Statistical and Thermal Physics Using Python

Anders Malthe-Sørenssen[1]
Dag Kristian Dysthe[1]

[1]Department of Physics, University of Oslo, Norway

Sep 2, 2021

Preface

This book was developed as a series of lecture notes for the course Fys2160 Statistical and thermal physics at the University of Oslo, Norway. The motivation for the book project was to develop a text that integrates the use of computational methods deeply into the exposition, examples and exercises. In particular, the aim is to demonstrate the whole problem solution process, from the initial complex physics problem, through simplified models that have analytical solutions and to more complex and realistic models that defy analytical approaches, but that can be addressed using computational models. Our experience is that this approach more closely reflects the real work processes found in research and industry, and that the approach make the curriculum more open — there are more processes and phenomena that can be addressed, and there are many ways to do that. It also allows students to experience that various approaches have various advantages and disadvantages.

A particular challenge when introducing programming and computational methods is to find a reasonable and productive way to include the computer programs. Here, we have selected a particular programming language, python, and used this exclusively throughout the book. However, there are two versions of this book, one using Python and one using Matlab. Everything but the programs and the associated syntax-specific comments are the same.

We have also found it useful to include simplifications, analytical and computational methods in the same worked examples and exercises. This provides the students with a complete worked example that also can be

used as a basis for exploring new phenomena. The code can be found both in the book and can be downloaded from the book website[1].

All the examples and some of the exercises in the book have been compiled as iPython notebooks, and you can find and use the notebooks at the book notebook site[2].

Since the contents of this course is considered an advanced course, the tradition in most texts is to not include complete worked examples. We have found this practice unfortunate, since worked examples are powerful pedagogical tools, and research have shown that going through a worked example before addressing exercises yourself will improve the learning outcomes. You will therefore find some worked examples in this book, and additional sets of worked examples in the worked example site[3].

Statistics is an essential part of statistical physics. We have therfore included an introductory chapter on statistics. At the University of Oslo this introduction to statistics served as an important part of the overall learning outcomes of the bachelor degree, where all physics courses provide introduction to specific statistical terms. In this text we introduce basic probability theory, including most common probability densities and distributions, basic combinatorics, expected values and variance as well as estimators. We also introduce the concept of Monte Carlo modelling, both in the basic introduction to statistics, but also as an important tool to understand physics.

We have selected several computational approaches that are used throughout the textbooks. We introduce the essentials of Molecular Dynamics modeling, since this is used to gain insight into the irreversible nature of systems with many particles and to provide basic estimates for the properties of real gases, liquids, and solids. We also provide a general introduction to Monte Carlo methods and use these methods systematically throughout the text. We start from a general lattice-based model that is then reused for many different systems. This also strengthens a central learning outcome — that physics often is about recognizing how a new model can be mapped onto a well known model for which we already know the solution. In addition, we also introduce various algorithmic models, which are stochastic or Monte Carlo type, but not always following a Metropolis type algorithm. And we will use continuum methods to address heat and particular transport in diffusion or flow problems.

[1] `http://folk.uio.no/malthe/fys2160/index.html`
[2] `http://folk.uio.no/malthe/fys2160/notebooks`
[3] `http://folk.uio.no/malthe/fys2160/workedexamples`

August 2021 *Anders Malthe-Sørenssen and Dag Kristian Dysthe*

Contents

Preface v

1 Introduction 1
1.0.1 Thermodynamics 3
1.0.2 The great success of statistical mechanics 4
1.0.3 Integrated numerical methods 4
1.0.4 Molecular dynamics modeling 5
1.0.5 Learning Physics 6
1.0.6 Prerequisites 6
1.0.7 Structure of this book 7
1.0.8 Key questions 7

2 The Arrow of Time 9

2.1 One and many particles 9
2.1.1 Reversible bouncing of a solid ball 10
2.1.2 Irreversible bouncing of an elastic ball 10
2.1.3 More degrees of freedom 13

2.2 Approach to equilibrium — molecular dynamics 14
2.2.1 Approach to equilibrium 16

2.3 Approach to equilibrium — algorithmic model 20
2.3.1 Model of atoms on the left side 20
2.3.2 Simplifying the system — part 1 21
2.3.3 Simplifying the system — part 2 22

2.3.4 Fluctuations and the number of atoms, N 23
2.3.5 Average value of n 24
2.3.6 Dependence on initial conditions 25

2.4 Approach to equilibrium — theoretical model 26

2.5 Summary 27

2.6 Exercises 28
2.6.1 Discussion questions 28
2.6.2 Problems 30

3 Getting started with Molecular Dynamics modeling . 31

3.1 Atomic modeling basics 32
3.1.1 Lennard-Jones potential 33
3.1.2 Initial conditions 34
3.1.3 Boundary conditions 35
3.1.4 Integrating the equations of motion 36

3.2 Simple implementation of a molecular dynamics simulator . . 37
3.2.1 Non-dimensional equations of motion 37
3.2.2 Initialization of system 38
3.2.3 Integration of the equations of motion 41

3.3 Running a molecular dynamics simulation 42
3.3.1 Installation of LAMMPS 42
3.3.2 Starting a simluation in LAMMPS 44
3.3.3 Visualizing the results of a simulation 48

3.4 Analyzing the results of a simulation 49
3.4.1 Matlab implementation 49
3.4.2 Python implementation 51
3.4.3 Results 52

3.5 Advanced models 52
3.5.1 Coarsening 53
3.5.2 Hemoglobin in water 54
3.5.3 Lipid bilayers 54
3.5.4 Fracture of Silicon crystal 54
3.5.5 Water permeation in a carbon nanotube 54

3.6 Summary 55

3.7 Exercises 55
3.7.1 Problems 55

4 Probability and Statistics 57

4.1 Motivating example — fluctuations in a gas 58

4.2 Probabilities 61
4.2.1 Statistical experiments 61
4.2.2 Generating a sequence of random numbers 61
4.2.3 Measuring the frequency of occurrence 62
4.2.4 Frequencies and probabilities 63
4.2.5 Example: Probability of throwing a single dice 64
4.2.6 Properties of probabilities 66
4.2.7 Probability for simultaneous outcomes 68
4.2.8 Example: Throwing two dice 70

4.3 Expected value and Variance 71
4.3.1 Average 71
4.3.2 Expected value 71
4.3.3 The expected value is linear 72
4.3.4 Expected value of a function $g(N)$ 73
4.3.5 Example: Expected value of a dice throw 73
4.3.6 Variance and Standard deviation 74
4.3.7 Standard deviation 75
4.3.8 Example: Variance and standard deviation of a dice throw 76

4.4 Toward a theory for $n(t)$ for the MD model 76
4.4.1 One atom model 77
4.4.2 Two-atom model 77
4.4.3 Many-atom system 79
4.4.4 Comparing theory and simulation 81

4.5 The binomial distriubtion 85
4.5.1 Example: Coin toss 86
4.5.2 Properties of the binomial distribution 86
4.5.3 Expected value and variance of the binomial distribution 87
4.5.4 Continuum limit of the binomial distribution 89
4.5.5 Continuum limit — Analytical derivation 89
4.5.6 Continuum limit — Symbolic derivation 93
4.5.7 Probabilities and probability densities 95

4.6 Key Concepts in Combinatorics 96

4.7 Probability distributions 99
4.7.1 Uniform distribution 99

4.7.2 Exponential distribution 99
4.7.3 Poisson distribution 104
4.7.4 Universal distributions — The Central Limit Theorem 105
4.7.5 Extreme value statistics 108

4.8 Summary 110

4.9 Exercises 110

5 A thermodynamical example 123

5.1 Ideal gas 124
5.1.1 Ideal gas — Results from simulations 124
5.1.2 Ideal gas — Results from theory 127
5.1.3 Ideal gas — equipartition principle 128

5.2 Summary 129

5.3 Exercises 130

6 Micro- and Macro-states 131

6.1 States of the model gas 132
6.1.1 Micro- and macro-states for two dice 133
6.1.2 Definition of micro- and macro-states 133
6.1.3 Multiplicity of the macrostate of the gas 134

6.2 Behavior of a solid — Observations 137

6.3 Behavior of a solid — Theory 140
6.3.1 Einstein crystal model 141
6.3.2 Two Einstein crystals in contact 144
6.3.3 Time development of the Einstein crystal — Monte Carlo model 148
6.3.4 Time development of the Einstein crystal — Theoretical argument 151
6.3.5 Approximate multiplicities using Stirling's formula . . . 152

6.4 The ideal gas — Theory 157
6.4.1 Microstates of a single particle 158
6.4.2 Microstates of two non-interacting particles 160
6.4.3 Microstates of N non-interacting particles 161
6.4.4 Interaction between two ideal gases 162

6.5 Thermal equilibrium, Entropy and Temperature 165
6.5.1 Equilibrium of the Einstein crystal 166

6.5.2 Temperature in the Einstein crystal 169
6.5.3 Example: Entropy and energy of the Einstein crystal . 171
6.5.4 Example: Entropy and energy of the ideal gas 172
6.5.5 Example: Entropy of expansion 174
6.5.6 Example: Entropy of mixing 175
6.5.7 Example: Gibb's paradox 176

6.6 The laws of thermodynamics 177
6.6.1 Example: Entropy and heat 180

6.7 Mechanical equilibrium — Pressure 181
6.7.1 Example: Pressure of ideal gas 184

6.8 The thermodynamic identity 185
6.8.1 The thermodynamic identity and partial derivatives . . 186
6.8.2 Entropy, Heat and Work 187
6.8.3 Process description 187

6.9 Diffusive equilibrium — Chemical potential 188
6.9.1 Thermodynamical identity 190
6.9.2 Example: Chemical potential of the Einstein crystal . . 191
6.9.3 Example: Chemical potential of the ideal gas 191

6.10 Simulations of the microcanonical system 192
6.10.1 Molecular dynamics simulations 192
6.10.2 Stochastic simulations 193
6.10.3 Monte Carlo simulations 196

6.11 Summary 199

6.12 Exercises 200

7 The canonical system 203

7.1 Boltzmann statistics in the Einstein crystal 204
7.1.1 Einstein crystal – numerical example 204
7.1.2 Einstein crystal in heat bath — Theory 208
7.1.3 Comparison with numerical model 210

7.2 The partition function and Boltzmann statistics 211
7.2.1 Average values and fluctuations 212
7.2.2 Example: Two-state system 213
7.2.3 Example: Fluctuations and heat capacity 216
7.2.4 Sharpness of the canonical system 216
7.2.5 Example: Einstein crystal in the canonical system . . . 217

7.2.6 Example: Single spin system 219
7.2.7 Example: N-spin system 220
7.2.8 Example: Ising spin system 221
7.2.9 Degenerate states 224
7.2.10 Example: System with degenerate states 224

7.3 Thermodynamics in the canonical ensemble 225
7.3.1 Helmholtz free energy 226
7.3.2 A microscopic expression for Helmholtz free energy . . 227
7.3.3 Helmholz free energy decreases in approach to equilbrium 229
7.3.4 Helmholtz free energy is a minimum in equilibrium . . 236
7.3.5 Maxwell relations 237

7.4 Composite systems 237
7.4.1 Aspects of motion 237
7.4.2 System of two particles 238
7.4.3 System of many particles 240
7.4.4 Example: Ideal gas in the canonical ensemble 241
7.4.5 Example: Comparing microcanonical and canonical analysis 244
7.4.6 Example: Extension of a long polymer 248
7.4.7 Example: Ising model in 1d and 2d 251
7.4.8 Example: Grain boundaries 251

7.5 On the great importance of kT 251

7.6 Transformation between ensembles 252
7.6.1 Example: Fuel cell 255
7.6.2 Example: Muscle cell 256

7.7 Equipartition and velocity distributions 257
7.7.1 Equipartition theorem 257
7.7.2 Maxwell Velocity Distribution 258

7.8 Statistical entropy 262

7.9 General Monte-Carlo framework 263

7.10 Simulations of a canonical system 263
7.10.1 Monte Carlo estimates of integrals 264
7.10.2 Monte Carlo method for a canonical system 266
7.10.3 Detailed balance 272

7.11 Summary 273

7.12 Exercises 273
7.13 Projects 279

8 Thermodynamics and Heat engines 281
8.1 Thermodynamics 282
8.2 Engines 284
8.2.1 A simple piston 285
8.2.2 Net work and efficiency of the machine 288
8.2.3 Limits on efficiency 290
8.3 The Carnot cycle 292
8.3.1 Carnot cycle for ideal gas 293
8.4 Carnot refrigerator 296
8.5 Realistic model for Stirling Engine 297
8.6 Additional aspects of heat engines and thermodynamics (not releavnt for exam) 302
8.6.1 Path dependence of Heat and Work 302
8.6.2 Heat and work at constant temperature and constant pressure 303
8.6.3 Chemical Work 304
8.6.4 Joule-Thompson effect 304
8.7 Summary 304
8.8 Exercises 304
8.9 Projects 307

9 Chemical potential and Gibbs Distribution 309
9.1 Chemical potential 310
9.1.1 Chemical potential in the canonical system 310
9.1.2 Thermodynamic identity for diffusive systems 313
9.1.3 Chemical potential for multi-particle systems 314
9.1.4 Example: Chemical potential of the Einstein crystal . . 314
9.1.5 Example: Chemical potential of the Ideal gas 314
9.1.6 Potential energy and the chemical potential 315
9.1.7 Example: Barometric pressure formula 317
9.1.8 Example: Batteries 318

9.1.9 Example: Helmholtz free energy for system in potential field . . . 318
9.1.10 Example: Modeling potential gradients . . . 319

9.2 Phase equilibrium at constant volume . . . 321
9.2.1 Equilibrium between an ideal gas and an ideal crystal 321
9.2.2 Vapor pressure . . . 324

9.3 Gibbs Free Energy . . . 325
9.3.1 Definition of Gibbs free energy . . . 326
9.3.2 Minimum Free Energy principle . . . 327
9.3.3 Gibbs free energy and the chemical potential . . . 329
9.3.4 Gibbs free energy for a system of several species . . . 329

9.4 Phase coexistence at constant pressure . . . 330
9.4.1 Phase coexistence in a simple model . . . 330
9.4.2 General phase coexistence . . . 336
9.4.3 General phase coexistence and Gibbs free energy . . . 339
9.4.4 Clausius-Clapeyron's equation . . . 340
9.4.5 Example: Transition from graphite to diamond . . . 343
9.4.6 Latent heat and enthalpy . . . 343

9.5 Van der Waals fluid model . . . 343
9.5.1 Molecular dynamics simulation of Lennard-Jones gas . 344
9.5.2 Modifications to the ideal gas law . . . 347
9.5.3 Properties of the van der Waals system . . . 352
9.5.4 Surface tension . . . 358
9.5.5 Lattice gases . . . 359
9.5.6 Stability . . . 359

9.6 Solutions . . . 360
9.6.1 Ideal mixture . . . 361
9.6.2 Binary mixtures . . . 365
9.6.3 Dilute solutions . . . 366
9.6.4 Example: Vapor pressure above a dilute solution . . . 367
9.6.5 Example: Osmotic pressure . . . 369
9.6.6 Example: Change in boiling point above a dilute solution . . . 371
9.6.7 Example: Henry's law and the solubility of gases . . . 371
9.6.8 Non-ideal mixtures . . . 371
9.6.9 Surface growth processes . . . 371

9.7 Chemical reactions . . . 371

9.7.1 Chemical equilibrium for Ideal Gas systems 373
9.7.2 Example: Disassociation of hydrogen 374
9.7.3 Example: pH and the Ionization of water 375
9.7.4 Example: Kinetics, reaction rates, and catalytic processes 375
9.7.5 Example: Dissolved Oxygen 376
9.7.6 Example: Charge distribution of the Hemoglobin Molecule 376

9.8 Old Stuff 377
9.8.1 Intensive, Extensive Variables, and G 377
9.8.2 Multi-component systems 377
9.8.3 Matlab MD script 378

9.9 Summary 380

9.10 Exercises 380

9.11 Projects 380

10 Fermi and Bose gases 381

10.1 Gibbs factor and Gibbs sum 381
10.1.1 Averages using Gibbs sum 384
10.1.2 Gibb's sum for systems with several species of particles 385
10.1.3 Example: CO poisoning 386
10.1.4 Example: Diffusive equilibrium in the Einstein crystal 388
10.1.5 Example: Vacancies 388

10.2 Fermi-Dirac and Bose-Einstein distribution 388
10.2.1 Bose-Einstein distribution 391
10.2.2 Classical limit 394

10.3 Fermi gases 396
10.3.1 Zero temperature limit 396
10.3.2 Density of states 401
10.3.3 Behavior at $T = 0$ 404
10.3.4 Qualitative behavior at low temperatures 405
10.3.5 Quantitative description of low temperature behavior 408

10.4 Photon Systems 410
10.4.1 Entropy of the photon gas 413
10.4.2 Radiation from a black body 413

10.5 Phonon Systems 414

10.6 Boson Gas and Einstein Condensation 418
10.6.1 Spacing of levels 418
10.6.2 Approaching $T = 0$ 419

References 423

Index 425

List of Exercises

Exercise 2.1: Direction of time for a bouncing dimer 28
Exercise 2.2: Time development for gas with two atom types 28
Exercise 2.3: One-way hole 29
Exercise 2.4: Two-size particles and one-size hole 29
Exercise 2.5: Irreversible processes 29
Exercise 2.6: Bouncing point masses 30
Exercise 2.7: Example of irreversible processes 30
Exercise 2.8: Lifting a ball 30
Exercise 2.9: Atomic modeling of mixing 30
Exercise 3.1: Evolution of density 55
Exercise 3.2: Temperature variation 55
Exercise 3.3: Coarsening 56
Exercise 4.1: Deck of cards 110
Exercise 4.2: Dice 111
Exercise 4.3: Casino in St. Petersburg 111
Exercise 4.4: Coin-flip game 112
Exercise 4.5: Waiting time statistics 112
Exercise 4.6: Poisson statistics 113
Exercise 4.7: Gamma distribution 113
Exercise 4.8: Stirling's formula 114
Exercise 4.9: Conducting strip 114
Exercise 4.10: Velocity distribution in Argon gas 116
Exercise 4.11: The cumulative distribution function 116
Exercise 4.12: Sharpness of the temperature 117
Exercise 4.13: Central limit theorem for a dice 118

Exercise 4.14: Poisson distribution 118
Exercise 4.15: Multi-nomial distribution 118
Exercise 4.16: Conditional probabilities 119
Exercise 4.17: Conditional probability expansion 119
Exercise 4.18: Bayes theorem 120
Exercise 4.19: Joint probability distributions 121
Exercise 4.20: Joint probability density 121
Exercise 4.21: Maximum likelihood estimators 121
Exercise 4.22: Gamma distribution 121
Exercise 5.1: Limit of ideal gas law 130
Exercise 6.1: Microstates and macrostates of a random walk 200
Exercise 6.2: Flipping 30 coins 200
Exercise 6.3: Stirling's approximation 201
Exercise 6.4: Entropy 201
Exercise 6.5: Gas in one- and two-dimensions 201
Exercise 7.1: Numerical values for the partition function 273
Exercise 7.2: Changing the energy scale 273
Exercise 7.3: Thermal excitations in the Hydrogen atom in the Sun 273
Exercise 7.4: Thermal excitations of the water molecule 274
Exercise 7.5: Three-level system 274
Exercise 7.6: Partition function of two systems 275
Exercise 7.7: Arrhenius processes 275
Exercise 7.8: Energy fluctuations 275
Exercise 7.9: Hydrogen gas 276
Exercise 7.10: One-dimensional ideal gas 277
Exercise 7.11: Heat capacities of an ideal gas 277
Exercise 7.12: Compressibility of an ideal gas 278
Exercise 7.13: Equipartion for a linear degree of freedom 278
Exercise 7.14: Speed of an oxygen molecule 278
Exercise 7.15: Escape velocities 278
Exercise 7.16: Informational entropy 279
Exercise 8.1: Adiabatic process 304
Exercise 8.2: Adiabatic compressibility 305
Exercise 8.3: Crossing adiabatic curves 305
Exercise 8.4: Work of a gas 306
Exercise 8.5: String engine 306
Exercise 8.6: Three-point cycle 306
Exercise 9.1: Maxwell relations 380
Exercise 9.2: Compressibility of the van der Waals gas 380

Introduction 1

Thermal physics is the physics of macroscopic objects — objects that consist of many individual elements such as atoms, molecules or large compounds. In thermal physics we bring the tools from other parts of physics together and introduce entirely new concepts needed to address fundemental and practical problems. We combine mechanics, electromagnetics and quantuum mechanics to describe processes on macroscopic scales. Thermal physics is needed to address many of the large scale phenomena seen around you from the scales of nanometers to lightyears. Thermal physics is about many of the questions we wonder about: How does you blood carry oxygen? How does an ionic pump in a cell in your body work? What determines chemical equilibria? Why is ice frozen while water flows? How can we desalinate water, generate energy, conserve energy, or understand the biological and geological processes responsible for the world we live in? Thermal physics is also about limits: How much energy can we get from heat? How does the universe evolve in the really long run? Thermal physics is about fundamental questions: What are the processes responsible for the complex patterns and structures in the Earth's crust? Why do many processes only spontaneously occur in one direction in time? Can we build a machine that generates energy from nothing? Why is more different?

Thermal and statistical physics are about the wonders of the world, but as with all physics, we need tools to address and understand the wonders. However, as you develop these tools and improve your knowledge and insights into thermal and statistical physics, you will see that these tools allow you to pose further questions and gain further insights. But they

also allow you to be more precise with both your observations of nature and in your ability to predict and utlize nature. Thermal physics contains many of the elements needed to innovate to address some of the key questions that face us, and the Earth, now: How can we generate, store and distribute energy efficiently and in environmentally sustainable ways? How can we generate clean water and air? How can we generate energy efficient homes, cities and societies?

Thermal physics is the physics of systems of many. It may therefore be tempting to ask what the limits of thermal physics are. What happens when we get better at engineering systems that are very, very small. Can we then escape from some of the limitations of statistical physics? Is the physics of nano-scale systems fundamentally different from that of macroscopic systems? Is there a need for new physical concepts on the meso-scale, on the scale between individual atoms and molecules and the thermal physics scale of many atoms? These are questions and concepts that are still under development, as we develop better computational and experimental tools to address systems with thousands to billions of atoms.

Many of the key concepts of thermal and statistical physics can also be applied in completely different fields, but this should be done with care and caution. If you have a background in statistics you may indeed recoqnize many of the concepts introduced in this text, such as the strong similarity between the partition function and the likelihood function. Indeed, concepts from statistical physics, and particular algorithmic modeling approaches, are being applied across fields, n areas such as life science, epidemology, economics, sociology, finance, and marketing. This illustrates that the concepts you learn in statistical physics are useful and powerful, and that many of their applications are yet to be discovered.

The goal of this text is to provide you with a first introduction to key elements of thermal and statistical physics for students with no background in thermodynamics or statistical physics. The approach of this book is to start from the microscopic and develop the thermodynamic concepts from there. This means that we will start from an atomic picture, similar to what you have already learned to master in your basic courses in mechanics and quantuum mechanics and move slowly towards thermodynamics. I believe this coupling between the microscopic and the macroscopic to be conceptually important, indeed one of the great achievements of physics, but it is fully possible to apply thermodynamic principles without understanding the underlying microscopic foundations. When we have motivated thermodynamics from statistical physics, we will

continue to apply and use thermodynamic concepts to address processes without always reverting to a microscopic picture. If you are primarily interested in the applictions, you will therefore have to bear with me through the first, conceptual introduction.

1.0.1 Thermodynamics

Thermal and statistical physics is the study of macroscopic objects with the aim of understanding and predicting their behavior from the microscopic interactions between the particles making up the system. In our study of thermal physics we will discover new laws that extend beyond the laws of energy conservation or Newton's laws that we already know from mechanics, but these new laws are only valid for systems with many particles. A wonderful result in statistical physics is that many collective properties of macroscopic systems do not depend on the microscopic details. For example, all liquids have common properties that do not depend on the details of the interatomic interactions of the atoms or molecules making up the liquid. The behavior is the same if the liquid is made of metal, glass, water, polymers (macromolecules) or liquid hydrogen! There are common properties of matter that simply depends on the statistical behavior of systems of many particles, and thermal and statistical physics is the study of these properties. *Thermal physics* and *thermodynamics* contain the laws describing macroscopic objects and *statistical physics* is the theoretical framework that allows us to derive the laws of thermodynamics from our understanding of the underlying microscopic interactions.

In thermodynamics we introduce new physical concepts such as temperature, thermal energy (heat), and heat capacity. We introduce the concepts needed to discuss thermodynamic processes and thermodynamic equilibrium. We introduce the basic notions needed to characterize the equilibrium chemical processes such as free energies and chemical potentials and use these concepts also to discuss phase transitions and mixtures, including solutions.

Macroscopic objects have their own laws and behaviors which we must understand, learn, and preferrably develop a theoretical foundation for. This is the aim of *thermodynamics*. Thermodynamics can be developed as an axiomatic system, based on a few basic laws, but it can also be developed from a microscopic model of a system — using statistical mechanics.

1.0.2 The great success of statistical mechanics

Statistical mechanics is the theory that allows us to demonstrate the validity of thermodynamics and allows us to calculate many of the fundamental thermodynamic properties of a system. For example, we will later show how we can find the ideal gas law from the quantuum mechanical behavior of a particle in a box. Indeed, statistical mechanics is a great success in physics from both a conceptual and a practical point of view. It gives us insight into the laws of thermodynamics, since we can derive them, and it gives us tools to calculate real material properties from first principles. Statistical mechanics also gives insight into why some of the microscopic details are exactly that — details that do not matter — and why some details are important. In this text you, we start from a microscopic picture and derive the laws of thermodynamics before we apply and extend them.

1.0.3 Integrated numerical methods

This text grew from a need to have a text that integrates the use of numerical methods into the exposition, examples and exercises of a course in thermal physics. Computations have always been at the center of physics. It can indeed be argued that Newton invented calculus in order to calculate motion. But with the advent of computers and with their rapid growth in computational power, computating is now an integrated part of all aspects of physics, and it is also central to how physical is developed and applied in research and industry. However, in most textbooks, only analytical tools are introduced and taught. In this text, I have attempted to fully integrate the use of computational tools. This provides us with some advantages, since we are not limited to problems that can be solved analytically, but it does add new challenges as new computational methods need to be learned and mastered.

Integration of computational methods allows us to develop examples and exercises that more closely follow a work-flow similar to that used in research and industry, and it allows us to use realistic data for analysis. It will also teach you computational skills that we expect to be useful later.

The basic philosophy of this text is that the students should implement and run simulations themselves instead of using packages or applets that can be tweaked. We have found that this provides students with useful tools and insight, but also that it is very fun and inspiring to be able

to generate realistic physical simulations yourself and to visualize and understand your own data. Again, this comes at a cost, since you will not be able to develop as advanced code as a commercial software package can provide, but we believe that the satisfaction you gain by implementing it yourself is more important and enjoyable.

We have also integrated the use of symbolic calculations in python in both the exposition and the examples. This is to introduce you to a useful tool in its natural context — this is how physicits and engineers actually solve problems. We use the symbolic toolbox in python in order to keek everything within a single programming framework, even if we know that for example Mathematica provides a better set of tools for many symbolic problems.

This text is part of a series of texts, where each text introduces key computational concepts relevant for the particular type of physics introduced in the text. For thermal and statistical physics, we will introduce new concepts in algorithmic and stochastic modelling, molecular dynamics modeling, statistical descriptions of large data sets, optimization methods needs for minimization of free energies, and continuum modeling methods needed for kinetic theory and transport.

1.0.4 Molecular dynamics modeling

A particular focus of this text, will be the introduction of molecular dynamics (MD) modeling. Molecular dynamics is the integration of the trajectories of many interacting atoms forward in time. There are now many commercial and open source tools that allow you to use such simulations to measure thermodynamic properties of many types of systems, and molecular dynamics may therefore be a tool that is practically useful for you in your future professional career. In addition, it is a tool that allows us to observe the collective effects and to measure the statistical properties of systems of many atoms, and hence molecular dynamics also provide a powerful method to gain intuition into the concepts in statistical and thermal physics. Molecular dynamics is also a tool that is used to understand basic atomic and molecular processes in chemistry, material science, and life science. With the advent of ever more powerful techniques that combine quantuum mechanical calculations with molecular dynamics simulations of many atoms, simulations tools are starting to become real predicitive tools that can be used for calculations of material properties, rates of chemical processes, or the interaction of

chemical compounds with biologically active molecules or cell machinery. Indeed, molecular dynamics simulations is now a routine part of research and development in areas such as the development of novel materials, catalysis in chemical processing or in pharmaceutical research.

We have chosen to use the open source tool LAMMPS[1] to integrate the motion of atoms. However, our focus is on the primary principles and on the analysis of data from simulations in order to gain a firm understanding of the key concepts. It is therefore simple to replace this with another tool of your choice.

1.0.5 Learning Physics

The typical advice you get on how to learn physics is usually based on solid, pedagogical research. You learn physics by doing physics, by trying to apply the concepts introduced in new situations. This is why most textbooks in physics contains many worked examples and exercises: You learn best by first reading the exposition, working your way through a worked example, then work on the shorted exercises before you start working on larger projects. This book therefore contains these three elements:

Worked examples: Worked examples are used to motivate the theory, and are then integrated into the text, or are provided as examples as to how to solve an exercise or a problem. They typically contain both analytical, symbolic and computational methods.

Exercises: Short problems that adress a key concept.

Projects: Long problems that combine all the key concepts introduced in a chapter into a coherent story. These problems will typically contain both analytical, symbolic, and computational approaches and various levels of refinement of a theory.

1.0.6 Prerequisites

An introductory course in thermal and statistical physics typically follows courses in mechanics, electromagnetism, and quantuum physics. However, this text does not require any significant knowledge of electromagnetism — the few concepts needed, such as for electromagnetic waves, will be

[1] `http://lammps.org`

introduced when needed. Basic concept from mechanics such as Newton's laws, mechanical energy and work is used throughout the text.

A few key concepts from quantuum mechanics is needed, but we will not need to solve the Schrodinger equation. The concepts needed from quantuum mechanics is the quantization of energy states for a particle in a box (used for an ideal gas), for a two-state system, and for an harmonic oscillator. In order to calculate the rotational behavior of a diatomic gas, you also need to apply the formula for the rotational energy states. However, in all these cases no derivations are needed, we only apply a given formula for the quantization of states. In addition, we need the concepts of Fermions and Bosons to describe Fermi and Bose gases, but the only concept we use is that two Fermions cannot be in the same quanuum mechanical state.

It is not necessary to have an introductory programming course in order to follow this text, but it may be an advantage to have some experience in scripting in either Python or Matlab.

1.0.7 Structure of this book

1.0.8 Key questions

In our studies of thermal and statistical physics, we will ask and answer many fundamental as well as practical questions:

- What is temperature?
- Can a machine run forever?
- Why does a movie look strange when played backwards?
- How efficient can an engine or a heat-exchanger work?
- Can we generate electricity by lowering the temperature of the sea?
- How much energy does it take to boil a liter of water?
- Why does water droplets form on the outside of may cold glass of water?
- Why do clouds form?

These and many other questions will be addressed in this text. Stay with us!

The Arrow of Time

2

Abstract In this chapter we introduce the notion of the arrow in time and provide the first insights into the difference between a system with few and many particle. We introduce heterogeity in the distribution of energy in simple mechanical models, such as a bouncing ball and in a monatomic gas. For the monatomic gas, we introduce two modelling approaches that will be used frequently throughout the book: Molecular dynamics modeling and algorithmic modeling. We show how a system starting outside equilibrium approaches equilibrium irreversibly, relating the behavior of many atoms to the arrow of time.

Thermal and statistical physics is about the behavior of systems of many particles. But in what way is a system of many particles different from that of a few particles. Mainly in the number of particles! There are concepts and effects that only make sense in systems with many particles. Let us now try to gain intuition into what the effects of many particles are.

2.1 One and many particles

Let us start exploring from a simple macroscopic system consisting of either a single particle or subdivided into many particlesd. A system you are well aquinted with from mechanics is that of a bouncing ball. Does the arrow of time appear in the dynamics of a bouncing ball?

2.1.1 Reversible bouncing of a solid ball

Fig. 2.1b illustrates a classical problem in mechanics — a single ball bouncing on an elastic floor. The ball is affected by two potentials: Gravity $U_G = mgz$, and a linear spring modeling the interaction with the floor, $U_F = -(1/2)kz^2$ when $z < 0$ and 0 when $z > 0$. The figure illustrates the motion of the center of mass of this ball. This curve is completely symmetric in time — if we reverse time, $t \to -t$, the curve would look the same. You would not be able to see if we play a movie of the ball backward or forward. This is because the motion of this individual particle is reversible in time. You can see this directly from Newton's second law:

$$m\frac{d^2\mathbf{r}}{dt^2} = -\nabla U(\mathbf{r}) = m\frac{d^2\mathbf{r}}{d(-t)^2} \ . \tag{2.1}$$

We get the same equation if we replace t with $-t$. Indeed, this is also obvious from the conservation of energy, which gives

$$\frac{1}{2}mv^2 = U(\mathbf{r}) \ \Rightarrow v = \pm\sqrt{2U(\mathbf{r})/m} \ , \tag{2.2}$$

where there are two possible signs corresponding to two possible directions of movement.

For a single particle (affected by conservative forces) motion is *reversible in time*, and it would not look unphysical if we played a movie of the motion backward. This will always be the case for the systems we are looking at here — the motion of individual particles will always be conservative and reversible. Irreversibility comes from many particles. But how?

2.1.2 Irreversible bouncing of an elastic ball

Now, let us study our first system with many particles: Instead of modeling a solid ball, we study a ball that deforms elastically consisting of a line of masses connected with springs as shown in Fig. 2.1b. Only the bottom-most mass interacts with the floor, just like before, and the overall mass and the initial height is the same as before. What happens if we drop this ball? Fig. 2.1c shows the position of the center of mass, $z_{\rm cm}(t)$ as a function of time. Hmmm. This does not look the same as in Fig. 2.1b. The ball does not bounce back up to its original height, but bounces around at a lower height. What has happened?

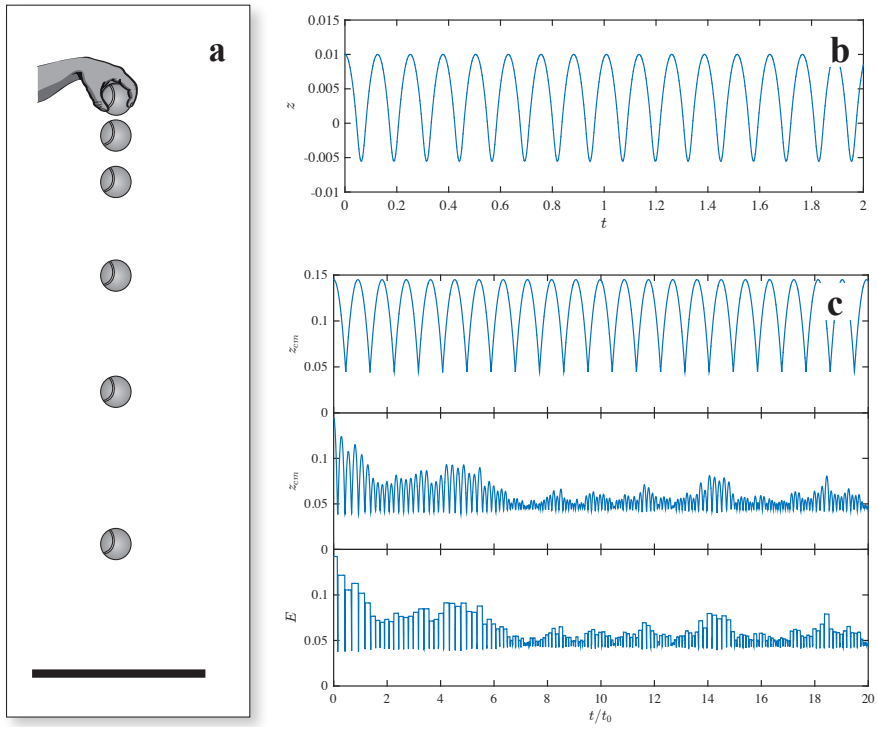


Fig. 2.1 The motion of a single ball bouncing on an elastic floor shown in a plot of $z(t)$ **a** Illustration of the experiment. **b** Ball modelled as a point particle. **c** Ball modelled as two connected particles.

The total energy in the system has not changed. The motion is conservative and the change in energy is negligible. However, we can divide the energy into two parts, the energy of the center of mass and the internal energy in the form of motion relative the center of mass and of internal compression or elongation of the springs. Fig. 2.1c shows the center-of-mass energy, $E_{\rm cm}$ and the internal energy $\Delta E = E_{\rm TOT} - {\rm e}_{\rm cm}$ as a function of time. We see that initially all the energy is the in the center of mass motion, but after a short while, energy is distributed between the center of mass motion and the internal energy.

The initial point therefore seems to be a special point, whereas configurations further on seems more representative. We are not surprised as the system bounces a bit up and down as time progresses, but we would be very surprised if the system suddenly bounced so that all the energy ended up in the center of mass motion, and no energy was in the internal motion, corresponding to what would happen if we played a movie of

the motion backwards. Notice that it is not physically impossible to go backwards. Indeed, if we reversed all the velocities after a short time, the system would return to its initial state where all the energy is in the center of mass motion.

It seems very *unlikely* that the system would move to this special state, corresponding to the initial state. *Improbable*, but not *impossible*. And we are now at the core of the arrow of time, motion forward in time is not the only possibility, but the most probable possibility. And we see that the system develops towards a state where the energy is distributed throughout all the possible degrees of freedom in the system, and that this state corresponds to an *equilibrium*. Still, all these words are vague and without a precise meaning, but we can see the initial contours of the difference between one particle and many particles. In systems with many particles there are many ways to distribute the energy, and the more probable states corresponds to a homogeneous distribution of energy.

We are now onto the new laws of physics of many particles. The are related to likelihoods — to probabilities — and they are therefore different than the law's of Newton, which are absolute laws. We will start to address the differences in a simple atomic system — an atomic gas.

Finding the motion of an elastic object: We model a one-dimensional elastically deforming ball as a set of masses, m, connected by springs of spring constant k. The springs are enumerated $i = 1, \ldots, N$ from the bottom to the top. The bottom-most mass also interacts with the floor. The force on particle i is then

$$m_i a_i = F_i = -m_i g - k\left(z_{i+1} - z_i - b\right) + k\left(z_i - z_{i-1}\right) - F_w \ , \quad (2.3)$$

where $F_w = -k_w z_i$ when for $i = 1$ when $z_i < 0$ and 0 otherwise. The motion is solved using the following program:

```
import numpy as np
import matplotlib.pyplot as plt
N = 10 # Nr of elements
k = 100000 # Internode spring constant
kw = 1000000 # Node-wall spring constant
d = 0.1 #Size in m
m = 0.1 #Mass in kg
t = 20.0 # Simulation time
g = 9.81 # Acc of gravity
z0 = 0.1 # Initial height
dt = 0.00001 # Timestep
```

```
nstep = int(np.ceil(t/dt))
z = np.zeros((N,nstep),float)
vz = np.zeros((N,nstep),float)
z[:,0] = d*(np.array(range(N)))/N+z0 # Initial positions
l = d/N # Distance between nodes
for i in range(nstep-1): # Calculate motion
    dz = np.diff(z[:,i])-l
    F = -k*np.append(0.0,dz) + k*np.append(dz,0.0) - m*g # Internode forces
    F[0] = F[0] - kw*z[0,i]*(z[0,i]<0) # Bottom wall
    a = F/m
    vz[:,i+1] = vz[:,i] + a*dt
    z[:,i+1] = z[:,i] + vz[:,i+1]*dt
# Plot results
time = np.linspace(0,nstep*dt,nstep)
for j in range(0,N,2):
    plt.plot(time,z[j,:])
```

Notice the use of vectorized math to find all the differences $z_{i+1} - z_i$ at once using 'diff'. The resulting motion is shown in Fig. 2.1c.

2.1.3 More degrees of freedom

The behavior of many particles becomes even clearer in a system with more degrees of freedom, as illustrated in Fig. 2.2, where we have simulated a two-dimensional ball bouncing against a flat surface. The ball bounces lower and lower as time progresses, but the fluctuations are now smaller, probably because we have much more particles to distribute the energy amongst. If we played the motion in this system backwards — if we reversed time — it would look strange and unphysical.

Energy is conserved in this system for both forward and backward motion. This means that energy conservation is not the only law that is important for macroscopic behavior. We need more laws! The new laws are related to how the energy is distributed in the object. From Fig. 2.2 we see that in the first few moments as the ball falls down, all the energy is in one particular degree of freedom, the motion of the center of mass of the ball. But as the ball bounces, the energy is distributed among many degrees of freedom. We will learn more about how we describe the distribution of energy between different degrees of freedom throughout this book, through the introduction of the concept of *entropy*. However, first we simply need to acknowledge that energy conservation is not enough to describe the arrow of time. We need something more!

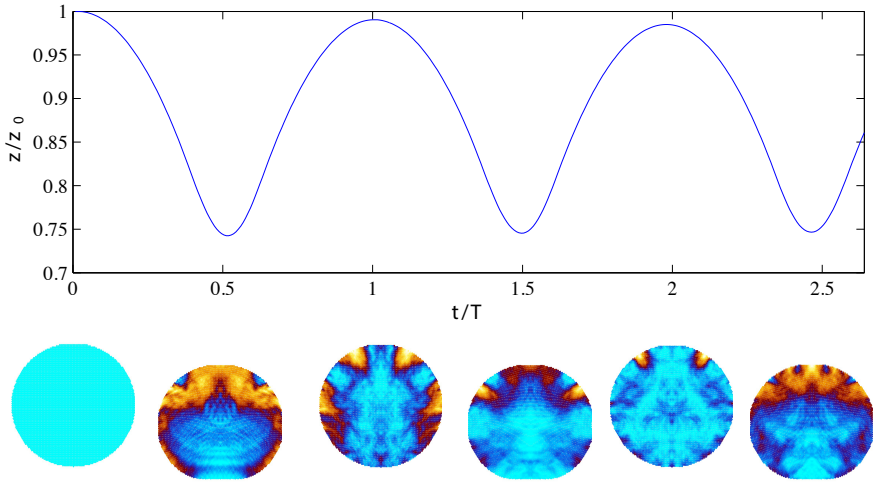


Fig. 2.2 Illustration of a computer simulation of a bouncing ball. Energy is conserved in the simulation, yet the ball does not bounce as high in its second bounce as its starting point. The color scale indicates the magnitude of the velocity. We see that initially, the whole ball has the same velocity, but after the ball has hit the ground, the distribution of velocities is heterogeneous.

2.2 Approach to equilibrium — molecular dynamics

Let us see how much insight we can gain from the atomic hypothesis and the motion of all the atoms in a system. Given the precise motion of all the atoms, we should be able to understand the behavior of the macroscopic system. However, as we shall see, it turns out not to be that simple. Description or direct calculation does not always provide insights by themselves. But we may be able to build understanding based on description and calculation.

Here, we use a molecular dynamics simulation program to find the motion of a set of atoms and then analyze the motion. You can read more about these tools in chapter (3).

How do we determine the time development of a system of atoms? This is a problem in quantum mechanics: The motion and interaction of atoms, including the full motion of the electrons, needs to be determined from the basic equations of quantum mechanics. However, it turns out that we can learn very much without a quantuum mechanical approach. The atomic hypothesis in itself combined with Newtonian mechanics and a simplified description of the interactions between atoms provide valuable insights as well as good predictions of the behavior of macroscopic systems.

How can we find the interactions between atoms? Again, we may use quantuum mechanics to find the effective potential energy for a system of many atoms. Here, we will start from a very simple model for atom-atom interactions: We will assume that the energy only only depends on the distance between pairs of atoms.

One of the simplest models for atom-atom interactions comes from a description of the interactions between nobel-gas atoms, but is widely used to model many types of atomic interactions. There are two main contributions to the interaction between two noble-gas atoms:

- There is an attractive force due to a dipole-dipole interaction. The potential for this interaction is proportional to $(1/r)^6$, where r is the distance between the atoms. This interaction is called the *van der Waals* interaction and we call the corresponding force the *van der Waals* force.
- There is a repulsive force which is a QM effect due to the possibility of overlapping electron orbitals as the two atoms are pushed together. We use a power-law of the form $(1/r)^n$ to represent this interaction. It is common to choose $n = 12$, which gives a good approximation to the behavior of Argon.

