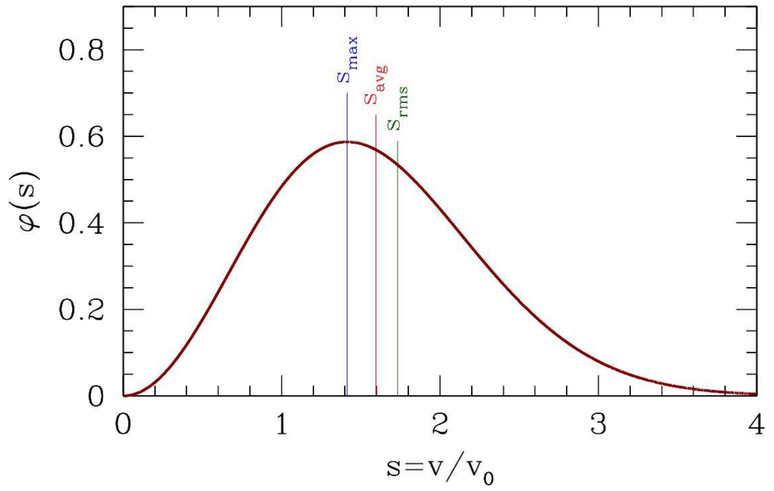


Figure 4.6: Maxwell distribution of speeds $\varphi(v/v_0)$. The most probable speed is $v_{\text{max}} = \sqrt{2}\,v_0$. The average speed is $v_{\text{avg}} = \sqrt{8/\pi}\,v_0$. The RMS speed is $v_{\text{rms}} = \sqrt{3}\,v_0$.

The distribution $\varphi(s)$ is shown in fig. 4.6. Computing averages, we have

$$C_k \equiv \langle s^k \rangle = \int\limits_0^\infty ds\, s^k\, \varphi(s) = \frac{2}{\sqrt{\pi}} \cdot 2^{k/2}\, \Gamma\left(\tfrac{3}{2} + \tfrac{k}{2}\right) \quad . \tag{4.181}$$

Thus, $C_0 = 1$, $C_1 = \sqrt{8/\pi}$, $C_2 = 3$, *etc.* The speed averages are

$$\left\langle v^k \right\rangle = C_k \left(\frac{k_{\rm B} T}{m}\right)^{k/2} \quad . \tag{4.182}$$

Note that the average *velocity* is $\langle \boldsymbol{v} \rangle = 0$, but the average *speed* is $\langle v \rangle = \sqrt{8 k_{\rm B} T/\pi m}$. The speed distribution is plotted in fig. 4.6.

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 ξ 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 ξ 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:

$$\zeta(\beta) \equiv \int_{-\infty}^{\infty} du\, e^{-\beta K u^2/2} = \left(\frac{2\pi}{K\beta}\right)^{1/2} \quad , \tag{4.183}$$

where $\beta = 1/k_{\rm B}T$. The contribution to the Helmholtz free energy is then

$$\Delta F_u = -k_{\rm B}T \ln \zeta = \tfrac{1}{2} k_{\rm B}T \ln\left(\frac{K}{2\pi k_{\rm B}T}\right) \quad , \tag{4.184}$$

and therefore the contribution to the internal energy E is

$$\Delta E_u = -\frac{\partial \ln \zeta}{\partial \beta} = \tfrac{1}{2} k_{\rm B}T \quad . \tag{4.185}$$

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 $E = \frac{1}{2} f N k_{\rm B}T$, and $C_V = \frac{1}{2} N k_{\rm B}$. 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

$$\zeta(\beta) = 2\int_0^{\infty} du\, e^{-\beta K u^\sigma/2} = \frac{2^{1+\sigma^{-1}}\,\Gamma(1/\sigma)}{\sigma K^{1/\sigma}}\, \beta^{-1/\sigma} \tag{4.186}$$

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

$$\hat{H} = -\mu_0 H \sum_{j=1}^{N_{\rm s}} \sigma_j \quad , \tag{4.187}$$

where now we write $\hat{H}$ for the Hamiltonian, to distinguish it from the external magnetic field H, and μ_0 is the magnetic moment per particle. We treat this system within the ordinary canonical ensemble. The partition function is

$$Z = \sum_{\sigma_1} \cdots \sum_{\sigma_{N_\mathsf{s}}} e^{-\beta \hat{H}} = \zeta^{N_\mathsf{s}} \quad , \tag{4.188}$$

where ζ is the single particle partition function:

$$\zeta = \sum_{\sigma=\pm 1} e^{\mu_0 H \sigma / k_\mathrm{B} T} = 2 \cosh\!\left(\frac{\mu_0 H}{k_\mathrm{B} T}\right) \quad . \tag{4.189}$$

The Helmholtz free energy is then

$$F(T, H, N_\mathsf{s}) = -k_\mathrm{B} T \ln Z = -N_\mathsf{s} \, k_\mathrm{B} T \ln\!\left[2 \cosh\!\left(\frac{\mu_0 H}{k_\mathrm{B} T}\right)\right] \quad . \tag{4.190}$$

The magnetization is

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T,\, N_\mathsf{s}} = N_\mathsf{s} \, \mu_0 \tanh\!\left(\frac{\mu_0 H}{k_\mathrm{B} T}\right) \quad . \tag{4.191}$$

The energy is

$$E = \frac{\partial}{\partial \beta} \left(\beta F\right) = -N_\mathsf{s} \, \mu_0 H \tanh\!\left(\frac{\mu_0 H}{k_\mathrm{B} T}\right) \quad . \tag{4.192}$$

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 σ is

$$P_\sigma = \frac{e^{\beta \mu_0 H \sigma}}{e^{\beta \mu_0 H} + e^{-\beta \mu_0 H}} \quad . \tag{4.193}$$

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:

$$\langle \sigma \rangle = P_+ - P_- = \tanh\!\left(\frac{\mu_0 H}{k_\mathrm{B} T}\right) \quad . \tag{4.194}$$

At low temperatures $T \ll \mu_0 H / k_\mathrm{B}$, we have $P_\uparrow \approx 1 - e^{-2\mu_0 H / k_\mathrm{B} T}$. At high temperatures $T > \mu_0 H / k_\mathrm{B}$, the two polarizations are equally likely, and $P_\sigma \approx \frac{1}{2}\left(1 + \frac{\sigma \mu_0 H}{k_\mathrm{B} T}\right)$.

The *isothermal magnetic susceptibility* is defined as

$$\chi_T = \frac{1}{N} \left(\frac{\partial M}{\partial H}\right)_T = \frac{\mu_0^2}{k_\mathrm{B} T} \, \mathrm{sech}^2\!\left(\frac{\mu_0 H}{k_\mathrm{B} T}\right) \quad . \tag{4.195}$$

(Typically this is computed per unit volume rather than per particle.) At $H = 0$, we have $\chi_T = \mu_0^2 / k_\mathrm{B} T$, 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 $\mathcal{E} = \mathcal{E}(S, M, N_\mathsf{s})$. It is obtained from the free energy $F(T, H, N_\mathsf{s})$ by a double Legendre transform:

$$\mathcal{E}(S, M, N_\mathsf{s}) = F(T, H, N_\mathsf{s}) + TS + HM \quad . \tag{4.196}$$

Now from eqn. 4.190 we derive the entropy

$$S = -\frac{\partial F}{\partial T} = N_\mathsf{s} k_\mathrm{B} \ln\left[2\cosh\left(\frac{\mu_0 H}{k_\mathrm{B} T}\right)\right] - \frac{N_\mathsf{s} \mu_0 H}{T} \tanh\left(\frac{\mu_0 H}{k_\mathrm{B} T}\right) \quad . \tag{4.197}$$

Thus, using eqns. 4.190 and 4.191, we obtain $\mathcal{E}(S, M, N_\mathsf{s}) = 0$.

The potential confusion here arises from our use of the expression $F(T, H, N_\mathsf{s})$. In thermodynamics, it is the Gibbs free energy $G(T, p, N_\mathsf{s})$ 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. 4.192 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_s spins, each of which can be either up $(+)$ or down $(-)$. Let N_σ be the number of sites with spin σ, where $\sigma = \pm 1$. Clearly $N_+ + N_- = N_\mathsf{s}$. 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

$$\Omega = \binom{N_\mathsf{s}}{N_+} = \frac{N_\mathsf{s}!}{N_+!\,N_-!} \quad . \tag{4.198}$$

Using Stirling's expression $\ln K! = K \ln K - K + \mathcal{O}(\ln K)$ for large K, we have

$$\begin{aligned} S = k_\mathrm{B} \ln \Omega &= (N_\mathsf{s} \ln N_\mathsf{s} - N_\mathsf{s}) - (N_+ \ln N_+ - N_+) - (N_- \ln N_- - N_-) \\ &= -N_\mathsf{s} k_\mathrm{B} \Big\{ x \ln x + (1-x) \ln(1-x) \Big\} \end{aligned} \tag{4.199}$$

in the thermodynamic limit: $N_\mathsf{s} \to \infty$, $N_+ \to \infty$, $x = N_+/N_\mathsf{s}$ constant. Now the magnetization is $M = (N_+ - N_-)\mu_0 = (2N_+ - N_\mathsf{s})\mu_0$, hence if we define the maximum energy $E_0 \equiv N\mu_0 H$, then

$$\frac{E}{E_0} = -\frac{M}{N\mu_0} = 1 - 2x \qquad \Longrightarrow \qquad x = \frac{E_0 - E}{2E_0} \quad . \tag{4.200}$$

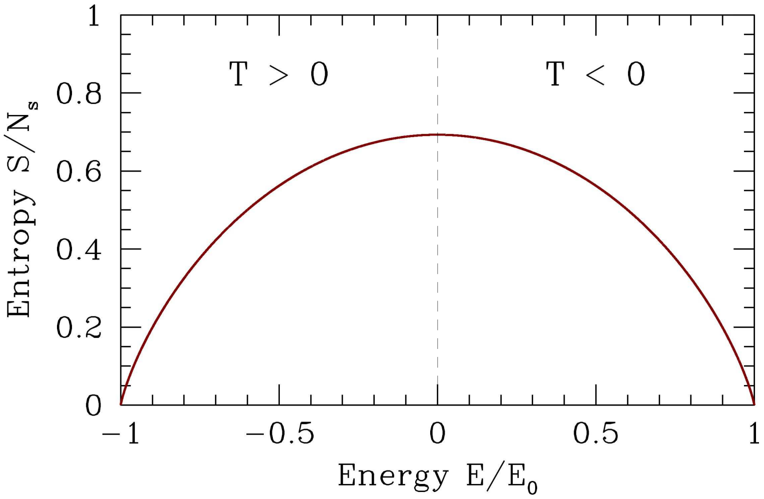


Figure 4.7: When entropy decreases with increasing energy, the temperature is negative. Typically, kinetic degrees of freedom prevent this peculiarity from manifesting in physical systems.

We therefore have

$$S(E,N_{\mathsf{s}}) = -N_{\mathsf{s}}\,k_{\mathrm{B}}\left[\left(\frac{E_0-E}{2E_0}\right)\ln\!\left(\frac{E_0-E}{2E_0}\right)+\left(\frac{E_0+E}{2E_0}\right)\ln\!\left(\frac{E_0+E}{2E_0}\right)\right] \quad . \tag{4.201}$$

We now have

$$\frac{1}{T}=\left(\frac{\partial S}{\partial E}\right)_{\!N_{\mathsf{s}}}=\frac{\partial S}{\partial x}\,\frac{\partial x}{\partial E}=\frac{N_{\mathsf{s}}k_{\mathrm{B}}}{2E_0}\,\ln\!\left(\frac{E_0-E}{E_0+E}\right) \quad . \tag{4.202}$$

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 $S(E,N_{\mathsf{s}})$ which is an increasing function of energy. In this system, $S(E,N_{\mathsf{s}})$ achieves a maximum of $S_{\mathrm{max}} = N_{\mathsf{s}}k_{\mathrm{B}}\ln 2$ 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. 4.192 : 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.

4.8.3 Adsorption

`PROBLEM`: A surface containing N_{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 Δ, and physical constants.

`SOLUTION:` The grand partition function for the surface is

$$\Xi_{\text{surface}} = e^{-\Omega_{\text{surface}}/k_{\text{B}}T} = \sum_{j=0}^{N_{\mathsf{s}}} \binom{N_{\mathsf{s}}}{j} e^{j(\mu+\Delta)/k_{\text{B}}T} = \big(1 + e^{\mu/k_{\text{B}}T}\, e^{\Delta/k_{\text{B}}T}\big)^{N_{\mathsf{s}}} \quad . \tag{4.203}$$

The fraction of occupied sites is

$$f = \frac{\langle \hat{N}_{\text{surface}} \rangle}{N_{\mathsf{s}}} = -\frac{1}{N_{\mathsf{s}}} \frac{\partial \Omega_{\text{surface}}}{\partial \mu} = \frac{e^{\mu/k_{\text{B}}T}}{e^{\mu/k_{\text{B}}T} + e^{-\Delta/k_{\text{B}}T}} \quad . \tag{4.204}$$

Since the surface is in equilibrium with the gas, its fugacity $z = \exp(\mu/k_{\text{B}}T)$ 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 $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\text{B}}T}$ is the thermal wavelength. Thus, the grand partition function, for indistinguishable particles, is

$$\Xi_{\text{gas}} = \exp\Big(V\lambda_T^{-3}\, e^{\mu/k_{\text{B}}T}\Big) \quad . \tag{4.205}$$

The gas density is

$$n = \frac{\langle \hat{N}_{\text{gas}} \rangle}{V} = -\frac{1}{V} \frac{\partial \Omega_{\text{gas}}}{\partial \mu} = \lambda_T^{-3}\, e^{\mu/k_{\text{B}}T} \quad . \tag{4.206}$$

We can now solve for the fugacity: $z = e^{\mu/k_{\text{B}}T} = n\lambda_T^3$. Thus, the fraction of occupied adsorption sites is

$$f = \frac{n\lambda_T^3}{n\lambda_T^3 + e^{-\Delta/k_{\text{B}}T}} \quad . \tag{4.207}$$

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 Δ, 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 $n \to 0$, 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

$$\hat{h} = \frac{\boldsymbol{p}^2}{2m} + \frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I\sin^2\!\theta} - \mu_0 \boldsymbol{E} \cdot \hat{\boldsymbol{n}} \quad . \tag{4.208}$$

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 ϕ, θ, and ψ. For a symmetric top, the Euler angles (θ,ϕ) are, respectively, the polar and azimuthal angles of the symmetry axis, and ψ is the angle of rotations 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_{nuc}^2, where a_{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 $k_{\text{B}}T$ 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

$$d\mu = \frac{dx\,dp_x}{h}\cdot\frac{dy\,dp_y}{h}\cdot\frac{dz\,dp_z}{h}\cdot\frac{d\theta\,dp_\theta}{h}\cdot\frac{d\phi\,dp_\phi}{h} \tag{4.209}$$

and the single molecule partition function is

$$\begin{aligned}\zeta(T,V,E) &= \int\! d\mu\, e^{-\hat{h}/k_{\text{B}}T} = \frac{V}{\lambda_T^3}\int\limits_0^{2\pi}\! d\phi \int\limits_0^{\pi}\! d\theta \int\limits_{-\infty}^{\infty}\!\frac{dp_\theta}{h}\int\limits_{-\infty}^{\infty}\!\frac{dp_\phi}{h}\; e^{-\beta p_\theta^2/2I}\, e^{-\beta p_\phi^2/2I\sin^2\!\theta}\, e^{-\beta u(\theta)} \\ &= \frac{V}{\lambda_T^3}\cdot\frac{Ik_{\text{B}}T}{\hbar^2}\int\limits_0^{\pi}\! d\theta\,\sin\theta\, e^{-\beta u(\theta)} = \underbrace{\frac{V}{\lambda_T^3}}_{\zeta_{\text{trans}}}\cdot\underbrace{\frac{Ik_{\text{B}}T}{\hbar^2}}_{\zeta_{\text{rot}}}\cdot\underbrace{\frac{\sinh(\mu_0 E/k_{\text{B}}T)}{\mu_0 E/k_{\text{B}}T}}_{\zeta_{\text{pol}}} \quad ,\end{aligned} \tag{4.210}$$

where

$$\zeta_{\text{trans}}(T,V) = \frac{V}{\lambda_T^3} \quad , \quad \zeta_{\text{rot}}(T) = \frac{Ik_{\text{B}}T}{\hbar^2} \quad , \quad \zeta_{\text{pol}}(T,E) = \frac{\sinh(\mu_0 E/k_{\text{B}}T)}{\mu_0 E/k_{\text{B}}T} \tag{4.211}$$

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,

$$\begin{aligned}F(T,V,N,E) &= -Nk_{\text{B}}T\log\big[\zeta(T,V,E)/N\big] - Nk_{\text{B}}T \\ &= -Nk_{\text{B}}T\left\{\log\!\left(\frac{V}{N\lambda_T^3}\right) + \log\!\left(\frac{2Ik_{\text{B}}T}{\hbar^2}\right) + \log\!\left(\frac{\sinh(\mu_0E/k_{\text{B}}T)}{\mu_0E/k_{\text{B}}T}\right) + 1\right\} \quad .\end{aligned} \tag{4.212}$$

The average dipole moment per molecule is given by

$$\langle \boldsymbol{d}\rangle = \mu_0\,\langle\hat{\boldsymbol{n}}\rangle = -\frac{1}{N}\left(\frac{\partial F}{\partial \boldsymbol{E}}\right)_{T,V,N} = \mu_0\left\{\text{ctnh}\!\left(\frac{\mu_0E}{k_{\text{B}}T}\right) - \frac{k_{\text{B}}T}{\mu_0E}\right\}\widehat{\boldsymbol{E}} \quad , \tag{4.213}$$

where $\widehat{\boldsymbol{E}} = \hat{\boldsymbol{z}}$.

Let's work out the physics in the GCE. The grand partition function is given by

$$\begin{aligned}\Xi(T,V,\mu,E) &= \sum_{N=0}^{\infty} z^N Z(T,V,N,E) = \exp\!\big(z\,\zeta_{\text{trans}}\,\zeta_{\text{rot}}\,\zeta_{\text{pol}}\big) \\ &= \exp\!\left\{\left(\frac{z\,V}{N\lambda_T^3}\right)\left(\frac{2Ik_{\text{B}}T}{\hbar^2}\right)\left(\frac{\sinh(\mu_0E/k_{\text{B}}T)}{\mu_0E/k_{\text{B}}T}\right)\right\} \quad ,\end{aligned} \tag{4.214}$$

where $z = \exp(\mu/k_{\rm B}T)$ is the fugacity. Thus, the grand potential is

$$\begin{aligned} \Omega(T,V,\mu,E) &= -k_{\rm B}T \log \Xi(T,V,\mu,E) = -k_{\rm B}T\, z\, \zeta_{\rm trans}\, \zeta_{\rm rot}\, \zeta_{\rm pol} \\ &= -k_{\rm B}T \bigg(\frac{z\,V}{N\lambda_T^3}\bigg)\bigg(\frac{2Ik_{\rm B}T}{\hbar^2}\bigg)\bigg(\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}\bigg) \quad . \end{aligned} \tag{4.215}$$

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

$$\boldsymbol{D} = \mu_0 \bigg\langle \sum_{j=1}^N \boldsymbol{d}_j \bigg\rangle = -\bigg(\frac{\partial \Omega}{\partial \boldsymbol{E}}\bigg)_{T,V,\mu} = k_{\rm B}T\, z\, \zeta_{\rm trans}\, \zeta_{\rm rot} \bigg(\frac{\partial \zeta_{\rm pol}}{\partial \boldsymbol{E}}\bigg)_{T,V,\mu} \quad . \tag{4.216}$$

The average number of particles is

$$N = -\bigg(\frac{\partial \Omega}{\partial \mu}\bigg)_{T,V,\boldsymbol{E}} = z\, \zeta_{\rm trans}\, \zeta_{\rm rot}\, \zeta_{\rm pol} \tag{4.217}$$

and therefore

$$\boldsymbol{d} = \frac{\boldsymbol{D}}{N} = k_{\rm B}T \bigg(\frac{\partial \log \zeta_{\rm pol}}{\partial \boldsymbol{E}}\bigg)_{T,V,\mu} = \mu_0 \bigg\{ {\rm ctnh}\bigg(\frac{\mu_0 E}{k_{\rm B}T}\bigg) - \frac{k_{\rm B}T}{\mu_0 E} \bigg\}\, \hat{\boldsymbol{E}} \quad . \tag{4.218}$$

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 $N_{\sf m}$ monomers is placed under a tension τ. We then have

$$\hat{H} = \sum_{j=1}^{N_{\sf m}} \Big[\varepsilon_{\rm A}\, \delta_{\sigma_j,+1} + \varepsilon_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad . \tag{4.219}$$

Similarly, the length is

$$\hat{L} = \sum_{j=1}^{N_{\sf m}} \Big[\ell_{\rm A}\, \delta_{\sigma_j,+1} + \ell_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad . \tag{4.220}$$

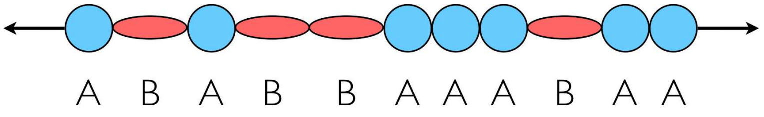


Figure 4.8: The monomers in wool are modeled as existing in one of two states. The low energy undeformed state is A, and the higher energy deformed state is B. Applying tension induces more monomers to enter the B state.

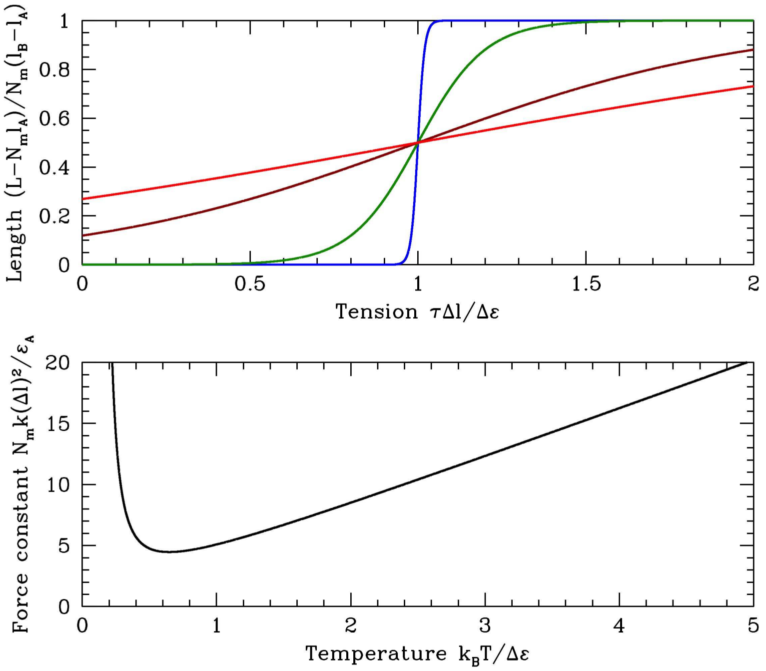


Figure 4.9: Upper panel: length $L(\tau,T)$ for $k_\mathrm{B}T/\tilde{\varepsilon} = 0.01$ (blue), 0.1 (green), 0.5 (dark red), and 1.0 (red). Bottom panel: dimensionless force constant $N_\mathsf{m}(\Delta\ell)^2 k/\varepsilon_\mathrm{A}$ *versus* dimensionless temperature $k_\mathrm{B}T/\Delta\varepsilon$.

The Gibbs partition function is $Y = \mathsf{Tr}\; e^{-\hat{K}/k_\mathrm{B}T}$, with $\hat{K} = \hat{H} - \tau\hat{L}$:

$$\hat{K} = \sum_{j=1}^{N_\mathsf{m}} \Big[\tilde{\varepsilon}_\mathrm{A}\,\delta_{\sigma_j,+1} + \tilde{\varepsilon}_\mathrm{B}\,\delta_{\sigma_j,-1}\Big] \quad , \tag{4.221}$$

where $\tilde{\varepsilon}_\mathrm{A} \equiv \varepsilon_\mathrm{A} - \tau\ell_\mathrm{A}$ and $\tilde{\varepsilon}_\mathrm{B} \equiv \varepsilon_\mathrm{B} - \tau\ell_\mathrm{B}$. At $\tau = 0$ the A state is preferred for each monomer, but when τ exceeds τ^*, defined by the relation $\tilde{\varepsilon}_\mathrm{A} = \tilde{\varepsilon}_\mathrm{B}$, the B state is preferred. One finds

$$\tau^* = \frac{\varepsilon_\mathrm{B} - \varepsilon_\mathrm{A}}{\ell_\mathrm{B} - \ell_\mathrm{A}} \quad . \tag{4.222}$$

Once again, we have a set of N_m noninteracting spins. The partition function is $Y = \zeta^{N_\mathsf{m}}$, where ζ is the single monomer partition function, $\zeta = \mathsf{Tr}\;\exp(-\beta\hat{h})$, where

$$\hat{h} = \tilde{\varepsilon}_\mathrm{A}\,\delta_{\sigma_j,1} + \tilde{\varepsilon}_\mathrm{B}\,\delta_{\sigma_j,-1} \tag{4.223}$$

is the single "spin" Hamiltonian. Thus,

$$\zeta = \mathsf{Tr}\; e^{-\beta\hat{h}} = e^{-\beta\tilde{\varepsilon}_\mathrm{A}} + e^{-\beta\tilde{\varepsilon}_\mathrm{B}} \quad , \tag{4.224}$$

It is convenient to define the differences

$$\Delta\varepsilon = \varepsilon_\mathrm{B} - \varepsilon_\mathrm{A} \quad , \quad \Delta\ell = \ell_\mathrm{B} - \ell_\mathrm{A} \quad , \quad \Delta\tilde{\varepsilon} = \tilde{\varepsilon}_\mathrm{B} - \tilde{\varepsilon}_\mathrm{A} \tag{4.225}$$

in which case the partition function Y and Gibbs free energy G are

$$\begin{aligned} Y(T,\tau,N_{\mathsf{m}}) &= e^{-N_{\mathsf{m}}\beta\,\tilde\varepsilon_{\rm A}}\Big[1+e^{-\beta\Delta\tilde\varepsilon}\Big]^{N_{\mathsf{m}}} \\ G(T,\tau,N_{\mathsf{m}}) &= N_{\mathsf{m}}\,\tilde\varepsilon_{\rm A} - N_{\mathsf{m}}\,k_{\rm B}T\ln\Big[1+e^{-\Delta\tilde\varepsilon/k_{\rm B}T}\Big] \end{aligned} \tag{4.226}$$

The average length is

$$L = \langle \hat L\rangle = -\bigg(\frac{\partial G}{\partial\tau}\bigg)_{T,N_{\mathsf{m}}} = N_{\mathsf{m}}\,\ell_{\rm A} + \frac{N_{\mathsf{m}}\,\Delta\ell}{e^{(\Delta\varepsilon-\tau\Delta\ell)/k_{\rm B}T}+1} \quad . \tag{4.227}$$

The polymer behaves as a spring, and for small τ the spring constant is

$$k = \frac{\partial\tau}{\partial L}\bigg|_{\tau=0} = \frac{4k_{\rm B}T}{N_{\mathsf{m}}(\Delta\ell)^2}\,\cosh^2\!\bigg(\frac{\Delta\varepsilon}{2k_{\rm B}T}\bigg) \quad . \tag{4.228}$$

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 $L = N_{\mathsf{m}}\,\ell_{\rm A}$ At high temperatures $k_{\rm B}T$ dominates and both A and B configurations are equally likely. The mean length is then $L = \frac{1}{2}N_{\mathsf{m}}(\ell_{\rm A}+\ell_{\rm B})$.

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 α, specifying rotational, vibrational, and electronic degrees of freedom. The single particle Hamiltonian is then

$$\hat h = \frac{\boldsymbol{p}^2}{2m} + \hat h_{\rm int} \quad , \tag{4.229}$$

with

$$\hat h\,|\,\boldsymbol{k}\,,\,\alpha\,\rangle = \bigg(\frac{\hbar^2\boldsymbol{k}^2}{2m} + \varepsilon_\alpha\bigg)\,|\,\boldsymbol{k}\,,\,\alpha\,\rangle \quad . \tag{4.230}$$

The partition function is

$$\zeta = \mathsf{Tr}\;e^{-\beta\hat h} = \sum_{\boldsymbol{p}} e^{-\beta\boldsymbol{p}^2/2m}\sum_j g_j\,e^{-\beta\varepsilon_j} \quad . \tag{4.231}$$

Here we have replaced the internal label α with a label j of energy eigenvalues, with g_j being the degeneracy of the internal state with energy ε_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

$$\boldsymbol{p} = \bigg(\frac{2\pi\hbar n_1}{L_1}\,,\,\frac{2\pi\hbar n_2}{L_2}\,,\,\ldots\,,\,\frac{2\pi\hbar n_d}{L_d}\bigg) \quad , \tag{4.232}$$

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:

$$\sum_{\boldsymbol{p}} \longrightarrow \int \frac{d^d\!p}{\Delta p_1 \cdots \Delta p_d} \tag{4.233}$$

where the volume in momentum space of an elementary rectangle is

$$\Delta p_1 \cdots \Delta p_d = \frac{(2\pi\hbar)^d}{L_1 \cdots L_d} = \frac{(2\pi\hbar)^d}{V} \quad . \tag{4.234}$$

Thus,

$$\zeta = V \!\int\! \frac{d^d\!p}{(2\pi\hbar)^d}\, e^{-\boldsymbol{p}^2/2mk_\mathrm{B}T} \sum_j g_j\, e^{-\varepsilon_j/k_\mathrm{B}T} = V\lambda_T^{-d}\, \xi(T) \tag{4.235}$$

where

$$\xi(T) = \sum_j g_j\, e^{-\varepsilon_j/k_\mathrm{B}T} \tag{4.236}$$

is the *internal coordinate partition function*. The N-particle ordinary canonical partition function is then

$$Z_N = \frac{1}{N!}\bigg(\frac{V}{\lambda_T^d}\bigg)^{\!N} \xi^N(T) \quad . \tag{4.237}$$

Using Stirling's approximation, we find the Helmholtz free energy $F = -k_\mathrm{B}T \ln Z$ is

$$F(T,V,N) = -Nk_\mathrm{B}T \Bigg[\ln\!\bigg(\frac{V}{N\lambda_T^d}\bigg) + 1 + \ln \xi(T) \Bigg] \quad , \tag{4.238}$$

which entails the ideal gas law $pV = Nk_\mathrm{B}T$ since $dF = -S\,dT - p\,dV + \mu\,dN$. Making the Legendre transform to $G(T,p,N) = F(T,V,N) + pV$, we have

$$G(T,p,N) = Nk_\mathrm{B}T \ln\!\bigg(\frac{p\,\lambda_T^d}{k_\mathrm{B}T}\bigg) + N\varphi(T) \quad , \tag{4.239}$$

where

$$\varphi(T) = -k_\mathrm{B}T \ln \xi(T) \tag{4.240}$$

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

$$S = -\bigg(\frac{\partial G}{\partial T}\bigg)_{\!p,N} = Nk_\mathrm{B} \Bigg[\ln\!\bigg(\frac{k_\mathrm{B}T}{p\lambda_T^d}\bigg) + 1 + \tfrac{1}{2}d \Bigg] - N\varphi'(T) \quad , \tag{4.241}$$

and therefore the heat capacities are

$$\begin{aligned} C_{p,N} &= T\bigg(\frac{\partial S}{\partial T}\bigg)_{\!p,N} = \big(\tfrac{1}{2}d + 1\big)Nk_\mathrm{B} - NT\varphi''(T) \\ C_{V,N} &= T\bigg(\frac{\partial S}{\partial T}\bigg)_{\!V,N} = \tfrac{1}{2}dNk_\mathrm{B} - NT\varphi''(T) \quad , \end{aligned} \tag{4.242}$$

and we have $C_{p,N} - C_{V,N} = Nk_\mathrm{B}$. 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

$$\hat{h}_{\text{int}} = \hat{h}_{\text{rot}} + \hat{h}_{\text{vib}} + \hat{h}_{\text{elec}} \tag{4.243}$$

as a sum over internal Hamiltonians governing rotational, vibrational, and electronic degrees of freedom. Then

$$\xi_{\text{int}}(T) = \xi_{\text{rot}}(T) \cdot \xi_{\text{vib}}(T) \cdot \xi_{\text{elec}}(T) \quad . \tag{4.244}$$

Associated with each class of excitation is a characteristic temperature Θ. Rotational and vibrational temperatures of a few common molecules are listed in table tab. 4.1.

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 $\vec{\mathsf{L}} = (\mathsf{L}_{\text{a}}, \mathsf{L}_{\text{b}}, \mathsf{L}_{\text{c}})$ as the angular momentum vector about instantaneous *body-fixed* principal axes. Thus $\mathsf{L}_{\text{a}} = \hat{\boldsymbol{n}}_{\text{a}} \cdot \boldsymbol{L}$, $\mathsf{L}_{\text{b}} = \hat{\boldsymbol{n}}_{\text{b}} \cdot \boldsymbol{L}$, and $\mathsf{L}_{\text{c}} = \hat{\boldsymbol{n}}_{\text{c}} \cdot \boldsymbol{L}$, where $\hat{\boldsymbol{n}}_{\text{a.b,c}}(t)$ are the instantaneous principal axes. We take $\hat{\boldsymbol{n}}_{\text{c}}$ to be the symmetry axis of the molecule. The rotational Hamiltonian is then

$$\begin{aligned}\hat{h}_{\text{rot}} &= \frac{\mathsf{L}_{\text{a}}^2 + \mathsf{L}_{\text{b}}^2}{2I_1} + \frac{\mathsf{L}_{\text{c}}^2}{2I_3} \\ &= \frac{\hbar^2 L(L+1)}{2I_1} + \left(\frac{1}{2I_3} - \frac{1}{2I_1}\right) \mathsf{L}_c^2 \quad .\end{aligned} \tag{4.245}$$

Note that

$$\big[L^\mu\,,\,\mathsf{L}_c\big] = n_{\text{c}}^\nu \big[L^\mu\,,\,L^\nu\big] + \big[L^\mu\,,\,n_{\text{c}}^\nu\big] L^\nu = i\epsilon_{\mu\nu\lambda}\, n_{\text{c}}^\nu\, L^\lambda + i\epsilon_{\mu\nu\lambda}\, n_{\text{c}}^\lambda\, L^\nu = 0 \quad , \tag{4.246}$$

which is equivalent to the statement that $\mathsf{L}_{\text{c}} = \hat{\boldsymbol{n}}_{\text{c}} \cdot \boldsymbol{L}$ is a rotational scalar. We can therefore simultaneously specify the eigenvalues of $\{\boldsymbol{L}^2, L^z, \mathsf{L}_{\text{c}}\}$, which form a complete set of commuting observables (CSCO)[9]. The eigenvalues of L^z are $m\hbar$ with $m \in \{-L, \ldots, L\}$, while those of L_{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,

$$\Theta = \frac{\hbar^2}{2I_1 k_{\text{B}}} \qquad , \qquad \widetilde{\Theta} = \frac{\hbar^2}{2I_3 k_{\text{B}}} \quad . \tag{4.247}$$

Prolateness then means $\widetilde{\Theta} > \Theta$. We conclude that the rotational partition function for an axisymmetric molecule is given by

$$\xi_{\text{rot}}(T) = \sum_{L=0}^{\infty} (2L+1)\, e^{-L(L+1)\,\Theta/T} \sum_{k=-L}^{L} e^{-k^2\,(\widetilde{\Theta}-\Theta)/T} \tag{4.248}$$

---

[9] 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.

molecule	$\Theta_{\rm rot}({\rm K})$	$\Theta_{\rm vib}({\rm K})$
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.

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

$$\xi_{\rm rot}(T) = \sum_{L=0}^{\infty} (2L+1)\, e^{-L(L+1)\,\Theta/T} \quad . \tag{4.249}$$

When $T \ll \Theta$, only the first few terms contribute, and

$$\xi_{\rm rot}(T) = 1 + 3\, e^{-2\Theta/T} + 5\, e^{-6\Theta/T} + \ldots \tag{4.250}$$

In the high temperature limit, we have a slowly varying summand. The *Euler-MacLaurin summation formula* may be used to evaluate such a series:

$$\sum_{k=0}^{n} F_k = \int\limits_0^n dk\, F(k) + \tfrac{1}{2}\big[F(0)+F(n)\big] + \sum_{j=1}^{\infty} \frac{\mathsf{B}_{2j}}{(2j)!}\Big[F^{(2j-1)}(n) - F^{(2j-1)}(0)\Big] \tag{4.251}$$

where B_j is the $j^{\rm th}$ *Bernoulli number* where

$$\mathsf{B}_0 = 1 \quad , \quad \mathsf{B}_1 = -\tfrac{1}{2} \quad , \quad \mathsf{B}_2 = \tfrac{1}{6} \quad , \quad \mathsf{B}_4 = -\tfrac{1}{30} \quad , \quad \mathsf{B}_6 = \tfrac{1}{42} \quad . \tag{4.252}$$

Thus,

$$\sum_{k=0}^{\infty} F_k = \int\limits_0^\infty dx\, F(x) + \tfrac{1}{2}F(0) - \tfrac{1}{12}F'(0) - \frac{1}{720}\, F'''(0) + \ldots \quad . \tag{4.253}$$

We have $F(x) = (2x+1)\, e^{-x(x+1)\Theta/T}$, for which $\int\limits_0^\infty dx\, F(x) = T/\Theta$, hence

$$\xi_{\rm rot} = \frac{T}{\Theta} + \frac{1}{3} + \frac{1}{15}\,\frac{\Theta}{T} + \frac{4}{315}\left(\frac{\Theta}{T}\right)^2 + \ldots \quad . \tag{4.254}$$

Recall that $\varphi(T) = -k_{\rm B}T\ln\xi(T)$. We conclude that $\varphi_{\rm rot}(T) \approx -3k_{\rm B}T\, e^{-2\Theta/T}$ for $T \ll \Theta$ and $\varphi_{\rm rot}(T) \approx -k_{\rm B}T\ln(T/\Theta)$ 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 Θ. Including the first few terms, we have

$$\begin{aligned}
\Delta C_V(T \ll \Theta) &= 12\, Nk_{\rm B}\left(\frac{\Theta}{T}\right)^2 e^{-2\Theta/T} + \ldots \\
\Delta C_V(T \gg \Theta) &= Nk_{\rm B}\left\{1 + \frac{1}{45}\left(\frac{\Theta}{T}\right)^2 + \frac{16}{945}\left(\frac{\Theta}{T}\right)^3 + \ldots\right\} \quad .
\end{aligned} \tag{4.255}$$

Note that C_V overshoots its limiting value of Nk_{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

$$\varepsilon(\mathsf{L}_{\mathrm{a}},\mathsf{L}_{\mathrm{b}},\mathsf{L}_{\mathrm{c}})=\frac{\mathsf{L}_{\mathrm{a}}^2}{2I_1}+\frac{\mathsf{L}_{\mathrm{b}}^2}{2I_2}+\frac{\mathsf{L}_{\mathrm{c}}^2}{2I_3}\quad . \tag{4.256}$$

We then have

$$\xi_{\mathrm{rot}}(T)=\frac{1}{g_{\mathrm{rot}}}\int\!\frac{d\mathsf{L}_{\mathrm{a}}\,d\mathsf{L}_{\mathrm{b}}\,d\mathsf{L}_{\mathrm{c}}\,d\phi\,d\theta\,d\psi}{(2\pi\hbar)^3}\,e^{-\varepsilon(\mathsf{L}_{\mathrm{a}}\,\mathsf{L}_{\mathrm{b}}\,\mathsf{L}_{\mathrm{c}})/k_{\mathrm{B}}T}\quad , \tag{4.257}$$

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_{rot} accounts for physically indistinguishable orientations of the molecule brought about by rotations, which can happen when more than one of the nuclei is the same. We then have

$$\xi_{\mathrm{rot}}(T)=\left(\frac{2k_{\mathrm{B}}T}{\hbar^2}\right)^{3/2}\sqrt{\pi I_1 I_2 I_3}\quad . \tag{4.258}$$

This leads to $\Delta C_V=\frac{3}{2}Nk_{\mathrm{B}}$.

4.9.4 Vibrations

Vibrational frequencies are often given in units of inverse wavelength, such as cm^{-1}, called a *wavenumber*. To convert to a temperature scale T^*, we write $k_{\mathrm{B}}T^*=h\nu=hc/\lambda$, hence $T^*=(hc/k_{\mathrm{B}})\,\lambda^{-1}$, and we multiply by

$$\frac{hc}{k_{\mathrm{B}}}=1.436\,\mathrm{K}\cdot\mathrm{cm}\quad . \tag{4.259}$$

For example, infrared absorption ($\sim 50\ \mathrm{cm}^{-1}$ to $10^4\ \mathrm{cm}^{-1}$) reveals that the 'asymmetric stretch' mode of the $\mathrm{H_2O}$ 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

$$\hat{h}=\frac{p^2}{2m}+\tfrac{1}{2}m\omega^2q^2=\hbar\omega\big(a^\dagger a+\tfrac{1}{2}\big)\quad . \tag{4.260}$$

In general there are many vibrational modes, hence many normal mode frequencies ω_α. We then must sum over all of them, resulting in

$$\xi_{\mathrm{vib}}=\prod_\alpha \xi_{\mathrm{vib}}^{(\alpha)}\quad . \tag{4.261}$$

For each such normal mode, the contribution is

$$\begin{aligned}\xi&=\sum_{n=0}^{\infty}e^{-(n+\frac{1}{2})\hbar\omega/k_{\mathrm{B}}T}=e^{-\hbar\omega/2k_{\mathrm{B}}T}\sum_{n=0}^{\infty}\Big(e^{-\hbar\omega/k_{\mathrm{B}}T}\Big)^n\\&=\frac{e^{-\hbar\omega/2k_{\mathrm{B}}T}}{1-e^{-\hbar\omega/k_{\mathrm{B}}T}}=\frac{1}{2\sinh(\Theta/2T)}\quad ,\end{aligned} \tag{4.262}$$

where $\Theta = \hbar\omega/k_{\rm B}$. Then

$$\varphi = k_{\rm B}T \ln\Big(2\sinh(\Theta/2T)\Big) = \tfrac{1}{2}k_{\rm B}\Theta + k_{\rm B}T\ln\big(1-e^{-\Theta/T}\big) \quad . \tag{4.263}$$

The contribution to the heat capacity is

$$\Delta C_V = Nk_{\rm B}\bigg(\frac{\Theta}{T}\bigg)^2 \frac{e^{\Theta/T}}{(e^{\Theta/T}-1)^2} = \begin{cases} Nk_{\rm B}\,(\Theta/T)^2\,\exp(-\Theta/T) & (T\to 0) \\ Nk_{\rm B} & (T\to\infty) \quad . \end{cases} \tag{4.264}$$

4.9.5 Two-level systems : Schottky anomaly

Consider now a two-level system, with energies ε_0 and ε_1. We define $\Delta \equiv \varepsilon_1 - \varepsilon_0$ and assume without loss of generality that $\Delta > 0$. The partition function is

$$\zeta = e^{-\beta\varepsilon_0} + e^{-\beta\varepsilon_1} = e^{-\beta\varepsilon_0}\big(1+e^{-\beta\Delta}\big) \quad . \tag{4.265}$$

The free energy is

$$f = -k_{\rm B}T\ln\zeta = \varepsilon_0 - k_{\rm B}T\ln\big(1+e^{-\Delta/k_{\rm B}T}\big) \quad . \tag{4.266}$$

The entropy for a given two level system is then

$$s = -\frac{\partial f}{\partial T} = k_{\rm B}\ln\big(1+e^{-\Delta/k_{\rm B}T}\big) + \frac{\Delta}{T}\cdot\frac{1}{e^{\Delta/k_{\rm B}T}+1} \tag{4.267}$$

and the heat capacity is $= T\,(\partial s/\partial T)$, *i.e.*

$$c(T) = \frac{\Delta^2}{k_{\rm B}T^2}\cdot\frac{e^{\Delta/k_{\rm B}T}}{\big(e^{\Delta/k_{\rm B}T}+1\big)^2} \quad . \tag{4.268}$$

Thus,

$$c\,(T\ll\Delta) = \frac{\Delta^2}{k_{\rm B}T^2}\,e^{-\Delta/k_{\rm B}T} \quad , \qquad c\,(T\gg\Delta) = \frac{\Delta^2}{4k_{\rm B}T^2} \quad . \tag{4.269}$$

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:

$$C(T) = \sum_i \widetilde{c}\,(\Delta_i/k_{\rm B}T) = N\int\limits_0^\infty d\Delta\,P(\Delta)\,\widetilde{c}(\Delta/T) \quad , \tag{4.270}$$

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

$$\int\limits_0^\infty d\Delta\,P(\Delta) = 1 \quad . \tag{4.271}$$

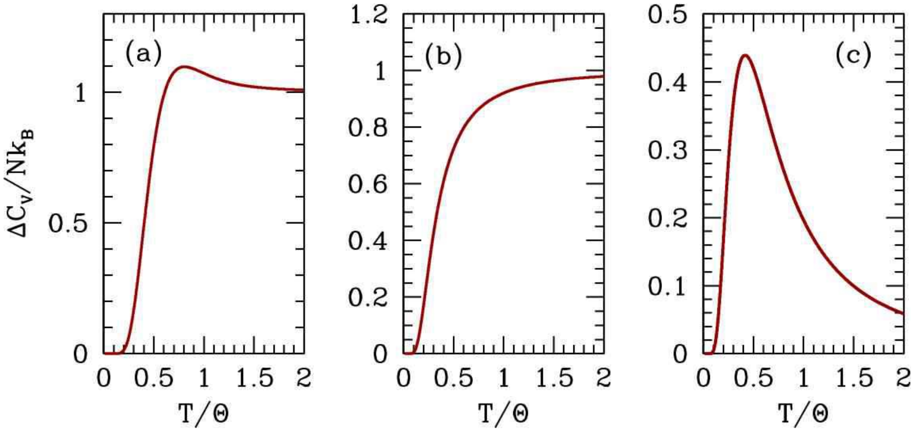


Figure 4.10: Heat capacity per molecule as a function of temperature for (a) heteronuclear diatomic gases, (b) a single vibrational mode, and (c) a single two-level system.

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 $(\mathrm{KBr})_{1-x}(\mathrm{KCN})_x$ 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 orientation 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_BT 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' ($A = 2$ (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 ${}^{2J+1}L_S$ 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 $\hat{H} = \mathit{\Lambda} \boldsymbol{L} \cdot \boldsymbol{S}$, where $\mathit{\Lambda}$ is the Russell-Saunders coupling. If the last shell is less than or equal to half-filled, then $\mathit{\Lambda} > 0$ and the ground state has $J = |L - S|$. If the last shell is more than half-filled, the coupling is *inverted*, *i.e.* $\mathit{\Lambda} < 0$, and the ground state has $J = L + S$.[10]

The electronic contribution to ξ is then

$$\xi_{\text{elec}} = \sum_{J=|L-S|}^{L+S} (2J+1)\, e^{-\Delta\varepsilon(L,S,J)/k_{\text{B}}T} \tag{4.272}$$

where

$$\Delta\varepsilon(L,S,J) = \tfrac{1}{2}\mathit{\Lambda}\Big[J(J+1) - L(L+1) - S(S+1)\Big] \quad . \tag{4.273}$$

At high temperatures, $k_{\text{B}}T$ is larger than the energy difference between the different J multiplets, and we have $\xi_{\text{elec}} \sim (2L+1)(2S+1)\, e^{-\beta\varepsilon_0}$, where ε_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 $\xi_{\text{elec}} \sim (2J+1)\, e^{-\beta\varepsilon_0}$. If, in addition, there is a nonzero nuclear spin I, then we also must include a factor $\xi_{\text{nuc}} = (2I+1)$, 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 $\xi_{\text{elec}} \cdot \xi^{(1)}_{\text{nuc}} \cdot \xi^{(2)}_{\text{nuc}}$, 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 angular momentum I_{tot}. If I_{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 ${}^1\mathit{\Sigma}^+_g$.[11]

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

$$\xi_{\text{rot}}(T) = g_g\, \zeta_g(T) + g_u\, \zeta_u(T) \quad , \tag{4.274}$$

where

$$\zeta_g(T) = \sum_{L \text{ even}} (2L+1)\, e^{-L(L+1)\,\Theta_{\text{rot}}/T} \quad , \quad \zeta_u(T) = \sum_{L \text{ odd}} (2L+1)\, e^{-L(L+1)\,\Theta_{\text{rot}}/T} \quad . \tag{4.275}$$

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

---

[10] See *e.g.* §72 of Landau and Lifshitz, *Quantum Mechanics*.

[11] *Ibid.* §86.

$2I$	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.

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

$$\frac{N^{\text{ortho}}_{\text{H}_2}}{N^{\text{para}}_{\text{H}_2}} = \frac{g_u\,\zeta_u(T)}{g_g\,\zeta_g(T)} = \frac{3\,\zeta_u(T)}{\zeta_g(T)} \qquad , \qquad \frac{N^{\text{ortho}}_{\text{D}_2}}{N^{\text{para}}_{\text{D}_2}} = \frac{g_g\,\zeta_g(T)}{g_u\,\zeta_u(T)} = \frac{2\,\zeta_g(T)}{\zeta_u(T)} \quad . \tag{4.276}$$

Incidentally, how do we derive the results in Tab. 4.2? The total nuclear angular momentum I_{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

$$I \otimes I = 0 \oplus 1 \oplus 2 \oplus \cdots \oplus 2I \quad . \tag{4.277}$$

The right hand side of the above equation lists all the possible multiplets. Thus, $I_{\text{tot}} \in \{0, 1, \ldots, 2I\}$. Now let us count the total number of states with *even* I_{tot}. If $2I$ is even, which is to say if I is an integer, we have

$$g_g^{\text{even}} = \sum_{n=0}^{I} \Big\{ 2 \cdot (2n) + 1 \Big\} = (I+1)(2I+1) \quad , \tag{4.278}$$

because the degeneracy of each multiplet is $2I_{\text{tot}} + 1$. It follows that

$$g_u^{\text{even}} = (2I+1)^2 - g_g^{\text{even}} = I(2I+1) \quad . \tag{4.279}$$

On the other hand, if $2I$ is odd, which is to say I is a half odd integer, then

$$g_g^{\text{odd}} = \sum_{n=0}^{I-\frac{1}{2}} \Big\{ 2 \cdot (2n) + 1 \Big\} = I(2I+1) \quad . \tag{4.280}$$

It follows that

$$g_u^{\text{odd}} = (2I+1)^2 - g_g^{\text{odd}} = (I+1)(2I+1) \quad . \tag{4.281}$$

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

$$F^*\big(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_k\big) \equiv F(x_1, \ldots, x_n) + \sum_{j=1}^{k} \lambda_j\, G_j(x_1, \ldots, x_n) \tag{4.282}$$

which is a function of the $(n+k)$ variables $\{x_1,\ldots,x_n;\lambda_1,\ldots,\lambda_k\}$, where the quantities $\{\lambda_1,\ldots,\lambda_k\}$ are *Lagrange undetermined multipliers*. We now *freely* extremize the extended function F^*:

$$\begin{aligned} dF^* &= \sum_{\sigma=1}^{n} \frac{\partial F^*}{\partial x_\sigma}\,dx_\sigma + \sum_{j=1}^{k} \frac{\partial F^*}{\partial \lambda_j}\,d\lambda_j \\ &= \sum_{\sigma=1}^{n} \left(\frac{\partial F}{\partial x_\sigma} + \sum_{j=1}^{k} \lambda_j\,\frac{\partial G_j}{\partial x_\sigma} \right) dx_\sigma + \sum_{j=1}^{k} G_j\,d\lambda_j = 0 \end{aligned} \tag{4.283}$$

This results in the $(n+k)$ equations

$$\begin{aligned} \frac{\partial F}{\partial x_\sigma} + \sum_{j=1}^{k} \lambda_j\,\frac{\partial G_j}{\partial x_\sigma} &= 0 \qquad (\sigma=1,\ldots,n) \\ G_j &= 0 \qquad (j=1,\ldots,k) \quad . \end{aligned} \tag{4.284}$$

The interpretation of all this is as follows. The first n equations in 4.284 can be written in vector form as

$$\boldsymbol{\nabla} F + \sum_{j=1}^{k} \lambda_j\,\boldsymbol{\nabla} G_j = 0 \quad . \tag{4.285}$$

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

$$\Sigma = \big\{ \boldsymbol{x} \in \mathbb{R}^n \,|\, G_j(\boldsymbol{x}) = 0 \ \forall\, j \in \{1,\ldots,k\} \big\} \tag{4.286}$$

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 Σ. 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 Σ, *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

$$G(a,h) = 2\pi a + \frac{h^2}{b} - \ell = 0 \quad . \tag{4.287}$$

Here, b and ℓ are constant parameters, each of which has dimensions of length.