The combined model is called the Lennard-Jones potential (see Fig. 2.3):

$$U(r) = 4\epsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right) . \tag{2.4}$$

The force is given as the derivative of the potential.

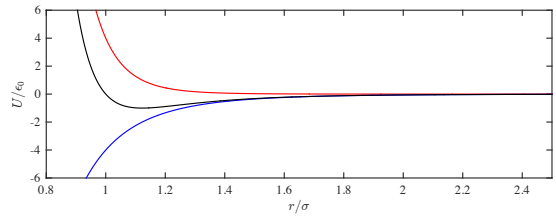


Fig. 2.3 Plot of the Lennard-Jones potential showing the repulsive part (*blue*), the attractive van-der-Waals interaction (*red*) and the combined Lennard-Jones poential (*black*).

We will use this potential as a basic tool to understand the behavior of macroscopic objects. To solve the equations of motion we will use molecular dynamics simulation package called LAMMPS, which is an efficient package for integrating the equations of motion of the system.

(Learn more about molecular dynamics and how to set up and analyze simulations in Chap. 3).

2.2.1 Approach to equilibrium

Let us address how a gas approaches equilibrium by starting the system outside equilibrium and studying the approach to equilibrium in detail.

We want to model a gas of atoms: a system where the atoms are rather far apart from each other. (We will be more specific later). We start with a box for the gas and divide the box into two equal parts. We fill one part of the box with atoms and keep the other part empty. The atoms start with random velocities. (But with zero total momentum). From $t = 0$ we allow the system to develop in time.

We run this simulation for 2000 timesteps in LAMMPS using

Terminal

```
lammps < in.gastwosection01.lmp
```

where the input file is in.gastwosection01.lmp[1]

```
# 2d Lennard-Jones gas
units lj
dimension 2
atom_style atomic
lattice hex 0.25
region mybox block 0 20 0 10 -0.5 0.5
create_box 1 mybox
region 2 block 0 10 0 10 -0.5 0.05
create_atoms 1 region 2
mass 1 1.0
velocity all create 0.5 87287
pair_style lj/cut 2.5
pair_coeff 1 1 1.0 1.0 2.5
neighbor 0.3 bin
neigh_modify every 20 delay 0 check no
fix 1 all nve
dump 1 all custom 10 twosec01.lammpstrj id type x y z vx vy vz
restart 2000 mydata.restart01 mydata.restart011
thermo 100
run 2000
```

The initial setup is shown in Fig. 2.4a. After 2000 timesteps (Fig. 2.4b) the system has expanded to fill almost the whole area, but we can still see that there are more atoms on the left side and the atoms still appear to

[1] http://folk.uio.no/malthe/fys2160/in.gastwosection01.lmp

be lumped together inhomogeneously. The resulting dynamics is shown in the movie of the simulation in gastwosection01.mp4[2].

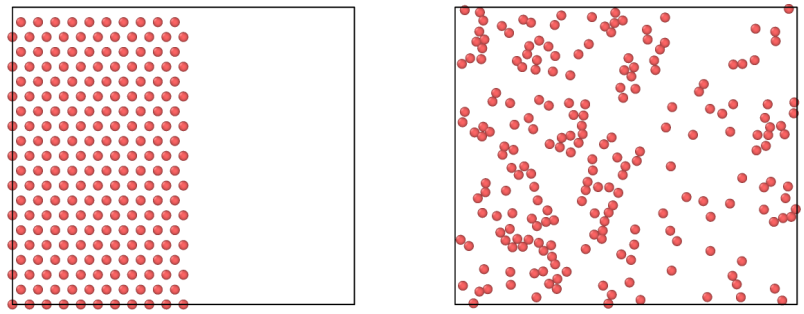


Fig. 2.4 Snap-shot from VMD of initial configuration of simulation of gas system (*left*) and after 2000 timesteps (*right*).

What if we ran the simulation backwards in time, starting from the state after 2000 timestep? The resulting behavior can be seen by running the simulator using in.gastwosection11.lmp[3], which results in the movie shown in gastwosection11.mp4[4].

```
# 2d Lennard-Jones gas
read_restart mydata.restart01
# Reverse all velocities
variable vx atom -vx
variable vy atom -vy
velocity all set v_vx v_vy NULL
fix all nve
dump 1 all custom 10 twosec11.lammpstrj id type x y z vx vy vz
restart 2000 mydata.restart11 mydata.restart111
thermo 100 # Output thermodyn variables every 100 timesteps
run 2000 # Number of timesteps
```

It *almost* gets back to where it started, with some small discrepancies due to numerical errors. The resulting behavior seen in the movie is clearly very different from the behavior we would expect. Indeed the behavior seems unphysical: We would not expect the atoms suddenly to organize themselves so that they all are on one side. Instead, we would expect the atoms to continue to be well mixed, and this is also what we would see if we continued to run the simulation forward in time. If we run the simulation for 10000 timesteps starting from the ordered

[2] http://folk.uio.no/malthe/fys2160/gastwosection01.mp4
[3] http://folk.uio.no/malthe/fys2160/in.gastwosection11.lmp
[4] http://folk.uio.no/malthe/fys2160/gastwosection11.mp4

state at $t = 0$, as is done by the input file in.gastwosection02.lmps[5] and visualized in the movie gastwosection02.mp4[6], we do not observe that the particles organize themselves regularly. Instead, we simply see that the particles mix homogeneously, and remain mixed while they move randomly about. Somehow, this mixed state seems more probable than the segregated state we started at, and the development in the movie in gastwosection02.mp4[7] appears to be physically reasonable. At least it is not clearly unphysical, as for the segregating system.

How can we quantify this behavior? We can measure the number of particles (or the relative number of particles) on the left and the right half of the system as a function of time. Let us use $n(t)$ for the number of atoms on the left side. Since the number of atoms is conserved and equal to the total number of atoms N, the number of atoms on the right side is $N - n(t)$. It is therefore enough to characterize the behavior of the number of atoms on the left side of the system.

How can we measure this from the data? We can read the data from the simulation directly into python using the Atomic Simulation Environment[8] Python package. to read and interpret the data from the files. You can learn more about how this works in Chap. 3. Here, I will simply use these routines to read data from the simulation.

To start using `ase` you must install it using:

Terminal

```
pip install ase
```

We use the following script to measure the number of atoms on the left side of the system:

```
#start1
import numpy as np
import matplotlib.pyplot as plt
from ase.io import read
traj = read("twosec02.lammpstrj", format="lammps-dump-text", index=":")
n = len(traj) # Number of timesteps
nleft = np.zeros(n)
# Find size of simulation box
firstframe = traj[0]
cellshape = firstframe.get_cell()
halfsize = cellshape[0][0]*0.5
# Count number of particles on left side for each timestep
```

[5] http://folk.uio.no/malthe/fys2160/gastwosection02.lmps
[6] http://folk.uio.no/malthe/fys2160/gastwosection02.mp4
[7] http://folk.uio.no/malthe/fys2160/gastwosection02.mp4
[8] https://wiki.fysik.dtu.dk/ase/

```
it = 0
for frame in traj:
    x = frame.get_positions() # positions of atoms
    xi = x[:,0] # x-coordinates of atoms
    nleft[it] = np.sum(xi<halfsize)
    it = it + 1
plt.plot(nleft)
#end1
np.savetxt('tmp.d', (t,Npart[0,:]))
```

and the resulting plot that shows the number of particle on the left is shown in Fig. 2.5. Now, this is interesting. We see that the system clearly moves in a particular direction. But it is also clear that we need more data to be really sure about what is happening. Let us run the simulation for 50000 timesteps and replot the data. (Using in.gastwosection20.lmps[9] and nrleft02.py[10]). The resulting plot in Fig. 2.5 now much more convincingly shows what happens: We see that $n(t)$ converges towards an approximately constant value and then fluctuates around this value.

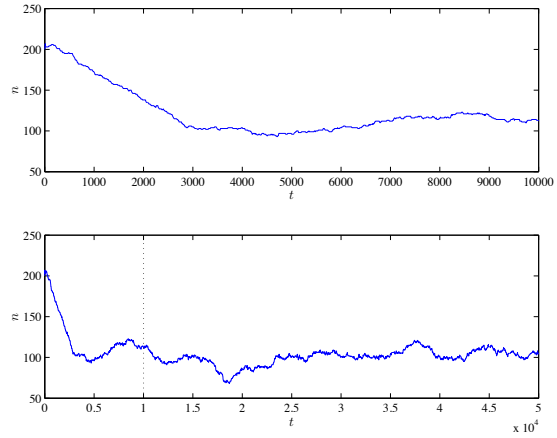


Fig. 2.5 Plot of the number of atoms on the left half (n) as a function of time (t) for 10000 timesteps (top) and for 50000 timesteps (bottom)

From this plot it is also clear that time is not fully reversible. It looks like we might move back and forth a bit in time, but if we go too far back we clearly end up in a situation that is different – all the atoms suddenly segrate. This seems unphysical. It is clearly not impossible – since it did indeed happen and would happen if we were able to run the system perfectly backwards in time. (You can achieve this numerically

[9] http://folk.uio.no/malthe/fys2160/in.gastwosection20.lmps
[10] http://folk.uio.no/malthe/fys2160/nrleft02.py

by having a sufficiently small time-step in your simulation). However, it seems to be unlikely, in the sense that even a tiny change in the positions or velocities, such as that generated by rounding errors, would mean that we would not be able to reverse the simulation. If you take the 50000 timestep simulation and reverse it, you will not get back to the segregated system. Instead, the system will fluctuate around the average value with no large excursion.

This arrow of time is central to our understanding of macrosopic systems. And it is clearly a new law — a law that comes in addition to conservation of momentum or conservation of energy.

2.3 Approach to equilibrium — algorithmic model

The atomic model directly illustrates the arrow of time, and we can use molecular dynamics to directly measure the approach towards an equilibrium situation: From the trajectories of all the atoms we can measure a macroscopic quantity such as $n(t)$. But we can gain even more insight into the process by trying the extract the most important part of the process — the part of the process that really introduced the arrow of time — and leave out all the unnecessary details. We do this by introducing a simplified model that only contains the most important feature of the process — we introduce an *algorithmic model* of the process. This method of defining a model which is as simple as possible is a common method used in statistical physics, which we will return to many times throughout this book.

How would such as simplified, algorithmic look like? What is the most important features of the process? It is usually not easy to know what the most essential feature are, and often you would need to experiment and test many different model before you arrive at the simplest (or at least a very simple) model that reproduces the main effects. However, here we will simply pick the right model immediately, unfairly giving you the impression that this is simple in practice. It is not.

2.3.1 Model of atoms on the left side

Let us try to analyze the core features of the system: What is the system? It is a closed box with an imaginary wall in the middle. Initially, there are N atoms on one side of the wall. On the other side of the wall there

are no atoms. As the atoms move about they get mixed, so that the distribution of atoms becomes more homogeneous after some time.

We could make this system more idealized by introducing an actual wall in the middle of the system. A wall with a small hole. This would also slow down the equilibration process, since it would take some time for the atoms to hit the hole and thus move to the other side to homogenize the distribution of atoms. However, we expect this system to retain many of the features of the original system: All the atoms starts on one side and as time progresses we expect the system to become more homogeneous. We could also model this system using molecular dynamics. But it would take a long time to find its behavior, since it would take a long time for an atom to actually hit the hole. We would spend a lot of time modeling atomic motion as the atoms bounced around inside the box, and we would spend little time modelling the actual process of interest, where an atom moves from one side of the box to the other. Instead, we should simplify the process and introduce an *algorithmic model* that only captures the most important part of the process — the motion of an atom from one box to another.

2.3.2 Simplifying the system — part 1

First, let us make a few simplifying assumptions about the system. Let us assume that all the atoms are independent. (This turns out to be a reasonable assumption and indeed the assumption we will use to define an ideal gas). With this assumption we get a new picture of what happens in the system. Each of the two boxes consists of many atoms that bounces around independenty. This means that each atom will have a given probability to go through the hole every time it hits the dividing wall. Another way to formulate this is that for each atom there is a certain probability per unit time that it will pass through the hole and move to the other side. Since the atoms are independent and identical, this probability must be the same for all the atoms. Now we have transformed the problem into a rule-based model, which we can use to find the behavior of the system:

- The system consists of N indepedent atoms seperated by a wall with a small hole. The position of each atom is given by s_i, so that it atom i is in the left box, $s_i = 1$, and if it is in the right box $s_i = 0$.
- For each atom, there is a probability $p = R\Delta t$ that the atom will move through the hole from one box to the other during a time interval Δt.

- There are $n(t) = \sum_i s_i$ atoms on the left side and $N - n(t)$ on the right side.

This means that we can formulate a simple rule for how the system develops in time. At a time t, the state of the system is given by the positions $s_i(t)$ of all the N atoms. In the short time interval Δt from t to $t + \Delta t$, each atom i has a probability $p = R\Delta t$ to move from the box it currently is in and into the other box[11], that it, each atom has a probability p to change its state s_i into the opposite state.

2.3.3 Simplifying the system — part 2

However, this model is still more complicated than necessary, because for a small time step Δt, we will spend a lot of time not doing anything. Instead, we could divide time into time intervals Δt, which corresponds to the typical time it takes for an atom to pass through the hole. Then, the rule would be:

- At each time interval Δt, corresponding to the typical time it takes for an atom to pass through the hole, we choose an atom at random and move it to the other side.

This is a possible model, but it is *still* a bit too elaborate. We can make the model even simpler and more to the point, because we are not interested in where each atom is – we are only interested in how many atoms there are at each side. We are only interested in n: the number of atoms on the left side, and $N - n$, the number of atoms on the right side. If we pick an atom at random an move it to the other side, the probability that we pick an atom on the left side is n/N. (And the probability that we pick an atom on the right side is $(N - n)/N$.) The probability that we move an atom from the left to the right hand side is therefore n/N. We therefore have arrived at a simple algorithm:

- Draw a random number r between 0 and 1.
- If $r \leq n/N$ we move an atom from left to right ($n \to n-1$), otherwise we move an atom from right to left ($n \to n+1$).

[11] You may ask whether these probabilities always must be the same: Does the probability for an atom to move from the left to the right have to be the same as the probability to move from the right to the left? We will return to this question later. For now we will assume that the system is symmetric, and that there is no difference between the two sides.

- Increase time by a unit Δt corresponding to the typical time it takes for an atom to pass through the hole.

We use this algorithm to study the approach towards equilibrium. The algorithm is implemented in the following program:

```
import numpy as np
import matplotlib.pyplot as plt
N = 210 # Number of particles
nstep = 4000 # Number of steps
n = np.zeros(nstep)
n[0] = 210 # Initial number on left side
for i in range(1,nstep):
    r = np.random.rand()
    if (r<n[i-1]/N):
        n[i] = n[i-1] - 1 # Move atom from left to right
    else:
        n[i] = n[i-1] + 1 # Move atom from right to left
plt.plot(range(0,nstep),n/N)
plt.xlabel('$t$'),plt.ylabel('$n/N$')
```

If we run this program for $N = 210$, which corresponds to the number of atoms in the molecular dynamics simulations, we get the $n(t)$ curve shown in Fig. 2.6.

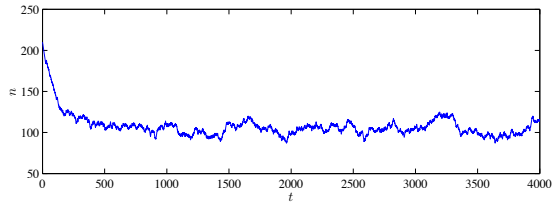


Fig. 2.6 Plot of the number of atoms on the left half (n) as a function of time (t) for 4000 timesteps of the discrete algorithm.

2.3.4 Fluctuations and the number of atoms, N

This simplified model allows us to study the approach to equilibrium and the fluctuations at equilibrium efficiently, for example by addressing the dependency on the number of atoms, N. What happens as we increase the number of atoms? Based on our discussions boave, we expect the system to become more irreversible, even if we still do not know what we precisely mean by that. Let us therefore look at the data. We implement a short script to plot the behavior for various values of N:

```
import numpy as np
```

```
import matplotlib.pyplot as plt
N = np.array([16,32,64]) # Number of particles
nstep = 4000 # Number of steps
for j in range(0,len(N)):
    NN = N[j]
    n = np.zeros(nstep)
    n[0] = NN # Initial number on left side
    for i in range(1,nstep):
        r = np.random.rand()
        if (r<n[i-1]/NN):
            n[i] = n[i-1] - 1 # Move atom from left to right
        else:
            n[i] = n[i-1] + 1 # Move atom from right to left
    plt.plot(range(0,nstep),n)
```

The resulting behaviors for $N = 20, 40, 60$ and 80 are shown in Fig. 2.7. We see that the fluctuations becomes smaller when N increases. This means that the system stays closer to the equilibrium value for n — corresponding to the average n over long times. The probability for a large excursion becomes smaller when the number of particles N becomes larger. This also means that when the number of particles is very small, it is fully possible to have a fluctuation where all the particles are on the left side ($n = 1$), but as N becomes larger, this becomes more and more unlikely. This illustrates the principle of the physical law we are exploring here: A behavoir far from the average becomes less and less probable — not impossible, only improbable — as the number of atoms increase. The physical laws for macroscopic systems are therefore a consequence of the many number of atoms in a typical macroscopic system. In a liter of gas (at room temperature and atmospheric pressure) there are typicall 10^{23} gas molecules. A very large number. And it is the largeness of this number that ensures the regularity of the physical laws. With $N = 10^{23}$ the fluctuations becomes very, very small indeed.

We also observe a clear transition from the initial state at $t = 0$ to the equilibrium state where there are (small) fluctuations around an average value of n. And we see that the transition time becomes longer for a larger system. For now, this will only be an interesting observation, but we may return to such transient effect later on.

2.3.5 Average value of n

What is the equilibrium value of n in this system? This is not a clear question, since we have not defined what *equilibrium* means. However, from our data in Fig. 2.7 it is clear that $n(t)$ approaches some value after

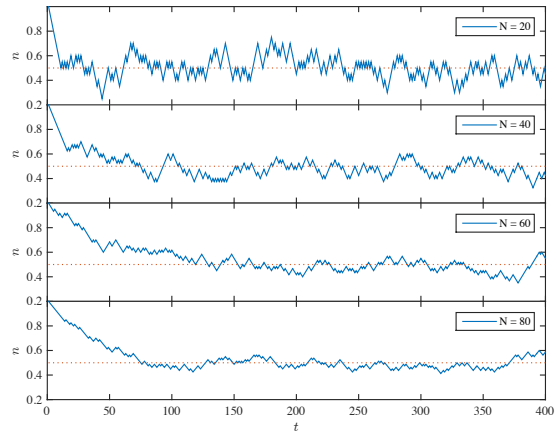


Fig. 2.7 Plot of $n(t)/N$ for various system sizes N.

a time, although it continues to fluctuate around this value. What is this value? We could find this value as the average value of n, after an initial transient time of 1000 steps, using the function `average`,

```
>> np.average(n[1000:])
0.50569523809523831
```

This is very close to $n(t)/N = 1/2$, which is what we would expect theoretically. Why? Because there is nothing special about the left hand side compared to the right hand side: They are equally large boxes and otherwise similar. We therefore expect the number of atoms on the left and on the right hand side to be equal, on average, and hence $n/N = (N-n)/N$, which gives $n/N = 1/2$. In practice, we are not going to measure exacly the same number of atoms on each side, since there are small fluctuations.

Test your understanding: What do you think will happen if we gradually increase N? Make a short program to plot the standard deviation of $n(t)$ as a function of N and comment on the results.

2.3.6 Dependence on initial conditions

Fig. 2.7 shows that the system approaches an equilibrium state after a short transition time, where the transition time appears to increase with the number of atoms, N, in the system. After this transition time, the system appears to have stabilized, with small fluctuations around an average value, $\langle n \rangle / N = 1/2$. During the short transition time, the initial conditions influence the state of the system: The value of $n(t)$ depends

on the initial value, $n(0)$, over some interval τ_0. This is also illustrated in Fig. 2.8, which shows the behavior for three different initial conditions, $n(0) = 1$, $n(0) = 1/2$ and $n(0) = 0$.

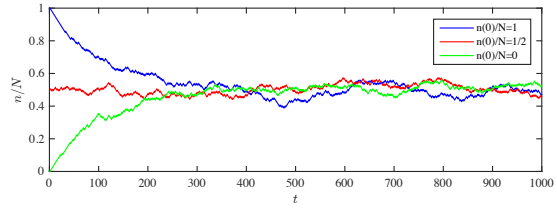


Fig. 2.8 Plot of $n(t)/N$ for various initial conditions $n(0)$: $n(0) = 1$ (*blue*), $n(0) = 1/2$ (*red*), $n(0) = 1/2$ (*green*).

We could call this transition time the memory extent for the system: The system remembers its initial conditions for some time interval τ_0, and after this time, it has reached the equilibrium state. We expect that the behavior of the system does not depend on the initial state after this transition time. This means that when the system reaches its equilibrium state, the behavior does no longer depend on the initial condition — the behavior is instead general, depending on the properties of the system when it is in the equilibrium state.

2.4 Approach to equilibrium — theoretical model

From these models we observe that macroscopic quantities, such as n, may become independent of time and of the initial state: If only the number of particles N is large enough the fluctuations will be very small and the quantity will be approximately independent of time. In the example above, we found that the gas system approaches $n(t) = 1/2$ independently of the initial condition of the gas. However, this prompts us to return to our main question: Why do these systems have a particular direction of time? In what way is the situation where all the atoms are on one side ($n(0)/N = 1$) different from the states with $n/N \simeq 1/2$ observed later, when the system has reached what we call equilibrium?

First, let us introduce a way to describe the state of the system. We use the term *microstate* to describe a particular arrangement of the atoms. This could be the positions of all the atoms, $\mathbf{r}_i$, or we could simplify it a bit further, and only specify if an atom is on the left or on the right side of the box. We can denote this by introducing $s_i = 1$ when atom i is on the left side, and $s_i = 0$ when atom i is on the right side. We can

describe the microstate of the system by listing all the s_i for all the i atoms in the system. Using this notation we can write a microstate as

$$(s_1, s_2, s_3, \ldots, s_N) \,, \tag{2.5}$$

For the $N = 2$ system there are 4 possible microstates:

Microstate	n	Description
$(1, 1)$	2	(atom 1 on the left, atom 2 on the left)
$(1, 0)$	1	(atom 1 on the left, atom 2 on the right)
$(0, 1)$	1	(atom 1 on the right, atom 2 on the left)
$(1, 1)$	0	(atom 1 on the right, atom 2 on the right)

Of these microstates, there is one microstate with $n = 2$, one microstate with $n = 0$ and two microstates with $n = 1$.

Now, if we assume that the system cycles through all the microstates and spends the same amount of time in each microstate, we would assume that the probability for each of the microstates are the same. Each microstate is equally probable. (This is a fundamental assumption. We will return with more arguments as to why this is true later).

This assumption would mean that the system will take the value $n = 2$, $n = 1$, $n = 1$, and $n = 0$ with the same frequency and the same probability, which will mean that the value $n = 1$ is twice as likely as the states $n = 2$ and $n = 0$. For the $N = 2$ the difference between the average value $n = 1$ and the initial condition where all atoms are on one side, $n = 2$ or $n = 0$, is not that large. But as N increases we will see that the state with $n = N$ becomes increasingly unlikely. For large N almost all the states occur near $n = N/2$ and there are very few states with $n = N$ or $n = 0$. That is what we are going to mean by equilibrium: It is the most probably state. But not the most probably microstate, because any microstate is equally probable, but the most probably macrostate: The most probably value of n in this case. We will dive deeper into these arguments and these distinctions in the next chapters.

(You may already now realize that the probability distribution for n is given by the binomial formula, $P(n, N) = N!/(n!(N - n)!)(1/2)^N$, which is very sharp around $n = N/2$ as N becomes large. We will address these issues in detail in the following chapters).

2.5 Summary

- On an atomic level all laws are reversible in time.

- On a macroscopic level, time has a specific direction.
- Energy conservation is not a sufficient principle. We need new principles describing the distribution of energy.
- We can understand the direction of time as a development towards a most likely state in the system.

2.6 Exercises

2.6.1 Discussion questions

Exercise 2.1: Direction of time for a bouncing dimer

Fig. 2.9 shows a bouncing dimer ($N = 2$). However, in one of the figures the time axis has been reversed. Which figure shows the correct progression of time and what arguments can you use to find out?

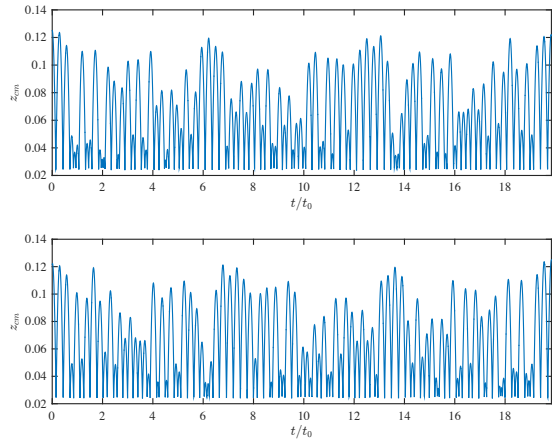


Fig. 2.9 Plot $z_{\mathrm{cm}}(t)$ for a bouncing dimer, that is a mass-spring system with two ($N = 2$) masses. Which figure shows the correct progression of time?

Exercise 2.2: Time development for gas with two atom types

A system consists of two types of atoms, A and B, in gas form in a rectangular box. Initially, all the A atoms are on the left side of the box and all the B atoms are on the right side. Which of the figures in Fig. 2.10 represent realistic time developments of the system? Explain your answer.

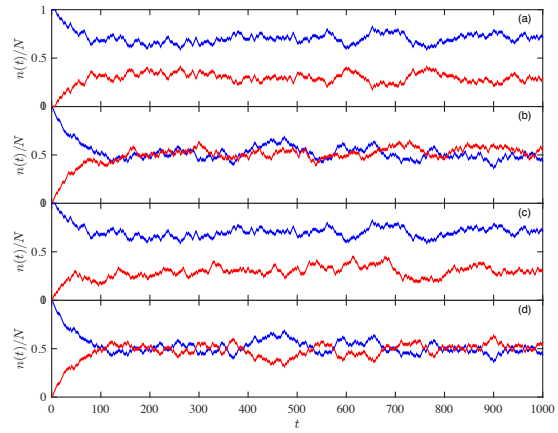


Fig. 2.10 Plot $n_A(t)$ (*blue*) and $n_B(t)$ (*red*) for a two-atom system. However, only one of the figures are from a realistic simulation.

Exercise 2.3: One-way hole

A box consists of atoms of one type. Initially, all the atoms are on the left side. There is a hole separating the left from the right side and the atoms can only progress from the left to the right through this hole. Sketch the time development of $n(t)$. Does this behavior seem realistic to you?

Exercise 2.4: Two-size particles and one-size hole

A box consists of two atoms, A and B, where atom A is smaller than atom B. Initially, the atoms are uniformly distributed throughout the box. Suddenly, we place a sieve in the middle of the box, so that only the A atoms may pass through the sieve, but not the B atoms. Sketch the time development of the number of particles $n_A(t)$ ($n_B(t)$) of type A (type B) on the left side as a function of time.

Exercise 2.5: Irreversible processes

List a few processes from your kitchen that clearly are not reversible and explain why you think the processes are irreversible.

2.6.2 Problems

Exercise 2.6: Bouncing point masses

Discuss the difference between one bouncing point mass and a system consisting of two bouncing point-particles connected by a spring. Calculate the time development for a one-dimensional system and comment on the behavior.

Exercise 2.7: Example of irreversible processes

List more examples of processes that only can go in one direction. Try to articulate why the reverse process does not happen spontaneously.

Exercise 2.8: Lifting a ball

If you take a ball from a table and place it on the floor – can this process spontaneously occur in reverse? If you take a ball from the floor and place it on the table – can this process spontaneously occur in reverse? Explain.

Exercise 2.9: Atomic modeling of mixing

We start a molecular dynamics simulation from a gas state where all the atoms are distributed uniformly in space with random initial velocities. Color the left half red and the right half blue. What is the time development of the number of red particles in the left side? Will you ever return to the starting position?

Getting started with Molecular Dynamics modeling

3

Abstract In this chapter we provide a quick introduction to molecular dynamics modeling. In molecular dynamics the motion of a set of atoms is determined from a model for the inter-atom interactions. We demonstrate the basic physical formulation for a Lennard-Jones model for a gas and provide a python implementation of the molecular dynamics algorithm. This python implementation is too slow for any practical application, and we therefore introduce an open-source integrator to determine the motion of all the atoms. Installation and use of the LAMMPS simulator is described in detail. The simulation produces a set of trajectories for all the atoms in the model, and we also demonstrate how to read these trajectories into python and use this data-set to characterize the behavior of realistic systems.

Statistical mechanics allows us to go from the atomic hypotesis to theories for macroscopic behavior. We will do this by developing strongly simplified models of the physical systems we are interested in: the ideal gas and the ideal crystal models. However, we can also study the behavior of a system in detail by calculating and following the motion of each atom in the system. In this book we will use a modeling tool called molecular dynamics. This is a useful tool that you may indeed use in your professional career as a researcher. Also, molecular dynamics allow you to pose and answer many questions about both atomic scale systems and many-particle systems — questions with non-trivial answers as well as real research questions. The tools we introduce here are indeed the tools that are used to understand nano-scale structures of materials, to understand and engineer chemical processes such as catalysis, to model

and understand biological processes and to understand the interaction between drugs and living organisms. Molecular dynamics modeling is an important field of research and engineering, and our quick introduction here is in no way exhaustive, but should give you a quick introduction for the capabilities of the tool.

This chapter starts by motivating the use of classical mechanics to understand atomic motion. We introduce the Lennard-Jones model and standard boundary and initial conditions for atomic modeling studies. We introduce standard integration methods and a very simple illustrative implementation in python. A more practical approach is then introduced using the LAMMPS software package, provide installation and simulation instructions for modeling the Lennard-Jones system using this package. We also demonstrate how to visualize and analyse the results of a simulation. Finally, we also demonstrate how LAMMPS can be used to model more realistic systems, such as water, biological and mineral systems. (These finals parts are interesting, but not central to exposition of this book, and can be skipped on a first reading).

3.1 Atomic modeling basics

How do we determine the motion of a system of atoms? This is really a problem in quantum mechanics — the motion and interaction of atoms, including the full motion of the electrons, needs to be determined from the basic equations of quantum mechanics. This is fully possible for small systems with present technology. However, to address the behavior of macroscopic systems, we need thousand, millions, or billions of atoms to get representative results from our calculations. However, if we want to model thousands to millions of atoms we need a different and more computationally efficient approach.

The use of the classical approximation to describe the motion of a system of atoms is, fortunately, able to capture many of the features of macroscopic systems. In this approximation, we use Newton's laws of motion and a well-chosen description of the forces between atoms to find the motion of the atoms. The forces are described by the potential energies of the atoms given their positions. The potential energy functions can, in principle, be determined from quantum mechanical calculations because the forces and the potential energies depend on the states of the electrons — it is the electrons that form the bonds between atoms and are the origin of the forces between atoms. However, in the classical

approximation we parametrize this problem: We construct simplified models for the interactions that provide the most important features of the interactions.

3.1.1 Lennard-Jones potential

These parametrized models can be simple or complicated. They can include only pair interactions, three-particle interactions, or even many-particle interactions. Here, we will primarily use simple models, since the statistical and thermal effects we are interested in do not depend strongly on the details of the system. One of the simplest models for atom-atom interactions is a representation of the interactions between nobel-gas atoms, such as between two Argon atoms. For the interaction between two noble-gas atoms we have two main contributions:

- There is an attractive force due to a dipole-dipole interaction. The potential for this interactions is proportional to $(1/r)^6$, where r is the distance between the atoms. This interaction is called the *van der Waals* interaction and we call the corresponding force the *van der Waals* force.
- There is a repulsive force which is a quantum mechanical effect due to the possibility of overlapping electron orbitals as the two atoms are pushed together. We use a power-law of the form $(1/r)^n$ to represent this interaction. It is common to choose $n = 12$, which gives a good approximation for the behavior of Argon.

The combined model is called the **Lennard-Jones potential**:

$$U(r) = 4\epsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right) . \tag{3.1}$$

Here, ϵ is a characteristic energy, which is specific for the atoms we are modeling, and σ is a characteristic length.

The Lennard-Jones potential and the corresponding force $F(r)$ is illustrated in Fig. 2.3. We see that the Lennard-Jones potential reaches its minimum when

$$F(r) = -\frac{d}{dr}U(r) = 0 \,, \tag{3.2}$$

which occurs at

$$F(r) = 24\,\epsilon_0 \left((\sigma/r)^6 - 2\,(\sigma/r)^{12} \right) = 0 \;\Rightarrow\; \frac{r}{\sigma} = 2^{1/6} \,. \tag{3.3}$$

We will use this potential to model the behavior of an Argon system. However, the Lennard-Jones potential is often used not only as a model for a noble gas, but as a fundamental model that reproduces behavior that is representative for systems with many particles. Indeed, Lennard-Jones models are often used as base building blocks in many interatomic potentials, such as for the interaction between water molecules and methane and many other systems where the interactions between molecules or between molecules and atoms is simplified (coarse grained) into a single, simple potential. Using the Lennard-Jones model you can model 10^2 to 10^6 atoms on your laptop and we can model 10^{10}-10^{11} atoms on large supercomputers. However, if you are adventurous you may also model other systems, such as water, graphene, or complex biological systems using this or other potentials as we demonstrate in Sect. 3.5

3.1.2 Initial conditions

An atomic (molecular) dynamics simulation starts from an initial configuration of atoms and determines the trajectories of all the atoms. The initial condition for such a simulation consists of all the positions, $\mathbf{r}_i(t_0)$ and velocities $\mathbf{v}_i(t_0)$ at the initial time t_0. In order to model a realistic system, it is important to choose the initial configuration with some care. In particular, since most potentials such as the Lennard-Jones potential increase very rapidly as the interatomic distance r goes to zero, it is important not to place the atoms too close to each other. We therefore often place the atoms regularly in space, on a lattice, with initial random velocities.

We generate a lattice by first constructing a *unit cell* and then copying this unit cell to each position of a lattice to form a regular pattern of unit cells. (The unit cell may contain more than one atom). Here, we will use cubic unit cells. For a cubic unit cell of length b with only one atom in each unit cell, we can place the atom at $(0, 0, 0)$ in the unit cell and generate a cubic lattice with distances b between the atoms by using this cubic unit cell to build a lattice. This is illustrated for a two-dimensional system in Fig. 3.1. Such a lattice is called a *simple cubic lattice*.

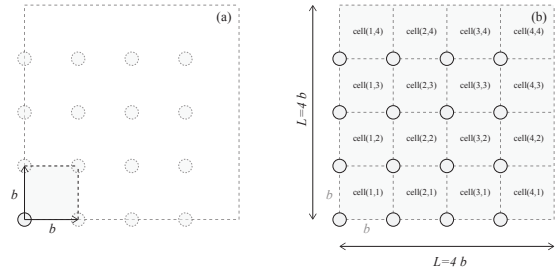


Fig. 3.1 **a** Illustration of a unit cell for a square lattice. **b** A system consisting of 4×4 unit cells, where each of the cells are marked and indexed for illustration.

However, for a Lennard-Jones system we know (from other theoretical, numerical and experimental studies) that the equilibrium crystal structure is not a simple cubic lattice, but a face centered cubic lattice. This is a cubic lattice, with additional atoms added on the center of each of the faces of the cubes. The unit cell for a face centered cubic lattice is illustrated in Fig. 3.2. We will use this as our basis for a simulation and then select a lattice size b so that we get a given density of atoms. The whole system will then consist of $L \times L \times L$ cells, each of size $b \times b \times b$ and with 4 atoms in each cell.

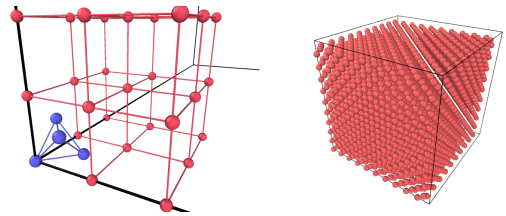


Fig. 3.2 *Left* Illustration of a unit cell for a face centered cubic lattice. Unit cell atoms illustrated in blue and the base position of other cells shown in red. *Right* A system consisting of $10 \times 10 \times 10$ unit cells.

3.1.3 Boundary conditions

A typical molecular model of a liquid of Argon molecules is illustrated in Fig. 3.3a. In this case, we have illustrated a small system of approximately $10 \times 10 \times 10$ atom diameters in size. Below, you will learn how to set up and simulate such systems on your laptop. Unfortunately, you will not be able to model macroscopically large systems — neither on your laptop nor on the largest machine in the world. A liter of gas at room temperature typically contains about 10^{23} atoms, and this is simply beyond practical computational capabilities.

But it is possible with a small system to gain some insights into how very large, even infinite, systems behave? One of the problems with the

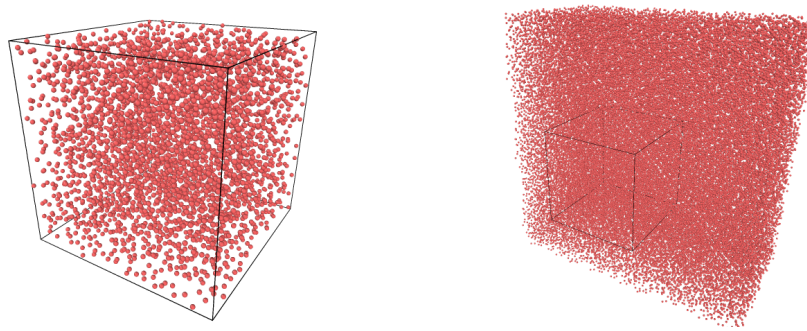


Fig. 3.3 **a** Visualization of a $10 \times 10 \times 10$ simulation of an Argon liquid. **b** Visualization of a 10×10 simulation of an Argon gas, showing the actual simulation area in the center box, and 8 copies of the center box illustrating the principles of the periodic boundary conditions.

$10 \times 10 \times 10$ system above is the external boundaries. But we can fool the system into believing it is infinite by applying what we call periodic boundary conditions. If the atoms on the left of the system do not see emptiness to their left, but instead see the right hand side of the system, as if the system is wrapped around a cylinder, the system will look like it is inifinite. This is illustrated in Fig. 3.3b. This convention of periodic boundary conditions is usually applied in all simulations in order to avoid dealing with boundary conditions. (There may be some possibly problematic aspects of periodic boundaries, but we will not address or worry about these here).

3.1.4 Integrating the equations of motion

How do we determine the behavior of the system? We solve the equations of motion. For molecular dynamics simulations we usually use an algorithm called the Velocity-Verlet, which is approximately like the forward Euler method, but it is very well suited for conservative forces. The velocity is calculated at both time t and at intermediate times $t + \Delta t/2$, where Δt is the time-step, whereas the forces are only calculated at the full time-steps, t, $t + \Delta t$, $t + 2\Delta t$ etc. The most time-consuming part of the calculation is the calculation of the forces. We therefore want to limit the number of times we calculate the forces and still have as high precision as possible in our calculations. The Velocity Verlet algorithm ensures a good trade-off between precision in the integration algorithm and the number of times forces are calculated.

In the Velocity-Verlet algorithm the positions $\mathbf{r}_i(t)$ and velocities $\mathbf{v}_i(t)$ of all the atoms, $i = 1, \ldots, N$, are propagated forward in time according to the algorithm:

$$\mathbf{v}_i(t + \Delta t/2) = \mathbf{v}_(t) + \mathbf{F}_i(t)/m_i \Delta t/2 \tag{3.4}$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_(t) + \mathbf{v}_i(t + \Delta t/2) \tag{3.5}$$

$$\mathbf{F}_i(t + \Delta t) = -\nabla V(\mathbf{r}_i(t + \Delta t)) \tag{3.6}$$

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_(t + \Delta t/2) + \mathbf{F}_i(t + \Delta t)/m_i \Delta t/2 \ , , \tag{3.7}$$

This method has very good properties when it comes to energy conservation, and it does, of course, preserve momentum perfectly.

3.2 Simple implementation of a molecular dynamics simulator

(*This is currently only implemented in Matlab — a Python version will be coming soon*) How can we implement the full simulation procedure in python? We need to set up the initial configuration of atoms, integrate the motion for a given number of time-steps, and then output the results to a file that can be read by standard visualization programs. Here, we provide a full implementation to show how a molecular dynamics simulation is implemented. However, this is mainly for illustrational purposes so that you can see how it is done. We will actually use a state-of-the-art open-source simulation package to perform our simulations.

3.2.1 Non-dimensional equations of motion

However, all the quantities in a molecular dynamics simulations are very small. It is therefore usual to introduce measurement units that are adapted to the task. For the Lennard-Jones model we usually use the intrinsic length and energy scale of the model as the basic units of length and energy. This means that we measure lengths in units of σ and energies in units of ϵ_0. A vector $\mathbf{r}'_i$ in the simulation is therefore related to the real-world length $\mathbf{r}_i$ through

$$\mathbf{r}_i = \sigma \, \mathbf{r}'_i \ \Leftrightarrow \ \mathbf{r}'_i = \mathbf{r}_i/\sigma \ . \tag{3.8}$$

Similarly, we can introduce a Lennard-Jones time, $\tau = \sigma\sqrt{m/\epsilon}$, where m is the mass of the atoms, and the Lennard-Jones temperature $T_0 = \epsilon/k_B$.

Using these notations, we can rewrite the equations of motion for the Lennard-Jones system using the non-dimensional position and time, $\mathbf{r}'_i = \mathbf{r}_i/\sigma$ and $t' = t/\tau$:

$$m\frac{d^2}{dt^2}\mathbf{r}_i = \sum_j 24\epsilon_0 \left((\sigma/r_{ij})^6 - 2\,(\sigma/r_{ij})^{12}\right)\left(\mathbf{r}_{ij}/r_{ij}^2\right) \ , \tag{3.9}$$

to become

$$\frac{d^2}{d(t')^2}\mathbf{r}'_i = \sum_j 24\left(r_{ij}^{-6} - 2r_{ij}^{-12}\right)\left(\mathbf{r}'_{ij}/r_{ij}'^2\right) \ . \tag{3.10}$$

Notice that this equation is general. All the specifics of the system is now part of the characteristic length, time and energy scales σ, τ, and ϵ_0. For Argon $\sigma = 0.3405\mu$m, and $\epsilon = 1.0318 \cdot 10^{-2}$eV, and for other atoms you need to find the corresponding parameter values.

Quantity	Equation	Conversion factor	Value for Argon
Length	$x' = x/L_0$	$L_0 = \sigma$	0.3405μm
Time	$t' = t/tau$	$\tau = \sigma\sqrt{m/\epsilon}$	$2.1569 \cdot 10^3$ fs
Force	$F' = F/F_0$	$F_0 = m\sigma/\tau^2 = \epsilon/\sigma$	$3.0303 \cdot 10^{-1}$
Energy	$E' = E/E_0$	$E_0 = \epsilon$	$1.0318 \cdot 10^{-2}$ eV
Temperature	$T' = T/T_0$	$T_0 = \epsilon/k_B$	119.74 K

3.2.2 Initialization of system

First, we need to initialize the system by generating a lattice of atoms with random velocities. We write a small script to generate a lattice and write it to disk. The lattice consists of units cells of size b (measure in units of σ, as are everything in the simulation). If a unit cell starts at $\mathbf{r}_0$, then the 4 atoms in the units cell are at positions $\mathbf{r}_0$, $\mathbf{r}_0 + (b/2, b/2, 0)$, $\mathbf{r}_0 + (b_2, 0, b_2)$, and $\mathbf{r}_0 + (0, b/2, b/2)$. If a system consists of $L \times L \times L$ such cells it is simple to create the system: We simply loop through all L^3 positions of $\mathbf{r}_0$ and for each such position we add atoms at the four positions relative to $\mathbf{r}_0$. This is done using the following script, which is explained in more detail below:

```
L = 5; % Lattice size
b = 2.0; % Size of unit cell (units of sigma)
v0 = 1.0; % Initial kinetic energy scale
N = 4*L^3; % Nr of atoms
r = zeros(N,3);
```

```
v = zeros(N,3);
bvec = [0 0 0; b/2 b/2 0; b/2 0 b/2; 0 b/2 b/2];
ip = 0;
% Generate positions
for ix = 0:L-1
    for iy = 0:L-1
        for iz = 0:L-1
            r0 = b*[ix iy iz]; % Unit cell base position
            for k = 1:4
                ip = ip + 1; % Add particle
                r(ip,:) = r0 + bvec(k,:);
            end
        end
    end
end
% Generate velocities
for i = 1:ip
    v(i,:) = v0*randn(1,3);
end
% Output to file
writelammps('mymdinit.lammpstrj',L*b,L*b,L*b,r,v);
```

In addition, we need the following function to write the data to file:

```
function writelammps(filename,Lx,Ly,Lz,r,v)
%WRITELAMMPS Write data to lammps file
fp = fopen(filename,'w');
s = size(r);
ip = s(1);
fprintf(fp,'ITEM: TIMESTEP\n');
fprintf(fp,'0\n');
fprintf(fp,'ITEM: NUMBER OF ATOMS\n');
fprintf(fp,'%d\n',ip); % Nr of atoms
fprintf(fp,'ITEM: BOX BOUNDS pp pp pp\n');
fprintf(fp,'%f %f\n',0.0,Lx); % box size, x
fprintf(fp,'%f %f\n',0.0,Ly); % box size, y
fprintf(fp,'%f %f\n',0.0,Lz); % box size, z
fprintf(fp,'ITEM: ATOMS id type x y z vx vy vz\n');
for i = 1:ip
    fprintf(fp,'%d %d %f %f %f %f %f %f \n',...
            i,1,r(i,:),v(i,:));
end
fclose(fp);
end
```

Notice that we use the vectors $\mathbf{b}_k$ to describe the four positions relative to the origin, $\mathbf{r}_0$ of each cell:

```
bvec = [0 0 0; b/2 b/2 0; b/2 0 b/2; 0 b/2 b/2];
...
r0 = b*[ix iy iz]; % Unit cell base position
for k = 1:4
    ip = ip + 1; % Add particle
    r(ip,:) = r0 + bvec(k,:);
```

```
end
```

Also notice that we use a variable `ip` to keep track of the particle index for the atom we are currently adding to the system. This is simpler than calculating the particle number each time. That is, we can simply use

```
    ip = ip + 1; % Add particle
```

instead of the more elaborate and less transparent

```
    ip = ix*4*L*L + iy*4*L + iz
```

to calculate the current particle number each time a new atom is added. Finally, we add a random velocity vector, with a normal (Gaussian) distribution of velocities with a standard deviation of v_0 using the `randm` function. This can be done using a loop or with a single, vectorized command:

```
for i = 1:ip
    v(i,:) = v0*randn(1,3);
end
#
```

or

```
v = v0*randn(ip,3);
```

Finally, the state is saved to the file `mymdinit.lammpstrj`. The resulting file, which contains the complete atomic state of the system, looks like this:

```
ITEM: TIMESTEP
0
ITEM: NUMBER OF ATOMS
500
ITEM: BOX BOUNDS pp pp pp
0.000000 10.000000
0.000000 10.000000
0.000000 10.000000
ITEM: ATOMS id type x y z vx vy vz
1 1 0.000000 0.000000 0.000000 -0.306633 -2.732455 1.612753
2 1 1.000000 1.000000 0.000000 0.099804 0.487968 -1.545902
3 1 1.000000 0.000000 1.000000 -0.500267 0.777696 0.028699
4 1 0.000000 1.000000 1.000000 -0.404407 -0.867741 0.626161
...
```

where we only have included the first four of 500 atoms. This file can then be used as a starting point for a simulation. The output from a simulation will have a similar format, providing the state of the atomic system which we can then analyze in details.

3.2.3 Integration of the equations of motion

Starting from the initial configuration, we integrate the equations of motion to find the particle trajectories and output them to file for further analysis. First, we write a short function to read the initial configuration from file, and then we integrate and write results to a file.

The function readlammps[1] used to read from a LAMMPS trajectory file will be reused many times throughout this book, hence we strive to make it sufficiently general and sophisticated. You can find the listing and discussion of this program in the Appendix XXX Here, we simply notice that we can call it as

```
[Lx,Ly,Lz,r,v]=readlammps('mymdinit.lammpstrj');
```

In order to read in the initial state. The following program integrates the equation of motion:

```
[Lx,Ly,Lz,r,v]=readlammps('mymdinit.lammpstrj');
L = [Lx Ly Lz]; s = size(r); N = s(1);
t = 3.0; dt = 0.001; n = ceil(t/dt);
a = zeros(N,3);  % Store calculated accelerations
for i = 1:n-1 % Loop over timesteps
    a(:,:) = 0;
    for i1 = 1:N
        for i2 = i1+1:N
            dr = r(i,i1,:) - r(i,i2,:);
            for k = 1:3  % Periodic boundary conditions
                if (dr(k)>L(k)/2) then
                    dr(k) = dr(k) - L(k);
                end
                if (dr(k)<-L(k)/2) then
                    dr(k) = dr(k) + L(k);
                end
            end
            rr = dot(dr,dr);
            aa = -24*(2*(1/rr)^6-(1/rr)^3)*dr/rr;
            a(i1,:) = a(i1,:) + aa(1); % from i2 on i1
            a(i2,:) = a(i2,:) - aa(2); % from i1 on i2
        end
    end
    v(i+1,:,:) = v(i,:,:) + a*dt;
    r(i+1,:,:) = r(i,:,:) + v(i+1,:,:)*dt;
    % Periodic boundary conditions
    for i1 = 1:N
        for k = 1:3
            if (r(i+1,i1,k)>L(k))
                r(i+1,i1,k) = r(i+1,i1,k) - L(k);
            end
             if (r(i+1,i1,k)<0)
```

[1] http://folk.uio.no/malthe/fys2160/readlammps.m

```
                r(i+1,i1,k) = r(i+1,i1,k) + L(k);
            end
        end
    end
end
writelammps('mymddump.lammpstrj',Lx,Ly,Lz,r,v);
```

Notice how the periodic boundary conditions are implemented. They need to be included both when the relative distance $\Delta\mathbf{r}_{ij}$ between two atoms are calculated and when the positions of the atoms are updated during the integration step.

The main idea here is to show you the structure and principles of a molecular dynamics program, as well as how data is handled and stored in files. This python program is not practically useful because python is too slow, in particular for a program using many nested loops. We will therefore not use this code, but instead use a professionally developed open-source molecular dynamics integrator to find the time development of the atomic system. The principles are the same, but the computational efficiency is many orders of magnitude greater for a professional code.

3.3 Running a molecular dynamics simulation

There are several efficient packages that solves the equations of motion for a molecular dynamics simulation. The packages allow us to model a wide variety of systems, atoms and molecules, and are efficienty implemented on various computing platforms, making use of modern hardware on your laptop or desktop or state-of-the-art supercomputing fascilities. We use a particular tool developed at Sandia National Laboratories called LAMMPS[2].

3.3.1 Installation of LAMMPS

If you want to be able to reproduce the simulations performed here you will need to install LAMMPS on your computer. This is very simple if you have a Mac or an Ubuntu system — for a windows system you will have to follow the installation instructions found at the web-site for LAMMPS. You can find all the documentation of LAMMS here[3].

[2] `http://lammps.sandia.gov/`
[3] `http://lammps.sandia.gov/`

Mac installation. On a mac you should be able to install LAMMS using Homebrew or Conda.

Using **Homebrew** (Homebrew[4]) LAMMPS is installed with:

Terminal

```
brew tap homebrew/science
brew install lammps
```

If you want to use the parallell implementation of LAMMPS you can install this version using

Terminal

```
brew tap homebrew/science
brew install lammps --with-mpi
```

Using **Conda** (MacPorts[5]) LAMMPS is installed with:

Terminal

```
conda config --add channels conda-forge
conda create -n my-lammps-env

conda activate my-lammps-env
conda install lammps
```

Ubuntu installation. You can install a recent version of LAMMPS with:

Terminal

```
sudo apt-get install lammps
```

Python interface installation. In addition you should install the Atomic Simulation Environment[6] for Python which we use to read simulation data into Python. You can install this using `pip`:

Terminal

```
pip install ase
```

and it should be avaiable also for Jupyter notebooks.

[4] `http://brew.sh/`
[5] `https://www.conda.io/`
[6] `https://wiki.fysik.dtu.dk/ase/index.html`

3.3.2 Starting a simluation in LAMMPS

If you have successfully installed LAMMPS, you are ready to start your first molecular dynamics simulations. The LAMMPS simulator reads its instructions on how to run a simulation from an input file with a specific syntax. Here, we will set up a two-dimensional simulation of a Lennard-Jones system using the file `in.myfirstmd`:

```
# 2d Lennard-Jones gas
units lj
dimension 2
boundary p p p
atom_style atomic

lattice hex 0.75
region simbox block 0 20 0 10 -0.1 0.1
create_box 1 simbox
create_atoms 1 box

mass 1 1.0
velocity all create 2.5 87287

pair_style lj/cut 2.5
pair_coeff 1 1 1.0 1.0 2.5

neighbor 0.3 bin
neigh_modify every 20 delay 0 check no

fix 1 all nve

dump 1 all custom 10 dump.lammpstrj id type x y z vx vy vz
thermo 100
run 5000
```

The simulation is run from the command line in a Terminal. Notice that the file `in.myfirstmd` must be in your current directory. I suggest creating a new directory for each simulation, copying the file `in.myfirstmd` into the directory and modifying the file to set up your simulation, before starting the simulation with:

Terminal

```
lmp_serial < in.myfirstmd
```

This simulation should only take a few seconds. It produces output in the Terminal and two files: `dump.lammpstrj`, which contains all the data from the simulation, and `log.lammps`, which contains a copy of what was output to your terminal. Fig. 3.4 illustrates the positions of the atoms initially and at the end of the simulation.

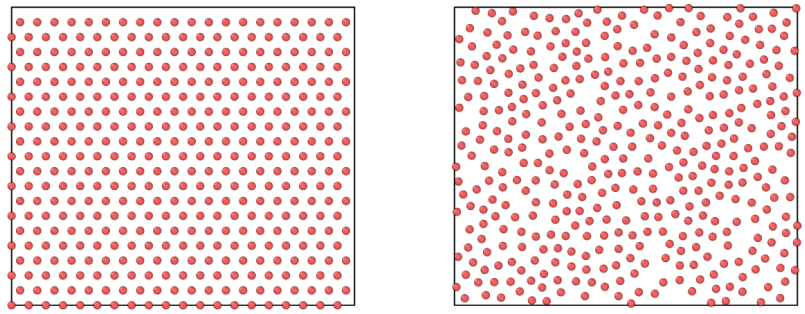


Fig. 3.4 (*Left*) Illustration of the initial hexagonal lattice generated in the simulation. (*Right*) Illustration of the final positions of the atoms after 5000 timesteps.

The input file `in.myfirstmd` consists of a series of commands to be interpreted by LAMMPS. Here, we look at what these do in detail. (You can skip this at first reading, and return when you wonder what the parameters actually do).

```
# 2d Lennard-Jones gas
```

This line starts with a `#` and is a comment that is ignored by the program.

```
units lj
dimension 2
boundary p p p
atom_style atomic
```

This block describes general features of the simulation:

The `units lj` command selects Lennard-Jones units, which were introduced in Sect. 3.2.1. This means that lengths are measured in units of σ, energies in units of ϵ_0, time in units of $\tau = \sigma\sqrt{m/\epsilon}$, and temperature in terms of $T_0 = \epsilon/k_B$. For Argon $\sigma = 0.3405\mu$m, and $\epsilon = 1.0318 \cdot 10^{-2}$eV. Other atomic models will have other parameters.

The `dimension` command specifies the dimensionality of the simulation: 2 or 3. Here we run a 2d simulation.

The `boundary` command specifies boundary conditions to be applied. Here we have periodic boundaries in the x-, y-, and z-directions.

The `atom_style` command specifies the complexity of the description of each atom/particle. Here, we will use the simplest description,

`atomic`, which is used for noble gases and coarse-grained simulation models.

```
lattice hex 0.75
region simbox block 0 20 0 10 -0.1 0.1
create_box 1 simbox
create_atoms 1 box
```

This block sets up the dimensions of the 20×10 simulation box and fills the box with atoms with a given packing fraction.

The `lattice` command generates a lattice of points. This does, surprisingly enough, not actually generate any atoms, it only generates a set of positions in space where atoms will be generated when we generate atoms. The type `hex` specifies a two-dimensional lattice of hexagonal shape, so that each atom has six nearest neighbors. And the number `0.75` is called the scale and is the reduced density, ρ', when we have chosen LJ units for the simulation. (Notice that the scale is interpreted differently if we do not use LJ units, see the LAMMPS documentation for more information).

The `region` command defines a region which is a `block` extending over $0 < x < 20$, $0 < y < 10$, $-0.1 < z < 0.1$ We give this region the name `simbox`.

The `create_box` command now actually creates the simulation box based on the spatial region we called `simbox`. The simulation box will only contain 1 (one) type of atoms, hence the number `1`.

The `create_atoms` finally fills the simulation box we have defined using the lattice we have defined with atoms of type `1`.

```
mass 1 1.0
velocity all create 2.5 87287
```

This block defines the material properties of the atoms and defines their initial velocities.

The `mass` command defines that atoms of type `1` will have a mass of `1.0` relative to the mass of the Lennard-Jones model. This means that all atoms have mass 1 in the Lennard-Jones units. This means that the masses of all the atoms are the same as the mass m used in the non-dimensionalization of the Lennard-Jones model.

The `velocity` command generates random velocities (using a Gaussian distribution) so that the initial temperature for `all` atom types in the system is `2.5` in the dimensionless Lennard-Jones units. The last, strange integer number `87287` is the seed for the random

number generator used to generate the random numbers. As long as you do not change the seed number you will always generate same initial distribution of velocities for this simulation.

```
pair_style lj/cut 2.5
pair_coeff 1 1 1.0 1.0 2.5
```

This block specifies the potential between the atoms.

The `pair_style` command specifies the we want to use a Lennard-Jones potential with a cut-off that is of length `2.5`. What does this mean? It means that since the Lennard-Jones potential falls so quickly to zero as the distance between the atoms increase, we will approximate interaction to be zero when the atoms are further than 2.5 away from each other (measured in Lennard-Jones units, that is in units of σ). The simulator ensures that both the potential and the force is continuous across the transition. There are many other types of force fields that you may use — take a look at the documentation of LAMMPS for ideas and examples.

The `pair_coeff` command specifies the parameters of the Lennard-Jones model. The two first numbers, `1 1`, specifies that we describe the interaction of atoms of type 1 with atoms of type 1. And the parameters of the Lennard-Jones model are `1.0 1.0 2.5`. This means that The interaction between an atom of type 1 with an atom of type 1 has a σ-value corresponding 1.0 times the the general σ-value (hence the first number 1.0), and a ϵ_0-value corresponding to 1.0 times the overall ϵ-value (hence the second number 1.0). The cut-off for this interaction is 2.5 — the same value as we specified above.

```
neighbor 0.3 bin
neigh_modify every 20 delay 0 check no
```

This block contains information about how the simulator is to calculate the interaction between the atoms using list of neighbors that are updated at regular intervals. You do not need to change these parameters, but changing them will typically not have any effects on the simulation results, only on the efficiency of the simulations.

```
fix 1 all nve
```

This one-line block specifies what type of simulation we are performing on the atoms. This is done by one or more `fix` commands that can be applied to regions of atoms. Here, the `fix`, which we

call `1` (you can choose numbers or names for identity), is applied to `all` particles and specifies that the simulation is run at constant `nve`, that is, at constant number of particles (`n`), constant volume (`v`, meaning that the simulation box does not change during the simulation), and constant energy (`e`). You may be surprised by the constant energy part. Does the integration algorithm ensure that the energy is constant. Yes, it does. However, there can be cases where we want to add energy to a particular part of the system, and in that case the basic interation algorithm still conserves energy, but we add additional terms that may change the total energy of the system.

```
dump 1 all custom 10 dump.lammpstrj id type x y z vx vy vz
thermo 100
run 5000
```

This block specifies simulation control, inclusing output and the number of time-steps to simulate.

The `dump` command tells the simulator to output the state. The `1` is the name we give this dump — it could also have been given a name such as `mydump`. We specify that `all` atoms are to be output using a `custom` output format, with output every `10` time-steps to the file `dump.lammpstrj`, and the 'id type x y z vx vy vz' list specifies what is output per atom.

The `thermo` command specifies that output to the Terminal and to the log file, `log.lammps`, is provided every `100` timesteps.

The `run` command starts the simulation and specifies that it will run for `5000` timesteps.

3.3.3 Visualizing the results of a simulation

It is good practice to look at the results of the simulation. Use for example VMD[7] or Ovito[8] to visualize the results. Here, let us demonstrate the use of VMD. First, open VMD. Then open the `File -> New Molecule` menu item. You find the `dump.lammpstrj` file that were generated by your simulation run, and press load. The whole time sequence of your simulation is now loaded in VMD. However, the default visualization

[7] `http://www.ks.uiuc.edu/Research/vmd/`
[8] `http://www.ovito.org/`

mode is usually not that clear. You fix that by going to the main menu again, and select `Graphics -> Representations...` In the window that opens you change the `Drawing method` to `CPK` and you select `Bond Radis` to be 0.0. Now, your visualization should show the time dynamics of an ensemble of atoms. Enjoy!

3.4 Analyzing the results of a simulation

The results from the simulations can be analyzed using built-in tools from LAMMPS. We demonstrate these tools by measuring the number of particle on the left side of the box as a function of time, by extracting that data from the simulation files.

We have illustrated the text-format output files from LAMMPS above. The file may contain data from may timesteps, t. For each timestep, there are a few initial lines describing the system, followed by a list of properties for each atom. We have chosen the custom output form

```
dump 1 all custom 10 dump.lammpstrj id type x y z vx vy vz
```

Each line of atom properties contains data of the form `id type x y z vx vy vz`. This means that we will get access to both the atom position and the atom velocity for each of the atoms if we read the data in this file. Our strategy will then be to use the atom position x_i and compare it with $L_x/2$, where L_x is the dimensions of the simulation box, measured using the same units as used for the positions x_i of the atoms. How can we read the data from the LAMMPS dump files. We could read our own input functions, but it is simpler to use already developed code, which is distributed with the LAMMPS distribution and the other libraries mentioned above.

3.4.1 Matlab implementation

For matlab, we use the tools found in the `tools/matlab` directory, which was part of the lammps installation. (If you cannot find these after installing LAMMS, you can always download LAMMS again as a tarball, extract the tarball, and find the programs in the resulting directory). You need to copy the file `readdump_all.m` and `readdump_one.m` to your current working directory — the directory where you ran the simulation and where the file `dump.lammpstrj` is located.

First, we read data from all the timesteps into matlab:

```
data = readdump_all('dump.lammpstrj');
```

We need the time values t_i at each of the timesteps and the number of time-steps, `nt`:

```
t = data.timestep;
nt = length(t);
```

We set up an array, `nleft`, to store $n(t)$, the number of atoms on the left side as a function of time:

```
nleft = zeros(nt,1);
```

The size of the simulation box is stored in the variable `box`, and we store $L_x/2$ in the variable `halfsize`:

```
box = data.x_bound;
halfsize = 0.5*box(:,2);
```

Now, we loop through all the timesteps. For each timestep we extract all the x-positions for all the atoms in a list `xit`. Next, we find a list of all the atoms that are on the left side, that is, all the atoms with an x-position smaller than $L_x/2$. This is done by the `find` command. Finally, we count how many elements are in this list. This is the number of atoms that are on the left side of the box. We store this number in the array `nleft` for this timestep, and reloop to address the next timestep:

```
for it = 1:nt
   xit = data.atom_data(:,3,it);
   jj = find(xit<halfsize(it));
   nleft(it) = length(jj);
end
```

The complete program reads:

```
%start1
data = readdump_all('dump.lammpstrj');
t = data.timestep;
nt = length(t);
nleft = zeros(nt,1);
box = data.x_bound;
halfsize = 0.5*box(:,2);
for it = 1:nt
   xit = data.atom_data(:,3,it);
   jj = find(xit<halfsize(it));
   nleft(it) = length(jj);
end
plot(t,nleft), xlabel('t'), ylabel('n')
```

A short and simple program to handle a complicated set of data. This should provide you with a good basis for measuring various properties of the simulated system using Python.

3.4.2 Python implementation

For Python we use the `ase` package which you can import and use in your python scripts.

First, we read data from all the timesteps into Python:

```
import numpy as np
import matplotlib.pyplot as plt
from ase.io import read
traj = read("twosec02.lammpstrj", format="lammps-dump-text", index=":")
```

Then number of timesteps is given as the number of elements in the `traj` array:

```
n = len(traj) # Number of timesteps
```

We set up an array, `nleft`, to store $n(t)$, the number of atoms on the left side as a function of time:

```
nleft = np.zeros(n)
```

The size of the simulation box is found from the cell size, which is found using the method `get_cell`, and we store $L_x/2$ in the variable `halfsize`:

```
firstframe = traj[0]
cellshape = firstframe.get_cell()
halfsize = cellshape[0][0]*0.5
# Count number of particles on left side for each timestep
```

Now, we loop through all the timesteps. For each timestep we extract all the x-positions for all the atoms in a list `xi`. Next, we sum up the number of atoms that are on the left side, that is, all the atoms with an x-position smaller than $L_x/2$. This is done by the `np.sum` command. We store this number in the array `nleft` for this timestep, and reloop to address the next timestep:

```
it = 0
for frame in traj:
    x = frame.get_positions() # positions of atoms
    xi = x[:,0] # x-coordinates of atoms
    nleft[it] = np.sum(xi<halfsize)
    it = it + 1
```

The complete program reads:

```
#start1
import numpy as np
import matplotlib.pyplot as plt
from ase.io import read
traj = read("twosec02.lammpstrj", format="lammps-dump-text", index=":")
n = len(traj) # Number of timesteps
nleft = np.zeros(n)
# Find size of simulation box
firstframe = traj[0]
cellshape = firstframe.get_cell()
halfsize = cellshape[0][0]*0.5
# Count number of particles on left side for each timestep
it = 0
for frame in traj:
    x = frame.get_positions() # positions of atoms
    xi = x[:,0] # x-coordinates of atoms
    nleft[it] = np.sum(xi<halfsize)
    it = it + 1
plt.plot(nleft)
```

A short and simple program to handle a complicated set of data. This should provide you with a good basis for measuring various properties of the simulated system using Python.

3.4.3 Results

The resulting plot is shown in fig. 3.5. If you wonder why the number of atoms on the left is varying, how it is varying, and how to describe how it is varying — this indeed in the topic of this book so just read on!

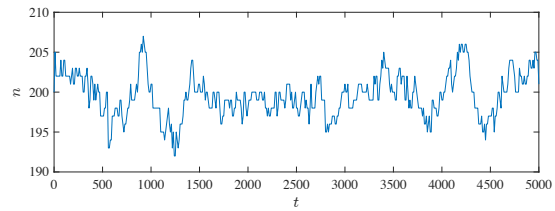


Fig. 3.5 Plot of $n(t)$, the number of atoms on the left side of the simulation box, as a function of time, where the time is measured in the number of timesteps in the molecular dynamics simulation.

3.5 Advanced models

Molecular dynamics simulation packages such as LAMMPS are professional tools that are used for research and development. Here, I provide

a few examples of more advanced use of LAMMPS that may inspire you to try the tool also on your own.

3.5.1 Coarsening

Out first example is an extension of the first two-dimensional simulation you performed above. What happens if we start the system with a homogeneous distribution of atoms, but with a low initial energy? In addition, we keep the average kinetic energy in the system approximately constant. (This corresponds, as you will learn later, approximately to keeping the temperature in the system constant):

```
units lj
dimension 2
boundary p p p
atom_style atomic

lattice hex 0.5
region simbox block 0 80 0 40 -0.1 0.1
create_box 1 simbox
create_atoms 1 box

mass 1 1.0
velocity all create 0.05 87287

pair_style lj/cut 2.5
pair_coeff 1 1 1.0 1.0 2.5

neighbor 0.3 bin
neigh_modify every 20 delay 0 check no

fix 1 all nvt temp 0.25 0.25 1.0
dump 1 all atom 1000 dump.lammpstrj
thermo 100
run 50000
```

The resulting behavior shown in Fig. 3.6 is a phase separation: The system separates into a liquid and a gas phase, demonstrating a phenomenon called spinoidal decomposition. We will discuss the behavior of such systems using both molecular dynamics models and algorithmic models later.

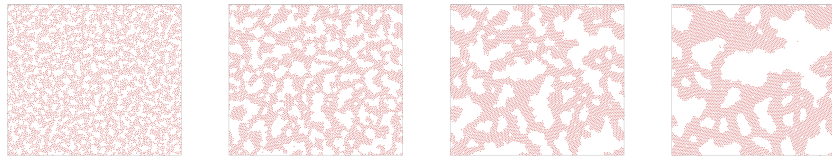


Fig. 3.6 Snapshots from a simulation of a Lennard-Jones liquid, with a low initial density. A slowening coarsening is observed.

3.5.2 Hemoglobin in water

3.5.3 Lipid bilayers

3.5.4 Fracture of Silicon crystal

3.5.5 Water permeation in a carbon nanotube

3.6 Summary

- Molecular dynamics simulations model how atomic systems propagate forward in time by integrating the equations of motion based on a given interaction potential between the atoms.
- We have introduced a commonly used interaction potential, the Lennard-Jones potential: $V(r) = 4\epsilon\left((\sigma/r)^{-12} - (\sigma/r)^{-6}\right)$, where the behavior is determined by the energy scale ϵ_0 and the length scale σ.
- We run molecular dynamics simulations using custom code, public or commercial codes, which include efficient integrators.
- A typical simulation models a system with constant energy, volume, and number of particles.
- Molecular dynamics simulations are routinely used as part of research and development tasks across disciplines, and are important tools for applications such as the development of new materials, chemical catalysis optimization and drug development.

3.7 Exercises

3.7.1 Problems

Exercise 3.1: Evolution of density

We will here study a $10 \times 10 \times 10$ system of Argon atoms, as done above. Initiate the system with all particles in the left 1/3 of the system. We call the number of particles in the left, center and middle thirds of the system $n_1(t)$, $n_2(t)$, and $n_3(t)$ respectively.

a) How do you expect n_1, n_2, and n_3 to develop in time?

b) Measure $n_1(t)$, $n_2(t)$, and $n_3(t)$ and compare with your predictions.

Exercise 3.2: Temperature variation

a) Measure the maximum and minimum temperature, T_{max} and T_{min} respectively for a simulation of $L = 10 \times 10 \times 10$ atoms, and calculate $\Delta T = T_{\text{max}} - T_{\text{min}}$.

b) Measure ΔT for three different system size, $L = 4$, $L = 8$, and $L = 16$. Comment on the results.

Exercise 3.3: Coarsening

Start a LAMMPS simulation with the following init file:

```
# 3d Lennard-Jones melt
units       lj
atom_style  atomic
lattice     fcc 0.2
region      box block 0 10 0 10 0 10
create_box  1 box
create_atoms    1 box
mass        1 1.0
velocity    all create 0.1 87287
pair_style  lj/cut 2.5
pair_coeff  1 1 1.0 1.0 2.5
neighbor    0.3 bin
neigh_modify    every 20 delay 0 check no
fix 1 all nvt temp 0.05 0.05 1.0
dump        id all atom 50 dump.lammpstrj
thermo      100
run     10000
```

(This simulation is performed at constant temperature and not at constant energy). Visualize the time development of the system and comment on what you see.

Probability and Statistics

4

Abstract This chapter will introduce you to some of the most common and useful tools from statistics, tools that you will need throughout your career. We will start from an introductory example — trying to understand the data you generated by running a molecular dynamics simulations. This is a good starting point, since statistics is about describing real data. The data from a realistic molecular dynamics simulation can be quite overwhelming. The data-sets become extremely big, so we need effective tools to pick out the most important aspects. Here, we introduce the concept of probabilities, the basic properties of probabilities, such as the rule of addition and the rule of multiplication. We introduce probability distributions and densities. We introduce the expected value and the variance and their estimators, the average and the standard deviation. We introduce the binomial, normal and exponential distribution. We demonstrate that the distribution of a sum of independent variables usually are normally distributed, independently of the details of the distribution of each event in the sum. Finally, we use these tools to develop a simple theory for the number of particles on the left side of a gas.

Why study statistics? Statistics provides us with the tools to describe real data, whereas probability theory provides us with the theoretical underpinning for statistics. Here, we will start with a practical approach to probability theory, based on the idea of frequencies of events, and use this to develop the fundamental laws of probabilities. This approach is natural for a physicist and is also the most direct approach from a

computational perspective. Again, we will introduce both theory and the computational tools needed to test the theory and to make accurate measurements for large data sets.

Probability theory and statistics. Probability theory is a classic field that has occupied many of the great mathematicians and physicists throughout history. Probability theory is the tool we need to address a world of real data and unknown processes, and the tool we need to describe the behavior of systems with many particles, where the physical laws no longer are absolute, but only represent the overwhelmingly most probable outcome of an experiment or an observation. However, we will start our venture into statistics and probability theory from observations of real data — in our case the data from measuring the number of atoms on the left side of a box filled with a gas.

4.1 Motivating example — fluctuations in a gas

A molecular dynamics simulations of a gas produces fluctuations from the approximately random motion of the atoms. Therefore, measures of physical quantities will also fluctuate. For example, the number of atoms on the left side of the box, $n(t)$, will vary in time. How can we characterize these variations?

Generating data for the gas. First, we follow the procedure from Chap. 3 to model a two-dimensional gas using the script:

```
# 2d Lennard-Jones gas
units lj
dimension 2
atom_style atomic
lattice hex 0.10
region box block 0 20 0 10 -0.1 0.1
create_box 1 box
create_atoms 1 box
mass 1 1.0
velocity all create 2.5 87287
pair_style lj/cut 2.5
pair_coeff 1 1 1.0 1.0 2.5
neighbor 0.3 bin
neigh_modify every 20 delay 0 check no
fix 1 all nve
dump 1 all custom 10 gasstat01.lammpstrj id type x y z vx vy vz
thermo 100
run 50000
```

which is run by

Terminal

```
lmp_serial < in.gasstatistics01
```

We apply the method from Sect. 3.4 to extract $n(t)$, the number of atoms in the left half of the system ($x < L_x/2$) as a function of time t:

```
import numpy as np
import matplotlib.pyplot as plt
from ase.io import read
traj = read("gasstat01.lammpstrj", format="lammps-dump-text", index=":")
n = len(traj) # Number of timesteps
nleft = np.zeros(n)
t = np.zeros(n)
# Find size of simulation box
firstframe = traj[0]
cellshape = firstframe.get_cell()
halfsize = cellshape[0][0]*0.5
# Count number of particles on left side for each timestep
it = 0
for frame in traj:
    x = frame.get_positions() # positions of atoms
    xi = x[:,0] # x-coordinates of atoms
    nleft[it] = np.sum(xi<halfsize)
    t[it] = it
    it = it + 1
plt.plot(t,nleft), xlabel('$t$ (timesteps)'), ylabel('$n$')
np.savetxt('ndata.d',(t,nleft))
```

The result is a sequence of numbers $n_i = n(t_i)$ for a sequence of timesteps i, as illustrated in Fig. 4.1. The resulting sequence t_i, n_i is written to the file `ndata.d` using the command `savetxt` for simple retrieval and analysis.

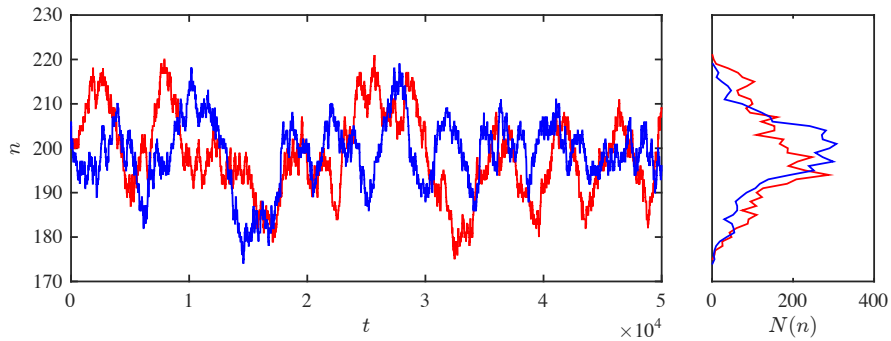


Fig. 4.1 Plot of the number of atoms on the left half (n) as a function of time (t) for 50000 timesteps for two runs with different random seeds for the initial velocity distribution.

General features of $n(t)$. Now we can simply read in the data from the file `ndata.d`:

```
import numpy as np
t,n = np.loadtxt('ndata.d')
```

What does this sequence of numbers tell us about the gas? Usually in physics, we try to understand systems and data by building models that reproduce the main features of the observed data. Our goal in this chapter is to develop such a model. But how do we know if the model is any good?

First, we notice that the data has some randomness to it. If we did another simulation, but with a different random seed for the initial velocities of the atoms, we would expect the behavior to be different — at least not identical — but also to retain some of the features. This is illustrated in Fig. 4.1, where two measured curves $n(t)$ are illustrated. The curves have similar features, but are clearly not identical. What general features of $n(t)$ should a model predict or explain? Since the curves are not identical, our model will clearly not reproduce $n(t)$ in detail, but instead it should reproduce some of its general properties.

Frequencies of particular values of n. For example, we could characterize $n(t)$ by how often it visits a particular value of n: We count the number of times, $N(x)$, when there are x atoms on the left side, that is, how many times $n(t) = x$ in a given time interval. This corresponds to what we call a histogram:

```
import matplotlib.pyplot as plt
plt.hist(n)
```

A plot of $N(x)$ is shown in Fig. 4.1, in the same figure as the data-set. This gives us a first insight into the frequency of occurrence of the various values of n. And we immediately observe that not all n-values occur equally frequently: values close to $n = 200$ occur more frequently that values further away from $n = 200$.

There is a whole field of science dedicated to studying frequencies of occurrence and histograms of frequencies — the field of probability theory and statistics. This chapter therefore contains a brief introduction to statistics and probability theory, before we develop a model for the molecular dynamics data.

4.2 Probabilities

4.2.1 Statistical experiments

The sequence n_i of atoms on the left side of the gas can be thought of as a sequence of statistical experiments, where the result of each experiment i is a number n_i. Let us first develop a better understanding of statistical experimenents by modeling and studying such sequences.

In statistics, the term statistical experiment has a precise meaning:

> A **statistical experiment** is a trial with a number of possible outcomes or events. The result of an experiment is called an **outcome** or an **event** and the set of all possible outcomes is called the **sample space**.

Typically, we will consider statistical experiments that are **uncorrelated**, meaning that if we know the result of the previous experiment, this will not help us know the result of the next experiment. (However, many physical systems, such as the gas, are **correlated**, because the state of the system may not have changed completely between two subsequent measurements). For example, throwing a fair die or flipping a fair coin are considered uncorrelated experiments. The outcome of such an experiment is the number, m_i, on the die.

4.2.2 Generating a sequence of random numbers

We can perform a numerical statistical experiment by using the pseudo-random number generator[1] in python. We generate a random number between 1 and 6 by:

```
import numpy as np
np.random.randint(0,6)
```

```
3
```

(Here, the first argument is the smallest random integer and the second argument is the largest random integer, and an optional third argument

[1] The random number generator only produces *pseudo-random* numbers based on a deterministic algorithm that generates a sequence of apparently random numbers. However, for all our purposes these numbers appear to be random.

is the number of outcomes to produce). We can use the same method to generate a sequence of 4 such experiments by:

```
np.random.randint(0,6,4)+1
```

```
array([1, 4, 4, 2])
```

The result is a sequence, n_i, $i = 1, \ldots, M$ of M outcomes from a numerical statistical experiment. How can we understand this experiment by analyzing the outcomes?

4.2.3 Measuring the frequency of occurrence

First, let us measure the frequency of occurrence of a particular outcome: We count how many times a given value x occurs in the sequence of numbers, n_i, by generating a histogram. A histogram records how many times, $N(x)$ a value x occurs in the interval from x_j to $x_{j+1} = x_j + \Delta x$, where Δx is called the *bin width* of the histogram. python has automatic functions for finding the histogram for a data set. We generate 1000 random numbers between 1 and 8 and find the histogram using:

```
n = np.random.randint(0,8,1000)+1
[Nx,edges] = np.histogram(n,8,range=(0.5,8.5))
print(Nx)
```

```
array([127, 131, 128, 135, 116, 140, 127,  96])
```

```
print(edges)
```

```
[ 0.5  1.5  2.5  3.5  4.5  5.5  6.5  7.5  8.5]
```

(Notice the range used in order to ensure there are 8 bins of size 1 each). Here, the `edges` are the x_j values describing the edges of the bins. We can find the centers of the bins and plot the results using:

```
x = (edges[1:]+edges[:-1])/2
print(x)
```

```
[ 1.  2.  3.  4.  5.  6.  7.  8.]
```

```
plt.plot(x,Nx)
```

We can compare this with a traditional bar-type histogram plot using

```
plt.plot(x,Nx)
plt.hist(n,bins=8)
```

The resulting plot is shown in Fig. 4.2. We see that the frequencies are approximately, but not exactly, the same for all the outcomes. How can we interpret the frequency of occurrence?

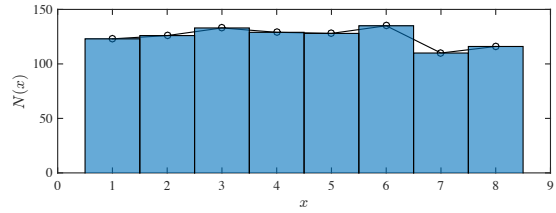


Fig. 4.2 Plot of the histogram of 1000 outcomes from an experiment where we choose a random number between 1 and 8.

4.2.4 Frequencies and probabilities

When we analyze a statistical experiment, or data from experiments or simulations, we will assume that an experiment reflects an underlying process which has given properties. We estimate these properties using various **estimators**. For a sequence of outcomes n_i we can measure the frequency $F(x)$ by

We observe a sequence of M outcomes, n_i, $i = 1, \ldots, M$, from a statistical experiment. The **frequency** $F(x)$ is defined as how many times, N_x, we have measured x in the sequence — that is the number of times $n_i = x$ — divided by the number of elements, M, in the sequence:

$$F(x) = \frac{N_x}{M} \,. \tag{4.1}$$

The frequency is an estimator for an underlying property of the process — the *probability* of an outcome. We can define the probability as the frequency when we do infinitely many experiments:

The **probability**, $P(n = x)$, of an outcome x is the limit of the frequency of the outcome when we perform infinitely many experiments:

$$P(n = x) = \lim_{M\to\infty} F(x) = \lim_{M\to\infty} \frac{N_x}{M} \,. \tag{4.2}$$

This limit is not a mathematical limit as known from calculus, but instead a limit of inifintely many experiments. We call this the **frequency definition of probability**.

This is not the only possible definition of probability. In statistics you may meet other definitions. But it is a practical and useful definition for our purposes. Also, it reflects the important difference between an *underlying property* of a process, such as the probability for an outcome, and the *estimator for a property*, such as the frequency.

4.2.5 Example: Probability of throwing a single dice

We throw a fair dice with six sides. What is the probability to throw a 1*?*

Numerical experiment. First, we perform a numerical statistical experiment to gain intuition. We generate M throws of a die with a result between 1 and 6 and plot the frequency, $F(x) = N(x)/M$, for experiments with increasing values of M:

```
import numpy as np
import matplotlib.pyplot as plt
Mval = np.array([10,100,1000,10000,100000])
nM = len(Mval)
for im in range(nM):
    M = Mval[im]
    m = np.random.randint(0,6,M)+1
    [Nx,edges] = np.histogram(m,6,range=(0.5,6.5))
    Fx = Nx/M
    x = (edges[1:]+edges[:-1])/2
    plt.plot(x,Fx,'-o',label='L = '+str(M))
plt.xlabel('x'), plt.ylabel('F(x)'), plt.legend()
```

The results are shown in Fig. 4.3.

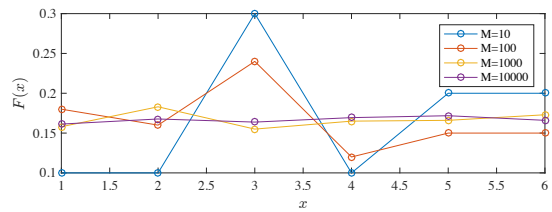


Fig. 4.3 Plot of the frequency of occurrences, $F(x)$ for $x = 1, 2, 3, 4, 5, 6$ for $M = 10, 100, 1000$, and 10000 throws.

We see that the frequencies, $F(x) = N(x)/M$ converges as M increases. We expect the frequencies to converge towards the underlying probabilities, $P(x)$. What are the values of $F(x)$ for $M = 10000$?

```
print(Fx)
```

```
array([0.1613, 0.1677, 0.1638, 0.1694, 0.1718, 0.1660])
```

Both these results, and the histogram plots in Fig. 4.3, indicate that the probability is a constant, the same, for all possible outcomes x. This is what we would expect from a *fair* dice. Otherwise, we would know that the dice preferred a particular outcome, and it would not be fair.

Theretical argument. We can also devise a theoretical argument for the value of the probability $P(x)$ for a throw of the die. From the definition of probability, the probability to get i is

$$P(i) = \frac{N_i}{M} \,, \tag{4.3}$$

where M is the number of throws and i may be 1,2,3,4,5 or 6 — valid in the limit when M becomes large. From this definition we see that probabilities are *normalized*:

$$\sum_i P(i) = \frac{\sum_i N_i}{M} = \frac{M}{M} = 1 \,, \tag{4.4}$$

where $\sum N_i = M$ because in each throw one and only one of the 6 possible outcomes are possible. This gives us *the normalization property of probabilities*:

$$\sum_i P(i) = 1 \,, \tag{4.5}$$

when any two outcomes i cannot occur at the same time and the sum is over all possible outcomes.

We can use this property to calculate the probability to throw a 1. This follows from two characteristics of the experiments. First, the dice is fair, which means that $P(i)$ are the same for all i, $P(i) = \text{const} = p$. Second, we use the normalization condition:

$$\sum_{i=1}^{6} P(i) = \sum_{i=1}^{6} P = p \sum_{i=1}^{6} = p \cdot 6 = 1 \;\Rightarrow\; p = \frac{1}{6} \,, \tag{4.6}$$

This example uses a standard set of tricks that you may want to remember: If all outcomes are equally probable we can assume they all have the same probability p, and the probabilities are normalized.

We can also illustrate another basic principle of probabilities that also serves as a useful trick: The probability of something *not happening* can be found from the normalization condition. For example, what is the probability of not throwing a one in a single throw?

Here, we use a trick and split all possible outcomes into $P(\text{one})$ and $P(\text{not one})$. Since one of these two outcomes must occur, they must be normalized, so that

$$P(\text{one}) + P(\text{not one}) = 1 \,, \tag{4.7}$$

where we know that $P(\text{one}) = p = 1/6$. We can therefore find $P(\text{not one})$ from

$$P(\text{not one}) = 1 - P(\text{one}) = 1 - p = 1 - \frac{1}{6} = \frac{5}{6} \,, \tag{4.8}$$

4.2.6 Properties of probabilities

From this example, we have learned several new properties of probabilities that are general:

Normalization: Probabilities are **normalized** so that the sum of the probabilities of all the possible outcomes is 1:

$$\sum_x P(x) = 1 \,, \tag{4.9}$$

when any two outcomes x cannot occur at the same time and the sum is over all possible outcomes x. This is called the **normalization condition of probabilities**.

This can be proven from the frequency definition and that the total number of observations, $\sum_x N(x)$ must be the same as the total number of experiments, M:

$$\sum_x P(x) = \sum_x \frac{N(x)}{M} = \frac{1}{M} \sum_x N(x) = \frac{M}{M} = 1 \,. \tag{4.10}$$

Complement rule: We can always divide all the possible outcomes of an experiment into two parts, that an outcome A occurs and that the outcome A does not occurs. Either of these outcomes must occur, and they cannot occur at the same time. The normalization condition therefore gives:

$$P(\text{A}) + P(\text{not A}) = 1 \;\Rightarrow\; P(\text{A}) = 1 - P(\text{not A})\ . \tag{4.11}$$

Sets of outcomes. Notice that outcomes from an experiment can be grouped in various ways. For example, when we throw a six-sided dice, we can group the set of all possible outcomes into two set: set A = 1,2, and set B = 3,4,5,6. These two sets are *independent*, meaning that set A cannot occur at the same time as set B, and they span the total *sample space*, the set of all possible outcomes. The normalization rule can therefore be applied to sets A and B: $P(A) + P(B) = 1$. (It is customary to write the probability for a set A as $P(A)$).