Following Lagrange's method, we define the extended function

$$V^*(a,h,\lambda) \equiv V(a,h) + \lambda\, G(a,h) \quad , \tag{4.288}$$

and set

$$\begin{aligned} \frac{\partial V^*}{\partial a} &= 2\pi a h + 2\pi\lambda = 0 \\ \frac{\partial V^*}{\partial h} &= \pi a^2 + 2\lambda\,\frac{h}{b} = 0 \\ \frac{\partial V^*}{\partial \lambda} &= 2\pi a + \frac{h^2}{b} - \ell = 0 \quad . \end{aligned} \tag{4.289}$$

Solving these three equations simultaneously gives

$$a = \frac{2\ell}{5\pi} \quad , \quad h = \sqrt{\frac{b\ell}{5}} \quad , \quad \lambda = -\frac{2}{5^{3/2}\pi}\, b^{1/2}\,\ell^{3/2} \quad , \quad V^* = \frac{4}{5^{5/2}\,\pi}\,\ell^{5/2}\,b^{1/2} \quad . \tag{4.290}$$

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

$$\hat{h} = -J\,\sigma_1\sigma_2 - \mu_0 H\,(\sigma_1 + \sigma_2) \quad . \tag{4.291}$$

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 $|\!\downarrow\uparrow\rangle$[12].

Suppose there are N_{d} dimers. Then the OCE partition function is $Z = \zeta^{N_{\mathsf{d}}}$, 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

$$\zeta = \mathrm{Tr}\, e^{-\hat{h}/k_{\rm B}T} = 2\,e^{-J/k_{\rm B}T} + 2\,e^{J/k_{\rm B}T}\cosh\!\left(\frac{2\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.292}$$

Thus, the free energy is

$$F(T,H,N_{\mathsf{d}}) = -N_{\mathsf{d}}\,k_{\rm B}T\ln 2 - N_{\mathsf{d}}\,k_{\rm B}T\ln\!\left[e^{-J/k_{\rm B}T} + e^{J/k_{\rm B}T}\cosh\!\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)\right] \quad . \tag{4.293}$$

---

[12] *Nota bene* we are concerned with classical spin configurations only – there is no superposition of states allowed in this model.

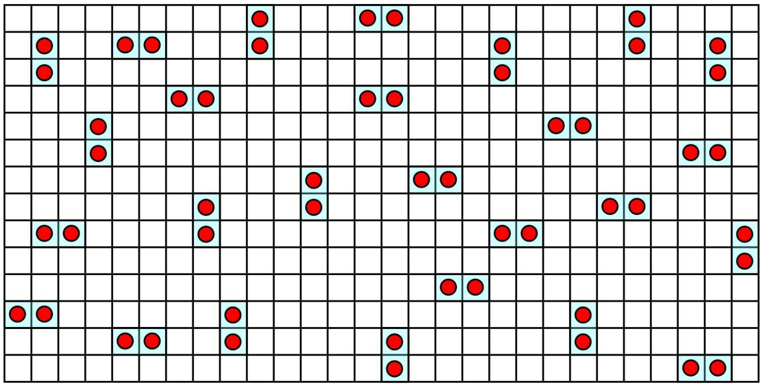


Figure 4.11: A model of noninteracting spin dimers on a lattice. Each red dot represents a classical spin for which $\sigma_j = \pm 1$.

The magnetization is

$$M = -\bigg(\frac{\partial F}{\partial H}\bigg)_{T,N_\mathsf{d}} = 2N_\mathsf{d}\,\mu_0 \cdot \frac{e^{J/k_\mathrm{B}T}\sinh\Big(\frac{2\mu_0 H}{k_\mathrm{B}T}\Big)}{e^{-J/k_\mathrm{B}T} + e^{J/k_\mathrm{B}T}\cosh\Big(\frac{2\mu_0 H}{k_\mathrm{B}T}\Big)} \tag{4.294}$$

It is instructive to consider the zero field isothermal susceptibility per spin,

$$\chi_T = \frac{1}{2N_\mathsf{d}}\,\frac{\partial M}{\partial H}\bigg|_{H=0} = \frac{\mu_0^2}{k_\mathrm{B}T} \cdot \frac{2\,e^{J/k_\mathrm{B}T}}{e^{J/k_\mathrm{B}T} + e^{-J/k_\mathrm{B}T}} \quad . \tag{4.295}$$

The quantity $\mu_0^2/k_\mathrm{B}T$ 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 $J \gg k_\mathrm{B}T$, then we obtain

$$\chi_T(J \gg k_\mathrm{B}T) \approx \frac{2\mu_0^2}{k_\mathrm{B}T} \quad . \tag{4.296}$$

This has the following simple interpretation. When $J \gg k_\mathrm{B}T$, the spins of each dimer are effectively locked in parallel. Thus, each dimer has an effective magnetic moment $\mu_\mathrm{eff} = 2\mu_0$. On the other hand, there are only half as many dimers as there are spins, so the resulting Curie susceptibility per spin is $\frac{1}{2} \times (2\mu_0)^2/k_\mathrm{B}T$.

When $-J \gg k_\mathrm{B}T$, the spins of each dimer are effectively locked in one of the two antiparallel configurations. We then have

$$\chi_T(-J \gg k_\mathrm{B}T) \approx \frac{2\mu_0^2}{k_\mathrm{B}T}\,e^{-2|J|/k_\mathrm{B}T} \quad . \tag{4.297}$$

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

$$\hat{h} = -\mu_0 H\,\sigma + \Delta(1 - \sigma^2) \quad . \tag{4.298}$$

We can also interpret this as describing a spin if $\sigma = \pm 1$ and a vacancy if $\sigma = 0$. The parameter Δ then represents the vacancy formation energy. The single particle partition function is

$$\zeta = \mathsf{Tr}\; e^{-\beta \hat{h}} = e^{-\beta\Delta} + 2\cosh(\beta\mu_0 H) \quad . \tag{4.299}$$

With N distinguishable noninteracting spins (*e.g.* at different sites in a crystalline lattice), we have $Z = \zeta^N$ and

$$F \equiv Nf = -k_B T \ln Z = -N k_B T \ln\Big[e^{-\beta\Delta} + 2\cosh(\beta\mu_0 H)\Big] \quad , \tag{4.300}$$

where $f = -k_B T \ln \zeta$ is the free energy of a single particle. Note that

$$\hat{n}_V = 1 - \sigma^2 = \frac{\partial \hat{h}}{\partial \Delta} \qquad , \qquad \hat{m} = \mu_0\, \sigma = -\frac{\partial \hat{h}}{\partial H} \quad , \tag{4.301}$$

are the vacancy number and magnetization, respectively. Thus,

$$n_V = \big\langle\, \hat{n}_V \,\big\rangle = \frac{\partial f}{\partial \Delta} = \frac{e^{-\Delta/k_B T}}{e^{-\Delta/k_B T} + 2\cosh(\mu_0 H/k_B T)} \tag{4.302}$$

and

$$m = \big\langle\, \hat{m} \,\big\rangle = -\frac{\partial f}{\partial H} = \frac{2\mu_0\, \sinh(\mu_0 H/k_B T)}{e^{-\Delta/k_B T} + 2\cosh(\mu_0 H/k_B T)} \quad . \tag{4.303}$$

At weak fields we can compute

$$\chi_T = \frac{\partial m}{\partial H}\bigg|_{H=0} = \frac{\mu_0^2}{k_B T} \cdot \frac{2}{2 + e^{-\Delta/k_B T}} \quad . \tag{4.304}$$

We thus obtain a modified Curie law. At temperatures $T \ll \Delta/k_B$, the vacancies are frozen out and we recover the usual Curie behavior. At high temperatures, where $T \gg \Delta/k_B$, 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 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

$$\ln(N!) = N \ln N - N + \mathcal{O}(\ln N) \quad .$$

(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 $N_\downarrow$:

$$N_\uparrow + N_0 + N_\downarrow = Q + N_0 = N \quad . \tag{4.305}$$

The total energy of the system is $E = -WQ$.

(a) The number of states available to the system is

$$\Omega = \frac{N!}{N_\uparrow!\,N_0!\,N_\downarrow!} \quad . \tag{4.306}$$

Fixing Q and M, along with the above constraint, is enough to completely determine $\{N_\uparrow, N_0, N_\downarrow\}$:

$$N_\uparrow = \tfrac{1}{2}\,(Q+M) \quad , \qquad N_0 = N - Q \quad , \qquad N_\downarrow = \tfrac{1}{2}\,(Q-M) \quad , \tag{4.307}$$

whence

$$\Omega(Q,M) = \frac{N!}{\big[\frac{1}{2}(Q+M)\big]!\,\big[\frac{1}{2}(Q-M)\big]!\,(N-Q)!} \quad . \tag{4.308}$$

The statistical entropy is $S = k_\mathrm{B} \ln \Omega$:

$$S(Q,M) = k_\mathrm{B} \ln(N!) - k_\mathrm{B} \ln\big[\tfrac{1}{2}(Q+M)!\big] - k_\mathrm{B} \ln\big[\tfrac{1}{2}(Q-M)!\big] - k_\mathrm{B} \ln\big[(N-Q)!\big] \quad . \tag{4.309}$$

(b) Now we invoke Stirling's rule,

$$\ln(N!) = N \ln N - N + \mathcal{O}(\ln N) \quad , \tag{4.310}$$

to obtain

$$\begin{aligned} \ln\Omega(Q,M) &= N\ln N - N - \tfrac{1}{2}(Q+M)\ln\big[\tfrac{1}{2}(Q+M)\big] + \tfrac{1}{2}(Q+M) \qquad (4.311)\\ &\qquad - \tfrac{1}{2}(Q-M)\ln\big[\tfrac{1}{2}(Q-M)\big] + \tfrac{1}{2}(Q-M) - (N-Q)\ln(N-Q) + (N-Q) \\ &= N\ln N - \tfrac{1}{2}Q\ln\Big[\tfrac{1}{4}(Q^2-M^2)\Big] - \tfrac{1}{2}M\ln\!\left(\frac{Q+M}{Q-M}\right) \end{aligned}$$

(4.312)

Combining terms,

$$\ln\Omega(Q,M) = -Nq\ln\Big[\tfrac{1}{2}\sqrt{q^2-m^2}\Big] - \tfrac{1}{2}Nm\ln\!\left(\frac{q+m}{q-m}\right) - N(1-q)\ln(1-q) \quad , \tag{4.313}$$

where $Q = Nq$ and $M = Nm$. Note that the entropy $S = k_\mathrm{B}\ln\Omega$ is extensive. The statistical entropy per site is thus

$$s(q,m) = -k_\mathrm{B}\,q\ln\Big[\tfrac{1}{2}\sqrt{q^2-m^2}\Big] - \tfrac{1}{2}k_\mathrm{B}\,m\ln\!\left(\frac{q+m}{q-m}\right) - k_\mathrm{B}\,(1-q)\ln(1-q) \quad . \tag{4.314}$$

The temperature is obtained from the relation

$$\begin{aligned}\frac{1}{T}&=\left(\frac{\partial S}{\partial E}\right)_{M}=\frac{1}{W}\left(\frac{\partial s}{\partial q}\right)_{m}\\&=\frac{1}{W}\ln(1-q)-\frac{1}{W}\ln\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right]\quad.\end{aligned}\tag{4.315}$$

Thus,

$$T=\frac{W/k_{\mathrm{B}}}{\ln\left[2(1-q)/\sqrt{q^2-m^2}\right]}\quad.\tag{4.316}$$

(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 $T=W/k_{\mathrm{B}}\ln(2q^{-1}-2)$ and $T<0$ for all $q\in\left(\frac{2}{3},1\right]$. The reason for this strange state of affairs is that the entropy S is bounded, and is not an monotonically increasing function of the energy E (or the dimensionless quantity Q). The entropy is maximized for $N\uparrow=N_0=N_{\downarrow}=\frac{1}{3}$, 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 the interface between the fluids, relative to equilibrium. The potential energy is a sum of gravitational and surface tension terms, with

$$U_{\mathrm{grav}}=\int\! d^2x\int_0^z\! dz'\,\Delta\rho\, g\, z'\qquad,\qquad U_{\mathrm{surf}}=\tfrac{1}{2}\sigma\int\! d^2x\,(\boldsymbol{\nabla} z)^2\quad.\tag{4.317}$$

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

$$T=\int\! d^2x\int\! d^2x'\,\tfrac{1}{2}\mu(\boldsymbol{x},\boldsymbol{x}')\,\frac{\partial z(\boldsymbol{x},t)}{\partial t}\,\frac{\partial z(\boldsymbol{x}',t)}{\partial t}\quad.\tag{4.318}$$

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

$$z(\boldsymbol{x})=\big(L_x\,L_y\big)^{-1/2}\sum_{\boldsymbol{k}} z_{\boldsymbol{k}}\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\quad,\tag{4.319}$$

where the wavevectors $\boldsymbol{k}$ are quantized according to

$$\boldsymbol{k}=\frac{2\pi n_x}{L_x}\,\hat{\boldsymbol{x}}+\frac{2\pi n_y}{L_y}\,\hat{\boldsymbol{y}}\quad,\tag{4.320}$$

with integer n_x and n_y, if we impose periodic boundary conditions (for calculational convenience). The Lagrangian is then

$$L = \frac{1}{2}\sum_{\boldsymbol{k}} \Big[\mu_{\boldsymbol{k}}\,\big|\dot z_{\boldsymbol{k}}\big|^2 - \big(g\,\Delta\rho + \sigma\boldsymbol{k}^2\big)\,\big|z_{\boldsymbol{k}}\big|^2\Big] \quad , \tag{4.321}$$

where

$$\mu_{\boldsymbol{k}} = \int\! d^2\!x\;\mu\big(|\boldsymbol{x}|\big)\,e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \quad . \tag{4.322}$$

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

$$p_{\boldsymbol{k}} = \frac{\partial L}{\partial \dot z^*_{\boldsymbol{k}}} = \mu_{\boldsymbol{k}}\,\dot z_{\boldsymbol{k}} \qquad , \qquad p^*_{\boldsymbol{k}} = \frac{\partial L}{\partial \dot z_{\boldsymbol{k}}} = \mu_{\boldsymbol{k}}\,\dot z^*_{\boldsymbol{k}} \tag{4.323}$$

The Hamiltonian is then

$$\begin{aligned}\hat H &= {\sum_{\boldsymbol{k}}}' \Big[p_{\boldsymbol{k}}\,z^*_{\boldsymbol{k}} + p^*_{\boldsymbol{k}}\,z_{\boldsymbol{k}}\Big] - L \\ &= {\sum_{\boldsymbol{k}}}' \left[\frac{|p_{\boldsymbol{k}}|^2}{\mu_{\boldsymbol{k}}} + \big(g\,\Delta\rho + \sigma\boldsymbol{k}^2\big)\,|z_{\boldsymbol{k}}|^2\right] \quad ,\end{aligned} \tag{4.324}$$

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:

$$\begin{aligned}Z &= {\prod_{\boldsymbol{k}}}' \int\!\frac{d^2\!p_{\boldsymbol{k}}\,d^2\!z_{\boldsymbol{k}}}{(2\pi\hbar)^2}\;e^{-|p_{\boldsymbol{k}}|^2/\mu_{\boldsymbol{k}} k_{\rm B}T}\;e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \\ &= {\prod_{\boldsymbol{k}}}' \left(\frac{k_{\rm B}T}{2\hbar}\right)^{\!2}\left(\frac{\mu_{\boldsymbol{k}}}{g\,\Delta\rho + \sigma\boldsymbol{k}^2}\right) \quad .\end{aligned} \tag{4.325}$$

Thus,

$$F = -k_{\rm B}T \sum_{\boldsymbol{k}} \ln\!\left(\frac{k_{\rm B}T}{2\hbar\varOmega_{\boldsymbol{k}}}\right) \quad , \tag{4.326}$$

where[13]

$$\varOmega_{\boldsymbol{k}} = \left(\frac{g\,\Delta\rho + \sigma\boldsymbol{k}^2}{\mu_{\boldsymbol{k}}}\right)^{\!1/2} \quad . \tag{4.327}$$

is the normal mode frequency for surface oscillations at wavevector $\boldsymbol{k}$. For deep water waves, it is appropriate to take $\mu_{\boldsymbol{k}} = \Delta\rho\big/|\boldsymbol{k}|$, 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

$$\begin{aligned}\big\langle\, |z_{\boldsymbol{k}}|^2 \,\big\rangle &= \int\! d^2\!z_{\boldsymbol{k}}\;|z_{\boldsymbol{k}}|^2\;e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \bigg/ \int\! d^2\!z_{\boldsymbol{k}}\;e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \\ &= \frac{k_{\rm B}T}{g\,\Delta\rho + \sigma\boldsymbol{k}^2} \quad .\end{aligned} \tag{4.328}$$

---

[13] 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.

Note that this result does not depend on $\mu_{\boldsymbol{k}}$, *i.e.* on our choice of kinetic energy. One defines the *correlation function*

$$\begin{aligned} C(\boldsymbol{x}) \equiv \big\langle\, z(\boldsymbol{x})\, z(0) \,\big\rangle &= \frac{1}{L_x L_y} \sum_{\boldsymbol{k}} \big\langle\, |z_{\boldsymbol{k}}|^2 \,\big\rangle\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = \int \frac{d^2k}{(2\pi)^2} \left(\frac{k_{\rm B}T}{g\,\Delta\rho + \sigma \boldsymbol{k}^2} \right) e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \\ &= \frac{k_{\rm B}T}{4\pi\sigma} \int\limits_0^\infty \! dq\, \frac{e^{ik|\boldsymbol{x}|}}{\sqrt{q^2+\xi^2}} = \frac{k_{\rm B}T}{4\pi\sigma}\, K_0\big(|\boldsymbol{x}|/\xi\big) \quad , \end{aligned} \tag{4.329}$$

where $\xi = \sqrt{g\,\Delta\rho/\sigma}$ is the correlation length, and where $K_0(z)$ is the Bessel function of imaginary argument. The asymptotic behavior of $K_0(z)$ for small z is $K_0(z) \sim \ln(2/z)$, 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 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

$$\mathrm{H} \rightleftharpoons \mathrm{p}^+ + \mathrm{e}^- \tag{4.330}$$

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

$$n_{\rm H} = (1-x)\,n \qquad , \qquad n_{\rm p} = xn \qquad , \qquad n_{\rm e} = xn \quad . \tag{4.331}$$

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)

$$\mu_{\rm H} = \mu_{\rm p} + \mu_{\rm e} \quad . \tag{4.332}$$

The single particle partition function for each species is

$$\zeta = \frac{g^N}{N!} \left(\frac{V}{\lambda_T^3} \right)^N e^{-N\varepsilon_{\rm int}/k_{\rm B}T} \quad , \tag{4.333}$$

where g is the degeneracy and $\varepsilon_{\rm int}$ the internal energy for a given species. We have $\varepsilon_{\rm int} = 0$ for p and e, and $\varepsilon_{\rm int} = -\Delta$ for H, where $\Delta = e^2/2a_{\rm B} = 13.6\,$eV, the binding energy of hydrogen. Neglecting hyperfine splittings[14], we have $g_{\rm H} = 4$, while $g_{\rm e} = g_{\rm p} = 2$ because each has spin $S = \frac{1}{2}$. Thus, the

---

[14] The hyperfine splitting in hydrogen is on the order of $(m_{\rm e}/m_{\rm p})\,\alpha^4\, m_{\rm e}c^2 \sim 10^{-6}\,$eV, which is on the order of $0.01\,$K. Here $\alpha = e^2/\hbar c$ is the fine structure constant.

associated grand potentials are

$$\begin{aligned}
\Omega_\mathrm{H}(T,V,\mu_\mathrm{H}) &= -g_\mathrm{H}\, V k_\mathrm{B} T\, \lambda_{T,\mathrm{H}}^{-3}\, e^{(\mu_\mathrm{H}+\Delta)/k_\mathrm{B}T} \\
\Omega_\mathrm{p}(T,V,\mu_\mathrm{p}) &= -g_\mathrm{p}\, V k_\mathrm{B} T\, \lambda_{T,\mathrm{p}}^{-3}\, e^{\mu_\mathrm{p}/k_\mathrm{B}T} \qquad (4.334) \\
\Omega_\mathrm{e}(T,V,\mu_\mathrm{e}) &= -g_\mathrm{e}\, V k_\mathrm{B} T\, \lambda_{T,\mathrm{e}}^{-3}\, e^{\mu_\mathrm{e}/k_\mathrm{B}T} \quad ,
\end{aligned}$$

where

$$\lambda_{T,a} = \sqrt{\frac{2\pi\hbar^2}{m_a k_\mathrm{B} T}} \qquad (4.335)$$

for species a. The corresponding number densities are

$$n = \frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = g\,\lambda_T^{-3}\, e^{(\mu-\varepsilon_\mathrm{int})/k_\mathrm{B}T} \quad , \qquad (4.336)$$

and the fugacity $z = e^{\mu/k_\mathrm{B}T}$ of a given species is thus given by

$$z = g^{-1} n\, \lambda_T^3\, e^{\varepsilon_\mathrm{int}/k_\mathrm{B}T} \quad . \qquad (4.337)$$

We now invoke $\mu_\mathrm{H} = \mu_\mathrm{p} + \mu_\mathrm{e}$, which says $z_\mathrm{H} = z_\mathrm{p}\, z_\mathrm{e}$, or

$$g_\mathrm{H}^{-1} n_\mathrm{H}\, \lambda_{T,\mathrm{H}}^3\, e^{-\Delta/k_\mathrm{B}T} = \big(g_\mathrm{p}^{-1} n_\mathrm{p}\, \lambda_{T,\mathrm{p}}^3\big)\big(g_\mathrm{e}^{-1} n_\mathrm{e}\, \lambda_{T,\mathrm{e}}^3\big) \quad , \qquad (4.338)$$

which yields

$$\left(\frac{x^2}{1-x}\right) n\tilde\lambda_T^3 = e^{-\Delta/k_\mathrm{B}T} \quad , \qquad (4.339)$$

where $\tilde\lambda_T = \sqrt{2\pi\hbar^2/m^* k_\mathrm{B} T}$, with $m^* = m_\mathrm{p} m_\mathrm{e}/m_\mathrm{H} \approx m_\mathrm{e}$. Note that

$$\tilde\lambda_T = a_\mathrm{B} \sqrt{\frac{4\pi m_\mathrm{H}}{m_\mathrm{p}}}\, \sqrt{\frac{\Delta}{k_\mathrm{B} T}} \quad , \qquad (4.340)$$

where $a_\mathrm{B} = 0.529\,\text{Å}$ is the Bohr radius. Thus, we have

$$\left(\frac{x^2}{1-x}\right)\cdot (4\pi)^{3/2}\,\nu = \left(\frac{T}{T_0}\right)^{3/2} e^{-T_0/T} \quad , \qquad (4.341)$$

where $T_0 = \Delta/k_\mathrm{B} = 1.578\times 10^5\,\mathrm{K}$ and $\nu = n a_\mathrm{B}^3$. Consider for example a temperature $T = 3000\,\mathrm{K}$, for which $T_0/T = 52.6$, and assume that $x = \frac{1}{2}$. We then find $\nu = 1.69\times 10^{-27}$, corresponding to a density of $n = 1.14\times 10^{-2}\,\mathrm{cm}^{-3}$. At this temperature, the fraction of hydrogen atoms in their first excited (2s) state is $x' \sim e^{-T_0/2T} = 3.8\times 10^{-12}$. This is quite striking: half the hydrogen atoms are completely dissociated, which requires an energy of Δ, 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^{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^{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^{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

The quantum mechanical Hamiltonian for a many particle but non-interacting system may be written as

$$\hat{H}=\sum_\alpha \varepsilon_\alpha\,\hat{n}_\alpha \quad , \tag{5.1}$$

where $\hat{n}_\alpha$ is the number of particles in the single particle quantum state $|\,\alpha\,\rangle$ with energy ε_α. 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

$$\hat{n}_\alpha\,|\,\vec{n}\,\rangle = n_\alpha\,|\,\vec{n}\,\rangle \tag{5.2}$$

and

$$\hat{H}\,|\,\vec{n}\,\rangle=\sum_\alpha n_\alpha\,\varepsilon_\alpha\,|\,\vec{n}\,\rangle \quad . \tag{5.3}$$

The eigenvalues n_α 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 $Z(T,N)$ is then

$$Z(T,N)=\sum_{\{n_\alpha\}} e^{-\beta\sum_\alpha n_\alpha\varepsilon_\alpha}\,\delta_{N\,,\sum_\alpha n_\alpha} \quad , \tag{5.5}$$

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

$$\begin{aligned}\varXi(T,\mu)&=\sum_N e^{\beta\mu N}\,Z(T,N)\\ &=\sum_{\{n_\alpha\}} e^{-\beta\sum_\alpha n_\alpha\varepsilon_\alpha}\,e^{\beta\mu\sum_\alpha n_\alpha}=\prod_\alpha\Bigg(\sum_{n_\alpha} e^{-\beta(\varepsilon_\alpha-\mu)\,n_\alpha}\Bigg) \quad .\end{aligned} \tag{5.6}$$

Note that the grand partition function $\varXi$ takes the form of a product over contributions from the individual single particle states. Recall that the grand potential $\varOmega$ is given by $\varOmega=-k_\mathrm{B}T\log\varXi$, *i.e.* $\varXi=\exp(-\varOmega/k_\mathrm{B}T)$.

By the way, our notation here is a bit loose and possibly misleading. Formally we should write $Z = Z(T, N, \{X_j\})$ where each X_j is an independent extensive variable changes in which under equilibrium conditions result in thermodynamic work. Similarly, $\Xi = \Xi(T, \mu, \{X_j\})$.

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[1]

$$\begin{aligned} \Xi^{\mathsf{BE}} &= \prod_\alpha \frac{1}{1-e^{-(\varepsilon_\alpha-\mu)/k_{\mathrm{B}}T}} \\ \Omega^{\mathsf{BE}} &= -k_{\mathrm{B}}T\log\Xi^{\mathsf{BE}} = k_{\mathrm{B}}T\sum_\alpha \log\Big(1-e^{-(\varepsilon_\alpha-\mu)/k_{\mathrm{B}}T}\Big) \end{aligned} \tag{5.8}$$

for Bose-Einstein statistics and

$$\begin{aligned} \Xi^{\mathsf{FD}} &= \prod_\alpha \Big(1+e^{-(\varepsilon_\alpha-\mu)/k_{\mathrm{B}}T}\Big) \\ \Omega^{\mathsf{FD}} &= -k_{\mathrm{B}}T\log\Xi^{\mathsf{FD}} = -k_{\mathrm{B}}T\sum_\alpha \log\Big(1+e^{-(\varepsilon_\alpha-\mu)/k_{\mathrm{B}}T}\Big) \end{aligned} \tag{5.9}$$

for Fermi-Dirac statistics. We can combine these expressions into one, writing

$$\Omega^{\mathsf{B/F}}(T,V,\mu) = \pm k_{\mathrm{B}}T\sum_\alpha \log\Big(1\mp e^{-(\varepsilon_\alpha-\mu)/k_{\mathrm{B}}T}\Big) \quad , \tag{5.10}$$

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 α is

$$\langle\hat{n}_\alpha\rangle = \frac{\partial\Omega^{\mathsf{B/F}}}{\partial\varepsilon_\alpha} = \frac{1}{e^{(\varepsilon_\alpha-\mu)/k_{\mathrm{B}}T}\mp 1} \qquad , \qquad n^{\mathsf{B/F}}_\varepsilon(T,\mu) = \frac{1}{e^{(\varepsilon-\mu)/k_{\mathrm{B}}T}\mp 1} \quad . \tag{5.11}$$

Thus $n^{\mathsf{B/F}}_\varepsilon(T,\mu)$ is the average occupation of single particle states of energy ε for bosons (B) and fermions (F) at temperature T and chemical potential μ. The total particle number is then

$$N^{\mathsf{B/F}}(T,V,\mu) = \sum_\alpha \frac{1}{e^{(\varepsilon_\alpha-\mu)/k_{\mathrm{B}}T}\mp 1} \quad . \tag{5.12}$$

We will henceforth write $n_\alpha(\mu,T) = \langle\hat{n}_\alpha\rangle$ for the thermodynamic average of this occupancy.

---

[1] Note that convergence of the partition sum for bosons requires $\exp((\mu-\varepsilon_\alpha)/k_{\mathrm{B}}T) < 1$, which is to say that $\mu < \min_\alpha \varepsilon_\alpha$ for all α.

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 ε_α. What is the partition function? Is it

$$Z \overset{?}{=} \sum_{\alpha_1} \cdots \sum_{\alpha_N} e^{-\beta\,(\varepsilon_{\alpha_1} + \,\varepsilon_{\alpha_2} + \,\ldots\, + \,\varepsilon_{\alpha_N})} = \zeta^N \quad , \tag{5.14}$$

where ζ 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_α, 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 ζ^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

$$\text{degree of overcounting} \quad = \quad \frac{N!}{\prod_\alpha n_\alpha!} \quad ,$$

which is the number of ways we can rearrange the labels α_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_α 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 $\varepsilon_\alpha - \mu \gg k_\mathrm{B}T$, which is to say $\mu \ll \varepsilon_0 - k_{\rm B}T$ where ε_0 is the minimum single particle energy. Consequently,

$$\begin{aligned} \Omega^{\sf B/F} &= \pm k_{\rm B}T \sum_\alpha \log\Big(1 \mp e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T}\Big) \\ &\approx -k_{\rm B}T \sum_\alpha e^{-(\varepsilon_\alpha-\mu)/k_{\rm B}T} \equiv \Omega^{\sf MB} \quad . \end{aligned} \tag{5.17}$$

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 Single particle density of states

The *single particle density of states per unit volume* $g(\varepsilon)$ is defined as

$$g(\varepsilon) = \frac{1}{V} \sum_\alpha \delta(\varepsilon - \varepsilon_\alpha) \quad . \tag{5.19}$$

The concept of density of states is an important one and the student should develop some facility with it. Note that the dimensions of $g(\varepsilon)$ and $\big[g(\varepsilon)\big] = E^{-1}L^{-d}$, where E stands for energy and L for length. We may now write

$$\Omega^{\sf B/F}(T,V,\mu) = \pm V k_{\rm B}T \int\limits_{-\infty}^{\infty} d\varepsilon\; g(\varepsilon)\; \log\Big(1 \mp e^{-(\varepsilon-\mu)/k_{\rm B}T}\Big) \quad . \tag{5.20}$$

For particles with a dispersion $\varepsilon(\boldsymbol{k})$, with $\boldsymbol{p} = \hbar\boldsymbol{k}$, we have

$$g(\varepsilon) = {\sf g} \int \frac{d^d\!k}{(2\pi)^d}\, \delta\big(\varepsilon - \varepsilon(\boldsymbol{k})\big) = \frac{{\sf g}\,\Omega_d}{(2\pi)^d}\, \frac{k^{d-1}}{d\varepsilon/dk} \quad . \tag{5.21}$$

Here, ${\sf g} = 2S+1$ is the spin degeneracy, and $\Omega_d = 2\pi^{d/2}/\Gamma(d/2)$ is the surface area of a unit sphere in d dimensions, with $\Gamma(z)$ the gamma function. Furthermore, we assume that $\varepsilon(\boldsymbol{k})$ is both isotropic and monotonically increasing function in k. Then we have

$$g_{d=1}(\varepsilon) = \frac{\sf g}{\pi}\,\frac{dk}{d\varepsilon} \quad , \quad g_{d=2}(\varepsilon) = \frac{\sf g}{2\pi}\,k\,\frac{dk}{d\varepsilon} \quad , \quad g_{d=3}(\varepsilon) = \frac{\sf g}{2\pi^2}\,k^2\,\frac{dk}{d\varepsilon} \quad . \tag{5.22}$$

In order to obtain $g(\varepsilon)$ as a function of the energy ε one must invert the dispersion relation $\varepsilon = \varepsilon(k)$ to obtain $k = k(\varepsilon)$. A quick way to derive the result in eqn. 5.21 is to write

$$g(\varepsilon)\, d\varepsilon = {\sf g}\, \frac{d^d\!k}{(2\pi)^d} = \frac{{\sf g}\,\Omega_d}{(2\pi)^d}\, k^{d-1}\, dk \quad . \tag{5.23}$$

For a general isotropic power law dispersion $\varepsilon(\boldsymbol{k}) = A\,|\boldsymbol{k}|^\sigma$, this yields

$$\varepsilon(\boldsymbol{k}) = A\,k^\sigma \qquad \Rightarrow \qquad g(\varepsilon) = \frac{{\sf g}\,\Omega_d}{(2\pi)^d\sigma}\, A^{-d/\sigma}\, \varepsilon^{\frac{d}{\sigma}-1}\, \Theta(\varepsilon) \quad . \tag{5.24}$$

For a ballistic dispersion $\varepsilon(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m$, we have[2]

$$g(\varepsilon) = \frac{\mathsf{g}}{\Gamma(d/2)}\left(\frac{m}{2\pi\hbar^2}\right)^{d/2} \varepsilon^{(d-2)/2}\,\Theta(\varepsilon) \quad , \tag{5.25}$$

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.

Note that

$$n^{\mathsf{B/F}}_\varepsilon(T,\mu) = \frac{1}{e^{(\varepsilon-\mu)/k_\mathrm{B}T} \mp 1} \quad . \tag{5.26}$$

This result is valid independent of the form of $g(\varepsilon)$. The average total number of particles is then

$$N^{\mathsf{B/F}}(T,V,\mu) = V\!\int\limits_{-\infty}^{\infty}\! d\varepsilon\; g(\varepsilon)\; \overbrace{\frac{1}{e^{(\varepsilon-\mu)/k_\mathrm{B}T} \mp 1}}^{n^{\mathsf{B/F}}_\varepsilon(T,\mu)} \quad , \tag{5.27}$$

which does depend on $g(\varepsilon)$.

5.3 Expansion in Powers of the Density

5.3.1 Expansion in powers of the fugacity

From eqn. 5.27, we have that the number density $n = N/V$ is

$$n(T,z) = \int\limits_{-\infty}^{\infty}\! d\varepsilon\; \frac{g(\varepsilon)}{z^{-1}\,e^{\varepsilon/k_\mathrm{B}T} \mp 1} = \sum_{j=1}^{\infty} C_j(T)\, z^j \quad , \tag{5.28}$$

where $z = \exp(\mu/k_\mathrm{B}T)$ is the *fugacity* and

$$C_j(T) = (\pm 1)^{j-1}\!\int\limits_{-\infty}^{\infty}\! d\varepsilon\; g(\varepsilon)\, e^{-j\varepsilon/k_\mathrm{B}T} \quad . \tag{5.29}$$

Note that $[C_j] = V^{-1}$ for all j. From $\Omega = -pV$ and our expression above for $\Omega(T,V,\mu)$, we have

$$p(T,z) = \mp\, k_\mathrm{B}T\!\int\limits_{-\infty}^{\infty}\! d\varepsilon\; g(\varepsilon)\; \log\!\Big(1 \mp z\, e^{-\varepsilon/k_\mathrm{B}T}\Big) = k_\mathrm{B}T \sum_{j=1}^{\infty} j^{-1} C_j(T)\, z^j \quad . \tag{5.30}$$

Eqns. 5.28 and 5.30 express $n(T,z)$ and $p(T,z)$ as power series in the fugacity z, with T-dependent coefficients. In principal, we can eliminate z using eqn. 5.28, writing $z = z(T,n)$ as a power series in

---

[2] 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 $\Delta\varepsilon \gg \delta\varepsilon$.

the number density n, and substitute this into eqn. 5.30 to obtain an equation of state $p = p(T, n)$ of the form

$$p(T, n) = n\, k_{\rm B} T \Big(1 + B_2(T)\, n + B_3(T)\, n^2 + \ldots\Big) \quad . \tag{5.31}$$

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.

Unfortunately, there is no general analytic expression for the virial coefficients $B_j(T)$ in terms of the expansion coefficients $n_j(T)$. One can grind through in an order-by-order calculation, with increasing labor at each successive order. Proceeding thusly, the first few virial coefficients are given by

$$B_1 = 1 \qquad , \qquad B_2 = -\frac{C_2}{2\, C_1^2} \qquad , \qquad B_3 = \frac{C_2^2}{C_1^4} - \frac{2\, C_3}{3\, C_1^3} \quad . \tag{5.32}$$

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,

$$\bigg(p + \frac{aN^2}{V^2}\bigg)(V - Nb) = N k_{\rm B} T \quad . \tag{5.33}$$

This may be recast as

$$\begin{aligned} p &= \frac{n k_{\rm B} T}{1 - bn} - an^2 \\ &= n k_{\rm B} T + \big(b\, k_{\rm B} T - a\big)\, n^2 + k_{\rm B} T\, b^2 n^3 + k_{\rm B} T\, b^3 n^4 + \ldots \quad , \end{aligned} \tag{5.34}$$

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 $k \geq 3$.

5.3.2 Ballistic dispersion

For the ballistic dispersion $\varepsilon(\boldsymbol{p}) = \boldsymbol{p}^2/2m$ we computed the density of states in eqn. 5.25. One finds

$$C_j(T) = (\pm 1)^{j-1} \frac{\mathsf{g}\, \lambda_T^{-d}}{\Gamma(d/2)} \int\limits_0^\infty \! dt\, t^{\frac{d}{2}-1}\, e^{-jt} = (\pm 1)^{j-1} j^{-d/2}\, \mathsf{g}\, \lambda_T^{-d} \quad , \tag{5.35}$$

where $\lambda_T = \sqrt{2\pi\hbar^2/m k_{\rm B} T}$ is the thermal wavelength.

We may also write

$$n(T, z) = \pm \mathsf{g}\, \lambda_T^{-d}\ \mathrm{Li}_{\frac{d}{2}}(\pm z) \tag{5.36}$$

and

$$p(T, z) = \pm \mathsf{g}\, k_{\rm B} T\, \lambda_T^{-d}\ \mathrm{Li}_{\frac{d}{2}+1}(\pm z) \quad , \tag{5.37}$$

where

$$\mathsf{Li}_s(z) \equiv \sum_{n=1}^{\infty} \frac{z^n}{n^s} \tag{5.38}$$

is the *polylogarithm function*. Note that $\mathsf{Li}_s(z)$ obeys a recursion relation in its index, *viz.*

$$z\,\frac{\partial}{\partial z}\,\mathsf{Li}_s(z) = \mathsf{Li}_{s-1}(z) \quad , \tag{5.39}$$

and that

$$\zeta(s) = \mathsf{Li}_s(1) = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{5.40}$$

is the Riemann zeta function.

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 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

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T.V} = 0 \quad . \tag{5.41}$$

The number distribution, from eqn. 5.11, is then

$$n_\varepsilon(T) = \frac{1}{e^{\varepsilon/k_\mathrm{B}T} - 1} \quad . \tag{5.42}$$

The grand partition function for a system of particles with $\mu = 0$ is

$$\Omega(T,V) = V k_\mathrm{B}T \int\limits_{-\infty}^{\infty} \! d\varepsilon\, g(\varepsilon)\, \log\left(1 - e^{-\varepsilon/k_\mathrm{B}T}\right) \quad , \tag{5.43}$$

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.*

$$g(\varepsilon) = \frac{\mathsf{g}\Omega_d}{(2\pi)^d}\,\frac{k^{d-1}}{d\varepsilon/dk} = \left(\frac{\mathsf{g}\,\Omega_d}{(2\pi)^d\,\sigma A^{d/\sigma}}\right) \varepsilon^{\frac{d}{\sigma}-1}\,\Theta(\varepsilon) \quad . \tag{5.44}$$

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

$$n(T) = C\int\limits_0^\infty \! d\varepsilon\,\frac{\varepsilon^{r-1}}{e^{\varepsilon/k_\mathrm{B}T}-1} = C\,\zeta(r)\,\Gamma(r)\,(k_\mathrm{B}T)^r \tag{5.45}$$

and

$$p(T) = -C\int\limits_0^\infty \! d\varepsilon\,\varepsilon^{r-1}\log\!\big(1-e^{-\varepsilon/k_\mathrm{B}T}\big) = C\,\zeta(r+1)\,\Gamma(r)\,(k_\mathrm{B}T)^{r+1} \quad . \tag{5.46}$$

To work out the above integrals, first change variables to $t = \varepsilon/k_\mathrm{B}T$, then expand in powers of $\exp(-t)$, integrate over t, and express the sums in terms of Riemann zeta (see eqn. 5.40). Dividing these two equations, we obtain the equation of state

$$p = \frac{\zeta(r+1)}{\zeta(r)}\,nk_\mathrm{B}T \quad . \tag{5.47}$$

To find the entropy, we use Gibbs-Duhem:

$$d\mu = 0 = -s\,dT + v\,dp \quad\Longrightarrow\quad s = v\,\frac{dp}{dT} \quad , \tag{5.48}$$

where s is the entropy per particle and $v = n^{-1}$ is the volume per particle. In d space dimensions,

$$s(T) = (r{+}1)\,\frac{\zeta(r{+}1)}{\zeta(r)}\,k_\mathrm{B} \quad . \tag{5.49}$$

The entropy per particle is constant. The internal energy is

$$E = -\frac{\partial\log\Xi}{\partial\beta} = -\frac{\partial}{\partial\beta}\big(\beta pV\big) = rpV \quad , \tag{5.50}$$

and hence the energy per particle is

$$\varepsilon = \frac{E}{N} = rpv = \frac{r\,\zeta(r{+}1)}{\zeta(r)}\,k_\mathrm{B}T \quad . \tag{5.51}$$

For the photon gas, we have $\varepsilon(\boldsymbol{k}) = \hbar ck$, whence $\sigma = 1$, $r = d$, $A = \hbar c$, and

$$g(\varepsilon) = \bigg(\frac{2\mathrm{g}\,\pi^{d/2}}{\Gamma(d/2)(\hbar c)^d}\bigg)\,\varepsilon^{d-1}\,\Theta(\varepsilon) \quad . \tag{5.52}$$

In $d = 3$ dimensions the degeneracy is $\mathrm{g} = 2$, which is the number of independent polarization states. Thus, for the photon gas in $d = 3$ dimensions, we have $\mathrm{g} = 2$ and $C = 1/\pi^2\hbar^3c^3$, and

$$n(T) = \frac{2\,\zeta(3)}{\pi^2}\,\frac{(k_\mathrm{B}T)^3}{(\hbar c)^3} \quad , \quad p(T) = \frac{2\,\zeta(4)}{\pi^2}\,\frac{(k_\mathrm{B}T)^4}{(\hbar c)^3} \quad . \tag{5.53}$$

It turns out that $\zeta(3) = 1.20206$ and $\zeta(4) = \pi^4/90$. We also have $\hbar c/k_\mathrm{B} = 0.2290\,\mathrm{cm}\cdot\mathrm{K}$, and thus $k_\mathrm{B}T/\hbar c = 4.3755\,T[\mathrm{K}]\,\mathrm{cm}^{-1}$. We then obtain

$$p = 0.90039\,nk_\mathrm{B}T \quad , \quad n(T) = 20.286\times T^3[\mathrm{K}^3]\,\mathrm{cm}^{-3} \quad . \tag{5.54}$$

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 λ. The cavity modes of the electromagnetic radiation have quantized wavevectors, even within classical electromagnetic theory, given by

$$\boldsymbol{k} = \left(\frac{2\pi n_x}{L_x}\,,\,\frac{2\pi n_y}{L_y}\,,\,\frac{2\pi n_z}{L_z}\right) \quad . \tag{5.55}$$

   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,

$$V\left(\frac{\partial E}{\partial V}\right)_{\!S} = \lambda\left(\frac{\partial E}{\partial \lambda}\right)_{\!S} = -\tfrac{1}{3}E \quad . \tag{5.56}$$

   Thus,

$$p = -\left(\frac{\partial E}{\partial V}\right)_{\!S} = \frac{E}{3V} \quad , \tag{5.57}$$

   as we found in eqn. 5.50. Since $E = E(T,V)$ is extensive, we must have $p = p(T)$ alone.

2. Since $p = p(T)$ alone, we have

$$\begin{aligned}\left(\frac{\partial E}{\partial V}\right)_{\!T} &= \left(\frac{\partial E}{\partial V}\right)_{\!p} = 3p \\ &= T\left(\frac{\partial p}{\partial T}\right)_{\!V} - p \quad ,\end{aligned} \tag{5.58}$$

   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,

$$T\,\frac{dp}{dT} = 4p \quad \Longrightarrow \quad p(T) = A\,T^4 \quad , \tag{5.59}$$

   where A is a constant. Thus, we recover the temperature dependence found microscopically in eqn. 5.46.

3. Given an energy density E/V, the differential energy flux emitted in a direction θ relative to a surface normal is

$$dj_\varepsilon = c \cdot \frac{E}{V} \cdot \cos\theta \cdot \frac{d\Omega}{4\pi} \quad , \tag{5.60}$$

   where $d\Omega$ is the differential solid angle. Thus, the power emitted per unit area is

$$\frac{dP}{dA} = \frac{cE}{4\pi V}\int\limits_0^{\pi/2}\! d\theta \int\limits_0^{2\pi}\! d\phi\ \sin\theta \cdot \cos\theta = \frac{cE}{4V} = \tfrac{3}{4}\,c\,p(T) \equiv \sigma\,T^4 \quad , \tag{5.61}$$

where $\sigma = \frac{3}{4} cA$, with $p(T) = AT^4$ as we found above. From quantum statistical mechanical considerations, we have

$$\sigma = \frac{\pi^2 k_{\rm B}^4}{60\, c^2\, \hbar^3} = 5.67 \times 10^{-8}\, \frac{\rm W}{\rm m^2\, K^4} \tag{5.62}$$

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

$$P_{\rm incident} = \frac{\pi R_{\rm e}^2}{4\pi a_{\rm e}^2} \cdot \sigma T_\odot^4 \cdot 4\pi R_\odot^2 = \frac{R_{\rm e}^2\, R_\odot^2}{a_{\rm e}^2} \cdot \pi \sigma T_\odot^4 \quad . \tag{5.63}$$

The power radiated by the earth is

$$P_{\rm radiated} = \sigma T_{\rm e}^4 \cdot 4\pi R_{\rm e}^2 \quad . \tag{5.64}$$

Setting $P_{\rm incident} = P_{\rm radiated}$, we obtain

$$T_{\rm e} = \left(\frac{R_\odot}{2\, a_{\rm e}}\right)^{1/2} T_\odot \quad . \tag{5.65}$$

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*.

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)

$$\mathcal{N}(\nu, T)\, d\nu = \frac{2\, V}{8\pi^3} \cdot \frac{d^3\!k}{e^{\hbar ck/k_{\rm B}T} - 1} = \frac{V}{\pi^2} \cdot \frac{k^2\, dk}{e^{\hbar ck/k_{\rm B}T} - 1} \quad . \tag{5.66}$$

We therefore have

$$\mathcal{N}(\nu, T) = \frac{8\pi V}{c^3} \cdot \frac{\nu^2}{e^{h\nu/k_{\rm B}T} - 1} \quad . \tag{5.67}$$

Since a photon of frequency ν carries energy $h\nu$, the energy per unit frequency $\mathcal{E}(\nu)$ is

$$\mathcal{E}(\nu, T) = \frac{8\pi h V}{c^3} \cdot \frac{\nu^3}{e^{h\nu/k_{\rm B}T} - 1} \quad . \tag{5.68}$$

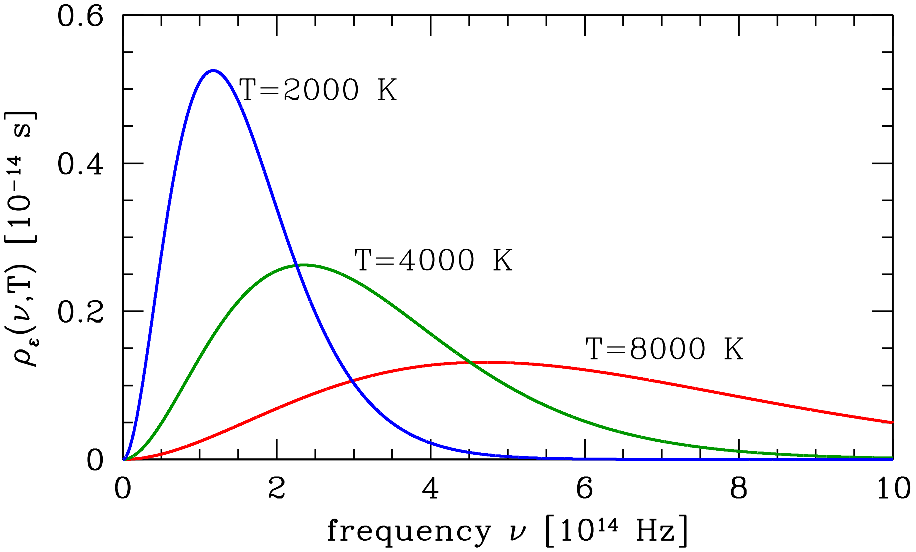


Figure 5.1: Spectral density $\rho_\varepsilon(\nu,T)$ for blackbody radiation at three temperatures.

Note what happens if Planck's constant h vanishes, as it does in the classical (Maxwell-Boltzmann) limit. The denominator can then be written

$$e^{h\nu/k_\mathrm{B}T} - 1 = \frac{h\nu}{k_\mathrm{B}T} + \mathcal{O}(h^2) \tag{5.69}$$

and

$$\mathcal{E}_\mathrm{MB}(\nu,T) = \lim_{h\to 0} \mathcal{E}(\nu) = V \cdot \frac{8\pi k_\mathrm{B}T}{c^3}\,\nu^2 \quad . \tag{5.70}$$

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 ν part of the integral, which in the optical spectrum is the ultraviolet portion. With quantization, the Bose-Einstein factor imposes an effective ultraviolet cutoff $k_\mathrm{B}T/h$ on the frequency integral, and the total energy, as we found above, is finite:

$$E(T) = \int_0^\infty \! d\nu \, \mathcal{E}(\nu) = 3pV = V \cdot \frac{\pi^2}{15} \, \frac{(k_\mathrm{B}T)^4}{(\hbar c)^3} \quad . \tag{5.71}$$

We can define the spectral density $\rho_\varepsilon(\nu)$ of the radiation as

$$\rho_\varepsilon(\nu,T) \equiv \frac{\mathcal{E}(\nu,T)}{E(T)} = \frac{15}{\pi^4} \, \frac{h}{k_\mathrm{B}T} \, \frac{(h\nu/k_\mathrm{B}T)^3}{e^{h\nu/k_\mathrm{B}T} - 1} \tag{5.72}$$

so that $\rho_\varepsilon(\nu,T)\, d\nu$ is the fraction of the electromagnetic energy, under equilibrium conditions, between frequencies ν and $\nu + d\nu$, *i.e.* $\int_0^\infty d\nu\, \rho_\varepsilon(\nu,T) = 1$. In fig. 5.1 we plot this for three different temperatures. The maximum occurs at a frequency $\nu^* = s\, k_\mathrm{B}T/h$, where

$$\frac{d}{ds}\left(\frac{s^3}{e^s - 1}\right) = 0 \qquad \Longrightarrow \qquad \frac{s}{1 - e^{-s}} = 3 \qquad \Longrightarrow \qquad s = 2.82144 \quad . \tag{5.73}$$

Thus $\nu^* = 58.8\,\mathrm{GHz} \cdot T[\mathrm{K}]$.