Sets do not need to be independent. For example, we could define three sets A = 1,2,3, B = 3,4 and C = 4,5,6. These sets span the sample space, but they are not independent. We can therefore not apply the normalization rule to these sets: $P(A) + P(B) + P(C) \neq 1$.

Random variable. If the outcome of an experiment is a number, such as for the throw of a die, we call the outcome a *random variable*, and usually denote it with a *upper case symbol* for the random variable and a lower case sumbol for the outcome of an experiment. For example, for the throw of a die, we would use the symbol N for the number shown in the die, and the symbol n for a particle outcome. The probability to observe the outcome n when we measure the random variable N can therefore be written as the probability for N to be equal to n: $P(N = n)$.

Rule of addition. For two independent sets of outcomes, A and B, the probability that A or B occurs is given as the sum of their probabilities:

Addition rule: For two independent sets A and B, the probability for A or B is

$$P(\text{A or B}) = P(\text{A}) + P(\text{B})\ . \tag{4.12}$$

This rule can be proven using the frequency definition of probabilities: If we perform M experiments, the number of results in set A is $N(A)$

and the number of results in set B is $N(B)$. The number of results in A or B is therefore $N(A)+N(B)$ and the probability for A or B is therefore $P(\text{A or B}) = (N(A) + N(B))/M = P(A) + P(B)$.

We can extend this rule also to cases where A and B not are independent:

General addition rule: For two sets A and B, the probability for A or B is

$$P(\text{A or B}) = P(\text{A}) + P(\text{B}) - P(\text{A and B}) \; . \tag{4.13}$$

Where the last term is zero if A and B are independent, because in that case they cannot both occur at the same time.

4.2.7 Probability for simultaneous outcomes

What would happen if we threw a dice two times? How can we find the probability to throw a 1 in the first throw and a 6 in the second throw? This corresponds to asking questions about the outcomes from two experiments. We could write the probability for both of these events to occur at the same time as:

$$P\left(n_1 = 1 \text{ and } n_2 = 6\right) \; , \tag{4.14}$$

We could call the event that we get a 1 in the first throw A and the event that we get a 6 in the second throw B. We can then write the probability for both to occur as $P(A \text{ and } B)$. This is usually written at $P(A \cap B)$, using the intersection sign $\cap$ because the outcome is the intersection of A and B, the set of outcomes that include both A and B.

Numerical approach. First, we estimate the probability for (A and B) numerically: We draw two random numbers, n_1 and n_2, many times, and count how many times, N, we get a $n_1 = 1$ and $n_2 = 6$ at the same time. First, we generate the numbers n_1 and n_2:

```
import numpy as np
n1 = np.random.randint(0,6,1000)+1
n2 = np.random.randint(0,6,1000)+1
```

We can find all the values that are equal to 1 or 6 using the equal to operator

```
i1 = 1*(n1==1)
i2 = 1*(n2==6)
```

Where we have multiplied with one to ensure the result is an integer and not `True/False`. Then we find how many times $n_1 = 1$ and $n_2 = 1$ at the same time by summing `i1` and `i2`. Where this sum is equal to 2, both $n_1 = 1$ and $n_2 = 6$. We then simply have to count how many times the sum is equal to 2 to find N, the number of times $n_1 = 1$ and $n_2 = 6$ at the same time, and the frequency, $F = N/M$, of this:

```
m = 1*(i1+i2)==2
N = sum(m)
F = N/1000.0
print(F)
```

```
0.034
```

The result of this experiment is therefore $p \simeq F = N/M = 0.034$.

Theoretical approach. How can we develop a theoretical answer for this problem? There are several approaches we can follow. We can use the same approach as we used numerically above: We can look at all the possible outcomes from the experiment which consists of throwing a dice two times. How many outcomes are there in total? The first die has 6 possible outcomes, and the second die has 6 outcomes. In total, there are $6 \times 6 = 36$ different outcomes. (Notice in particular, that the outcome where the first die is a 2 and the second is a 4 is different from the outcome where the first die is a 4 and the second die is a 2. We have to count both these). All these outcomes are equally likely, and will have the same probability, if the die is fair. This means that the probability p for any such outcome is found from the normalization rule:

$$\sum_{i=1}^{36} p = p\,36 = 1 \;\Rightarrow\; p = 1/36\,. \tag{4.15}$$

The probability for a particular outcome, $n_1 = 1$ and $n_2 = 6$ is the probability for one of these 36 outcomes, and hence the probability for this is 1/36.

This approach is very robust. We can easily use the same approach to find the probability for the first die to be equal to 1 or 2 and the second die to be equal to 6. Can you find the argument and the answer?

Alternative theoretical approach. Notice that we could have used an alternative approach: In order to get a 1 in the first throw and then a 6 in the second throw, we see that there are 6 possibilities for the first

throw, and only 1 of these give a 1. That is, the probability to get a 1 in the first throw is $p_1 = 1/6$. Then, in the second throw, there are six possible outcomes, but only one of them give a 6. Thus the probability to get a 6 in the second throw is $p_2 = 1/6$. What is the probability to get both, that is a 1 in the first throw and a 6 in the second throw. The total number of outcomes is 6×6 and the number of outcomes that give $n_1 = 1$ and $n_2 = 6$ is 1×1. Hence, the probability for both is

$$p = \frac{1 \times 1}{6 \times 6} = \frac{1}{6}\,\frac{1}{6} = p_1\, p_2 \ . \tag{4.16}$$

This indicates that the probability for (A and B) is the product of the probability for each of the events — given that the events cannot occur at the same time. (If they could occur at the same time, that is, that they are not independent, then we could not find the total number of outcomes by multiplying the number of outcomes in throw 1 and the number of outcomes in throw 2).

Rule of multiplication. This rule is indeed a general law in statistics, which we call the rule of multiplication:

Rule of multiplication: For two *independent* events A and B, the probability to observe A and B at the same time is the product of the probability to observe A and the probability to observe B:

$$P(\mathrm{A} \cap \mathrm{B}) = P(\mathrm{A})\, P(\mathrm{B}) \ , \tag{4.17}$$

4.2.8 Example: Throwing two dice

Let us answer various questions for the results of two fair dice.
What is the probability to throw 1,2 or 3 in the first and 4,5 or 6 in the second throw?

We can use the rule of multiplication to find this. The probability to throw 1,2 or 3 in the first throw can be found from the addition rule. The outcomes 1,2 or 3 cannot occur at the same time. Hence the probability to observe 1,2 or 3 is the sum of the probability for each, $p_1 = P(1) + P(2) + P(3) = p + p + p = 1/6 + 1/6 + 1/6 = 3/6$. Similarly, the probability to throw 4,5 or 6 in the second throw is $p_2 = 3/6$. The rule of multiplication then gives that the probability for both is $p_1\, p_2 = (3/6)\,(3/6) = 9/36 = 1/4$.

What is the probability that the sum of the dice is 7?

We can use the rule of addition to find this. We need to find all the possible outcomes that add to 7. Each of the outcomes have the same probability (1/36) and they cannot occur at the same time. What outcomes add to seven? We write the outcome as two numbers where $(1, 2)$ means that the first throw is a 1 and the second is a 2, whereas $(2, 1)$ means that the first throw is a 2 and the second throw is a 1. The outcomes that add to 7 are $(1, 6)$, $(2, 5)$, $(3, 4)$, $(4, 3)$, $(5, 2)$, and $(6, 1)$, six outcomes in total. The probability that the sum is 7 is therefore $P(n_1 + n_2 = 7) = 6\,(1/36) = 1/6$.

What is the probability that at least one of the dice show a 1?

This problem can be solved in several ways. We could count the number of outcomes (n_1, n_2) where at least one of the numbers is a 1: There are 11 such outcomes. This means that probability for at least one 1 is 12/36. Can you think of a way to use the rule of multiplication? (Be careful to ensure that the events you are multiplying cannot occur at the same time!)

4.3 Expected value and Variance

We have now found that we can describe a random process by the probability for a given outcome. However, there are even simpler descriptions that can be measured directly from the sequence of measurements, the *average* and the *standard deviation*:

4.3.1 Average

For a sequence of outcomes n_i, $i = 1, \ldots, M$, the **average** of the outcome is defined as the arithmetic average, $\bar{n}$, of the outcomes:

$$\bar{n} = \frac{1}{M} \sum_{i=1}^{M} n_i \ , \tag{4.18}$$

4.3.2 Expected value

The average is an estimator for an underlying property of the process, which we call the *expected value* of the random variable N. We can see the

relation between the estimator and the underlying property by replacing the sum in the average by a sum over all possible outcomes, x, and the number of times each outcome x occurs, N_x:

$$\bar{n} = \sum_{i=1}^{M} \frac{n_i}{M} = \sum_x \frac{N_x\, x}{M} = \sum_x \left(\frac{N_x}{M}\right) x = \sum_x F(x)\, x \;, \tag{4.19}$$

In the limit of infinitely many measurements, the frequency $F(x) \to P(x)$, and we can replace the sum with a sum over probabilities

$$E(N) = \lim_{M\to\infty} \sum_x F(x)\, x = \sum_x P(x)\, x \;, \tag{4.20}$$

where we have introduced the notation $E(N)$ for this asymptotic value of the average. This is indeed what we will use as the *definition of the expected value*:

Expected value of a random variable: The expected value $E(N)$ of a random variable N is defined as:

$$E(N) = \sum_n P(n)\, n \;, \tag{4.21}$$

where $P(n)$ is the probability to observe outcome n.

We will also use the notation $\langle N \rangle = E(N)$ for the expected value.

The best way to estimate the expected value is to use the average as an *estimator* for the expected value. It can be shown that the average may be considered a best possible estimator for the expected value — using a specific definition of the word best. (You will learn more about what characterizes good estimators and how to find them in basic courses in statistics).

4.3.3 The expected value is linear

The expected value is *linear*. This means that the expected value of $Y = aX + b$ is

$$E(Y) = E(aX + b) = a\, E(X) + b \;. \tag{4.22}$$

We can prove this using the definition of the expected value in (4.21):

$$E(Y) = E(aX + b) = \sum_y P(y)\, y = \sum_x P(x)(ax + b) \tag{4.23}$$

$$= a \sum_x P(x)\, x + n \sum_x P(x) = aE(X) + b\,. \tag{4.24}$$

4.3.4 Expected value of a function $g(N)$

We will often be interested in the expected value of simple functions of random variables, such as N^2 or $\sin(N)$. We can find the expected value from the definition:

$$E(g(N)) = \langle g(N) \rangle = \sum_x g(N(x))\, P(x). \tag{4.25}$$

And we can calculate the average of a function $g(n)$ directly from its values:

$$\overline{g(N)} = \frac{1}{M} \sum_{i=1}^{M} g(n_i)\,. \tag{4.26}$$

4.3.5 Example: Expected value of a dice throw

Analytical solution of expected value. Let us use this to find the expected value for the throw of a dice. The result of a throw is the random variable N, and the expected value is therefore $\langle N \rangle = E(N)$. We find $\langle N \rangle$ from the probability $P(n)$ to observe the value n:

$$\begin{aligned} \langle N \rangle = E(N) &= \sum_{n=1}^{6} P(n)\, n = \sum_{n=1}^{6} p\, n = \sum_{n=1}^{6} \frac{1}{6}\, n \\ &= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2}\,. \end{aligned} \tag{4.27}$$

Notice that this is an exact result that characterizes the underlying process — it is not a measurement.

Numerical measurement of average value. We can calculate the average value from may dice throws from a sequence of observations, n_j, $j = 1, \ldots, M$:

$$\bar{n} = \frac{1}{M} \sum_{j=1}^{M} n_i\,, \tag{4.28}$$

where the sum is over all the measurements M. If we use python to estimate the average of the sequence of die throws found in Sect. 4.2.5 using using the formula in (4.28), we get:

```
import numpy as np
M = 10000
m = np.random.randint(0,6,M)+1
myavg = sum(m)/len(m)
print(myavg)
```

```
3.5044
```

The average is implemented as the standard function `average` in python:

```
myavg2 = np.average(m)
print(myavg2)
```

```
3.5044
```

Notice that this estimated value is close to, but not exactly equal to, the theoretical value.

4.3.6 Variance and Standard deviation

In addition to the expected value, we may want to describe the variation around the expected value. Does the outcomes typically occur very close to or far away from the expected value. We could characterize this by the expected value of the deviation from the expected value of the random variable N:

$$E\left(N - E(N)\right) = \sum_x P(x)\left(x - E(N)\right) \tag{4.29}$$

$$= \underbrace{\sum_x P(x)\, x}_{=E(N)} - \sum_x P(x)E(N) \tag{4.30}$$

$$= E(N) - E(N)\underbrace{\sum_x P(x)}_{=1} \tag{4.31}$$

$$= E(N) - E(N) = 0\ . \tag{4.32}$$

Unfortunately this did not work, since the result is always zero. Instead, we may use the absolute value of the deviation:

$$E\left(|N - E(N)|\right) = \sum_x P(x)\,|x - E(N)|\ . \tag{4.33}$$

This does not become zero. However, it is more common to characterize the deviation by the expected value of the square of the deviation:

$$E\left((N - E(N))^2\right) \ . \tag{4.34}$$

We can simplify the expression by expanding the paranthesis:

$$\begin{aligned} E\left((N - E(N))^2\right) &= E\left(N^2 - 2NE(N) + (E(N))^2\right) \\ &= E(N^2) - 2E(N)E(N) + (E(N))^2 \\ &= E(N^2) - (E(N))^2 \ . \end{aligned} \tag{4.35}$$

The average of N^2 is called the *variance* of the random variable N:

The **variance** of a random variable N is defined as the expected value of N^2:

$$\mathrm{Var}(N) = E(N^2) \ . \tag{4.36}$$

Here, we will use the term *deviation* for the deviation from the expected value:

$$E\left((N - E(N))^2\right) = \sigma^2 = E(N^2) - (E(N))^2 \ . \tag{4.37}$$

For a random variable with zero expected value, the deviation and the variance is the same.

4.3.7 Standard deviation

We *estimate* the deviation, which can be used to find the variance, using the standard deviation:

Standard deviation: We measure the standard deviation, $\bar{\sigma}$, of a sequence of M outcomes n_i using:

$$\bar{\sigma} = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}(n_i - \bar{n})^2} \ . \tag{4.38}$$

where $\bar{n}$ is the average, which also is estimated from the data.

(There is a subtle statistical reason that we divide by $M-1$ and not by M in this sum if we use the estimated value $\bar{n}$ instead of the exact value $\langle N \rangle$ in this sum.)

4.3.8 Example: Variance and standard deviation of a dice throw

Analytical solution of the variance. We found that for a single dice, the probability to get the outcome i is $P(i) = p = 1/6$ for $i = 1, \ldots, 6$. We can find the deviation for a single throw of a dice, N, through

$$\begin{aligned} &\langle N^2 \rangle - (\langle N \rangle)^2 \\ &\qquad = \sum_{i=1}^{6} P(i) i^2 - \left(\sum_{i=1}^{6} P(i) i \right)^2 \\ &\qquad = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(\frac{7}{2} \right)^2 \\ &\qquad = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}\,, \end{aligned} \tag{4.39}$$

Numerical estimation of the standard deviation. This is again an exact result. We can *estimate* the standard deviation directly from the 10000 dice throws directly using the formula in (4.35):

```
myavg = np.average(m)
mystd = np.sqrt(sum(m*m)/(len(m)-1.0)-myavg*myavg)
print(mystd)
```

```
1.7072142923487961
```

Or you can use the built-in function `std` directly

```
std(m)
```

```
1.7072142923487961
```

4.4 Toward a theory for $n(t)$ for the MD model

Now, let us use what we have learned so far to develop a theory that can describe the probability $P(n)$ to observe a particular value n in the molecular dynamics simulation of the gas/liquid system.

4.4.1 One atom model

Let us start by simplifying the system. First to one atom. In that case we guess that the atom has equal probability to be in the left or in the right hand half, since the system is symmetric and we have no reason to prefer the left to the right side. Hence we would expect there only to be two possible outcomes for n, $n = 0$ and $n = 1$ with equal probability. This is just like flipping a coin (throwing a 2-sided dice). From our discussions above, we know that the probability for a fair coin to show the number n corresponds to the probability of a fair, two-sided die to show the number n:

$$P(n) = 1/2, n = 0, 1 \ , \tag{4.40}$$

where we have already seen that this theory fits well with a numerical simulation. However, this model is too simple to predict $n(t)$ for a many-particle gas. Let us instead address a two-atom system.

4.4.2 Two-atom model

We make the model slightly more advanced by looking at two *independent* atoms, 1 and 2. We will assume that each atom is described by the random variables n_1 and n_2, which is 1 if the atom is on the left side and 0 if the atom is on the right side. We assume that these two values are independent and that their probabibilities for each of the atoms are the same as for a single atom in (4.40).

Using this notation, the total number of atoms on the left side is $n = n_1 + n_2$. How can we find the probability for the possible values of n? We address this by a numerical experiment and by a theoretical argument:

Numerical approach. We study the distribution of n by performing a numerical statistical experiment. We generate two sets of random numbers, $n_1(i)$ and $n_2(i)$, for $i = 1, \ldots, M$, and add them to get n. We can then find the probabilities for various values of n using the same method as we did above. First, we generate 2 arrays n_1 and n_2 each with M elements. We do this by generating a variable with $M \times 2$ random numbers so that the first column equals n_1 and the second column equals n_2:

```
import numpy as np
M = 1000
m = np.random.randint(0,2,(M,2))
print(m[0:5,:])
```

```
[[0 1]
 [1 0]
 [0 0]
 [1 1]
 [1 0]]
```

Where we have shown the first 5 elements for n_1 and n_2. We find n by adding the columns corresponding to n_1 and n_2. That is, we want to generate a new array with M elements, where each element corresponds to the sum along each row in the `m`-matrix. This is done by the `sum` command, but we must specify that the sum should be done along the second dimension of the array, so that we sum along the rows:

```
n = npsum(m,axis=1)
print(n[0:5])
```

```
[1 1 0 2 1]
```

(Notice that we have used the `numpy` version of `sum` in order to sum along the rows). We can now find and plot the histogram:

```
M = 10000
N = 2
m = np.random.randint(0,2,(M,N))
n = np.sum(m,axis=1)
import matplotlib.pyplot as plt
plt.hist(n,bins=3)
```

The resulting plot is shown Fig. 4.4.

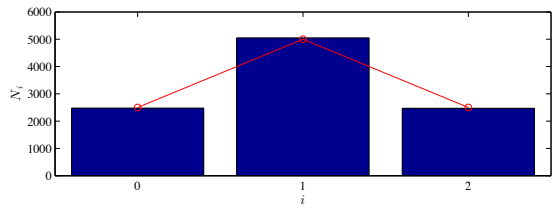


Fig. 4.4 Plot of the histogram for 10000 realizations of the sum of two coin tosses: $n = n_1 + n_2$, where $n_1 = 0, 1$ and $n_2 = 0, 1$.

Theoretical approach. We can find the probability $P(n)$ from a theoretical argument: There are four possible outcomes of (n_1, n_2), but only three possible outcomes of $n = n_1 + n_2$, as illustrated in the following table:

n_1	n_2	$n = n_1 + n_2$
0	0	0
1	0	1
0	1	1
1	1	2

If all the 4 outcomes of (n_1, n_2) are equally likely, with probability $p = 1/4$, then the probability for an outcome n is given as the number, $g(n)$ of (n_1, n_2)-states that give the same n-value, multiplied with the probability, p, per state, where $p = 1/M$, and $M = 2^2$ is the total number of (n_1, n_2)-states.

$$P(n) = g(n)p = \frac{1}{2^M} g(n) \ , \tag{4.41}$$

where the **multiplicity** $g(n)$ is read from the table above, and are tabulated in the following table:

n	$g(n)$	$P(n)$
0	1	1/4
1	2	2/4
2	1	1/4

We can compare these results directly with the histogram in Fig. 4.4. The histogram records how many experiments, N_n, that resulted in outcome n, when the total number of experiments were $M = 10000$. The probability $P(n)$ is approximately equal to N_n/M, hence we must compare $N_n = P(n)\, M$ with the values in the histogram, as done with the following code:

```
nn = np.array([0,1,2])
Pn = np.array([0.25, 0.5, 0.25])
Nn = Pn*M
plt.plot(nn,Nn)
plt.hist(n,bins=3)
```

The resulting correspondence is excellent, as seen in Fig. 4.4.

4.4.3 Many-atom system

How can we generalize this theoretical approach to many atoms? For the two-atom system we counted the number of possible outcomes, that is, all the possible (n_1, n_2) values: $(0, 0)$, $(1, 0)$, $(0, 1)$, $(1, 1)$. There are four outcomes in total. We call such an outcome a **microstate**. There are four microstates. Each such microstate is equally probable — just like when we threw two dice. The possible states for n are $n = 0$, $n = 1$, and $n = 2$. We call these states **macrostates**. The macrostates do not have the same probability, because there may be a different number of microstates in each macrostate. To find the probability of a macrostate, we have to count how the number of microstates in a particular macrostate.

For two atoms there are four microstates in total, each with probability 1/4. The macrostate $n = 0$ can only result from one microstate, $n = 1$ from two, and $n = 2$ from one.

This method can easily be generalized to any number of atoms. For N atoms, each microstate can be described as a sequence of N numbers, $(n_1, n_2, \ldots, n_N)$, for example $(1, 0, 0, 1, 0, 1, 1, 0, 0, 0, \ldots, 1)$, and n will be the sum, $n = \sum_i n_i$, of the values in the sequence. To find the probability of a macrostate, we must count how many microstates results in a particular macrostate. To do this, we use a clever trick: We realize that the value of n only depends on how many of the n_i in the microstate are 1. It does not matter where they are in the sequence. We can therefore find all microstates that give a value n by finding out how many ways we can place n ones in a sequence of N numbers. (The remaining numbers in the sequence will be zeros).

We have now mapped the problem onto a well know problem in combinatorics with a simple solution[2]. We have n ones (and $N-n$ zeros). Let us decide where to place the ones. The first one can be placed in N positions in the sequence, the second in $N-1$ positions and so on until we reach $N-n+1$. This results in $N!/(N-n)!$ different ways to place the ones.

However, using this method we have counted many microstates several times, because the ones are identical and therefore the order in which they are placed into the sequence is inconsequential. We can show this in the two-atom system. With the proposed counting method, we count one state where we first place a 1 at the first position ($n_1 = 1$), and then we place a one at the second position ($n_2 = 1$). This is the state $(1, 1)$. However, we also count a state where we first place a 1 at the second position ($n_2 = 1$) and then we place a one at the first position ($n_1 = 1$), giving the state $(1, 1)$. But these two states are identical! And here we have counted these states twice. We must therefore correct for this by dividing by the number of times we have counted the same state. For a sequence with n ones, such as for this sequence with 4 ones: $(1, 0, 1, 1, 0, 1)$, we must divide by all the possible ways we could have generated this state, which corresponds to all the possible ways we can organize n numbers, which is $n!$. This means that the total number of microstates is

$$\Omega(n, N) = \frac{N!}{(N-n)!\,n!} = \binom{N}{n} , \tag{4.42}$$

[2] This is a common approach in physics that you will meet many times in your career as a physicist.

where we have introduced the multiplicity $\Omega(n, N)$ and the usual notation used for this combinatorical result.

The total number of microstates is 2^N, and all these microstates are equally likely, so that the probabilty for n is

$$P(n) = \binom{N}{n} \frac{1}{2^N} = \frac{N!}{n!(N-n)!} \frac{1}{2^N} \,. \tag{4.43}$$

4.4.4 Comparing theory and simulation

We have now developed a theoretical description of the system, and the theory provides a prediction for the distribution of n — the number of atoms in the left half of the system — given the assumption that the position of each atom is independent of all the position of all the other atoms.

Reading data from simulations. Let us compare the theory directly with the measurements. First, we compare the simulated result from 5001 i-values, n_i, (from a simulation of 50000 time-steps in in.gasstatistics01[3]) and $N = 400$ atoms. The resulting trajectory gasstat01.lammpstrj[4] is analyzed by the script gasplotnt01.py[5]. You can find the data in the file ndata01.d[6], which was generated by the script. We read this data, extract the n-values, and plot $n(t)$

```
import numpy as np
import matplotlib.pyplot as plt
t,n = np.loadtxt('ndata01.d');
plt.plot(t,n)
```

The resulting plot of $n(t)$ is shown in the top of Fig. 4.5.

Estimating the probability distribution $P(n)$. We estimate the probability distribution $P(n)$ for the n using the histogram function: We count the number of times, N_n, that $n(t) = n$, and then divide by the total number of observations, M (called `nt` in the code):

$$P(n) \simeq \frac{N_n}{M} \,. \tag{4.44}$$

The results depend on the number of bins used for the histogram. We can display the histogram with 10 and 20 bins using

[3] http://folk.uio.no/malthe/fys2160/in.gasstatistics01
[4] http://folk.uio.no/malthe/fys2160/gasstat01.lammpstrj
[5] http://folk.uio.no/malthe/fys2160/gasplotnt01.py
[6] http://folk.uio.no/malthe/fys2160/ndata01.d

```
plt.hist(n,bins=10)
plt.hist(n,bins=20)
```

The resulting histograms are shown in Fig. 4.5. Notice that the number of counts in each bin goes down when the number of bins is increased. These histograms only measure the number of outcomes in each bin, where the bin may span over a size Δn corresponding to many n-values. However, we would like to estimate $P(n)$ as N_n/M, where N_n is the number of outcomes that are equal to n. This can be done either by ensuring that the bin width is n, or we need to divide by the bin size in order to find the estimate for the probability $P(n)$ across the bin: $P(n) \simeq N(n, n+\Delta n)/(M\,\Delta n)$. Why do we need to divide by Δn? If we count $N(n, n+\Delta n)$, we have really counted

$$N(n, n+\Delta n) = N(n) + N(n+1) + N(n+2) + \ldots + N(n+\Delta n)\ . \quad (4.45)$$

If we now divide by M, the total number of outcomes, we get

$$\frac{N(n, n+\Delta n)}{M} = \frac{N(n)}{M} + \frac{N(n+1)}{M} + \frac{N(n+2)}{M} + \ldots + \frac{N(n+\Delta n)}{M}$$

$$(4.46)$$

$$\simeq P(n) + P(n+1) + P(n+2) + \ldots + P(n+\Delta n) \quad (4.47)$$

$$= \Delta n\,\frac{1}{\Delta n}\,(P(n) + P(n+1) + P(n+2) + \ldots + P(n+\Delta n)) \quad (4.48)$$

$$= \Delta n\,\overline{P(n+\Delta n/2)}\ . \quad (4.49)$$

We therefore need to divide by Δn to find the average probability over the interval from n to $n+\Delta n$, and we need to find the centers of the bins, $n+\Delta n/2$, in order to plot the data in this point. We can calculate the probabilities and the bin centers directly from the output from the `histogram`-function:

```
Nn,edges = np.histogram(n,bins=10)
ncenter = 0.5*(edges[1:]+edges[:-1])
dn = edges[1:]-edges[:-1]
nt = len(n)
plt.plot(ncenter,Nn/dn/nt)
```

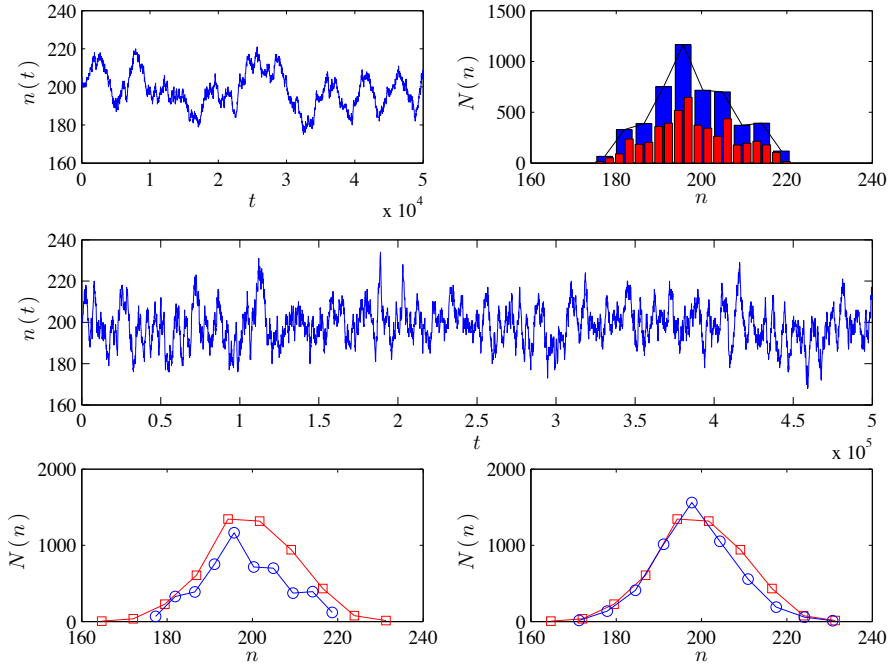


Fig. 4.5 Plot of $n(t)$ from the simulation and n_i from 5001 random microstates from 50,000 timesteps (top left) and 500,000 timesteps (middle). (Top right) Histogram of $n(t)$ for 50,000 timesteps with 10 and 20 bins. (Bottom left) Histogram of $n(t)$ for 50,000 timesteps (blue) and for a simulation with 5001 random outcomes using the binomial model (red). (Bottom right) Histogram of $n(t)$ for 500,000 timesteps (blue) and for a simulation with 5001 random outcomes using the binomial model (red).

Comparison with theory — numerical estimates. We now have the tools needed to compare the observed results from the molecular dynamics simulations with our simple theory for $n(t)$. The number of observations from the molecular dynamics simulations is 5001. To compare with the observed results, we generate set of possible values of n for $N = 400$ atoms using the theory we have developed so far: Each of the N atoms can either be on the left side ($X_i = 1$) or on the right side ($X_i = 0$) with equal probability, and we count the total number of atoms, n, on the left side

$$n = \sum_{i=1}^{N} X_i \ , \tag{4.50}$$

We generate 5001 such values randomly and compare the two histograms:

```
ni = randint(0,2,(5001,400))
nt = sum(ni,axis=1)
```

```
hist(nt), hist(n), show()
```

The resulting plot is shown in Fig. 4.5. Hmmm. This did not look like a very good fit. What has gone wrong? It may be that the initial correlations in the data were too strong, so that not all the atoms really are active and can change places between each measurement from the simulation. Let us increase the time interval between measurements of n_i, but keep the number of observations the same. We rerun with the input file in.gasstatistics02[7] resulting in gasstat02.lammpstrj[8], producing the data file ndata02.d[9], which we plot using

```
t2,n2 = np.loadtxt('ndata02.d')
plt.hist(nt), plt.hist(n2)
```

The resulting sequence $n(t)$ of n-values is shown in Fig. 4.5. We see that the correlations in the $n(t)$ signal are now less prominent, although there are still some correlations present. However, the histograms in the bottom right part of Fig. 4.5 now show much better correspondence, indicating that the theory we have developed provides a good explanation of the behavior. Indeed, the correspondence between observations and theory is surprisingly good given the simplicity of the model. We may therefore conclude that the model captures the most important features of the behavior of $n(t)$.

Comparison with theory — calculated probabilities. Now, let us also compare directly with the probabilities we have calculated for n:

$$P(n) = \frac{N!}{n!(N-n)!} 2^{-N} \,, \tag{4.51}$$

We use the `scipy.stats.binom.pmf` function to calculate $P(n)$ as a function of n. (This uses the scipy[10] package). In order to compare directly with the histogram, we notice that $P(n) = N_n/M$, and therefore $N_n = M\,P(n)$. We also choose many bins (301) to ensure that each value of n falls into one and only one bin, so that $P(n) = N_n/M$, where N_n is the number of outcomes in the bin that includes n. This is done by the following script:

```
import scipy, scipy.stats
import numpy as np
import matplotlib.pyplot as plt
```

[7] http://folk.uio.no/malthe/fys2160/in.gasstatistics02
[8] http://folk.uio.no/malthe/fys2160/gasstat02.lammpstrj
[9] http://folk.uio.no/malthe/fys2160/ndata02.d
[10] http://www.scipy.org

```
M = 5001
Nn,edges = np.histogram(n,bins=10)
ncenter = 0.5*(edges[1:]+edges[:-1])
dn = edges[1:]-edges[:-1]
xn = scipy.linspace(0,400,401)
pmf = scipy.stats.binom.pmf(xn,400,0.5)
plt.plot(ncenter,Nn/dn/M,xn,pmf)
```

The resulting plot is shown in Fig. 4.6. The correspondence is not spectacular, but decent, showing that the theory we have developed gives a reasonably good description of $n(t)$ from the molecular dynamics simulations.

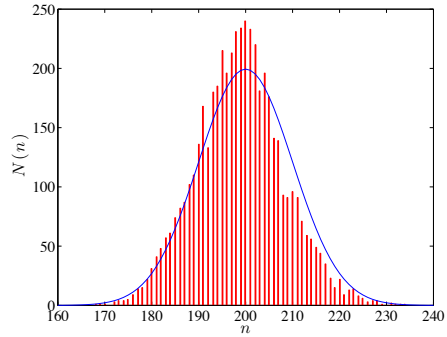


Fig. 4.6 Plot of the histogram for $n(t)$ from the simulation and n_i from a binomial distribution.

4.5 The binomial distriubtion

The binomial distribution fits well to the observed data for the molecular simulation. Indeed, the binomial distribution provides a good description for any sum of independent events, such as the sum of several dice, the results of many flips of a coin, or the behavior of the number of atoms in a part of a gas.

In general, the binomial distribution describes the case where we have a sequence of N *identical*, *independent* trials with discrete outcomes X_i, $i = 1, \ldots, N$ – often called a *Bernoulli process*. The simplest case is when X_i is binary, such as either 0 or 1. We have looked at a fair coin, but even when the coin is not fair, such as if the case if we have probability p for $X_i = 1$ and $q = 1 - p$ for $X_i = 0$, the results are described by the binomial distribution.

Binomial distribution: For N independent binary random variables, X_i, the sum $Z = \sum_{i=1}^{N} X_i$, is described by a binomial distribution, $P(z)$:

$$P(Z = z) = \binom{N}{z} p^z q^{N-z} , \tag{4.52}$$

This formula reproduces our results from above for a fair dice with $p = q = 1 - p = 1/2$.

4.5.1 Example: Coin toss

We can use the binomial distribution to answer questions about coin tosses. Let us assume we flip a fair coin, so that $p = q = 1/2$.

(a). If we flip the coin N times, what is the probability that all the flips are ones?

Getting all the flips to be ones corresponds to $n = N$. We can use the formula to find the probability for this:

$$P(N, N) = \frac{N!}{N!(N-N)!} 2^{-N} = 2^{-N} . \tag{4.53}$$

(b). If we flip the coin $N = 5$ times, what is the probability to get exactly 3 heads? Again, we can apply the formula directly, now with $n = 3$ heads and $N = 5$ flips:

$$P(N = 5, n = 3) = \frac{5!}{3!(5-3)!} 2^{-5} = \frac{5}{16} . \tag{4.54}$$

4.5.2 Properties of the binomial distribution

We can gain intuition about the binomial distribution by studying numerical experiments of $N = 10$, 20, 40, and 80 trials when X_i is 0 or 1 with equal probability, and then measuring $Z = \sum_i X_i$. This is done by generating a set of M random sequences of N steps:

$$X_{i,j} \begin{cases} 0 & (1-p) \\ 1 & p \end{cases} \quad i = 1, \ldots, N \ , \ j = 1, \ldots, M \ , \tag{4.55}$$

and

$$z_j = \sum_{i=1}^{N} X_{i,j} \ , \ j = 1, \ldots, M \ , \tag{4.56}$$

This gives M outcomes z_j, and we can then collect statistics about these outcomes. We generate M outcomes z_j of length N in the array `z` using

```
import numpy as np
xi = np.random.randint(0,2,(M,N))
z = np.sum(xi,axis=1)
```

Again, we use histograms to estimate $P(z) \simeq N_z/M$, where N_z is how many times we observed z in M outcomes:

```
import numpy as np
Nvalues = np.array([10,20,40,80]) # Number of flips
M = 100000 # Number of samples
for ival in range(0,len(Nvalues)):
    print(ival)
    N = Nvalues[ival]
    ni = np.random.randint(0,2,(M,N))
    nt = np.sum(ni,axis=1)
    y,binEdges = np.histogram(nt,bins=100)
    bincenters = 0.5*(binEdges[1:]+binEdges[:-1])
    j = np.where(y>0) # find(y>0)
    plt.plot(bincenters[j],y[j]/(1.0*M),'-o',label='N ='+str(N))
plt.xlabel('n')
plt.ylabel('P(n)')
plt.legend()
```

We have used 100 bins. This is more than the number of possible values, which span from 0 to 80 for the case where $N = 80$ and over a smaller range when N is smaller. This was done to ensure that each bin only contains one possible outcome. Now, if we use 100 bins for the case when $N = 20$, there are only 21 possible outcomes, and many of the bins will be empty. We therefore only plot the values in the bins that contain at least one outcome. This is done by the `find` function, which returns an array with indicies for all the element where the histogram is non-zero. The resulting plots of the estimated probabilities for $N = 10, 20, 40, 80$ are shown in Fig. 4.7.

4.5.3 Expected value and variance of the binomial distribution

From the plots we see that both the average value and the width of $P(z; N)$ increases with N. Let us see how this compares with the theoretical values for the expected value and the variance of the binomial distribution. We

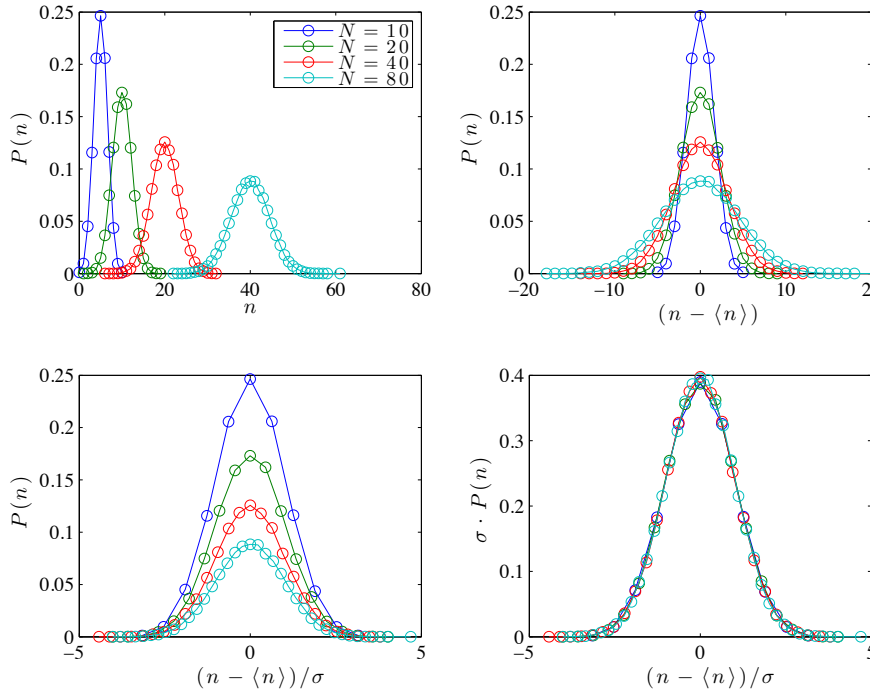


Fig. 4.7 Plot of the $P(n)$ as a function of n for $N = 10, 20, 40$, and 80.

can calculate the theoretical expected value of the binomial probability distribution using the definition of the expected value:

$$\begin{aligned}\langle n\rangle &= \sum_{n=0}^{N} nP(N,n) = \sum_{n=0}^{N} \binom{N}{n} np^n q^{N-n} \\ &= p\frac{d}{dp}\sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} \\ &= p\frac{d}{dp}\left(p+q\right)^N = pN\left(p+q\right)^{N-1} = Np\ ,\end{aligned} \tag{4.57}$$

Similarly, you can use the same approach to show that the deviation is:

$$\left\langle (n-\langle n\rangle)^2\right\rangle = Npq\ , \tag{4.58}$$

We test these theoretical predictions directly on the simulated results. First, let us measure the deviation from the average, $n - \langle n\rangle = n - Np$, and find the histogram as a function of the average. As we see in Fig. 4.7

the distribution now appears symmetric around $n - Np = 0$, which validates the theory for the average.

What about the width of the distribution? The theoretical prediction is that the width is $\sigma = \sqrt{Npq}$. We can test this by rescaling the n-scale with σ by plotting $P(n)$ as a function of $(n - \langle n \rangle)/\sigma$ where $\sigma = \sqrt{Npq}$. The resulting plot is shown in Fig. 4.7. The widths in these figures are correct, but the heights are now different. Experimentation shows that we can rescale the height also by multiplying with σ as shown in Fig. 4.7.

Can we understand how we arrived at this particular way of rescaling the distributions? Yes – this we can understand by addressing the continuum limit of the distribution.

4.5.4 Continuum limit of the binomial distribution

The binomial distribution is valid for discrete values z. However, as N becomes large, the range of possible outcomes for the sum becomes wide and we may ask what the limit of the distribution is for large N?

In this case, we need an approximate expression for $N!$ when N is large. This is given by Stirling's approximation, which is

$$\ln N! \simeq \ln \sqrt{2\pi N} + N \left(\ln N - 1 \right) \ , \tag{4.59}$$

The second term contains the factor N, which typically will be much larger than the $\ln N$ term in the first factor. We will therefore often only include the second $N \ln N - N$ term, but here we will also include the smaller first factor.

Here, we will prove that when N becomes large, the binomial distribution approaches that of a Gaussian or a Normal distribution

$$P(N, n) = \frac{N!}{n!(N-n)!} 2^{-N} \simeq \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{((N/2)-n)^2}{2(N/4)}} \ . \tag{4.60}$$

We will provide you with two ways to approach this derivation, using purely analytical techniques and using a combination of analytical and symbolic methods.

4.5.5 Continuum limit — Analytical derivation

We will study the binomial distribution when n is close to its average value $\langle n \rangle = Np = N/2$. To simplify the notation we will here use $\langle n \rangle = \bar{n}$.

We introduce the relative deviation from the average h, defined as

$$h = \frac{n - \bar{n}}{\bar{n}} \ , \tag{4.61}$$

where we will look at the case when n is close to its average so that h is small. We solve for n:

$$\bar{n}h = n - \bar{n} \ \Rightarrow \ n = \bar{n} + \bar{n}h = \bar{n}\,(1 + h) \ . \tag{4.62}$$

Similarly, we can express $N - n$ in terms of $\bar{n}$ and h:

$$N - n = N - \bar{n}\,(1 + h) = 2\bar{n} - \bar{n} - \bar{n}h = \bar{n}\,(1 - h) \ . \tag{4.63}$$

We can use these terms to simplify the binomial distribution

$$P(N, n) = \frac{N!}{n!(N - n)!} 2^{-N} = \frac{N!}{(\bar{n}\,[1 + h)]!\,[\bar{n}\,(1 - h)]!} 2^{-N} \ . \tag{4.64}$$

Let us now look at $\ln P(N, n)$ and gradually introduce Stirling's formula for the faculty functions

$$\begin{aligned}
\ln P(N, n) &= \ln\,(2\pi N)^{1/2} + N \ln N - N && (4.65)\\
&\quad - \ln\,(2\pi n)^{1/2} - n \ln n + n && (4.66)\\
&\quad - \ln\,(2\pi (N - n))^{1/2} - (N - n) \ln(N - n) + (N - n) && (4.67)\\
&\quad - N \ln 2 \ , && (4.68)
\end{aligned}$$

where we notice that the terms

$$-\,N + n + (N - n) = 0 \ , \tag{4.69}$$

cancel. As you do more of these approximations by youself, you will notice that this is always the case when you apply Stirling's formula to the binomial formula.

Let us look at the terms that include the 2π factor, and introduce $n = \bar{n}(1 + h)$ and $(N - n) = \bar{n}(1 - h)$. Notice that we have still not made any other approximations than to apply Stirling's formula.

$$\ln(2\pi N)^{1/2} - \ln(2\pi n)^{1/2} - \ln(2\pi(N-n))^{1/2} \tag{4.70}$$
$$= \frac{1}{2}\ln(2\pi N) - \frac{1}{2}\ln(2\pi\bar{n}(1+h)) - \frac{1}{2}\ln(2\pi\bar{n}(1-h)) \tag{4.71}$$
$$= \frac{1}{2}\ln\frac{2\pi N}{2\pi\bar{n}(1+h)\,2\pi\bar{n}(1-h)} \tag{4.72}$$
$$= \frac{1}{2}\ln\frac{N}{2\pi\bar{n}^2(1-h^2)}\,. \tag{4.73}$$

Now we assume that $h \ll 1$ and we insert that $\bar{n} = Np = N/2$, getting

$$\frac{1}{2}\ln\frac{N}{2\pi\bar{n}^2(1-h^2)} \simeq \frac{1}{2}\ln\frac{N}{2\pi\bar{n}^2} = \frac{1}{2}\ln\frac{N}{2\pi(N/2)^2}\,, \tag{4.74}$$

where we will not insert $\sigma^2 = Npq = N/4$:

$$\frac{1}{2}\ln\frac{1}{2\pi(N/4)} = \ln\frac{1}{\sqrt{2\pi\sigma^2}}\,. \tag{4.75}$$

We have then been able to simplify $\ln P(N,n)$ significantly:

$$\ln P(N,n) = \ln\left(2\pi\sigma^2\right)^{\frac{1}{2}} + N\ln N - n\ln n - (N-n)\ln(N-n) - N\ln 2\,. \tag{4.76}$$

Now, let us simplify further by introducing $n = \bar{n}(1-h)$ and $(N-n) = \bar{n}(1+h)$. Let us first simplify the $-n\ln N - (N-n)\ln(N-n)$ terms:

$$-n\ln n - (N-n)\ln(N-n) \tag{4.77}$$
$$= -\bar{n}(1+h)\ln\bar{n}(1+h) - \bar{n}(1-h)\ln\bar{n}(1-h) \tag{4.78}$$
$$= -\bar{n}(1+h)\left(\ln\bar{n} + \ln(1+h)\right) - \bar{n}(1-h)\left(\ln\bar{n} + \ln(1-h)\right) \tag{4.79}$$
$$= -2\bar{n}\ln\bar{n} - \bar{n}(1+h)\ln(1+h) - \bar{n}(1-h)\ln(1-h) \tag{4.80}$$
$$= -2\frac{N}{2}\ln\frac{N}{2} - \bar{n}(1+h)\ln(1+h) - \bar{n}(1-h)\ln(1-h) \tag{4.81}$$
$$= -N\ln N + N\ln 2 - \bar{n}(1+h)\ln(1+h) - \bar{n}(1-h)\ln(1-h)\,. \tag{4.82}$$

Here, we see that the terms $-N\ln N + N\ln 2$ cancel the terms $N\ln N - N\ln 2$ in $\ln P(N,n)$. We are left with the last two terms. Now, we assume that $h \ll 1$, and we can therefore use that $\ln(1+h) \simeq h$ and $\ln(1-h) \simeq -h$, getting:

$$- \bar{n}(1+h)\ln(1+h) - \bar{n}(1-h)\ln(1-h) \tag{4.83}$$
$$\simeq -\bar{n}(1+h)(h) - \bar{n}(1-h)(-h) \tag{4.84}$$
$$= -2\bar{n}h^2 \, . \tag{4.85}$$

We can now put all this together into $\ln P(N,n)$, getting

$$\ln P(N,n) = \ln \frac{1}{\sqrt{2\pi\sigma^2}} - 2\bar{n}h^2 \, , \tag{4.86}$$

where we insert $h = (n-\bar{n})/\bar{n}$ in the last term:

$$-2\bar{n}h^2 = -2\bar{n}\left(\frac{n-\bar{n}}{\bar{n}}\right)^2 = -2\frac{(n-\bar{n})^2}{\bar{n}} \, . \tag{4.87}$$

We insert for $\bar{n} = Np = (N/2) = 2(N/4) = 2\sigma^2$:

$$\ln P(N,n) = \ln \frac{1}{\sqrt{2\pi\sigma^2}} - 2\frac{(n-\bar{n})^2}{2\sigma^2} \, , \tag{4.88}$$

If we now insert $N/4 = Npq = \sigma^2$, we see that the result of all of these calculations are that:

$$P(N,n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{s^2}{2(N/4)}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{s^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{s}{\sigma}\right)^2} \, , \tag{4.89}$$

where

$$s = \frac{N}{2} - n = \bar{n} - n \, . \tag{4.90}$$

We can therefore rewrite the result in an even more general form:

$$\boxed{P(N,n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{n-\bar{n}}{\sigma}\right)^2} \, .} \tag{4.91}$$

We now use this result to understand the scaling we found to work in fig. 4.7. We see that if we rewrite the equation as

$$P(N,n)\sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{n-\bar{n}}{\sigma}\right)^2} \, , \tag{4.92}$$

we would expect all the data to fall onto a common, universal curve – given as the Gaussian distribution – as we indeed observe in Fig. 4.7.

4.5.6 Continuum limit — Symbolic derivation

Also when we use symbolic tools, we need both to have a good mathematical intuition, and we often also need to rewrite and simplify expressions ourselves. In particular, we often need to interpret limits, such as the limit of large N, ourselves using our mathematical experience and intuition.

First, we realize that the general expression for $P(N, n)$ is difficult to work with. It is simpler to work with some variable s, which is close to zero, instead of the variable n which is close to $N/2$, since we know from experience that this makes the algebra simpler when we derive a Taylor expansion around $s = 0$. We therefore introduce $s = (N/2) - n$, and rewrite the binomial probability as

$$P(N, n) = \frac{N!}{(N/2 - u)!\,(N/2 + u)!} 2^{-N} \,. \tag{4.93}$$

Also, we know from experience that with all these factors, it may be simpler to work with the logarithm of the probability, $\ln P(N, n)$. We introduce these quantities symbolically

```
import sympy as sy
u = sy.symbols('u')
N = sy.symbols('N')
P = sy.factorial(N)/(sy.factorial((N/2)-u)*sy.factorial((N/2)+u))*2**N
lnP = sy.log(P)
```

We are then ready to start symbolic manipulation. First, we would like to Taylor expand `lnP` around $u = 0$. We start from the lowest orders (second order):

```
Taylor2_lnP = sy.series(lnP,u,n=2)
print(sy.simplify(Taylor2_lnP))
```

```
-log(pi)/2 + log(gamma(N/2 + 1/2)/gamma(N/2 + 1)) + O(u**2)
```

Ooops. There is no u-dependence! We know from experience that this typically means that the first order terms cancel. We need to include higher order terms in the Taylor expansion:

```
Taylor3_lnP = sy.series(lnP,u,n=3)
print(Taylor3_lnP)
```

```
log(2**N*factorial(N)/(factorial(N/2))**2) + ...
u**2*(-gamma(N/2 + 1)*polygamma(0, N/2 + 1)**2/factorial(N/2) - ...
gamma(N/2 + 1)*polygamma(1, N/2 + 1)/factorial(N/2) + ...
gamma(N/2 + 1)**2*polygamma(0, N/2 + 1)**2/(factorial(N/2))**2) + ...
O(u**3)
```

A bit messy. Let us simplify:

```
sy.simplify(Taylor3_lnP)
```

```
log(2**N*factorial(N)/(factorial(N/2))**2) - ...
u**2*polygamma(1, N/2 + 1) + O(u**3)
```

That is more like it! Now, we use our mathematical insight to simplify further. We realize that N is a large number and that $N/2+1 \simeq N/2$. The `polygamma` is the derivative of the digamma function. We can look up the digamma-function, $\Psi(x)$, finding that it can be approximated as

$$\Psi(x) \simeq \ln x - (1/2x) - (1/12x^2) + \dots \tag{4.94}$$

In the limit of large x we can use $\Psi(x) \simeq \ln x$. The notation `polygamma(1,N/2+1)` means the first (hence the number 1) derivative of the digamma-function. The `polygamma(1,N/2+1)`-term is therefore approximately $d \ln x/dx = 1/x$. We insert $x = N/2$, getting

$$\ln P = \ln(1/2)^N + \ln \frac{N!}{(N/2)!\,(N/2)!} - \frac{u^2}{(N/2)} \; . \tag{4.95}$$

This expression already shows us how P depends on u:

$$P(N,u) = C(N)\, e^{-\frac{u^2}{(N/2)}} \; . \tag{4.96}$$

Where the prefactor $C(N)$ is

$$\ln C(N) = \ln(1/2)^N + \ln \frac{N!}{(N/2)!\,(N/2)!} = -N \ln 2 + \ln N! - 2 \ln(N/2)! \; . \tag{4.97}$$

In order to sort out the prefactor, $C(N)$, we need to use Stirling's approximation $\ln x! \simeq \ln \sqrt{2\pi x} + x \ln x - x$, which is valid for large x. We apply Stirling's approximation to the prefactor:

$$\begin{aligned} \ln C(N) &= -N \ln 2 + \ln N! - 2 \ln(N/2)! \\ &\simeq -N \ln 2 + \ln \sqrt{2\pi N} + N \ln N - N \\ &\quad - 2 \left(\ln \sqrt{2\pi (N/2)} + (N/2) \ln(N/2) - (N/2) \right) \; . \end{aligned} \tag{4.98}$$

This is again well suited for symbolic simplification.

```
lnC= -N*log(sym(2))+log(sqrt(2*pi*N))+N*log(N)-N- ...
      2*(log(sqrt(2*pi*(N/2)))+(N/2)*log(N/2)-(N/2))
simplify(lnC)
```

```
-log(sqrt(N)) - log(pi)/2 + log(2)/2
```

This is as good as it gets. We see that $\ln C = -(1/2)\ln(\pi N/2) = -\ln\sqrt{2\pi N/4}$. We insert this prefactor into $P(N,u)$ in (4.96), getting

$$P(N,u) = \frac{1}{\sqrt{2\pi N/4}} e^{\frac{-u^2}{2(N/4)}} \ . \tag{4.99}$$

This result can be generalized by realizing that $\bar{n} = N/2$ and $\sigma^2 = N/4$. We can therefore write u as $u = \bar{n} - n$, and $P(N,n)$ as

$$P(N,n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\bar{n}-n}{\sigma}\right)^2} \ . \tag{4.100}$$

This is exactly the same result as we found above, but now derived with a combination of analytical and symbolic tools.

4.5.7 Probabilities and probability densities

There is a small, but very important, distinction between the $P(N,n)$ we introduced for the bionomial distribution, and the $P(N,n)$ we have introduced now, because the new version is defined for all possible values of n, not only for discrete ones. This means that the probability to observe a specific value for n is really zero. We can only specify the probability to observe n in some interval from n to $n+dn$:

$$P(n \text{ is between } n \text{ and } n+dn) = P(N,n)dn \ , \tag{4.101}$$

We now realize that our notation is confusing, and we should clean it up. Let us introduce the notation that the **random variable** Z is the result of a binomial experiment. The probability for the observed value of Z, z, to be in the range from n to $n+dn$ is then written as

$$P(n < Z < n+dn) = f_Z(n)dn \ , \tag{4.102}$$

where we call $f_Z(z)$ the **probability density** for the random variable Z.

Notice that the probability density $f_Z(z)$ is normalized not by a sum, but by an integral:

$$\int_{-\infty}^{\infty} f_Z(z)dz = 1 \ . \tag{4.103}$$

For the case of the Gaussian approximation of the binomial distribution, we found that the probability density was

$$f_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\bar{n}-n}{\sigma}\right)^2} . \tag{4.104}$$

The normalization integral is therefore

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\bar{n}-n}{\sigma}^2\right)} = 1 . \tag{4.105}$$

We solve this by changing the integration variable, first by introducing $u = n - \bar{n}$, which does not change the integral:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{u}{\sigma}\right)^2} = 1 . \tag{4.106}$$

Then we change to $z = u/\sigma$, and the integral becomes

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} = 1 , \tag{4.107}$$

which you can look up and check that indeed is equal to one.

4.6 Key Concepts in Combinatorics

Both in statistics and in physics we frequently need to count configurations. In statistics we need to count all the possible outcomes of an experiment and in statistical physics we need to count all the possible states a system can be in. It may seem that each problem you meet needs a special approach, but usually we can use a standard trick in problem solving — we can map a new problem we are working on onto a problem we already know the solution to. Indeed, this is one of the most common techniques in physics, and what maps mathematics and physics so powerful. There are not that many different problems out there! However, in order to use such a technique, you need to know some of the classic counting problems. Here, we will introduce you to three classic problems arising when you select random elements from a bag with k balls numbered from 1 to k.

Ordered sequence with replacement. First, we draw a sequence of numbers by drawing a random ball from the bag, noting down the number, placing the ball back into the bag, and repeating this process

N times. How many different such sequences, $(n_1, n_2, n_3, \ldots, n_N)$, are there when $n_i = 1, 2, \ldots, k$? This corresponds to asking how possible outcomes are there if we throw a dice 5 times, or how many possible outcomes are there if we flip a coin 8 times. The result is a sequence of numbers, $(1, 6, 3, 5, 1)$, or (H, T, H, H, T, T, H, T), and we ask how many such (ordered) sequences are there?

For the first number, there are k possibilities. For each such possibility, there are k possibilities for the second number, and so on, so that the total number of possibilities is $\Omega = k\,k_{\cdot} \ldots k$ (N times), or $\Omega = k^N$. There are 6^5 possible outcomes if we throw 5 dice, and there are 2^8 possible outcomes if we flip a coin.

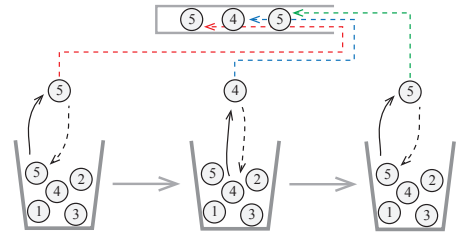


Fig. 4.8 Illustration of an ordered sequence with replacement.

Ordered sequence without replacement. What if we perform the same experiment, but we do not place a ball back into the bag after it has been selected. This means that a given number now just can occur once in the sequence. We still draw the numbers in sequence and consider the sequence $(1, 5, 3)$ different from $(5, 3, 1)$. How many such sequences are there when there are N numbers in the sequence and k different balls in the bag? This corresponds to asking how many ways can you place students at the various chairs in class or how many ways can ways can we select runners for the first, second and third place in a marathon — in all these cases the position in the sequence matters.

How many possible configurations are there of such a sequence. For the first number, there are k possible values. However, the value that occured in the first number, cannot occur as the second number. There are therefore only $k - 1$ values available for the second number, $k - 2$ for the third number, etc. The number of ways to select a sequence of N such numbers is therefore $\Omega = k\,(k-1)\,(k-2)\,\ldots\,(k-(N-1))$, which we also can write as $\Omega = k!/(k-N)!$.

Unordered set without replacement. Now, what if we do not care about the order in which we draw the elements? We still draw N balls

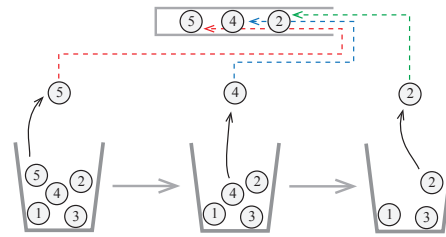


Fig. 4.9 Illustration of an ordered sequence without replacement.

from the bag, but we do not care about the order in which they were selected. This means that we only care about the set of drawn elements and not the sequence. Hence if we draw $(1, 5, 3)$ or if we draw $(5, 3, 1)$ this both corresponds to drawing the set $\{1, 3, 5\}$. How many such sets can be drawn if we draw N elements from a bag with k different elements? This corresponds to asking how many 5-cards hands can we draw from a full deck of cards or how many sets of possible lotto numbers there are?

How many possible such sets are there? We know that there are $k!/(k - N)!$ ordered sequences. But if we count all the ordered sequences, we count both $(1, 5, 3)$ and $(5, 3, 1)$ as two occurrences, even though they are part of the same set. We have therefore counted way too many sets. How many different sequences are there that results in the same set? This corresponds to the number of ways we can order the N numbers $(1, 3, 5)$, that is the number of ordered sequences of N elements, which is $N!$. We can therefore find the number of sets by dividing the number of sequences by $N!$, giving us $\Omega = k!/(N!\,(k - N)!)$.

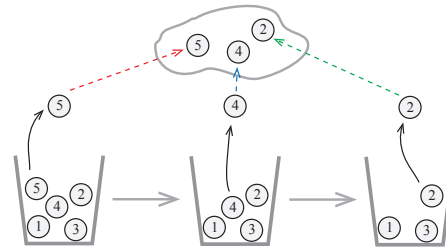


Fig. 4.10 Illustration of an unordered sequence with replacement.

4.7 Probability distributions

4.7.1 Uniform distribution

The uniform distribution describes a random process with values between a and b so that the probability $P(z, z+dz) = f_Z(z)dz$ does not depend on the value of z as long a z is between a and b.

This means that $f_Z(z) = c$ is a constant. We can find the constant from the normalization condition of the density:

$$\int_a^b f_Z(z)dz = c(b-a) = 1 \;\Rightarrow\; c = \frac{1}{b-a}\,. \tag{4.108}$$

Uniform distribution: The distribution:

$$f_Z(z) = \begin{cases} 0 & z < a \\ \frac{1}{b-a} & a \le z \le b \\ 0 & b < z \end{cases}\,, \tag{4.109}$$

is called the *uniform distribution*.

The uniform distribution is illustrated in Fig. 4.11.

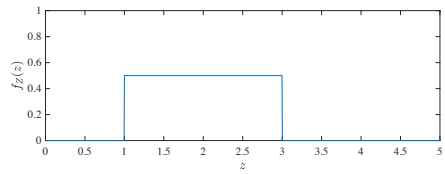


Fig. 4.11 Illustration of a uniform distribution from $a = 1$ to $b = 3$.

The expected value of a random variable Z which is uniformly distributed from a to b is:

$$\langle Z\rangle = \int_a^b z f_Z(z)dz = \int_a^b \frac{z}{b-a}dz = \frac{1}{2}\frac{(b^2-a^2)}{b-a} = \frac{1}{2}(b+a)\,. \tag{4.110}$$

4.7.2 Exponential distribution

Exponential distribution: The distribution

$$f_Z(z) = \frac{1}{c} e^{-z/c} \,, \tag{4.111}$$

for positive z is called the exponential distribution. It is commonly used to describe waiting time statistics.

Normalization. We check that the distribution is normalized:

$$\int_0^\infty \frac{1}{c} e^{-z/c} dz = \int_0^\infty e^{-u} du = 1 \,. \tag{4.112}$$

Expected value. The expected value of Z is

$$\langle Z \rangle = \int_0^\infty \frac{1}{c} z e^{-z/c} dz = c \int_0^\infty e^{-u} du = c \,. \tag{4.113}$$

Interpretation of the exponential distribution. We can understand the origin of the exponential distribution as a waiting time distribution from a simple example: A process occurs at a constant, but very low rate r, so that the probability for an event to occur in a time interval from t to $t + dt$ is a constant r multiplied with dt, rdt. What is the probability for the event to occur for the first time in the time interval from t to $t + dt$? For this to happen, the even must not occur for the time t and then occur in the time interval from t to $t + dt$. We can divide the time t into small units Δt. The probability for an event to occur in such a small time interval is $r\Delta t$, and there are $n = t/\Delta t$ such interval from 0 to t. The probability for no event to occur from 0 to t is then

$$P = (1 - r\Delta t)^n = \left(1 - \frac{r \cdot t}{n}\right)^n \to e^{-rt} \,, \tag{4.114}$$

when $\Delta t \to 0$ and $n = (t/\Delta t) \to \infty$. The probability for the event to occur for the first time in a time interval from t to $t + dt$ is P, which is the probability for no event to occur before t, multiplied by rdt, which is the probability that the event occurs in the interval from t to $t + dt$. This gives the probability for the event to occur for the first time in the inverval from t to $t + dt$ to be:

$$f_T(t)dt = e^{-rt} r dt \,. \tag{4.115}$$

We can therefore interpret $f_T(t)$ as the probability density for the waiting time T, and we may interpret the exponential distribution as a waiting time distribution with a rate r or a characteristic time $\tau = 1/r$.

Comparison with numerical simulation. We can test this theory by generating a simple data-set for waiting time statistics. Let us assume that a light bulb has a probability $p = 1/1000$ to fail during an hour. What is the distribution of the lifetime of such light bulbs?

Let us generate a numerical experiment to measure this, and then compare the results of the numerical simulation with the theory. First, how can we find the number of trials needed to observe the first event, when the probability of an even in each trial is p? We could do this using a loop, which continues until we have achieved success, and then count how many attempts were needed:

```
n = 0
while (rand(1,1)>p):
    n=n+1
```

This is not a very efficient algorithm in a vectorized language such as python, but we have chosen to implement it like this because the structure of the algorithm is very clear. We then need to repeat this experiment `nsamp` times and collect the results in a variable `tval`. This is implemented in the following program:

```
#start1
import numpy as np
import matplotlib.pyplot as plt
p = 1.0/1000.0
nsamp = 10000
tval = np.zeros(nsamp)
for i in range(nsamp):
    j = 0
    while (np.random.rand()>p):
        j = j + 1
    tval[i] = j
```

Estimating the probability density. How to estimate the probability density for T, the waiting time? We do this following the frequency definition of probability: By counting how many events, N_t, are in an interval from t to $t + dt$. The probability for T to be in the interval from t to $t + dt$ is then $f_T(t)dt = N_t/N$, where N is the total number of experiments. The number N_t are produced in a histogram, and we can then find the probability density through

$$f_T(t) = \frac{N_t}{N\,dt}\,. \tag{4.116}$$

Notice that the interval size, dt, appears in the denominator! This factor is often forgotten by students, but it is essential in order to find the probability *density*.

First, let us find the histogram using bins that are 1 second in width. We do this by choosing the number of bins used to correspond to the largest value of T observed using

```
[n,t] = np.histogram(tval,bins=int(max(tval)))
tmid = 0.5*(t[:-1]+t[1:])
```

where the variable `n` now contains the counts, N_t, and `tmid` contains the midpoints of the bins. Since the bins all have the same size, which is a feature of the `histogram` function, we can find the bin size dt from the first two values of `t`:

```
dt = t[1]-t[0]
```

We can finally estimate the probability density using (4.116):

```
Pt = n/(nsamp*dt)
```

Notice that this is a vectorized command, which calculates the values for `Pt` for all the values in `n`. Essentially, a command like this corresponds to

```
for i in range(len(n)):
    Pt[i] = n[i]/(nsamp*dt)
```

We can then plot the calculated probability density and compare with the theoretical values found above:

$$f_T(t) = pe^{-pt}\,, \tag{4.117}$$

which is done by

```
theory = p*np.exp(-p*tmid)
```

which again is a vectorized command that calculates the values for `theory` for all values in the `t`-array. The results are plotted by

```
plot(tmid,Pt,'o',tmid,theory,'-')
```

and the resulting distribution is shown to the left in Fig. 4.12.

Unfortunately, this plot is really messy! It is not easy to observe the behavior because there are too many data points and too much noise. What can we do to improve this? We can broaden the bins, so that there are more points in each bin – we need to change the bin size dt. For

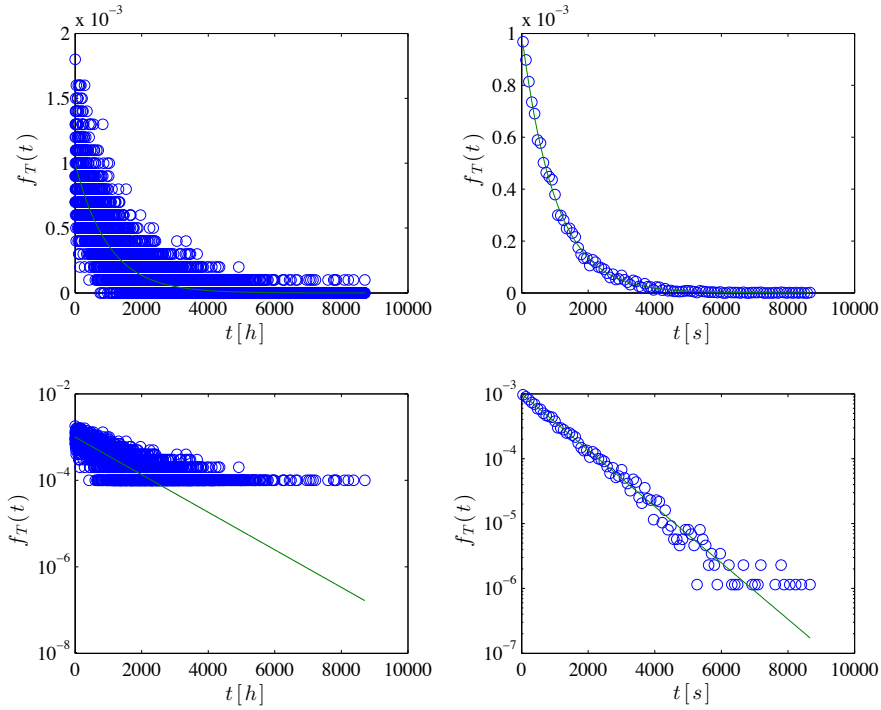


Fig. 4.12 Plot of the measured probability denstiy $f_T(t)$ for the waiting time distribution for a process occuring with a probabity $p = 1/1000$ per hour. (Left) Bin size correspoding to the smallest possible resolution in t. (Right) With 100 bins.

example, we can analyze the data using only 100 bins. However, we then need to be very careful to remember to include the bin size dt when we estimate the probability density. This was done using

```
[n,t] = np.histogram(tval,bins=100)
tmid = 0.5*(t[:-1]+t[1:])
dt = t[1]-t[0]
Pt = n/(nsamp*dt)
theory = p*np.exp(-tmid*p)
plt.plot(tmid,Pt,'o',label='Observation')
plt.plot(tmid,theory,'-',label='Theory')
plt.legend()
```

And the resulting plots are shown to the right in Fig. 4.12. The results are now much clearer. And we can now also see that the theory fits perfectly. The data is indeed well described by the exponential distribution.

4.7.3 Poisson distribution

The Poisson distribution appears when we look at the number of event to happen in a given interval for a random process that occurs at a constant rate. For example, the number of alpha particles emitted in a given time frame, if the rate of emission is constant, follows a Poisson distribution, or the number of childbirths in a given time internval.

Making a simplified model. Let us develop a theory for the number of events, n, occuring in a time interval T, when the expected number of events is λ. First, we divide the time interval into N many small intervals Δt. In each time interval, the typical number of events would be λ/N, and when N is very much larger than λ this corresponds to the probability for an event to occur, $p = \lambda/N$. As we divide the time interval into more and more subintervals, the probability for an event to occur in a subinterval, p, becomes smaller and smaller, and we can therefore assume that the probability for more than one event to occur in a subinterval is negligible. Let us use this to estimate the distribution of n, the number of event occuring during the time interval T?

Finding the distribution from the model. The number of events, n, occuring in N subintervals is the sum of the number of events, n_i, occuring in each subinterval i:

$$n = \sum_{i=1}^{N} n_i \ , \tag{4.118}$$

where the probability for n_i to be 1 is $p = \lambda/N$ and the probability for n_i to be 0 is $1 - p$. This means that n is described by a binomial distribution:

$$P(n) = \frac{N!}{n!(N-n)!)} p^n (1-p)^{N-n} \ , \tag{4.119}$$

where we can now insert $p = \lambda/N$:

$$P(n) = \frac{N!}{n!(N-n)!} \left(\frac{\lambda}{N}\right)^n \left(1 - \frac{\lambda}{N}\right)^{N-n} \tag{4.120}$$

$$= \frac{N!}{n!(N-n)!N^n} \lambda^n \left(1 - \frac{\lambda}{N}\right)^N \left(1 - \frac{\lambda}{N}\right)^n \tag{4.121}$$

$$= \frac{1}{n!} \frac{N!}{(N-n)!N^n} \lambda^n \left(1 - \frac{\lambda}{N}\right)^N \left(1 - \frac{\lambda}{N}\right)^n \ , \tag{4.122}$$

where we notice that when $N \to \infty$:

$$\frac{N!}{(N-n)!N^x} \to 1 \;\Rightarrow\; \left(1-\frac{\lambda}{N}\right)^N \to e^{-\lambda} \;\Rightarrow\; \left(1-\frac{\lambda}{N}\right)^n \to 1\,. \tag{4.123}$$

This gives

$$P(n) = \frac{\lambda^n e^{-\lambda}}{n!}\,, \tag{4.124}$$

for the probability for the number n of event in an interval with an average number of events λ, where the events are rare and independent.

Poisson distribution: The probability to observe n events in a time interval with an average number of events, λ is called the Poisson distribution:

$$P(\lambda; n) = \frac{\lambda^n e^{-\lambda}}{n!}\,, \tag{4.125}$$

Discussion of the Poisson distribution. The Poisson distribution describes processes where there are a large number of possible events (trials), but where the probability of each event is low. The distribution describes how many events occur in a given time interval. This can for example be used to describe the number of photons detected by a telescope or a camera in astronomy, the number of cars arriving at a traffic light, the decay of radiactive nuclei, or the number of mutations on a strand of DNA.

Normalization. Let us check that the Poisson distribution is normalized. We sum over all possible values of n:

$$\sum_{n=0}^{\infty} P(\lambda; n) = \sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!} = e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\,, \tag{4.126}$$

where we recognize the second term as the Taylor expansion of the exponential function, e^{λ}, which proves that the probability is normalized.

We leave it as an exercise to find the expected value and variance of the distribution.

4.7.4 Universal distributions — The Central Limit Theorem

Here, we will demonstrate that transition from the Binomial to the Gaussian distribution for large N is not special for the Binomial distribution,

but is a very general result. Indeed, the sum of any set of independent random variables (with finite variance) will always follow a Gaussian distribution. We call this distribution the Normal distribution. This is a fascinating result with wide ramifications: The behavior of the sum is independent of the behavior of the individual components. The collective behavior is general and not dependent on the details. This is a general feature in statistical systems — that behaviors are universal and independent of the details — and a central part of physics is to discover what types of universal exist and what universality class a given process belongs to.

The binomial distribution becomes Gaussian for large N. We found above that a sum Z of Bernoulli variables:

$$Z = \sum_{i=1}^{N} X_i \quad \text{where} \quad P(X_i = 1) = p \ , \ P(X_i = 0) = q \ , \ p + q = 1 \ . \tag{4.127}$$

follows the Binomial distribution:

$$P(Z = z) = P(N, z) = \binom{N}{z} p^z q^{N-z} \ , \tag{4.128}$$

when p is a constant as N increases. As N becomes large, we found that this distribution approaches the Gaussian distribution:

$$P(N, z) \simeq \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{z-\bar{n}}{\sigma}\right)^2} \ , \tag{4.129}$$

where $\bar{n} = N/2$ and $\sigma^2 = \sqrt{N/4}$ is the expected value and variance of Z. (Notice that when p goes as $1/N$ as N increases, the sum Z is Poisson distributed instead — the details matter!)

The Central Limit Theorem. The result we found for the Binomial distribution is a special case of a very general theorem, called the central limit theorem. We will not prove this theorem here, but state it and explains its applications.

A random variable Z which is a sum of a sequence of N independent, identically distributed random variables X_i:

$$Z = \frac{1}{N} \sum_{i=1}^{N} X_i \ , \tag{4.130}$$

is described by a Gaussian or **Normal** distribution when N becomes large as long as the standard deviation σ of each X_i if finite.

Central limit theorem: The random variable $Z = \sum_{i=1}^{N} X_i$, where all the X_i have the same distribution with expected value $E(X_i) = \mu$ and variance σ^2, will for large N approach the distribution:

$$f_Z(z) = \frac{1}{\sqrt{2\pi N\sigma^2}} e^{-\frac{1}{2}\left(\frac{z-N\mu}{\sqrt{N\sigma^2}}\right)^2} . \tag{4.131}$$

This distribution is called the **Normal distribution**. We say that Z is Normally distributed with expected value $E(Z) = N\mu$ and variance $\mathrm{Var}(()Z) = N\sigma^2$.

This means that if we introduce the variable $U = (Z - N\mu)/\sqrt{N\sigma^2}$, then the distribution of U is what we call Normally distributed with expected value 0 and deviation 1. We write this as

$$U \sim N(0, 1) . \tag{4.132}$$

Application to the Binomial distribution. We can apply this theorem to the Binomial distribution for $p = 1/2$. In this case we have that X_i is a Bernoulli trial with probabilities $p = q = 1/2$. Therefore $E(X_i) = 1/2$ and $\sigma^2 = 1/4$ and the distribution for Z is

$$f_Z(z) = \frac{1}{\sqrt{2\pi N\sigma^2}} e^{-\frac{1}{2}\left(\frac{z-N/2}{\sqrt{N/4}}\right)^2} , \tag{4.133}$$

which is the same as we found by various methods above.

The central limit theorem is general. Notice that the central limit theorem is general. It is valid for any sum of independent processes. This is why the Normal distribution is so common, because many of the processes we observe around us is the result of a sum of many individual processes and hence may be approximated as the sum of many random variables. However, there are some limitations to its use. In particular, we need to ensure that the two basic assumptions are valid: The variables must be independent and their variance must be finite. **ams 1**: Provide examples here.

Sharpness increases with number. We expect that if we throw a dice ten thousand times and take the average value, we would get a better approximation to the real average — the expected value — than if we only did it ten times. Now we have the tools to quantify this intuition. If we throw a six-sided dice N times, the average value is $\bar{Z} = (1/N)\sum_i X_i$. We notice that $\bar{Z} = Z/N$, where Z is the sum of X_i. Now, this is a random variable, and we know that as N becomes large, it will be (approximately) Normally distributed. The expected value of $\bar{Z}$ is

$$E(\bar{Z}) = (1/N)E(Z) = (1/N)\, N\, \mu = \mu\ , \tag{4.134}$$

where $E(Z) = N\mu$ and μ is the expected value of a single throw. And the variance is

$$\mathrm{Var}(()\bar{Z}) = \mathrm{Var}(()Z/N) = (1/N)^2\mathrm{Var}(()Z) = \sigma^2/N\ . \tag{4.135}$$

Therefore, we see that as N increases, the variation in $\bar{Z}$ around the expected value μ will become smaller. Indeed, the standard deviation will go as $1/\sqrt{N}$.

4.7.5 Extreme value statistics

The central limit theorem describes the behavior of the average value of a sequence of identical variables, X_i. However, we can also ask other questions, such as what is the minimal or the maximum value of the sequence of N elements. This would for example be relevant if we were interested in the stength of a chain — the strength would be determined by the element with the smallest strength. What is the distribution of such elements?

Numerical estimate. We could address this numerically. We generate a set of elements that are uniformly distributed between 0 and 1, find the minimum of this, and then perform this experiment many times, and characterize the resulting distribution.

We generate 10000 samples using the following program.

```
# Extreme value statistics
import numpy as np
import matplotlib.pyplot as plt
Nsamp = 10000
M = 1000
u = np.random.rand(M,nsamp)
umin = np.min(u,axis=1)
```

```
[n,x] = np.histogram(umin,bins=100)
xmid = 0.5*(x[:-1]+x[1:])
ft = n/nsamp
plt.subplot(2,1,1)
plt.plot(xmid,ft,'o')
plt.xlabel('u'),plt.ylabel('P(u)')
plt.subplot(2,1,2)
plt.semilogy(xmid,ft,'o'),
plt.xlabel('u'), plt.ylabel('P(u)')
```

The results in Fig. 4.13 show that the results look like an exponential distribution. Can we find a theory for this?

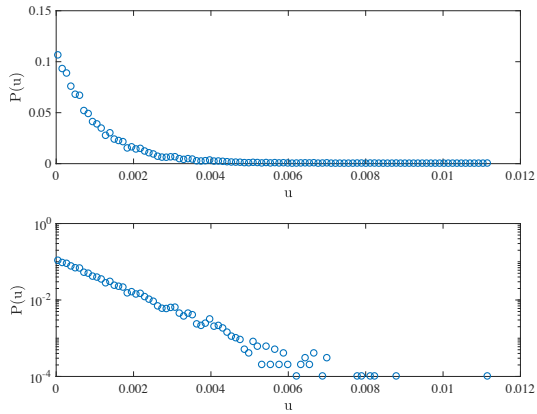


Fig. 4.13 Plot of the probability $P(u)$ for the smallest value $u = \min\{X_i\}$ among $N = 1000$ X_i values that are uniformly distributed from 0 to 1.

Theoretical approach. We start from a sequence of N identically distributed random variables, X_i, and find the minimum u_N. This means that u_N must be smaller than or equal to all the X_i. The probability to be smaller than or equal to X_i is

$$P(X \leq x) = \int_0^x dx = x \; . \tag{4.136}$$

The probability for u to be smaller than each of the N X_i is therefore

$$P(u) = (P(X < u))^N = u^N \; . \tag{4.137}$$

ams 2: This argument did not work out properly, check textbook on statistics.

General result for extreme value statistics. ams 3: Do we want this part. The result is somewhat complicated with the Gumbel, Weibull and Frechet laws.

4.8 Summary

Statistical experiment. A **statistical experiment** is a trial with a given set of possible outcomes.

Outcome. A result n_i of a staticial experiment is called an **outcome**.

Probability. The frequency based definition of the probability p_i for the event that $n = x$ is that $P(n = x) = N_x/M$, where N_x is the number of times that $n = x$ was observed in M events, in the limit when M becomes large.

Random variable. A random variable Z describes the result of an experiment. The outcome of a particular experiment is z.

Probability density. The probability for a set of discrete outcomes is $P_Z(z)$. The probability for an outcome z in the range from z to $z + dz$ from a continous set of outcomes is $f_Z(z)dz$, where $f_Z(z)$ is called the probability density.

Expected value. The **expected value** of Z is $\langle Z \rangle = \sum_z z P_Z(z)$ for a discrete set of outcomes and $\langle Z \rangle = \int z f_Z(z) dz$ for a continuous set of outcomes. The expected value is estimated by the **average**, $\bar{z} = (1/M) \sum_{i=1}^{M} z_i$.

Standard deviation. The variation in Z is characterized by the variance σ^2 and the variation $E((Z - E(Z))^2)$. The variance is estimated from the **standard deviation** $\overline{\sigma^2} = (1/M) \sum (z_i - \langle z \rangle)^2$ or $\overline{\sigma^2} = (1/(M - 1)) \sum_i (z_i - \bar{z}^2)^2$.

4.9 Exercises

Exercise 4.1: Deck of cards

We will address various probabilities for drawing cards from a deck of cards.

a) What is the probability to draw an ace of spades?

b) What is the probability to draw a two of diamonds?

c) What is the probability to draw a black card?

d) What is the probability to draw a spade?

e) What is the probability to draw a card that is not a spade?

f) What is the probability to draw two spades from a full deck?

g) What is the probability to draw five cards from a full deck and get four cards of equal value?

Exercise 4.2: Dice

We are using normal, fair, six-sided dice:

a) What is the probability to get a 1?

b) What is the probability to not get a 1?

c) If we throw two dice, what is the probability that both dice show 6?

d) If we throw two dice, what is the probability that the sum is 3?

e) If we throw two dice, what is the probability that the sum is j for all j?

f) If we throw two dice two times, what is the probability to get two six'es at least once?

g) If we throw four dice, what is the probability to get at least two six'es?

h) If we throw two dice, what is the probability that both dice show the same?

i) If we throw two dice, what is the probability that the sum is odd? And even?

j) If we throw one dice six times, what is the probability that we get at least one six?

Exercise 4.3: Casino in St. Petersburg

In the Casino in old St. Petersburg they had a special game. You paid a fee to participate, and then you started to flip a coin, and you continued flipping until you got heads. If heads came up for the first time after N flips, you won 2^N rubles. You can assume that the probability to flip a head is p.

a) Find the probability $P(N)$ for the coin to show heads for the first time at flip N.

b) Show that the probability $P(N)$ is normalized.

c) Find the average number of flips needed to get the first heads. where we have introduced $q = (1 - p)$. We use a common trick to find such sums, by realizing that the sum can be written as

$$\sum_{n=0}^{\infty} nq^n = \sum_{n=0}^{\infty} q\frac{d}{dq}q^n = q\frac{d}{dq}\sum_{n=0}^{\infty} q^n = q\frac{d}{dq}\frac{1}{1-q} = \frac{q}{(1-q)^2} = \frac{q}{p^2} \,, \tag{4.138}$$

which we insert back to get

$$\langle N \rangle = 1 + p\frac{q}{p^2} = 1 + \frac{1-p}{p} \,. \tag{4.139}$$

d) Find the average return for a player in this game. Show that the return approaches infinity unless p is larger than $1/2$.

e) If $p = 0.6$, what must the entry fee for the game be to ensure that the Casino does not loose money?

Exercise 4.4: Coin-flip game

Two persons are playing a simple coin-flip game. They flip a coin in turn, and whoever first gets a heads wins the game.

a) Write a program to model $M = 10000$ such games. Estimate the probability for the person who starts flipping to win the game.

b) Find the probability for the first person to flip (the person who starts flipping) to win the game.

Exercise 4.5: Waiting time statistics

Many problems in physics are about waiting — waiting for an event to occur. For example, we often ask questions of the type — what is the probability for an event X to happen for the first time after N attempts. Concepts such as mean-free-path and mean-free-time are often associated with such processes.

As an example of a waiting time distribution problem, let us address electrons with a narrow distribution of kinetic energies around E

($E \gg kT$). These electrons are injected into a material, where they have a certain probability to the scattered. The probabilitiy for scattering depends on the number of scattering centers the electron passes. This gives a probability $p\Delta x$ to be scattered after a short distance Δx.

a) When the electron starts from $x = 0$, show that the probability, $P(x)$, for the electron to reach x without being scattered is e^{-px}.

b) What is the probability for the electron to be scattered in the interval from x to $x + dx$?

c) How far does the electron travel on average before it is scattered?

d) What is the probability for the electron to travel for a time t without getting scattered?

e) What is the average time the electron travels before being scattered? This is called the average-free-time between collisions.

Exercise 4.6: Poisson statistics

We assume that you are studying a rare event such as the chance of getting the number 1000 if you spin a wheel of fortune with 1000 numbers from 1 to 1000. To make our results more general we call the probability to get 1000 p.

a) If you spin N times, what is the probability to get 1000 exactly one time.

b) What is the general probability $P(N, n)$ to get 1000 n times in N spins. What is the average value of n, $\bar{n}$?

c) Show that the expression can be approximated as

$$P(N, n) = \frac{\bar{n}^n e^{-\bar{n}}}{n!} \,, \tag{4.140}$$

when N approaches infinity and $p \ll 1$.

Exercise 4.7: Gamma distribution

Let us address an unfair coin with a probability p for heads. We flip the coin until we have obtained n heads, which concludes the experiment. The result of the experiment is N, the number of flips.

a) Find the probability $P(N, n)$ that we need to flip N times to get n heads.

b) Show that the probability is normalized.

c) Find an expression for the average N and the standard deviation of N.

Exercise 4.8: Stirling's formula

We often use Stirling's formula as an approximation for the factorial. However, the formula is only supposed to be a good approximation for $N \gg 1$. Let us see how good the approximation is for smaller N.

Stirling's approximation gives

$$N! = N^N \sqrt{2\pi N} \exp\left(-N + \frac{1}{12N} + \mathcal{O}\left(\frac{1}{N^2}\right)\right) , \tag{4.141}$$

or

$$\ln N! = \frac{1}{2} \ln(2\pi N) + N(\ln N - 1) + \frac{1}{12N} + \mathcal{O}\left(\frac{1}{N^2}\right) , \tag{4.142}$$

However, we often use simpler versions of the formula, such as

$$S_1(N) = N^N e^{-N} \ , \ S_2(N) = \sqrt{2\pi N} S_1(N) \ , \ S_3(N) = e^{\frac{1}{12N}} S_2(N) \ . \tag{4.143}$$

a) Write a program to compare the values for $N = 1, 2, 4, 8, 16, 32, 64$. Plot the three resulting functions and the exact values.

b) Find the relative error from using $N! \simeq S_3(N)$ for $N = 10^{20}$.

c) How large should N be for $\ln S_1(N)$ to represent $\ln N!$ with a relative error of less than 10^{-6}?

Exercise 4.9: Conducting strip

An experiment may result in two possible outcomes, heads with probability p and tails with probability $q = 1 - p$.

a) What is the probability that N independent flips all results in heads?

b) What is the probability for n heads and $N - n$ tails in N flips?

c) From the result above, find the average number of heads, $\bar{n}$ and the variance

$$(n - \bar{n})^2 . \tag{4.144}$$

Now, let us consider a strip consisting of N squares as illustrated in fig. 4.14. The squares are electrically conductive with probability p. Two conducting squares next to each other provide a conductive path.