5.4.5 What if the sun emitted ferromagnetic spin waves?

We saw in eqn. 5.60 that the power emitted per unit surface area by a blackbody is σT^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

$$dP = dA \cdot \int \frac{d^3\!k}{(2\pi)^3}\,\Theta(\cos\theta)\cdot\varepsilon(\boldsymbol{k})\cdot\frac{1}{\hbar}\frac{\partial\varepsilon(\boldsymbol{k})}{\partial k_z}\cdot\frac{1}{e^{\varepsilon(\boldsymbol{k})/k_{\mathrm{B}}T}-1} \quad . \tag{5.74}$$

Let us assume an isotropic power law dispersion of the form $\varepsilon(\boldsymbol{k}) = Ck^\alpha$. Then after a straightforward calculation we obtain

$$\frac{dP}{dA} = \widetilde{\sigma}\,T^{2+2\alpha^{-1}} \quad , \tag{5.75}$$

where the expression for $\widetilde{\sigma}$ is left as an exercise to the reader. One can check that for $\mathrm{g}=2$, $C=\hbar c$, and $\alpha=1$ that this result reduces to that of eqn. 5.62. 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

$$T_{\mathrm{e}} = \bigg(\frac{R_\odot}{2\,a_{\mathrm{e}}}\bigg)^{\alpha/(1+\alpha)} T_\odot \quad . \tag{5.76}$$

With $R_\odot/2a_{\mathrm{e}} = 2.32\times10^{-3}$, assuming the same value for $T_\odot = 5780\,\mathrm{K}$, and with $\alpha=2$, we obtain $T_{\mathrm{e}} = 101\,\mathrm{K}$.

5.5 The Ideal Bose Gas

5.5.1 General formulation for noninteracting systems

Recall that the grand partition function for noninteracting bosons is given by

$$\Xi = \prod_\alpha \bigg(\sum_{n_\alpha=0}^{\infty} e^{\beta(\mu-\varepsilon_\alpha)n_\alpha}\bigg) = \prod_\alpha \Big(1-e^{\beta(\mu-\varepsilon_\alpha)}\Big)^{-1} \quad , \tag{5.77}$$

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

$$n(T,\mu) = -\frac{1}{V}\bigg(\frac{\partial\Omega}{\partial\mu}\bigg)_{T,V} = \frac{1}{V}\sum_\alpha \frac{1}{e^{\beta(\varepsilon_\alpha-\mu)}-1} = \int_{\varepsilon_0}^{\infty} \! d\varepsilon\; g(\varepsilon)\,\overbrace{\frac{1}{e^{(\varepsilon-\mu)/k_{\mathrm{B}}T}-1}}^{n^{\mathrm{B}}_\varepsilon(T,\mu)} \quad , \tag{5.78}$$

where $g(\varepsilon) = V^{-1}\sum_\alpha \delta(\varepsilon-\varepsilon_\alpha)$ is the density of single particle states per unit volume. We assume that $g(\varepsilon)=0$ for $\varepsilon<\varepsilon_0$; typically $\varepsilon_0=0$, as is the case for any dispersion of the form $\varepsilon(\boldsymbol{k}) = A\,|\boldsymbol{k}|^r$, for example. However, in the presence of a magnetic field, we could have $\varepsilon(\boldsymbol{k},\sigma) = A\,|\boldsymbol{k}|^r - g\mu_0 H\sigma$, in which case $\varepsilon_0 = -g\mu_0|H|$.

Clearly $n(T,\mu)$ is an increasing function of both T and μ. At fixed T, the maximum possible value for $n(T,\mu)$, called the *critical density* $n_\text{c}(T)$, is achieved for $\mu = \varepsilon_0$, *i.e.*

$$n_\text{c}(T) = \int\limits_{\varepsilon_0}^{\infty}\! d\varepsilon \, \frac{g(\varepsilon)}{e^{(\varepsilon-\varepsilon_0)/k_\text{B}T} - 1} = \int\limits_0^\infty\! d\omega \, \frac{g(\varepsilon_0+\omega)}{e^{\omega/k_\text{B}T} - 1} \quad . \tag{5.79}$$

The above integral converges provided $g(\varepsilon_0) = 0$, assuming $g(\varepsilon)$ is continuous[3]. If $g(\varepsilon_0) > 0$, the integral diverges, and $n_\text{c}(T) = \infty$. In this latter case, one can always invert the equation for $n(T,\mu)$ to obtain the chemical potential $\mu(T,n)$. In the former case, where the $n_\text{c}(T)$ is finite, we have a problem – what happens if $n > n_\text{c}(T)$?

In the former case, where $n_\text{c}(T)$ is finite, we can equivalently restate the problem in terms of a *critical temperature* $T_\text{c}(n)$, defined by the equation $n_\text{c}(T_\text{c}) = n$. For $T < T_\text{c}$, 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 are discrete, and separate out the contribution from the lowest energy state ε_0. *I.e.* we write

$$n(T,\mu) = \overbrace{\frac{1}{V}\,\frac{\text{g}_0}{e^{\beta(\varepsilon_0-\mu)} - 1}}^{n_0} + \overbrace{\int\limits_0^\infty\! d\omega \, \frac{g(\varepsilon_0+\omega)}{e^{\beta\omega} - 1}}^{n'} \quad , \tag{5.80}$$

where g_0 is the degeneracy of the single particle state with energy ε_0. We assume that n_0 is finite, which means that $N_0 = n_0 V$ is extensive. We say that the particles have *condensed* into the one-body state $|\,\psi_0\,\rangle$ with energy ε_0 . The quantity n_0 is the *condensate number density*. The remaining particles, with number density n', are said to comprise the *overcondensate*. With the total density n fixed, we have $n = n_0 + n'$. Note that n_0 finite means that μ is infinitesimally close to ε_0:

$$\mu = \varepsilon_0 - k_\text{B}T\log\!\bigg(1 + \frac{\text{g}_0}{n_0 V}\bigg) \approx \varepsilon_0 - \frac{\text{g}_0 k_\text{B}T}{n_0 V} \quad . \tag{5.81}$$

Note also that if $\varepsilon_0 - \mu$ is finite, then $n_0 \propto V^{-1}$ is infinitesimal.

Thus, for $T < T_\text{c}(n)$, we have $\mu = \varepsilon_0$ with $n_0 > 0$, and

$$n(T,n_0) = n_0 + \overbrace{\int\limits_0^\infty\! d\omega \, \frac{g(\varepsilon_0+\omega)}{e^{\beta\omega} - 1}}^{n_\text{c}(T)} \quad . \tag{5.82}$$

For $T > T_\text{c}(n)$, we have $n_0 = 0$ and

$$n(T,\mu) = \int\limits_0^\infty\! d\omega \, \frac{g(\varepsilon_0+\omega)}{e^{\beta(\omega+\varepsilon_0-\mu)} - 1} \quad . \tag{5.83}$$

---

[3] 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 η is the lower cutoff. But for any power law density of states $g(\varepsilon) \propto \varepsilon^r$ with $r > 0$, the integral converges.

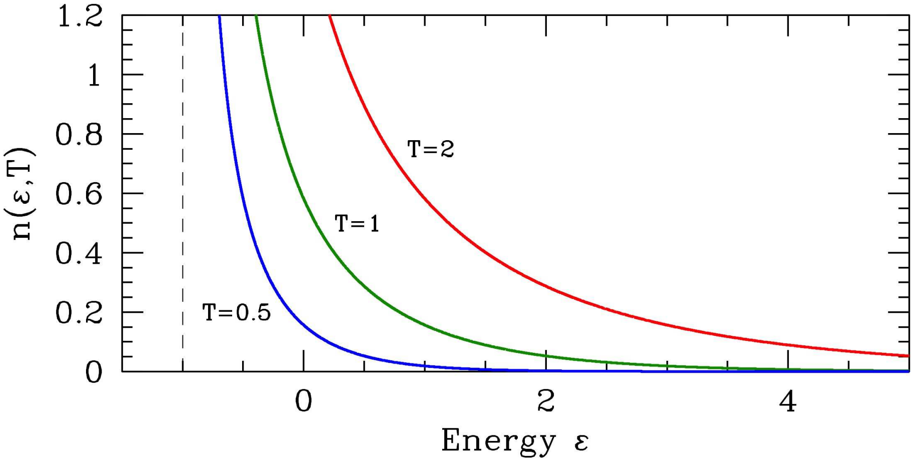


Figure 5.2: The Bose distribution $n^{\mathsf{B}}_{\varepsilon}(T,\mu)=1\big/\big[e^{(\varepsilon-\mu)/k_{\mathrm{B}}T}-1\big]$ *vs.* ε at fixed $\mu=-1$ ($k_{\mathrm{B}}\equiv 1$).

The equation relating T_{c} and n is

$$n=\int\limits_0^\infty \! d\omega\,\frac{g(\varepsilon_0+\omega)}{e^{\omega/k_{\mathrm{B}}T_{\mathrm{c}}}-1}\quad . \tag{5.84}$$

5.5.2 Ballistic dispersion

We already derived, in §5.3.2, 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

$$\begin{aligned} n(T,z)&=\mathsf{g}\,\lambda_T^{-d}\ \mathsf{Li}_{\frac{d}{2}}(z)\\ p(T,z)&=\mathsf{g}\,k_{\mathrm{B}}T\,\lambda_T^{-d}\ \mathsf{Li}_{\frac{d}{2}+1}(z), \end{aligned} \tag{5.85}$$

where $\lambda_T=\sqrt{2\pi\hbar^2/mk_{\mathrm{B}}T}$ is the thermal wavelength, and where 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 $\mathsf{Li}_s(z)=\sum_{m=1}^\infty z^m/m^s$ is the polylogarithm function. For bosons with a spectrum bounded below by $\varepsilon_0=0$, the fugacity takes values on the interval $z\in[0,1]$[4]. Note that $\mathrm{Li}_s(z=1)=\zeta(s)$, which is Riemann's zeta function; $\zeta(s)$ is finite for $s>1$.

Clearly $n(T,z)=\mathsf{g}\,\lambda_T^{-d}\ \mathsf{Li}_{d/2}(z)$ is an increasing function of z for fixed T. In fig. 5.3 we plot the function $\mathsf{Li}_s(z)$ *versus* z for three different values of s. We note that the maximum value $\mathsf{Li}_s(z=1)$ is *finite* if $s>1$. Thus, for $d>2$, there is a *maximum density* $n_{\mathrm{max}}(T)=\mathsf{g}\,\zeta(d/2)\,\lambda_T^{-d}$ which is an increasing function of temperature T. Put another way, if we fix the density n, then there is a *critical temperature* T_{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

---

[4] It is easy to see that the chemical potential for noninteracting bosons can never exceed the minimum value ε_0 of the single particle dispersion.

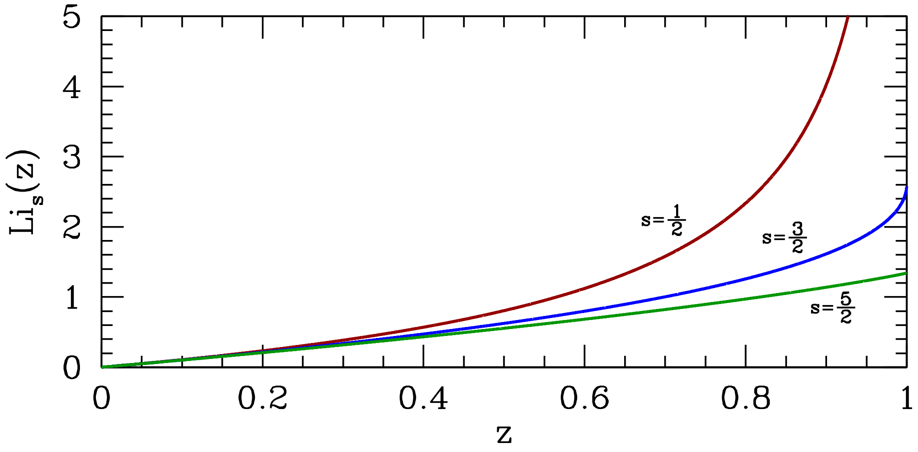


Figure 5.3: The polylogarithm function $\mathrm{Li}_s(z)$ *versus* z for $s = \frac{1}{2}$, $s = \frac{3}{2}$, and $s = \frac{5}{2}$. Note that $\mathrm{Li}_s(1) = \zeta(s)$ diverges for $s \le 1$.

relation

$$n = \mathsf{g}\,\zeta\big(\tfrac{d}{2}\big)\,\lambda_{T_\mathrm{c}}^{-d/2} = \mathsf{g}\,\zeta\big(\tfrac{d}{2}\big)\left(\frac{mk_\mathrm{B}T_\mathrm{c}}{2\pi\hbar^2}\right)^{d/2} \qquad \Longrightarrow \qquad k_\mathrm{B}T_\mathrm{c} = \frac{2\pi\hbar^2}{m}\left(\frac{n}{\mathsf{g}\,\zeta\big(\frac{d}{2}\big)}\right)^{2/d} \quad . \tag{5.86}$$

What happens for $T < T_\mathrm{c}$?

As shown above in §5.5, we must separate out the contribution from the lowest energy single particle mode, which for ballistic dispersion lies at $\varepsilon_0 = 0$. Thus writing

$$n = \frac{1}{V}\,\frac{1}{z^{-1} - 1} + \frac{1}{V}\sum_{\substack{\alpha \\ (\varepsilon_\alpha > 0)}} \frac{1}{z^{-1}\,e^{\varepsilon_\alpha/k_\mathrm{B}T} - 1} \quad , \tag{5.87}$$

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

$$N_0 = V n_0 = \frac{1}{z^{-1} - 1} \qquad \Longrightarrow \qquad z = \frac{1}{1 + N_0^{-1}} \quad . \tag{5.88}$$

The chemical potential is given by

$$\mu = k_\mathrm{B}T \log z = -k_\mathrm{B}T \log\big(1 + N_0^{-1}\big) \approx -\frac{k_\mathrm{B}T}{N_0} \to 0^- \quad . \tag{5.89}$$

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

$$p_0 = -\frac{k_\mathrm{B}T}{V}\log(1 - z) = \frac{k_\mathrm{B}T}{V}\log(1 + N_0) \to 0^+ \quad . \tag{5.90}$$

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 α by the usual integral over $\boldsymbol{k}$. We then have

$$T < T_{\rm c} \quad : \quad n = n_0 + \mathsf{g}\,\zeta\big(\tfrac{d}{2}\big)\,\lambda_T^{-d} \qquad , \qquad p = \mathsf{g}\,\zeta\big(\tfrac{d}{2}{+}1\big)\,k_{\rm B}T\,\lambda_T^{-d} \tag{5.91}$$

and

$$T > T_{\rm c} \quad : \quad n = \mathsf{g}\,\mathsf{Li}_{\frac{d}{2}}(z)\,\lambda_T^{-d} \qquad , \qquad p = \mathsf{g}\,\mathsf{Li}_{\frac{d}{2}+1}(z)\,k_{\rm B}T\,\lambda_T^{-d} \quad . \tag{5.92}$$

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

$$\frac{n_0(T)}{n} = 1 - \frac{\mathsf{g}\,\zeta\big(\frac{d}{2}\big)}{n\,\lambda_T^d} = 1 - \bigg(\frac{T}{T_{\rm c}(n)}\bigg)^{d/2} \quad . \tag{5.93}$$

Energy and heat capacity

Let us compute the internal energy E for the ideal Bose gas. We have

$$\frac{\partial}{\partial\beta}\,(\beta\Omega) = \Omega + \beta\,\frac{\partial\Omega}{\partial\beta} = \Omega - T\,\frac{\partial\Omega}{\partial T} = \Omega + TS \tag{5.94}$$

and therefore

$$\begin{aligned} E = \Omega + TS + \mu N &= \mu N + \frac{\partial}{\partial\beta}\,(\beta\Omega) \\ &= V\Big(\mu\, n - \frac{\partial}{\partial\beta}\,(\beta p)\Big) = \tfrac{1}{2} d\,\mathsf{g} V k_{\rm B}T\,\lambda_T^{-d}\,\mathsf{Li}_{\frac{d}{2}+1}(z) \quad . \end{aligned} \tag{5.95}$$

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.

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

$$N = \begin{cases} N_0 + \zeta(3/2)\,V\lambda_T^{-3} & (T < T_{\rm c}) \\ V\lambda_T^{-3}\,\mathsf{Li}_{3/2}(z) & (T > T_{\rm c}) \end{cases} \quad , \tag{5.96}$$

and the energy is

$$E = \tfrac{3}{2}\,k_{\rm B}T\,\frac{V}{\lambda_T^3}\,\mathsf{Li}_{5/2}(z) = \tfrac{3}{2}\,pV \quad . \tag{5.97}$$

For $T < T_{\rm c}$, we have $z = 1$ and

$$C_{V,N} = \bigg(\frac{\partial E}{\partial T}\bigg)_{V,N} = \tfrac{15}{4}\,\zeta(5/2)\,\frac{V}{\lambda_T^3}\,k_{\rm B} \quad . \tag{5.98}$$

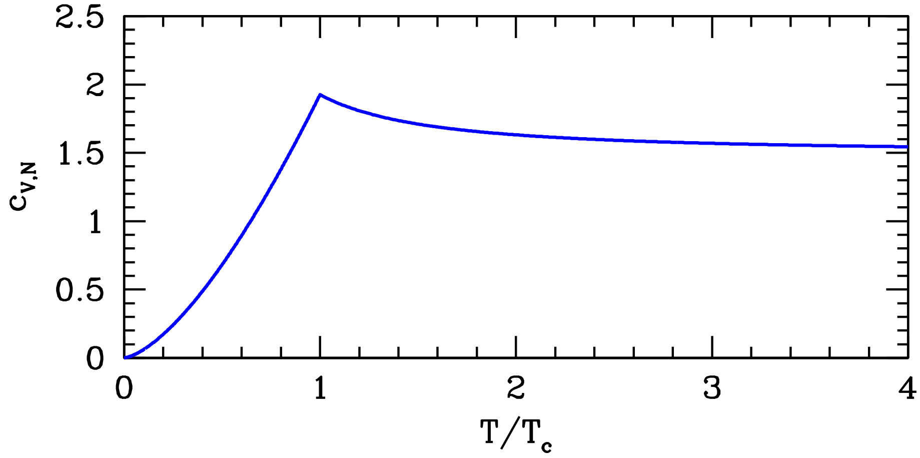


Figure 5.4: Molar heat capacity of the ideal Bose gas (units of R). Note the cusp at $T = T_c$.

The molar heat capacity is therefore

$$c_{V,N}(T,n) = N_A \cdot \frac{C_{V,N}}{N} = \tfrac{15}{4}\,\zeta(5/2)(n\,\lambda_T^3)^{-1} R \quad . \tag{5.99}$$

For $T > T_c$, we have

$$dE\big|_V = \tfrac{15}{4}\,k_B T\,\text{Li}_{5/2}(z)\,\frac{V}{\lambda_T^3}\cdot\frac{dT}{T} + \tfrac{3}{2}\,k_B T\,\text{Li}_{3/2}(z)\,\frac{V}{\lambda_T^3}\cdot\frac{dz}{z} \quad , \tag{5.100}$$

where we have invoked eqn. 5.39. Taking the differential of N, we have

$$dN\big|_V = \tfrac{3}{2}\,\text{Li}_{3/2}(z)\,\frac{V}{\lambda_T^3}\cdot\frac{dT}{T} + \text{Li}_{1/2}(z)\,\frac{V}{\lambda_T^3}\cdot\frac{dz}{z} \quad . \tag{5.101}$$

We set $dN = 0$, which fixes dz in terms of dT, resulting in

$$c_{V,N}(T,z) = \tfrac{3}{2}R\cdot\left[\frac{\frac{5}{2}\,\text{Li}_{5/2}(z)}{\text{Li}_{3/2}(z)} - \frac{\frac{3}{2}\,\text{Li}_{3/2}(z)}{\text{Li}_{1/2}(z)}\right] \quad . \tag{5.102}$$

To obtain $c_{V,N}(T,n)$, we must then invert the relation

$$n(T,z) = \lambda_T^{-3}\,\text{Li}_{3/2}(z) \tag{5.103}$$

in order to obtain $z(T,n)$, and then insert this into eqn. 5.102. The results are shown in fig. 5.4. 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 $T/T_c(n) \propto T\,n^{-2/3}$. Second, the high temperature limit is $\frac{3}{2}R$, which is the classical value. Finally, there is a *cusp* at $T = T_c(n)$. For another example, see §**??**.

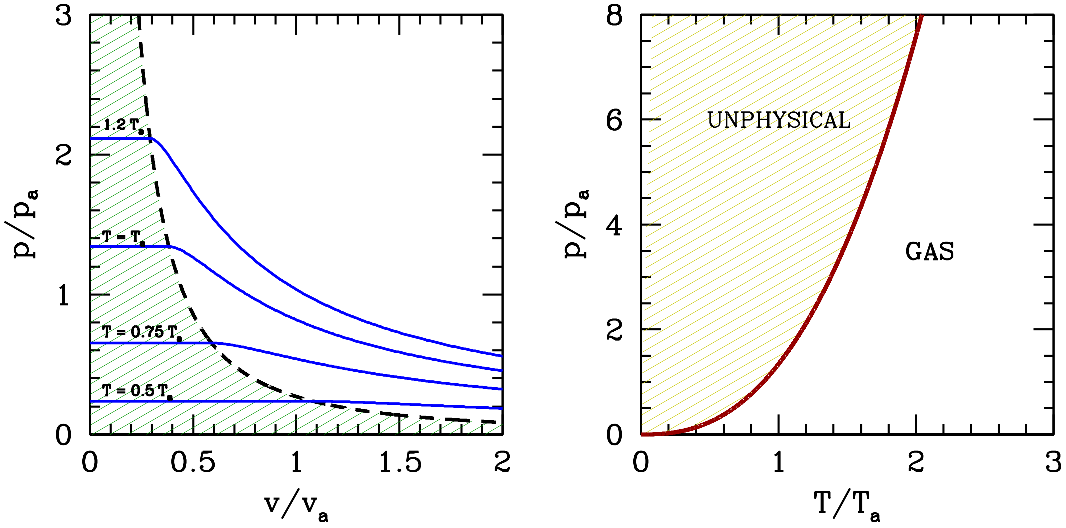


Figure 5.5: Phase diagrams for the ideal Bose gas. Left panel: (p,v) plane. The solid blue curves are isotherms, and the green hatched region denotes $v < v_{\rm c}(T)$, where the system is partially condensed. Right panel: (p,T) plane. The solid red curve is the coexistence curve $p_{\rm c}(T)$, along which Bose condensation occurs. No distinct thermodynamic phase exists in the yellow hatched region above $p = p_{\rm c}(T)$.

Entropy

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:

$$\begin{aligned} \frac{S(T<T_{\rm c},V,N)}{Nk_{\rm B}} &= \big(\tfrac{d}{2}+1\big)\,\zeta\big(\tfrac{d}{2}+1\big)\cdot\frac{\textsf{g}}{n\lambda_T^d} \\ \frac{S(T>T_{\rm c},V,N)}{Nk_{\rm B}} &= \big(\tfrac{d}{2}+1\big)\,\frac{\textsf{Li}_{\frac{d}{2}+1}(z)}{\textsf{Li}_{\frac{d}{2}}(z)} - \log z \quad , \end{aligned} \tag{5.104}$$

with $n = N/V$. In the second equation above, one must invert $n\lambda_T^d = \textsf{g}\,\textsf{Li}_{d/2}(z)$ 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

$$Z_{\rm CIG}(T,V,N) = \frac{1}{N!}\bigg(\frac{\textsf{g}V}{\lambda_T^d}\bigg)^N \qquad \Rightarrow \qquad \frac{S_{\rm CIG}(T,V,N)}{Nk_{\rm B}} = \log(n\lambda_T^d/\textsf{g}) + \tfrac{d}{2} + 1 \quad . \tag{5.105}$$

For the ideal Bose gas in the Maxwell-Boltzmann limit, where $z \to 0$ and $\textsf{Li}_s(z) = z + 2^{-s}z^2 + \ldots$, we have $z = n\lambda_T^d/\textsf{g}$, 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 $T \to 0$.

5.5.3 Isotherms for the ideal Bose gas

Let a be some length scale and define

$$v_a = a^3 \quad , \quad p_a = \frac{2\pi\hbar^2}{ma^5} \quad , \quad T_a = \frac{2\pi\hbar^2}{ma^2 k_B} \tag{5.106}$$

Then we have

$$\frac{v_a}{v} = \left(\frac{T}{T_a}\right)^{3/2} \mathsf{Li}_{3/2}(z) + v_a\, n_0 \quad , \quad \frac{p}{p_a} = \left(\frac{T}{T_a}\right)^{5/2} \mathsf{Li}_{5/2}(z) \quad , \tag{5.107}$$

where $v = V/N$ is the volume per particle[5] and n_0 is the condensate number density; n_0 vanishes for $T \geq T_c$, where $z = 1$. One identifies a critical volume $v_c(T)$ by setting $z = 1$ and $n_0 = 0$, leading to $v_c(T) = v_a\,(T/T_a)^{3/2}$. For $v < v_c(T)$, we set $z = 1$ in eqn. 5.107 to find a relation between v, T, and n_0. For $v > v_c(T)$, we set $n_0 = 0$ in eqn. 5.107 to relate v, T, and z. Note that the pressure is independent of volume for $T < T_c$. The isotherms in the (p, v) plane are then flat for $v < v_c$. This resembles the coexistence region familiar from our study of the thermodynamics of the liquid-gas transition. The situation is depicted in fig. 5.5. In the (T, p) plane, we identify $p_c(T) = p_a(T/T_a)^{5/2}$ 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,

$$\left(\frac{dp}{dT}\right)_{\text{coex}} = \frac{s_2 - s_1}{v_2 - v_1} = \frac{\ell}{T\,\Delta v} \quad , \tag{5.108}$$

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. 5.5. There is no meaning to the shaded region where $p > p_c(T)$. Nevertheless, it is tempting to associate the curve $p = p_c(T)$ 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

$$s = -\frac{1}{N}\left(\frac{\partial\Omega}{\partial T}\right)_V = \tfrac{5}{2}\,\zeta(5/2)\,k_B\,v\,\lambda_T^{-3} = \frac{\tfrac{5}{2}\,\zeta(5/2)}{\zeta(3/2)}\,k_B\left(1 - \frac{n_0}{n}\right) \quad . \tag{5.109}$$

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 $v' = v_c(T)$. 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 $p > p_c(T)$. Ideal Bose gas condensation is a second order transition, and not a first order transition.

---

[5] Note that in the thermodynamics chapter we used v to denote the molar volume, $N_A V/N$.

5.5.4 Properties of liquid ^{4}He

The λ-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 $T = T_\lambda = 2.17\,\mathrm{K}$ (at $p = 1$ atm). The λ-transition, so named for the λ-shaped anomaly in the specific heat in the vicinity of the transition, as shown in fig. 5.7, is continuous (*i.e.* second order). A sketch of the low temperature phase diagram is shown in fig. 5.6. At atmospheric pressure, helium liquefies at $T_\mathrm{l} = 4.2\,\mathrm{K}$.

If we pretend that ^{4}He is a noninteracting Bose gas, then from a density of $n = 2.2 \times 10^{22}\,\mathrm{cm}^{-3}$, we obtain a Bose-Einstein condensation temperature $T_\mathrm{c}(n) = \frac{2\pi\hbar^2}{m}\big(n/\zeta(3/2)\big)^{2/3} = 3.16\,\mathrm{K}$, which is in the right ballpark. The specific heat $C_p(T)$ is found to be singular at $T = T_\lambda$, with

$$C_p(T) = A\,\big|T - T_\lambda(p)\big|^{-\alpha} \quad . \tag{5.110}$$

α is an example of a *critical exponent*. 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. 5.4 for 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 §**??**). 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 α is almost zero but in fact slightly negative, with $\alpha = -0.0127 \pm 0.0003$ in the best experiments (Lipa *et al.*, 2003). The λ transition is most definitely *not* an ideal Bose gas condensation. Theoretically, in the parlance of critical phenomena, IBG condensation and the λ-transition in ^{4}He lie in different *universality classes*[6]. Unlike the IBG, the condensed phase in ^{4}He is a distinct thermodynamic phase, known as a *superfluid*. Note that $C_p(T < T_\mathrm{c})$ for the IBG is not even defined, since for $T < T_\mathrm{c}$ we have $p = p(T)$ and therefore $dp = 0$ requires $dT = 0$.

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 ρ_s. In ^{4}He, for example, at $T = 0$ the condensate fraction is only about 8%, while the superfluid fraction $\rho_\mathrm{s}/\rho = 1$. The distinction between N_0 and ρ_s is very interesting but lies beyond the scope of this course.

One aspect of the superfluid state is its complete absence of viscosity. 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

---

[6] IBG condensation is in the universality class of the spherical model. The λ-transition is in the universality class of the XY model.

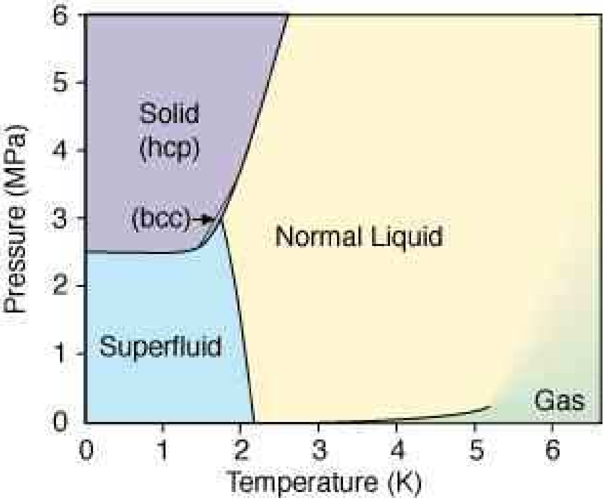


Figure 5.6: Phase diagram of ^{4}He. All phase boundaries are first order transition lines, with the exception of the normal liquid-superfluid transition, which is second order. (*Source: University of Helsinki*)

component is that it has zero energy density. Therefore even though there is a transfer of particles across the plug, there is no energy exchange, and therefore a temperature gradient across the plug can be maintained[7]. In the U-tube experiment depicted in fig. 5.8, this results in a height difference between the two arms of the U-tube.

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 $c = 238\,\text{m/s}$.[8] The 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 Δh 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

$$\begin{aligned} p = -\Omega/V &= -k_{\mathrm{B}}T \int \frac{d^3\!k}{(2\pi)^3} \log\left(1 - e^{-\hbar ck/k_{\mathrm{B}}T}\right) \\ &= -\frac{(k_{\mathrm{B}}T)^4}{(\hbar c)^3} \frac{4\pi}{8\pi^3} \int\limits_0^\infty du\, u^2 \log(1 - e^{-u}) = \frac{\pi^2}{90} \frac{(k_{\mathrm{B}}T)^4}{(\hbar c)^3} \quad . \end{aligned} \tag{5.111}$$

Let's assume $T = 1\,\text{K}$. We'll need the density of liquid helium, $\rho = 148\,\text{kg/m}^3$. We thus obtain

$$\begin{aligned} \frac{dh}{dT} &= \frac{2\pi^2}{45} \left(\frac{k_{\mathrm{B}}T}{\hbar c}\right)^3 \frac{k_{\mathrm{B}}}{\rho g} \\ &= \frac{2\pi^2}{45} \left(\frac{(1.38 \times 10^{-23}\,\text{J/K})(1\,\text{K})}{(1.055 \times 10^{-34}\,\text{J}\cdot\text{s})(238\,\text{m/s})}\right)^3 \times \frac{(1.38 \times 10^{-23}\,\text{J/K})}{(148\,\text{kg/m}^3)(9.8\,\text{m/s}^2)} \simeq 32\,\text{cm/K} \quad , \end{aligned} \tag{5.112}$$

---

[7] Recall that two bodies in thermal equilibrium will have identical temperatures *if they are free to exchange energy*.

[8] The phonon velocity c is slightly temperature dependent.

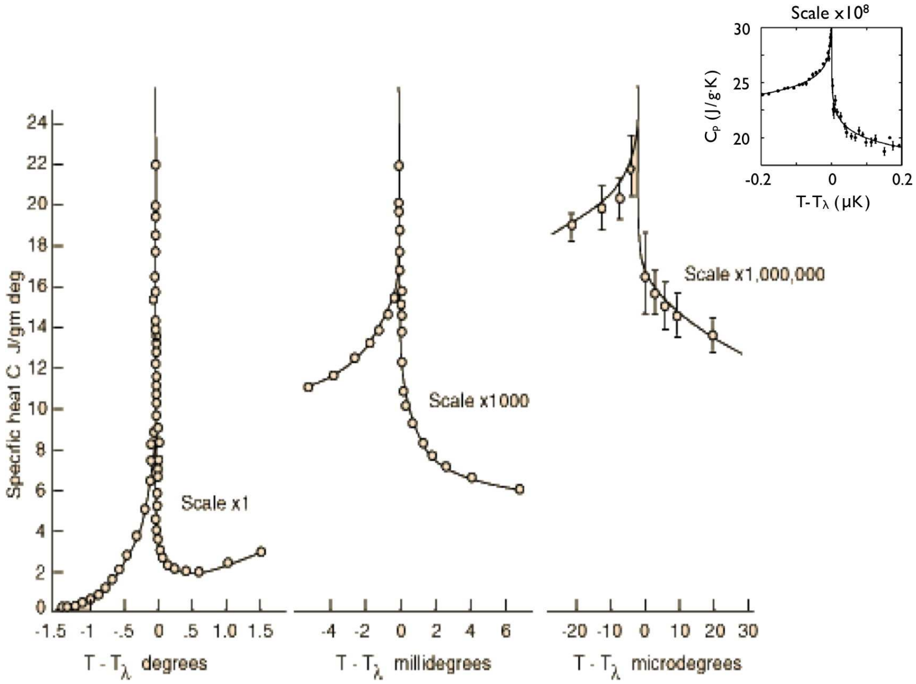


Figure 5.7: Specific heat of liquid ^{4}He in the vicinity of the λ-transition. Data from M. J. Buckingham and W. M. Fairbank, in *Progress in Low Temperature Physics*, C. J. Gortner, ed. (North-Holland, 1961). Inset at upper right: more recent data of J. A. Lipa *et al.*, *Phys. Rev. B* **68**, 174518 (2003) performed in zero gravity earth orbit, to within $\Delta T = 2\,$nK of the transition.

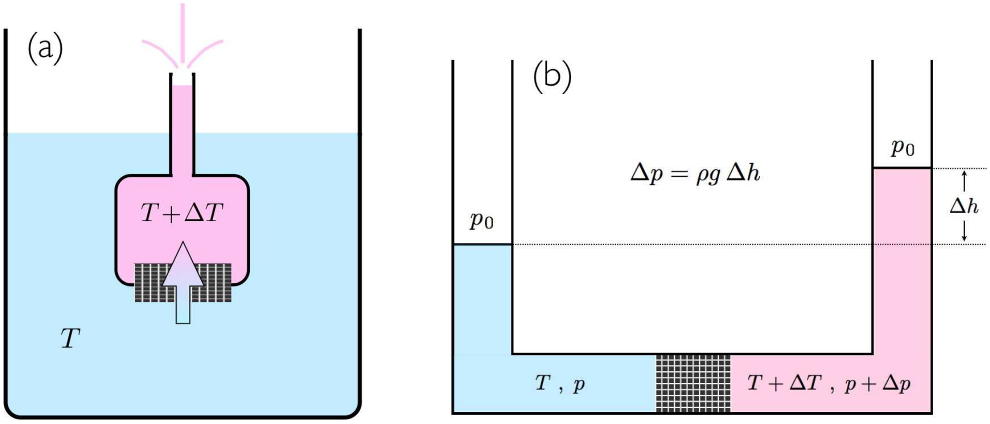


Figure 5.8: The fountain effect. In each case, a temperature gradient is maintained across a porous plug through which only superfluid can flow. This results in a pressure gradient which can result in a fountain or an elevated column in a U-tube.

a very noticeable effect!

Andronikashvili experiment, superfluid density, and condensate fraction

At temperatures below T_λ, 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. 5.9. 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 θ 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, $I = I_{\text{app}} + T_{\text{He}}$, where I_{app} is the moment of inertia of the apparatus and I_{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 ν 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,

$$I_{\text{He}}(T) \approx \tfrac{1}{2}\pi R^4 h_{\text{eff}}\,\rho_n(T) \quad , \tag{5.113}$$

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_{eff} are the radius and height of the helium liquid between the plates. The experimental results are sketched in panel (e) of fig. 5.9. As the temperature T decreases below T_λ, the moment of inertia $I_{\text{He}}(T)$ 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* ρ_{s}/ρ, 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, Ψ_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.*

$$\int \prod_{i=1}^{N} d^3x_i \, \big|\Psi_0(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)\big|^2 = 1 \quad . \tag{5.114}$$

The condensate fraction f_0 can be expressed in terms of Ψ_0 as the following integral:

$$f_0 = \frac{N_0}{N} = \lim_{|\boldsymbol{x}-\boldsymbol{x}'|\to\infty} \frac{n_1(\boldsymbol{x}, \boldsymbol{x}')}{n} \quad , \tag{5.115}$$

where $n_1(\boldsymbol{x}, \boldsymbol{x}')$ is the off-diagonal one-body density matrix in the ground state[9] and n is the total ^{4}He

---

[9] See S. Moroni and M. Boninsegni, *J. Low Temp. Phys.* **136**, 129 (2004).

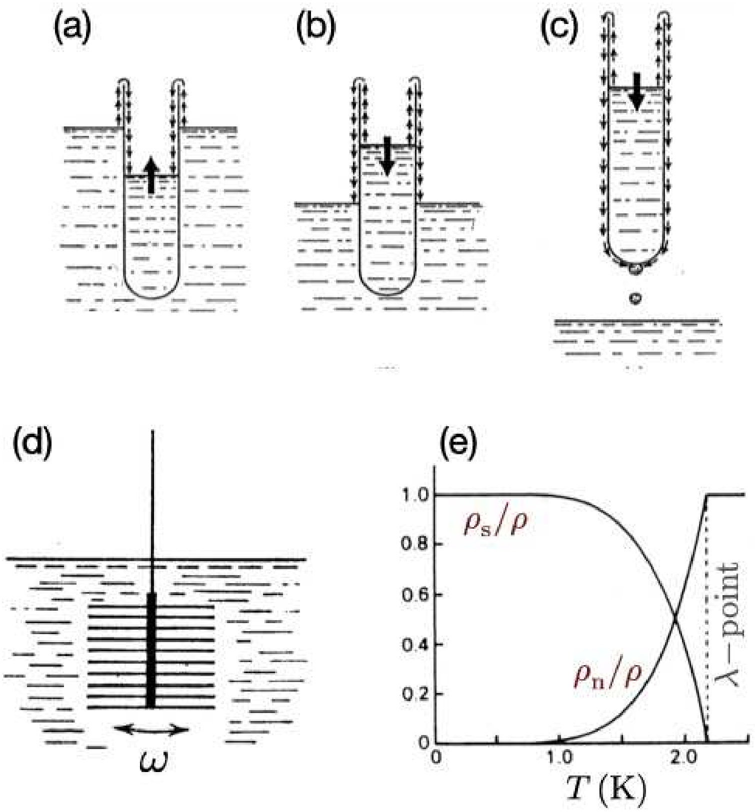


Figure 5.9: Strange properties of superfluid ^{4}He: (a)-(c) Driven by van der Waals forces, superfluid ^{4}He can climb walls and a container of superfluid can empty itself. (d) Andronikashvili's experiment (1946), in which a set of metallic disks immersed in a fluid is used as a torsional oscillator. Below T_λ, the period of the oscillations depends on temperature. (e) Normal and superfluid density fractions as a function of temperature.

number density

$$n_1(\boldsymbol{x},\boldsymbol{x}') = \int \prod_{i=2}^{N} d^3x_i \; \Psi_0^*(\boldsymbol{x},\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N)\,\Psi_0(\boldsymbol{x}',\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N) \quad . \tag{5.116}$$

Theoretically, an estimate of f_0 requires an accurately approximated ground state wavefunction Ψ_0. The simplest trial wavefunctions for ^{4}He are of the Jastrow pair product form,

$$\Psi_0^{\mathrm{var}}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N) = \prod_{j<k}^{N} f\big(|\boldsymbol{x}_j - \boldsymbol{x}_k|\big) \quad , \tag{5.117}$$

where the pair function $f(r)$ is roughly constant for $r > d$, and rapidly decays to zero for $r < d$, with $d \approx 4.2\,\text{Å}$ the diameter of a helium atom. 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

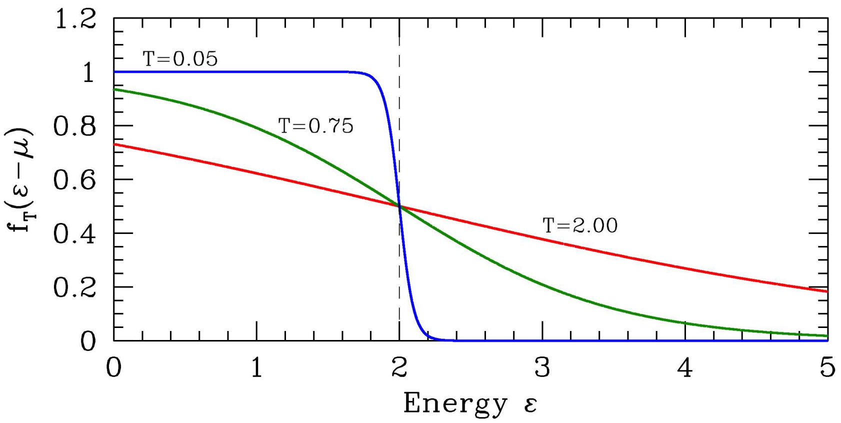


Figure 5.10: The Fermi distribution, $f_T(\varepsilon-\mu)=1/\big[e^{(\varepsilon-\mu)/k_BT}+1\big]$. Here we have set $k_B=1$ and taken $\mu=2$, with $T=\frac{1}{20}$ (blue), $T=\frac{3}{4}$ (green), and $T=2$ (red). In the $T\to 0$ limit, $f_{T\to 0}(\varepsilon-\mu)$ approaches a step function $\Theta(\mu-\varepsilon)$.

gases, discussed in the next section, the interactions are weak, $f_0\approx 1$, and tthe 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.

5.5.5 The Fermi distribution

We define the function

$$f_T(\omega)\equiv\frac{1}{e^{\omega/k_BT}+1}\quad, \tag{5.118}$$

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 ω. As $T\to 0$, $f_T(\omega)$ approaches a step function $\Theta(-\omega)$. The average number of particles in a state of energy ε in a system at temperature T and chemical potential μ is $f_T(\varepsilon-\mu)=n^{\mathsf{F}}_\varepsilon(T,\mu)$. In fig. 5.10 we plot $f_T(\varepsilon-\mu)$ *versus* ε 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.

5.5.6 $T=0$ 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*: $\varepsilon_F=\mu(T=0)$. 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 $\varepsilon(\boldsymbol{k})=\varepsilon_F$, 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 ε_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_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_F and Fermi energy ε_F for the IFG with a ballistic dispersion $\varepsilon(\boldsymbol{k}) = \hbar^2\boldsymbol{k}^2/2m$. The number density is

$$n = \mathsf{g}\int\!\frac{d^d\!k}{(2\pi)^d}\,\Theta(k_\mathrm{F} - k) = \frac{\mathsf{g}\,\Omega_d}{(2\pi)^d}\cdot\frac{k_\mathrm{F}^d}{d} \tag{5.119}$$

and thus

$$n^{(d=1)} = \mathsf{g}\,k_\mathrm{F}/\pi \quad , \quad n^{(d=2)} = \mathsf{g}\,k_\mathrm{F}^2/4\pi \quad , \quad n^{(d=3)} = \mathsf{g}\,k_\mathrm{F}^3/6\pi^2 \quad , \tag{5.120}$$

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 $n(k_\mathrm{F})$ is independent of the dispersion relation, so long as it remains isotropic. Inverting the above expressions, we obtain $k_\mathrm{F}(n) = 2\pi(dn/\mathsf{g}\,\Omega_d)^{1/d}$:

$$k_\mathrm{F}^{(d=1)} = \pi n/\mathsf{g} \quad , \quad k_\mathrm{F}^{(d=2)} = (4\pi n/\mathsf{g})^{1/2} \quad , \quad k_\mathrm{F}^{(d=3)} = (6\pi^2 n/\mathsf{g})^{1/3} \quad . \tag{5.121}$$

The Fermi energy in each case, for ballistic dispersion, is given by

$$\varepsilon_\mathrm{F} = \frac{\hbar^2 k_\mathrm{F}^2}{2m} = \frac{2\pi^2\hbar^2}{m}\left(\frac{d\,n}{\mathsf{g}\,\Omega_d}\right)^{2/d} \quad , \tag{5.122}$$

and so

$$\varepsilon_\mathrm{F}^{(d=1)} = \frac{\pi^2\hbar^2 n^2}{2\mathsf{g}^2 m} \quad , \quad \varepsilon_\mathrm{F}^{(d=2)} = \frac{2\pi\hbar^2\,n}{\mathsf{g}\,m} \quad , \quad \varepsilon_\mathrm{F}^{(d=3)} = \frac{\hbar^2}{2m}\left(\frac{6\pi^2 n}{\mathsf{g}}\right)^{2/3} \quad . \tag{5.123}$$

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

$$g(\varepsilon_\mathrm{F}) = \frac{\mathsf{g}\,\Omega_d}{(2\pi)^d}\cdot\frac{m k_\mathrm{F}^{d-2}}{\hbar^2} = \frac{d}{2}\cdot\frac{n}{\varepsilon_\mathrm{F}} \quad . \tag{5.124}$$

That $g(\varepsilon_\mathrm{F})$ must be a numerical factor multiplied by n/ε_F is obvious on dimensional grounds.

For the electron gas, we have $\mathsf{g} = 2$. In a metal, one typically has $k_\mathrm{F} \sim 0.5\,\text{Å}^{-1}$ to $2\,\text{Å}^{-1}$, and $\varepsilon_\mathrm{F} \sim 1\,\mathrm{eV} - 10\,\mathrm{eV}$. 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

$$\varepsilon(k_x, k_y) = -2t\,\cos(k_x a) - 2t\,\cos(k_y a) \quad , \tag{5.125}$$

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 $\varepsilon(k_x, k_y) = \varepsilon_\mathrm{F}$. When ε_F achieves its minimum value of $\varepsilon_\mathrm{F}^\mathrm{min} = -4t$, 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

$$\varepsilon(k_x, k_y) = -4t + ta^2\left(k_x^2 + k_y^2\right) - \tfrac{1}{12}\,ta^4\left(k_x^4 + k_y^4\right) + \ldots \quad . \tag{5.126}$$

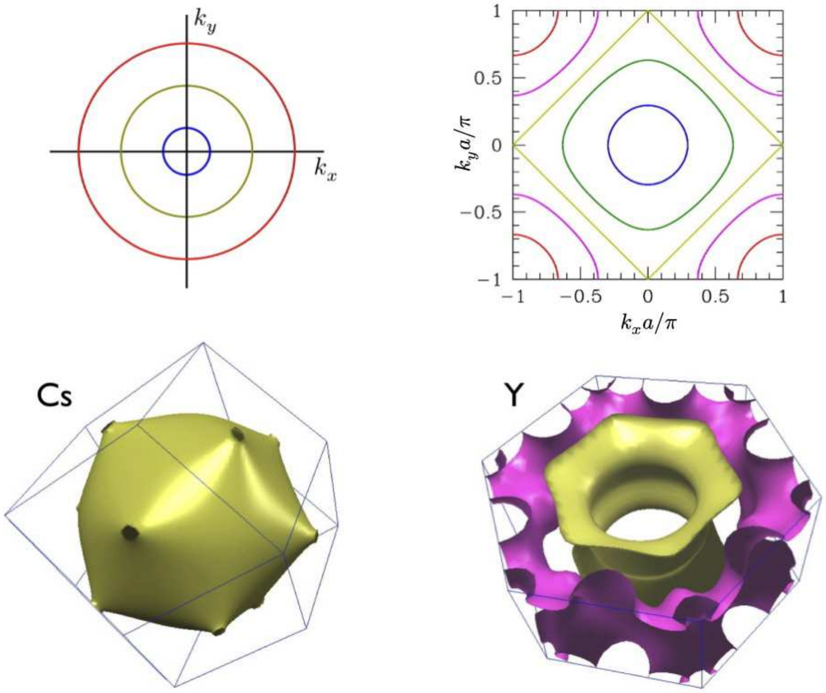


Figure 5.11: Fermi surfaces for two and three-dimensional structures. The wavevector $\boldsymbol{k}$ ranges over the first Brillouin zone in each case. Upper left: free particles in two dimensions. Upper right: ‘tight binding’ electrons on a square lattice. Lower left: Fermi surface for cesium, which 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/*)

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 $k_{\mathrm{F}}^2 = (\varepsilon_{\mathrm{F}} + 4t)/ta^2$. 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 $\varepsilon_{\mathrm{F}} = 0$, the Fermi surface is itself a square. As ε_{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. 5.11.

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. 5.11. 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[10]

$$\hat{H} = \frac{\boldsymbol{p}^2}{2m} + \mu_{\rm B} H\sigma \quad , \tag{5.127}$$

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 $\mu_{\rm B}/k_{\rm B} = 6.717 \times 10^{-5}\,{\rm K/G}$. where 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

$$\frac{\hbar^2 k^2_{{\rm F}\uparrow}}{2m} + \mu_{\rm B} H = \frac{\hbar^2 k^2_{{\rm F}\downarrow}}{2m} - \mu_{\rm B} H \quad , \tag{5.128}$$

which says

$$k^2_{{\rm F}\downarrow} - k^2_{{\rm F}\uparrow} = \frac{2eH}{\hbar c} \quad . \tag{5.129}$$

The total density is

$$n = \frac{k^3_{{\rm F}\uparrow}}{6\pi^2} + \frac{k^3_{{\rm F}\downarrow}}{6\pi^2} \qquad \Longrightarrow \qquad k^3_{{\rm F}\uparrow} + k^3_{{\rm F}\downarrow} = 6\pi^2 n \quad . \tag{5.130}$$

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

$$H_{\rm c} = \frac{\hbar c}{2e} \cdot \big(6\pi^2 n\big)^{1/3} \quad . \tag{5.131}$$

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. 5.11.

5.5.7 The Sommerfeld expansion

In dealing with the ideal Fermi gas, we will repeatedly encounter integrals of the form

$$\mathcal{I}(T,\mu) \equiv \int\limits_{-\infty}^{\infty} d\varepsilon\, f(\varepsilon-\mu)\,\phi(\varepsilon) \quad . \tag{5.132}$$

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).

---

[10] 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.

We start by defining

$$\varPhi(\varepsilon) \equiv \int\limits_{-\infty}^{\varepsilon} d\varepsilon' \, \phi(\varepsilon') \tag{5.133}$$

so that $\phi(\varepsilon) = \varPhi'(\varepsilon)$. We then have

$$\mathcal{I}(T,\mu) = \int\limits_{-\infty}^{\infty} d\varepsilon \, f(\varepsilon - \mu) \, \frac{d\varPhi}{d\varepsilon} = -\int\limits_{-\infty}^{\infty} d\varepsilon \, f'(\varepsilon) \, \varPhi(\mu + \varepsilon) \quad , \tag{5.134}$$

where we assume $\varPhi(-\infty) = 0$. Next, we invoke Taylor's theorem, to write

$$\varPhi(\mu + \varepsilon) = \sum_{n=0}^{\infty} \frac{\varepsilon^n}{n\,!} \, \frac{d^n\varPhi}{d\mu^n} = \exp\left(\varepsilon \, \frac{d}{d\mu}\right) \varPhi(\mu) \quad . \tag{5.135}$$

This last expression involving the exponential of a differential operator may appear overly formal but it proves extremely useful. Since

$$f'(\varepsilon) = -\frac{1}{k_{\rm B}T} \, \frac{e^{\varepsilon/k_{\rm B}T}}{\big(e^{\varepsilon/k_{\rm B}T} + 1\big)^2} \quad , \tag{5.136}$$

we define

$$\widetilde{\mathcal{I}}(T,\mu) = \int\limits_{-\infty}^{\infty} dv \, \frac{e^{vD}}{(e^v + 1)(e^{-v} + 1)} \, \varPhi(\mu) \quad , \tag{5.137}$$

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:[11]

$$\begin{aligned} \int\limits_{-\infty}^{\infty} dv \, \frac{e^{vD}}{(e^v + 1)(e^{-v} + 1)} &= 2\pi i \sum_{n=1}^{\infty} \text{Res}\left[\frac{e^{vD}}{(e^v + 1)(e^{-v} + 1)}\right]_{v=(2n+1)i\pi} \\ &= -2\pi i \sum_{n=0}^{\infty} D \, e^{(2n+1)i\pi D} = -\frac{2\pi i D \, e^{i\pi D}}{1 - e^{2\pi i D}} = \pi D \, \csc \pi D \quad . \end{aligned} \tag{5.138}$$

Thus,

$$\widetilde{\mathcal{I}}(T,\mu) = \pi D \csc(\pi D) \, \varPhi(\mu) \quad , \tag{5.139}$$

which is to be understood as the differential operator $\pi D \csc(\pi D)$ acting on the function $\varPhi(\mu)$. Appealing once more to Taylor's theorem, we have

$$\pi D \csc(\pi D) = 1 + \frac{\pi^2}{6} \, (k_{\rm B}T)^2 \, \frac{d^2}{d\mu^2} + \frac{7\pi^4}{360} \, (k_{\rm B}T)^4 \, \frac{d^4}{d\mu^4} + \ldots \quad . \tag{5.140}$$

---

[11] 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: $\text{Res} = -D\, e^{(2n+1)i\pi D}$.

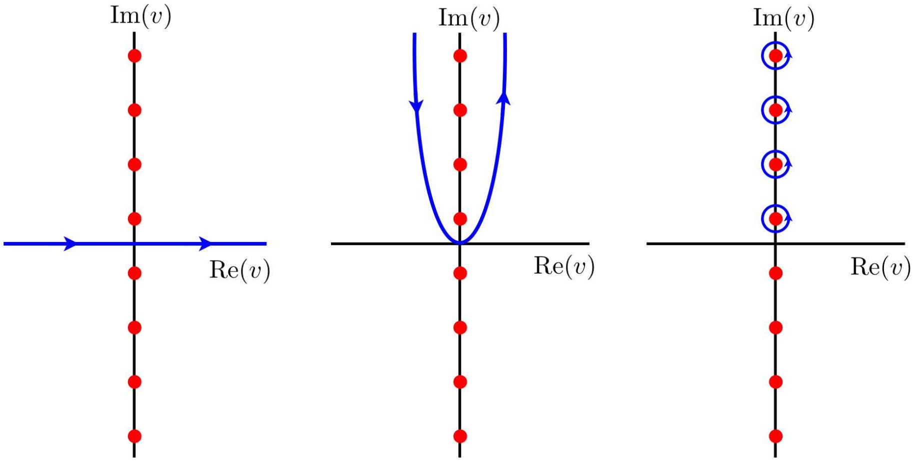


Figure 5.12: Deformation of the complex integration contour in eqn. 5.138.

Thus,

$$\widetilde{\mathcal{I}}(T,\mu)=\int\limits_{-\infty}^{\infty} d\varepsilon\, f(\varepsilon-\mu)\,\phi(\varepsilon)=\int\limits_{-\infty}^{\mu} d\varepsilon\,\phi(\varepsilon)+\frac{\pi^2}{6}\,(k_{\rm B}T)^2\,\phi'(\mu)+\frac{7\pi^4}{360}\,(k_{\rm B}T)^4\,\phi'''(\mu)+\ldots \quad . \tag{5.141}$$

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

$$n=\int\limits_{-\infty}^{\infty} d\varepsilon\, g(\varepsilon)\, f(\varepsilon-\mu)=\int\limits_{-\infty}^{\mu} d\varepsilon\, g(\varepsilon)+\frac{\pi^2}{6}\,(k_{\rm B}T)^2\,g'(\mu)+\ldots \quad . \tag{5.142}$$

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

$$\begin{aligned} n&=\int\limits_{-\infty}^{\varepsilon_{\rm F}+\delta\mu} d\varepsilon\, g(\varepsilon)+\frac{\pi^2}{6}\,(k_{\rm B}T)^2\,g'(\varepsilon_{\rm F}+\delta\mu)+\ldots \\ &=\int\limits_{-\infty}^{\varepsilon_{\rm F}} d\varepsilon\, g(\varepsilon)+g(\varepsilon_{\rm F})\,\delta\mu+\frac{\pi^2}{6}\,(k_{\rm B}T)^2\,g'(\varepsilon_{\rm F})+\ldots \quad , \end{aligned} \tag{5.143}$$

from which we derive

$$\delta\mu = -\frac{\pi^2}{6}\,(k_{\rm B}T)^2\,\frac{g'(\varepsilon_{\rm F})}{g(\varepsilon_{\rm F})} + \mathcal{O}(T^4) \quad . \tag{5.144}$$

Note that $g'/g = (\log g)'$. For a ballistic dispersion, assuming g $= 2$,

$$g(\varepsilon) = 2\int\!\frac{d^3\!k}{(2\pi)^3}\,\delta\bigg(\varepsilon - \frac{\hbar^2k^2}{2m}\bigg) = \frac{m\,k(\varepsilon)}{\pi^2\hbar^2}\bigg|_{k(\varepsilon)=\frac{1}{\hbar}\sqrt{2m\varepsilon}} \tag{5.145}$$

Thus, $g(\varepsilon) \propto \varepsilon^{1/2}$ and $(\log g)' = \frac{1}{2}\,\varepsilon^{-1}$, so

$$\mu(n,T) = \varepsilon_{\rm F} - \frac{\pi^2}{12}\,\frac{(k_{\rm B}T)^2}{\varepsilon_{\rm F}} + \ldots \quad , \tag{5.146}$$

where $\varepsilon_{\rm F}(n) = \frac{\hbar^2}{2m}(3\pi^2 n)^{2/3}$.

Specific heat

The energy of the electron gas is

$$\begin{aligned}\frac{E}{V} = \int\limits_{-\infty}^{\infty}\!\! d\varepsilon\; g(\varepsilon)\,\varepsilon\,f(\varepsilon-\mu) &= \int\limits_{-\infty}^{\mu}\!\! d\varepsilon\; g(\varepsilon)\,\varepsilon + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,\frac{d}{d\mu}\Big(\mu\,g(\mu)\Big) + \ldots \\ &= \int\limits_{-\infty}^{\varepsilon_{\rm F}}\!\! d\varepsilon\; g(\varepsilon)\,\varepsilon + g(\varepsilon_{\rm F})\,\varepsilon_{\rm F}\,\delta\mu + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,\varepsilon_{\rm F}\,g'(\varepsilon_{\rm F}) + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,g(\varepsilon_{\rm F}) + \ldots \\ &= \varepsilon_0 + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\,g(\varepsilon_{\rm F}) + \ldots \quad ,\end{aligned} \tag{5.147}$$

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,

$$C_{V,N} = \bigg(\frac{\partial E}{\partial T}\bigg)_{V,N} = \frac{\pi^2}{3}\,V k_{\rm B}^2\,T\,g(\varepsilon_{\rm F}) \equiv V\gamma\,T \quad , \tag{5.148}$$

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

$$c_V = \frac{N_{\rm A}}{N}\cdot C_V = \frac{\pi^2}{3} R\cdot\frac{k_{\rm B}T\,g(\varepsilon_{\rm F})}{n} = \frac{\pi^2}{2}\bigg(\frac{k_{\rm B}T}{\varepsilon_{\rm F}}\bigg)\,R \quad , \tag{5.149}$$

where in the last expression on the RHS we have assumed a ballistic dispersion, for which

$$\frac{g(\varepsilon_{\rm F})}{n} = \frac{\mathsf{g}\,mk_{\rm F}}{2\pi^2\hbar^2}\cdot\frac{6\pi^2}{\mathsf{g}\,k_{\rm F}^3} = \frac{3}{2\,\varepsilon_{\rm F}} \quad . \tag{5.150}$$

The molar heat capacity in eqn. 5.149 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[12]: Starting with eqn. 5.144 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

$$\bigg(\frac{\partial s}{\partial n}\bigg)_{\!T} = \frac{\pi^2}{3}\,k_{\rm B}^2 T\,\frac{\partial g(\varepsilon_{\rm F})}{\partial n} + \mathcal{O}(T^3) \quad , \tag{5.151}$$

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

$$\mathcal{I}(T,\mu) = \int\limits_0^\infty \! d\varepsilon\; f(\varepsilon - \mu) = \mu + k_{\rm B}T\log\!\big(1 + e^{-\mu/k_{\rm B}T}\big) \quad . \tag{5.152}$$

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

$$\widetilde{\mathcal{I}}(T,\mu) = \tfrac{1}{2}\mu^2 + \frac{\pi^2}{6}(k_{\rm B}T)^2 \quad , \tag{5.153}$$

while the exact result is

$$\mathcal{I}(T,\mu) = \tfrac{1}{2}\mu^2 + \frac{\pi^2}{6}(k_{\rm B}T)^2 + (k_{\rm B}T)^2\sum_{j=1}^\infty \frac{(-1)^j}{j^2}\,e^{-j\mu/k_{\rm B}T} \quad , \tag{5.154}$$

which follows from the polylogarithm identity

$$\mathsf{Li}_2(z) + \mathsf{Li}_2(1/z) = -\tfrac{1}{2}\big[\log(-z)\big]^2 - \frac{\pi^2}{6} \quad . \tag{5.155}$$

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[13]. This is a generic state of affairs for the case where $\phi(\varepsilon)$ is a finite order polynomial in ε.

---

[12] I thank my colleague Tarun Grover for this observation.

[13] Once again I thank my colleague Tarun Grover for pointing this out to me.

5.5.8 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, μ_0 its magnetic moment, and $\boldsymbol{S}$ is the vector of quantum mechanical spin operators satisfying $\big[S^\alpha\,,\,S^\beta\big] = i\hbar\epsilon_{\alpha\beta\gamma}\,S^\gamma$, *i.e.* SU(2) commutation relations. The Hamiltonian for a single particle is then

$$\hat{H} = \frac{1}{2m^*}\Big(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}\Big)^2 - \boldsymbol{H}\cdot\boldsymbol{m} = \frac{1}{2m^*}\Big(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}\Big)^2 + \frac{g}{2}\,\mu_{\rm B} H\,\sigma \quad , \tag{5.156}$$

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 ($m = 511\,{\rm keV}$) 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

$$\varepsilon_\sigma(\boldsymbol{k}) = \frac{\hbar^2\boldsymbol{k}^2}{2m^*} + \tilde{\mu}_{\rm B} H\,\sigma \quad . \tag{5.157}$$

At $T = 0$, we have the results of §5.5.6. At finite T, we once again use the Sommerfeld expansion. We then have

$$\begin{aligned} n &= \int\limits_{-\infty}^{\infty}\! d\varepsilon\; g_\uparrow(\varepsilon)\, f(\varepsilon-\mu) \;+\; \int\limits_{-\infty}^{\infty}\! d\varepsilon\; g_\downarrow(\varepsilon)\, f(\varepsilon-\mu) \\ &= \tfrac{1}{2}\int\limits_{-\infty}^{\infty}\! d\varepsilon\,\Big\{ g(\varepsilon - \tilde{\mu}_{\rm B}H) + g(\varepsilon + \tilde{\mu}_{\rm B}H)\Big\}\, f(\varepsilon-\mu) \\ &= \int\limits_{-\infty}^{\infty}\! d\varepsilon\,\Big\{ g(\varepsilon) + (\tilde{\mu}_{\rm B}H)^2\, g''(\varepsilon) + \ldots\Big\}\, f(\varepsilon-\mu) \quad . \end{aligned} \tag{5.158}$$

We now invoke the Sommerfeld expension to find the temperature dependence:

$$\begin{aligned} n &= \int\limits_{-\infty}^{\mu}\! d\varepsilon\; g(\varepsilon) + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\, g'(\mu) + (\tilde{\mu}_{\rm B}H)^2\, g'(\mu) + \ldots \\ &= \int\limits_{-\infty}^{\varepsilon_{\rm F}}\! d\varepsilon\; g(\varepsilon) + g(\varepsilon_{\rm F})\,\delta\mu + \frac{\pi^2}{6}\,(k_{\rm B}T)^2\, g'(\varepsilon_{\rm F}) + (\tilde{\mu}_{\rm B}H)^2\, g'(\varepsilon_{\rm F}) + \ldots \quad . \end{aligned} \tag{5.159}$$

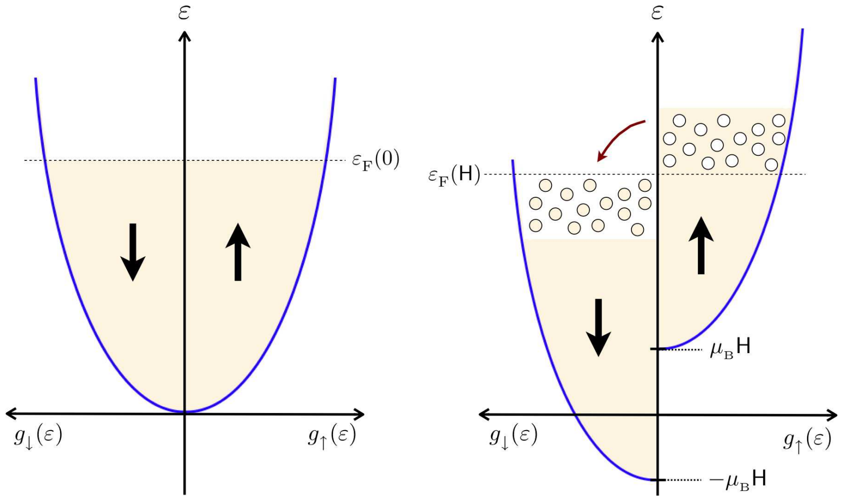


Figure 5.13: Fermi distributions in the presence of an external Zeeman-coupled magnetic field.

Note that the density of states for spin species σ 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

$$\delta\mu(T,n,H) = -\bigg\{\frac{\pi^2}{6}\,(k_{\rm B}T)^2 + (\tilde\mu_{\rm B}H)^2\bigg\}\,\frac{g'(\varepsilon_{\rm F})}{g(\varepsilon_{\rm F})} + \ldots \quad . \tag{5.160}$$

We next compute the difference $n_\uparrow - n_\downarrow$ in the densities of up and down spin electrons:

$$\begin{aligned} n_\uparrow - n_\downarrow &= \int\limits_{-\infty}^{\infty}\! d\varepsilon\,\Big\{g_\uparrow(\varepsilon) - g_\downarrow(\varepsilon)\Big\}\, f(\varepsilon - \mu) \\ &= \tfrac{1}{2}\int\limits_{-\infty}^{\infty}\! d\varepsilon\,\Big\{g(\varepsilon - \tilde\mu_{\rm B}H) - g(\varepsilon + \tilde\mu_{\rm B}H)\Big\}\, f(\varepsilon-\mu) \\ &= -\tilde\mu_{\rm B}H\cdot\pi D\csc(\pi D)\, g(\mu) + \mathcal{O}(H^3) \quad . \end{aligned} \tag{5.161}$$

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

$$M = -\tilde\mu_{\rm B}(n_\uparrow - n_\downarrow) = \tilde\mu^2_{\rm B}\, g(\varepsilon_{\rm F})\, H \quad . \tag{5.162}$$

in which the magnetic susceptibility is

$$\chi = \bigg(\frac{\partial M}{\partial H}\bigg)_{T,N} = \tilde\mu^2_{\rm B}\, g(\varepsilon_{\rm F}) \quad . \tag{5.163}$$

This is called the *Pauli paramagnetic susceptibility*.

Landau diamagnetism

When orbital effects are included, the single particle energy levels are given by

$$\varepsilon(n,k_z,\sigma) = (n+\tfrac{1}{2})\hbar\Omega_{\rm c} + \frac{\hbar^2 k_z^2}{2m^*} + \tilde\mu_{\rm B} H\,\sigma \quad . \tag{5.164}$$

Here n is a Landau level index, and $\Omega_{\rm c} = eH/m^* c$ is the *cyclotron frequency*. Note that

$$\frac{\tilde\mu_{\rm B} H}{\hbar\Omega_{\rm c}} = \frac{ge\hbar H}{4mc}\cdot\frac{m^* c}{\hbar e H} = \frac{g}{4}\cdot\frac{m^*}{m} \quad . \tag{5.165}$$

Accordingly, we define the ratio $r \equiv (g/2)\times(m^*/m)$. We can then write

$$\varepsilon(n,k_z,\sigma) = \big(n+\tfrac{1}{2}+\tfrac{1}{2}r\sigma\big)\hbar\Omega_{\rm c} + \frac{\hbar^2 k_z^2}{2m^*} \quad . \tag{5.166}$$

The grand potential is then given by

$$\varOmega = -\frac{HA}{\phi_0}\cdot L_z\cdot k_{\rm B}T\int\limits_{-\infty}^{\infty}\!\frac{dk_z}{2\pi}\sum_{n=0}^\infty\sum_{\sigma=\pm 1}\log\Big[1+e^{\mu/k_{\rm B}T}\,e^{-(n+\frac{1}{2}+\frac{1}{2}r\sigma)\hbar\Omega_{\rm c}/k_{\rm B}T}\,e^{-\hbar^2k_z^2/2m^* k_{\rm B}T}\Big] \quad . \tag{5.167}$$

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

$$\frac{HA}{\phi_0}\cdot L_z\cdot k_{\rm B}T = \hbar\Omega_{\rm c}\cdot\frac{V}{\lambda_T^3} \quad , \tag{5.168}$$

hence we can write

$$\varOmega(T,V,\mu,H) = \hbar\Omega_{\rm c}\sum_{n=0}^\infty\sum_{\sigma=\pm 1} Q\big((n+\tfrac{1}{2}+\tfrac{1}{2}r\sigma)\hbar\Omega_{\rm c}-\mu\big) \quad , \tag{5.169}$$

where

$$Q(\varepsilon) = -\frac{V}{\lambda_T^2}\int\limits_{-\infty}^{\infty}\!\frac{dk_z}{2\pi}\,\log\Big[1+e^{-\varepsilon/k_{\rm B}T}e^{-\hbar^2k_z^2/2m^*k_{\rm B}T}\Big] = \frac{V}{\lambda_T^3}\,\mathsf{Li}_{3/2}\big(e^{-\varepsilon/k_{\rm B}T}\big) \quad . \tag{5.170}$$

We now invoke the Euler-MacLaurin formula,

$$\sum_{n=0}^\infty F(n) = \int\limits_0^\infty\! dx\; F(x) + \tfrac{1}{2}\,F(0) - \tfrac{1}{12}\,F'(0) + \ldots \quad , \tag{5.171}$$

resulting in

$$\begin{aligned}\varOmega = \sum_\sigma\int\limits_{-\infty}^{\infty}\! d\varepsilon\; Q(\varepsilon-\mu)\,\Theta\big(\varepsilon-\tfrac{1}{2}(1+\sigma r)\hbar\Omega_{\rm c}\big) \qquad\qquad (5.172)\\ +\sum_\sigma\Big[\tfrac{1}{2}\,\hbar\Omega_{\rm c}\,Q\big(\tfrac{1}{2}(1+\sigma r)\hbar\Omega_{\rm c}-\mu\big) - \tfrac{1}{12}\,(\hbar\omega_{\rm c})^2\,Q'\big(\tfrac{1}{2}(1+\sigma r)\hbar\Omega_{\rm c}-\mu\big)+\ldots\Big] \quad .\end{aligned}$$

We next expand in powers of the magnetic field H to obtain

$$\Omega(T,V,\mu,H) = 2\int\limits_0^\infty d\varepsilon\, Q(\varepsilon-\mu) + \left(\tfrac{1}{4}r^2 - \tfrac{1}{12}\right)(\hbar\Omega_{\rm c})^2\, Q'(-\mu) + \ldots \quad . \tag{5.173}$$

Thus, the magnetic susceptibility is

$$\begin{aligned} \chi = -\frac{1}{V}\frac{\partial^2\Omega}{\partial H^2} &= \left(r^2 - \tfrac{1}{3}\right)\cdot \tilde\mu_{\rm B}^2 \cdot (m/m^*)^2 \cdot \left\{ -\frac{2}{V}\,Q'(-\mu)\right\} \\ &= \left(\frac{g^2}{4} - \frac{m^2}{3m^{*2}}\right)\cdot \tilde\mu_{\rm B}^2\cdot n^2\kappa_T \quad , \end{aligned} \tag{5.174}$$

where κ_T is the isothermal compressibility[14]. 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*.

Exercise : Show that $-\frac{2}{V}\,Q'(-\mu) = n^2\kappa_T$.

---

[14] We've used $-\frac{2}{V}Q'(\mu) = -\frac{1}{V}\frac{\partial^2\Omega}{\partial\mu^2} = n^2\kappa_T$.

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^{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^{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 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.

6.2.1 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

$$\begin{aligned} Z(T,V,N) &= \frac{1}{N!}\int\prod_{i=1}^{N}\frac{d^d\!p_i\,d^d\!x_i}{h^d}\,e^{-\hat{H}/k_\text{B}T} \\ &= \frac{\lambda_T^{-Nd}}{N!}\int\prod_{i=1}^{N}d^d\!x_i\,\exp\bigg(-\frac{1}{k_\text{B}T}\sum_{i<j}u\big(|\boldsymbol{x}_i-\boldsymbol{x}_j|\big)\bigg) \quad . \end{aligned} \tag{6.1}$$

Here, we have assumed a many body Hamiltonian of the form

$$\hat{H}=\sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m}+\sum_{i<j}u\big(|\boldsymbol{x}_i-\boldsymbol{x}_j|\big) \quad , \tag{6.2}$$

in which massive nonrelativistic particles interact via a two-body central potential. As before, $\lambda_T=\sqrt{2\pi\hbar^2/mk_\text{B}T}$ is the thermal wavelength. We can now write

$$Z(T,V,N)=\lambda_T^{-Nd}\,Q_N(T,V) \quad , \tag{6.3}$$

where the *configuration integral* $Q_N(T,V)$ is given by

$$Q_N(T,V)=\frac{1}{N!}\int\!d^d\!x_1\cdots\int\!d^d\!x_N\,\prod_{i<j}e^{-\beta u(r_{ij})} \quad . \tag{6.4}$$

There are no general methods for evaluating the configurational integral exactly.

6.2.2 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

$$u(x-x')=\begin{cases}\infty & \text{if } |x-x'|<a \\ 0 & \text{if } |x-x'|\geq a \quad .\end{cases} \tag{6.5}$$

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

$$v(x) = \begin{cases} \infty & \text{if } x < \frac{1}{2}a \\ 0 & \text{if } \frac{1}{2}a \le x \le L - \frac{1}{2}a \\ \infty & \text{if } x > L - \frac{1}{2}a \quad . \end{cases} \tag{6.6}$$

The configuration integral of the 1D Tonks gas is given by

$$Q_N(T,L) = \frac{1}{N!} \int\limits_0^L \! dx_1 \cdots \int\limits_0^L \! dx_N \, \chi(x_1, \ldots, x_N) \quad , \tag{6.7}$$

where $\chi = e^{-U/k_\text{B}T}$ 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 χ 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

$$\begin{aligned} Q_N(T,L) &= \int\limits_{a/2}^{Y_1} \! dx_1 \int\limits_{x_1+a}^{Y_2} \! dx_2 \cdots \int\limits_{x_{N-1}+a}^{Y_N} \! dx_N = \int\limits_{a/2}^{Y_1} \! dx_1 \int\limits_{x_1+a}^{Y_2} \! dx_2 \cdots \int\limits_{x_{N-2}+a}^{Y_{N-1}} \! dx_{N-1} \, \big(Y_{N-1} - x_{N-1}\big) \\ &= \frac{1}{2} \int\limits_{a/2}^{Y_1} \! dx_1 \int\limits_{x_1+a}^{Y_2} \! dx_2 \cdots \int\limits_{x_{N-3}+a}^{Y_{N-2}} \! dx_{N-2} \, \big(Y_{N-2} - x_{N-2}\big)^2 = \cdots = \frac{1}{k!} \int\limits_{a/2}^{Y_1} \! dx_1 \int\limits_{x_1+a}^{Y_2} \! dx_2 \cdots \int\limits_{x_{N-k-1}+a}^{Y_{N-k}} \! dx_{N-k} \, \big(Y_{N-k} - x_{N-k}\big)^k \\ &= \frac{1}{N!} \big(Y_1 - \tfrac{1}{2}a\big)^N = \frac{1}{N!} \, (L - Na)^N \quad . \end{aligned} \tag{6.8}$$

The partition function is $Z(T,L,N) = \lambda_T^{-N} \, Q_N(T,L)$, and so the free energy is

$$F = -k_\text{B}T \log Z = -Nk_\text{B}T \Bigg\{ -\log\lambda_T + 1 + \log\bigg(\frac{L}{N} - a\bigg) \Bigg\} \quad , \tag{6.9}$$

where we have used Stirling's rule to write $\log N! \approx N \log N - N$. The pressure is

$$p = -\frac{\partial F}{\partial L} = \frac{nk_\text{B}T}{1 - na} \quad , \tag{6.10}$$

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, $pL = Nk_\text{B}T$, is replaced by $pL_\text{eff} = Nk_\text{B}T$, where $L_\text{eff} = L - Na$ is the 'free' volume obtained by subtracting the total "excluded volume" Na from the original volume L.

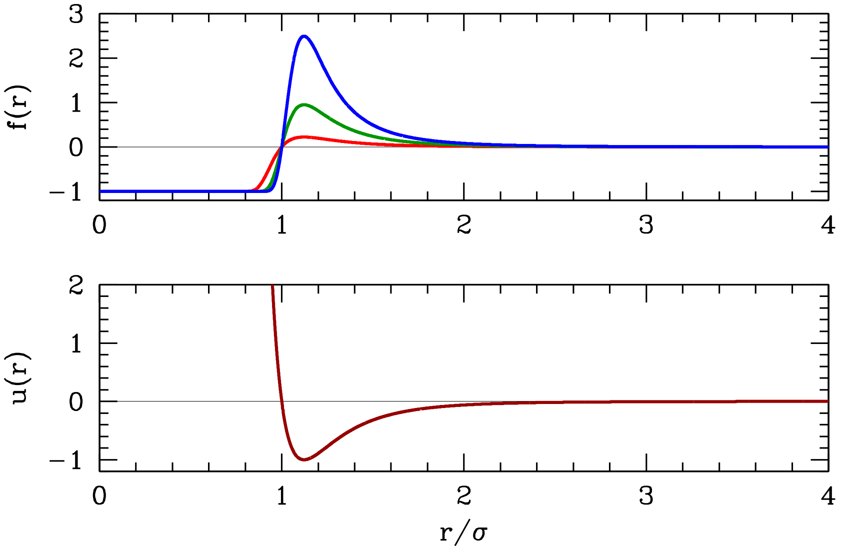


Figure 6.1: Bottom panel: Lennard-Jones potential $u(r) = 4\epsilon\left(x^{-12} - x^{-6}\right)$, with $x = r/\sigma$ and $\epsilon = 1$. Note the weak attractive tail and the strong repulsive core. Top panel: Mayer function $f(r,T) = e^{-u(r)/k_B T} - 1$ for $k_B T = 0.8\,\epsilon$ (blue), $k_B T = 1.5\,\epsilon$ (green), and $k_B T = 5\,\epsilon$ (red).

Note the similarity here to the van der Waals equation of state, $(p + av^{-2})(v - b) = RT$, where $v = N_A V/N$ is the molar volume. Defining $\tilde{a} \equiv a/N_A^2$ and $\tilde{b} \equiv b/N_A$, we have

$$p + \tilde{a}n^2 = \frac{nk_B T}{1 - \tilde{b}n} \quad , \tag{6.11}$$

where $n = N_A/v$ 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.

6.2.3 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

$$f(r) = e^{-\beta u(r)} - 1 \quad , \tag{6.12}$$

called the *Mayer function* after Josef Mayer. We may now write

$$Q_N(T,V) = \frac{1}{N!}\int\! d^d\!x_1 \cdots \int\! d^d\!x_N \prod_{i<j} \big(1 + f_{ij}\big) \quad . \tag{6.13}$$

A typical potential we might consider is the semi-phenomenological Lennard-Jones potential,

$$u(r) = 4\,\epsilon \left\{ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right\} \quad . \tag{6.14}$$

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. 6.1.

The Lennard-Jones potential 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 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. 6.13 as

$$\prod_{i<j} \big(1 + f_{ij}\big) = 1 + \sum_{i<j} f_{ij} + \sum_{\substack{i<j\,,\,k<l \\ (ij)\neq(kl)}} f_{ij}\, f_{kl} + \ldots \quad . \tag{6.15}$$

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. 6.2. 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. 6.2, the contribution to the configurational integral would be

$$\begin{aligned} \Delta Q = \frac{V^{N-11}}{N!} &\int\! d^d\!x_1\, d^d\!x_4\, d^d\!x_7\, d^d\!x_9\; f_{1,4}\, f_{4,7}\, f_{4,9}\, f_{7,9} \\ &\times \int\! d^d\!x_2\, d^d\!x_5\, d^d\!x_6\; f_{2,5}\, f_{2,6} \times \int\! d^d\!x_3\, d^d\!x_{10}\; f_{3,10} \times \int\! d^d\!x_8\, d^d\!x_{11}\; f_{8,11} \quad . \end{aligned} \tag{6.16}$$

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

$$N = \sum_{\gamma} m_\gamma\, n_\gamma \quad , \tag{6.17}$$

where γ ranges over all possible connected subgraphs, and

m_γ = number of connected subgraphs of type γ in the unlabeled graph

n_γ = number of vertices in the connected subgraph γ .

$$\big(f_{1,4}\, f_{4,7}\, f_{4,9}\, f_{7,9}\big)\, \big(f_{2,5}\, f_{2,6}\big)\, \big(f_{3,10}\big)\, \big(f_{8,11}\big)$$

$$\Updownarrow$$

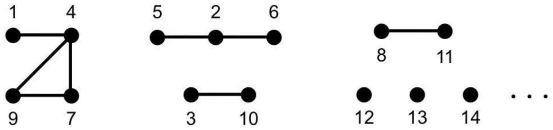


Figure 6.2: Diagrammatic interpretation of a term involving a product of eight Mayer functions.

Note that the single vertex • 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 thing the answer is simply $N!$, however this is too big, because different assignments of the labels to the vertices may not result in a distinct graph. To see this, consider the examples in fig. 6.3. 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 γ there may be permutations which leave the integrand unchanged, as shown in the second and third examples in fig. 6.3. We define the *symmetry factor* s_γ as the number of permutations of the labels which leaves a given connected subgraphs γ invariant. Examples of symmetry factors are shown in fig. 6.4. 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 figur e in the plane by $180°$ 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

$$\begin{aligned} Z &= \frac{\lambda_T^{-Nd}}{N!} \sum_{\{m_\gamma\}} \frac{N!}{\prod m_\gamma!\, s_\gamma^{m_\gamma}} \cdot \prod_\gamma \bigg(\int \! d^d\!x_1 \cdots d^d\!x_{n_\gamma} \prod_{i<j}^{\gamma} f_{ij} \bigg)^{m_\gamma} \delta_{N\,,\sum m_\gamma n_\gamma} \\ &= \sum_{\{m_\gamma\}} \prod_\gamma \frac{1}{m_\gamma!} \bigg(\frac{V b_\gamma(T)}{\lambda_T^d} \bigg)^{m_\gamma} \delta_{N\,,\sum m_\gamma n_\gamma} \end{aligned} \tag{6.18}$$

where the product $\prod_{i<j}^{\gamma} f_{ij}$ is over all links in the subgraph γ. The final Kronecker delta enforces the constraint $N = \sum_\gamma m_\gamma\, n_\gamma$. We have defined the *dimensionless cluster integrals* b_γ as

$$b_\gamma(T) \equiv \frac{1}{s_\gamma} \int \! \frac{d^d\!x_1}{\lambda_T^d} \cdots \int \! \frac{d^d\!x_{n_\gamma - 1}}{\lambda_T^d} \prod_{i<j}^{\gamma} f_{ij} \quad , \tag{6.19}$$

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

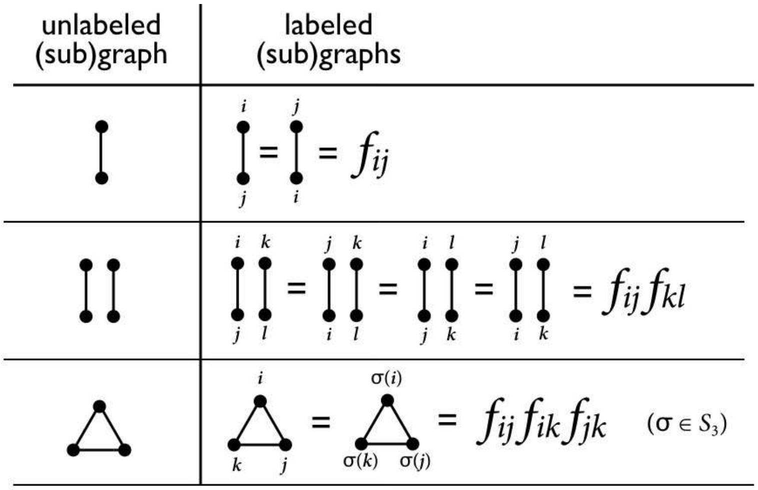


Figure 6.3: Different assignations of labels to vertices may not result in a distinct term in the expansion of the configuration integral.

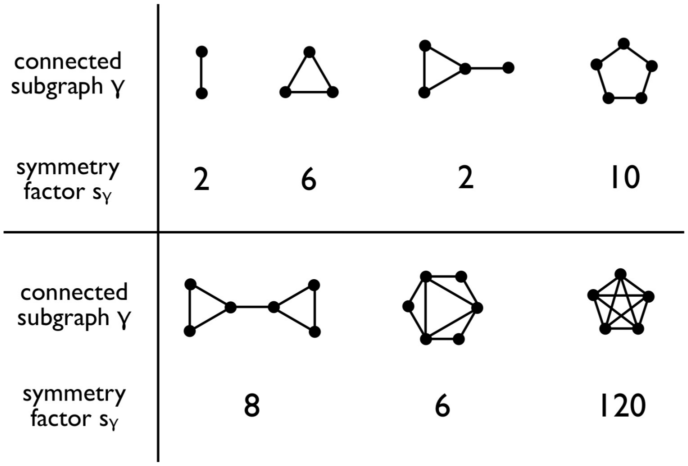


Figure 6.4: The symmetry factor s_γ for a connected subgraph γ is the number of permutations of its indices which leaves the term $\prod_{(ij)\in\gamma} f_{ij}$ invariant.

simultaneous translation of all the coordinate vectors by any constant vector, and hence the integral over the n_γ position variables contains exactly one factor of the volume, which yields factor of V within the round brackets in the second line of eqn. 6.18. Thus, each cluster integral is *intensive*[1], scaling as V^0.

If we compute the grand partition function, then the fixed N constraint is relaxed, and we can do the

---

[1] We assume that the long-ranged behavior of $f(r) \approx -\beta u(r)$ is integrable.

sums:

$$\begin{aligned}\Xi(T,V,\mu) &= \sum_{\{m_\gamma\}} \big(e^{\beta\mu}\big)^{\sum m_\gamma n_\gamma} \prod_\gamma \frac{1}{m_\gamma!} \bigg(\frac{V b_\gamma(T)}{\lambda_T^d}\bigg)^{m_\gamma} \\ &= \prod_\gamma \sum_{m_\gamma=0}^{\infty} \frac{1}{m_\gamma!} \bigg(\frac{V z^{n_\gamma}\, b_\gamma(T)}{\lambda_T^d}\bigg)^{m_\gamma} = \exp\bigg(V\lambda_T^{-d} \sum_\gamma z^{n_\gamma}\, b_\gamma\bigg) \quad ,\end{aligned} \tag{6.20}$$

where $z = \exp(\beta\mu)$ is the fugacity. Thus, since $\Omega = -k_{\rm B}T \log \Xi$,

$$\Omega(T,V,\mu) = -\frac{V k_{\rm B}T}{\lambda_T^d} \sum_\gamma z^{n_\gamma}\, b_\gamma(T) \quad , \tag{6.21}$$

and we can write

$$\begin{aligned} p &= k_{\rm B}T\, \lambda_T^{-d} \sum_\gamma z^{n_\gamma}\, b_\gamma(T) \\ n &= \lambda_T^{-d} \sum_\gamma n_\gamma\, z^{n_\gamma}\, b_\gamma(T) \quad ,\end{aligned} \tag{6.22}$$

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,

$$p = nk_{\rm B}T \Big\{ 1 + B_2(T)\, n + B_3(T)\, n^2 + \ldots \Big\} \quad . \tag{6.23}$$

Lowest order expansion

We may rather easily obtain

$$B_2(T) = -\lambda_T^d\, b_{-}(T) \qquad , \qquad B_3(T) = 2\lambda_T^{2d}\, b_{\triangle}(T) \quad , \tag{6.24}$$

with

$$b_{-}(T) = \tfrac{1}{2}\!\int\! \frac{d^d\!x_1}{\lambda_T^d}\, f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\big) = \tfrac{1}{2}\!\int\! \frac{d^d\!r}{\lambda_T^d}\, f(r) \tag{6.25}$$

and

$$\begin{aligned} b_{\wedge}(T) &= \tfrac{1}{2}\!\int\! \frac{d^d\!x_1}{\lambda_T^d} \!\int\! \frac{d^d\!x_2}{\lambda_T^d}\, f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\big)\, f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_3|\big) \\ &= \tfrac{1}{2}\!\int\! \frac{d^d\!r}{\lambda_T^d} \!\int\! \frac{d^d\!r'}{\lambda_T^d}\, f(r)\, f(r') = 2\big(b_{-}\big)^2 \end{aligned} \tag{6.26}$$

and

$$\begin{aligned} b_{\triangle}(T) &= \tfrac{1}{6}\!\int\! \frac{d^d\!x_1}{\lambda_T^d} \!\int\! \frac{d^d\!x_2}{\lambda_T^d}\, f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\big)\, f\big(|\boldsymbol{x}_1 - \boldsymbol{x}_3|\big)\, f\big(|\boldsymbol{x}_2 - \boldsymbol{x}_3|\big) \\ &= \tfrac{1}{6}\!\int\! \frac{d^d\!r}{\lambda_T^d} \!\int\! \frac{d^d\!r'}{\lambda_T^d}\, f(r)\, f(r')\, f\big(|\boldsymbol{r} - \boldsymbol{r}'|\big) \quad . \end{aligned} \tag{6.27}$$

Cookbook recipe

Just follow these simple steps:

- The pressure and number density are written as sums over unlabeled connected clusters γ, *viz.*

$$
\begin{aligned}
p &= k_{\rm B}T\,\lambda_T^{-d}\sum_\gamma z^{n_\gamma}\,b_\gamma(T)\\
n &= \lambda_T^{-d}\sum_\gamma n_\gamma\,z^{n_\gamma}\,b_\gamma(T)\quad,
\end{aligned}
\tag{6.28}
$$

  where $z=\exp(\beta\mu)$ is the fugacity.

- To compute the dimensionless cluster integral $b_\gamma(T)$, first draw the connected cluster γ with *unlabeled* vertices.

- Next, assign labels $1\,,\,2\,,\,\ldots\,,\,n_\gamma$ to the vertices, where n_γ is the total number of vertices in the cluster γ. 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_γ is the number of elements of the symmetric group S_{n_γ} 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

$$
b_\gamma(T)\equiv\frac{1}{s_\gamma}\int\!\frac{d^d\!x_1}{\lambda_T^d}\cdots\int\!\frac{d^d\!x_{n_\gamma-1}}{\lambda_T^d}\prod_{i<j}^{\gamma} f_{ij}\quad,
\tag{6.29}
$$

  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 mut 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,

$$
p=nk_{\rm B}T\Big\{1+B_2(T)\,n+B_3(T)\,n^2+\ldots\Big\}\quad,
\tag{6.30}
$$

  where

$$
B_k(T)=-(k-1)\,\lambda_T^{(k-1)d}\sum_{\gamma\in\varGamma_k} b_\gamma(T)\quad,
\tag{6.31}
$$

  with $\varGamma_k$ 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.

6.2.4 Examples

Hard sphere gas in three dimensions

The hard sphere potential is given by

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ 0 & \text{if } r > a \end{cases} \quad . \tag{6.32}$$

Here a is the *diameter* of the spheres. The corresponding Mayer function is then temperature independent, and given by

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ 0 & \text{if } r > a \end{cases} \quad . \tag{6.33}$$

We can change variables to obtain

$$\mathsf{b}_2(T) = \tfrac{1}{2}\int \frac{d^3\!r}{\lambda_T^3}\, f(r) = -\tfrac{2}{3}\pi a^3 \lambda_T^{-3} \quad . \tag{6.34}$$

The calculation of b_3 is more challenging. We have

$$\mathsf{b}_3 = \tfrac{1}{6}\int \frac{d^3\!\rho}{\lambda_T^3}\int \frac{d^3\!r}{\lambda_T^3}\, f(\rho)\, f(r)\, f\big(|\boldsymbol{r}-\boldsymbol{\rho}|\big) \quad . \tag{6.35}$$

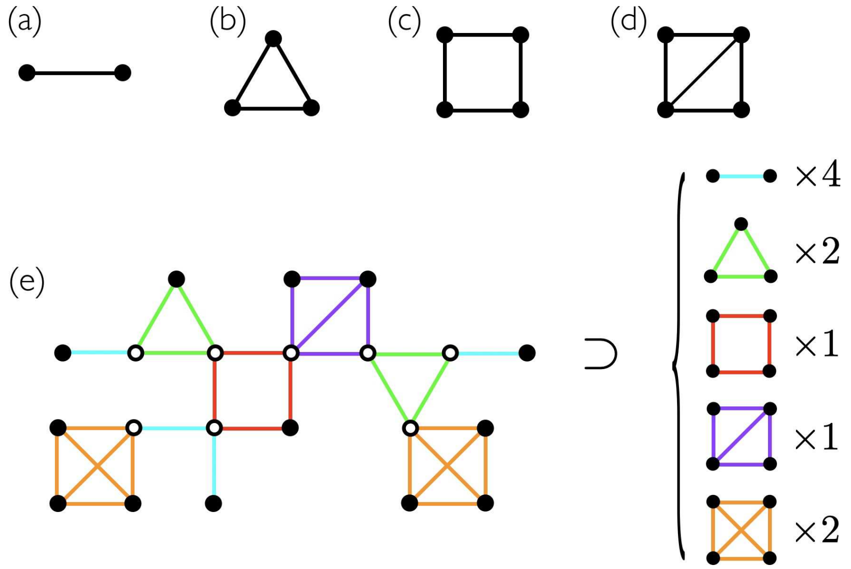


Figure 6.5: Connected *versus* irreducible clusters. Clusters (a) through (d) are *irreducible* in that they remain connected if any component site and its connecting links are removed. Cluster (e) is connected, but is *reducible*. Its integral b_γ is proportional to a product over its irreducible components, 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.

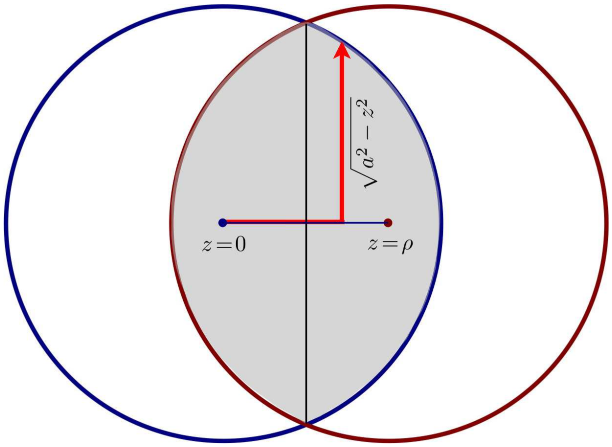


Figure 6.6: The overlap of hard sphere Mayer functions. The shaded volume is $\mathcal{V}$.

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 $\boldsymbol{\rho}$:

$$\begin{aligned}\mathcal{V} &= \int d^3\!r\, f(r)\, f\big(|\boldsymbol{r}-\boldsymbol{\rho}|\big) \\ &= 2\int\limits_{\rho/2}^{a} dz\, \pi(a^2 - z^2) = \tfrac{4\pi}{3}a^3 - \pi a^2\rho + \tfrac{\pi}{12}\,\rho^3 \quad .\end{aligned} \tag{6.36}$$

We then integrate over region $|\boldsymbol{\rho}| < a$, to obtain

$$\mathsf{b}_3 = -\tfrac{1}{6}\cdot 4\pi\lambda_T^{-6}\int\limits_0^a d\rho\,\rho^2\cdot\Big\{\tfrac{4\pi}{3}a^3 - \pi a^2\rho + \tfrac{\pi}{12}\,\rho^3\Big\} = -\tfrac{5\pi^2}{36}\,a^6\lambda_T^{-6} \quad . \tag{6.37}$$

Thus, we have

$$B_2(T) = -\lambda_T^3\,\mathsf{b}_2(T) = \tfrac{2\pi}{3}a^3 \qquad , \qquad B_3(T) = -2\lambda_T^6\,\mathsf{b}_3(T) = \tfrac{5\pi^2}{18}a^6 \tag{6.38}$$

and the equation of state is then

$$p = nk_{\rm B}T\,\Big\{1 + \tfrac{2\pi}{3}a^3 n + \tfrac{5\pi^2}{18}a^6 n^2 + \mathcal{O}(n^3)\Big\} \quad . \tag{6.39}$$

Weakly attractive tail

Suppose

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ -u_0(r) & \text{if } r > a \quad . \end{cases} \tag{6.40}$$

Then the corresponding Mayer function is

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ e^{\beta u_0(r)} - 1 & \text{if } r > a \quad . \end{cases} \tag{6.41}$$

Thus,

$$\mathsf{b}_2(T) = \tfrac{1}{2}\int \frac{d^3\!r}{\lambda_T^3}\, f(r) = -\tfrac{2\pi}{3} a^3 \lambda_T^{-3} + 2\pi\lambda_T^{-3} \int\limits_a^\infty \! dr\, r^2 \left[e^{\beta u_0(r)} - 1 \right] \quad . \tag{6.42}$$

Thus, the second virial coefficient is

$$B_2(T) = -\lambda_T^3\, \mathsf{b}_2(T) \approx \tfrac{2\pi}{3} a^3 - \frac{2\pi}{k_\mathrm{B} T} \int\limits_a^\infty \! dr\, r^2\, u_0(r) \quad , \tag{6.43}$$

where we have assumed $k_\mathrm{B} T \ll u_0(r)$. 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,

$$u(r) = \begin{cases} \infty & \text{if } r \le a \\ -\epsilon & \text{if } a < r < R \\ 0 & \text{if } r > R \quad . \end{cases} \tag{6.44}$$

Then the corresponding Mayer function is

$$f(r) = \begin{cases} -1 & \text{if } r \le a \\ e^{\beta\epsilon} - 1 & \text{if } a < r < R \\ 0 & \text{if } r > R \quad . \end{cases} \tag{6.45}$$

Writing $s \equiv R/a$, we have

$$\begin{aligned} B_2(T) = -\lambda_T^3\, \mathsf{b}_2(T) &= -\tfrac{1}{2}\int \! d^3\!r\, f(r) \\ &= -\frac{1}{2}\left\{ (-1)\cdot \tfrac{4\pi}{3} a^3 + \big(e^{\beta\epsilon} - 1\big)\cdot \tfrac{4\pi}{3} a^3 (s^3 - 1) \right\} \\ &= \tfrac{2\pi}{3} a^3 \left\{ 1 - (s^3 - 1)\big(e^{\beta\epsilon} - 1\big) \right\} \quad . \end{aligned} \tag{6.46}$$

To find the temperature T_0 where $B_2(T)$ changes sign, we set $B_2(T_0) = 0$ and obtain

$$k_\mathrm{B} T_0 = \epsilon \bigg/ \log\!\left(\frac{s^3}{s^3 - 1} \right) \quad . \tag{6.47}$$

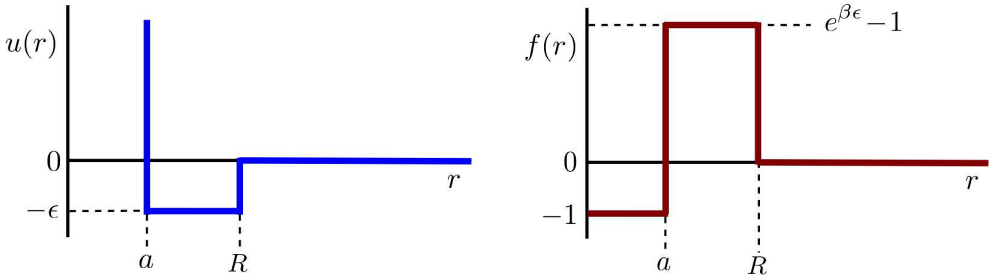


Figure 6.7: An attractive spherical well with a repulsive core $u(r)$ and its associated Mayer function $f(r)$.

Recall in our study of the thermodynamics of the Joule-Thompson 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

$$\Delta T = \int_{p_1}^{p_2} dp \left(\frac{\partial T}{\partial p}\right)_H \quad , \tag{6.48}$$

where

$$\left(\frac{\partial T}{\partial p}\right)_H = \frac{1}{C_p}\left[T\left(\frac{\partial V}{\partial T}\right)_p - V\right] \quad . \tag{6.49}$$

Appealing to the virial expansion, and working to lowest order in corrections to the ideal gas law, we have

$$p = \frac{N}{V}\,k_B T + \frac{N^2}{V^2}\,k_B T\,B_2(T) + \ldots \tag{6.50}$$

and we compute $\left(\frac{\partial V}{\partial T}\right)_p$ by setting

$$0 = dp = -\frac{Nk_B T}{V^2}\,dV + \frac{Nk_B}{V}\,dT - \frac{2N^2}{V^3}\,k_B T\,B_2(T)\,dV + \frac{N^2}{V^2}\,d\big(k_B T\,B_2(T)\big) + \ldots \quad . \tag{6.51}$$

Dividing by dT, we find

$$T\left(\frac{\partial V}{\partial T}\right)_p - V = N\left[T\,\frac{\partial B_2}{\partial T} - B_2\right] \quad . \tag{6.52}$$

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.*

$$\left.\frac{d\log B_2}{d\log T}\right|_{T^*} = 1 \quad . \tag{6.53}$$

If we approximate $B_2(T) \approx A - \frac{B}{T}$, then the inversion temperature follows simply:

$$\frac{B}{T^*} = A - \frac{B}{T^*} \qquad \Longrightarrow \qquad T^* = \frac{2B}{A} \quad . \tag{6.54}$$

6.3 Coulomb Systems : Plasmas and the Electron Gas

6.3.1 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

$$U = \tfrac{1}{2}\int\! d^d\!r \int\! d^d\!r'\, \rho(\boldsymbol{r})\, u(\boldsymbol{r}-\boldsymbol{r}')\, \rho(\boldsymbol{r}') \quad , \tag{6.55}$$

where $\rho(\boldsymbol{r})$ is the charge density and $u(\boldsymbol{r})$, which has the dimensions of (energy)/(charge)^2, satisfies

$$\nabla^2 u(\boldsymbol{r}-\boldsymbol{r}') = -4\pi\, \delta(\boldsymbol{r}-\boldsymbol{r}') \quad . \tag{6.56}$$

Thus,

$$u(\boldsymbol{r}) = \begin{cases} -2\pi\,|x-x'| & , \quad d=1 \\ -2\,\log|\boldsymbol{r}-\boldsymbol{r}'| & , \quad d=2 \\ |\boldsymbol{r}-\boldsymbol{r}'|^{-1} & , \quad d=3 \quad . \end{cases} \tag{6.57}$$

For discete particles, the charge density $\rho(\boldsymbol{r})$ is given by

$$\rho(\boldsymbol{r}) = \sum_i q_i\, \delta(\boldsymbol{r}-\boldsymbol{x}_i) \quad , \tag{6.58}$$

where q_i is the charge of the i^{th} particle. We will assume two types of charges: $q = \pm e$, with $e > 0$. The electric potential is

$$\phi(\boldsymbol{r}) = \int\! d^d\!r'\, u(\boldsymbol{r}-\boldsymbol{r}')\, \rho(\boldsymbol{r}') = \sum_i q_i\, u(\boldsymbol{r}-\boldsymbol{x}_i) \quad . \tag{6.59}$$

This satisfies the Poisson equation, $\nabla^2\phi(\boldsymbol{r}) = -4\pi\rho(\boldsymbol{r})$. The total potential energy can be written as

$$U = \tfrac{1}{2}\int\! d^d\!r\, \phi(\boldsymbol{r})\, \rho(\boldsymbol{r}) = \tfrac{1}{2}\sum_i q_i\, \phi(\boldsymbol{x}_i) \quad , \tag{6.60}$$

where it is understood that we omit self-interaction terms.