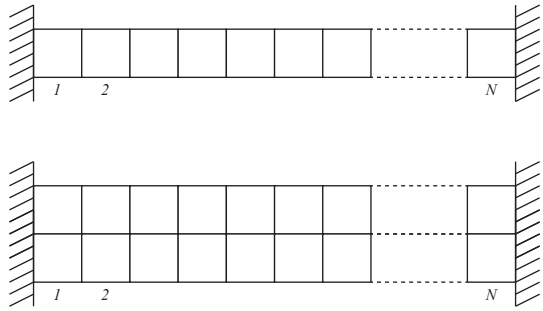


Fig. 4.14 Illustration of strip of conducting patches.

d) What is the probability for the strip to be conducting? What is the probability for one isolating square in the strip?

Now, we study two strips next to each other, that is, $2N$ squares. We assume that a conducting path can be formed both by two squares sharing a side and when touching at a corner.

e) What is the probability, P_c, for the double-strip to be conductive?

Let us assume that we have $2N$ squares, of which $N+M$ are conductive and differ from the others by a metallic surface. The $2N$ squares form a double-strip of length N.

f) What is the total number of different configurations Ω_{tot}?

g) What is the total number of conducting configurations, Ω_{cond}?

h) If we define

$$P_{\text{cond}} = \frac{\Omega_{\text{cond}}}{\Omega_{\text{tot}}} ; \tag{4.145}$$

and introduce $p = (N + M)/(2N)$, will we get the same situation as for P_c above or is the situation different? If this is the case, explain the difference.

Exercise 4.10: Velocity distribution in Argon gas

Let us use the molecular simulation approach to measure and characterize the velocity distribution in a gas.

a) The file in.gasexer010[11] can be used to simulate a two-dimensional gas. The resulting output file is found in gasexer010.lammpstrj[12]. Measure and plot the distribution of the x-component of the velocities of the atoms in the final time-step.

b) What is the average velocity $\bar{v}_x$ and the standard deviation of the velocities, σ_v? Comment on the results.

c) Suggest a function $P(v_x)$ that can fit the data and check how well it fits the measured data.

d) Compare the distribution of v_x and the distribution of $v = \sqrt{v_x^2 + v_y^2}$. Comment on the results. (Difficult) Can you predict the functional form of $P(v)$?

e) Run a sequence of simulations for $N = 200, 400, 800, 1600$ atoms and find the standard deviation, σ_v, as a function of N. Find a reasonable functional form for $\sigma_v(N)$. What happens when N becomes realistically large, such as $N = 10^{23}$?

f) Run a sequence of simulations for various values of the initial velocity (energy) used in the simulations $v_0 = 1.5, 2.0, 2.5, 3.0$, and 3.5. Find the standard deviation σ_v as a function of K/N, the average kinetic energy per particle. Comment on the result. What does this tell you about the temperature-dependence of the velocity distribution?

Exercise 4.11: The cumulative distribution function

While the probability desity $f_Z(z)$ for a random variable Z gives the probability for Z to be in the range from z to $z + dz$:

$$P(z < Z < z + dz) = f_Z(z)dz \, , \tag{4.146}$$

the **cumulative distribution function**, $F_Z(z)$, of Z gives the probability for Z to be smaller than z:

$$F_Z(z) = P(Z \leq z) = F_Z(z) = \int_{-\infty}^{z} f_Z(z)dz \, . \tag{4.147}$$

[11] `http://folk.uio.no/malthe/fys2160/in-gasexer010`
[12] `http://folk.uio.no/malthe/fys2160/gasexer010.lammpstrj`

a) Show that we can find the probability density $f_Z(z)$ from the derivative of the cumulative distribution function:

$$f_Z(z) = \frac{d}{dz} F_Z(z) \, . \tag{4.148}$$

b) Show that the probability for Z to be larger than z is

$$P(Z > z) = 1 - F_Z(z) \, . \tag{4.149}$$

Let us assume that Z is uniformly distributed on 0 to 1.

c) What is the cumulative distribution function for a random variable Z, which is uniformly distributed on the interval from 0 to 1, that is, for which

$$f_Z(z) = \begin{cases} 1 & 0 < z < 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.150}$$

We now introduce a random variable $Y = Z^a$.

d) Show that the cumulative distribution for Y is $F_Y(y) = 1 - F_Z(y^{1/a})$.

e) Show that the probability density for Y is $f_Y(y) = (1/a)y^{1/a-1}$ for $0 < y < 1$. (This means that we have found a way to generate a random variable with a power-law distribution from a uniformly distributed variable. We often use the cumulative distribution when we want to find the results of variable transformations on random variables).

Exercise 4.12: Sharpness of the temperature

We measure the temperature of a system of many particles from the average kinetic energy of the particles:

$$T = c \frac{1}{N} \sum_i K_i \, , \tag{4.151}$$

where c is a constant prefactor that we will introduce later. For simplicity, assume that the kinetic energies are drawn from some random distribution which is the same for all the particles with average $E(K_i) = \mu$ and variance $\mathrm{Var}(()K_i) = \sigma^2$.

a) What is the expected value of the temperature?

b) What is the variance of the temperature?

c) What is the standard deviation of the temperature for a gas with 10^{23} atoms? (You can assume that σ is smaller than μ, $\sigma < \mu$, in order to make a coarse estimate). Comment on the results.

d) Sketch the probability distribution for T for various values of N and comment on the results.

Exercise 4.13: Central limit theorem for a dice

Let us check the central limit theorem for a single dice. We throw a six sided dice N times and measure the average value $\bar{Z} = (1/N)\sum_i X_i$, where X_i is the result of throw number i.

a) Write a program to find $\bar{Z}$ for a given value of N.

b) Find the standard deviation of $\bar{Z}$ for $N = 10^1, 10^2, 10^3, 10^4$ from $M = 10^5$ simulations. Plot the results. How does the results compare to the central limit theorem?

Exercise 4.14: Poisson distribution

The Poisson distribution for the number of events n in a interval with an average number of events λ is given as

$$P(\lambda; n) = \frac{\lambda^n e^{-\lambda}}{n!} \,. \tag{4.152}$$

a) Find the expected value, $E(n)$, of the Poisson distribution.

b) Find the variance, $\mathrm{Var}(()E)$, of the Poisson distribution.

Exercise 4.15: Multi-nomial distribution

For the binomial distribution, we throw a dice with two possible outcomes A and B and count how many times we get an A. For the trinomial distribution, we throw a dice with three possible outcomes, A, B and C, and ask, what is the probability to observe N_A of type A, N_B of type B and N_C of type C, where the total number of outcomes are $N = N_A + N_B + N_C$ and the probability to get A, B, or C in a given trial is p, q and r respectively, so that $p + q + r = 1$.

a) Show that the probability $P_N(N_A, N_B, N_C)$ is

$$P_N(N_A, N_B, N_C) = \frac{N!}{N_A!\, N_B!\, N_C!} p^{N_A} q^{N_B} r^{N_C} \ . \tag{4.153}$$

b) Find $E(N_A)$, $E(N_A^2)$, and $E((N_A - E(N_A))^2)$.

c) Generalize the probability to the multinomial distribution with M different species.

Exercise 4.16: Conditional probabilities

The *conditional probability* for an event A to occur, given that event B occurred, is written as $P(A|B)$. For example, if you throw a six sided dice, and you know that the result was even (event B), what is then the probability for the results to be 2 (event A)? This is $P(X = 2|X \text{ is even })$.

Simple example. We start from a simple example with an experiment with N equally likely outcomes. Assume that event A and B contain a and b outcomes, and that c is the number of outcomes that are both in A and B.

a) Draw a sketch of A, B and C = A ∩ B.

b) What is the probability of B given A in this case?

c) Argue that in the general case, also when the outcomes are not all equally likely, we have that $P(A|B) = P(A \cap B)/P(B)$ and show that this gives the same result as above.

d) The lifetime T of a LED lamp is described by the density $f_T(t) = (1/T)e^{(-t/T)}$. If the bulb survives a time T, what is the probability that it will die in the interval from $0.8T$ to $0.9T$?

e) If the probability to get a specific disease is $p = 0.05$, and the probability to die if you get the disease is $q = 0.01$, what is the probability for any person in the population to die from the disease?

Exercise 4.17: Conditional probability expansion

The following theorem describes how we can use conditional probabilities to decompose a probability in a similar way to what we do for a state in a quantum mechanical basis.

For a set A_i of event that are mutually exclusive ($A_i \cap A_j = \emptyset$), span the entire set of outcomes ($\cup_i A_i = S$), and with finite probabilities ($P(A_i) > 0$), then

$$P(B) = \sum_i P(B|A_i)P(A_i) \ , \tag{4.154}$$

for any set B.

Two-in-a-row. Let us study a simple game on a two-by-two matrix. Each site in the matrix is filled with a probability p. You win the game if you get two-in-a-row either vertically or horizontally. Let us find out what p must be for this game to have a 0.5 change of winning.

a) First, list up all the possible configurations, c. Can any two such configurations occurs at the same time?

b) For each configuration c, find the probability for this configuration when we assume that each site is filled independently with a probability p.

c) For each configuration c, what is the probability to get a two-in-a-row?

d) Find the probability for two-in-a-row, $P_2(p)$.

e) How would your calculation change if we also included the diagonals in the winning configurations?

Percolation on a 2×2 system.

f) Instead of playing two-in-a-row, we would like to find the probability to have a connected path of filled sites from the left side to the right side of the two-by-two matrix. Two sites are connected if they are nearest neighbors, that is, if they share a side. Two sites are not connected if they only share a corner. Repeat the argument above to find $\Pi(p)$, the probability for there to be a path from the left to the right.

g) Can you find the result also for a 1×1 system and for a 3×3 system? Plot all the results in the same plot and speculate as to what happens as the system size becomes very large.

Exercise 4.18: Bayes theorem

Bayes theorem states that for a set of mutually exclusive events A_i that span the whole sample space, we have for an event B that:

$$P(A_j|B) = \frac{P(B|A_i)}{\sum_i P(B|A_i)P(A_i)} \ . \tag{4.155}$$

a) You are screened for a rare disease that only 1 in 10000 has. The test that is used correctly identifies 90 of 100 of those who has the disease. However, the test gives one false positive out of every 1000 who take the test. What is the probability that you have the disease if the test is positive?

Exercise 4.19: Joint probability distributions

For two random variables X_1 and X_2, the probability for $X_1 = x_1$ and $X_2 = x_2$ at the same time is called the *joint probability distribution* $f(x_1, x_2) = P(X_1 = x_1 \text{ and } X_2 = x_2)$.

Flipping a coin: We flip a coin two times. The result of the first throw is X_1 and the second is X_2, where X_1, X_2 can be either 0 or 1 corresponding to heads and tails respectively.

a) List all the possible outcomes and find their probabilities. Find $f(x_1, x_2)$. Check that $f(x_1, x_2)$ is normalized.

b) Can you find a general rule for how to find $f(x_1, x_2)$ in this case? How can you generalize your result to N number of coins?

Drawing without replacement: We draw two balls from a bowl with four balls. (We do not place the first ball back into the bowl). Initially, in the bowl there are two balls with the number 0 on them, and two balls with the number 1.

c) List all the possible outcomes in a table and find their probabilities. Check that $f(x_1, x_2)$ is normalized.

d) Discuss differences and similarities between the coin and the ball experiments.

Exercise 4.20: Joint probability density

Exercise 4.21: Maximum likelihood estimators

ams 4: Introduce likelihood and maximum likelihood estimators, find maximum likelihood estimators for a set of uniform and normally distributed variables

Exercise 4.22: Gamma distribution

ams 5: Introduce the gamma distribution and its applications.

A thermodynamical example

5

Abstract We have now seen how a system approaches equilibrium using both molecular dynamics modeling and algorithmic modeling, but what characterizes the equilibrium of a system? Here, we briefly introduce the thermodynamic principles of equilibrium for an ideal gas. We measure the ideal gas law in molecular simulations and develop a simplified theory for the equation of state for the gas. We also quickly introduce a few key concepts that will reappear with full force later in the book.

Temperature. We start with a quick introduction to thermodynamcis with emphasis on an example – an (ideal) gas. Here, we will focus of on a few key concepts of the gas — in particular on the concept of temperature. What do we know about temperature? Not much, yet, because we have not introduced it rigorously. We do know that temperature is what we measure with a thermometer. However, this is not that useful. We might expect it to be related to energy. If all the atoms in a gas move faster — so that the average or total kinetic energy increases — we expect the gas to have a higher temperature.

The typical definition of temperature is that temperature is a *property* that is equal when two systems are in thermal equilibrium. If a highly agitated gas comes in contact with a less agitated gas, we expect the two gases to becomes equally agitated over time. We expect the temperature to become the same over time. And in the process the system with a lot of kinetic energy will tranfer it to the system with less kinetic energy.

Temperature and energy. However, notice that even though temperature is related to energy, it is not the same as energy. We can see that from two though experiments:

- If we increase the velocity of all particles we do not expect the temperature to increase.
- If we lift all the particles upwards in a gravitational field, increasing their energy, we do not expect the temperature to increase.

Temperature is not the same as energy, but it is related to the distribution of energy. One of our goals for this course is to develop a good understanding of temperature. Here, we will start from a simple example — how is the energy in a gas related to other properties of the gas?

5.1 Ideal gas

An ideal gas is a gas that has so low density that there are no interactions between the particles. It is simply a box filled with non-interacting particles (atoms or molecules). We can model an ideal gas using molecular dynamics in a system with very low density.

5.1.1 Ideal gas — Results from simulations

Let us find out how the pressure in the gas depends on other properties of the system. The pressure is a property we can measure in the gas and which may vary or at least fluctuate with time. The volume and the total energy on the other hand, are values we determine when we start the simulations, and we typically do not allow them to vary throughout the simulation: Total energy is conserved when the particles move in conservative force fields, and we keep the size of the box, and hence the volume, constant.

== Measuring pressure and total kinetic energy === If the gas is very dilute, all the particles are far away from each other most of the time, and the interaction energies from the interatomic potentials will be very small (since it decays rapidly with the distance between atoms). We may therefore approximate the total energy by the kinetic energy of the atoms. We will come back later to how to measure the pressure in a molecular dynamics simulation, but let us here assume that it is

measured from the average force on the walls of the system. We introduce reflective walls on all sides to contain the system.

We can set up a three-dimensional simulation of a dilute gas, just as we did before, and measure the volume $V = L_x \cdot L_y \cdot L_z$, the pressure P, as measured by the simulation program, and the total kinetic energy

$$K = \sum_{i=1}^{N} \frac{1}{2} m_i \left(v_{xi}^2 + v_{yi}^2 + v_{zi}^2 \right) \ , \tag{5.1}$$

which is also measured directly in the simulation program.

Setting up and running simulations. To characterize the behavior of a gas, we perform simulations and record measurements for various values of the initial velocities (and hence K) and volumes V of the system. We do this by changing the input file used by the simulation program, run the simulation, extract the volume, pressure and temperature from the log file produced by the simulation. This is implemented in the program findeqstategas01.py[1]:

```
import numpy as np
import matplotlib.pyplot as plt
import re
import os
import lammps_logfile
T = []
V = []
P = []
myvelocities = np.array([1.5,2.0,2.5,3.0,3.5])
myvolumes = np.array([0.010, 0.020, 0.040, 0.080])
for ivel in range(0,len(myvelocities)):
    for ivol in range(0,len(myvolumes)):
        # Change the word mydensity to myvolumes[ivol]
        infile = open("in.gasstatistics30",'r')
        intext = infile.read()
        infile.close()
        replacestring = "%f" % (myvolumes[ivol])
        intext2=intext.replace('mydensity',replacestring)
        # Change the word myvelocity to myvelocities[ivel]
        replacestring = "%f" % (myvelocities[ivel])
        intext3=intext2.replace('myvelocity',replacestring)
        infile = open("in.tmp",'w')
        infile.write(intext3)
        infile.close()
        # Run the simulator
        print("Executing lmp_serial < in.tmp")
        os.system("lmp_serial < in.tmp") # Run lammps
        # Extract data from logfile
```

[1] http://folk.uio.no/malthe/fys2160/findeqstategas01.py

```
        log = lammps_logfile.File("log.lammps")
        p = log.get("Press") # Get pressure from log file
        P.append(np.average(p)) # Add average pressure value
        v = log.get("Volume")
        V.append(np.average(v))
        t = log.get("Temp")
        T.append(np.average(t))
PV = np.array(P)*np.array(V)
plt.plot(T,PV,'o')
```

Notice the use of regular expressions to change the volume (actually the density) and initial velocities in a default file, and then write the resulting text into a new file which is read by the simulation. Notice also how we can run a program from the command line from python. This script runs a sequence of 20 molecular dynamics simulations, and the result is shown in Fig. 5.1.

Ideal gas law. The plot shows PV as a function of K. It is not easy to know how to plot such data to find general relationships. Often we have to make guesses and test them out, but you can get hints from dimensional analysis: We see that both PV and K have units of energy. However, when we plot the data in this way, we see that PV is approximately linearly dependent on K with a constant of proportionality equal to $3/2$:

$$PV = \frac{3}{2}K \ , \tag{5.2}$$

We may rewrite K as $N(K/N)$, where N is the number of atoms in the simulation. The kinetic energy per particle may be used as a measure of how energetic the gas is, and we could define the this to be the temperature, or proportional to the temperature with some factor of proportionality k_B to ensure that temperature has reasonable values for the melting point of water and the boiling point of water, since these are important scales for us. We therefore, for now, introduce the temperature T through $K/N = k_B T$, where k_B is called Boltzmann's constant:

$$k_B = 1.3806488(13) \cdot 10^{-23}\text{J/K} = 8.6173324(78) \cdot 10^{-5}\text{eV/K} \ . \tag{5.3}$$

We can therefore rewrite the equation of state as

$$PV = \frac{3}{2}Nk_B T \ , \tag{5.4}$$

This form of the result is general and true also in one and three dimensions. Indeed, what we have found here is the **ideal gas law**, or the **equation**

of state for the ideal gas. An **ideal gas** is a gas that is so dilute there are no collisions (no interactions) between the atoms.

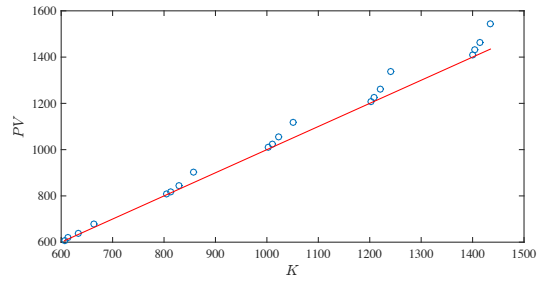


Fig. 5.1 Plot of the histogram for $n(t)$ from the simulation and n_i from a binomial distribution.

5.1.2 Ideal gas — Results from theory

We have now observed the ideal gas law in our simulations. Can we also develop a theoretical understanding and a theoretical argument for the ideal gas law? We can do that by making a simplified, microscopic model.

A collision reverses the velocity orthogonal to the wall. Let us start by addressing the motion of a single atom bouncing back and forth in a volume $V = L^3$. What is the average pressure P on the walls of the container due to these collisions?

If each bounce is conservative and the only change in momentum is in the direction normal to the wall, then the velocity in the direction normal to the wall is reversed in each collision. For a collision along the x-axis this corresponds to $v_x \to -v_x$, and a change $\Delta p_x = 2mv_x$ in momentum. The average force F_x over a time interval Δt is then

$$F_x = \frac{1}{\Delta t}\Delta p_x = \frac{1}{\Delta t} m \Delta v_x \ . \tag{5.5}$$

What time interval Δt should we use? We use the time between each collision. The particle travels a distance $2L$ before hitting the same wall again (in the mean time it has also hit the opposite wall) with a typical velocity v_x, hence $\Delta t = 2L/v_x$,

$$F_x = \frac{mv_x^2}{2L} \ , \tag{5.6}$$

and the average pressure on this wall of area $A = L^2$ from the collisions with a single particle is

$$P = \frac{F_x}{A} = \frac{-m(-2v_x^2)}{A2L} = \frac{mv_x^2}{AL} = \frac{mv_x^2}{V} \ . \tag{5.7}$$

If we now sum the effects of all N particles in the gas we get a contribution to the pressure from motion in the x-direction:

$$PV = Nmv_x^2 \ . \tag{5.8}$$

We get similar results for the (independent) motion in the y- and z-directions, and therefore:

$$PV = Nmv_x^2 + Nmv_y^2 + Nmv_z^2 \ . \tag{5.9}$$

We assume that all the velocities are approximately the same, $v_x \simeq v_y \simeq v_z$, and therefore:

$$PV = \frac{3}{2}N\frac{1}{2}mv_x^2 \ . \tag{5.10}$$

This looks very similar to the ideal gas law, $PV = Nk_BT$ which we found from the moleculary dynamics simulations. Therefore we have related temperature to the velocity of a particle:

$$\frac{1}{2}kT = \frac{1}{2}mv_x^2 \ . \tag{5.11}$$

5.1.3 Ideal gas — equipartition principle

We have therefore found that the average energy for each degree of freedom of motion is:

$$\langle\frac{1}{2}mv_x^2\rangle = \langle\frac{1}{2}mv_y^2\rangle = \langle\frac{1}{2}mv_z^2\rangle = \frac{1}{2}kT \ . \tag{5.12}$$

This is a general result, as we will prove later, and it is valid not only for the kinetic energy, but for all terms in the energy that are quadratic in position or velocity. We call this principle the **equipartition principle**:

> **Equipartition principle:** In equilibrium, every (quadratic) term in the sum for the energy of a system has an energy which is $(1/2)kT$

This simple principle can be applied to provide insight into the two fundamental model systems for gases and solids:

Simplfied model for a gas — ideal gas. An ideal gas where there are no interactions between the atoms/molecules. This means that the particles only have kinetic energy and no (internal) potential energy due to interparticle interactions. The total number of quadradic terms in the energy sum is therefore $3N$ — there are three quadratic terms for each particles. The equipartition principle therefore tells us that

$$U = \frac{3}{2}NkT \,, \tag{5.13}$$

If there are more degrees of freedom, such as from rotation or vibrations of molecules, these must also be included by including f, the number of degrees of freedom into the expression, and the expression for the energy of the gas becomes

$$U = \frac{f}{2}NkT \,, \tag{5.14}$$

where f is:

- $f = 3$ if the only degree of freedom per particle are motion along the three orthogonal directions.
- $f = 5$ if we include rotation in a diatomic gas. (Two new degrees of freedom to describe rotation).
- $f = 7$ if we also include vibration in a diatomic gas. (Two additional degrees of freedom to describe vibration).

Simplified model for a solid — ideal solid. In the simplest model for a solid, each particle in the solid can move in three directions. A solid with N atoms therefore have 6 degrees of freedom for each atom, three degrees of freedom for their positions relative to their equililibrium positions and three degrees of freedom for their velocities. This gives: 3

$$U = 3NkT \,, \tag{5.15}$$

for a solid.

Clickers - Equipartition

5.2 Summary

- An ideal gas is described by the ideal gas law $PV = Nk_BT$.

- Each degree of (quadratic) freedom for a system contributes an amount $(1/2)kT$ to the total energy in equilibrium.

5.3 Exercises

Exercise 5.1: Limit of ideal gas law

Set up simulations of Lennard-Jones atoms in three dimensions using the LAMMPS input file `in.nvt_3d`:

```
# Atomify commands start with #/
#/camera position -15 25 25
#/camera viewcenter 5 5 5
# Set atom size and color
#/atom 1 1.0 #ffffff
variable T equal 1.6
variable rho equal 0.8
units lj
dimension 3
lattice fcc $(v_rho)
region simbox block 0 5 0 5 0 5
create_box 1 simbox
create_atoms 1 box
mass 1 1.0
velocity all create $T 87287
pair_style lj/cut 2.5
pair_coeff 1 1 1.0 1.0 2.5
fix 1 all nvt temp $T $T 1
variable ideal_gas_pressure equal $T*$(v_rho)
variable pressure_ratio equal c_thermo_press/v_ideal_gas_pressure
run 100000
```

a) Systematically change the density of the system to determine for how high densities the ideal gas law is valid.

Micro- and Macro-states

6

Abstract In this chapter we introduce the concepts of micro- and macro-states.

We have seen that a gas develops towards a state where the number of particles is uniform throughout the gas. We measured the number of particles on the left side of the gas, and found that the system naturally evolves towards a state with a uniform distribution. While the number of particles in one half of the gas is fluctuating, the fluctuations becomes very small compared with the number of particles, and the number of particles increases. The behavior we observed for the gas is not a special case, but a very general behavior in all systems with many particles — the system will develop towards an equilibrium state. So far, we have only analyzed the number of particles — or the density — in a part of a system. In this chapter we will extend this analysis to address *thermal fluctuations*, fluctuations in energy, *mechanical fluctuations*, fluctuations in volume, and *density fluctuations*, fluctuations in the number of particles in a system.

Behavior of isolated systems. Our plan is to follow the same strategy as we introduced for the gas: We will study an isolated system — in molecular dynamics the number of particles, the total energy and the volume is constant, which means that the system is completely isolated. To understand what happens inside this isolated system we will divide the system in two, and study how the energy, the volume or the number of particles are distributed between the two halves, just like we have done for the gas system.

Equilibrium state. Already for the gas, we saw that the system evolved towards an *equilibrium* state, which corresponded to the most probable value for n, and that the development was *irreversible* when we started from a state that was far from the equilibrium state. Here, we will develop a better understanding for irreversible processes and the associated fundamental law, the second law of thermodynamics. We will also study the conditions for the *equilibrium* state.

Coupling between micro-scopic and macro-scopic behavior. However, the most important new concepts introduced in this chapter occurs in the coupling between the microscopic and the macroscopic. We will demonstrate how we can use our microscopic models to determine the macroscopic behavior of systems. We will be able to *predict* quantitative features of macroscopic systems such as the equation of state for an ideal gas or the heat capacity for a gas or a solid. This is made possible by a coupling from the microscopic to the macroscopic — a major achievement in physics. Using this coupling we will introduce the concepts of entropy, temperature, pressure, and chemical potential, which are the essential components of thermodynamics — the theory of the thermal behavior of macroscopic objects.

We start from our description of a gas — a description of both the microscopic and the macroscopic state.

6.1 States of the model gas

We described the gas on several levels: In the molecular dynamics model we described the detailed dynamics of all the particles. But this was too detailed. Instead, we introduced a simplified model where we described the position of each particle using a binary number, n_i, which shows if the particles is on the left side ($n_i = 1$) or not ($n_i = 0$). We call such a description a microscopic description, and we can describe the state of the system by providing the set of n_i values, $(n_1, n_2, n_3, \ldots, n_N)$. We call this set of numbers a *microstate* for the gas.

Given the microstate, that is the set of n_i for every particle i, we could calculate the number of particles on the left side, $n = \sum_i n_i$. We call the value of n a *macrostate* of the system. We found that there are many possible microstates that give the same result for n, that is, there are many possible microstates for a given macrostate.

6.1.1 Micro- and macro-states for two dice

If we throw two six-sided dice, the set of possible outcomes is the set of pairs or numbers, $(1,1)$, $(1,2)$, $(1,3)$ etc. representing the number on dice 1 and dice 2 respectively. That is, the number of possible states of the two-dice system is given by (n_1, n_2), where $n_1 = 1, 2, \ldots, 6$ and $n_2 = 1, 2, \ldots, 6$. These are the *microstates* of the two-dice system. How many microstates are there? This is problem from combinatorics and for two independent dice, there are $6 \times 6 = 36$ such (ordered) two-number states.

The *macrostates* of this system represent a group of microstates. There are many possible ways to divide the total set of outcomes into various groups — various macrostates. For example, we could introduce a set of macrostates corresponding to the sum, $n = n_1 + n_2$, of the two dice. In this case, the macrostate n can vary from 2 to 12. Each unique value of Z represents a macrostate. But there can be many microstates corresponding to a macrostate.

We call the number of microstates that give the same macrostate the *multiplicity* of the macrostate. For the two-dice system, all the microstates — all the possible pairs of numbers (n_1, n_2) are equally probable since the dice are fair and the two throws are independent of each other. The multiplicity of macrostate can therefore be used to find the probability of a macrostate: The probability of the macrostate is the number of microstates in the macrostate divided by the total number of microstates. If we introduce the notation $\Omega(n)$ for the multiplicity, then the probability for state n is

$$P(n) = \frac{\Omega(n)}{\Omega_{TOT}} = \frac{\Omega(n)}{\sum_{n=2}^{12} \Omega(n)} \,, \tag{6.1}$$

where we have used that the total number of microstates is the sum of the number of microstates for each of the possible values of n. Fig. 6.1 illustrates the number of states for each of the macrostates, and you should now be able to find the probability for each macrostate simply by counting the number of microstates.

6.1.2 Definition of micro- and macro-states

A **microstate** is a microscopic description of the degrees of freedom of the system on some level of precision. Typically, but not always,

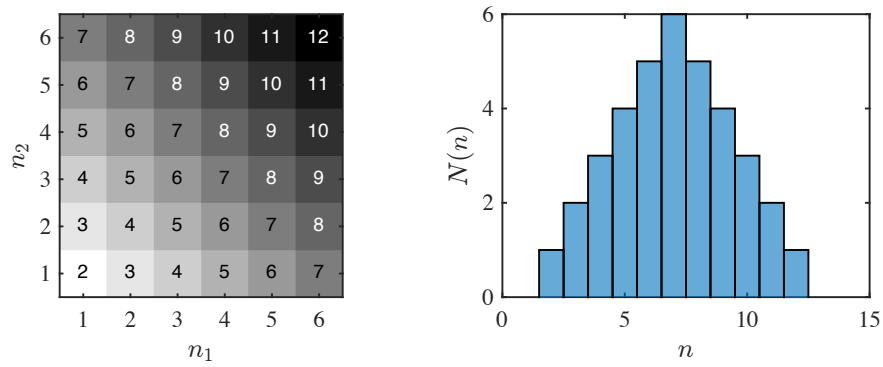


Fig. 6.1 *Left* Illustration of the microstates of two dice. The microstates that are part of the same macrostate $n = n_1 + n_2$ are given the same color. *Right* Plot of the histogram of the number of microstates for each macrostate n for two dice.

we can specify the microstate by specifying the state of each of the N elements in the system, $(n_1, n_2, \ldots, n_N)$.

A **macrostate** is a set of microstates. Usually, the macrostate is specified by a number which can be computed directly from the microstate. We call this number the *macrovariable*. Each microstate belongs to one and only one macrostate. The set of all possible macrostates therefore span all the possible microstates of the system.

The **multiplicity** $\Omega(N, z)$ of a macrostate described by the macrovariable Z, is the number of microstates in the macrostate characterized by the value z.

6.1.3 Multiplicity of the macrostate of the gas

We have already studied a particular macrostate of the ideal gas: The number of particles, n, on the left side of a box. For each possible value of n we found that there were an enormous amount of microstates. In order to make our concepts and the notation very clear, we introduce a more consistent notation: In our theoretical description of the system,

the microstates was a sequence of numbers describing the position of each atom. Each atom may be on the left side ($n_i = 1$) or on the right side ($n_i = 0$) for $i = 1, \ldots, N$. The total number of atoms on the left side was

$$n = \sum_{i=1}^{N} n_i \ , \tag{6.2}$$

We found that the probability distribution for n was given by the binomial distribution

$$P(N, n) = \Omega(N, n) 2^{-N} \ , \tag{6.3}$$

where the *multiplicity* $\Omega(N, n)$ represents the number of microstates that give the value n for n. We call this the number of microstates in macrostate n.

Sharpness of the multiplicity. The multiplicity for the number of particles on the left side of an ideal gas was given as

$$\Omega(N, n) = \binom{N}{n} \ . \tag{6.4}$$

We also found that this probability, and hence also the multiplicity, was very sharp around its average value. Indeed, the average value $\bar{n}$ and the standard deviation σ_n are given as

$$\bar{n} = \frac{N}{2} \ , \ \sigma_n = \sqrt{\frac{N}{4}} \ . \tag{6.5}$$

From this result we see that the standard deviation becomes very small compared with the average:

$$\frac{\sigma_n}{\bar{n}} = \frac{\sqrt{N}}{N} = N^{-1/2} \ , \tag{6.6}$$

which means that for a litre of gas at room temperature and atmosphere pressure, where N is of the order of 10^{23} we see that the standard deviation is 10^{-10} of the average value, which is well beyond the resolution of most measurement methods. This means that for all practical purposes the number of particles n is a constant.

Most probable macrostate. For the ideal gas system, we expect the system to cycle through all possible microstates: They are all equally probable. However, we also found that the initial state where all the atoms were on the left hand side did not appear spontaneously. Why? In

principle, all values of n are possible. In practice, there are most states near the average value of n. So many more states that if we move a few standard deviations away from the average, which is 10^{-11} of the average value, the probability to find such a state becomes negligible. This is the reason why we find the return to the intial state comic — it is extremely unlike to happen.

The development of the system away from $n = N$, where we started the simulation previously, towards $n = N/2$ is therefore effectively *irreversible*, because there are so many more states near $\bar{n} = N/2$ than near $n = N$ or $n = 0$. This also means that the system develops towards a macrostate where the number of particles per volume (the number density) is the same on each side (if the two sides are equally large). The system develops towards a homogeneous density.

Fluctuations in particle number. When we study the number of particles on the left side of the gas, we are addressing the variations, of the *fluctuations*, in the number of particles inside the system. Alternatively, we could say that there are fluctuations in the number of particles per unit volume, since the volume does not change, that is, there are fluctuations in the density of the gas.

These fluctuations are related to transport of particles: If the number of particles on the left side changes by Δn: $n(t + \Delta t) = n(t) - \Delta n$, the number on the right hand side increases by the same amount: The particles Δn are transported from the left hand side to the right hand side. This can occur because the systems are open to particle transport: There is no wall between the two halves.

Other fluctuations. Ok, we have a good description of the fluctuations of the number of particles, or the density, inside an isolated gas. However, we may be interested in variations in other macroscopic properties as well. For example. if the two halves of the system can transfer energy between each other, what are the fluctuations in energy in the two halves of the system? However, if we want to address fluctuations in energy, we would like to single out this fluctuation alone. We would therefore like to have a system where the number of particles on each side of the system is the same — identically the same — but where the total energy can be distributed in many ways between the two parts of the isolated system. This is the motivation to start working with other model system such as solids, where the particles do not move around significantly, but energy is free to flow throughout the system. We will therefore first address the fluctuations of energy in a simplified solid — in the Einstein crystal.

6.2 Behavior of a solid — Observations

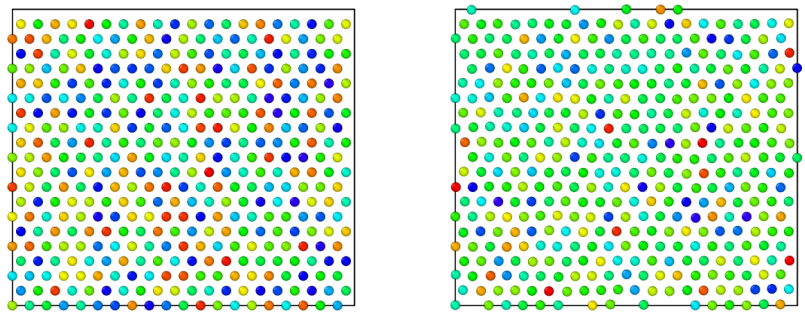


Fig. 6.2 Visualization of a solid of Argon atoms modelled using the Lennard-Jones potential at two different times. The colors indicate the velocities in the x-direction.

For a solid, the atoms are approximately fixed in space, but still vibrating. For this system energy is free to flow, but particles will remain in place. The solid is therefore well suited to study fluctuations and transport of energy.

Running a molecular dynamics simulation of a solid. Let us first start from a molecular model of a solid. We can address the behavior of a solid using the same model we used for the Argon system, but with smaller initial velocities and a smaller volume so that we ensure that the system is in its solid phase. You find reasonable input parameters for a simulation in in.solidstat02[1]:

```
# 2d Lennard-Jones solid
units       lj
dimension   2
atom_style  atomic
lattice     hex 1.05
region      box block 0 20 0 10 -0.1 0.1
create_box  1 box
create_atoms 1 box
mass        1 1.0
velocity    all create 0.5 87287
pair_style  lj/cut 2.5
pair_coeff  1 1 1.0 1.0 2.5
neighbor    0.3 bin
neigh_modify every 20 delay 0 check no
fix     1 all nve
dump 1 all custom 100 solidstat02.lammpstrj id type x y z vx vy vz
```

[1] http://folk.uio.no/malthe/fys2160/in.solidstat02

```
thermo      100
run         500000
```

You can run the simulation using

```
lmp_serial < in.solidstat02
```

The resulting atomic configuration is shown in Fig. 6.2. The atoms are organized in a regular lattice, a triangular lattice, which corresponds to the (equilibrium) crystal structure of the two-dimensional system. (It should here be noted that a two-dimensional system is somewhat unphysical, but it is simple to visualize. You can easily run all simulations in 3d without changing any of the general aspects of the results).

Measuring energy in the solid. Based on these simulations, we can measure the kinetic energy of the particles on the left side of the model. This does not represent the total energy for a crystal, since the potential energy part is non-negligible. But it does give a picture similar to what we had for the gas simulations. We read the data from the simulations into python, find all the particles on the left side, find the sum of the kinetic energies of all these particles, and plot the result using the script solidplotkt02.py[2]:

```
import numpy as np
import matplotlib.pyplot as plt
import re
import os
import lammps_logfile
traj = read("tmpdump.lammpstrj", format="lammps-dump-text", index=":")
nt = len(traj)
firstframe = traj[0]
cellshape = firstframe.get_cell()
vol = cellshape[0][0]*cellshape[1][1]
Ktot = [] # Kinetic energy
halfsize = 0.5*cellshape[0][0] # Box size in x-dir
# Loop over all timesteps
for frame in traj:
    x = frame.get_positions()
    v = frame.get_velocities()
    jj = np.where(x<halfsize)
    k = np.sum(0.5*(v[jj,0]*v[jj,0]+v[jj,1]*v[jj,1]))
    Ktot.append(k)
plt.plot(Ktot),xlabel('t'),ylabel('n(t)')
```

Fluctuations in the kinetic energy. The kinetic energy $K(t)$ as a function of time t is shown in Fig. 6.3. This plot looks similar what we saw

[2] http://folk.uio.no/malthe/fys2160/solidplotkt02.py

for $n(t)$ for the gas — there are fluctuations around a clear average value. We can characterize the size of the fluctuations by looking at the probability distribution $P(K)$ of the kinetic energy of the left-hand half of the system, which is also shown in Fig. 6.3. This plot suggests that the fluctuations also may by described by a Gaussian distribution. We will look more closely at this as we develop a theory for the behavior in the following, when we introduce the ideal crystal model for the behavior of a solid.

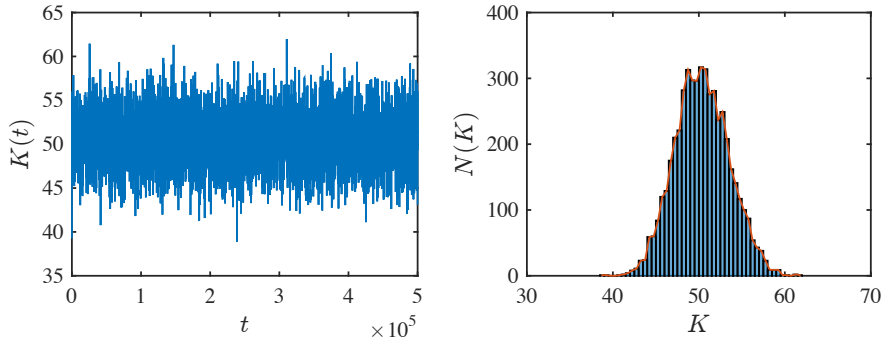
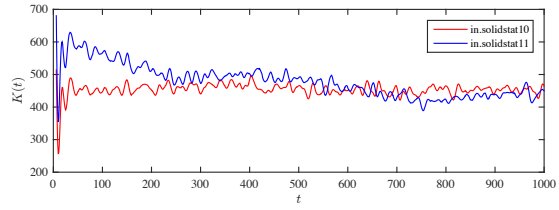


Fig. 6.3 *Left* Plot of the kinetic energy of the left-half of the Argon solid. *Right* Histogram of the kinetic energy K for the Argon system.

Fig. 6.4 Plot of $K(t)$ for added velocity (Black) and for the control simulation with no change (Red).



Towards equilibrium of the solid. We can also use this simulation framework to address how the system approaches equilibrium. We start a simulation where the left half of the system have a different average initial kinetic energy than the right half of the system. This is done by the input file in.solidstat10[3]:

```
# 2d Lennard-Jones gas
units       lj
dimension   2
```

[3] `http://folk.uio.no/malthe/fys2160/in.solidstat10`

```
atom_style  atomic
lattice     hex 1.05
region          box block 0 40 0 20 -0.5 0.5
create_box      1 box
create_atoms    1 box
region          2 block   0 20 0 20 -0.5 0.05
group           leftatoms region 2
mass            1 1.0
velocity        all create 0.5 87287
velocity        leftatoms create 1.5 282772
pair_style  lj/cut 2.5
pair_coeff  1 1 1.0 1.0 2.5
neighbor    0.3 bin
neigh_modify    every 20 delay 0 check no
fix     1 all nve
# Define output to file
dump    1 all custom 1 ss10.lammpstrj id type x y z vx vy vz
thermo  10
run 1000
```

The resulting plot of $K(t)$ is shown in Fig. 6.4. Again, we see that the system approaches a stationary state, corresponding to an equilibrium state, after a short transient. The system has a clear direction of time. What characterizes the equilibrium state? To answer this we need to develop a theory for the behavior of the solid.

6.3 Behavior of a solid — Theory

Let us now build a simplified model of the solid, inspired by the model for the ideal gas. Fig. 6.5 illustrates atomic arrangements in a solid, liquid and gas system. In the simulated solid, and in a real crystal where all the atoms are spaced regularly on a lattice, the atoms are fixed into a position on a lattice through the interactions with the other atoms in the crystal. The lattice configuration is a stable configuration as long as none of the atoms gain enough energy to break out of the local potential well they are situated in. A reasonable description of such a system would be that each atom vibrates around the minimum in an energy landscape, where the potential energy is due to the interactions with the other atoms. For simplicity, we can assume that the atoms do not directly interact, they only interact through their average potential field, but the potential does not vary as the atom vibrates. In this simplified picture, which is our theoretical approximation to the system, each atom in the lattice is affected by a local potential, and we can describe the vibrations around the minimum of this potential by the lowest-order approximation, as

illustrated in Fig. 6.5, where each atom is affected by a set of springs, and therefore acts as if it affected by a spring force. This corresponds to the most basic vibration system — the harmonic oscillator.

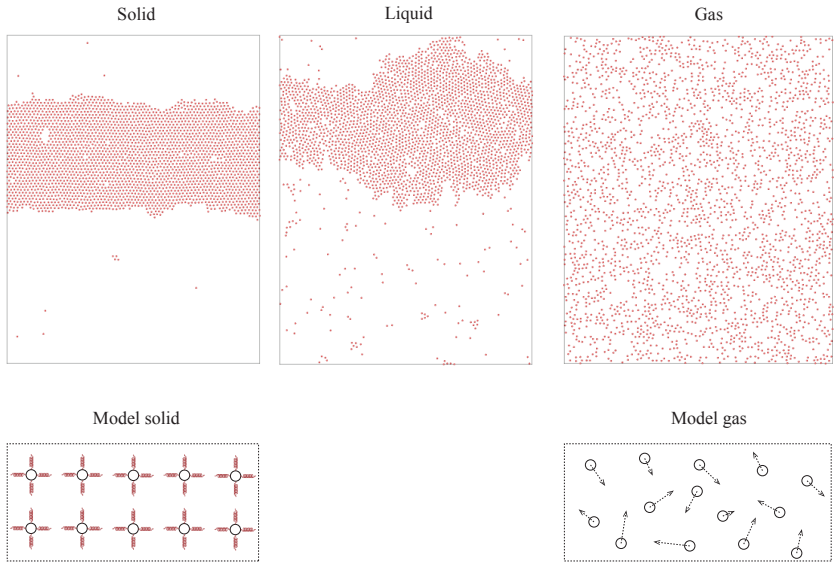


Fig. 6.5 Snapshots from simulations of a solid, a liquid, and a gas using the molecular dynamics model. Illustrations of the corresponding model systems.