6.3.2 Debye-Hückel theory

We now write the grand partition function:

$$\begin{aligned} \varXi(T,V,\mu_+,\mu_-) = \sum_{N_+=0}^{\infty} \sum_{N_-=0}^{\infty} & \frac{1}{N_+!}\, e^{\beta\mu_+ N_+}\, \lambda_+^{-N_+ d} \cdot \frac{1}{N_-!}\, e^{\beta\mu_- N_-}\, \lambda_-^{-N_- d} \\ & \cdot \int\! d^d\!r_1 \cdots \int\! d^d\!r_{N_{\text{tot}}}\; e^{-\beta U(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_{N_{\text{tot}}})} \quad , \end{aligned} \tag{6.61}$$

where $N_{\rm tot} = N_+ + N_-$. We adopt a *mean field* approach, known as *Debye-Hückel theory*, writing

$$\rho(\boldsymbol{r}) = \bar\rho(\boldsymbol{r}) + \delta\rho(\boldsymbol{r}) \qquad , \qquad \phi(\boldsymbol{r}) = \bar\phi(\boldsymbol{r}) + \delta\phi(\boldsymbol{r}) \quad . \tag{6.62}$$

We then have

$$\begin{aligned} U &= \tfrac{1}{2}\!\int\! d^d\!r \, \big[\bar\rho(\boldsymbol{r}) + \delta\rho(\boldsymbol{r})\big] \cdot \big[\bar\phi(\boldsymbol{r}) + \delta\phi(\boldsymbol{r})\big] \\ &= \tfrac{1}{2}\!\int\! d^d\!r \, \bar\phi(\boldsymbol{r}) \, \bar\rho(\boldsymbol{r}) + \overbrace{\tfrac{1}{2}\!\int\! d^d\!r \, \bar\phi(\boldsymbol{r}) \, \delta\rho(\boldsymbol{r}) + \tfrac{1}{2}\!\int\! d^d\!r \, \bar\rho(\boldsymbol{r}) \, \delta\phi(\boldsymbol{r})}^{\text{these two terms are the same}} + \tfrac{1}{2}\!\int\! d^d\!r \, \delta\rho(\boldsymbol{r}) \, \delta\phi(\boldsymbol{r}) \\ &= -\!\int\! d^d\!r \, \bar\phi(\boldsymbol{r}) \, \bar\rho(\boldsymbol{r}) + \int\! d^d\!r \, \bar\phi(\boldsymbol{r}) \, \rho(\boldsymbol{r}) + \overbrace{\tfrac{1}{2}\!\int\! d^d\!r \, \delta\rho(\boldsymbol{r}) \, \delta\phi(\boldsymbol{r})}^{\text{ignore fluctuation term}} \\ &= -U_0 + \int\! d^d\!r \, \bar\phi(\boldsymbol{r}) \, \rho(\boldsymbol{r}) + (\text{flucts})^2 \quad , \end{aligned} \tag{6.63}$$

where $U_0 \equiv \frac{1}{2}\!\int\! d^d\!r \, \bar\phi(\boldsymbol{r}) \, \bar\rho(\boldsymbol{r})$, and we where have used $\delta\rho = \rho - \bar\rho$. Thus we have

$$\Xi = e^{U_0/k_{\rm B}T} \exp\!\bigg(z_+ \lambda_+^{-d} \!\int\! d^d r_+ \, e^{-e\,\bar\phi(\boldsymbol{r}_+)/k_{\rm B}T} \bigg) \exp\!\bigg(z_- \lambda_-^{-d} \!\int\! d^d r_- \, e^{e\,\bar\phi(\boldsymbol{r}_-)/k_{\rm B}T} \bigg) \tag{6.64}$$

whence

$$\Omega(T,V,\mu_+,\mu_-) = -U_0 - k_{\rm B}T z_+ \lambda_+^{-d} \!\int\! d^d\!r \, \exp\!\bigg(-\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T} \bigg) - k_{\rm B}T z_- \lambda_-^{-d} \!\int\! d^d\!r \, \exp\!\bigg(+\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T} \bigg) \quad , \tag{6.65}$$

where

$$\lambda_\pm = \bigg(\frac{2\pi\hbar^2}{m_\pm k_{\rm B}T} \bigg) \qquad , \qquad z_\pm = \exp\!\bigg(\frac{\mu_\pm}{k_{\rm B}T} \bigg) \quad . \tag{6.66}$$

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

$$\frac{\delta U_0}{\delta\bar\phi(\boldsymbol{r})} = \bar\rho(\boldsymbol{r}) \quad . \tag{6.67}$$

We next demand that the free energy Ω is extremized with respect to the mean field $\bar\phi(\boldsymbol{r})$, *viz.*

$$0 = \frac{\delta\Omega}{\delta\bar\phi(\boldsymbol{r})} = -\bar\rho(\boldsymbol{r}) + e\,\lambda_+^{-d} \, z_+ \, \exp\!\bigg(-\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T} \bigg) - e\,\lambda_-^{-d} \, z_- \, \exp\!\bigg(+\frac{e\,\bar\phi(\boldsymbol{r})}{k_{\rm B}T} \bigg) \quad . \tag{6.68}$$

At $r \to \infty$, we assume charge neutrality and $\phi(\infty) = 0$. Thus

$$\lambda_+^{-d} \, z_+ = n_+(\infty) = \lambda_-^{-d} \, z_- = n_-(\infty) \equiv n_\infty \quad , \tag{6.69}$$

where n_∞ is the ionic density of either species at infinity. Therefore,

$$\rho(\boldsymbol{r}) = -2e\,n_\infty \sinh\!\bigg(\frac{e\,\phi(\boldsymbol{r})}{k_{\rm B}T} \bigg) \quad , \tag{6.70}$$

where we have dropped the bars on $\bar{\phi}$ and $\bar{\rho}$ for convenience. We now invoke Poisson's equation,

$$\nabla^2\phi = 8\pi e n_\infty \sinh(\beta e\phi) - 4\pi\rho_{\rm ext} \quad , \tag{6.71}$$

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

$$\nabla^2\phi = \kappa_{\rm D}^2\,\phi - 4\pi\rho_{\rm ext} \quad , \tag{6.72}$$

where

$$\kappa_{\rm D} = \bigg(\frac{8\pi n_\infty e^2}{k_{\rm B}T}\bigg)^{1/2} \qquad , \qquad \lambda_{\rm D} = \bigg(\frac{k_{\rm B}T}{8\pi n_\infty e^2}\bigg)^{1/2} \quad . \tag{6.73}$$

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,

$$\nabla^2\phi = \kappa_{\rm D}^2\,\phi - 4\pi Q\,\delta(\boldsymbol{r}) \quad . \tag{6.74}$$

Fourier transforming, we obtain

$$-\boldsymbol{q}^2\,\hat{\phi}(\boldsymbol{q}) = \kappa_{\rm D}^2\,\hat{\phi}(\boldsymbol{q}) - 4\pi Q \qquad \Longrightarrow \qquad \hat{\phi}(\boldsymbol{q}) = \frac{4\pi Q}{\boldsymbol{q}^2 + \kappa_{\rm D}^2} \quad . \tag{6.75}$$

Transforming back to real space, we obtain, in three dimensions, the Yukawa potential,

$$\phi(\boldsymbol{r}) = \int\!\frac{d^3\!q}{(2\pi)^3}\,\frac{4\pi Q\,e^{i\boldsymbol{q}\cdot\boldsymbol{r}}}{\boldsymbol{q}^2 + \kappa_{\rm D}^2} = \frac{Q}{r}\,e^{-\kappa_{\rm D} r} \quad . \tag{6.76}$$

This solution must break down sufficiently close to $r=0$, since the assumption $e\phi(r) \ll k_{\rm B}T$ is no longer valid there. However, for larger r, the Yukawa form is increasingly accurate.

For another example, consider an electrolyte held between two conducting plates, one at potential $\phi(x=0)=0$ 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

$$\phi(x) = A\,e^{\kappa_{\rm D} x} + B\,e^{-\kappa_{\rm D}\,x} \quad . \tag{6.77}$$

We fix the constants A and B by invoking the boundary conditions, which results in

$$\phi(x) = \frac{V\sinh(\kappa_{\rm D} x)}{\sinh(\kappa_{\rm D} L)} \quad . \tag{6.78}$$

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.

6.3.3 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

$$\varepsilon_{\rm F} = \frac{\hbar^2 k_{\rm F}^2}{2m} - e\phi(\boldsymbol{r}) \quad . \tag{6.79}$$

Thus, the Fermi wavevector $k_{\rm F}$ is spatially varying, according to the relation

$$k_{\rm F}(\boldsymbol{r}) = \left[\frac{2m}{\hbar^2}\Big(\varepsilon_{\rm F} + e\phi(\boldsymbol{r})\Big)\right]^{1/2} \quad . \tag{6.80}$$

The local electron number density is

$$n(\boldsymbol{r}) = \frac{k_{\rm F}^3(\boldsymbol{r})}{3\pi^2} = n_\infty \left(1 + \frac{e\phi(\boldsymbol{r})}{\varepsilon_{\rm F}}\right)^{3/2} \quad . \tag{6.81}$$

In the presence of a uniform compensating positive background charge $\rho_+ = en_\infty$, Poisson's equation takes the form

$$\nabla^2\phi = 4\pi e\, n_\infty \left[\left(1 + \frac{e\phi(\boldsymbol{r})}{\varepsilon_{\rm F}}\right)^{3/2} - 1\right] - 4\pi\rho_{\rm ext}(\boldsymbol{r}) \quad . \tag{6.82}$$

If $e\phi \ll \varepsilon_{\rm F}$, we may expand in powers of the ratio, obtaining

$$\nabla^2\phi = \frac{6\pi n_\infty e^2}{\varepsilon_{\rm F}}\,\phi \equiv \kappa_{\rm TF}^2\,\phi - 4\pi\rho_{\rm ext}(\boldsymbol{r}) \quad . \tag{6.83}$$

Here, $\kappa_{\rm TF}$ is the *Thomas-Fermi wavevector*,

$$\kappa_{\rm TF} = \left(\frac{6\pi n_\infty e^2}{\varepsilon_{\rm F}}\right)^{1/2} \quad . \tag{6.84}$$

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 $[\mathrm{Ar}]\,3d^{10}\,4s^1$. The $4s$ electron forms an energy band which contains the Fermi surface. Zinc has a configuration of

[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,

$$\phi(\boldsymbol{r}) = \frac{Q}{r}\, e^{-\kappa_{\rm TF} r} \quad . \tag{6.85}$$

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

$$\delta n(\boldsymbol{r}) = e\phi(\boldsymbol{r})\, g(\varepsilon_{\rm F}) \quad , \tag{6.86}$$

where $g(\varepsilon_{\rm F})$ is the density of states at the Fermi energy. Thus, in a metal, we should write

$$\nabla^2\phi = (-4\pi)(-e\,\delta n) = 4\pi e^2 g(\varepsilon_{\rm F})\,\phi = \kappa^2_{\rm TF}\,\phi \quad , \tag{6.87}$$

where

$$\kappa_{\rm TF} = \sqrt{4\pi e^2\, g(\varepsilon_{\rm F})} \quad . \tag{6.88}$$

The value of $g(\varepsilon_{\rm F})$ will depend on the form of the dispersion. For ballistic bands with an effective mass m^*, the formula in eqn. 6.83 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

$$\nabla^2\phi(\boldsymbol{r}) = 4\pi e \cdot \frac{(2m)^{3/2}}{3\pi^2\hbar^3}\Big(\varepsilon_{\rm F} + e\phi(\boldsymbol{r})\Big)^{3/2} - 4\pi Z e\,\delta(\boldsymbol{r}) \quad . \tag{6.89}$$

We assume an isotropic solution. It is then convenient to define

$$\varepsilon_{\rm F} + e\phi(\boldsymbol{r}) = \frac{Ze^2}{r}\cdot\chi(r/r_0) \quad , \tag{6.90}$$

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

$$\nabla^2\left\{\frac{Ze^2}{r}\cdot\chi(r/r_0)\right\} = \frac{1}{r_0^2}\,\frac{Ze^2}{r}\,\chi''(r/r_0) \quad , \tag{6.91}$$

thus we arrive at the *Thomas-Fermi equation*,

$$\chi''(t) = \frac{1}{\sqrt{t}}\,\chi^{3/2}(t) \quad , \tag{6.92}$$

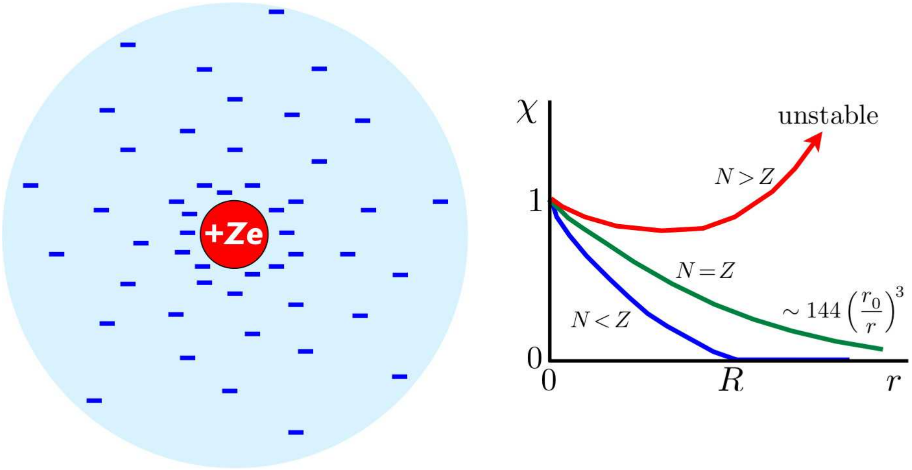


Figure 6.8: The Thomas-Fermi atom consists of a nuclear charge $+Ze$ surrounded by N electrons distributed in a cloud. The electric potential $\phi(\boldsymbol{r})$ felt by any electron at position $\boldsymbol{r}$ is screened by the electrons within this radius, resulting in a self-consistent potential $\phi(\boldsymbol{r}) = \phi_0 + (Ze^2/r)\,\chi(r/r_0)$.

with $r = t\,r_0$, provided we take

$$r_0 = \frac{\hbar^2}{2me^2}\left(\frac{3\pi}{4\sqrt{Z}}\right)^{2/3} = 0.885\,Z^{-1/3}\,a_{\rm B} \quad , \tag{6.93}$$

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

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi = \left[-\frac{Ze^2}{R^2}\,\chi(R/r_0) + \frac{Ze^2}{R\,r_0}\,\chi'(R/r_0)\right]\hat{\boldsymbol{r}} = \frac{(Z-N)\,e}{R^2}\,\hat{\boldsymbol{r}} \quad . \tag{6.94}$$

This requires

$$-t^*\,\chi'(t^*) = 1 - \frac{N}{Z} \quad . \tag{6.95}$$

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.

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* (3rd 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.

– 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. 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[2]

$$p_{\rm c}=\frac{a}{27\,b^2} \quad , \quad v_{\rm c}=3b \quad , \quad T_{\rm c}=\frac{8a}{27\,bR} \quad . \tag{7.5}$$

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] 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 $p_{\rm c} = p(T_{\rm c}, v_{\rm c})$.

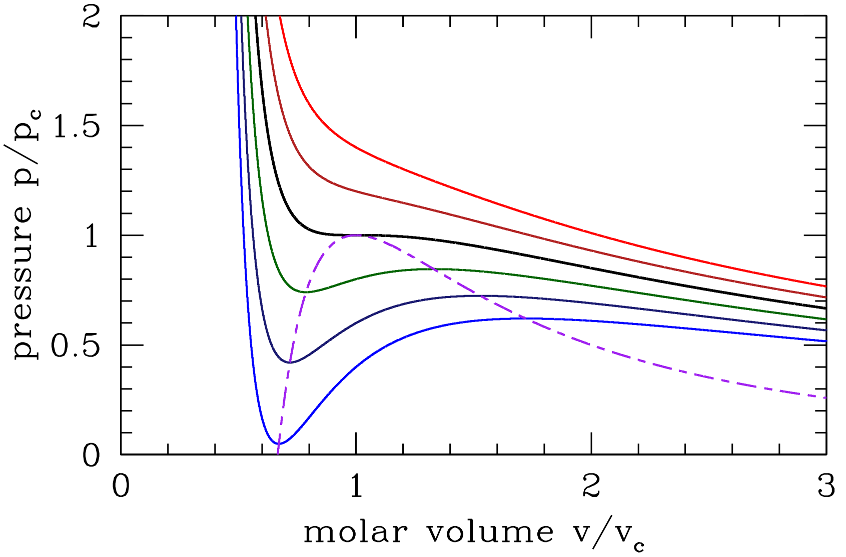


Figure 7.1: Pressure *versus* molar volume for the van der Waals gas at temperatures in equal intervals from $T = 1.10\,T_{\rm c}$ (red) to $T = 0.85\,T_{\rm c}$ (blue). The purple curve is $\bar{p}^*(\bar{v})$.

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

$$\bigg(\frac{\partial \varepsilon}{\partial v}\bigg)_T = T\bigg(\frac{\partial p}{\partial T}\bigg)_v - p = \frac{a}{v^2} \quad , \tag{7.8}$$

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 freedoms, 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[3],

$$f(T,v) = \tfrac{1}{2}\mathfrak{f}RT\Big(1 - \log\big(T/T_{\rm c}\big)\Big) - \frac{a}{v} - RT\log(v-b) - Ts_0 \quad . \tag{7.11}$$

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	7404	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 $(\partial^2 f/\partial v^2)_T < 0$ 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. The total volume and temperature thus remain fixed, but the free energy changes by an amount $\Delta f = \frac{1}{2}\big(\partial^2 f/\partial v^2\big)_T(\delta v)^2 < 0$. 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

$$\frac{\partial^2 f}{\partial v^2}\bigg|_T = -\frac{\partial p}{\partial v}\bigg|_T = \frac{1}{v\kappa_T} = 0 \quad , \tag{7.12}$$

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, $\big(\partial^2 f/\partial v^2\big)_T < 0$, is in fact too strong. That is, the system can be unstable even at molar volumes where $\big(\partial^2 f/\partial v^2\big)_T > 0$. The reason is shown graphically in fig. 7.2. At the fixed temperature T, for any molar volume v between $v_\mathrm{liquid} \equiv v_1$ and $v_\mathrm{gas} \equiv v_2$, 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

$$f = (1-x)\,f_1 + x\,f_2 \quad , \tag{7.14}$$

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.

---

[3] Don't confuse the molar free energy (f) with the number of molecular degrees of freedom ($\mathfrak{f}$)!

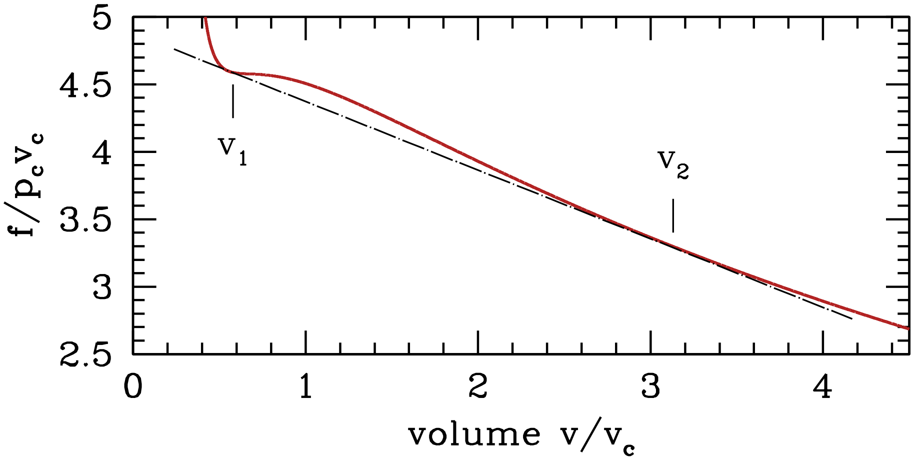


Figure 7.2: The molar free energy $f(T,v)$ of the van der Waals system at $T = 0.85\, T_c$. The dot-dashed black line shows the Maxwell construction connecting molar volumes $v_{1,2}$ on opposite sides of the coexistence curve.

The two equations which give us v_1 and v_2 are

$$\frac{\partial f}{\partial v}\bigg|_{v_1,T} = \frac{\partial f}{\partial v}\bigg|_{v_2,T} = \frac{f(T,v_2)-f(T,v_1)}{(v_2-v_1)} \quad . \tag{7.15}$$

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 $T > T_c$, the volume v decreases continuously as the pressure p increases. If $T < T_c$, 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*. Note that the volume discontinuity, $\Delta v = w_- \propto (T_c - T)^{1/2}$. This is an example of a *critical behavior* in which the *order parameter* ϕ, which in this case may be taken to be the difference $\phi = v_G - v_L$, behaves as a power law in $\big|T - T_c\big|$, where T_c is the *critical temperature*. In this case, we have $\phi(T) \propto (T_c - T)_+^{\beta}$, where $\beta = \frac{1}{2}$ is the exponent, and where $(T_c - T)_+$ is defined to be $T_c - T$ if $T < T_c$ and 0 otherwise. Recall

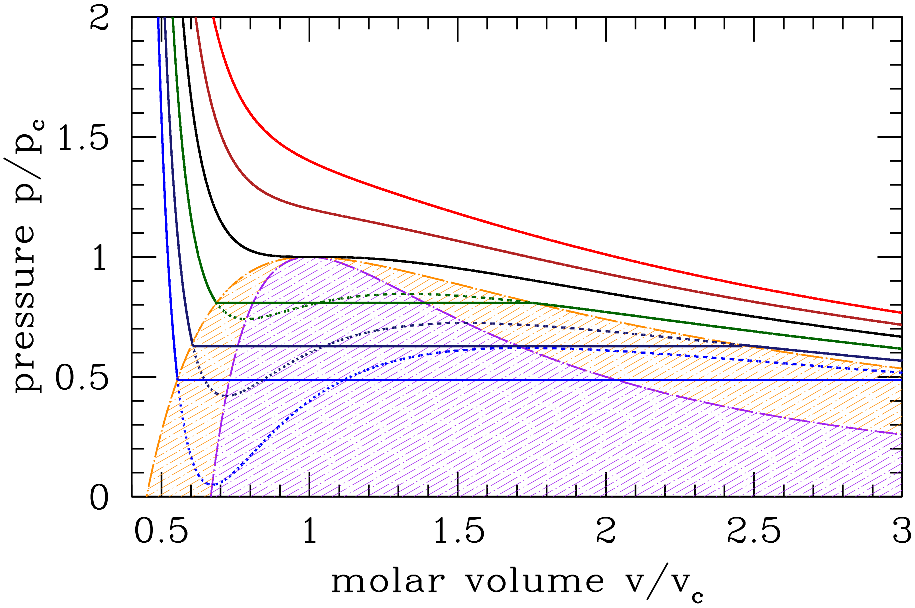


Figure 7.3: Pressure-volume isotherms for the van der Waals system, as in fig. 7.1, but corrected to account for the Maxwell construction. The boundary of the purple shaded region is the spinodal line $\bar{p}^*(\bar{v})$. The boundary of the orange shaded region is the stability boundary with respect to phase separation, and is called the *coexistence curve*.

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

Close to the critical point, the dimensionless equation of state may be written as $\pi = \pi(\epsilon, t)$, where

$$\bar{p} = 1 + \pi \quad , \quad \bar{v} = 1 + \epsilon \quad , \quad \bar{T} = 1 + t \quad , \tag{7.19}$$

where $\pi(0,0) = 0$. Equivalently,

$$\pi = \frac{p - p_{\rm c}}{p_{\rm c}} \quad , \quad \epsilon = \frac{v - v_{\rm c}}{v_{\rm c}} \quad , \quad t = \frac{T - T_{\rm c}}{T_{\rm c}} = \frac{\Theta}{T_{\rm c}} \quad . \tag{7.20}$$

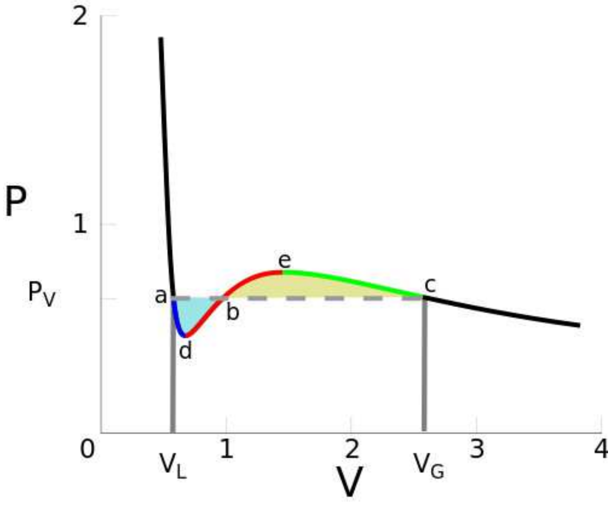


Figure 7.4: Maxwell construction in the (v, p) plane. The system is absolutely unstable between volumes v_{d} and v_{e}. For $v \in [v_{\mathrm{a}}, v_{\mathrm{d}}]$ of $v \in [v_{\mathrm{e}}, v_{\mathrm{c}}]$, the solution is unstable with respect to phase separation. Source: *Wikipedia*.

Here π, ϵ, 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

$$\begin{aligned}\pi(\epsilon, t) &= \frac{8(1+t)}{2+3\epsilon} - \frac{3}{(1+\epsilon)^2} - 1 \\ &= 4t - 6t\epsilon + 9\epsilon^2 t - \tfrac{3}{2}\epsilon^3 - \tfrac{27}{2}\epsilon^3 t + \tfrac{21}{4}\epsilon^4 + \tfrac{81}{4}\epsilon^4 t - \tfrac{99}{8}\epsilon^5 + \ldots \quad .\end{aligned} \tag{7.21}$$

from which we may derive

$$\epsilon_{\mathsf{L,G}}(t) = \mp\, 2\sqrt{-t} - \tfrac{5}{18}\, t + \ldots \quad . \tag{7.22}$$

We identify the difference $\Delta v \equiv v_{\mathsf{L}} - v_{\mathsf{G}}$ as the *order parameter* for the transition which occurs at $T = T_{\mathrm{c}}$. 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

$$\frac{\partial \pi}{\partial \epsilon} = -6t + 18\epsilon t - \tfrac{9}{2}\epsilon^2 - \tfrac{81}{2}\epsilon^2 t + 21\epsilon^3 + \ldots \quad . \tag{7.23}$$

For the spinodal, 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 $\bar{T} = T/T_{\mathrm{c}}$ is fairly well-fit to the curve[4]

$$\bar{n}_{\mathsf{L,G}}(t) = 1 \pm \tfrac{7}{4}(1 - \bar{T})^{1/3} + \tfrac{3}{4}(1 - \bar{T}) \tag{7.24}$$

---

[4] See M. Schwartz, `https://scholar.harvard.edu/files/schwartz/files/9-phases.pdf`.

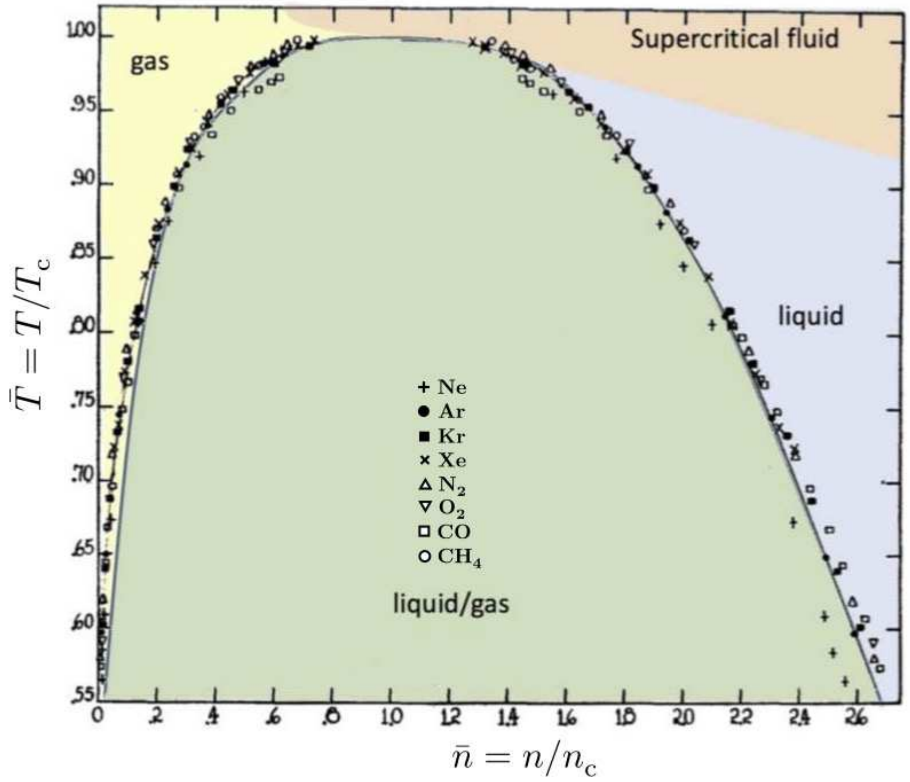


Figure 7.5: Universality of the liquid-gas transition for eight different atomic and molecular fluids, from E. A. Guggenheim, *J. Chem. Phys.* **13**, 253 (1945). Dimensionless temperature $\bar{T} = T/T_c$ *versus* dimensionless number density $\bar{n} = n/n_c = v_c/v$ is shown. The van der Waals / mean field theory gives $\Delta n = n_{\text{liquid}} - q_{\text{gas}} \propto (-t)^{1/2}$, 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.)

which shows that the critical exponent β is much closer to $\beta = \frac{1}{3}$ than to the vdW value $\beta^{\text{vdW}} = \frac{1}{2}$. 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[5], *"Over de Continuïteit van den Gas - en Vloeistoftoestand"* ("On the continuity of the gas and liquid state"). In his Nobel lecture[6], van der Waals writes of how he was inspired by Rudolf Clausius' 1857 treatise on the nature of heat,

---

[5] 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.

[6] See `http://www.nobelprize.org/nobel_prizes/physics/laureates/1910/waals-lecture.pdf`

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,

$$p = \frac{RT}{v-b} - \frac{a}{v^2} \quad . \tag{7.25}$$

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 ω_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 $\zeta_{\rm max} = \frac{\pi}{3\sqrt{2}} = 0.74078$. Dense random packing results in $\zeta_{\rm drp} = 0.634$. Expanding the vdW equation of state in inverse powers of v yields

$$p = \frac{RT}{v} + \left(b - \frac{a}{RT}\right) \cdot \frac{RT}{v^2} + \mathcal{O}\big(v^{-3}\big) \quad , \tag{7.26}$$

and we read of 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 $\zeta_{\sf cub} = \frac{\pi}{6}$.

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 $v_{\sf G} - v_{\sf L} \propto (T_{\rm c} - T)^{\beta}$ with noninteger exponent β 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

$$\hat{H}(\boldsymbol{p},\boldsymbol{x}) = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} + \sum_{i<j} u(\boldsymbol{x}_i - \boldsymbol{x}_j) \quad . \tag{7.27}$$

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

$$U = \sum_{i<j} u(\boldsymbol{x}_i - \boldsymbol{x}_j) = \tfrac{1}{2} \sum_{\boldsymbol{R}\neq\boldsymbol{R}'} V_{\boldsymbol{R}\boldsymbol{R}'}\, n_{\boldsymbol{R}}\, n_{\boldsymbol{R}'} \quad , \tag{7.28}$$

where $V_{\boldsymbol{R}\boldsymbol{R}'} \approx v(\boldsymbol{R} - \boldsymbol{R}')$, where R_k is the position at the center of cube k. The grand partition function is then approximated as

$$\Xi(T,V,\mu) \approx \sum_{\{n_{\boldsymbol{R}}\}} \bigg(\prod_{\boldsymbol{R}} \xi^{n_{\boldsymbol{R}}}\bigg) \exp\bigg(-\tfrac{1}{2}\beta \sum_{\boldsymbol{R}\neq\boldsymbol{R}'} V_{\boldsymbol{R}\boldsymbol{R}'}\, n_{\boldsymbol{R}}\, n_{\boldsymbol{R}'}\bigg) \quad , \tag{7.29}$$

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 λ_T^{-d} factor arises from the integration over the momenta. Note $\sum_{\boldsymbol{R}} n_{\boldsymbol{R}} = N$ is the total number of fluid particles, so

$$\prod_{\boldsymbol{R}} \xi^{n_{\boldsymbol{R}}} = \xi^N = e^{\beta\mu N}\, \lambda_T^{-Nd}\, a^{Nd} \quad . \tag{7.30}$$

Thus, we can write a lattice Hamiltonian,

$$\begin{aligned} \hat{H} &= \tfrac{1}{2} \sum_{\boldsymbol{R}\neq\boldsymbol{R}'} V_{\boldsymbol{R}\boldsymbol{R}'}\, n_{\boldsymbol{R}}\, n_{\boldsymbol{R}'} - k_{\rm B}T \log\xi \sum_{\boldsymbol{R}} n_{\boldsymbol{R}} \\ &= -\tfrac{1}{2} \sum_{\boldsymbol{R}\neq\boldsymbol{R}'} J_{\boldsymbol{R}\boldsymbol{R}'}\, \sigma_{\boldsymbol{R}}\, \sigma_{\boldsymbol{R}'} - H \sum_{\boldsymbol{R}} \sigma_{\boldsymbol{R}} + E_0 \quad , \end{aligned} \tag{7.31}$$

where $\sigma_{\boldsymbol{R}} \equiv 2n_{\boldsymbol{R}} - 1$ is a spin variable taking the possible values $\{-1,+1\}$, and

$$J_{\boldsymbol{R}\boldsymbol{R}'} = -\tfrac{1}{4} V_{\boldsymbol{R}\boldsymbol{R}'} \qquad , \qquad H = \tfrac{1}{2} k_{\rm B}T \log\xi - \tfrac{1}{4} {\sum_{\boldsymbol{R}'}}' V_{\boldsymbol{R}\boldsymbol{R}'} \qquad , \qquad E_0 = \tfrac{1}{2} \sum_{\boldsymbol{R}\neq\boldsymbol{R}'} V_{\boldsymbol{R}\boldsymbol{R}'} \quad , \tag{7.32}$$

where the prime on the sum indicates that $\boldsymbol{R}' = \boldsymbol{R}$ is to be excluded. For the Lennard-Jones system, $V_{\boldsymbol{R}\boldsymbol{R}'} = v(\boldsymbol{R} - \boldsymbol{R}') < 0$ is due to the attractive tail of the potential, hence $J_{\boldsymbol{R}\boldsymbol{R}'}$ is positive, which prefers alignment of the spins $\sigma_{\boldsymbol{R}}$ and $\sigma_{\boldsymbol{R}'}$. This interaction is therefore *ferromagnetic*. The spin Hamiltonian in eqn. 7.31 is known as the Ising model.

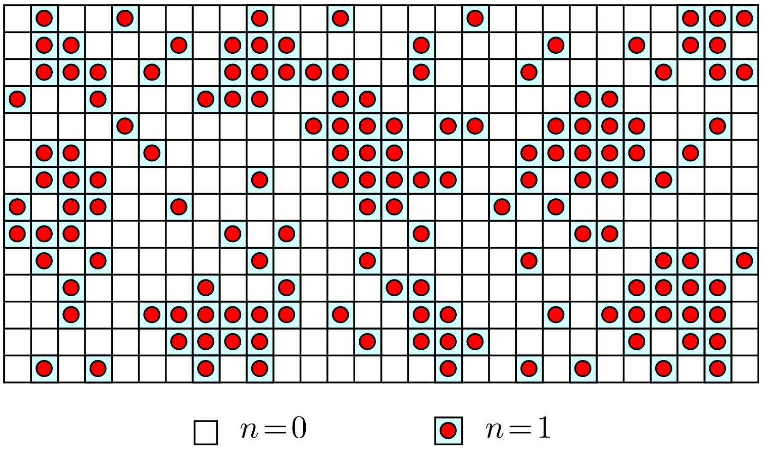


Figure 7.6: The lattice gas model. An occupied cell corresponds to $n = 1$ ($\sigma = +1$), and a vacant cell to $n = 0$ ($\sigma = -1$).

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[7]

$$p \longleftrightarrow H \qquad , \qquad v \longleftrightarrow m \quad ,$$

where m is the magnetization density, defined here to be the total magnetization M divided by the number of lattice sites N:[8]

$$m = \frac{M}{N} = \frac{1}{N} \sum_{\boldsymbol{R}} \langle \sigma_{\boldsymbol{R}} \rangle \quad . \tag{7.33}$$

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 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 $T < T_c$, 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

---

[7] H is more properly analogous to μ. 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 $p/k_B T$.

[8] Note the distinction between the number of lattice sites N and the number of occupied cells N. According to our definitions, $N = \frac{1}{2}(M + N)$.

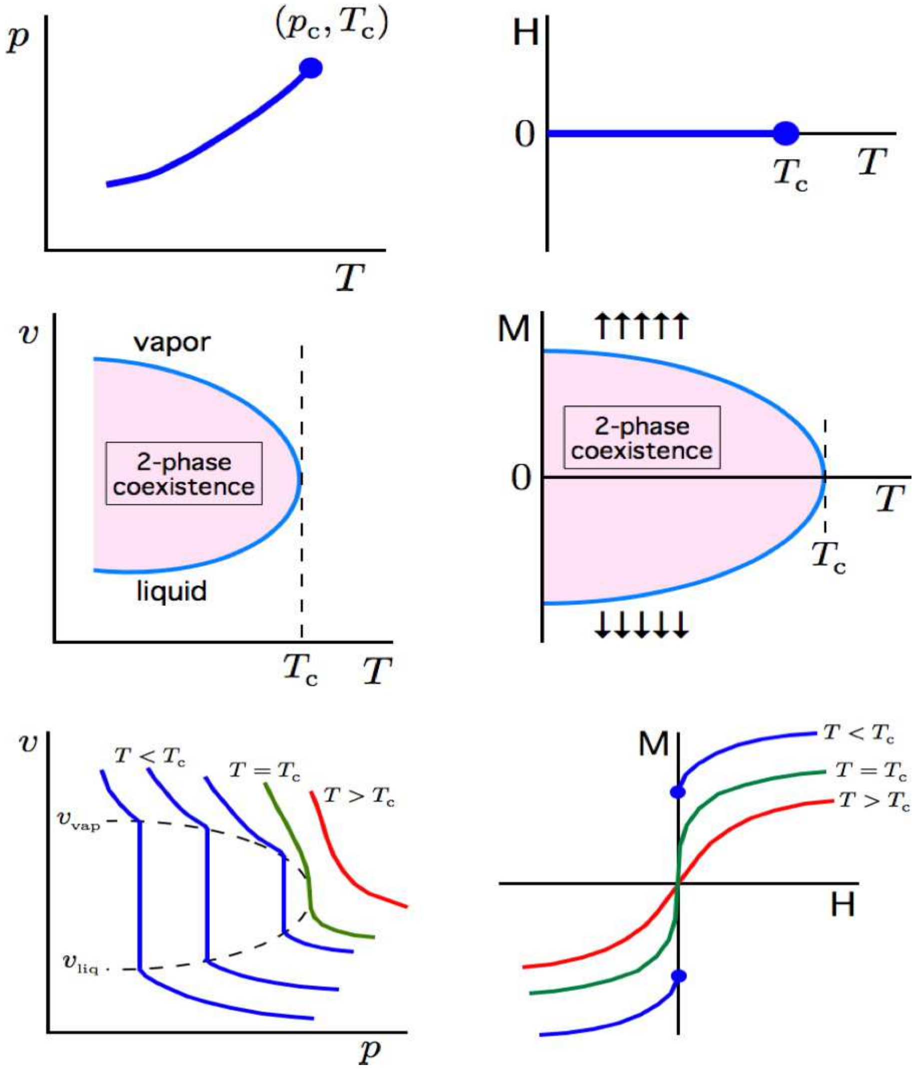


Figure 7.7: Comparison of the liquid-gas phase diagram with that of the Ising ferromagnet.

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 κ_T and the isothermal susceptibility χ_T:

$$\kappa_T = -\frac{1}{v}\left(\frac{\partial v}{\partial p}\right)_T \quad , \quad \chi_T = \left(\frac{\partial m}{\partial H}\right)_T \quad , \tag{7.34}$$

with $\kappa_T(T_c, p_c) = \infty$ and $\chi_T(T_c, H_c) = \infty$.

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 $\Psi \propto (v_{vap} - v_{liq})$, 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.*

$$m(T,H_c) \sim \big(T_c - T\big)_+^\beta \quad , \quad \chi(T,H_c) \sim |T-T_c|^{-\gamma} \quad , \quad C_M(T,H_c) \sim |T-T_c|^{-\alpha} \quad , \quad m(T_c,H) \sim \pm |H|^{1/\delta} \quad . \tag{7.35}$$

The quantities α, β, γ, and δ are the *critical exponents* associated with the transition. These exponents satisfy certain equalities, such as the Rushbrooke and Griffiths relations:

$$\alpha + 2\beta + \gamma = 2 \quad \text{(Rushbrooke)} \qquad , \qquad \beta + \gamma = \beta\delta \quad \text{(Griffiths)} \quad . \tag{7.36}$$

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 α, β, γ, and δ, one defines the *correlation length exponent* ν 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[9]. In the limit $T \to T_c$, one has

$$C(\boldsymbol{r},T,h) = r^{-(d-2+\eta)}\, \phi\big(r/\xi(T), H/\xi_H(T)\big) \quad , \tag{7.37}$$

where η is the *anomalous exponent*, $\phi(r/\xi, H/\xi_H)$ is a *scaling function*, $\xi(T) \propto |T - T_c|^{-\nu}$ is the *correlation length*, and $\xi_H(T) \propto |T - T_c|^{\Delta}$, 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 η and ν come additional exponent relations,

$$(2-\eta)\nu = \gamma \qquad , \qquad d\,\nu = 2 - \alpha \quad \text{(hyperscaling)} \quad . \tag{7.38}$$

Thus there are three relations among the six critical exponents α, β, γ, δ, η, and ν, which entails that there are three independent values among the six.

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 $dE = T\,dS + H\,dM + \mu NS$, we have

$$S\,dT + M\,dH + N\,d\mu = 0 \quad . \tag{7.39}$$

Now consider two magnetic phases in coexistence. We must have $d\mu_1 = d\mu_2$, hence

$$d\mu_1 = -s_1\,dT - m_1\,dH = -s_2\,dT - m_2\,dH = d\mu_2 \quad , \tag{7.40}$$

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,

$$\bigg(\frac{dH}{dT}\bigg)_{\text{coex}} = -\frac{s_1 - s_2}{m_1 - m_2} \quad . \tag{7.41}$$

---

[9] 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.

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 §2.12.2 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[10]. This leads to the Clapeyron relation,

$$\bigg(\frac{dp}{dT}\bigg)_{\rm coex} = \frac{s_2 - s_1}{v_2 - v_1} = \frac{\ell}{T\,\Delta v} \quad , \tag{7.42}$$

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 β-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

$$E_{ij} = \varepsilon_{\rm AA}\, P^{\rm A}_i\, P^{\rm A}_j + \varepsilon_{\rm BB}\, P^{\rm B}_i\, P^{\rm B}_j + \varepsilon_{\rm AB}\,\big(P^{\rm A}_i\, P^{\rm B}_j + P^{\rm B}_i\, P^{\rm A}_j\big) \quad , \tag{7.43}$$

---

[10] In §2.12.2 we considered, equivalently, the differential of the molar free energy $g = N_{\rm A}\mu$, with v the molar volume.

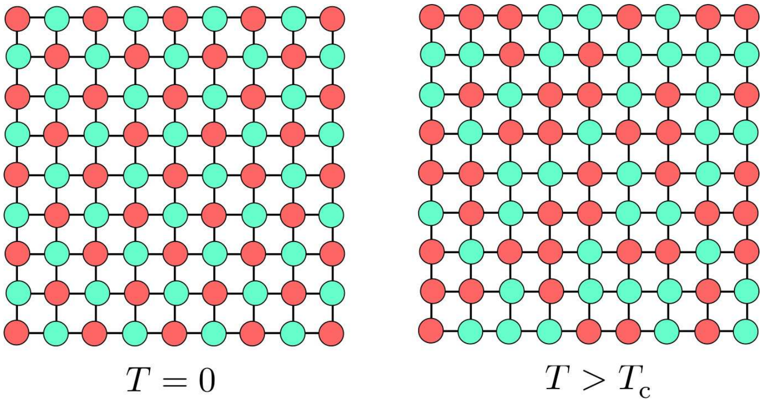


Figure 7.8: Order-disorder transition on the square lattice. Below $T = T_{\rm c}$, order develops spontaneously on the two $\sqrt{2} \times \sqrt{2}$ sublattices. There is perfect sublattice order at $T = 0$ (left panel).

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

$$\begin{aligned} \hat{H} = \sum_{\langle ij \rangle} E_{ij} &= \sum_{\langle ij \rangle} \Big\{ \tfrac{1}{4}\big(\varepsilon_{\rm AA} + \varepsilon_{\rm BB} - 2\varepsilon_{\rm AB}\big)\,\sigma_i\,\sigma_j + \tfrac{1}{4}\big(\varepsilon_{\rm AA} - \varepsilon_{\rm BB}\big)\,(\sigma_i + \sigma_j) + \tfrac{1}{4}\big(\varepsilon_{\rm AA} + \varepsilon_{\rm BB} + 2\varepsilon_{\rm AB}\big) \Big\} \\ &= -J \sum_{\langle ij \rangle} \sigma_i\,\sigma_j - H \sum_i \sigma_i + E_0 \quad , \end{aligned} \tag{7.44}$$

where the exchange constant J and the magnetic field H are given by

$$J = \tfrac{1}{4}\big(2\varepsilon_{\rm AB} - \varepsilon_{\rm AA} - \varepsilon_{\rm BB}\big) \qquad , \qquad H = \tfrac{1}{4}\big(\varepsilon_{\rm BB} - \varepsilon_{\rm AA}\big) \quad , \tag{7.45}$$

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

$$\begin{aligned} 2\varepsilon_{\rm AB} &> \varepsilon_{\rm AA} + \varepsilon_{\rm BB} \quad &\Longrightarrow \quad J > 0 \quad &\text{(ferromagnetic)} \\ 2\varepsilon_{\rm AB} &< \varepsilon_{\rm AA} + \varepsilon_{\rm BB} \quad &\Longrightarrow \quad J < 0 \quad &\text{(antiferromagnetic)} \quad . \end{aligned} \tag{7.46}$$

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,

$$\hat{H} = -J \sum_{\langle ij \rangle} \sigma_i\,\sigma_j - H \sum_i \sigma_i \quad , \tag{7.47}$$

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[11] and that the coupling $J_{ij} = J\big(|\boldsymbol{R}_i - \boldsymbol{R}_j|\big)$, where $\boldsymbol{R}_i$ is the position of the $i^{\rm th}$ spin.

On each site i we decompose σ_i into a contribution from its thermodynamic average and a fluctuation:

$$\sigma_i = \langle \sigma_i \rangle + \delta\sigma_i \quad . \tag{7.48}$$

---

[11] 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.

We will write $\langle\sigma_i\rangle \equiv m$, the local magnetization (dimensionless), and assume that m is independent of position i. Then

$$\begin{aligned}\sigma_i\,\sigma_j &= (m+\delta\sigma_i)\,(m+\delta\sigma_j)\\ &= m^2 + m\,(\delta\sigma_i+\delta\sigma_j)+\delta\sigma_i\,\delta\sigma_j \\ &= -m^2 + m\,(\sigma_i+\sigma_j)+\delta\sigma_i\,\delta\sigma_j \quad .\end{aligned} \tag{7.49}$$

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*

$$\hat{H}_{\rm MF} = \tfrac{1}{2}NzJ\,m^2 - \big(H+zJm\big)\sum_i\sigma_i \quad , \tag{7.50}$$

where N is the total number of lattice sites. 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'

$$H_{\rm eff} = H + zJm \quad . \tag{7.51}$$

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

$$m = \frac{e^{\beta H_{\rm eff}} - e^{-\beta H_{\rm eff}}}{e^{\beta H_{\rm eff}} + e^{-\beta H_{\rm eff}}} = \tanh\!\bigg(\frac{H+zJm}{k_{\rm B}T}\bigg) \quad . \tag{7.52}$$

It is a simple matter to solve for the free energy, given the noninteracting Hamiltonian $\hat{H}_{\rm MF}$. The partition function is

$$Z = \mathsf{Tr}\; e^{-\beta\hat{H}_{\rm MF}} = e^{-\frac{1}{2}\beta NzJ\,m^2}\bigg(\sum_\sigma e^{\beta(H+zJm)\sigma}\bigg)^N = e^{-\beta F} \quad . \tag{7.53}$$

We now define dimensionless variables:

$$f \equiv \frac{F}{NzJ} \quad , \quad \theta \equiv \frac{k_{\rm B}T}{zJ} \quad , \quad h \equiv \frac{H}{zJ} \quad , \tag{7.54}$$

and obtain the dimensionless free energy

$$f(m,\theta,h) = \tfrac{1}{2}m^2 - \theta\,\log\cosh\!\Big(\frac{m+h}{\theta}\Big) - \theta\log 2 \quad . \tag{7.55}$$

Differentiating with respect to m gives the mean field equation,

$$m = \tanh\!\Big(\frac{m+h}{\theta}\Big) \quad , \tag{7.56}$$

which is equivalent to the self-consistency requirement, $m = \langle\sigma_i\rangle$. In terms of the dimensionless variables θ and h, the physics is universal, and independent of details such as the magnitude of J, the value of z, and Boltzmann's constant.

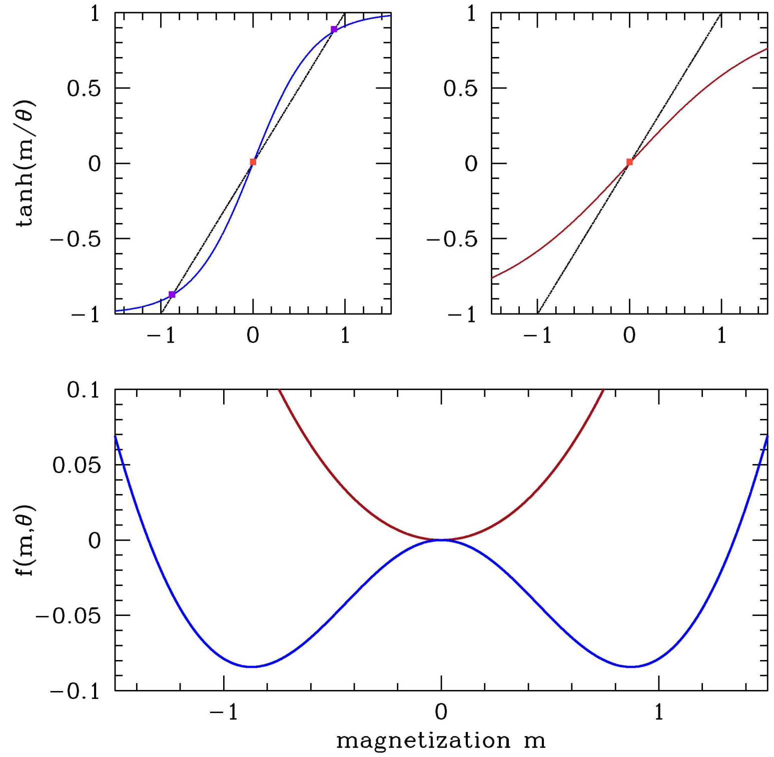


Figure 7.9: Results for $h = 0$. Upper panels: graphical solution to self-consistency equation $m = \tanh(m/\theta)$ at temperatures $\theta = 0.65$ (blue) and $\theta = 1.5$ (dark red). Lower panel: mean field free energy, with energy shifted by $\theta \log 2$ so that $f(m = 0, \theta) = 0$.

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 in a tanh function which gets steeper with decreasing dimensionless temperature θ. If, at $m = 0$, the slope of $\tanh(m/\theta)$ is smaller than unity, then the curve $y = \tanh(m/h)$ 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

$$f(m,\theta) = -\theta \log 2 + \tfrac{1}{2}m^2 - \theta\, \log\cosh\left(\frac{m}{\theta}\right) = f_0 + \tfrac{1}{2}(\theta - \theta_{\rm c})\, m^2 + \frac{m^4}{12} + \ldots \quad , \tag{7.57}$$

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 $T_{\rm c} = zJ\theta_{\rm c} = zJ$.

In the high temperature phase, $\theta > \theta_c$, 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 $\theta < \theta_c$, 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 $\theta > \theta_c$, the free energy $f(m, \theta)$ has a single minimum at $m = 0$. Below θ_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 $m^2 = 3(\theta_c - \theta) + \mathcal{O}\big((\Delta\theta)^2\big)$. Thus, we find for $|\theta - \theta_c| \ll 1$,

$$m(\theta, h = 0) = \pm\sqrt{3}\,\big(\theta_c - \theta\big)_+^{1/2} \quad , \tag{7.58}$$

where the $+$ subscript indicates that this solution is only for $\theta_c - \theta > 0$. 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 $\theta > \theta_c$ 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 $\theta \to \theta_c^-$ is denoted as β. That is, $m(\theta, h = 0) \propto (\theta_c - \theta)_+^{\beta}$ 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

$$f(\theta, h = 0) = \begin{cases} -\theta \log 2 & \text{if } \theta > \theta_c \\ -\theta \log 2 - \frac{3}{4}(\theta_c - \theta)^2 + \mathcal{O}\big((\theta_c - \theta)^4\big) & \text{if } \theta < \theta_c \quad . \end{cases} \tag{7.59}$$

Thus, if we compute the heat capacity, we find in the vicinity of $\theta = \theta_c$

$$c_V = -\theta\,\frac{\partial^2 f}{\partial \theta^2} = \begin{cases} 0 & \text{if } \theta > \theta_c \\ \frac{3}{2} & \text{if } \theta < \theta_c \quad . \end{cases} \tag{7.60}$$

Thus, the specific heat is *discontinuous* at $\theta = \theta_c$. We emphasize that our results here are only valid for $|\theta - \theta_c| \ll 1$. The general result valid for all θ (within our mean field theory) is[12]

$$c_V(\theta) = \frac{1}{\theta} \cdot \frac{m^2(\theta) - m^4(\theta)}{\theta - 1 + m^2(\theta)} \quad , \tag{7.61}$$

With this expression one can check both limits $\theta \to 0$ and $\theta \to \theta_c$. 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 θ^2. As a result, $c_V(\theta \to 0) = 0$, as expected. As $\theta \to \theta_c^-$, we have $m^2 = 3(\theta_c - \theta) + \ldots$ and we recover $c_V(\theta_c^-) = \frac{3}{2}$. In the theory of critical phenomena, $c_V(\theta) \propto |\theta - \theta_c|^{-\alpha}$ as $\theta \to \theta_c$. We see that mean field theory yields $\alpha = 0$.

---

[12] To obtain this result, one writes $f = f\big(\theta, m(\theta)\big)$ and then differentiates twice with respect to θ, 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 θ by the mean field equation.

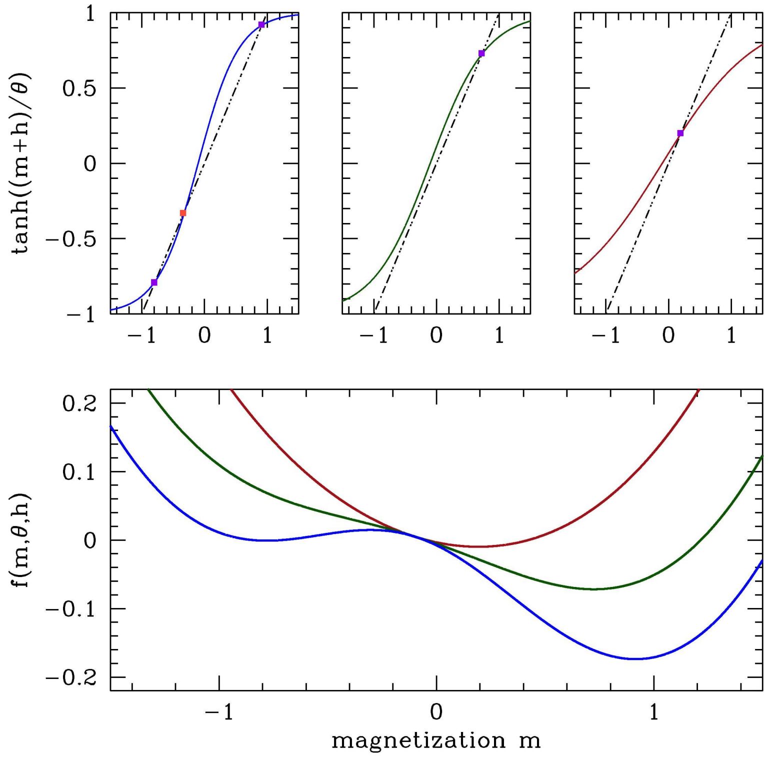


Figure 7.10: $f(m.\theta, h)$ for $h = 0.1$. Upper panels: graphical solution to the self-consistency equation $m = \tanh\big((m + h)/\theta\big)$ at temperatures $\theta = 0.65$ (blue), $\theta = 0.9$ (dark green), and $\theta = 1.5$ (dark red). Lower panel: mean field free energy, with energy shifted by $\theta \log 2$ so that $f(m = 0, \theta) = 0$.

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 $(\theta_c, h_c) = (1, 0)$. The mean field solution for $m(\theta, h)$ then be small, and we may expand the free energy from eqn. 7.55 in m and h, *viz.*

$$\begin{aligned} f(m, \theta, h) &= -\theta \, \log 2 + \tfrac{1}{2}(1 - \theta^{-1})\, m^2 + \frac{m^4}{12\, \theta^3} - \frac{hm}{\theta} \\ &= f_0 + \tfrac{1}{2}\, (\theta - \theta_c)\, m^2 + \tfrac{1}{12} m^4 - hm + \ldots \quad . \end{aligned} \tag{7.62}$$

Note that we have only gone to linear order in h. Setting $\partial f/\partial m = 0$, we obtain

$$\tfrac{1}{3} m^3 + (\theta - \theta_c) \cdot m - h = 0 \quad . \tag{7.63}$$

If $\theta > \theta_c$ then we have a solution $m = h/(\theta - \theta_c)$. The m^3 term can be ignored because it is higher order in h, and we have assumed $h \ll |\theta - \theta_c| \ll 1$. This is known as the *Curie-Weiss law*[13]. The magnetic

---

[13] 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

susceptibility behaves as

$$\chi(\theta) = \frac{\partial m}{\partial h} = \frac{1}{\theta - \theta_c} \propto |\theta - \theta_c|^{-\gamma} \quad , \tag{7.64}$$

where the magnetization critical exponent γ is $\gamma = 1$. If $\theta < \theta_c$ then while there is still a solution at $m = h/(\theta - \theta_c)$, it lies at a local maximum of the free energy, as shown in fig. 7.10. The minimum of the free energy occurs close to the $h = 0$ solution $m = m_0(\theta) \equiv \sqrt{3}\,(\theta_c - \theta)$, and writing $m = m_0 + \delta m$ we find δm to linear order in h as $\delta m(\theta, h) = h/2(\theta_c - \theta)$. Thus,

$$m(\theta, h) = \sqrt{3}\,(\theta_c - \theta)^{1/2} + \frac{h}{2(\theta_c - \theta)} \quad . \tag{7.65}$$

Once again, we find that $\chi(\theta)$ diverges as $|\theta - \theta_c|^{-\gamma}$ with $\gamma = 1$. The exponent γ on either side of the transition is the same.

Finally, we can set $\theta = \theta_c$ and examine $m(h)$. We find, from eqn. 7.63,

$$m(\theta = \theta_c, h) = (3h)^{1/3} \propto h^{1/\delta} \quad , \tag{7.66}$$

where δ is a new critical exponent. Mean field theory gives $\delta = 3$. Note that at $\theta = \theta_c = 1$ we have $m = \tanh(m + h)$, and inverting we find

$$h(m, \theta = \theta_c) = \tfrac{1}{2}\log\left(\frac{1+m}{1-m}\right) - m = \frac{m^3}{3} + \frac{m^5}{5} + \ldots \quad , \tag{7.67}$$

which is consistent with what we just found for $m(h, \theta = \theta_c)$.

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 α, β, γ, and δ 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 $d > d_u$, where d_u is the *upper critical dimension* of the theory. For the Ising model, $d_u = 4$, 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.

---

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, $\chi = C/(T - T_c)$.

Exponent	MFT	2D Ising (exact)	3D Ising (numerical)	binary liquid (3D)
α	0	0	$0.11008(1)$	$0.113(5)$
β	$1/2$	$1/8$	$0.326419(3)$	$0.316(8)$
γ	1	$7/4$	$1.237075(10)$	$1.240(7)$
δ	3	15	$4.78984(1)$	–
η	0	$1/3$	$0.036298(2)$	$0.016(7)$
ν	$1/2$	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 (α : D. Beysens and A. Bourgou, *Phys. Rev. A* **19**, 2407 (1979)), isobutyric acid and water (β : S. C. Greer, *Phys. Rev. A* **14**, 1770 (1976)), and 3-methylpentane-nitroethane (γ, η, ν : R. F. Chang *et al.*, *Phys. Rev. Lett.* **37**, 1481 (1976)).

Metastable states at $h \neq 0$

Consider the free energy $f(m,\theta,h)$ in eqn. 7.55, now for general θ not restricted to the immediate vicinity of θ_c. When $\theta < \theta_\mathrm{c}$ 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. 7.55, we have

$$\frac{\partial f}{\partial m} = m - \tanh\bigg(\frac{m+h}{\theta}\bigg) \qquad , \qquad \frac{\partial^2 f}{\partial m^2} = 1 - \theta^{-1}\mathrm{sech}^2\bigg(\frac{m+h}{\theta}\bigg) \quad . \tag{7.68}$$

Thus $h = \theta\tanh^{-1}(m) - m$, and using $\mathrm{sech}^2 x = 1 - \tanh^2 x$, we have $m^2 = 1-\theta$ and

$$h^*(\theta) = \sqrt{1-\theta} - \frac{\theta}{2}\log\Bigg(\frac{1+\sqrt{1-\theta}}{1-\sqrt{1-\theta}}\Bigg). \tag{7.69}$$

The solutions lie at $h = \pm h^*(\theta)$. For $\theta < \theta_\mathrm{c} = 1$ 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

$$\frac{1}{\Gamma}\frac{dm}{dt} = -\frac{\partial f}{\partial m} = \tanh\!\left(\frac{m+h}{\theta}\right) - m \quad , \tag{7.70}$$

where Γ 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:

$$\frac{df}{ds} = \frac{\partial f}{\partial m}\,\frac{dm}{ds} = -\left(\frac{\partial f}{\partial m}\right)^{\!2} \le 0 \quad . \tag{7.71}$$

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[14].

The phase flow for the equation $\dot{m} = -f'(m)$ is shown in fig. 7.11. As we have seen, for any value of h there is a temperature θ^* 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. 7.70, 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 θ is increased, this fixed point evolves smoothly. At $\theta = \theta^*$, the (metastable) local minimum and the local maximum coalesce and annihilate in a saddle-note bifurcation. However at $h = 0$ all three fixed points coalesce at $\theta = \theta_{\text{c}}$ and the bifurcation is a supercritical pitchfork. As a function of t 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

$$h(\theta, m) = \frac{\theta}{2}\log\!\left(\frac{1+m}{1-m}\right) - m \quad . \tag{7.72}$$

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. 7.69. Now imagine that $\theta < \theta_{\text{c}}$ 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

---

[14] Since $f(s)$ is never increasing, the extremum cannot be a local maximum.

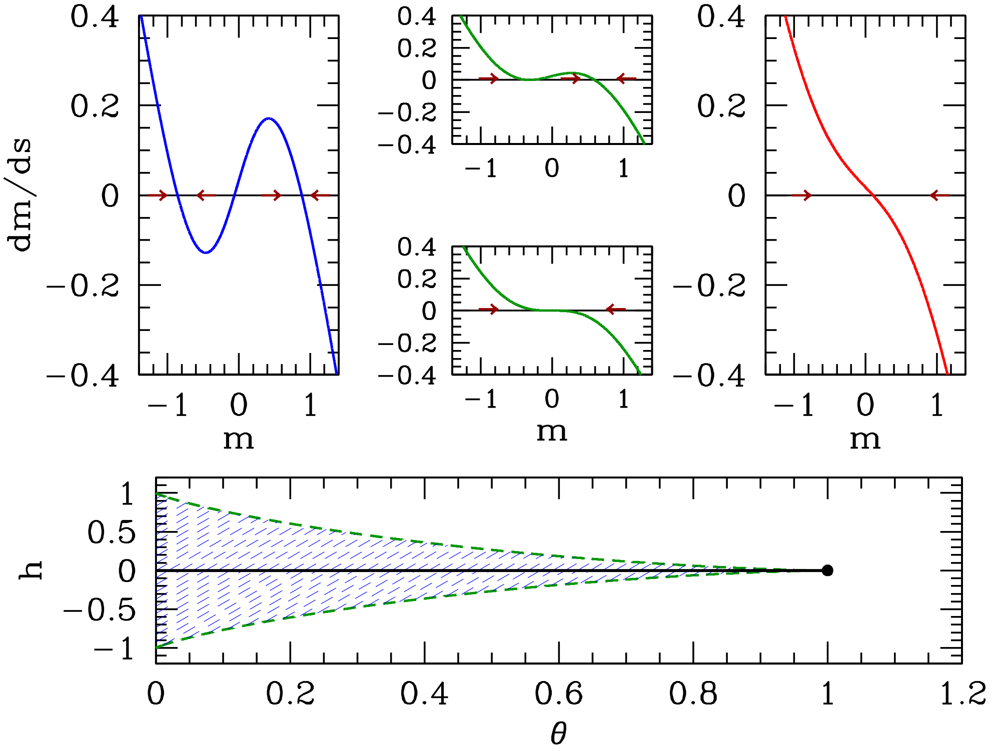


Figure 7.11: Dissipative magnetization dynamics $\dot{m} = -f'(m)$. Bottom panel shows $h^*(\theta)$ from eqn. 7.69. For (θ, h) within the blue shaded region, the free energy $f(m)$ has a global minimum plus a local minimum and a local maximum. Otherwise $f(m)$ has only a single global minimum. Top panels 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).

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 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, or changes in other externally determined parameters.

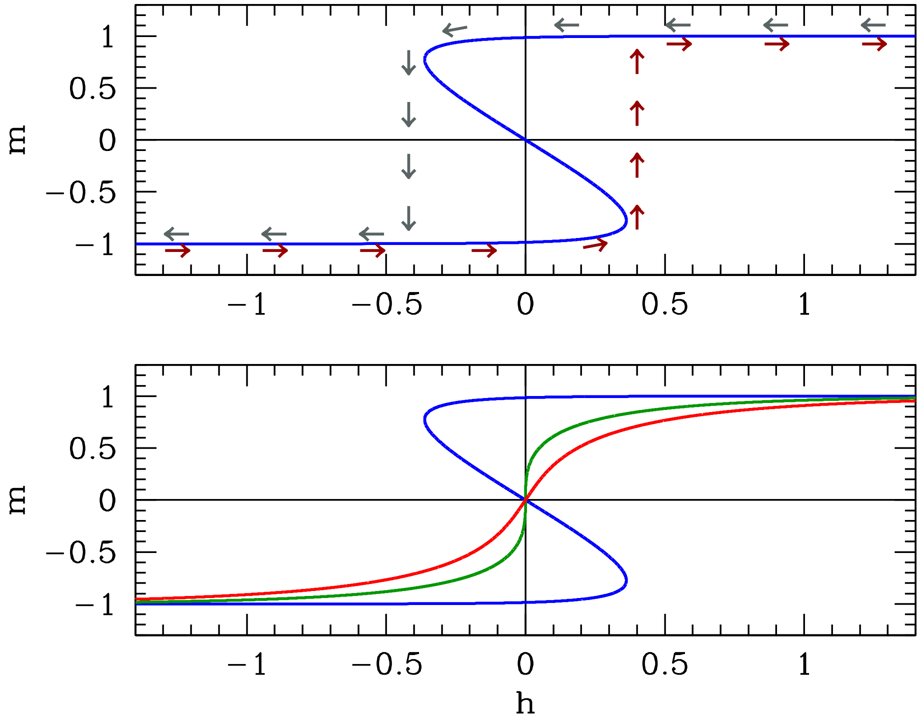


Figure 7.12: Top panel : hysteresis as a function of ramping the dimensionless magnetic field h at $\theta = 0.40$. Dark red arrows below the curve follow evolution of the magnetization on slow increase of h. Dark grey arrows above the curve follow evolution of the magnetization on slow decrease of h. 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 $t = 1.25$ (red).

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,

$$\hat{H} = -\tfrac{1}{2}\sum_{i\neq j} J_{ij}\,\sigma_i\,\sigma_j - H\sum_i \sigma_i \quad , \tag{7.73}$$

where $J_{ij} = J\big(|\boldsymbol{R}_i - \boldsymbol{R}_j|\big)$ is the coupling between spins on sites i and j[15]. 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 we include a factor of $\frac{1}{2}$.[16] The resulting mean field Hamiltonian is then

$$\hat{H}_{\text{MF}} = \tfrac{1}{2}N\hat{J}(0)\,m^2 - \big(H + \hat{J}(0)\,m\big)\sum_i \sigma_i \quad . \tag{7.74}$$

---

[15] 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.

[16] 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.

Here, $\hat{J}(\boldsymbol{q})$ is the lattice Fourier transform of the interaction function $J(\boldsymbol{R})$:[17]

$$\hat{J}(\boldsymbol{q}) = \sum_{\boldsymbol{R}} J(\boldsymbol{R})\, e^{-i\boldsymbol{q}\cdot\boldsymbol{R}} \quad . \tag{7.75}$$

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. 7.55, 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

$$\hat{H} = -\frac{J}{2N}\bigg(\sum_i \sigma_i\bigg)^2 - H\sum_k \sigma_k \quad . \tag{7.76}$$

The partition function is

$$Z = \mathop{\mathsf{Tr}}_{\{\sigma_i\}} \exp\left[\frac{\beta J}{2N}\bigg(\sum_i \sigma_i\bigg)^2 + \beta H \sum_i \sigma_i\right] \quad . \tag{7.77}$$

We now invoke the Gaussian integral,

$$\int\limits_{-\infty}^{\infty} dx\; e^{-\alpha x^2 - \beta x} = \sqrt{\frac{\pi}{\alpha}}\; e^{\beta^2/4\alpha} \quad . \tag{7.78}$$

Thus,

$$\exp\left[\frac{\beta J}{2N}\bigg(\sum_i \sigma_i\bigg)^2\right] = \bigg(\frac{N\beta J}{2\pi}\bigg)^{1/2} \int\limits_{-\infty}^{\infty} dm\; e^{-\frac{1}{2}N\beta J m^2 + \beta J m \sum_i \sigma_i} \quad , \tag{7.79}$$

and we can write the partition function as

$$Z = \bigg(\frac{N\beta J}{2\pi}\bigg)^{1/2} \int\limits_{-\infty}^{\infty} dm\; e^{-\frac{1}{2}N\beta J m^2} \bigg(\sum_\sigma e^{\beta(H+Jm)\sigma}\bigg)^N = \bigg(\frac{N}{2\pi\theta}\bigg)^{1/2} \int\limits_{-\infty}^{\infty} dm\; e^{-NA(m)/\theta} \quad , \tag{7.80}$$

where $\theta = k_{\rm B}T/J$, $h = H/J$, and

$$A(m) = \tfrac{1}{2}m^2 - \theta \log\left[2\cosh\bigg(\frac{h+m}{\theta}\bigg)\right] \quad . \tag{7.81}$$

---

[17] 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.

Since $N \to \infty$, we can perform the integral using the method of steepest descents. Thus, we must set

$$\frac{dA}{dm}\bigg|_{m^*} = 0 \quad \Longrightarrow \quad m^* = \tanh\left(\frac{m^* + h}{\theta}\right) \quad . \tag{7.82}$$

Expanding about $m = m^*$, we write

$$A(m) = A(m^*) + \tfrac{1}{2}A''(m^*)\,(m - m^*)^2 + \tfrac{1}{6}\,A'''(m^*)\,(m - m^*)^3 + \ldots \quad . \tag{7.83}$$

Writing $\nu = m - m^*$ and performing the integration, we obtain

$$\begin{aligned} Z &= \left(\frac{N}{2\pi\theta}\right)^{1/2} e^{-NA(m^*)/\theta} \int\limits_{-\infty}^{\infty} d\nu\, \exp\left[-\frac{NA''(m^*)}{2\theta}\,\nu^2 - \frac{NA'''(m^*)}{6\theta}\,\nu^3 + \ldots\right] \\ &= \frac{1}{\sqrt{A''(m^*)}}\, e^{-NA(m^*)/\theta} \cdot \Big\{1 + \mathcal{O}(N^{-1})\Big\} \quad . \end{aligned} \tag{7.84}$$

The corresponding free energy per site

$$f = \frac{F}{NJ} = A(m^*) + \frac{\theta}{2N}\,\log A''(m^*) + \mathcal{O}(N^{-2}) \quad , \tag{7.85}$$

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 M of a ferromagnet in zero external field but at finite temperature typically vanishes for temperatures $T > T_c$, where T_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.

7.5.1 Quartic free energy with Ising symmetry

The simplest example is the quartic free energy,

$$f(m, \theta, h = 0) = f_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 \quad , \tag{7.86}$$

where $f_0 = f_0(\theta)$, $a = a(\theta)$, and $b = b(\theta)$. Here, θ is a dimensionless measure of the temperature[18]. We assume $b > 0$, which is necessary if the free energy is to be bounded from below[19].

---

[18] For example in an Ising ferromagnet we might define $\theta = k_B T/\hat{J}(0)$, as before.

[19] 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.

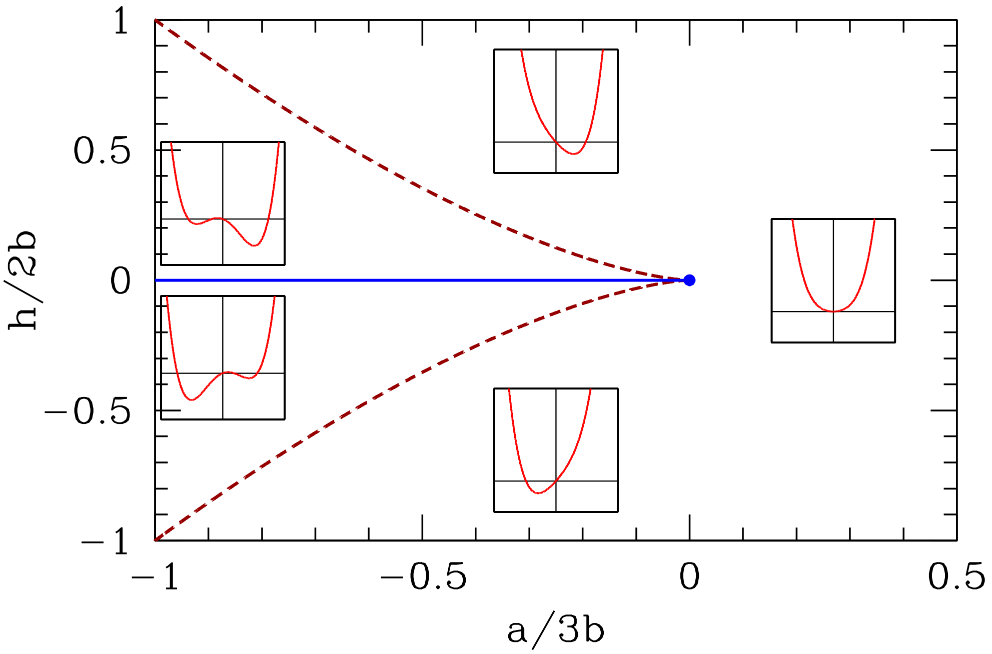


Figure 7.13: Phase diagram for the quartic Landau free energy $f = f_0 + \frac{1}{2}am^2 + \frac{1}{4}bm^4 - hm$, with $b > 0$. There is a first order line at $h = 0$ extending from $a = -\infty$ and terminating in a critical point at $a = 0$. For $|h| < h^*(a)$ (dashed red line) there are three solutions to the mean field equation, corresponding to one global minimum, one local minimum, and one local maximum. Insets show behavior of the free energy $f(m)$.

The equation of state, which relates the intensive quantities m and θ at $h = 0$, is then

$$\frac{\partial f}{\partial m} = 0 = am + bm^3 \quad , \tag{7.87}$$

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 θ_c where $a(\theta_c) = 0$. Minimizing f, we find

$$\begin{aligned} a < 0 \quad (\theta < \theta_c) \quad &: \quad f(\theta) = f_0 - \frac{a^2}{4b} \\ a > 0 \quad (\theta > \theta_c) \quad &: \quad f(\theta) = f_0 \quad . \end{aligned} \tag{7.88}$$

Thus $a(\theta)$ changes sign at $\theta = \theta_c$, where the free energy is continuous, since $a(\theta_c) = 0$. The specific heat, however, is discontinuous across the transition, with

$$c(\theta_c^+) - c(\theta_c^-) = -\theta_c \frac{\partial^2}{\partial\theta^2}\bigg|_{\theta=\theta_c} \left(\frac{a^2}{4b}\right) = -\frac{\theta_c \big[a'(\theta_c)\big]^2}{2b(\theta_c)} \quad . \tag{7.89}$$

The presence of a magnetic field h breaks the $\mathbb{Z}_2$ symmetry of $m \to -m$. The free energy becomes

$$f(m,\theta,h) = f_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 - hm \quad , \tag{7.90}$$

and the mean field equation is

$$bm^3 + am - h = 0 \quad . \tag{7.91}$$

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

$$h^*(a) = \pm \frac{2}{3^{3/2}} \, \frac{(-a)^{3/2}}{b^{1/2}} \quad , \tag{7.92}$$

or, equivalently,

$$a^*(h) = -\frac{3}{2^{2/3}} \, b^{1/3} \, |h|^{2/3}. \tag{7.93}$$

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

$$\begin{aligned} m(a,h) &= \frac{h}{a} - \frac{b}{a^4} \, h^3 + \mathcal{O}(h^5) & (a > 0) \\ &= \pm \frac{|a|^{1/2}}{b^{1/2}} + \frac{h}{2\,|a|} \pm \frac{3\,b^{1/2}}{8\,|a|^{5/2}} \, h^2 + \mathcal{O}(h^3) & (a < 0) \quad . \end{aligned} \tag{7.94}$$

7.5.2 Cubic terms in Landau theory : first order transitions

Next, consider a free energy with a cubic term,

$$f = f_0 + \tfrac{1}{2} a m^2 - \tfrac{1}{3} y m^3 + \tfrac{1}{4} b m^4 \quad , \tag{7.95}$$

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

$$\frac{\partial f}{\partial m} = 0 = am - ym^2 + bm^3 \quad . \tag{7.96}$$

This cubic equation factorizes into a linear and quadratic piece, and hence may be solved simply. The three solutions are $m = 0$ and

$$m = m_{\pm} \equiv \frac{y}{2b} \pm \sqrt{\left(\frac{y}{2b}\right)^2 - \frac{a}{b}} \quad . \tag{7.97}$$

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

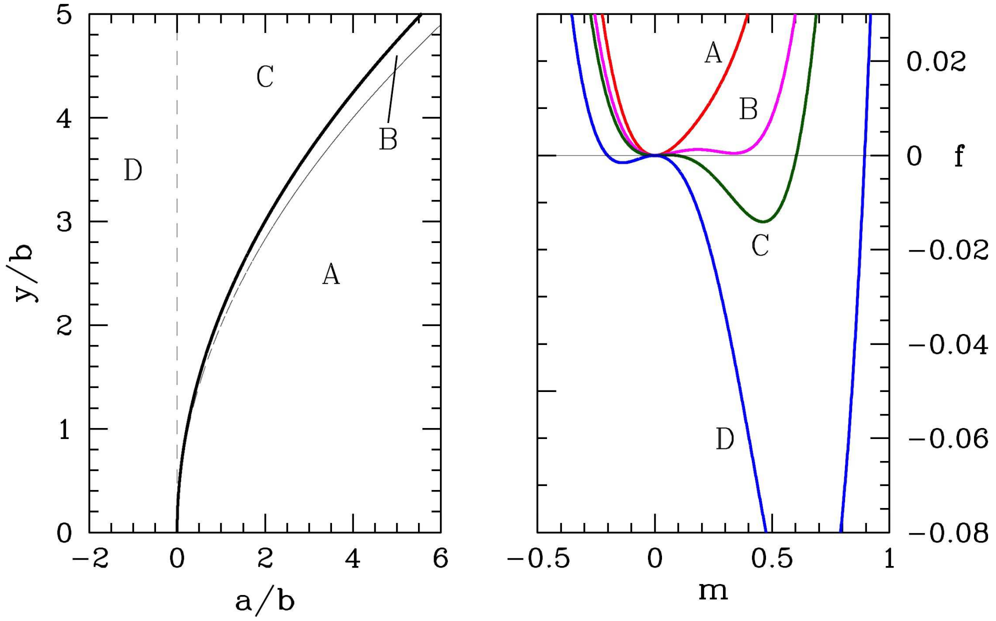


Figure 7.14: Behavior of the quartic free energy $f(m) = \frac{1}{2}am^2 - \frac{1}{3}ym^3 + \frac{1}{4}bm^4$. A: $y^2 < 4ab$; B: $4ab < y^2 < \frac{9}{2}ab$; C and D: $y^2 > \frac{9}{2}ab$. The thick black line denotes a line of first order transitions, where the order parameter is discontinuous across the transition.

$f(m_+)$[20]. Thus, we set

$$f(m) = f(0) \quad \Longrightarrow \quad \tfrac{1}{2}am^2 - \tfrac{1}{3}ym^3 + \tfrac{1}{4}bm^4 = 0 \quad , \tag{7.98}$$

and we now have two quadratic equations to solve simultaneously:

$$0 = a - ym + bm^2 \qquad , \qquad 0 = \tfrac{1}{2}a - \tfrac{1}{3}ym + \tfrac{1}{4}bm^2 = 0 \quad . \tag{7.99}$$

Eliminating the quadratic term gives $m = 3a/y$. Finally, substituting $m = m_+$ gives us a relation between a, b, and y:

$$y^2 = \tfrac{9}{2}\,ab \quad . \tag{7.100}$$

Thus, we have the following:

$$\begin{array}{rcl} a > \dfrac{y^2}{4b} & : & \text{1 real root } m = 0 \\ \dfrac{y^2}{4b} > a > \dfrac{2y^2}{9b} & : & \text{3 real roots; minimum at } m = 0 \\ \dfrac{2y^2}{9b} > a & : & \text{3 real roots; minimum at } m = \dfrac{y}{2b} + \sqrt{\left(\dfrac{y}{2b}\right)^2 - \dfrac{a}{b}} \end{array} \tag{7.101}$$

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

---

[20] We needn't waste our time considering the $m = m_-$ solution, since the cubic term prefers positive m.

at $m = m_+$, and a local maximum at $m = m_-$, with $m_+ > m_- > 0$. For $y^2/9b > a > 0$, there is a local minimum at $a = 0$, 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 $a_c = 2y^2/9b > 0$ and $m_c = 2y/3b$. 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 $T_c = T_0 + 2y^2/9\alpha b$.

7.5.3 Order parameter dynamics

Suppose we now impose some dynamics on the system, of the simple relaxational type

$$\frac{dm}{dt} = -\Gamma \, \frac{\partial f}{\partial m} \quad , \tag{7.102}$$

where Γ is a phenomenological kinetic coefficient. Assuming $y > 0$ and $b > 0$, it is convenient to adimensionalize by writing

$$m \equiv \frac{y}{b} \cdot u \qquad , \qquad a \equiv \frac{y^2}{b} \cdot r \qquad , \qquad t \equiv \frac{b}{\Gamma y^2} \cdot s \quad . \tag{7.103}$$

Then we obtain

$$\frac{du}{ds} = -\frac{\partial \varphi}{\partial u} \quad , \tag{7.104}$$

where the dimensionless free energy function is

$$\varphi(u) = \tfrac{1}{2} r u^2 - \tfrac{1}{3} u^3 + \tfrac{1}{4} u^4 \quad . \tag{7.105}$$

We see that there is a single control parameter, r. 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

$$u^* = 0 \qquad , \qquad u^* = \tfrac{1}{2} \pm \big(\tfrac{1}{4} - r\big)^{1/2} \quad . \tag{7.106}$$

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. 7.104, $u(s)$ 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

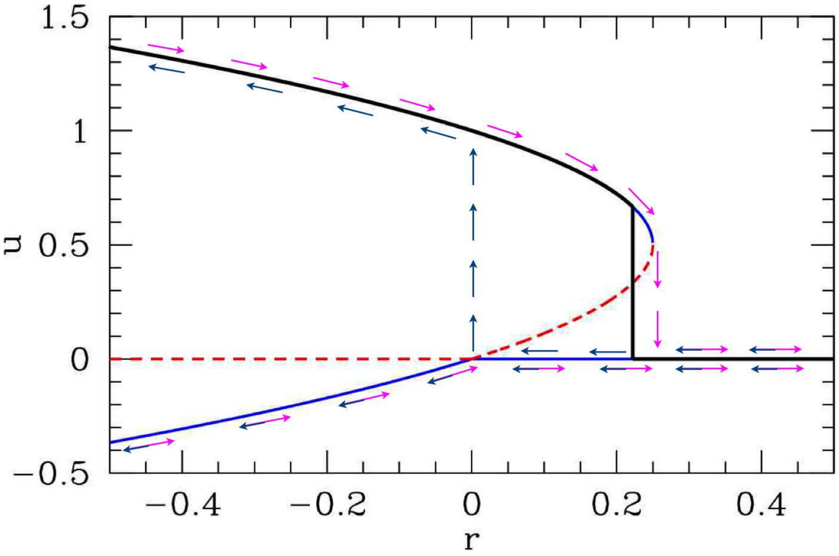


Figure 7.15: Fixed points for $\varphi(u) = \frac{1}{2}ru^2 - \frac{1}{3}u^3 + \frac{1}{4}u^4$ and flow under the dynamics $\dot{u} = -\varphi'(u)$. Solid curves represent stable fixed points and dashed curves unstable fixed points. Magenta arrows show behavior under slowly increasing control parameter 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 $u = \frac{2}{3}$.

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 $u_+ = \frac{1}{2} + \big(\frac{1}{4} - r\big)^{1/2}$. 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 $u_- = \frac{1}{2} - \big(\frac{1}{4} - r\big)^{1/2}$, 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

$$f = f_0 + \tfrac{1}{2}am^2 + \tfrac{1}{4}bm^4 + \tfrac{1}{6}cm^6 \quad , \tag{7.107}$$

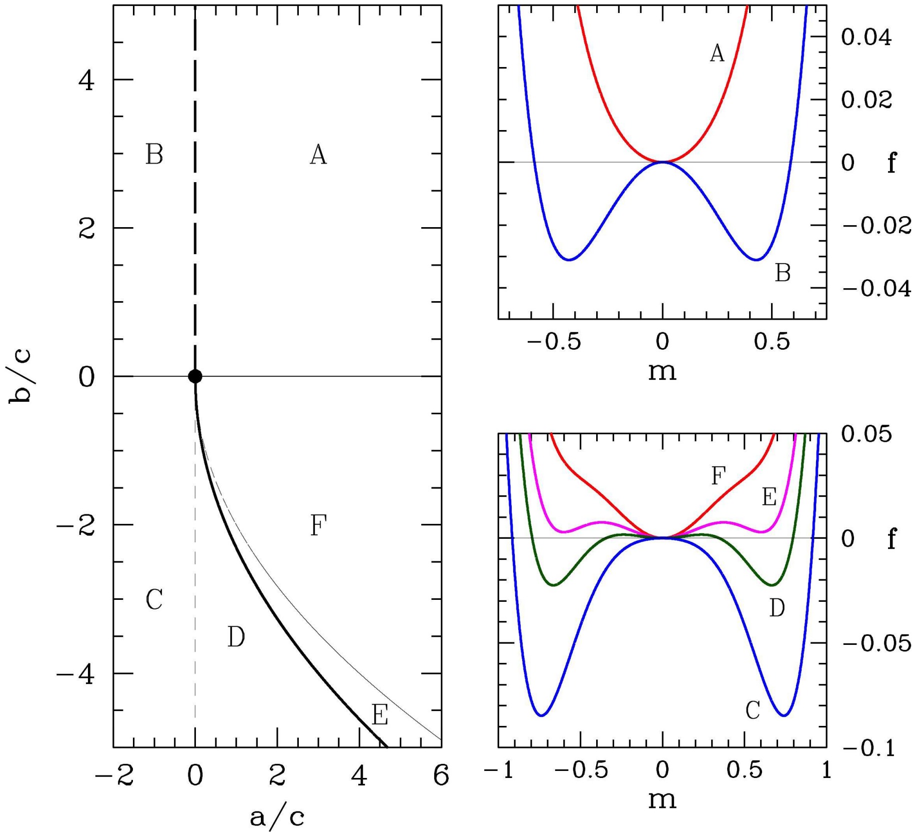


Figure 7.16: Behavior of the sextic free energy $f(m) = \frac{1}{2}am^2 + \frac{1}{4}bm^4 + \frac{1}{6}cm^6$. A: $a > 0$ and $b > 0$; B: $a < 0$ and $b > 0$; C: $a < 0$ and $b < 0$; D: $a > 0$ and $b < -\frac{4}{\sqrt{3}}\sqrt{ac}$; E: $a > 0$ and $-\frac{4}{\sqrt{3}}\sqrt{ac} < b < -2\sqrt{ac}$; F: $a > 0$ 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, $(a, b) = (0, 0)$.

with $c > 0$ for stability. We seek the phase diagram in the (a, b) plane. Extremizing f with respect to m,

$$\frac{\partial f}{\partial m} = 0 = m\,(a + bm^2 + cm^4) \quad , \tag{7.108}$$

which is a quintic with five solutions over the complex m plane. One solution is obviously $m = 0$. The other four are

$$m = \pm\sqrt{-\frac{b}{2c} \pm \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a}{c}}} \quad . \tag{7.109}$$

For each ± 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, since the $-$ choice for the $\pm$ sign under

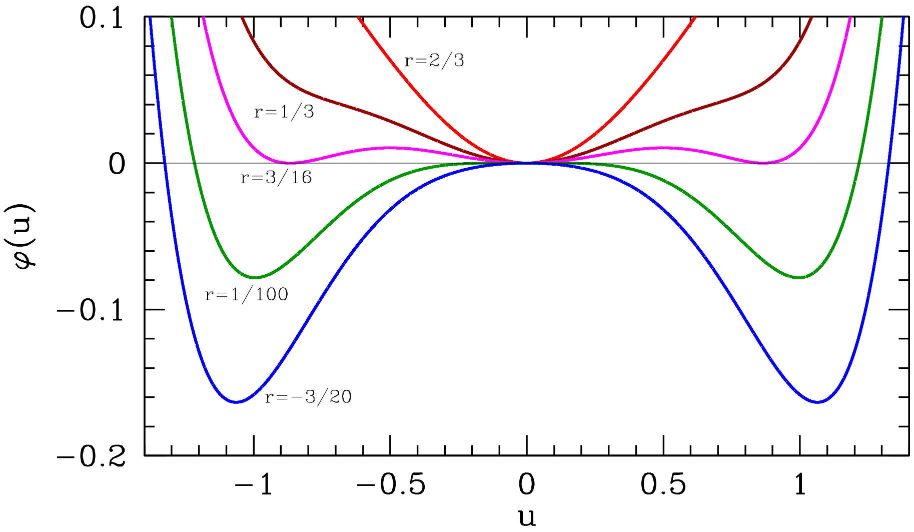


Figure 7.17: Free energy $\varphi(u) = \frac{1}{2}ru^2 - \frac{1}{4}u^4 + \frac{1}{6}u^6$ for different values of the control parameter r.

the radical leads to imaginary roots. One of the solutions is $m = 0$. The other two are

$$m = \pm\sqrt{-\frac{b}{2c} + \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a}{c}}} \quad . \tag{7.110}$$

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 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,

$$0 = \tfrac{1}{2}a + \tfrac{1}{4}bm^2 + \tfrac{1}{6}cm^4 \quad , \tag{7.111}$$

which together with equation 7.108 gives $b = -\frac{4}{\sqrt{3}}\sqrt{ac}$. Thus, we have the following, for fixed $a > 0$:

$$\begin{aligned} b > -2\sqrt{ac} \quad &: \quad \text{1 real root } m = 0 \\ -2\sqrt{ac} > b > -\tfrac{4}{\sqrt{3}}\sqrt{ac} \quad &: \quad \text{5 real roots; minimum at } m = 0 \\ -\tfrac{4}{\sqrt{3}}\sqrt{ac} > b \quad &: \quad \text{5 real roots; minima at } m = \pm\sqrt{-\frac{b}{2c} + \sqrt{\left(\frac{b}{2c}\right)^2 - \frac{a}{c}}} \end{aligned} \tag{7.112}$$

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*.

7.5.5 Hysteresis for the sextic potential

Once again, we consider the dissipative dynamics $\dot{m} = -\Gamma f'(m)$. We adimensionalize by writing

$$m \equiv \sqrt{\frac{|b|}{c}} \cdot u \qquad , \qquad a \equiv \frac{b^2}{c} \cdot r \qquad , \qquad t \equiv \frac{c}{\Gamma b^2} \cdot s \quad . \tag{7.113}$$

Then we obtain once again the dimensionless equation $du/ds = -\partial\varphi/\partial u$, where

$$\varphi(u) = \tfrac{1}{2} r u^2 \pm \tfrac{1}{4} u^4 + \tfrac{1}{6} u^6 \quad . \tag{7.114}$$

In the above equation, the coefficient of the quartic term is positive if $b > 0$ and negative if $b < 0$. That is, the coefficient is $\mathsf{sgn}(b)$. 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

$$\varphi(u) = \tfrac{1}{2} r u^2 - \tfrac{1}{4} u^4 + \tfrac{1}{6} u^6 \qquad , \qquad \varphi'(u) = u\,(r - u^2 + u^4) \quad . \tag{7.115}$$

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:

$$\begin{aligned} r < 0 \quad &: \quad u^* = 0 \quad , \; u^* = \pm\sqrt{\tfrac{1}{2} + \sqrt{\tfrac{1}{4} - r}} \\ 0 < r < \tfrac{1}{4} \quad &: \quad u^* = 0 \quad , \; u^* = \pm\sqrt{\tfrac{1}{2} + \sqrt{\tfrac{1}{4} - r}} \quad , \; u^* = \pm\sqrt{\tfrac{1}{2} - \sqrt{\tfrac{1}{4} - r}} \\ r > \tfrac{1}{4} \quad &: \quad u^* = 0 \quad . \end{aligned} \tag{7.116}$$

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

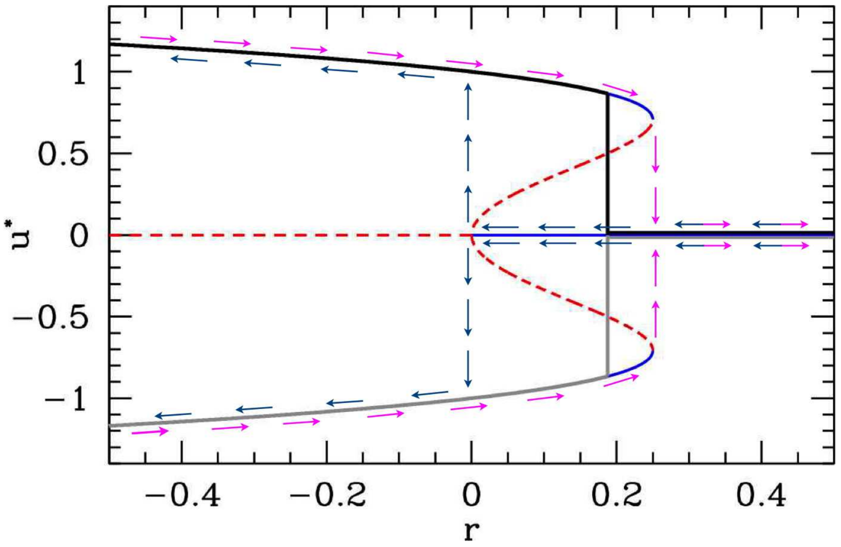


Figure 7.18: Fixed points $\varphi'(u^*) = 0$ for the sextic potential $\varphi(u) = \frac{1}{2}ru^2 - \frac{1}{4}u^4 + \frac{1}{6}u^6$, and dynamical flow (arrows) under $\dot{u} = -\varphi'(u)$. Solid curves show stable fixed points and dashed curves show unstable fixed points. The thick solid black and solid grey curves indicate the equilibrium thermodynamic values for u; 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 $r = \frac{3}{16}$.

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 $u^*(r) = \big(\frac{1}{2} + (\frac{1}{4} - r)^{1/2}\big)^{1/2}$ for all $r < 0$.

Now suppose we start increasing r (*i.e.* increasing temperature). The magnetization follows the stable fixed point $u^*(r) = \big(\frac{1}{2} + (\frac{1}{4} - r)^{1/2}\big)^{1/2}$ 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:

$$F = E - TS = \mathrm{Tr}\,(\varrho\,\hat{H}) + k_{\mathrm{B}}T\,\mathrm{Tr}\,(\varrho\,\log\varrho) \quad . \tag{7.117}$$

Here, ϱ is the *density matrix*[21]. 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 ϱ must be non-negative).

Our goal is to extremize the free energy subject to the various constraints on ϱ. Let us assume that ϱ is diagonal in the basis of eigenstates of $\hat{H}$, *i.e.*

$$\varrho = \sum_\gamma P_\gamma \,|\,\gamma\,\rangle\langle\,\gamma\,| \quad , \tag{7.118}$$

where P_γ is the probability that the system is in state $|\,\gamma\,\rangle$. Then

$$F = \sum_\gamma E_\gamma\, P_\gamma + k_{\mathrm{B}}T \sum_\gamma P_\gamma \log P_\gamma \quad . \tag{7.119}$$

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

$$F^*\big(\{P_\gamma\},\lambda\big) = F\big(\{P_\gamma\}\big) + \lambda\Big(\sum_\gamma P_\gamma - 1\Big) \quad , \tag{7.120}$$

and then freely extremize over both the probabilities $\{P_\gamma\}$ as well as the Lagrange multiplier λ. This yields the Boltzmann distribution,

$$P_\gamma^{\mathrm{eq}} = \frac{1}{Z}\,\exp(-E_\gamma/k_{\mathrm{B}}T) \quad , \tag{7.121}$$

where $Z = \sum_\gamma e^{-E_\gamma/k_{\mathrm{B}}T} = \mathrm{Tr}\, e^{-\hat{H}/k_{\mathrm{B}}T}$ is the canonical partition function, which is related to λ through

$$\lambda = k_{\mathrm{B}}T\,(\log Z - 1) \quad . \tag{7.122}$$

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 ϱ which depends on a set of variational parameters $\{x_\alpha\}$, and to minimize F with respect to this set.

---

[21] 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.

7.6.2 Variational density matrix for the Ising model

Consider once again the Ising model Hamiltonian,

$$\hat{H} = -\sum_{i<j} J_{ij}\,\sigma_i\,\sigma_j - H\sum_i \sigma_i \quad . \tag{7.123}$$

The states of the system $|\,\boldsymbol{\sigma}\,\rangle$ may be labeled by the values of the spin variables: $|\,\boldsymbol{\sigma}\,\rangle \longleftrightarrow |\,\sigma_1, \sigma_2, \ldots\rangle$. We assume the density matrix is diagonal in this basis, *i.e.*

$$\varrho_N\big(\boldsymbol{\sigma}\,\big|\,\boldsymbol{\sigma}'\big) \equiv \varrho_N(\boldsymbol{\sigma})\,\delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} \quad , \tag{7.124}$$

where $\delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} = \prod_i \delta_{\sigma_i,\sigma'_i}$ 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,

$$\varrho_N(\sigma_1, \sigma_2, \ldots) = \frac{1}{Z}\, e^{-\beta \hat{H}(\sigma_1,\ldots,\sigma_N)} \quad . \tag{7.125}$$

We now write a *trial density matrix* which is a product over contributions from independent single sites:

$$\varrho_N(\sigma_1, \sigma_2, \ldots) = \prod_i \varrho(\sigma_i) \quad , \tag{7.126}$$

where

$$\varrho(\sigma) = \Big(\frac{1+m}{2}\Big)\,\delta_{\sigma,1} + \Big(\frac{1-m}{2}\Big)\,\delta_{\sigma,-1} \quad . \tag{7.127}$$

Note that we've changed our notation slightly. We are denoting by $\varrho(\sigma)$ the corresponding diagonal element of the matrix

$$\varrho = \begin{pmatrix} \frac{1+m}{2} & 0 \\ 0 & \frac{1-m}{2} \end{pmatrix} \quad , \tag{7.128}$$

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 ϱ and ϱ_N are normalized with unit trace.

Note that ϱ and hence ϱ_N are appropriately normalized. The variational parameter here is m, which, if ρ 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

$$\langle \sigma_i \rangle = \sum_\sigma \varrho(\sigma)\,\sigma = m. \tag{7.129}$$

We may now evaluate the average energy:

$$\begin{aligned} E = \mathsf{Tr}\,(\varrho_N \hat{H}) &= -\sum_{i<j} J_{ij}\, m^2 - H \sum_i m \\ &= -\tfrac{1}{2} N \hat{J}(0)\, m^2 - N H m \quad , \end{aligned} \tag{7.130}$$

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

$$\begin{aligned} S &= -k_{\rm B}\,\mathsf{Tr}\,(\varrho_N\log\varrho_N) = -Nk_{\rm B}\,\mathsf{Tr}\,(\varrho\log\varrho) \\ &= -Nk_{\rm B}\bigg\{\Big(\frac{1+m}{2}\Big)\log\Big(\frac{1+m}{2}\Big)+\Big(\frac{1-m}{2}\Big)\log\Big(\frac{1-m}{2}\Big)\bigg\}\quad . \end{aligned} \tag{7.131}$$

We now define the dimensionless free energy per site: $f\equiv F/N\hat{J}(0)$. We have

$$f(m,\theta,h) = -\tfrac{1}{2}\,m^2 - hm + \theta\,\bigg\{\Big(\frac{1+m}{2}\Big)\log\Big(\frac{1+m}{2}\Big)+\Big(\frac{1-m}{2}\Big)\log\Big(\frac{1-m}{2}\Big)\bigg\}\quad , \tag{7.132}$$

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

$$\frac{\partial f}{\partial m} = 0 = -m - h + \frac{\theta}{2}\log\Big(\frac{1+m}{1-m}\Big)\quad . \tag{7.133}$$

Solving for m, we obtain

$$m = \tanh\left(\frac{m+h}{\theta}\right)\quad , \tag{7.134}$$

which is precisely what we found in eqn. 7.56. There is a second order phase transition at a critical temperature $\theta_{\rm c}=1$.