6.3.1 Einstein crystal model

Our model for an ideal crystal therefore contains the following components:

- The atoms in the crystal do no interact directly
- Each atom acts as a harmonic oscillator in a three-dimensional potential
- The total system consist of N such atoms
- The atoms in the system share the total energy in the system

This model is called the **Ideal crystal** or the **Einstein crystal**. This model is going to be our basic model for a solid, and it is going to be equally versatile as the ideal gas description of a gas.

Harmonic oscillator system. For a three-dimension system, each atom consists of three independent oscillators in the x, y, and z directions so

that a system with N oscillators really only consists of $N/3$ atoms. Each harmonic oscillator is subject to a potential

$$V(x) = \frac{1}{2}kx^2 \, . \tag{6.7}$$

where k is the "spring constant" and x is the deviation from the equilibrium position.

Quantized harmonic oscillator. From quantum mechanics we know that the energies of such a system is quantized, with possible values

$$\epsilon = h\nu \, n \, . \tag{6.8}$$

where $n = 0, 1, 2, ..$ can only take integer values.

Energy in a system of harmonic oscillators. An Einstein crystal is a system consisting of N such independent oscillators, where each oscillator has the same ν and therefore the same energy levels. What are the possible states of such a system of N oscillators? It depends on the total energy E available to the system. We measure the energy in units of $h\nu = \epsilon$, and we often use the symbol q for this dimensionless energy. Notice that the number q can only take *integer values*.

Four-oscillator system. Let us address a simplified system with $N = 4$ and $q = 2$. What are the possible states of the whole system under these conditions? We describe the states using a simple illustration where we have drawn the possible energy levels for the $N = 4$ oscillators in Fig. 6.6. There are generally two possibilities: (case A) One oscillator may be in energy level 2, and the other in energy level 0; or (case B) two oscillators may be in energy level 1 and the others in energy level 0. But how many possible ways can this occur in? We can describe the possible configurations by a sequence (n_1, n_2, n_3, n_4) where $n_i = 0, 1, 2, \ldots$ describes the state of oscillator i. For case A, the oscillator in energy level 2 can be placed in $N = 4$ possible places: (2,0,0,0), (0,2,0,0), (0,0,2,0), and (0,0,0,2). There are therefore 4 possible states of the system with one oscillator in energy level 2 and the rest in energy level 0. For case B, the two oscillators in level 1 can be placed in 6 possible configurations. You may be tempted to use the following argument to count the number of possible configurations: We can choose the first oscillator from 4 positions and the second from 3 positions, hence the total number of possibilities are $4 \cdot 3 = 12$. However, this would be wrong, since we have then counted all the actual configurations twice. If we place the first energy unit in oscillator one and the second in oscillator two, we get the state (1,1,0,0),

but this is the same state we would get if we place the first energy unit in oscillator two and the second in oscillator one: (1,1,0,0). The total number of possible states are therefore only $4 \cdot 3/2 = 6$: (1,1,0,0), (1,0,0,1),(0,0,1,1),(1,0,1,0),(0,1,0,1), and (0,1,1,0).

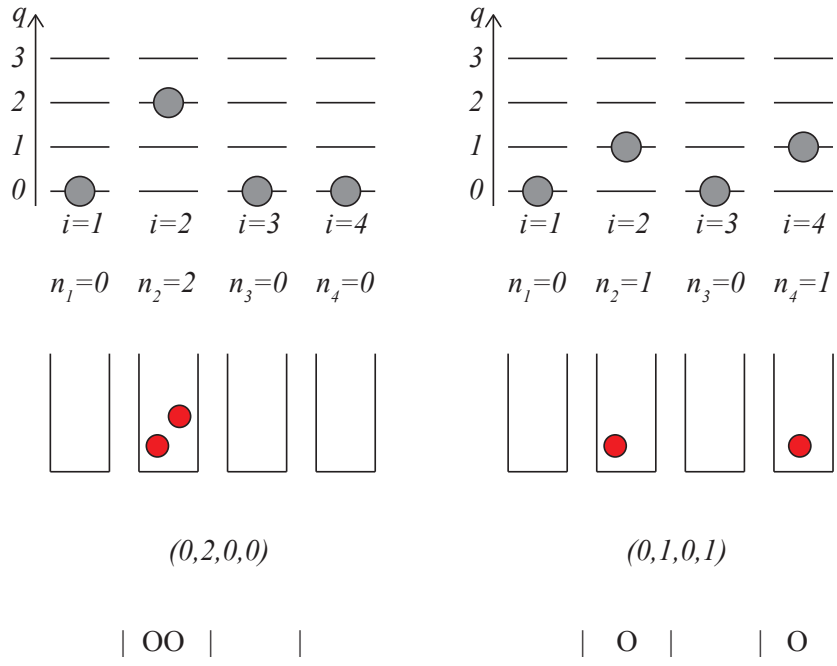


Fig. 6.6 Illustration of the Einstein crystal model.

Energy level diagram. It is customary to illustrate the states in an energy level diagram. Fig. 6.6 shows an illustration of the state (0,2,0,0) and (0,1,0,1). Another way to think about what states are possible is to think of q as the total number of energy units and N as the total number of boxes they can be placed in. The two states (0,2,0,0) and (0,1,0,1) is also illustrated in this way in Fig. 6.6 – by the number of energy units shown as balls in the boxes.

Microstates in the Einstein crystal. We call each sequence $(n_1, n_2, \ldots, n_N)$ a microstate of the Einstein crystal. The energy level or box representation in Fig. 6.6 therefore provides a basis for counting the number of possible microstates for a given set of oscillators N and a given total energy q of the system. However, there is an even better way to represent the states that makes counting simpler,

illustrated in the bottom of Fig. 6.6. We can represent the microstate by a sequence of energy units, 0, and walls, 1. If we only include the internal walls (the outer walls of the boxes cannot change position) we can list all possible microstates by listing all possible orderings of $N-1$ walls and q energy units. The two states (0,2,0,0) and (0,1,0,1) then get the following representation 10011 and 10110. Now, we have mapped the problem onto a problem we already know the answer to. (This is a very common method in physics — to find some smart way to map an unsolved problem onto a well-known and solved problem). We know how to count how many possible configurations we can have of q zeroes and $N-1$ ones: This corresponds to the number of ways we can place q zeros in $q+N-1$ positions. The number of microstates for the Einstein crystal with q energy units and N oscillators is therefore given as

$$\Omega(N,q) = \binom{N-1+q}{q} = \frac{(N-1+q)!}{q!(N-1)!}\,, \tag{6.9}$$

We call this the multiplicity of the Einstein crystal with N oscillators and energy q.

Question

Clickers: Einstein crystal 1

6.3.2 Two Einstein crystals in contact

Now we know how to count microstates. But with only a single Einstein crystal, the macrostate, the energy, q, in this case, is not that interesting. In the numerical simulations, we saw how the energy was fluctuating between two halves in a crystal. Can we apply the theory we have developed of the microstates in a model crystal to address this behavior?

Fig. 6.7 illustrates the kinetic energy in the molecular dynamics system. We divide the system into two parts, or two subsets of the oscillators, A and B. For example, we may choose part A to be the left half and part B as the right half as illustrated. We can now measure the energy, q_A, directly, and calculate the energy in system B, $q_B = q - q_A$. How can we make a model of the exchange between system A and B?

In our simulations, the two parts of the solids were connected. In the model, there are no direct interactions between the oscillators. However,

all the oscillators in a crystal share the total energy: If one oscillator in a crystal gets more energy, another oscillator gets less. Now, let us model systems A and B as two separate Einstein crystals with N_A and N_B oscillators and q_A and q_B units of energy in system A and B respectively. Given N_A, q_A we have found how to calculate the multiplicity of system A, and similarly for system B. However, we will also assume that the whole system (A and B together) is isolated so that the total energy, the total volume and the total number of particles is constant — this was also the underlying assumption for the molecular dynamics simulations.

Multiplicity of macrostate $q_A, q_B = q - q_A$. The total system therefore consists of two Einstein crystals sharing the total energy q so that $q = q_A + q_B$. The macrostate of this system is the energy q_A in system A. This is similar to the gas system, where the macrostate was the number of particles on the left side. Our plan is now to find the probability of a macrostate by counting the number of microstates for macrostate q_A and divide by the total number of microstates. But how can we find the multiplicity of macrostate q_A of the whole system?

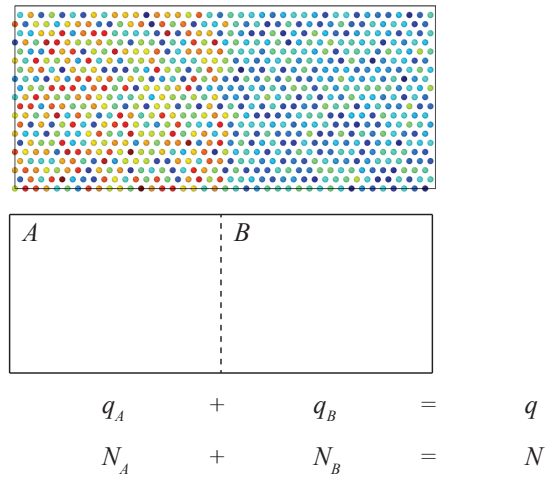


Fig. 6.7 Illustration of two parts A and B of an Einstein crystal.

First, we notice that of the system is in macrostate q_A for system A, it must be in macrostate $q_B = q - q_A$ in system B. It is therefore sufficient to provide q_A to describe the macrostate of the whole system. We already know how to find the number of microstates for Einstein crystals A and B respectively. What about the combined system? We can now use another assumption about the system: System A and system B are independent

of each other. This means that if I know the microstate of system A, which must be consistent with an energy q_A, I cannot predict what the microstate of system B is, except that it must be consistent with the energy q_B. This means that for each microstate in A, system B can be in any of its microstates. This means that the total number of microstates is the number in system A multiplied by the number in system B. If we call the multiplicity of system A, $\Omega_A(N_A, q_A)$, and the multiplicity of system B, $\Omega_B(N_B, q_B)$, then the multiplicity of the whole system in macrostate q_A is $\Omega = \Omega_A(N_A, q_A)\,\Omega_B(N_B, q_B)$.

Question

Clickers - Einstein crystal 2

We can calculate the multiplicities Ω_A, Ω_B and Ω for all the possible values of q_A numerically. For example, if $N_A = N_B = 5$, and $q = 10$, the the possible values of q_A are $0, 1, 2, \ldots, 10$. The multiplicities are calculated and plotted in the following script.

```
# Einstein solid NA - NB multiplicity
import numpy as np
import matplotlib.pyplot as plt
from scipy.special import comb
NA = 100
NB = 100
q = 200
N = NA + NB
# Find multiplicity of all macrostates
omegaA = np.zeros(q+1)
qAvalue = np.zeros(q+1)
omegaB = np.zeros(q+1)
omegaTOT = np.zeros(q+1)
# Loop through all macrostates and find multiplicity
for istate in range(0,q):
    qA = istate
    qAvalue[istate] = qA
    omegaA[istate] = comb(qA+NA-1,qA)
    qB = q - qA
    omegaB[istate] = comb(qB+NB-1,qB);
    omegaTOT[istate] = omegaA[istate]*omegaB[istate];
plt.plot(qAvalue,omegaTOT,'-o')
plt.xlabel('q'),plt.ylabel('W')
```

The resulting number of microstates are show in Fig. 6.8 and in the following table. We notice that the the multiplicity $\Omega = \Omega_A\,\Omega_B$ has a

maximum for $q_A = q/2 = 5$. But how can we use the *multiplicities* to find the *probability* of a macrostate?

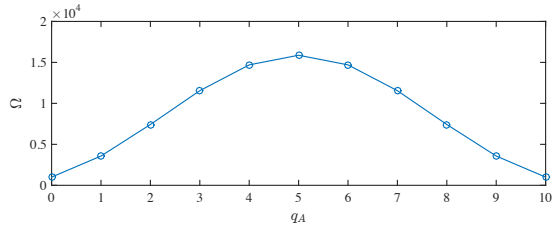


Fig. 6.8 Plot of Ω_{tot} as a function of q_A for an two Einstein crystals A and B in contact with $N_A = N_B = 5$ and $q = q_A + q_B = 10$.

q_A	q_B	Ω_A	Ω_B	Ω_{tot}
0	10	1	1001	1001
1	9	5	715	3575
2	8	15	495	7425
3	7	35	330	11550
4	6	70	210	14700
5	5	126	126	15876
6	4	210	70	14700
7	3	330	35	11550
8	2	495	15	7425
9	1	715	5	3575
10	0	1001	1	1001

Fundamental assumption of statistical mechanics. To find the probability of a macrostate, we need more information about the microstates. Are the microstates equally probable, as we assumed for a dice, or are there some microstates that are more probable than others? We need another assumption about the microstates — we need the *fundamental assumption of statistical mechanics*:

> **Fundamental assumption of statistical mechanics**: In an isolated system in thermal equilibrium, all accessible microstates are equally probable.

We cannot prove this assumption here, but we can understand where it comes from: At the microscopic level we expect any process that takes the system from state X to Y to be reversible, so that the probability for the reverse process is the same. This assumption, or aspect of physics, is called **the principle of detailed balance**. Another idea is that the system, as it is evolving in time like in our simulation, will span

all possible microstates if you just wait long enough. For now, we will simply assume that the fundamental assumption is correct, and that all microstates have the same probability.

Probabilities of the macrostate $q_A, q_B = q - q_A$**.** Now we can find the probability of a macrostate by simply summing the probabilities of all the microstates in the macrostates. How many microstates are there in a macrostate? This is the multiplicity $\Omega_A\,\Omega_B$ of the macrostate. What is the probability of a microstate? It is $p = 1/\Omega$, where Ω is the total number of microstates. And what is the total number of microstates? This is the number of microstates for each possible value of q_A, which also corresponds to the number of microstates in the whole system, which the multiplicity of an Einstein crystal with N oscillators and q units of energy:

$$\Omega = \sum_{q_A=0}^{q} \Omega_A(N_A, q_A)\,\Omega_B(N_B, \underbrace{q_B}_{q-q_A}) = \Omega(N, q)\,. \tag{6.10}$$

The probability for macrostate q_A is therefore

$$P(q_A) = \frac{\Omega_A(N_A, q_A)\,\Omega_B(N_B, q - q_A)}{\sum_{q_A=0}^{q} \Omega_A(N_A, q_A)\,\Omega_B(N_B, q - q_A)}\,. \tag{6.11}$$

The probability of macrostate q_A is proportional to $\Omega_A\,\Omega_B$, we can therefore also interpret the plot of $\Omega_A\,\Omega_B$ in Fig. 6.8 as a plot of the probabilities. We observe that the macrostates are not equally probable. For example, the $q_A = 0$ macrostate is clearly much more unlikely than $q_A/N_A = q_B/N_B$. How can we use such a result to interpret the development of an Einstein crystal with time? Let us address this by introducing a simplified model for the dynamics of an Einstein model, and by a theoretical argument based on the probabilities of macrostates

6.3.3 Time development of the Einstein crystal — Monte Carlo model

While the molecular dynamics simulation of a crystal lattice, illustrated in Fig. 6.3, shows the time development of the system, the Einstein crystal does not include any description of how the system develops in time. The Einstein crystal only describes the system in terms of the microstate, $(n_1, n_2, \ldots, n_N)$, of the system. If the system is in one particular microstate, we do not have any physical laws that tells us how the system can develop to another microstate. To include dynamics, we

need to add further assumptions to the model. When we defined the Einstein crystal we assumed that there were no interactions between the individual oscillators, whereas for the molecular dynamics simulations, the atoms had both specific positions in the crystal and they interacted with their neighbors. How could we extend the Einstein crystal model to include both simplified interactions and the relative positions of the atoms, so that we can model energy flow without including the full dynamics.

Global dynamics in the Einstein crystal. We can make a small modification to the Einstein crystal model to include some of these effects. The system consists of two parts, A and B, so that each atom/oscillator belongs to either A or B. As a first approximation, we introduce simplified interaction between the oscillators, by allowing energy to be randomly moved from one oscillator to another oscillator, while conserving the total energy. This means that we go from one microstate s_1 to another microstate s_2 with the same energy. Hence both these microstates are possible microstates for the system, and they are both equally probable, since all microstates are equally probable.

We introduce the following algorithm to imitate the random transmission of energy in the system:

- Select an oscillator (particle) at random, i_1. Let us attempt energy transport from this oscillator.
- Select another oscillator (particle) at random, i_2. This oscillator may receive energy from n_1.
- We transfer one unit of energy from i_1 to i_2, if oscillator i_1 has at least one unit of energy.

We can implement this algorithm using the following steps: (i) Generate an initial state of the Einstein crystal with energy q_A in part A and energy q_B in part B, (ii) Perform one step of the algorithm, (iii) Measure the new energy q_A (and $q_B = q - q_A$), (iv) Continue with step (ii). We generate the initial state by placing each of the q_A energy units in part A at random oscillators in part A and similarly for part B. This is implemented in the following program, which also plots both the state and the energy q_A of the system:

```
%matplotlib notebook
import numpy as np
import matplotlib.pyplot as plt
NA = 100 # Nr of oscillators
NB = 100
```

```
qA = 300 # Initial energy
qB =   0
q = qA + qB # Total energy
N = NA + NB # Total oscillators
state = np.zeros(N,float)  # Microstate of the system
# state[0:NA-1] is part A, state[NA:NA+NB-1] is part B
# Generate initial, random state
placeA = np.random.randint(0,NA,(qA,1))
for ip in range(len(placeA)):
    i = placeA[ip]
    state[i] = state[i] + 1
placeB = np.random.randint(0,NB,(qB,1))
for ip in range(len(placeB)):
    i = placeB[ip]
    state[i] = state[i] + 1
# Simulate state development
nstep = 100000
EA = np.zeros(nstep,float)
EB = np.zeros(nstep,float)
# Setup dynamic plotting
fig = plt.figure()
ax1 = fig.add_subplot(211)
ax2 = fig.add_subplot(212)
for istep in range(nstep):
    i1 = np.random.randint(0,N) # Select rand osc. i1
    if (state[i1]>0): # Does it have any energy ?
        i2 = np.random.randint(0,N) # Select rand osc. i2
        state[i2] = state[i2] + 1 # Transfer
        state[i1] = state[i1] - 1 #  energy
    # Output and display result dynamically
    ax1.clear()
    ax1.plot(np.arange(0,NA-1),state[0:NA-1],'b',np.arange(0,NB-1)+NA,state[NA:NA+NB-1],'r'
    EA[istep] = sum(state[0:NA-1])/NA
    EB[istep] = q - EA[istep]
    ax2.clear()
    ax2.plot(np.arange(0,istep),EA[0:istep],'-r',np.arange(0,istep),EB[0:istep],'-b');
    fig.canvas.draw()
```

Fig. 6.9 shows the resulting dynamics. In this case all the energy starts on the left hand side and is gradually transferred to an uniform distribution — as expected.

Local dynamics in the Einstein crystal. We could add another level of complexity to the model, by introducing local interactions. This is demonstrated in Sect. 6.10.

Sampling the space of all microstates. The model we have introduced here is not a fundamental model for the dynamics of the model. We have introduced an algorithmic, stochastic model for the process, but not a physical model. We cannot therefore interpret the progression of such a model directly as time. However, the model does move through

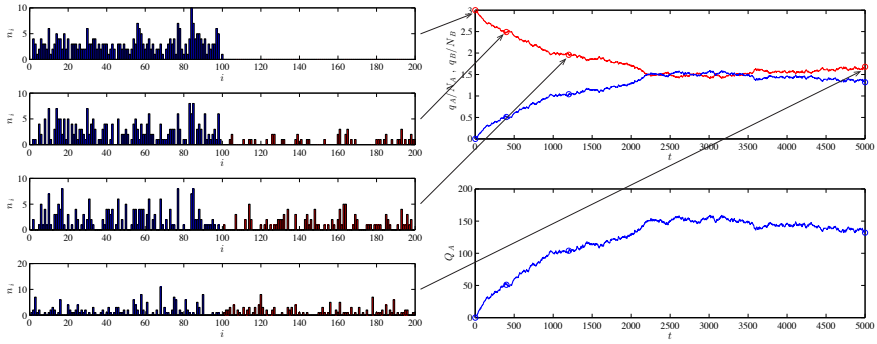


Fig. 6.9 Illustration of the time development of two Einstein crystals in thermal contact.

a sequence of microstates, all of which are equally probably. Therefore, if we study a long enough sequence we would expect the sequence of microstates to sample all possible microstates. We get some insight into whether we are far from equilibrium just be studying the development of q_A. From Fig. 6.9 we see that we initially are far from equilibrium, and hence only have seen non-representative microstates, but after many steps, the system approaches an equilibrium where q_A appears to have forgotten the details of the initial state. We will discuss such models and how to sample the space of all possible states in Sect. 6.10.

Approaching the stationary state. Both of these models show that the Einstein crystal will approach a stationary state with a uniform distribution of energy per oscillator: $q_A/N_A = q_B/N_B$. This corresponds to the macrostate with the maximum multiplicity. Why is that? How is this observation of the dynamics explained by the probabilities of the macrostates?

6.3.4 Time development of the Einstein crystal — Theoretical argument

How can we interpret the time development of the Einstein crystal using the probabilities of a macrostate in *equilibrium*? The probabilities we have found an expression for are the probabilities of a macrostate when the system have reached equilibrium. If the system starts with all the energy in system A, $q_A = q$, $q_B = 0$, we know that this state is very unlikely in equilibrium. If we just leave the system to itself, it seems reasonable that it will develop from one microstate to another,

exploring the space of all possible microstates. (This is actually another fundamental hypothesis called the ergodicity hypothesis). If we then observe the system after some time, when it has forgotten about its initial microstate, it will be in some random microstate. But since there are so many more microstates in the most probable macrostates than in the initial macrostate, it is much more likely that the system after a while will be in the most probable macrostates. And it will be very unlikely to find the system in the initial macrostate. The behavior therefore corresponds to an *irreversible* processes. Not because it is impossible to go back to the initial macrostate, but just because it is very unlikely to happen.

This also means that if the system starts with a much higher energy per oscillator in system A than in system B, $q_A/N_A \gg q_B/N_B$, energy will flow spontaneously from A to B, but not the other way around. (Except for small fluctuations). This is an explanation of *heat* — the flow of thermal energy from one side to another.

We have here essentially found the second law of thermodynamics: the law of increased multiplicity. The system develops towards a state of maximum (or increased) multiplicity. Simply because this is the most probable behavior. And as the system size increases it becomes increasingly more probable to be in the most probable macrostate and the probability for all other states goes to zero. But to gain insights into the probability of a macrostate for a large system, we need an improved expression for the multiplicity of macrostate q_A, an expression we can analyze mathematically.

6.3.5 Approximate multiplicities using Stirling's formula

Fig. 6.10 show the multiplicities of the macrostates q_A for two systems with $N_A = N_B = 10, q = 200$, and $N_A = N_B = 50, q = 1000$. In both cases, the maximum of the multiplicity occurs at $q_A = q_B = q/2$, but we see that for the system with larger N (and larger q) the multiplicity is sharper relative to the position of the maximum. But how sharp is it and how does the function look like around its peak? The multiplicity of macrostate q_A of the two-part Einstein crystal is $\Omega(q_A) = \Omega_A(q_A)\,\Omega_B(q_B)$ and the probability of the macrostate is $P(q_A) = \Omega(q_A)/\Omega_{TOT}$, where the multiplicity $\Omega_A(q_A) = (q_A + N_A - 1)!/(q_A!\,(N_A - 1)!)$. Unfortunately, it is difficult to discuss this expression because of the factorials. Instead, we would like an approximative formula which we can address using the

tool-box of calculus. How can we find this approximative formula? We will do this in two stages, first we will find an approximative expression for the mulitiplicity using Stirling's approximation, and then we will find an approximative expression for the probability of a macrostate.

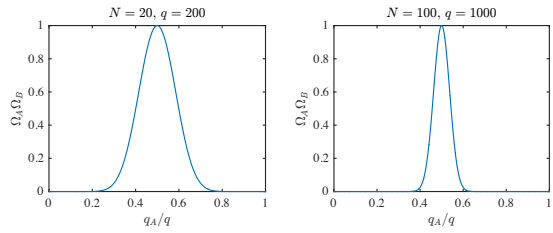


Fig. 6.10 Plot of $\Omega_A(q_A; q, N)\Omega_B(q_B = q - q_A; q, N)$ as a function of q_A for **a** $N_A = N_B = 10, q = 200$ **b** $N_A = N_B = 50, q = 1000$. Notice that the multiplicity becomes sharper for larger systems.

Theoretical argument: Approximation for the multiplicity. The multiplicity of an Einstein crystal is

$$\Omega(N, q) = \frac{(q + N - 1)!}{q!\,(N - 1)!}\ , \tag{6.12}$$

How can we approximate this expression? First, we employ the common trick of working with the logarithm of the multiplicity. Life becomes simpler when all the multiplications then become additions.

$$\ln \Omega(N, q) = \ln(q + N - 1)! - \ln q! - \ln(N - 1)!\ , \tag{6.13}$$

To simplify the factorials, we introduce Stirling's approximation for the factorial, which is good when the argument of the factorial is large:

$$\ln x! \simeq x \ln x - x + \frac{1}{2} \ln 2\pi x\ , \tag{6.14}$$

which for large values of x can be further approximated to

$$\ln x! \simeq x \ln x - x\ . \tag{6.15}$$

This is a good starting point to simplify the multiplicity. We can apply Stirling's approximation to each of the factorials in (6.12) if each of the factorials are large. That is, if q and N are large. In the limit of (very) large q and N we can also assume that $N - 1 \simeq N$. With these approximations we get:

$$\ln \Omega(N,q) = \ln(q+N-1)! - \ln q! - \ln(N-1)! \tag{6.16}$$
$$\simeq \ln(q+N)! - \ln q! - \ln N! \tag{6.17}$$
$$\simeq (q+N)\ln(q+N) - (q+N) - q\ln q + q - N\ln N + N \tag{6.18}$$
$$= (q+N)\ln(q+N) - q\ln q - N\ln N\ , \tag{6.19}$$

Now we assume that the number of units of energy is very large compared to the number of oscillators: $q \gg N$, which corresponds to assuming that we are in the classical limit where the quantum resolution into discrete energy levels are not important. We can then simplify $\ln(q+N)$ through:

$$\ln(q+N) = \ln\left[q\left(1+\frac{N}{q}\right)\right] = \ln q + \ln\left(1+\frac{N}{q}\right) \simeq \ln q\ , \tag{6.20}$$

where we have used the approximation $\ln(1+x) \simeq 0$, which is valid when $x \ll 1$. Plugging this back into the relation for $\ln \Omega$ gets rid of the $q\ln q$ terms, giving

$$\ln \Omega \simeq (q+N)\ln q - N\ln N - q\ln q \tag{6.21}$$
$$= q\ln q + N\ln q - N\ln N - q\ln q = N\ln\frac{q}{N}\ . \tag{6.22}$$

The multiplicity is then

$$\Omega \simeq e^{N\ln(q/N)} = \left(\frac{q}{N}\right)^N \ (q \gg N)\ , \tag{6.23}$$

This is the multiplicity of the Einstein crystal. We will see further on that this is a very useful expression since we can use it to calculate macroscopic properties of the Einstein crystal. Now, we will use this formula to show that the multiplicity function is indeed very sharp.

Sharpness of the multiplicity function. Armed with this appoximation for the multiplicity, we are ready to find an approximate expression for the multiplicity and probability of a macrostate for an Einstein crystal divided into two parts, A and B, that are in thermal contact. The multiplicity for the two-part system is

$$\Omega = \Omega_A(N_A,q_A)\,\Omega_B(N_B,q_B) = \left(\frac{q_A}{N_A}\right)^{N_A}\left(\frac{q_B}{N_B}\right)^{N_B} \tag{6.24}$$

For simplicity, let us address the case where $N_A = N_B$. The multiplicity is then

$$\Omega_A\,\Omega_B = \left(\frac{q_A}{N}\right)^N \left(\frac{q}{N}\right)^N = \left(\frac{q_A\,q_B}{N^2}\right)^N \,. \tag{6.25}$$

We saw in Fig. 6.10 that the multiplicity is sharp around $q_A = q/2$, we therefore express both q_A and q_B as small deviations from $q/2$ by introducing

$$q_A = \frac{q}{2} + x \;,\; q_B = \frac{q}{2} - x \;, \tag{6.26}$$

where x is a number much smaller than q (but it is still rather large). We insert this expression back into the multiplicity, getting

$$\Omega = \Omega_A\,\Omega_B = \left(\frac{q_A\,q_B}{N^2}\right)^N = N^{-2N}\left[\left(\frac{q}{2}+x\right)\left(\frac{q}{2}-x\right)\right]^N \tag{6.27}$$

$$= N^{-2N}\left[\left(\frac{q}{2}\right)^2 - x^2\right]^N \,. \tag{6.28}$$

Again, it is easier to work with the logarithm of this expression

$$\ln\Omega = \ln N^{-2N} + \ln\left[\left(\frac{q}{2}\right)^2 - x^2\right]^N \,. \tag{6.29}$$

We simplify the second part of this expression by pulling $(q/2)^2$ outside the paranthesis

$$\ln\left[\left(\frac{q}{2}\right)^2 - x^2\right]^N = N\ln\left[\left(\frac{q}{2}\right)^2\left(1-\left(\frac{2x}{q}\right)^2\right)\right] \tag{6.30}$$

$$= N\ln\left(\frac{q}{2}\right)^2 + N\ln\left(1-\left(\frac{2x}{q}\right)^2\right) \,. \tag{6.31}$$

Where $2x/q \ll 1$ and we can therefore use the approximation $\ln(1+u) \simeq u$ (when $u \ll 1$):

$$\ln\left[\left(\frac{q}{2}\right)^2 - x^2\right]^N \simeq N\ln\left(\frac{q}{2}\right)^2 - N\left(\frac{2x}{q}\right)^2 \,. \tag{6.32}$$

The total multiplicity is therefore

$$\Omega = N^{-2N}\,e^{N\ln(q/2)^2}e^{-N(2x/q)^2} = \Omega_{max}\,e^{-N(2x/q)^2} = \Omega_{max}\,e^{-\left(\frac{x}{q/(2\sqrt{N})}\right)^2} \,, \tag{6.33}$$

This is the multiplicity for the macrostate $q_A = q/2 + x$, and the probability for this macrostate is simply the multiplicity multiplied by a constant. The probability $P(q_A)$ is therefore a similar function

$$P(q_A) = \Omega_A\, \Omega_B/\Omega_{TOT} = P_{max}\, e^{-\left(\frac{x}{q/(2\sqrt{N})}\right)^2}\,, \tag{6.34}$$

We have found that the multiplicity and the probability is a Gaussian with a peak at $x = 0$, that is for $q_A = q/2$. This function falls off rapidly for both positive and negative x. How rapidly? The multiplicity has fallen to a value of $1/e$ when $N(2x/q)^2 = 1$, that is, when $x = q/(2\sqrt{N})$. It might appear that this value of x is a large number, since q is a large number. However, for realistic systems $N = 10^{22}$, which means that the multiplicity has fallen to $1/e$ after a deviation $x = q/(2\,10^{11})$, which is very small compared with the value of q. If the full graph, that is with the full range of possible q_A values from 0 to q, spanned from here to the Moon, the multiplicity of a system with would have fallen to $1/e$ over a distance of $3.86\,10^8\,\mathrm{m}/(2\,10^{11}) = 19\mathrm{mm}$. This also means that it is very unlikely to observe a value of q_A which is far away from $q/2$. The relative variations of q_A are extremely small in realistic systems — in practice these fluctuations are not measureable at all.

We call the limit when there are no measureable fluctuations away from the most likely macrostate **the thermodynamic limit**.

Comparison of exact and approximate results. We can test this approximation by comparing it with directly calculated values of the multiplicity. Fig. 6.11a shows plots of the multiplicity $\Omega(q_A; N, q)/\Omega_{\max}$ for $N = 10, 20, 40, 80$ and $q = 10\,N$. Fig, 6.11a shows direct plots of the multiplicity rescaled by its maximum value. Otherwise the systems with larger N and q would completely dominate the plot. The multiplicities have a clear maximum value and decay approximately symmetrically from each side of the maximum. The q_A value that gives maximum multiplicity increases as we increase N and q. We expect the maximum to be at $q_A = q/2$. Indeed, in Fig. 6.11b we see that if we plot $\Omega_{\max}$ as a function of $q_A - q/2$, the distributions are centered around zero. However, the widths of the distributions are also changing when we change q and N. How wide is the distrubution compared to the average or typical value of q_A? This is shown in Fig. 6.11, which shows $\Omega_{\max}$ as a function of $(q_A - q/2)/(q/2)$. Here, it is clear that the multiplicity becomes more and more narrowly distributed as N and q increases — the multiplicity

is becoming sharper, and deviations from the average value for q_A are becoming less and less likely.

How good is the approximation we found in (6.33)? First, we replot the multiplicities according to the theory in (6.33): We plot $\Omega/\Omega_{\max}$ as a function of $(q_A - (q/2))/(2\sqrt{N})$ in Fig. 6.11d. Yes! All the curves now fall onto the same curve, corresponding to the Gaussian form in (6.33). We call such a plot a data-collapse, and we often use such plots to demonstrate/validate our theories. Now, we can even compare with the Gaussian curve, we plot $\exp\left((q_A - q/2)/(2\sqrt{N})\right)$ in the same plot in Fig. 6.11d, and indeed the theoretical curve fits nicely with the observed curve — the approximation we have developed seems sound.

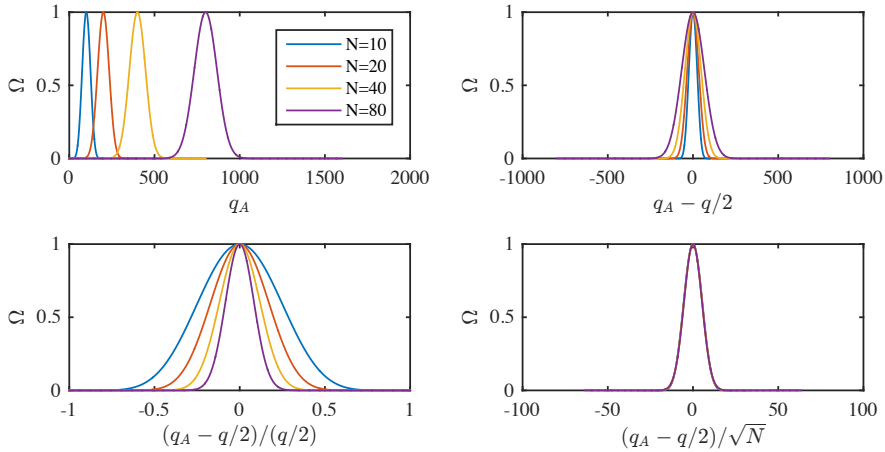


Fig. 6.11 Plot of $\Omega(q_A; q, N)/\Omega_{\max}$ as a function of **a** q_A, **b** $q_A - q/2$, **c** $(q_A - q/2)/(q/2)$, **d** $(q_A - q/2)/(2\sqrt{N})$.

6.4 The ideal gas — Theory

The main observations and conclusions we have made so far are valid for most systems in thermal contact or for fluctuations in parts within the system. The multiplicity function will typically be very sharp for large systems, meaning that the system will only have a reasonable probability to be in a very small fraction of the macrostates. Here, we will demonstrate in detail that this is also the case for an ideal gas — a simplified model of a gas.

Ideal gas. An ideal gas consists of N atoms (or molecules) in a box of size $L \times L \times$ as illustrated in Fig. 6.12. We assume that if the gas is thin, that is if the particles typically are far from each other, they do not interact much, except when they collide. As a first approximation, we will assume that the particles do not interact at all. This is our gas model, called an ideal gas, consisting of a set of particles moving inside a box without interactions. However, we do assume an implicit interaction — the atoms may exchange energy, but in such a way that the total energy of the system is conserved. In addition to this, we will encounter a two quantum mechanical features. First, if the particles are Fermions, then two particles cannot be in the same energy state. We will see what consequences this has further on. Second, the particle cannot be discerned, and therefore we cannot discern two states where we only have exchanged two atoms. This will have consequences for how many states there are for the system.

The **ideal gas** model consists of the following assumptions

- The gas consists of N particles
- The particles are confined to an $L \times L \times L$ box of volume V
- The particles do not interact directly
- The total energy of the gas is E, but the distribution of energy within the gas is not prescribed
- The particles are identical and cannot be discern from one another

To describe the ideal gas, we will start from a single particle, then discuss two particles, before finally moving to a system of N particles. Our goal is to find the multiplicity of the ideal gas and to show that the multiplicity function indeed is sharp also for the ideal gas.

6.4.1 Microstates of a single particle

We start by addressing the multiplicity of a single particle in a three-dimensional box. What are the possible states of a single particle? From quantum mechanics we know that the possible states of a particle in a $L \times L \times L$ box are quantized with energy levels

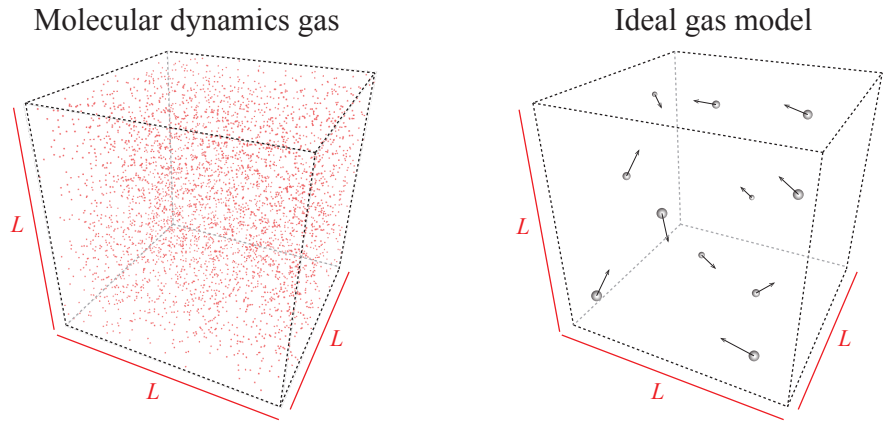


Fig. 6.12 **(Left)** Illustration of a molecular dynamics simulation of a thin gas. **Right** Illustration of the ideal gas model consisting of N non-interacting, identical particles in a $L \times L \times L$ box.

$$\epsilon_(n_x, n_z, n_z) = \frac{h^2}{8mL^2}\mathbf{n}\cdot\mathbf{n} = \frac{h^2}{8mL^2}\left(n_x^2 + n_y^2 + n_z^2\right) \ , \tag{6.35}$$

where m is the mass of the particle, L is the length of the box and $\mathbf{n} = (n_x, n_z, n_z)$ is related to the momentum

$$\mathbf{p} = \frac{h}{2L}\mathbf{n} \ . \tag{6.36}$$

Each of the numbers n_x, n_y, n_z represents state for the motion in the x, y, and z-direction respectively, and they are all positive integers. We can illustrate the possible states in a three-dimensional space with n_x, n_y, n_z along the axes as illustrated in Fig. 6.13. Each point such as $(1, 1, 1)$ or $(2, 3, 2)$ represent a state for system. In the $\mathbf{n}$ space there is a constant *density of states* corresponding to one state per unit volume. (Each $1 \times 1 \times 1$ volume has one state). We also see that because the energy depends on n^2, $E = (h^2/8mL^2)\, n^2$, all states that have the same distance to the origin in n-space have the same energy.

How can we use this to count the number of states in an ideal gas with a given total energy E, but with only one particle, $N = 1$? In this case, all the states with the same energy is on a the surface of sphere with radius n in $\mathbf{n}$-space. What is this radius? It is found from the relation between n and E for the system consisting of a single particle in three dimensions:

$$E = \frac{h^2}{8mL^2}n^2 \ \Rightarrow \ n = \frac{2L}{h}\sqrt{2mE} \tag{6.37}$$

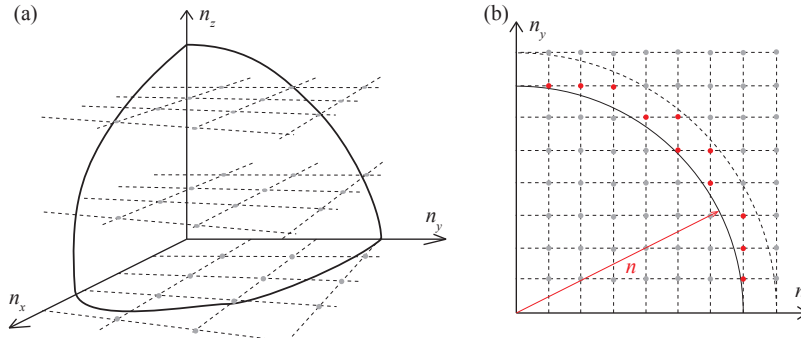


Fig. 6.13 Illustration of the states of a single particle in a three-dimensional box.

How can we count the corresponding number of states with this energy? We need to find the surface area of the corresponding sphere in **n**-space, remembering to include only 1/8 of this surface since we only address positive n-values. This is simple for the 3d system. The surface area is $4\pi n^2$, so that the number of states is:

$$\Omega_1 = \frac{1}{8}4\pi n^2 = \frac{1}{8}4\pi\frac{8mL^2}{h^2}E = \frac{4\pi L^2}{h^2}E = \frac{4\pi}{h^2}v^{2/3}E\ . \tag{6.38}$$

We notice that this result is a function of of E and V as it should, and a function of N, but $N = 1$ here. How can we generalize this result to more particles?

6.4.2 Microstates of two non-interacting particles

What happens if we have more than one particle, for example, two particles? Let us address this in the simplest case — a one-dimensional system. In this case each particle is described by a single quantum state n_i, n_1 for particle 1 and n_2 for particle 2, so that the state of the system can be written as the pair (n_1, n_2). When the total energy of the system is E, the possible states of the system is therefore all the states so that $E = h^2/(8mL^2)\,(n_1^2+n_2^2) = E_0\,(n_1^2+n_2^2)$, and we call this the multiplicity $\Omega_2(E)$ of the two-particle system. This means that for a total energy of $E = 4E_0$ there are two possible states, $(2,0)$, and $(0,2)$, and therefore $\Omega_2(4E_0) = 2$. Or is it? We have overlooked an important quantum mechanical effect!

Indistinguishable particles. This result is true of the particles are *distinguishable* from each other. They are distinguishable if we can place

a label on each of them so that we can see that particle 1 is in state 2 whereas particle 2 is in state 0. However, for a gas of identical particles (atoms, molecules), the particles are *indistinguishable*. We cannot discern the state $(2, 0)$ from the state $(0, 2)$ because the molecules are indentical. This means that we have counted too many states! We should not count all the possible ways that we can interchange the two particles. In general we know that for N particles, there are $N!$ possible ways to interchange them. For two particles, there are $2! = 2$ ways to interchange them. The correct multiplicity is therefore $\Omega_2(4E_0) = 2/2! = 1$.

Fermions or Bosons. In addition, we should also reflect on whether the particles are Fermions or Bosons. Why? Because if the particles are Fermions, there are additional restrictions. Two particles cannot be in the same state. This means that the state $(4, 4)$ is not a possible state for the two-particle system. We will return to this case later. For now, we will simply assume that we are in the limit when the energy is large, so that there are many more possible states available than there are particles, so that the number of restricted states for Fermions is negligible. These considerations will be more explicit in Chap. 7.

6.4.3 Microstates of N non-interacting particles

Distinguishable particles. For N independent, *distinguishable* particles all in a box of size $V = L \times L \times L$ the energy of the system is given as:

$$E = \frac{h^2}{8mL^2} \sum_{i=1} N(n_{xi}^2 + n_{yi}^2 + n_{zi}^2) = \frac{h^2}{8mL^2} \sum_{i=1}^{3N} n_i^2 \,. \tag{6.39}$$

Here, the sum is over $3N$ values of n_i. This means that we write N particles with 3 degrees of freedom as $3N$ particles with 1 degree of freedom. This expression reproduces the result for $N = 1$. We see that the energy E depends on the radius n in the $3N$-dimensional $\mathbf{n} - space$:

$$E = (h^2/(8mL^2))\, n^2 \;\Rightarrow\; n = (2L/h)\,\sqrt{2mE} \tag{6.40}$$

We can use the same approach as for a single particle two find the number of states with this energy. The density of states in the $3N$-dimensional $\mathbf{n}$-space is still 1: There is one state for each $1 \times 1 \times \ldots \times 1$ ($3N$ times) volume.