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

$$m(h,\theta) = \frac{h}{\theta-\theta_{\rm c}}\quad . \tag{7.135}$$

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 χ^{-1} *versus* θ and extrapolating to obtain the θ-intercept. In our case, $\chi(\theta)=(\theta-\theta_{\rm c})^{-1}$. For low θ 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

$$f(m,\theta,h) = -\theta\,\log 2 - hm + \tfrac{1}{2}\,(\theta-1)\,m^2 + \tfrac{\theta}{12}\,m^4 + \tfrac{\theta}{30}\,m^6 + \tfrac{\theta}{56}\,m^8 + \ldots\quad . \tag{7.136}$$

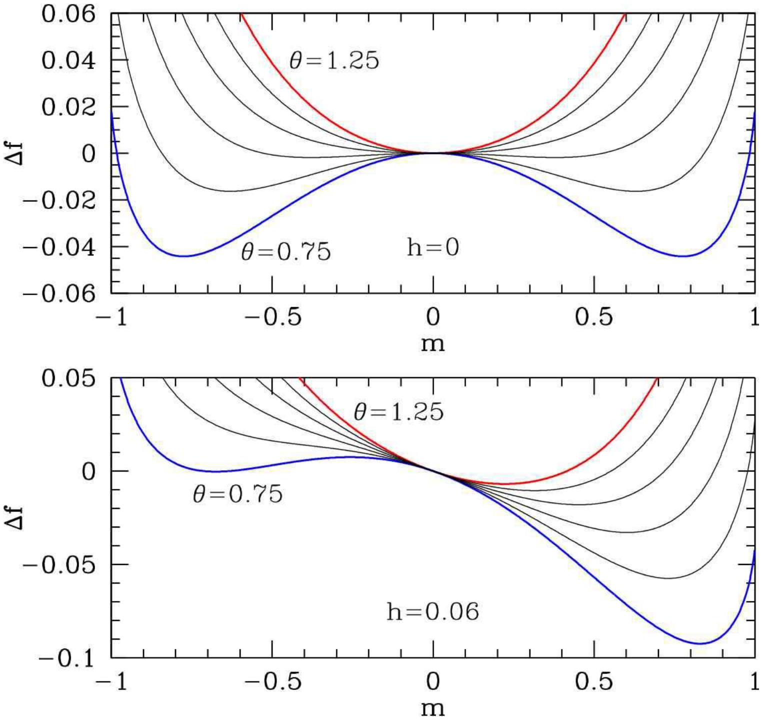


Figure 7.19: Variational field free energy $\Delta f = f(m, \theta, h) + \theta \, \log 2$ *versus* magnetization m at six equally spaced temperatures interpolating between 'high' ($\theta = 1.25$, red) and 'low' ($\theta = 0.75$, blue) values. Top panel: $h = 0$. Bottom panel: $h = 0.06$.

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 ϕ 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 ϕ is continuous across θ_c, the phase transition is said to be *second order.* When ϕ 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

$$\hat{H} = -\sum_{i<j} J_{ij}\, \delta_{\sigma_i,\sigma_j} - H \sum_i \delta_{\sigma_i,1} \quad . \tag{7.137}$$

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 integers from 1 to q, but it could be anything, such as

$$\sigma_i \in \{\text{tomato, penny, ostrich, Grateful Dead ticket from 1987, } \ldots\} \quad . \tag{7.138}$$

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 σ be ± 1. Then the quantity

$$\delta_{\sigma,\sigma'} = \tfrac{1}{2} + \tfrac{1}{2}\,\sigma\sigma' \tag{7.139}$$

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

$$\delta_{\sigma,\sigma'} = 1 + \tfrac{1}{2}\,\sigma\sigma' + \tfrac{3}{2}\,\sigma^2\sigma'^2 - (\sigma^2 + \sigma'^2) \quad . \tag{7.140}$$

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, σ^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 ϱ as a diagonal matrix with entries

$$\varrho(\sigma) = x\,\delta_{\sigma,1} + \left(\frac{1-x}{q-1}\right)\big(1 - \delta_{\sigma,1}\big) \quad , \tag{7.141}$$

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:

$$\begin{aligned} E = \mathsf{Tr}\,(\varrho_N \hat{H}) &= -\tfrac{1}{2}N\hat{J}(0)\bigg\{x^2 + \frac{(1-x)^2}{q-1}\bigg\} - NHx \\ S = -k_\mathrm{B}\,\mathsf{Tr}\,(\varrho_N \log \varrho_N) &= -Nk_\mathrm{B}\bigg\{x\log x + (1-x)\log\left(\frac{1-x}{q-1}\right)\bigg\} \quad . \end{aligned} \tag{7.142}$$

The dimensionless free energy per site is then

$$f(x,\theta,h) = -\tfrac{1}{2}\left\{x^2 + \frac{(1-x)^2}{q-1}\right\} + \theta\left\{x\log x + (1-x)\log\left(\frac{1-x}{q-1}\right)\right\} - hx \quad , \tag{7.143}$$

where $\theta = k_\mathrm{B}T/\hat{J}(0)$ and $h = H/\hat{J}(0)$. We now extremize with respect to x to obtain the mean field equation,

$$\frac{\partial f}{\partial x} = 0 = -x + \frac{1-x}{q-1} + \theta\,\log x - \theta\,\log\left(\frac{1-x}{q-1}\right) - h \quad . \tag{7.144}$$

Note that for $h = 0$, $x = q^{-1}$ is a solution, corresponding to a 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

$$h = \frac{q\,(q\theta - 1)}{q-1}\,u + \frac{q^3\,(q-2)\,\theta}{2\,(q-1)^2}\,u^2 + \mathcal{O}(u^3) \quad . \tag{7.145}$$

For weak fields, $|h| \ll 1$, and in the high temperature phase we have

$$u(\theta) = \frac{(q-1)\,h}{q\,(q\theta-1)} + \mathcal{O}(h^2) \quad , \tag{7.146}$$

which again is of the Curie-Weiss form, with a Curie temperature $\theta_\mathrm{C} = q^{-1}$. 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

$$f(u,\theta,h) = f_0 - \frac{qu^2}{2(q-1)} - hu + \theta\,(q^{-1}+u)\log(1+qu) - \theta\,(1-q^{-1}-u)\log\bigg(1 - \frac{qu}{q-1}\bigg) \quad , \tag{7.147}$$

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

$$\begin{aligned} f(u,\theta,h) = f_0 - hu &+ \frac{q^2\,(\theta - q^{-1})}{2\,(q-1)}\,u^2 - \frac{(q-2)\,q^3\,\theta}{6\,(q-1)^2}\,u^3 \\ &+ \frac{q^3\theta}{12}\Big[1 + (q-1)^{-3}\Big]u^4 - \frac{q^4\theta}{20}\Big[1 - (q-1)^{-4}\Big]u^5 + \frac{q^5\theta}{30}\Big[1 + (q-1)^{-5}\Big]u^6 + \ldots \quad . \end{aligned} \tag{7.148}$$

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 $q > 3$, 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. In Landau theory, the free energy expansion is truncated, and with $f(m) = \frac{1}{2}am^2 - \frac{1}{3}ym^3 + \frac{1}{4}bm^4$ we found that the first order transition occurs for $a = 2y^2/9b$. For the Potts model, the mean field critical point for $q > 2$ 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 θ_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:

$$\theta_c = \frac{q-2}{2(q-1)\log(q-1)} \qquad , \qquad u_c = u(\theta_c^-) = 1 - \frac{2}{q} \quad . \tag{7.149}$$

For $q = 2$ our Potts model mean field theory predicts a second order transition at $\theta_c = \frac{1}{2}$.[22]

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 $\phi_i \in [-\pi, \pi]$:

$$\hat{H} = -\frac{1}{2}\sum_{i \neq j} J_{ij} \cos\big(\phi_i - \phi_j\big) - H \sum_i \cos\phi_i \quad . \tag{7.150}$$

We write the (diagonal elements of the) full density matrix once again as a product:

$$\varrho_N(\phi_1, \phi_2, \ldots) = \prod_i \varrho(\phi_i) \quad . \tag{7.151}$$

Our goal will be to extremize the free energy with respect to the *function* $\varrho(\phi)$. To this end, we compute

$$E = \mathsf{Tr}\,(\varrho_N\,\hat{H}) = -\tfrac{1}{2} N \hat{J}(0) \left|\mathsf{Tr}\left(\varrho\, e^{i\phi}\right)\right|^2 - NH\,\mathsf{Tr}\big(\varrho\,\cos\phi\big) \quad . \tag{7.152}$$

The entropy is $S = -Nk_B\,\mathsf{Tr}\,(\varrho\,\log\varrho)$. Note that for any function $A(\phi)$, we have[23]

$$\mathsf{Tr}\left(\varrho\,A\right) \equiv \int_{-\pi}^{\pi} \frac{d\phi}{2\pi}\, \varrho(\phi)\, A(\phi) \quad . \tag{7.153}$$

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

$$F^* = F - Nk_B T\,\lambda\left(\mathsf{Tr}\,\varrho - 1\right) \quad . \tag{7.154}$$

---

[22] 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?

[23] The denominator of 2π 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.

Note that F^* is a *function* of the Lagrange multiplier λ and a *functional* of the density matrix $\varrho(\phi)$. The prefactor $Nk_{\rm B}T$ which multiplies λ 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 λ itself dimensionless and of order N^0. We now have

$$\begin{aligned}\frac{\delta F^*}{\delta\varrho(\phi)} = \frac{\delta}{\delta\varrho(\phi)}\bigg\{ &-\tfrac{1}{2}N\hat{J}(0)\,\Big|\mathsf{Tr}\big(\varrho\, e^{i\phi}\big)\Big|^2 - NH\,\mathsf{Tr}\big(\varrho\,\cos\phi\big) \\ &+ Nk_{\rm B}T\ \mathsf{Tr}\,\big(\varrho\,\log\varrho\big) - Nk_{\rm B}T\,\lambda\Big(\mathsf{Tr}\,\varrho - 1\Big)\bigg\} \quad .\end{aligned} \tag{7.155}$$

To this end, we note that

$$\frac{\delta}{\delta\varrho(\phi)}\,\mathsf{Tr}\,(\varrho\,A) = \frac{\delta}{\delta\varrho(\phi)}\int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\varrho(\phi)\,A(\phi) = \frac{1}{2\pi}\,A(\phi) \quad . \tag{7.156}$$

Thus, we have

$$\begin{aligned}\frac{\delta\tilde{F}}{\delta\varrho(\phi)} = -\tfrac{1}{2}N\hat{J}(0)\cdot\frac{1}{2\pi}\bigg[\mathop{\mathsf{Tr}}_{\phi'}\big(\varrho\, e^{i\phi'}\big)\,e^{-i\phi} + \mathop{\mathsf{Tr}}_{\phi'}\big(\varrho\, e^{-i\phi'}\big)\,e^{i\phi}\bigg] - NH\cdot\frac{\cos\phi}{2\pi} \\ + Nk_{\rm B}T\cdot\frac{1}{2\pi}\Big[\log\varrho(\phi) + 1\Big] - Nk_{\rm B}T\cdot\frac{\lambda}{2\pi} \quad .\end{aligned} \tag{7.157}$$

Now let us define

$$\mathop{\mathsf{Tr}}_{\phi}\big(\varrho\, e^{i\phi}\big) = \int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\varrho(\phi)\,e^{i\phi} \equiv m\,e^{i\phi_0} \quad . \tag{7.158}$$

We then have

$$\log\varrho(\phi) = \frac{\hat{J}(0)}{k_{\rm B}T}\,m\,\cos(\phi-\phi_0) + \frac{H}{k_{\rm B}T}\,\cos\phi + \lambda - 1. \tag{7.159}$$

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

$$\varrho(\phi) = \mathcal{C}\,\exp\bigg(\frac{H_{\rm eff}}{k_{\rm B}T}\,\cos\phi\bigg) \quad , \tag{7.160}$$

where $H_{\rm eff} = \hat{J}(0)\,m + H$ and $\mathcal{C} = e^{\lambda-1}$. The value of λ is then determined by invoking the constraint,

$$\mathsf{Tr}\,\varrho = 1 = \mathcal{C}\!\int\limits_{-\pi}^{\pi}\frac{d\phi}{2\pi}\,\exp\bigg(\frac{H_{\rm eff}}{k_{\rm B}T}\,\cos\phi\bigg) = \mathcal{C}\,I_0(H_{\rm eff}/k_{\rm B}T) \quad , \tag{7.161}$$

where $I_0(z)$ is the Bessel function. We are free to define $\varepsilon \equiv H_{\rm eff}/k_{\rm B}T$, and treat ε as our single variational parameter. We then have the normalized single site density matrix

$$\varrho(\phi) = \frac{\exp(\varepsilon\cos\phi)}{\int\limits_{-\pi}^{\pi}\frac{d\phi'}{2\pi}\,\exp(\varepsilon\cos\phi')} = \frac{\exp(\varepsilon\cos\phi)}{I_0(\varepsilon)} \quad . \tag{7.162}$$

We next compute the following averages:

$$\big\langle\, e^{\pm i\phi} \,\big\rangle = \int\limits_{-\pi}^{\pi} \frac{d\phi}{2\pi}\, \varrho(\phi)\, e^{\pm i\phi} = \frac{I_1(\varepsilon)}{I_0(\varepsilon)} \tag{7.163}$$

$$\big\langle \cos(\phi - \phi') \big\rangle = \mathsf{Re}\, \big\langle e^{i\phi}\, e^{-i\phi'} \big\rangle = \bigg(\frac{I_1(\varepsilon)}{I_0(\varepsilon)}\bigg)^2 \quad , \tag{7.164}$$

as well as

$$\mathsf{Tr}\,(\varrho\, \log \varrho) = \int\limits_{-\pi}^{\pi} \frac{d\phi}{2\pi}\, \frac{e^{\varepsilon \cos\phi}}{I_0(\varepsilon)} \Big\{ \varepsilon \cos\phi - \log I_0(\varepsilon) \Big\} = \varepsilon\, \frac{I_1(\varepsilon)}{I_0(\varepsilon)} - \log I_0(\varepsilon) \quad . \tag{7.165}$$

The dimensionless free energy per site is therefore

$$f(\varepsilon, \theta, h) = -\frac{1}{2} \bigg(\frac{I_1(\varepsilon)}{I_0(\varepsilon)}\bigg)^2 + (\theta\varepsilon - h)\, \frac{I_1(\varepsilon)}{I_0(\varepsilon)} - \theta\, \log I_0(\varepsilon) \quad , \tag{7.166}$$

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 $m = \theta\varepsilon - h = \big\langle\, e^{i\phi} \,\big\rangle$, *i.e.*

$$\theta\varepsilon - h = \frac{I_1(\varepsilon)}{I_0(\varepsilon)} \quad . \tag{7.167}$$

For small ε, we may expand the Bessel functions, using

$$I_\nu(z) = (\tfrac{1}{2} z)^\nu \sum_{k=0}^\infty \frac{(\frac{1}{4} z^2)^k}{k!\, \Gamma(k + \nu + 1)} \quad , \tag{7.168}$$

to obtain

$$f(\varepsilon, \theta, h) = \tfrac{1}{4} \left(\theta - \tfrac{1}{2}\right) \varepsilon^2 + \tfrac{1}{64} \left(2 - 3\theta\right) \varepsilon^4 - \tfrac{1}{2}\, h\varepsilon + \tfrac{1}{16}\, h\varepsilon^3 + \ldots \quad . \tag{7.169}$$

This predicts a second order phase transition at $\theta_{\rm c} = \frac{1}{2}$.[24] Note also the Curie-Weiss form of the susceptibility at high θ:

$$\frac{\partial f}{\partial \varepsilon} = 0 \quad \Longrightarrow \quad \varepsilon = \frac{h}{\theta - \theta_{\rm c}} + \ldots \quad . \tag{7.170}$$

Neglect of fluctuations method

Consider again the Hamiltonian of eqn. 7.150. 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,

$$\hat{H}^{\rm MF} = \tfrac{1}{2} N \hat{J}(0)\, |w|^2 - \tfrac{1}{2} \hat{J}(0)\, |w| \sum_i \left(w^* z_i + w z_i^*\right) - \tfrac{1}{2} H \sum_i \left(z_i^* + z_i\right) \tag{7.171}$$

---

[24] Note that the coefficient of the quartic term in ε is negative for $\theta > \frac{2}{3}$. At $\theta = \theta_{\rm c} = \frac{1}{2}$, the coefficient is positive, but for larger θ one must include higher order terms in the Landau expansion.

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

$$\hat{H}^{\rm MF} = \tfrac{1}{2} N \hat{J}(0)\, |w|^2 - \big(H + \hat{J}(0)\, |w|\big) \sum_i \cos\phi_i \quad . \tag{7.172}$$

The dimensionless free energy per site is then

$$f = \tfrac{1}{2}|w|^2 - \theta \log I_0\bigg(\frac{h + |w|}{\theta}\bigg) \quad . \tag{7.173}$$

Differentiating with respect to $|w|$, one obtains

$$|w| \equiv m = I_1\bigg(\frac{h+m}{\theta}\bigg) \bigg/ I_0\bigg(\frac{h+m}{\theta}\bigg) \quad , \tag{7.174}$$

which is the same equation as eqn. 7.167. The two mean field theories yield the same results in every detail (see §**??**).

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,

$$\hat{H} = -\tfrac{1}{2} \sum_{i,j} J_{ij}\, \sigma_i\, \sigma_j - \sum_k H_k\, \sigma_k \quad , \tag{7.175}$$

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

$$\begin{aligned} \frac{\partial Z}{\partial H_i} &= \beta\, \mathsf{Tr} \Big[\sigma_i\, e^{-\beta \hat{H}} \Big] = \beta Z \cdot \langle \sigma_i \rangle \\ \frac{\partial^2 Z}{\partial H_i\, \partial H_j} &= \beta^2\, \mathsf{Tr} \Big[\sigma_i\, \sigma_j\, e^{-\beta \hat{H}} \Big] = \beta^2 Z \cdot \langle \sigma_i\, \sigma_j \rangle \quad . \end{aligned} \tag{7.176}$$

Thus,

$$\begin{aligned} m_i &= -\frac{\partial F}{\partial H_i} = \langle \sigma_i \rangle \\ \chi_{ij} &= \frac{\partial m_i}{\partial H_j} = -\frac{\partial^2 F}{\partial H_i\, \partial H_j} = \frac{1}{k_{\rm B} T} \cdot \Big\{ \langle \sigma_i\, \sigma_j \rangle - \langle \sigma_i \rangle \, \langle \sigma_j \rangle \Big\} \quad . \end{aligned} \tag{7.177}$$

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

$$C_{ij} \equiv \langle\sigma_i\,\sigma_j\rangle - \langle\sigma_i\rangle\,\langle\sigma_j\rangle \quad . \tag{7.178}$$

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 $\delta\varrho$:

$$\varrho^{\rm var} = \varrho^{\rm eq} + \delta\varrho \quad . \tag{7.179}$$

Then if we calculate a correlator using the variational distribution, we have

$$\langle\sigma_i\,\sigma_j\rangle_{\rm var} = \mathsf{Tr}\left[\varrho^{\rm var}\,\sigma_i\,\sigma_j\right] = \mathsf{Tr}\left[\varrho^{\rm eq}\,\sigma_i\,\sigma_j\right] + \mathsf{Tr}\left[\delta\varrho\,\sigma_i\,\sigma_j\right] \quad . \tag{7.180}$$

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

$$F^{\rm var} = F^{\rm eq} + \sum_{\boldsymbol{\sigma}} \frac{\partial F}{\partial\varrho_{\boldsymbol{\sigma}}}\bigg|_{\varrho^{\rm eq}} \delta\varrho_{\boldsymbol{\sigma}} + \frac{1}{2}\sum_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} \frac{\partial^2 F}{\partial\varrho_{\boldsymbol{\sigma}}\,\partial\varrho_{\boldsymbol{\sigma}'}}\bigg|_{\varrho^{\rm eq}} \delta\varrho_{\boldsymbol{\sigma}}\,\delta\varrho_{\boldsymbol{\sigma}'} + \ldots \quad . \tag{7.181}$$

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 $\varrho(\boldsymbol{\sigma}) = \prod_i \varrho_i(\sigma_i)$, where

$$\varrho_i(\sigma_i) = \left(\frac{1+m_i}{2}\right)\delta_{\sigma_i,1} + \left(\frac{1-m_i}{2}\right)\delta_{\sigma_i,-1} \quad . \tag{7.182}$$

The variational energy $E = \mathsf{Tr}\,(\varrho\,\hat{H})$ is

$$E = -\tfrac{1}{2}\sum_{i,j} J_{ij}\,m_i\,m_j - \sum_i H_i\,m_i \tag{7.183}$$

and the entropy $S = -k_{\rm B}T\,\mathsf{Tr}\,(\varrho\log\varrho)$ is

$$S = -k_{\rm B}\sum_i \left\{ \left(\frac{1+m_i}{2}\right)\log\left(\frac{1+m_i}{2}\right) + \left(\frac{1-m_i}{2}\right)\log\left(\frac{1-m_i}{2}\right) \right\} \quad . \tag{7.184}$$

Setting the variation $\frac{\partial F}{\partial m_i} = 0$, with $F = E - TS$, we obtain the mean field equations,

$$m_i = \tanh\big(\beta J_{ij}\, m_j + \beta H_i\big) \quad , \tag{7.185}$$

where we use the summation convention: $J_{ij}\, m_j \equiv \sum_j J_{ij}\, m_j$. Suppose $T > T_c$ and m_i is small. Then we can expand the RHS of the above mean field equations, obtaining

$$\big(\delta_{ij} - \beta J_{ij}\big)\, m_j = \beta H_i \quad . \tag{7.186}$$

Thus, the susceptibility tensor χ is the inverse of the matrix $(k_B T\,\mathbb{1} - \mathbb{J})$:

$$\chi_{ij} = \frac{\partial m_i}{\partial H_j} = \big(k_B T\,\mathbb{1} - \mathbb{J}\big)^{-1}_{ij} \quad , \tag{7.187}$$

where $\mathbb{1}$ is the identity. Note also that so-called *connected averages* of the kind in eqn. 7.178 vanish identically if we compute them using our variational density matrix, since all the sites are independent, hence

$$\langle \sigma_i\, \sigma_j \rangle = \mathsf{Tr}\,\big(\varrho^{\rm var}\, \sigma_i\, \sigma_j\big) = \mathsf{Tr}\,\big(\varrho^{\rm var}_i\, \sigma_i\big)\, \mathsf{Tr}\,\big(\varrho^{\rm var}_j\, \sigma_j\big) = \langle \sigma_i \rangle \langle \sigma_j \rangle \quad , \tag{7.188}$$

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

$$\hat{m}(\boldsymbol{q}) = \frac{\hat{H}(\boldsymbol{q})}{k_B T - \hat{J}(\boldsymbol{q})} \equiv \hat{\chi}(\boldsymbol{q})\, \hat{H}(\boldsymbol{q}) \quad . \tag{7.189}$$

Once again, our definition of lattice Fourier transform of a function $\phi(\boldsymbol{R})$ is

$$\hat{\phi}(\boldsymbol{q}) \equiv \sum_{\boldsymbol{R}} \phi(\boldsymbol{R})\, e^{-i\boldsymbol{q}\cdot\boldsymbol{R}} \qquad , \qquad \phi(\boldsymbol{R}) = \Omega \int\limits_{\hat{\Omega}} \frac{d^d q}{(2\pi)^d}\, \hat{\phi}(\boldsymbol{q})\, e^{i\boldsymbol{q}\cdot\boldsymbol{R}} \quad , \tag{7.190}$$

where Ω 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

$$\begin{aligned} \hat{J}(\boldsymbol{q}) &= \sum_{\boldsymbol{R}} J(\boldsymbol{R}) \Big(1 - i\boldsymbol{q}\cdot\boldsymbol{R} - \tfrac{1}{2}(\boldsymbol{q}\cdot\boldsymbol{R})^2 + \ldots\Big) \\ &= \hat{J}(0)\cdot\Big\{1 - q^2 R_*^2 + \mathcal{O}(q^4)\Big\} \quad , \end{aligned} \tag{7.191}$$

where

$$R_*^2 = \frac{\sum_{\boldsymbol{R}} \boldsymbol{R}^2 J(\boldsymbol{R})}{2d\sum_{\boldsymbol{R}} J(\boldsymbol{R})} \quad . \tag{7.192}$$

Here we have assumed inversion symmetry for the lattice, in which case

$$\sum_{\boldsymbol{R}} R^\mu R^\nu J(\boldsymbol{R}) = \frac{1}{d}\cdot\delta^{\mu\nu} \sum_{\boldsymbol{R}} \boldsymbol{R}^2 J(\boldsymbol{R}) \quad . \tag{7.193}$$

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

$$\hat{\chi}(\boldsymbol{q}) = \frac{1}{k_{\rm B}(T-T_{\rm c}) + k_{\rm B}T_{\rm c}\,R_*^2\,\boldsymbol{q}^2 + \mathcal{O}(q^4)} = \frac{1}{k_{\rm B}T_{\rm c}\,R_*^2}\cdot\frac{1}{\xi^{-2}+q^2+\mathcal{O}(q^4)} \quad , \tag{7.194}$$

where

$$\xi(T) = R_* \cdot \left(\frac{T-T_{\rm c}}{T_{\rm c}}\right)^{-1/2} \tag{7.195}$$

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

$$m_i = \sum_j \chi_{ij}\,H_j \qquad , \qquad \chi_{ij} = \Omega \int\limits_{\hat{\Omega}} \frac{d^dq}{(2\pi)^d}\,\hat{\chi}(\boldsymbol{q})\,e^{i\boldsymbol{q}\cdot(\boldsymbol{R}_i-\boldsymbol{R}_j)} \quad . \tag{7.196}$$

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

$$\hat{\chi}^{-1}(\boldsymbol{q}) = k_{\rm B}T - \hat{J}(\boldsymbol{q}) = k_{\rm B}T - J\sum_{\boldsymbol{\delta}} e^{i\boldsymbol{q}\cdot\boldsymbol{\delta}} \quad , \tag{7.197}$$

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

$$\chi(\boldsymbol{R}, H=0) = \int\limits_{-\pi}^{\pi} \frac{d\theta_1}{2\pi} \cdots \int\limits_{-\pi}^{\pi} \frac{d\theta_d}{2\pi}\,\frac{e^{in_1\theta_1}\cdots e^{in_d\theta_d}}{k_{\rm B}T - (2J\cos\theta_1 + \ldots + 2J\cos\theta_d)} \quad , \tag{7.198}$$

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 §6.5.7 on Ornstein-Zernike theory[25]). For convenience we reiterate those results:

- In $d = 1$,

$$\chi_{d=1}(x) = \left(\frac{\xi}{2k_{\rm B}T_{\rm c}R_*^2}\right) e^{-|x|/\xi} \quad . \tag{7.199}$$

---

[25] 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 δn should be negative, while in the current discussion a positive magnetic field H prefers $m > 0$.

- In $d>1$, with $r\to\infty$ and ξ fixed,

  $$\chi^{\rm OZ}_d(\boldsymbol{r}) \simeq C_d \cdot \frac{\xi^{(3-d)/2}}{k_{\rm B}T\,R_*^2} \cdot \frac{e^{-r/\xi}}{r^{(d-1)/2}} \cdot \bigg\{1+\mathcal{O}\bigg(\frac{d-3}{r/\xi}\bigg)\bigg\} \quad , \tag{7.200}$$

  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 $\boldsymbol{r}$),

  $$\chi_d(\boldsymbol{r}) \simeq \frac{C'_d}{k_{\rm B}TR_*^2} \cdot \frac{e^{-r/\xi}}{r^{d-2}} \cdot \bigg\{1+\mathcal{O}\bigg(\frac{d-3}{r/\xi}\bigg)\bigg\} \quad . \tag{7.201}$$

  In $d=2$ dimensions we obtain

  $$\chi_{d=2}(\boldsymbol{r}) \simeq \frac{C'_2}{k_{\rm B}TR_*^2} \cdot \log\bigg(\frac{r}{\xi}\bigg)\, e^{-r/\xi} \cdot \bigg\{1+\mathcal{O}\bigg(\frac{1}{\log(r/\xi)}\bigg)\bigg\} \quad , \tag{7.202}$$

  where the C'_d are dimensionless constants.

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

$$C(\boldsymbol{r},T)=r^{-(d-2+\eta)}\,\phi\big(r/\xi(T)\big) \quad , \tag{7.203}$$

where η 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[26]. Thus, we've encountered in this section two additional critical exponents, the anomalous exponent η and the correlation length exponent ν, which we first met in §7.3.2, whose mean field values are $\eta=0$ and $\nu=\frac{1}{2}$.

---

[26] 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.

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^{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,

$$\varrho_N(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N)=\frac{1}{N!}\times\begin{cases} Z_N^{-1}\,e^{-\beta\hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} & \text{OCE} \\ \Xi^{-1}\,e^{\beta\mu N}\,e^{-\beta\hat{H}_N(\{\boldsymbol{p}_i\},\{\boldsymbol{x}_i\})} & \text{GCE} \end{cases}\quad, \tag{8.1}$$

We assume a Hamiltonian of the form

$$\hat{H}_N=\sum_{i=1}^N\frac{\hat{\boldsymbol{p}}_i^2}{2m}+\sum_{i=1}^N v(\hat{\boldsymbol{x}}_i)+\sum_{i<j}^N u(\hat{\boldsymbol{x}}_i-\hat{\boldsymbol{x}}_j). \tag{8.2}$$

Here $v(\boldsymbol{x})=U_{\text{ext}}(\boldsymbol{x})$ is due to external forces. In the context of transport theory, $v(\hat{\boldsymbol{x}})$ typically will denote the effect of an applied external field, *e.g.* $v(\hat{\boldsymbol{x}})=-q\boldsymbol{E}\cdot\hat{\boldsymbol{x}}$ for a particle of charge q in the presence of a uniform electric field $\boldsymbol{E}$. We write $\hat{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{p}}_i$ for the corresponding phase space variables, the position and momentum vectors for the i^{th} particle, respectively. The quantity

$$\varrho_N^{\text{eq}}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_N)\prod_{j=1}^N d\mu_j\quad, \tag{8.3}$$

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, witparticle #1 lying within d^3x_1 of $\boldsymbol{x}_1$ and having momentum within d^dp_1 of $\boldsymbol{p}_1$, *etc.* The temperature T and chemical potential μ are constants, independent of position. Note that $\varrho_N^{\text{eq}}(\{\boldsymbol{x}_i\},\{\boldsymbol{p}_i\})$ has units of h^{-N}.

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, $\varrho_N($ For 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* ℓ and the *collision time* τ 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

$$\begin{aligned} f(\boldsymbol{r},\boldsymbol{p};t)&=\sum_{i=1}^N\big\langle\,\delta\big(\boldsymbol{r}-\boldsymbol{x}_i(t)\big)\,\delta\big(\boldsymbol{p}-\boldsymbol{p}_i(t)\big)\,\big\rangle \\ &=Nh^{-d}\!\int\prod_{i=2}^N d\mu_i\,\varrho_N(\boldsymbol{r},\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N,\boldsymbol{p},\boldsymbol{p}_2,\ldots,\boldsymbol{p}_N,t)\quad. \end{aligned} \tag{8.4}$$

Note that $f(\boldsymbol{r},\boldsymbol{p},t)$ has dimensions of h^{-d}, and $f(\boldsymbol{r},\boldsymbol{p},t)\,d^3\!r\,d^3\!p$ is the average number of particles found within $d^3\!r$ of $\boldsymbol{r}$ and $d^3\!p$ of $\boldsymbol{p}$ at time t. The full k-particle density matrix, which for classical systems we may write in terms of its diagonal matrix elements in the $\big(\{\boldsymbol{x}_i\},\{\boldsymbol{p}_i\}\big)$ basis, is given by

$$\begin{aligned} f_k\big(\{\boldsymbol{x}_i\}\,,\{\boldsymbol{p}_i\}\,,t\big) &= \sum_{\{j_1,\ldots,j_k\}}{\!\!}' \big\langle \delta\big(\boldsymbol{x}_1-\hat{\boldsymbol{x}}_{j_1}(t)\big)\cdots\delta\big(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_{j_k}(t)\big)\,\delta\big(\boldsymbol{p}_1-\hat{\boldsymbol{p}}_{j_1}(t)\big)\cdots\delta\big(\boldsymbol{p}_k-\hat{\boldsymbol{p}}_{j_k}(t)\big)\big\rangle \\ &= \frac{N!\,h^{-kd}}{(N-k)!}\int\prod_{i=k+1}^{N} d\mu_i\;\varrho_N\big(\{\boldsymbol{x}_j\}\,,\,\{\boldsymbol{p}_j\}\,,\,t\big) \quad , \end{aligned} \tag{8.5}$$

where $\{\boldsymbol{x}_j\}$ as an argument of the s-body density matrix f_k denotes the ordered set $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k\}$, where $k\in\{1,\ldots,N\}$ (similarly for $\{\boldsymbol{p}_j\}$). 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

$$\int\prod_{i=1}^{k} d^d\!x_i\,d^d\!p_i\;f_k(\{\boldsymbol{x}_j\},\{\boldsymbol{p}_j\},t)=\frac{N!}{(N-k)!} \quad , \tag{8.6}$$

for all t. We write the one-body density matrix $f_1(\boldsymbol{x}_1,\boldsymbol{p}_1,t)\equiv f(\boldsymbol{r},\boldsymbol{p},t)$, where $\boldsymbol{r}=\boldsymbol{x}_1$ and $\boldsymbol{p}=\boldsymbol{p}_1$.

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:

$$f^0(\boldsymbol{r},\boldsymbol{p})=n\,(2\pi m k_{\rm B}T)^{-3/2}\,e^{-\boldsymbol{p}^2/2mk_{\rm B}T} \quad , \tag{8.7}$$

where $n=N/V$ or $n=n(T,\mu)$ is the particle density in the OCE or GCE. From the one-body distribution we can compute things like the particle current, $\boldsymbol{j}$, and the energy current, $\boldsymbol{j}_\varepsilon$:

$$\boldsymbol{j}(\boldsymbol{r},t)=\int\! d^d\!p\,f(\boldsymbol{r},\boldsymbol{p};t)\,\frac{\boldsymbol{p}}{m} \tag{8.8}$$

$$\boldsymbol{j}_\varepsilon(\boldsymbol{r},t)=\int\! d^d\!p\,f(\boldsymbol{r},\boldsymbol{p};t)\;\varepsilon(\boldsymbol{p})\,\frac{\boldsymbol{p}}{m} \quad , \tag{8.9}$$

where $\varepsilon(\boldsymbol{p})=\boldsymbol{p}^2/2m$. Clearly these currents both 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

$$đq=T\,ds=d\varepsilon-\mu\,dn \quad , \tag{8.10}$$

where s, ε, and n are entropy density, energy density, and particle density, respectively, and $đq$ is the differential heat density. This relation may be case as one among the corresponding current densities:

$$\boldsymbol{j}_q=T\boldsymbol{j}_s=\boldsymbol{j}_\varepsilon-\mu\,\boldsymbol{j} \quad . \tag{8.11}$$

Thus, in a system with no particle flow, $\boldsymbol{j}=0$ and the *heat current* $\boldsymbol{j}_q$ is the same as the energy current $\boldsymbol{j}_\varepsilon$.

When the individual particles are not point particles, they possess angular momentum as well as linear momentum. Following Lifshitz and Pitaevskii, we abbreviate $\varGamma=(\boldsymbol{p},\boldsymbol{L})$ for these two variables for the case of diatomic molecules, and $\varGamma = (\boldsymbol{p}, \boldsymbol{L}, \hat{\boldsymbol{n}} \cdot \boldsymbol{L})$ 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,

$$d\varGamma = \begin{cases} d^3\!p & \text{point particles} \\ d^3\!p\, L\, dL\, d\Omega_{\hat{\boldsymbol{L}}} & \text{diatomic molecules} \\ d^3\!p\, L^2\, dL\, d\Omega_{\hat{\boldsymbol{L}}}\, d\cos\vartheta & \text{symmetric tops} \quad , \end{cases} \tag{8.12}$$

where $\vartheta = \cos^{-1}(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{L}})$. We will call the set $\varGamma$ the 'kinematic variables'. The instantaneous number density at $\boldsymbol{r}$ is then

$$n(\boldsymbol{r}, t) = \int\! d\varGamma\, f(\boldsymbol{r}, \varGamma; t) \quad . \tag{8.13}$$

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. Recall that

$$f(\boldsymbol{r}, \boldsymbol{p}, t)\, d^3\!r\, d^3\!p \equiv \begin{cases} \text{\# of particles with positions within } d^3\!r \text{ of} \\ \boldsymbol{r} \text{ and momenta within } d^3\!p \text{ of } \boldsymbol{p} \text{ at time } t. \end{cases} \tag{8.14}$$

Thus, the units of $f(\boldsymbol{r}, \boldsymbol{p}, t)$ are those of inverse action, *i.e.* T/ML^2. We now ask how the distribution functions $f(\boldsymbol{r}, \boldsymbol{p}, t)$ 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:

$$\boldsymbol{u} = (\ \dot{x} \ \ ,\ \dot{y} \ \ ,\ \dot{z} \ \ ,\ \dot{p}_x \ \ ,\ \dot{p}_y \ \ ,\ \dot{p}_z\ ) \tag{8.16}$$

$$\boldsymbol{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial p_x}, \frac{\partial}{\partial p_y}, \frac{\partial}{\partial p_z}\right) \quad . \tag{8.17}$$

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

$$\frac{d\boldsymbol{r}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}} = \boldsymbol{v}(\boldsymbol{p}) \qquad , \qquad \frac{d\boldsymbol{p}}{dt} = -\frac{\partial H}{\partial \boldsymbol{r}} = \boldsymbol{F}_{\text{ext}} \quad , \tag{8.18}$$

where $\boldsymbol{F}$ is an external applied force. Here,

$$H(\boldsymbol{p},\boldsymbol{r}) = \varepsilon(\boldsymbol{p}) + U_{\text{ext}}(\boldsymbol{r}) \quad . \tag{8.19}$$

For example, under the influence of gravity, $U_{\text{ext}}(\boldsymbol{r}) = m\boldsymbol{g}\cdot\boldsymbol{r}$ and $\boldsymbol{F} = -\boldsymbol{\nabla} U_{\text{ext}} = -m\boldsymbol{g}$.

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

$$\frac{\partial f}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla} f = 0 \quad . \tag{8.20}$$

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.

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

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \dot{\boldsymbol{r}}\cdot\frac{\partial f}{\partial \boldsymbol{r}} + \dot{\boldsymbol{p}}\cdot\frac{\partial f}{\partial \boldsymbol{p}} = \left(\frac{df}{dt}\right)_{\text{coll}} \tag{8.21}$$

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. Suppose we evaluate the time-dependent distribution $f(\boldsymbol{r},\boldsymbol{p},t)$ along a particle trajectory, *i.e.* substituting $\boldsymbol{r}\to\boldsymbol{r}(t)$ and $\boldsymbol{p} = \boldsymbol{p}(t)$. Then

$$\frac{d}{dt} f\big(\boldsymbol{r}(t),\boldsymbol{p}(t),t\big) = \frac{\partial f}{\partial \boldsymbol{r}}\cdot\frac{d\boldsymbol{r}}{dt}\bigg|_{\{\boldsymbol{r}(t),\boldsymbol{p}(t),t\}} + \frac{\partial f}{\partial \boldsymbol{p}}\cdot\frac{d\boldsymbol{p}}{dt}\bigg|_{\{\boldsymbol{r}(t),\boldsymbol{p}(t),t\}} + \frac{\partial f}{\partial t}\bigg|_{\{\boldsymbol{r}(t),\boldsymbol{p}(t),t\}} = \frac{Df}{Dt}\bigg|_{\{\boldsymbol{r}(t),\boldsymbol{p}(t),t\}} \quad . \tag{8.22}$$

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

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{\text{str}} + \left(\frac{df}{dt}\right)_{\text{coll}} \quad , \tag{8.23}$$

where

$$\left(\frac{\partial f}{\partial t}\right)_{\text{str}} \equiv -\dot{\boldsymbol{r}}\cdot\frac{\partial f}{\partial \boldsymbol{r}} - \dot{\boldsymbol{p}}\cdot\frac{\partial f}{\partial \boldsymbol{p}} \tag{8.24}$$

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

$$\frac{\partial f}{\partial t} + \frac{\partial \varepsilon}{\partial \boldsymbol{p}}\cdot\frac{\partial f}{\partial \boldsymbol{r}} - \boldsymbol{\nabla} U_{\text{ext}}\cdot\frac{\partial f}{\partial \boldsymbol{p}} = 0 \quad . \tag{8.25}$$

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 $\boldsymbol{F}_{\rm ext}=0$. We then have

$$\frac{\partial f}{\partial t}+\boldsymbol{v}(\boldsymbol{p})\cdot\frac{\partial f}{\partial \boldsymbol{r}}=0 \quad . \tag{8.26}$$

Clearly, then, *any* function of the form

$$f(\boldsymbol{r},\boldsymbol{p},t)=\varphi\big(\boldsymbol{r}-\boldsymbol{v}(\boldsymbol{p})\,t\,,\,\boldsymbol{p}\big) \tag{8.27}$$

will be a solution to the collisionless Boltzmann equation, where $\boldsymbol{v}(\boldsymbol{p})=\frac{\partial\varepsilon}{\partial\boldsymbol{p}}$. One possible solution would be the Boltzmann distribution,

$$f(\boldsymbol{r},\boldsymbol{p},t)=C\,e^{\mu/k_{\rm B}T}\,e^{-\varepsilon(\boldsymbol{p})/k_{\rm B}T} \quad , \tag{8.28}$$

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 $\int d^d\!r\,d^d\!p\,f(\boldsymbol{r},\boldsymbol{p})=V\exp(\mu/k_{\rm B}T)=\langle\hat{N}\rangle$.

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

$$f(\boldsymbol{r},\boldsymbol{p},t)=C\,e^{-\left(\boldsymbol{r}-\frac{\boldsymbol{p}t}{m}\right)^2/2\sigma^2}\,e^{-\boldsymbol{p}^2/2\kappa^2} \quad . \tag{8.29}$$

Consider the one-dimensional version, and rescale position, momentum, and time so that

$$f(x,p,t)=C\,e^{-(\bar{x}-\bar{p}\,\bar{t})^2/2}\,e^{-\bar{p}^2/2} \quad . \tag{8.30}$$

Now consider the level sets of f, where $f(x,p,t)=C\,e^{-\alpha^2/2}$. The equation for these sets is

$$\bar{x}=\bar{p}\,\bar{t}\pm\sqrt{\alpha^2-\bar{p}^2} \quad . \tag{8.31}$$

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 α^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 $\theta=\text{ctn}^{-1}(t)$ with respect to the $\bar{x}$ axis. For $\bar{t}>0$, he particles at the outer edges of the cloud are more likely to be moving away from the center. See the sketches in fig. 8.1

Suppose we add in a constant external force $\boldsymbol{F}_{\rm ext}$. Then it is easy to show (and left as an exercise to the reader to prove) that any function of the form

$$f(\boldsymbol{r},\boldsymbol{p},t)=\varphi\bigg(\boldsymbol{r}-\frac{\boldsymbol{p}\,t}{m}-\frac{\boldsymbol{F}_{\rm ext}t^2}{2m}\,,\,\boldsymbol{p}-\frac{\boldsymbol{F}_{\rm ext}t}{m}\bigg) \tag{8.32}$$

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.

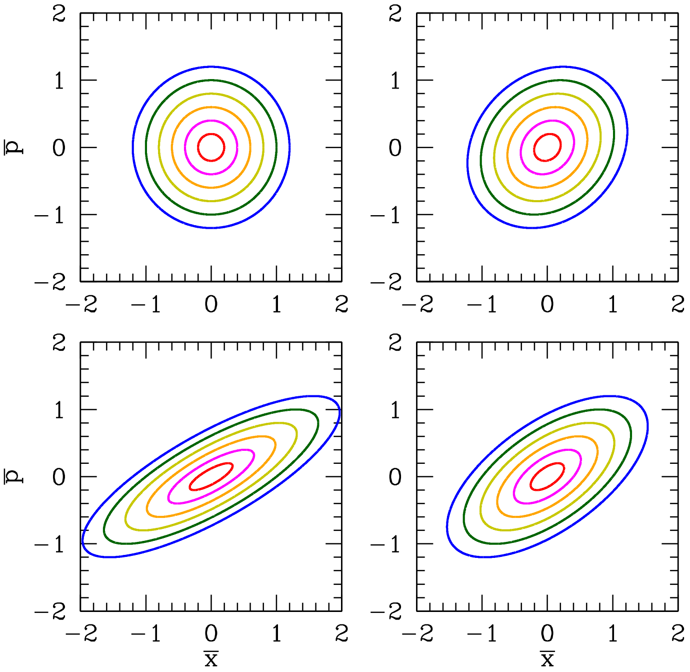


Figure 8.1: Level sets for a sample $f(\bar{x},\bar{p},\bar{t}) = C\,e^{-(\bar{x}-\bar{p}\bar{t})^2/2}e^{-\bar{p}^2/2}$, for values $f = C\,e^{-\alpha^2/2}$ with α in equally spaced intervals from $\alpha = 0.2$ (red) to $\alpha = 1.2$ (blue). The time variable $\bar{t}$ is taken to be $\bar{t} = 0.0$ (upper left), 0.2 (upper right), 0.8 (lower right), and 1.3 (lower left).

8.3.3 Collisional invariants

Consider a function $A(\boldsymbol{r},\boldsymbol{p})$ of position and momentum. Its average value at time t is

$$\langle A(t)\rangle = \int d^3\!r\, d^3\!p\; A(\boldsymbol{r},\boldsymbol{p})\, f(\boldsymbol{r},\boldsymbol{p},t) \quad . \tag{8.33}$$

Taking the time derivative,

$$\begin{aligned} \frac{dA}{dt} &= \int d^3\!r \int d^3\!p\; A(\boldsymbol{r},\boldsymbol{p})\, \frac{\partial f}{\partial t} \\ &= \int d^3\!r \int d^3\!p\; A(\boldsymbol{r},\boldsymbol{p}) \left\{ -\frac{\partial}{\partial \boldsymbol{r}} \cdot (\dot{\boldsymbol{r}} f) - \frac{\partial}{\partial \boldsymbol{p}} \cdot (\dot{\boldsymbol{p}} f) + \left(\frac{df}{dt}\right)_{\rm coll} \right\} \\ &= \int d^3\!r \int d^3\!p \left\{ \left(\frac{\partial A}{\partial \boldsymbol{r}} \cdot \frac{d\boldsymbol{r}}{dt} + \frac{\partial A}{\partial \boldsymbol{p}} \cdot \frac{d\boldsymbol{p}}{dt} \right) f + A(\boldsymbol{r},\boldsymbol{p}) \left(\frac{df}{dt}\right)_{\rm coll} \right\} \quad . \end{aligned} \tag{8.34}$$

Hence, if A is preserved by the dynamics between collisions, then[2]

$$\frac{d\langle A(t)\rangle}{dt} = \frac{\partial A}{\partial \boldsymbol{r}} \cdot \frac{d\boldsymbol{r}}{dt} + \frac{\partial A}{\partial \boldsymbol{p}} \cdot \frac{d\boldsymbol{p}}{dt} = 0 \quad . \tag{8.35}$$

---

[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.

We therefore have that the rate of change of $\langle A(t)\rangle$ is determined wholly by the collision integral

$$\frac{d\langle A(t)\rangle}{dt} = \int\! d^3\!r \int\! d^3\!p \; A(\boldsymbol{r},\boldsymbol{p}) \left(\frac{df}{dt}\right)_{\rm coll} \quad . \tag{8.36}$$

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 ($A = \varepsilon(\boldsymbol{p})$).

8.3.4 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 $|\varGamma\rangle$ scatters, in the presence of an external potential, to a state $|\varGamma'\rangle$. Recall that $\varGamma$ is an abbreviation for the set of kinematic variables, *e.g.* $\varGamma = (\boldsymbol{p},\boldsymbol{L})$ in the case of a diatomic molecule. For point particles, $\varGamma = (p_x, p_y, p_z)$ and $d\varGamma = d^3\!p$.

Single particle scattering

We now define the function $w\big(\varGamma' \,|\, \varGamma\big)$ such that

$$w\big(\varGamma' \,|\, \varGamma\big)\, f(\varGamma)\, d\varGamma\, d\varGamma' = \text{rate per unit volume to scatter } \big\{|\,\varGamma + d\varGamma\,\rangle\big\} \to \big\{|\,\varGamma' + d\varGamma'\,\rangle\big\} \tag{8.37}$$

at time t. Here we have suppressed the time dependence in the distribution $f(\varGamma,t)$, and typically we will presume $w\big(\varGamma' \,|\, \varGamma\big)$ to be independent of both $\boldsymbol{r}$ and t. By $\big\{|\,\varGamma + d\varGamma\,\rangle\big\}$ we mean states with momenta within $d^3\!p$ of $\boldsymbol{p}$ (more generally, within $d\varGamma = d^3\!p\, d^3\!L$ 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 $w\, d\varGamma$ are therefore $1/T$. The *differential scattering cross section* for single particle scattering is then

$$d\sigma = \frac{w\big(\varGamma' \,|\, \varGamma\big)}{n\,|\boldsymbol{v}|}\, d\varGamma' \quad , \tag{8.38}$$

where $\boldsymbol{v} = \boldsymbol{p}/m$ is the particle's velocity and n the density.

In computing the collision integral for the state $|\boldsymbol{r},\varGamma\rangle$, we must take care to sum over contributions from transitions *out of* this state, *i.e.* $|\varGamma\rangle \to |\varGamma'\rangle$, which reduce $f(\boldsymbol{r},\varGamma)$, and transitions *into* this state, *i.e.* $|\varGamma'\rangle \to |\varGamma\rangle$, which increase $f(\boldsymbol{r},\varGamma)$. Thus, for one-body scattering, we have

$$\frac{D}{Dt}\, f(\boldsymbol{r},\varGamma;t) = \int\! d\varGamma' \,\Big\{ w(\varGamma \,|\, \varGamma')\, f(\boldsymbol{r},\varGamma';t) - w(\varGamma' \,|\, \varGamma)\, f(\boldsymbol{r},\varGamma;t) \Big\} = \left(\frac{df}{dt}\right)_{\rm coll} \quad . \tag{8.39}$$

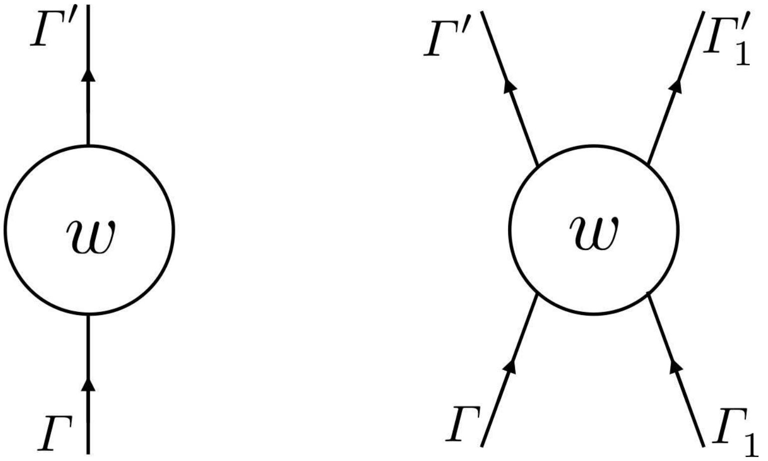


Figure 8.2: Left: single particle scattering process $|\varGamma\rangle \to |\varGamma'\rangle$. Right: two-particle scattering process $|\varGamma,\varGamma_1\rangle \to |\varGamma',\varGamma'_1\rangle$.

Two particle scattering

The second class is that of two-particle scattering processes, *i.e.* $|\,\{\boldsymbol{r},\varGamma\},\{\boldsymbol{r},\varGamma_1\}\,\rangle \to |\,\{\boldsymbol{r},\varGamma'\},\{\boldsymbol{r},\varGamma'_1\}\,\rangle$. We define the scattering function $w\big(\varGamma',\varGamma'_1\,|\,\varGamma,\varGamma_1\big)$ by

$$w\big(\varGamma',\varGamma'_1\,|\,\varGamma,\varGamma_1\big)\,f(\varGamma)\,f(\varGamma_1)\,d\varGamma\,d\varGamma_1\,d\varGamma'\,d\varGamma'_1 = \begin{cases}\text{rate per unit volume to scatter two particles}\\ \big\{|\,\varGamma+d\varGamma,\varGamma_1+d\varGamma_1\,\rangle\big\} \to \big\{|\,\varGamma+d\varGamma',\varGamma'_1+d\varGamma'_1\,\rangle\big\}\end{cases} \tag{8.40}$$

at time t. We assume that $w\big(\varGamma',\varGamma'_1\,|\,\varGamma,\varGamma_1\big)$ is independent of $\boldsymbol{r}$ and t. Thus the units of $w\,d\varGamma\,d\varGamma_1$ are again $1/T$, and the differential scattering cross section is

$$d\sigma = \frac{w\big(\varGamma,\varGamma_1\,|\,\varGamma',\varGamma'_1\big)}{|\boldsymbol{v}-\boldsymbol{v}_1|}\,d\varGamma'\,d\varGamma'_1 \quad . \tag{8.41}$$

For two-body scattering, we therefore have

$$\begin{aligned}\frac{D}{Dt}\,f(\boldsymbol{r},\varGamma;t) = \int\!d\varGamma_1\!\int\!d\varGamma'\!\int\!d\varGamma'_1\,\Big\{&w\big(\varGamma,\varGamma_1\,|\,\varGamma',\varGamma'_1\big)\,f_2(\boldsymbol{r},\varGamma';\boldsymbol{r},\varGamma'_1;t)\\ &-w\big(\varGamma',\varGamma'_1\,|\,\varGamma,\varGamma_1\big)\,f_2(\boldsymbol{r},\varGamma;\boldsymbol{r},\varGamma_1;t)\Big\} = \left(\frac{df}{dt}\right)_{\!\text{coll}} \quad .\end{aligned} \tag{8.42}$$

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. To close the equations, we make the *approximation*

$$f_2(\boldsymbol{r},\varGamma;\tilde{\boldsymbol{r}},\tilde{\varGamma};t) \approx f(\boldsymbol{r},\varGamma;t)\,f(\tilde{\boldsymbol{r}},\tilde{\varGamma};t) \quad . \tag{8.43}$$

We then have

$$\begin{aligned}\frac{D}{Dt}\, f(\boldsymbol{r},\mathit{\Gamma};t) = \int\! d\mathit{\Gamma}_1 \!\int\! d\mathit{\Gamma}' \!\int\! d\mathit{\Gamma}_1' \,\Big\{ w\big(\mathit{\Gamma},\mathit{\Gamma}_1 \,\big|\, \mathit{\Gamma}',\mathit{\Gamma}_1'\big)\, f(\boldsymbol{r},\mathit{\Gamma}';t)\, f(\boldsymbol{r},\mathit{\Gamma}_1';t) \\ - w\big(\mathit{\Gamma}',\mathit{\Gamma}_1' \,\big|\, \mathit{\Gamma},\mathit{\Gamma}_1\big)\, f(\boldsymbol{r},\mathit{\Gamma};t)\, f(\boldsymbol{r},\mathit{\Gamma}_1;t)\Big\} = \bigg(\frac{df}{dt}\bigg)_{\rm coll} \quad .\end{aligned} \tag{8.44}$$

We stress that in both cases we assume that any 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.

8.3.5 Detailed balance

Classical mechanics places some restrictions on the form of the kernel $w\big(\mathit{\Gamma},\mathit{\Gamma}_1 \,\big|\, \mathit{\Gamma}',\mathit{\Gamma}_1'\big)$. In particular, if $\mathit{\Gamma}^{\mathcal{T}} = (-\boldsymbol{p}, -\boldsymbol{L})$ denotes the kinematic variables under time reversal ($\mathcal{T}$), then

$$w\big(\mathit{\Gamma}',\mathit{\Gamma}_1' \,\big|\, \mathit{\Gamma},\mathit{\Gamma}_1\big) = w\big(\mathit{\Gamma}^{\mathcal{T}},\mathit{\Gamma}_1^{\mathcal{T}} \,\big|\, \mathit{\Gamma}'^{\mathcal{T}},\mathit{\Gamma}_1'^{\mathcal{T}}\big) \quad . \tag{8.45}$$

This is because the time reverse of the process $|\mathit{\Gamma},\mathit{\Gamma}_1\rangle \to |\mathit{\Gamma}',\mathit{\Gamma}_1'\rangle$ is $|\mathit{\Gamma}'^{\mathcal{T}},\mathit{\Gamma}_1'^{\mathcal{T}}\rangle \to |\mathit{\Gamma}^{\mathcal{T}},\mathit{\Gamma}_1^{\mathcal{T}}\rangle$.

In equilibrium, we must have

$$w\big(\mathit{\Gamma}',\mathit{\Gamma}_1' \,\big|\, \mathit{\Gamma},\mathit{\Gamma}_1\big)\; f^0(\mathit{\Gamma})\, f^0(\mathit{\Gamma}_1)\, d^4\mathit{\Gamma} = w\big(\mathit{\Gamma}^{\mathcal{T}},\mathit{\Gamma}_1^{\mathcal{T}} \,\big|\, \mathit{\Gamma}'^{\mathcal{T}},\mathit{\Gamma}_1'^{\mathcal{T}}\big)\; f^0(\mathit{\Gamma}'^{\mathcal{T}})\, f^0(\mathit{\Gamma}_1'^{\mathcal{T}})\, d^4\mathit{\Gamma}^{\mathcal{T}} \tag{8.46}$$

where

$$d^4\mathit{\Gamma} \equiv d\mathit{\Gamma}\, d\mathit{\Gamma}_1\, d\mathit{\Gamma}' d\mathit{\Gamma}_1' \qquad , \qquad d^4\mathit{\Gamma}^{\mathcal{T}} \equiv d\mathit{\Gamma}^{\mathcal{T}}\, d\mathit{\Gamma}_1^{\mathcal{T}}\, d\mathit{\Gamma}'^{\mathcal{T}} d\mathit{\Gamma}_1'^{\mathcal{T}} \quad . \tag{8.47}$$

Since $d\mathit{\Gamma} = d\mathit{\Gamma}^{\mathcal{T}}$ *etc.*, we may cancel the differentials above, and after invoking eqn. 8.45 and suppressing the common $\boldsymbol{r}$ label, we find

$$f^0(\mathit{\Gamma})\, f^0(\mathit{\Gamma}_1) = f^0(\mathit{\Gamma}'^{\mathcal{T}})\, f^0(\mathit{\Gamma}_1'^{\mathcal{T}}) \quad . \tag{8.48}$$

This is the condition of *detailed balance*. For the Boltzmann distribution, we have $f^0(\mathit{\Gamma}) = A\, e^{-\varepsilon/k_{\rm B}T}$, where A is a constant and where $\varepsilon = \varepsilon(\mathit{\Gamma})$ is the kinetic energy, *e.g.* $\varepsilon(\mathit{\Gamma}) = \boldsymbol{p}^2/2m$ in the case of point particles. Note that $\varepsilon(\mathit{\Gamma}^{\mathcal{T}}) = \varepsilon(\mathit{\Gamma})$. Detailed balance is satisfied because the kinematics of the collision requires energy conservation:

$$\varepsilon + \varepsilon_1 = \varepsilon' + \varepsilon_1' \quad . \tag{8.49}$$

Since momentum is also kinematically conserved, *i.e.*

$$\boldsymbol{p} + \boldsymbol{p}_1 = \boldsymbol{p}' + \boldsymbol{p}_1' \quad , \tag{8.50}$$

any distribution of the form

$$f^0(\mathit{\Gamma}) = A\, \exp\bigg(-\, \frac{\varepsilon(\mathit{\Gamma}) - \boldsymbol{V}\cdot\boldsymbol{p}}{k_{\rm B}T}\bigg) \tag{8.51}$$

also satisfies detailed balance, for any velocity parameter $\boldsymbol{V}$. This distribution is appropriate for gases which are flowing with average particle $\boldsymbol{V}$. Note that

$$\frac{\partial}{\partial \boldsymbol{p}}\big(\varepsilon - \boldsymbol{V}\cdot\boldsymbol{p}\big) = \boldsymbol{v} - \boldsymbol{V} \quad , \tag{8.52}$$

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, $\varGamma^{\mathcal{P}} = (-\boldsymbol{p}, \boldsymbol{L})$. Under the combined operation of $\mathcal{C} = \mathcal{P}\mathcal{T}$, we have $\varGamma^{\mathcal{C}} = (\boldsymbol{p}, -\boldsymbol{L})$. If the microscopic Hamiltonian is invariant under C, then we must have

$$w\big(\varGamma', \varGamma'_1 \,\big|\, \varGamma, \varGamma_1\big) = w\big(\varGamma^{c}, \varGamma^{c}_1 \,\big|\, \varGamma'^{c}, \varGamma'^{c}_1\big) \quad . \tag{8.53}$$

For point particles, invariance under T and P then means

$$w(\boldsymbol{p}', \boldsymbol{p}'_1 \,|\, \boldsymbol{p}, \boldsymbol{p}_1) = w(\boldsymbol{p}, \boldsymbol{p}_1 \,|\, \boldsymbol{p}', \boldsymbol{p}'_1) \quad , \tag{8.54}$$

and therefore the collision integral takes the simplified form,

$$\frac{Df(\boldsymbol{p})}{Dt} = \int\! d^3\!p_1 \!\int\! d^3\!p' \!\int\! d^3\!p'_1 \; w(\boldsymbol{p}', \boldsymbol{p}'_1 \,|\, \boldsymbol{p}, \boldsymbol{p}_1)\, \Big\{ f(\boldsymbol{p}')\, f(\boldsymbol{p}'_1) - f(\boldsymbol{p})\, f(\boldsymbol{p}_1) \Big\} = \bigg(\frac{df}{dt}\bigg)_{\!\rm coll} \quad , \tag{8.55}$$

where we have suppressed both $\boldsymbol{r}$ and t variables.

The most general statement of detailed balance is

$$\frac{f^0(\varGamma')\, f^0(\varGamma'_1)}{f^0(\varGamma)\, f^0(\varGamma_1)} = \frac{w\big(\varGamma', \varGamma'_1 \,\big|\, \varGamma, \varGamma_1\big)}{w\big(\varGamma, \varGamma_1 \,\big|\, \varGamma', \varGamma'_1\big)} \quad . \tag{8.56}$$

Under this condition, the collision term vanishes for $f = f^0$, which is the equilibrium distribution.

8.3.6 Kinematics and cross section

We can rewrite eqn. 8.55 in the form

$$\frac{Df(\boldsymbol{p})}{Dt} = \int\! d^3\!p_1 \!\int\! d\varOmega \; |\boldsymbol{v} - \boldsymbol{v}_1| \, \frac{d\sigma}{d\varOmega} \Big\{ f(\boldsymbol{p}')\, f(\boldsymbol{p}'_1) - f(\boldsymbol{p})\, f(\boldsymbol{p}_1) \Big\} \quad , \tag{8.57}$$

where $\frac{d\sigma}{d\varOmega}$ is the *differential scattering cross section*. Recall the definition from scattering theory in classical mechanics,

$$d\sigma = \frac{\text{\# of particles scattered into solid angle } d\varOmega \text{ per unit time}}{\text{incident flux}} \quad . \tag{8.58}$$

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 ϑ^{-4} as $\vartheta \to 0$.

8.3.7 $\mathcal{H}$-theorem

To peek ahead, we are about to prove the following. Let

$$\begin{aligned} h(\boldsymbol{r},t) &= \int\! d^3\!p\, f(\boldsymbol{r},\boldsymbol{p},t)\, \log\big[f(\boldsymbol{r},\boldsymbol{p},t)/f^0\big] \\ j(\boldsymbol{r},t) &= \int\! d^3\!p\, f(\boldsymbol{r},\boldsymbol{p},t)\, \log\big[f(\boldsymbol{r},\boldsymbol{p},t)/f^0\big]\, \frac{d\boldsymbol{r}}{dt} \quad . \end{aligned} \tag{8.61}$$

Here f^0 can be any constant which has the appropriate dimensions of A^{-3}, where A stands for 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,

$$\mathcal{H}(t) = \int\! d^3\!r\; h(\boldsymbol{r},t) \quad \Rightarrow \quad \frac{d\mathcal{H}}{dt} \le 0 \quad . \tag{8.63}$$

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

$$\rho_\varphi(\boldsymbol{r},t) \equiv \int\! d^3\!p\, f(\boldsymbol{r},\boldsymbol{p},t)\, \varphi\big(f(\boldsymbol{r},\boldsymbol{p},t)\big) \quad , \tag{8.64}$$

where $f = f(\boldsymbol{r},\boldsymbol{p},t)$ and $\varphi(f)$ is arbitrary. At this point, $\varphi(\boldsymbol{p},f)$ is arbitrary. We now compute

$$\begin{aligned} \frac{\partial \rho_\varphi}{\partial t} &= \int d^3\!p\, \frac{\partial (f\varphi)}{\partial t} = \int d\Gamma\, \frac{\partial (f\varphi)}{\partial f}\,\frac{\partial f}{\partial t} \\ &= \int d^3\!p\, \frac{\partial (f\varphi)}{\partial f} \left\{ -\,\dot{\boldsymbol{r}}\cdot\frac{\partial f}{\partial \boldsymbol{r}} - \dot{\boldsymbol{p}}\cdot\frac{\partial f}{\partial \boldsymbol{p}} + \left(\frac{df}{dt}\right)_{\rm coll}\right\} \\ &= \int d^3\!p \left\{ -\,\dot{\boldsymbol{r}}\cdot\frac{\partial (f\varphi)}{\partial \boldsymbol{r}} - \dot{\boldsymbol{p}}\cdot\frac{\partial (f\varphi)}{\partial \boldsymbol{p}} + \frac{\partial (f\varphi)}{\partial f}\left(\frac{df}{dt}\right)_{\rm coll}\right\} \end{aligned} \tag{8.65}$$

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

$$\begin{aligned} \frac{\partial \rho_\varphi}{\partial t} &= \int d^3\!p \left\{ -\,\dot{\boldsymbol{r}}\cdot\frac{\partial (f\varphi)}{\partial \boldsymbol{r}} - \frac{\partial \dot{\boldsymbol{r}}}{\partial \boldsymbol{r}}\,(f\varphi) + \frac{\partial (f\varphi)}{\partial f}\left(\frac{df}{dt}\right)_{\rm coll}\right\} \\ &= -\frac{\partial}{\partial \boldsymbol{r}} \int d^3\!p\, f\varphi\, \dot{\boldsymbol{r}} + \int d^3\!p\, \frac{\partial (f\varphi)}{\partial f}\left(\frac{df}{dt}\right)_{\rm coll} \quad . \end{aligned} \tag{8.66}$$

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

$$\begin{aligned} \boldsymbol{j}_\varphi(\boldsymbol{r},t) &= \int d^3\!p\, f(\boldsymbol{r},\boldsymbol{p},t)\,\varphi\big(f(\boldsymbol{r},\boldsymbol{p},t)\big)\,\boldsymbol{v}(\boldsymbol{p}) \\ \sigma_\varphi(\boldsymbol{r},t) &= \int d^3\!p\, \frac{\partial (f\varphi)}{\partial f}\bigg|_{f(\boldsymbol{r},\boldsymbol{p},t)} \left(\frac{df}{dt}\right)_{\rm coll} \end{aligned} \tag{8.68}$$

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 σ_φ under the assumption that f satisfies the Boltzmann equation with two particle scattering. Thus,

$$\begin{aligned} \sigma_\varphi(\boldsymbol{r},t) &= \int d^3\!p \int d^3\!p_1 \int d^3\!p' \int d^3\!p'_1 \Big\{ w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big)\, f(\boldsymbol{p}) f(\boldsymbol{p}_1)\,\chi(\boldsymbol{p}) - w\big(\boldsymbol{p},\boldsymbol{p}_1\,|\,\boldsymbol{p}',\boldsymbol{p}'_1\big)\, f(\boldsymbol{p}') f(\boldsymbol{p}'_1)\,\chi(\boldsymbol{p}')\Big\} \\ &= \int d^3\!p \int d^3\!p_1 \int d^3\!p' \int d^3\!p'_1\, w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big)\, f(\boldsymbol{p}) f(\boldsymbol{p}_1)\,\big(\chi(\boldsymbol{p}) - \chi'(\boldsymbol{p})\big) \quad , \end{aligned} \tag{8.69}$$

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

$$w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big) = w\big(\boldsymbol{p}'_1,\boldsymbol{p}'\,|\,\boldsymbol{p}_1,\boldsymbol{p}\big) \quad , \tag{8.71}$$

which allows us to write

$$\sigma = \tfrac{1}{2}\int\! d^3\!p\!\int\! d^3\!p_1\!\int\! d^3\!p'\!\int\! d^3\!p'_1\; w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big)\, f(\boldsymbol{p}) f(\boldsymbol{p}_1)\,\big(\chi(\boldsymbol{p})+\chi(\boldsymbol{p}_1)-\chi(\boldsymbol{p}')-\chi(\boldsymbol{p}'_1)\big) \quad , \tag{8.72}$$

This shows that $\sigma_\varphi = 0$ if $\chi(\boldsymbol{p})$ is a collisional invariant.

Now let us fix $\varphi(f) = \log(f/f^0)$ and evaluate the source term $\sigma \equiv \sigma_{\varphi=\log(f/f^0)}$. We have

$$\sigma = -\tfrac{1}{2}\int\! d^3\!p\!\int\! d^3\!p_1\!\int\! d^3\!p'\!\int\! d^3\!p'_1\; w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big)\, f(\boldsymbol{p}') f(\boldsymbol{p}'_1)\cdot x(\boldsymbol{p},\boldsymbol{p}_1\,|\,\boldsymbol{p}',\boldsymbol{p}'_1)\log x(\boldsymbol{p},\boldsymbol{p}_1\,|\,\boldsymbol{p}',\boldsymbol{p}'_1) \quad , \tag{8.73}$$

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

$$\int\! d^3\!p'\!\int\! d^3\!p'_1\; w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big) = \int\! d^3\!p'\!\int\! d^3\!p'_1\; w\big(\boldsymbol{p},\boldsymbol{p}_1\,|\,\boldsymbol{p}',\boldsymbol{p}'_1\big) \tag{8.74}$$

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

$$\begin{aligned} 0 &= \int\! d^3\!p\!\int\! d^3\!p_1\!\int\! d^3\!p'\!\int\! d^3\!p'_1\; w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big)\,\Big(f(\boldsymbol{p})f(\boldsymbol{p}_1) - f(\boldsymbol{p}')f(\boldsymbol{p}'_1)\Big) \\ &= \int\! d^3\!p\!\int\! d^3\!p_1\!\int\! d^3\!p'\!\int\! d^3\!p'_1\; w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big)\, f(\boldsymbol{p}')f(\boldsymbol{p}'_1)\,\Big\{x(\boldsymbol{p},\boldsymbol{p}_1\,|\,\boldsymbol{p}',\boldsymbol{p}'_1) - 1\Big\} \quad . \end{aligned} \tag{8.75}$$

Multiplying this result by $\frac{1}{2}$ and adding it to the previous equation for $\dot{h}$, we arrive at our final result,

$$\sigma = -\tfrac{1}{2}\int\! d^3\!p\!\int\! d^3\!p_1\!\int\! d^3\!p'\!\int\! d^3\!p'_1\; w\big(\boldsymbol{p}',\boldsymbol{p}'_1\,|\,\boldsymbol{p},\boldsymbol{p}_1\big)\, f(\boldsymbol{p}')f(\boldsymbol{p}'_1)\;\,(x\log x - x + 1) \quad , \tag{8.76}$$

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

$$\frac{\partial h(\boldsymbol{r},t)}{\partial t} + \boldsymbol{\nabla}\cdot j(\boldsymbol{r},t) = \sigma(\boldsymbol{r},t) \le 0 \quad . \tag{8.77}$$

Boltzmann's $\mathcal{H}$ function is the space integral of the local density $h(\boldsymbol{r})$: $\mathcal{H} = \int\! d^3\!r\; h(\boldsymbol{r})$.

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 ε are the only collisional invariants, then $\log f^0$ must be expressible in terms of them. Thus,

$$-\log f^0 = \frac{\varepsilon(\boldsymbol{p}) - \boldsymbol{V}\cdot\boldsymbol{p} - \mu}{k_{\rm B}T} \quad , \tag{8.80}$$

where μ, $\boldsymbol{V}$, and T are constants which parameterize the equilibrium distribution $f^0(\boldsymbol{p})$, corresponding to the chemical potential, flow velocity, and temperature, respectively.

---

[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.

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 is 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 δf must be added in order to obtain a solution. The Boltzmann equation is written

$$\left(\frac{\partial}{\partial t} + \frac{\boldsymbol{p}}{m} \cdot \frac{\partial}{\partial \boldsymbol{r}} + \boldsymbol{F} \cdot \frac{\partial}{\partial \boldsymbol{p}}\right)(f^0 + \delta f) = \left(\frac{df}{dt}\right)_{\rm coll} \quad . \tag{8.81}$$

The RHS of this equation must be of order δf 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

$$f^0(\boldsymbol{r}, \varGamma) = C \exp\left(\frac{\mu - \varepsilon(\varGamma) + \boldsymbol{V} \cdot \boldsymbol{p}}{k_{\rm B} T}\right) \quad , \tag{8.82}$$

where $\mu = \mu(\boldsymbol{r}, t)$, $T = T(\boldsymbol{r}, t)$, and $\boldsymbol{V} = \boldsymbol{V}(\boldsymbol{r}, t)$ 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. We also have

$$-\frac{\partial f^0}{\partial \varepsilon} = \frac{f^0}{k_{\rm B} T} \quad , \tag{8.85}$$

which we shall invoke further on below. When f^0 is the Maxwell-Boltzmann distribution, we have

$$f^0(\boldsymbol{p}) = n\,(2\pi m k_{\rm B} T)^{-3/2}\, e^{-\boldsymbol{p}^2/2mk_{\rm B}T} \tag{8.86}$$

is normalized so that $\int\! d^3\!r \int\! d^3\!p\; f^0(\boldsymbol{p}) = N$.

We wish to evaluate one of the contributions to the LHS of eqn. 8.81:

$$\begin{aligned} \frac{\partial f^0}{\partial t} + \frac{\boldsymbol{p}}{m} \cdot \frac{\partial f^0}{\partial \boldsymbol{r}} + \boldsymbol{F} \cdot \frac{\partial f^0}{\partial \boldsymbol{p}} = \left(-\frac{\partial f^0}{\partial \varepsilon}\right) \bigg\{ &\frac{1}{n}\,\frac{\partial p}{\partial t} + \frac{\varepsilon - h}{T}\,\frac{\partial T}{\partial t} + m\boldsymbol{v} \cdot \Big[(\boldsymbol{v} \cdot \boldsymbol{\nabla})\,\boldsymbol{V}\Big] \\ &+ \boldsymbol{v} \cdot \left(m\,\frac{\partial \boldsymbol{V}}{\partial t} + \frac{1}{n}\,\boldsymbol{\nabla} p\right) + \frac{\varepsilon - h}{T}\,\boldsymbol{v} \cdot \boldsymbol{\nabla} T - \boldsymbol{F} \cdot \boldsymbol{v} \bigg\} \quad . \end{aligned} \tag{8.87}$$

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,

$$\frac{\partial \log n}{\partial t} = \frac{\partial \log p}{\partial t} - \frac{\partial \log T}{\partial t} = -\boldsymbol{\nabla}\cdot\boldsymbol{V} \quad , \tag{8.88}$$

where we have invoked the ideal gas law $n = p/k_{\mathrm{B}}T$ 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

$$\frac{\partial(ns)}{\partial t} + \boldsymbol{\nabla}\cdot(ns\boldsymbol{V}) = n\bigg(\frac{\partial s}{\partial t} + \boldsymbol{V}\cdot\boldsymbol{\nabla} s\bigg) + s\bigg(\frac{\partial n}{\partial t} + \boldsymbol{\nabla}\cdot(n\boldsymbol{V})\bigg) = 0 \quad . \tag{8.89}$$

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

$$ds = \bigg(\frac{\partial s}{\partial T}\bigg)_{\!p} dT + \bigg(\frac{\partial s}{\partial p}\bigg)_{\!T} dp = \frac{c_p}{T}\,dT - \frac{k_{\mathrm{B}}}{p}\,dp \quad , \tag{8.90}$$

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,

$$\frac{c_p}{k_{\mathrm{B}}}\,\frac{\partial \log T}{\partial t} - \frac{\partial \log p}{\partial t} = 0 \quad . \tag{8.91}$$

We now have in eqns. 8.88 and 8.91 two equations in the two unknowns $\partial \log T/\partial t$ and $\partial \log p/\partial t$, yielding

$$\frac{\partial \log T}{\partial t} = -\frac{k_{\mathrm{B}}}{c_V}\,\boldsymbol{\nabla}\cdot\boldsymbol{V} \quad , \qquad \frac{\partial \log p}{\partial t} = -\frac{c_p}{c_V}\,\boldsymbol{\nabla}\cdot\boldsymbol{V} \quad . \tag{8.92}$$

Thus eqn. 8.87 becomes

$$\frac{\partial f^0}{\partial t} + \frac{\boldsymbol{p}}{m}\cdot\frac{\partial f^0}{\partial \boldsymbol{r}} + \boldsymbol{F}\cdot\frac{\partial f^0}{\partial \boldsymbol{p}} = \bigg(-\frac{\partial f^0}{\partial \varepsilon}\bigg)\bigg\{\frac{\varepsilon - h}{T}\,\boldsymbol{v}\cdot\boldsymbol{\nabla} T + m\,v_\alpha v_\beta\,\mathcal{Q}_{\alpha\beta} + \frac{h - Tc_p - \varepsilon}{c_V/k_{\mathrm{B}}}\,\boldsymbol{\nabla}\cdot\boldsymbol{V} - \boldsymbol{F}\cdot\boldsymbol{v}\bigg\} \quad , \tag{8.93}$$

where $\varepsilon = \varepsilon(\varGamma)$ and

$$\mathcal{Q}_{\alpha\beta} = \frac{1}{2}\bigg(\frac{\partial V_\alpha}{\partial x_\beta} + \frac{\partial V_\beta}{\partial x_\alpha}\bigg) \quad . \tag{8.94}$$

Therefore, the Boltzmann equation takes the form

$$\frac{\partial\,\delta f}{\partial t} + \bigg\{\frac{\varepsilon(\varGamma) - h}{T}\,\boldsymbol{v}\cdot\boldsymbol{\nabla} T + m\,v_\alpha v_\beta\,\mathcal{Q}_{\alpha\beta} - \frac{\varepsilon(\varGamma) - h + Tc_p}{c_V/k_{\mathrm{B}}}\,\boldsymbol{\nabla}\cdot\boldsymbol{V} - \boldsymbol{F}\cdot\boldsymbol{v}\bigg\}\,\frac{f^0}{k_{\mathrm{B}}T} = \bigg(\frac{df}{dt}\bigg)_{\mathrm{coll}} \quad . \tag{8.95}$$

---

[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.

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 δf 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,

$$\left(\frac{df}{dt}\right)_{\rm coll}=-\frac{f-f^0}{\tau}=-\frac{\delta f}{\tau}\quad. \tag{8.96}$$

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 τ 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

$$\frac{\partial\,\delta f}{\partial t}=-\frac{\delta f}{\tau}\qquad\Longrightarrow\qquad \delta f(\boldsymbol{r},\boldsymbol{p},t)=\delta f(\boldsymbol{r},\boldsymbol{p},0)\,e^{-t/\tau}\quad. \tag{8.97}$$

The distribution f then relaxes to the equilibrium distribution f^0 on a time scale τ. We note that this approximation is obviously flawed in that all quantities – even the collisional invariants – relax to their equilibrium values on the scale τ. 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.

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

$$\langle\,|\boldsymbol{v}-\boldsymbol{v}'|\,\rangle=\int\! d^3\!v\!\int\! d^3\!v'\;P(\boldsymbol{v})\,P(\boldsymbol{v}')\,|\boldsymbol{v}-\boldsymbol{v}'|\quad, \tag{8.98}$$

where $P(\boldsymbol{v})$ is the Maxwell velocity distribution,

$$P(\boldsymbol{v})=\left(\frac{m}{2\pi k_{\rm B}T}\right)^{3/2}\exp\left(-\frac{m\boldsymbol{v}^2}{2k_{\rm B}T}\right)\quad, \tag{8.99}$$

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

$$\bar v_{\rm rel}\equiv\langle\,|\boldsymbol{v}-\boldsymbol{v}'|\,\rangle=\frac{4}{\sqrt{\pi}}\left(\frac{k_{\rm B}T}{m}\right)^{1/2}\quad. \tag{8.100}$$

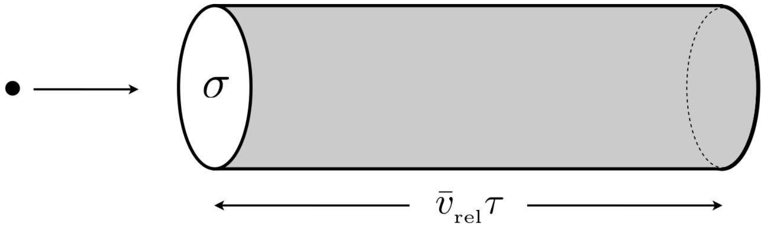


Figure 8.3: Graphic representation of the equation $n\,\sigma\,\bar{v}_{\rm rel}\,\tau = 1$, which yields the scattering time τ in terms of the number density n, average particle pair relative velocity $\bar{v}_{\rm rel}$, and two-particle total scattering cross section σ. The equation says that on average there must be one particle within the tube.

Note that $\bar{v}_{\rm rel} = \sqrt{2}\,\bar{v}$, where $\bar{v}$ is the average particle speed. Let σ 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

$$\frac{1}{\tau} = n\,\bar{v}_{\rm rel}\,\sigma \quad . \tag{8.101}$$

The particle *mean free path* is then

$$\ell = \bar{v}\,\tau = \frac{1}{\sqrt{2}\,n\,\sigma} \quad . \tag{8.102}$$

While the scattering length is not temperature-dependent within this formalism, the scattering time *is* T-dependent, with

$$\tau(T) = \frac{1}{n\,\bar{v}_{\rm rel}\,\sigma} = \frac{\sqrt{\pi}}{4n\sigma}\left(\frac{m}{k_{\rm B}T}\right)^{1/2} \quad . \tag{8.103}$$

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. 8.95, and using the relaxation time approximation for the collision integral, we have

$$\delta f = -\frac{\tau(\varepsilon - c_p\,T)}{k_{\rm B}T^2}\,(\boldsymbol{v}\cdot\boldsymbol{\nabla} T)\,f^0 \quad . \tag{8.104}$$

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:

$$\rho_A(\boldsymbol{r},t) = \int\! d^3\!p\,A(\boldsymbol{r},\boldsymbol{p})\,f(\boldsymbol{r},\boldsymbol{p},t) \quad , \tag{8.105}$$

For the energy (thermal) current, we let $A = \varepsilon\,v_\alpha = \varepsilon\,p_\alpha/m$, in which case $\rho_A = j_\alpha$. Note that $\int d^3\!p\,\boldsymbol{p}\,f^0 = 0$ since f^0 is isotropic in $\boldsymbol{p}$ even when μ and T depend on $\boldsymbol{r}$. Thus, only δf enters into the calculation of the various currents. Thus, the energy (thermal) current is

$$j_\varepsilon^\alpha(\boldsymbol{r}) = \int\! d^3\!p\, \varepsilon\, v^\alpha\, \delta f = -\frac{n\tau}{k_\mathrm{B} T^2}\, \langle\, v^\alpha v^\beta\, \varepsilon\, (\varepsilon - c_p\, T)\,\rangle\, \frac{\partial T}{\partial x^\beta} \quad , \tag{8.106}$$

where the repeated index β is summed over, and where momentum averages are defined relative to the equilibrium distribution, *i.e.*

$$\langle\, \phi(\boldsymbol{p})\,\rangle = \int\! d^3\!p\, \phi(\boldsymbol{p})\, f^0(\boldsymbol{p}) \Big/ \!\! \int\! d^3\!p\, f^0(\boldsymbol{p}) = \!\int\! d^3\!v\; P(\boldsymbol{v})\, \phi(m\boldsymbol{v}) \quad . \tag{8.107}$$

In this context, it is useful to invoke the identity $d^3\!p\, f^0(\boldsymbol{p}) = n\, d^3\!v\; P(\boldsymbol{v})$, where

$$P(\boldsymbol{v}) = \left(\frac{m}{2\pi k_\mathrm{B} T}\right)^{3/2} e^{-m(\boldsymbol{v}-\boldsymbol{V})^2/2k_\mathrm{B} T} \tag{8.108}$$

is the Maxwell velocity distribution.

Note that if $\phi = \phi(\varepsilon)$ is a function of the energy, and if $\boldsymbol{V} = 0$, then

$$d^3\!p\, f^0(\boldsymbol{p}) = n\, d^3\!v\; P(\boldsymbol{v}) = n\, \mathcal{P}(\varepsilon)\, d\varepsilon \quad , \tag{8.109}$$

where

$$\mathcal{P}(\varepsilon) = \frac{2}{\sqrt{\pi}}\, (k_\mathrm{B} T)^{-3/2}\, \varepsilon^{1/2}\, e^{-\varepsilon/k_\mathrm{B} T} \quad , \tag{8.110}$$

is the Maxwellian distribution expressed in terms of ε, and is normalized according to $\int\limits_0^\infty \! d\varepsilon\, \mathcal{P}(\varepsilon) = 1$.[6] Averages with respect to this distribution are given by

$$\langle\, \phi(\varepsilon)\,\rangle = \int\limits_0^\infty \! d\varepsilon\; \phi(\varepsilon)\, \mathcal{P}(\varepsilon) = \frac{2}{\sqrt{\pi}}\, (k_\mathrm{B} T)^{-3/2} \int\limits_0^\infty \! d\varepsilon\; \varepsilon^{1/2}\, \phi(\varepsilon)\, e^{-\varepsilon/k_\mathrm{B} T} \quad . \tag{8.111}$$

For homogeneous $\phi(s) = \varepsilon^s$ we have

$$\langle\, \varepsilon^s\,\rangle = \frac{2}{\sqrt{\pi}}\, \Gamma\big(s + \tfrac{3}{2}\big)\, (k_\mathrm{B} T)^s \quad . \tag{8.112}$$

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. 8.106.. We then have $\boldsymbol{j}_\varepsilon = -\kappa\, \boldsymbol{\nabla} T$, with

$$\kappa = \frac{2n\tau}{3mk_\mathrm{B} T^2}\, \langle\, \varepsilon^2 \big(\varepsilon - c_p\, T\big)\,\rangle = \frac{5n\tau k_\mathrm{B}^2 T}{2m} = \frac{\pi}{8} n\ell\bar{v}\, c_p \quad , \tag{8.113}$$

where $c_p = \frac{5}{2} k_\mathrm{B}$ and $\bar{v}^2 = 8k_\mathrm{B} T/\pi m$. The quantity κ is called the *thermal conductivity*. Note that $\kappa \propto T^{1/2}$.

---

[6] $\mathcal{P}(\varepsilon)$ can also be recognized as the normalized product of the Boltzmann energy distribution $\exp(-\varepsilon/k_\mathrm{B} T)$ and the density of states $g(\varepsilon)$. In $d = 3$ we have $g(\varepsilon) \propto \varepsilon^{1/2}$.

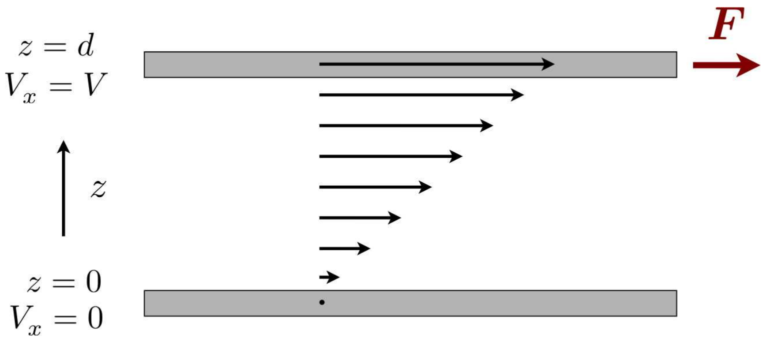


Figure 8.4: *Gedankenexperiment* to measure shear viscosity η in a fluid. The lower plate is fixed. The viscous drag force per unit area on the upper plate is $F_{\text{drag}}/A = -\eta V/d$. This must be balanced by an applied force F.

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 $\boldsymbol{F} = F\hat{\boldsymbol{x}}$ 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 $F_{\text{drag}}/A = -\eta V/d$, where $V/d = \partial V_x/\partial z$ is the velocity gradient and η is the *shear viscosity*. In steady state, the applied force balances the drag force, *i.e.* $F + F_{\text{drag}} = 0$. Clearly in the steady state the net momentum density of the fluid does not change, and is given by $\frac{1}{2}\rho V\hat{\boldsymbol{x}}$, where ρ 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_α in the direction of increasing x_β. For a gas in motion with average velocity $\boldsymbol{V}$, we have

$$\begin{aligned}\Pi_{\alpha\beta} &= nm\,\langle\,(V_\alpha + v'_\alpha)(V_\beta + v'_\beta)\,\rangle \\ &= nm\,V_\alpha V_\beta + nm\,\langle\, v'_\alpha v'_\beta\,\rangle \\ &= nm\,V_\alpha V_\beta + \tfrac{1}{3} nm\,\langle\, \boldsymbol{v}'^2\,\rangle\,\delta_{\alpha\beta} = \rho\,V_\alpha V_\beta + p\,\delta_{\alpha\beta} \quad ,\end{aligned} \tag{8.114}$$

where $\boldsymbol{v}'$ is the particle velocity in a frame moving with velocity $\boldsymbol{V}$, and where we have invoked the ideal gas law $p = nk_{\text{B}}T$. The mass density is $\rho = nm$.

When $\boldsymbol{V}$ is spatially varying,

$$\Pi_{\alpha\beta} = p\,\delta_{\alpha\beta} + \rho\,V_\alpha V_\beta - \tilde{\sigma}_{\alpha\beta} \quad , \tag{8.115}$$

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:

$$\begin{aligned}\tilde{\sigma}_{\alpha\beta} &= \eta\bigg(\frac{\partial V_\alpha}{\partial x_\beta} + \frac{\partial V_\beta}{\partial x_\alpha} - \tfrac{2}{3}\,\boldsymbol{\nabla}\!\cdot\!\boldsymbol{V}\,\delta_{\alpha\beta}\bigg) + \zeta\,\boldsymbol{\nabla}\!\cdot\!\boldsymbol{V}\,\delta_{\alpha\beta} \\ &= 2\eta\,\Big(\mathcal{Q}_{\alpha\beta} - \tfrac{1}{3}\,\mathsf{Tr}\,(\mathcal{Q})\,\delta_{\alpha\beta}\Big) + \zeta\ \mathsf{Tr}\,(\mathcal{Q})\,\delta_{\alpha\beta} \quad .\end{aligned} \tag{8.116}$$

The coefficient of the traceless component is η, known as the *shear viscosity*. The coefficient of the component proportional to the identity is ζ, known as the *bulk viscosity*. The full stress tensor $\sigma_{\alpha\beta}$ contains a contribution from the pressure:

$$\sigma_{\alpha\beta} = -p\,\delta_{\alpha\beta} + \tilde{\sigma}_{\alpha\beta} \quad . \tag{8.117}$$

The differential force dF_α that a fluid exerts on 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 β. We will now compute the shear viscosity η using the Boltzmann equation in the relaxation time approximation.

Appealing again to eqn. 8.95, with $\boldsymbol{F} = 0$ and $h = c_p T$, we find

$$\delta f = -\frac{\tau}{k_{\rm B}T}\Bigg\{m\,v_\alpha v_\beta\,\mathcal{Q}_{\alpha\beta} + \frac{\varepsilon - c_p\,T}{T}\,\boldsymbol{v}\cdot\boldsymbol{\nabla}T - \frac{\varepsilon}{c_V/k_{\rm B}}\,\boldsymbol{\nabla}\cdot\boldsymbol{V}\Bigg\}\,f^0 \quad . \tag{8.118}$$

We assume $\boldsymbol{\nabla}T = \boldsymbol{\nabla}\cdot\boldsymbol{V} = 0$, and we compute the momentum flux:

$$\begin{aligned}\Pi_{xz} &= n\!\int\! d^3\!p\;p_x v_z\,\delta f = -\frac{nm^2\tau}{k_{\rm B}T}\,\mathcal{Q}_{\alpha\beta}\,\langle\,v_x\,v_z\,v_\alpha\,v_\beta\,\rangle \\ &= -\frac{n\tau}{k_{\rm B}T}\bigg(\frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x}\bigg)\,\langle\,mv_x^2\cdot mv_z^2\,\rangle = -n\tau k_{\rm B}T\bigg(\frac{\partial V_z}{\partial x} + \frac{\partial V_x}{\partial z}\bigg) \quad .\end{aligned} \tag{8.119}$$

Thus, if $V_x = V_x(z)$, we have

$$\Pi_{xz} = -n\tau k_{\rm B}T\,\frac{\partial V_x}{\partial z} \tag{8.120}$$

from which we read off the viscosity,

$$\eta = nk_{\rm B}T\tau = \frac{\pi}{8}nm\ell\bar{v} \quad . \tag{8.121}$$

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 $\partial V_x/\partial z$

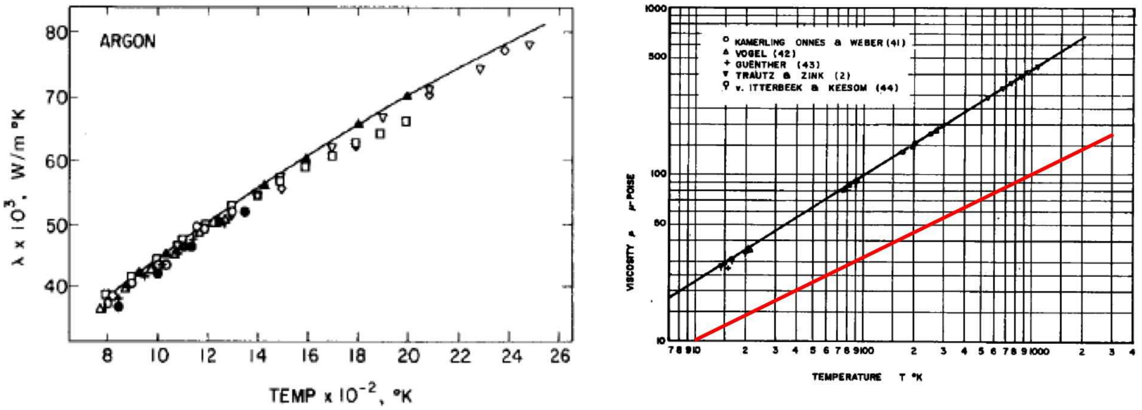


Figure 8.5: Left: thermal conductivity (λ in figure) of Ar between $T = 800\,\mathrm{K}$ and $T = 2600\,\mathrm{K}$. The best fit to a single power law $\lambda = aT^b$ results in $b = 0.651$. Source: G. S. Springer and E. W. Wingeier, *J. Chem Phys.* **59**, 1747 (1972). Right: log-log plot of shear viscosity (μ 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).

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.

8.5.5 Oscillating external force

Suppose a uniform oscillating external force $\boldsymbol{F}_{\text{ext}}(t) = \boldsymbol{F}\, e^{-i\omega t}$ is applied. For a system of charged particles, this force would arise from an external electric field $\boldsymbol{F}_{\text{ext}} = q\boldsymbol{E}\, e^{-i\omega t}$, where q is the charge of each particle. We'll assume $\boldsymbol{\nabla} T = 0$. The Boltzmann equation is then written

$$\frac{\partial f}{\partial t} + \frac{\boldsymbol{p}}{m} \cdot \frac{\partial f}{\partial \boldsymbol{r}} + \boldsymbol{F} e^{-i\omega t} \cdot \frac{\partial f}{\partial \boldsymbol{p}} = -\frac{f - f^0}{\tau} \quad . \tag{8.122}$$

We again write $f = f^0 + \delta f$, and we assume δf is spatially constant. Thus,

$$\frac{\partial\, \delta f}{\partial t} + \boldsymbol{F} e^{-i\omega t} \cdot \boldsymbol{v}\, \frac{\partial f^0}{\partial \varepsilon} = -\frac{\delta f}{\tau} \quad . \tag{8.123}$$

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

$$\delta f(t) = -\frac{\tau\, e^{-i\omega t}}{1 - i\omega\tau}\, \frac{\partial f^0}{\partial \varepsilon}\, \boldsymbol{F} \cdot \boldsymbol{v} \quad . \tag{8.124}$$

We now compute the particle current:

$$\begin{aligned} j_\alpha(\boldsymbol{r},t) &= \int\! d^3\!p\, \boldsymbol{v}\, \delta f = \frac{\tau\, e^{-i\omega t}}{1-i\omega\tau} \cdot \frac{F_\beta}{k_{\rm B}T}\!\int\! d^3\!p\, f^0(\boldsymbol{p})\, v_\alpha\, v_\beta \\ &= \frac{\tau\, e^{-i\omega t}}{1-i\omega\tau} \cdot \frac{nF_\alpha}{3k_{\rm B}T}\!\int\! d^3\!v\; P(\boldsymbol{v})\, \boldsymbol{v}^2 = \frac{n\tau}{m} \cdot \frac{F_\alpha\, e^{-i\omega t}}{1-i\omega\tau} \quad . \end{aligned} \tag{8.125}$$

If the particles are electrons, with charge $q = -e$, then the electrical current is $(-e)$ times the particle current. We then obtain

$$j^{(\rm elec)}_\alpha(t) = \frac{ne^2\tau}{m} \cdot \frac{E_\alpha\, e^{-i\omega t}}{1-i\omega\tau} \equiv \sigma_{\alpha\beta}(\omega)\, E_\beta\, e^{-i\omega t} \quad , \tag{8.126}$$

where

$$\sigma_{\alpha\beta}(\omega) = \frac{ne^2\tau}{m} \cdot \frac{1}{1-i\omega\tau}\, \delta_{\alpha\beta} \tag{8.127}$$

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 μ only, and the final result for the conductivity tensor $\sigma_{\alpha\beta}(\omega)$ is unchanged.

8.5.6 Quick and dirty calculation of transport coefficiencs

Suppose we have some averaged intensive quantity ϕ 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 ϕ across some surface whose equation is $dz = 0$. If the mean free path is ℓ, then the value of ϕ for particles crossing this surface in the $+\hat{\boldsymbol{z}}$ direction is $\phi(z - \ell\cos\theta)$, where θ 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 ϕ through this surface is then

$$\begin{aligned} \boldsymbol{j}_\phi &= n\hat{\boldsymbol{z}} \int\limits_{v_z>0}\! d^3\!v\, P(\boldsymbol{v})\, v_z\, \phi(z - \ell\cos\theta) + n\hat{\boldsymbol{z}} \int\limits_{v_z<0}\! d^3\!v\, P(\boldsymbol{v})\, v_z\, \phi(z + \ell\cos\theta) \\ &= -n\ell\, \frac{\partial\phi}{\partial z}\, \hat{\boldsymbol{z}} \!\int\! d^3\!v\; P(\boldsymbol{v})\, \frac{v_z^2}{v} = -\tfrac{1}{3} n\bar{v}\ell\, \frac{\partial\phi}{\partial z}\, \hat{\boldsymbol{z}} \quad , \end{aligned} \tag{8.128}$$

where $\bar{v} = (8k_{\rm B}T/\pi m)^{1/2}$ is the average particle speed. If the z-dependence of ϕ comes through the dependence of ϕ on the local temperature T, then we have

$$\boldsymbol{j}_\phi = -\tfrac{1}{3}\, n\ell\bar{v}\, \frac{\partial\phi}{\partial T}\, \boldsymbol{\nabla} T \equiv -K\, \boldsymbol{\nabla} T \quad , \tag{8.129}$$

where

$$K = \tfrac{1}{3} n\ell\bar{v}\, \frac{\partial\phi}{\partial T} \tag{8.130}$$

is the transport coefficient. If $\phi = \langle\varepsilon\rangle$, then $\partial\phi/\partial T = c_p$, where c_p is the heat capacity per particle at constant pressure. We then find $\boldsymbol{j}_\varepsilon = -\kappa\,\boldsymbol{\nabla} T$ with thermal conductivity $\kappa = \frac{1}{3} n\ell\bar{v}\,c_p$. Our Boltzmann equation calculation yielded the same result (eqn. 8.113), but with a prefactor of $\frac{\pi}{8}$ instead of $\frac{1}{3}$.

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

$$j^z_{p_x} = \Pi_{xz} = -\tfrac{1}{3} nm\ell\bar{v}\,\frac{\partial V_x}{\partial z} \quad , \tag{8.131}$$

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. The ratio of the coefficients is $K_{\text{QDC}}/K_{\text{BRT}} = \frac{8}{3\pi} = 0.849$ in both cases[7].

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

$$nc_p\,\frac{\partial T}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j}_q = 0 \quad . \tag{8.132}$$

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,

$$\frac{\partial T}{\partial t} = \frac{\kappa}{nc_p}\,\nabla^2 T \quad . \tag{8.133}$$

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 $\kappa = nk_{\text{B}}T\tau c_p/m$. Thus, we find $a = k_{\text{B}}T\tau/m$ 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 $\eta = nk_{\text{B}}T\tau$ from the relaxation time approximation calculation, hence $\nu = k_{\text{B}}T\tau/m$. The ratio ν/a, called the *Prandtl number*, $\text{Pr} = \eta c_p/m\kappa$, is dimensionless. According to our calculations, $\text{Pr} = 1$. According to table 8.1, most monatomic gases have $\text{Pr} \approx \frac{2}{3}$.

---

[7] Here we abbreviate QDC for *'quick and dirty calculation'* and BRT for *'Boltzmann equation in the relaxation time approximation'*.

Gas	η (μPa $\cdot$ s)	κ (mW/m $\cdot$ K)	c_p/k_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 $T = 293\,\text{K}$ and $p = 1\,\text{atm}$. (Source: Table 1.1 of Smith and Jensen, with data for triatomic gases added.)

8.6 The Equations of Hydrodynamics

We now derive the equations governing fluid flow. The equations of mass and momentum balance are

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot(\rho\,\boldsymbol{V}) = 0 \tag{8.134}$$

$$\frac{\partial(\rho\,V_\alpha)}{\partial t} + \frac{\partial\Pi_{\alpha\beta}}{\partial x^\beta} = 0 \quad , \tag{8.135}$$

where

$$\Pi_{\alpha\beta} = \rho\,V_\alpha V_\beta + p\,\delta_{\alpha\beta} - \overbrace{\left\{\eta\left(\frac{\partial V_\alpha}{\partial x_\beta} + \frac{\partial V_\beta}{\partial x_\alpha} - \tfrac{2}{3}\,\boldsymbol{\nabla}\!\cdot\!\boldsymbol{V}\,\delta_{\alpha\beta}\right) + \zeta\,\boldsymbol{\nabla}\!\cdot\!\boldsymbol{V}\,\delta_{\alpha\beta}\right\}}^{\tilde\sigma_{\alpha\beta}} \quad . \tag{8.136}$$

Substituting the continuity equation into the momentum balance equation, one arrives at

$$\rho\,\frac{\partial\boldsymbol{V}}{\partial t} + \rho\,(\boldsymbol{V}\cdot\boldsymbol{\nabla})\boldsymbol{V} = -\boldsymbol{\nabla}p + \eta\,\nabla^2\boldsymbol{V} + (\zeta + \tfrac{1}{3}\eta)\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{V}) \quad , \tag{8.137}$$

which, together with continuity, are known as the *Navier-Stokes equations*. These equations are supplemented by an equation describing the conservation of energy,

$$T\frac{\partial s}{\partial T} + T\,\boldsymbol{\nabla}\cdot(s\boldsymbol{V}) = \tilde\sigma_{\alpha\beta}\,\frac{\partial V_\alpha}{\partial x^\beta} + \boldsymbol{\nabla}\cdot(\kappa\boldsymbol{\nabla}T) \quad , \tag{8.138}$$

where $\tilde\sigma_{\alpha\beta}$ is the viscosity stress tensor of eqn. 8.116 (also identified in eqn. 8.136 above). Note that the LHS of eqn. 8.137 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 inviscid Navier-Stokes equations,

$$\frac{\partial \boldsymbol{V}}{\partial t} + (\boldsymbol{V} \cdot \boldsymbol{\nabla})\boldsymbol{V} = -\frac{1}{\rho}\,\boldsymbol{\nabla} p + \nu\,\nabla^2 \boldsymbol{V} \quad . \tag{8.139}$$

Since $[\eta] = M/LT$ and $[\rho] = M/L^3$, we have $[\eta] = L^2/T$. At $T = 20^\circ\mathrm{C}$,

$$\nu_{\mathrm{H_2O}} = 10^{-2}\,\mathrm{cm}^2/\mathrm{sec} \qquad , \qquad \nu_{\mathrm{air}} = 0.15\,\mathrm{cm}^2/\mathrm{sec} \qquad , \qquad \nu_{\mathrm{glycerine}} = 6.8\,\mathrm{cm}^2/\mathrm{sec} \quad . \tag{8.140}$$