[4] We can therefore estimate the number of states by estimating the "area" of a sphere with radius n in $3N$ dimesions. (We must also remember to only include positive values for all the n_i in $\mathbf{n} = (n_1, n_2, n_3, \ldots, n_{3N})$).

The general formula for the surface area of a d-dimensional sphere is:

$$A = \frac{2\pi^{d/2}}{\left(\frac{d}{2}-1\right)!} r^{d-1} \,. \tag{6.41}$$

We use this expression with $d = 3N$. In addition, we also need to ensure that only the n_i-values in the first "octant" ($3N$-tant) are used, that is we must divide by the number of "quadrants" $= 2^d = 2^{3N}$:

$$\Omega_{\text{disting}} = \frac{1}{2^{3N}} \frac{2\pi^{3N/2}}{\left(\frac{3N}{2}-1\right)!} \left(\frac{2L}{h}(2mE)^{1/2}\right)^{3N-1} \,, \tag{6.42}$$

Indistinguishable particles. This the number of states of a system of N distinguishable particles. However, an ideal gas consists of N *indistinguishable* particles? How should the multiplicity be modified? Based on our argument for two-particles above, we realize that we have counted too many states. We must therefore divide by the number of ways we can rearrange the N particles between the states, which is $N!$. The multiplicity for a gas of indistinguishable particles is therefore

$$\Omega = \Omega_{\text{indisting}} = \frac{1}{N!}\Omega_{\text{disting}} = \frac{1}{N!}\frac{1}{2^{3N}} \frac{2\pi^{3N/2}}{\left(\frac{3N}{2}-1\right)!} \left(\frac{2L}{h}(2mE)^{1/2}\right)^{3N-1} \,. \tag{6.43}$$

We can simplify this expression by using that for very large N, $3N - 1 \simeq 3N$. However, in the first few examples, we will only be interested in the V and E depedence of this expression, which is simple:

$$\Omega = f(N) V^N E^{3N/2} \,. \tag{6.44}$$

6.4.4 Interaction between two ideal gases

We have now found the multiplicity for an ideal gas of N atoms, volume V, and energy E. We can use this expression to address the macrostate

[4] This is true if we do not include spin. For particles with spin 1/2 there will be two states per such volume, one state with spin up and one state with spin down. We will return to this effect in Chap. 10.

of two ideal gases in equilibrium, or equivalently, to find the distribution of energy inside an ideal gas.

Just like we did for the ideal crystal, we can divide the system into two parts A and B as illustrated in Fig. 6.14. The whole system is isolated, but the two parts are not thermally insulated. This means that the sum of the energies in A and B is constant, $E_A + E_B = E$. We assume that the two parts cannot change shape and that there is no transport of particles from one part to the other. Thus, the number of particles and the volumes of each of the two parts are constant and equal to N_A, N_B and V_A and V_B. For simplicity we assume that the system is divided into two equal parts: $N_A = N_B = N/2$ and $V_A = V_B = V/2$.

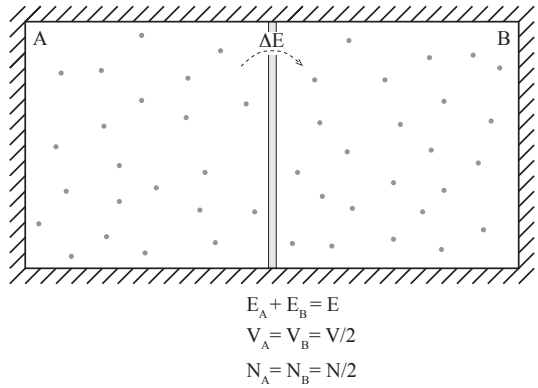


Fig. 6.14 An isolated ideal gas is divided into two equal-sized parts, A and B. The two parts are separated by a partition that allows energy to pass through, but it is immobile (so that the volumes do not change), and impermeable (so that the number of particles on each side is constant).

The question we want to answer is: What is the multiplicity of the macrostate described by E_A, $\Omega(E_A)$? In general, the multiplicity of the combined system is the product of the multiplicities of each of the systems:

$$\Omega(E_A; E, N, V) = \Omega_A\, \Omega_B = f(N_A) f(N_B)\, (V_A V_B)^N\, (E_A E_B)^N \ , \quad (6.45)$$

The only variation occurs in E_A and $E_B = E - E_A$. Just like we did for the Einstein crystal, we can develop a theory for this variation for small fluctuations in the energy. Let us look at a variation

$$E_A = \bar{E}_A + \Delta E \ , \ E_B = \bar{E}_B - \Delta E \ , \quad (6.46)$$

where $\bar{E}_A = \bar{E}_B = E/2$ when $N_A = N_B$ and $V_A = V_B$. The two systems are then equal, and we would expect each system to have the same equilibrium energy given by $\bar{E}_A$ and $\bar{E}_B$. We insert $E_A = E/2 + \Delta E$ and

$E_B = E/2 - \Delta E$ into the multiplicity, getting

$$\Omega(\Delta E) = C\left(\left(\frac{E}{2}\right)^2 - \Delta E^2\right)^N , \tag{6.47}$$

To simplify the algebra, we work on the logarithm of $\Omega(\Delta E)$, getting

$$\begin{aligned}\ln \Omega(\Delta E) &= \ln C + N \ln\left(\left(\frac{E}{2}\right)^2 - \Delta E^2\right) \\ &= \ln C + N \ln\left(\frac{E}{2}\right)^2 \left(1 - \left(\frac{2\Delta E}{E}\right)^2\right) \\ &= \ln C + 2N \ln(E/2) - N\left(\frac{2\Delta E}{E}\right)^2 ,\end{aligned} \tag{6.48}$$

which gives that

$$\Omega(\Delta E) = \Omega_m e^{-\left(\frac{\sqrt{N}\Delta E}{(E/2)}\right)^2} = \Omega_m e^{-\left(\frac{\Delta E}{(E/\sqrt{N})}\right)^2} , \tag{6.49}$$

This is Gaussian form for the multiplicity function, just like we found for the Einstein crystal. This means that also for the ideal gas, the multiplicity function is very sharp. Just as we found for the Einstein crystal, the width of the multiplicity function is proportional to $\sqrt{N}$ while the average value is proportional to E which is proportional to N, hence the relative value of the width goes like $1/\sqrt{N}$. For realistic values of N, such as $N = 10^{20}$, this relative width becomes a very small number, typically 10^{-10}, which means that we cannot discern the actual value of the energy E_A in the ideal gas from its average value $E/2$ unless we can measure E_A to a precision of 10^{-10}, which is practically impossible.

This means that it is very unlikely to find the system in a state that is significantly different from $E/2$, where the multiplicity is at its maximum. If the system started away from this value, for example by having much more energy in part A than in part B, the system would evolve towards the most probable macrostate, where $E_A = E/2$. This is an example of a general principle, the second law of thermodynamics.

6.5 Thermal equilibrium, Entropy and Temperature

We have found that for the two models system we have addressed, the ideal gas and the ideal (Einstein) crystal, the most likely macrostate is very sharp and any fluctuation away from the most likely macrostate is extremely unlikely. If we start from a microstate that corresponds to a macrostate that is away from the most likely macrostate the system will develop toward the most likely macrostate, and therefore toward a state with higher multiplicity — simply from the laws of probability.

This is what we call **the second law of thermodynamics**: The multiplicity of a system increases. What characterizes the state it evolves towards? The system evolves towards the macrostate with the largest multiplicity, and this state corresponds to the stationary state or the **equilibrium state** of the system.

Characteristics of the equilibrium state. What characterizes this equilibrium state — the state with maximum multiplicity? Let us find the maximum of the multiplicity for two Einstein crystals in contact. The multiplicity of a macrostate with energy q_A is

$$\Omega(N_A, q_A)\Omega(N_B, q_B) = \Omega(N_A, q_A)\Omega(N_B, q - q_A) \tag{6.50}$$

where $q = q_A + q_B$ is conserved. We can find the extremum of this function by differentiation with q_A:

$$\frac{d}{d\,q_A}\Omega(N_A, q_A)\Omega(N_B, q - q_A) = 0\ , \tag{6.51}$$

$$\frac{d\Omega(N_A, q_A)}{d\,q_A}\Omega(N_B, q - q_A) + \Omega(N_A, q_A)\frac{d\Omega(N_B, q - q_A)}{d\,q_A} = 0\ , \tag{6.52}$$

$$\frac{\partial\Omega(N_A, q_A)}{\partial\,q_A}\Omega(N_B, q - q_A) - \Omega(N_A, q_A)\frac{\partial\Omega(N_B, q_B)}{\partial\,q_B} = 0\ , \tag{6.53}$$

$$\frac{\partial\Omega(N_A, q_A)}{\partial\,q_A}\Omega(N_B, q_B) = \Omega(N_A, q_A)\frac{\partial\Omega(N_B, q_B)}{\partial\,q_B}\ , \tag{6.54}$$

$$\frac{1}{\Omega(N_A, q_A)}\frac{\partial\Omega(N_A, q_A)}{\partial\,q_A} = \frac{1}{\Omega(N_B, q_B)}\frac{\partial\Omega(N_B, q_B)}{\partial\,q_B}\ , \tag{6.55}$$

$$\frac{\partial \ln\Omega(N_A, q_A)}{\partial\,q_A} = \frac{\partial \ln\Omega(N_B, q_B)}{\partial\,q_B}\ . \tag{6.56}$$

$$\frac{\partial \ln\Omega(N_A, q_A)}{\partial\,q_A} = \frac{\partial \ln\Omega(N_B, q_B)}{\partial\,q_B}\ . \tag{6.57}$$

Definition of Entropy. This equation, (6.57), suggests an interesting relation between the logarithms of the multiplicities for the q_A-value that corresponds to the maximum multiplicity, that is, for the macrostate that is most probable. We introduce the term **entropy** to describe the logarithm of the multiplicity of the system:

Entropy: The entropy of an isolated system with given N, V, and E is given as

$$S = k \ln \Omega(N, V, E) \,, \tag{6.58}$$

where k is called the Boltmann constant, $k = 1.3806488(13)\, 10^{-23}\mathrm{J/K}$, and $\Omega(N, V, E)$ is the multiplicity of the system.

Relation between a microscopic and a macroscopic description. The definition of entropy relates the *microscopic description* of the world, given in the form of the multiplicity function, with the *macroscopic description* in the form of entropy. This definition of the entropy will allow us to calculate the macroscopic, thermodynamic properties of a system from a microscopic theory of the system.

Entropy and the second law of thermodynamics. We have seen that isolated systems will develop so that the multiplicity increases and they will develop toward a state with maximal multiplicity. We can now reformulate this law in terms of entropy: Isolated systems will develop so that the entropy increases, and in equilibrium the system will be in a macrostate with maximum entropy.

Properties of the Entropy. We see from the definition of the entropy that it is **additive**:Since the entropy is defined as the logarithm of the multiplicities, the entropy of two independent systems are the sums of their individual entropies, because the total multiplicity is the product of the individual multiplicities:

$$S = S_A + S_B \text{ when } \Omega = \Omega_A \cdot \Omega_B \,. \tag{6.59}$$

6.5.1 Equilibrium of the Einstein crystal

Let us see how the concept of entropy is related to the equilibrium of a system through an example we already know in detail — the Einstein crystal. For this system we know the multiplicity for a macrostate q_A,

where q_A of the energy is in part A of the system and $q_B = q - q_A$ is in part B. The multiplicity for the macrostate q_A is

$$\Omega(q_A) = \Omega(N_A, q_A)\,\Omega(N_B, q_B) = \frac{(q_A + N_A - 1)!}{q_A!(N_A - 1)!}\,\frac{(q_B + N_B - 1)!}{(q_B!(N_B - 1)!}\,. \tag{6.60}$$

And the entropy of the system when it is in macrostate q_A is therefore

$$S(q_A) = \ln \Omega(q_A) = \ln \frac{(q_A + N_A - 1)!}{q_A!(N_A - 1)!}\,\frac{(q_B + N_B - 1)!}{(q_B!(N_B - 1)!}\,. \tag{6.61}$$

We have already plotted the multiplicity of this system in Fig. 6.8, and we found that it had a sharp maximum. Let us redo this, but instead plot the Entropy $S(q_A)$ of the system as a function of the macrostate q_A. This is done using the a similar program, but we now take the logarithm of the multiplicity. The following prorgam find the multiplicity and the entropy as a function of q_A for two Einstein crystals in thermal contact with $N_A = 300$ and $N_B = 200$ and $q = 100$ [5]

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.special import comb
NA = 300
NB = 200
q = 200
multA = np.zeros(q+1,float)
multB = np.zeros(q+1,float)
mult =  np.zeros(q+1,float)
N = NA + NB
qvalue = np.array(range(q+1))
for ik in range(len(qvalue)):
    qA = qvalue[ik]
    qB = q - qA
    multA[ik] = comb(qA+NA-1,qA)
    multB[ik] = comb(qB+NB-1,qB)
    mult[ik] = multA[ik]*multB[ik]
SA = np.log(multA)
SB = np.log(multB)
STOT = SA + SB
plt.plot(qvalue,SA,'-r',qvalue,SB,'-b',qvalue,STOT,':k')
plt.xlabel('q_A'), plt.ylabel('S')
```

The resulting entropies of system A, $S_A(q_A)$, system B, $S_B(q - q_A)$, and of the total system, $S_{TOT}(q_A) = S_A(q_A) + S_B(q - q_B)$, is plotted in Fig. 6.15.

[5] We have here chosen N_A and N_B not to be equal so that the position of the maximum is not at the center of the figure. This makes the arguments clearer.

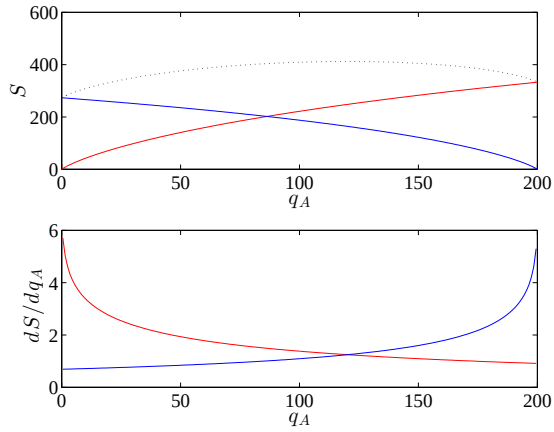


Fig. 6.15 Plot of the entropies S_A and S_B of two coupled Einstein crystals, and plots of the derivatives of the entropies with respect to q_A.

Entropy and equilibrium. We can use the plot of $S_{TOT}(q_A)$ and $S_A(q_A)$ in Fig. 6.15 to better understand the condition for equilibrium. Equilibrium occurs at the q_A-value where the entropy (and therefore the multiplicity) is maximum. This point is easily recognized in Fig. 6.15a. However, we can also find the equilibrium from the equilibrium condition in (6.57):

$$\frac{\partial \ln \Omega(N_A, q_A)}{\partial q_A} = \frac{\partial \ln \Omega(N_B, q_B)}{\partial q_B} \,, \tag{6.62}$$

which we can we translate into a condition for the entropies

$$\frac{\partial S(N_A, q_A)}{\partial q_A} = \frac{\partial S(N_B, q_B)}{\partial q_B} \,, \tag{6.63}$$

that is, the derivatives of the entropies of the two subsystems are the same when in equilibrium.

The derivatives of the entropies are shown in Fig. 6.15b. How can we interpret the condition in (6.63)? We can use the plot of $\partial S_A/\partial q_A$ to see what happens if system A is in a macrostate with a q_A-value that is smaller than the equilibrium value, such as the point marked with a circle in Fig. 6.15b. In this case, we see from the plot of $\partial S_A/\partial q_A$ that if we increase q_A by one, S_A increases and S_B decreases. However, we also see from the two curves that S_A increases more than S_B decreases, because the slope of $S_A(q_A)$ is steeper than that of $S_B(q_A)$, that is

$$\frac{\partial S(N_A, q_A)}{\partial q_A} > \frac{\partial S(N_B, q_B)}{\partial q_B} \,, \tag{6.64}$$

in this point. A small increase in q_A by Δq: $q_A \to q_A + \Delta q$, therefore results in an increase in the total entropy of the system

$$\Delta S_{TOT} = \Delta S_A + \Delta S_B = \frac{\partial S_A}{\partial q_A}\Delta q_A + \frac{\partial S_B}{\partial q_B}\Delta q_B = \left(\frac{\partial S_A}{\partial q_A} - \frac{\partial S_B}{\partial q_B}\right)\Delta q \tag{6.65}$$

since $\Delta q_A = \Delta q$ and $\Delta q_B = -\Delta q_A = -\Delta q$ since $q_B = q - q_A$ and q is a constant when the small amount Δq of energy is transferred from part B to part A. Since the entropy must increase as the system evolves towards equilibrium, we see that a small transfer of energy from part B to part A will occur when q_A is smaller than the equilibrium value.

Test your understanding: See if you can formulate a similar argument when q_A is larger than the equilibrium value.

It is first when the two slopes are equal that the system is in equilibrium. What does the slope of S_A tell us? It tells us how much the entropy (or multiplicity) changes when we add a small amount of energy. It tells us which way the system will develop, because it will develop toward larger entropy (multiplicity). For $q_A < \bar{q}_A$ the system will gain more entropy (multiplicity) by increasing q_A and decreasing q_B. Since the system, most likely, develops towards larger entropy (multiplicity) it will develop this way. Not always, but most probably.

6.5.2 Temperature in the Einstein crystal

Now, we can interpret the slope, $\partial S_A/\partial q_A$. This is what is equal when two systems are in thermal equilibrium. Since we also know that the temperature is equal in two systems in thermal equilibrium, it means that this relation is related to the temperature. How is it related to the temperature? The simplest assumption would be that it is a simple function of temperature. (If we look at the dimensions, we would get a strong hint, but this, does, of course depend on how we have chosen the dimensions of k_B, which means that this argument becomes circular.)

When the slope, $\partial S/\partial q$, of the entropy as a function of energy is larger in system A than in system B, $\partial S_A/\partial q_A > \partial S_B/\partial q_B$, energy will spontaneously flow from system B to system A, since this will lower the total energy. Since we also know that thermal energy tends to flow from a system with high temperature to a system with low temperature, we expect the temperature to be low when the slope is high. One possible functional relationship with this behavior is to assume that

the temperature is inversely proportional to the slope. This will indeed be how we define the temperature:

Temperature in an isolated system — a system with given (N, V, E):

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} , \tag{6.66}$$

where we have included the N and V to show that we keep them constant as we are changing E.

This is our definition of temperature. We will in the following see that this is indeed consistent with our other possible definitions of temperature, such as from the average kinetic energy per particle in an ideal gas.

Connection between the microscopic and the macroscopic. We now have all the tools we need to connect the microscopic theories we have developed and the measurements we have done with macroscopic properties such as temperature, energy, and heat. This allows us to connect between the microscopic and the macroscopic worlds. This is the coupling between *statistical mechanics* and *thermodynamics*. Thermodynamics is the rules and laws we use to describe thermal and mechanical properties of macroscopic systems, whereas statistical mechanics is the theory that allows us to calculate and prove thermodynamic principles, laws and properties from a microscopical foundation.

We will in the following use this coupling to calculate the entropy of a system, $S(N, V, E)$ from a microscopic model: It may be from a theoretical model where we can count all the states or from a numerical simulation such as a molecular dynamics simulation. As soon as we have the function $S(N, V, E)$ we can start applying thermodynamics to this function, to understand relations between energy, temperature, volume, pressure, and the number of particles in our system. For this development, for the application of thermodynamics, it is not important how we have found $S(N, V, E)$, we just need this function as a starting point for our calculations. Let us therefore see how we can use the theories we have developed to find entropies, temperatures and energies for the ideal gas and the ideal crystal systems.

Microscopic.

- Multiplicity: $\Omega(N, V, E)$
- An isolated system develops towards larger multiplicify

- In equilibrium, an isolated system is in the macrostate with maximum multiplicity

Macroscopic.

- Entropy: $S = k_B \ln \Omega(N, V, E)$
- An isolated system develops towards larger entropy
- In equilibrium, an isolated sytem is in the macrostate with maximum entropy
- Temperature: $(1/T) = (\partial S/\partial E)_{N,V}$

Clickers: Thermal 01

6.5.3 Example: Entropy and energy of the Einstein crystal

We can now start connecting the microscopic and the macroscopic worlds. We have developed methods to find the multiplicity of an ideal gas and an Einstein crystal for a system with given N, V, and E. From this we can calculate the entropy, and from the entropy we can calculate the temperature.

The multiplicity of the Einstein crystal is

$$\Omega = \left(\frac{e\,q}{N}\right)^N \ , \tag{6.67}$$

in the limit when q and N is large, and $q \gg N$. The internal energy, E, is related to q by $E = q\epsilon$, where ϵ is the energy unit for the energy states of the harmonic oscillator. The entropy is

$$S = k \ln \Omega = Nk \left(\ln q + \ln(E/\epsilon) - \ln N\right) \ , \tag{6.68}$$

$$S = Nk \ln E - Nk \ln(\epsilon\, N) + Nk \ . \tag{6.69}$$

And the temperature is:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right) = \frac{Nk}{E} \ , \tag{6.70}$$

which gives

$$E = NkT \ . \tag{6.71}$$

This is what we would expect from the **equipartition theorem**, since N is the number of oscillators, and each oscillator has two degerees of freedom, hence, in equilibrium the energy is kT per oscillator and NkT in total.

From this we can also predict the **heat capacity**. The heat capacity is the amount of thermal energy, heat, which we need to transfer to a system in order to increase its temperature by one unit (Kelvin), which is given as the derivative of the energy with respect to temperature:

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{N,V} = \frac{\partial}{\partial T}\left(NkT\right) = Nk \; . \tag{6.72}$$

We call this the heat capacity at constant volume, since we keep both the volume and the number of particles constant when we take the partial derivative.

What we have found here now is a theoretical prediction of the heat capacity of a solid. The prediction is that it only depends on the number of oscillators, N, in the solid. This result can be compared with experimental or simulational results to check its validity.

6.5.4 Example: Entropy and energy of the ideal gas

We can perform a similar calculation for the ideal gas. In this case the multiplicity was a bit more complicated:

$$\Omega = \frac{1}{N!}\frac{1}{2^{3N}}\frac{2\pi^{3N/2}}{\left(\frac{3N}{2}-1\right)!}\left(\frac{2L}{h}(2mE)^{1/2}\right)^{3N-1} \; . \tag{6.73}$$

Here, we will assume that N is very large, so that $N-1 \simeq N$, which simplifies the expression to:

$$\Omega \simeq \frac{1}{N!}\frac{1}{2^{3N}}\frac{2\pi^{3N/2}}{\left(\frac{3N}{2}\right)!}\left(\frac{2L}{h}(2mE)^{1/2}\right)^{3N} \; . \tag{6.74}$$

This is further simplified to

$$\Omega = 2\frac{1}{N!\,(\frac{3}{2}N)!}\left(\frac{2\pi mEL^2}{h^2}\right)^{3N/2} \; . \tag{6.75}$$

And the entropy is

$$\begin{aligned}\frac{S}{k} &= \ln \Omega \\ &= -\left(N \ln N - N\right) - \left(\frac{3N}{2} \ln \frac{3N}{2} - \frac{3N}{2}\right) + N \ln \left(\frac{2\pi m E L^2}{h^2}\right)^{3/2} \\ &= N \left(-\left(-1\right) - \left(-\frac{3}{2}\right) + \ln \left(\frac{V}{N}\right) + \ln \left(\frac{4\pi m E}{3N h^2}\right)^{3/2}\right) \\ &= N \left(\ln \frac{V}{N} \left(\frac{4\pi m E}{3N h^2}\right)^{3/2} + \frac{5}{2}\right) .\end{aligned} \tag{6.76}$$

This equation is called Sackur-Tetrode's equation.

Sackur-Tetrode's equation for the entropy of an ideal gas:

$$S = N k_B \left(\ln \frac{V}{N} \left(\frac{4\pi m E}{3N h^2}\right)^{3/2} + \frac{5}{2}\right) . \tag{6.77}$$

We can use this to find the energy E of the ideal gas, by finding the temperature:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} . \tag{6.78}$$

Since N and V are constants in this derivative, we only need to include the terms that include the energy:

$$S(N, V, E) = g(N, V) + Nk \ln E^{3/2} = g(N, V) + \frac{3}{2} Nk \ln E , \tag{6.79}$$

where $g(N, V)$ contains all the factors that do not depend on E. It is now easy to find the temperature:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{3}{2} Nk \frac{1}{E} , \tag{6.80}$$

which gives

$$E = \frac{3}{2} NkT . \tag{6.81}$$

Again, this corresponds to the **equipartition principle**. The ideal gas has N atoms, and each atom has three degrees of freedom: classically this corresponds to three independent directions for the velocities, and

quantum-mechanically this corresponds to the three independent quantum number n_x, n_y, n_z we used to characterize the states of the ideal gas. Each degree of freedom contributes with $kT/2$, which gives $3NkT/2$ for the whole gas of N gas atoms.

6.5.5 Example: Entropy of expansion

We can use Sackur-Tetrode's equation to gain more insight into the behavior of the ideal gas. First, let us look at what happens if we allow an ideal gas to expand without changing its energy or the number of particles as illustrated in Fig. 6.16. We only change the volume. You can think of an isolated box of size V_2, where the gas initially is confined to a smaller part, V_1, of the box. Then an internal barrier is removed, and the gas is allowed to fill the entire volume, while the total energy and the total number of particles is conserved since the whole system is isolated. The change (increase, as always) in entropy in this case by going from a volume V_1 to a volume V_2 is:

$$\Delta S = Nk\left(\ln V_2 - \ln V_1\right) = Nk \ln \frac{V_2}{V_1} \,. \tag{6.82}$$

In this case, the expansion of the system was *isoenergetic* — done at constant energy of the whole system. However, since for an ideal gas, the energy and the temperature are related by $E = (3/2)NkT$, we see that when the number of particles is constant, having a constant energy corresponds to having a constant temperature. In this particular case, the change in entropy we have found therefore also corresponds to an *isothermal* expansion — an expasion done at constant temperature. (*Iso* here means *at the same* or constant).

In this case, the process is clearly also *irreversible* because the entropy of the isolated system increases during the processes. Indeed, we would be very surprised if we placed the dividing wall back into the system, and all the gas particles spontaneously concentrated in the initial volume, V_1, by itself. Such a process in an isolated system would be violating the second law of thermodynamics, because entropy would decrease, and also appears to violate our basic intuitions about valid or probable processes in the world.

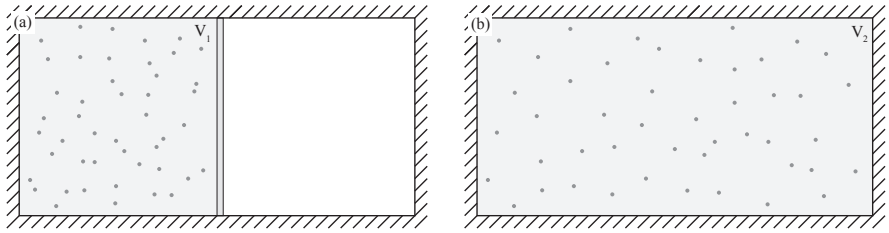


Fig. 6.16 Illustration of an expansion of an ideal gas while the gas is in an isolated box. The system is isolated, and the gas expands from V_1 to V_2 after the internal wall was removed from the system.

6.5.6 Example: Entropy of mixing

What if we take two gases, A and B, each initially occupying a volume V, and then allow the two gases to mix? Let us assume that the two gases are inside a common, isolated system, so that their total number of particles, energy and volumes are conserved. Then an internal barrier between the two gases is removed, so that gas A can move into the volume originally occupied by gas B and vice versa as illustrated in Fig. 6.17. If the two gases are not interacting, and the two gases are not the same, we can consider each process independently. First gas A expands into twice its volume, leading to an entropy change

$$\Delta S_A = N_A k \ln \frac{V_{A,2}}{V_{A,1}} \,, \tag{6.83}$$

and similarly for gas B:

$$\Delta S_B = N_B k \ln \frac{V_{B,2}}{V_{B,1}} \,, \tag{6.84}$$

If $N_A = N_B = N$ and $V_2 = 2V_1$ for both A and B, we get

$$\Delta S_A = Nk \ln 2 \,, \; \Delta S_B = Nk \ln 2 \,, \tag{6.85}$$

and the total entropy change is

$$\Delta S_{TOT} = \Delta S_{\text{mix}} = \Delta S_A + \Delta S_B = 2Nk \ln 2 \,, \tag{6.86}$$

where we have used the additive property of entropies. This entropy is called the *entropy of mixing* of the system.

Notice that we here assumed that the two gases were different. Otherwise, nothing would happen — there would not be any change in entropy if we open a separation between two equal gases each of volume V.

We also clearly see that this process is *irreversible* for this isolated system since the entropy is increasing. We would be very surprised if, after the system was completely mixed, gas A spontaneously segregated into the left side and gas B segregated into the right side. Indeed, such as processes in an isolated system would violated the second law of thermodynamics, since the total entropy of the isolated system would increase. Your intuition is therefore perfectly aligned with the second law of thermodynamics in this case.

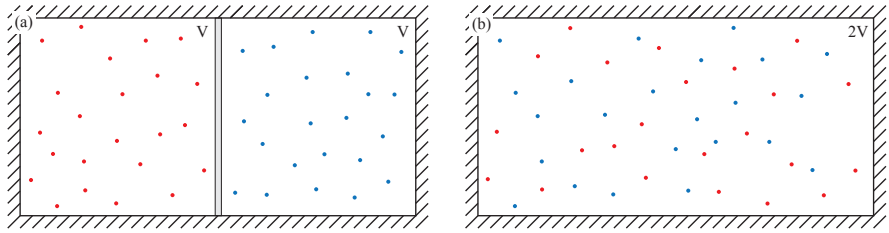


Fig. 6.17 Illustration of the mixing of two gases A and B, each with initial volume V.

6.5.7 Example: Gibb's paradox

Sackur-Tetrode's equation is carefully constructed to ensure that the entropy does not change if we mix two identical gases. What is the change in entropy if we have two identical gases each with N particles and volume V and at the same energy, E, and then we remove the barrier between them?

If the gases were different, we know that the change in entropy would correspond to the entropy of mixing. If the gases are identical, there should not be any change in entropy. What does the Sackur-Tetrode equation give us?

Before the partition is removed, the entropy of each part is $S(N, V, E)$ so that the total entropy of the system is $2\,S(N, V, E)$. After the partition is removed, the system consists of a gas with $2N$ particles, $2V$ volume and $2E$ energy with entropy $S(2N, 2V, 2E)$. According to Sackur-Tetrode's equation the entropy is

$$S(2N, 2V, 2E) = (2N)k \left(\ln \frac{2V}{2N} \left(\frac{4\pi m(2E)}{3(2N)h^2} \right)^{3/2} + \frac{5}{2} \right) . \tag{6.87}$$

We see that the two ratios $2V/2N = V/N$ and $2E/2N = E/N$, and that the total entropy therefore is

$$S(2N, 2V, 2E) = 2Nk \left(\ln \frac{V}{N} \left(\frac{4\pi mE}{3Nh^2} \right)^{3/2} + \frac{5}{2} \right) = 2S(N, V, E) . \tag{6.88}$$

Sackur-Tetrode's equation therefore correctly predicts that there is no change in entropy when we remove such an internal barrier between two identical gases. However, in order for this to be the case, it was essential to include the $N!$-term in the multiplicity, because it is this term that ends up as the N in the V/N term. Without the $N!$-term, we would instead have

$$S(N, V, E) = Nk \left(\ln \left(V \left(\frac{4\pi mE}{3Nh^2} \right)^{3/2} \right) + \frac{3}{2} \right) . \tag{6.89}$$

and the entropy of a double system, $S(2N, 2V, 2E)$, would be:

$$S(2N, 2V, 2E) = 2Nk \left(\ln \left(2V \left(\frac{4\pi m2E}{3N2h^2} \right)^{3/2} \right) + \frac{3}{2} \right) \tag{6.90}$$

The difference between $2S(N, V, E)$ and $S(2N, 2V, 2E)$ would then be $2Nk \ln 2$, which is the entropy of mixing. The Sackur-Tetrode equation is therefore carefully made to ensure that this does not occur. This issue was first raised by Gibbs, and is called *Gibbs paradox*.

We notice that the entropy scales in a simple way when we change all the variables with a factor b: $S(bN, bV, bE) = bS(N, V, E)$. We call quantities that scale in such as simple way *extensive quantities*.

6.6 The laws of thermodynamics

We now have the basis for the two first laws of thermodynamics, and we have both a microscopic and a macroscopic interpretation of the laws of thermodynamics.

In the *microscopic view*, we know that a system with constant N, V, and E has constant energy. This is indeed how we have set the system up. On an atomic scale, the total energy is the sum of the kinetic and

potential energies of all the atoms in the system. and this energy is conserved. This corresponds to the **first law of thermodynamics** – the conservation of energy.

In the microscopic view, we characterize a system by the multiplicity $\Omega(N, V, E)$, which we have been able to calculate for simple systems such as the ideal gas and the ideal crystal. We have also found two basic laws:

- The system develops in time to the macrostate with the maximum multiplicity
- In equilibrium, the system in the macrostate with the maximum multiplicity

We have found that we can characterize a system by its microstates. We assume that all microstates are equally probable. There are much more microstates for some macrostates, and these most likely macrostates will be the most likely observed macrostates in the system. Indeed, for realistic system sizes — i.e. $N = 10^{20}$ — deviations from the maximum values of the macrostates are extremely unlikely. The standard deviation of the distribution of macrostates is about 10^{-10} of the value at the maximum. These distributions are therefore so sharp that we can characterize the system by the value of the most likely macrostate alone.

These two laws are called the second law of thermodynamics. Here, they have a microscopic formulation. How can we formulate similar laws on the macroscale?

The first law of thermodynamics is easily transferrable. The total energy of the system is converved. We call this the internal energy E. Any change in the internal energy must come from an external source: In the form of mechanical work W performed on the system or in the form of thermal energy Q transferred into the system. Hence, we can formulate the first law of thermodynamics:

First law of thermodynamics:

$$\Delta E = W + Q \,, \tag{6.91}$$

Macroscopically, we have introduced the quantity S for the entropy of the system, and entropy is related to the multiplicity of the system with constant N, V, and E by

$$S = k \ln \Omega(N, V, E) \,. \tag{6.92}$$

The two microscopic laws that (i) the system evolves towards larger multiplicity and (ii) that in equilibrium the system has maximum multiplicity therefore corresponds to

The second law of thermodynamics:
An isolated system with constant N, V, E develops towards larger entropy, $\Delta S \geq 0$, and in equilibrium, the entropy is at its maximal value.

These are the fundamental laws we need to understand the macroscopic world. We can use our microscopic understanding to understand the foundation of the macroscopic laws, and to calculate the relationships for quantities such as the entropy.

In addition, we usually supplement with two additional laws:

The zeroth law of thermodynamics:
If systems A and B are in thermal equilibrium and systems B and C are in thermal equilibrium, then system A and C are also in thermal equilibrium.

This follows from the equation for thermal equilibrium - and temperature: If $T_A = T_B$ and $T_B = T_C$ then, indeed, $T_A = T_C$.

In addition, we will later also introduce a third law of thermodynamics, which allows us to set an absolute scale for the entropy.

The third law of thermodynamics: The entropy of a system approaches a constant as the temperature approaches zero.

These are the fundamental laws of thermodynamics. From these laws we will be able to develop the complete theory of thermodynamics. We will now used these laws, along with our microscopical theory that allows us to calculate the entropy and the temperature of an isolated system, in order to address the behavior of realistic, macroscopic systems.

Clickers: Thermal02

6.6.1 Example: Entropy and heat

Let us now use the laws of thermodynamics to discuss a process without resorting to a microscopic description. Let us address the entropy change for a real process where thermal energy, heat, is transferred from one system to another.

First, we need to be precise in how we define the system we study. We start with an isolated system that consists of two subsystems, A and B. Initially, the two systems have temperatures T_A and T_B, energies E_A and E_B, and entropies, S_A and S_B. The systems are also isolated from each other. But what happens if we now transfer a small amount of thermal energy, ΔE, from A to B, so that the total energy in the system is conserved. This means that the change of energy in system A is $\Delta E_A = -\Delta E$ and the change of energy in system B is $\Delta E_B = \Delta E$.

Change in entropy. The entropy change in the whole system is

$$\begin{aligned} \Delta S &= \Delta S_A + \Delta S_B \\ &= \left(\frac{\partial S_A}{\partial E_A}\right)_{N_A} \Delta E_A + \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B} \Delta E_B \\ &= \left(\frac{\partial S_A}{\partial E_A}\right)_{N_A} (-\Delta E) + \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B} \Delta E \end{aligned} \tag{6.93}$$

We can now use the general relation between temperature and the derivative of the entropy for each of the two systems:

$$\frac{1}{T_A} = \left(\frac{\partial S_A}{\partial E_A}\right)_{N_A} , \quad \frac{1}{T_B} = \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B} . \tag{6.94}$$

The total change in entropy is therefore:

$$\Delta S = \Delta S_A + \Delta S_B = \left(-\frac{1}{T_A} + \frac{1}{T_B}\right) \Delta E . \tag{6.95}$$

Because the whole system is isolated, the total entropy must increase (or be the same) in this process. This means that T_A must be larger than T_B for this process to be thermodynamically possible. (Meaning probable: Any other process would be very unlikely). This is, of course, in perfect alignment with our intuition: Heat is transferred from objects with high temperature to objects with lower temperature and not the opposite way.

Heat and entropy change. How is the heat transferred between the two systems related to the entropy change in the systems? Since there is no other energy transfer than heat between systems A and B, we know that

the change in energy corresponds to heat transferred. From the first law of thermodynamics we know that $\Delta E_A = Q_A + W_A$, and since $W_A = 0$, we see that $Q_A = \Delta E_A$. Similarly, for system B, $Q_B = \Delta E_B$. It is now simple to relate the entropy change and the change in energy for each of the systems. As long as the systems do not change volume or particle number, we can use the definition of temperature to find the change in entropy:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} \Rightarrow dS = \frac{dE}{T} \ . \tag{6.96}$$

This means that we can estimate the change in entropy from the heat:

$$dS = \frac{Q}{T} \ . \tag{6.97}$$

or we can estimate the heat from the change in entropy:

$$Q = TdS \ . \tag{6.98}$$

Numerical example. For example, if a hot object, A, at $T_A = 1500\,\mathrm{K}$ is put in contact with a cold object, B, at $T_B = 300\,\mathrm{K}$, and the thermal energy transfered during $60\,\mathrm{s}$ is $150\,\mathrm{J}$, then we can calculate the changes in entropy through:

$$\Delta S_A = \frac{-150\,\mathrm{J}}{1400\,\mathrm{K}} = -0.1\,\mathrm{J/K} \ , \tag{6.99}$$

and

$$\Delta S_B = \frac{150\,\mathrm{J}}{300\,\mathrm{K}} = 0.5\,\mathrm{J/K} \ . \tag{6.100}$$

We see that the total change in entropy is indeed positive, as it must for this process to be possible according to the second law of thermodynamics.

6.7 Mechanical equilibrium — Pressure

So far we have developed a theory for a fluctuation in the energy of a system: We divided an isolated system into two parts, and allowed a small amount of energy, ΔE, to be transported from one part to another, and found the equilibrium state as a function of the derivative of the entropy. Now, let us see if we can expand this argument also to systems where we allow also the volume of the system to vary and futher on, also

the number of particles to vary. But we will now use a thermodynamics argument without resorting to the microscale. Can we pull this off?

We address the behavior of an isolated system divided into two parts A and B. The systems are divdided by a piston: A wall that can transfer thermal energy and mechanical energy by moving. The wall is not permeable, so the number of particles is constant on each side of the wall. The total energy and volume is conserved: $E_A + E_B = E$ and $V_A + V_B = V$, as well as the number of particles in each subsystem N_A, N_B, as illustrated in Fig. 6.18.

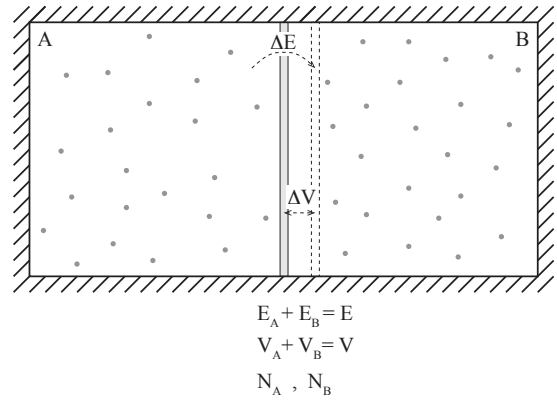


Fig. 6.18 Illustration of an isolated system divided into two parts by a piston so that thermal energy and volume may be exchanged between the systems, but the total energy and volume is conserved.

Intuitively, what would we expect to determine the equilibrium in such as system? We already know that for thermal equilibrium to occur, the two systems must have the same temperature, otherwise we could increase entropy by transferring some energy from one part to another. But what would determine mechanical equilibrium? We would expect the piston to be in equilibrium only if the force acting from each part of the system on the piston sum to zero, so that the net force on the piston is zero. This occurs only if the pressure in each of the parts are the same, since the force from part A on the piston is $F_A = p_A A_A$, where F_A is the force and p_A is the pressure, and similarly for part B, $F_B = p_B A_B$. Here, the areas on each side of the piston are the same, so that the pressures also must be the same. Equilibrium would therefore require equal pressures.

What are the equilibrium condition from a thermodynamical perspective? We know from the second law of thermodynamics, that an isolated system — the whole system including both part A and B — is in equilibrium when the entropy of the whole system is maximum. The

total entropy is the sum of the entropies of each subsystem:

$$S_{TOT} = S_A(N_A, V_A, E_A) + S_B(N_B, V_B, E_B) \,, \tag{6.101}$$

where $V_A + V_B = V$, $E_A + E_B = E$m N_A, and N_B are constants, so that $V_B = V - V_A$ and $E_B = E - E_A$. The entropy is therefore a function of the macrostate described by E_A and V_A, and we want to find the macrostate that maximizes the entropy. This corresponds to finding the maximum of a two-dimensional function, which occurs as a maximum along both axes, as illustrated in Fig. 6.19. The maximum (or extremum) occurs when $dS = 0$, that is, when

$$dS = dS_A + dS_B = \left(\frac{\partial S_A}{\partial E_A}\right)_{N_A,V_A} dE_A + \left(\frac{\partial S_A}{\partial V_A}\right)_{N_A,E_A} dV_A \tag{6.102}$$

$$+ \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B,V_B} dE_B + \left(\frac{\partial S_B}{\partial V_B}\right)_{N_B,E_B} dV_B = 0 \,, \tag{6.103}$$

where we now insert that $dE_B = -dE_A$ and $dV_B = -dV_A$ (from $E_A + E_B = E = \text{const.}$ and $V_A + V_B = V = \text{const.}$), getting:

$$dS = \left(\left(\frac{\partial S_A}{\partial E_A}\right)_{N_A,V_A} - \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B,V_B}\right) dE_A \tag{6.104}$$

$$+ \left(\left(\frac{\partial S_A}{\partial V_A}\right)_{N_A,E_A} - \left(\frac{\partial S_B}{\partial V_B}\right)_{N_B,E_B}\right) dV_A = 0 \,. \tag{6.105}$$

For this to be true for any small dE_A and dV_A, we get the two conditions:

$$\left(\frac{\partial S_A}{\partial E_A}\right)_{N_A,V_A} = \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B,V_B} \,, \quad \left(\frac{\partial S_A}{\partial V_A}\right)_{N_A,E_A} = \left(\frac{\partial S_B}{\partial V_B}\right)_{N_B,E_B} \,. \tag{6.106}$$

The first condition is the same condition we found for a purely thermal contact: The temperatures in the two parts must be the same in equilibrium. We guess that the second term is related to the pressures in the two parts. But how is it related to the pressure? We can gain some insight into this by looking at the units of this expression: dS/dV. We notice that entropy has dimensions J/K=Nm/K, V has dimensions m^3, and pressure has dimensions $\mathrm{N/m}^2$. The derivative of S with respect to V therefore has dimensions $(\mathrm{N\,m\,/\,K})\,/\,\mathrm{m}^3$, that is $(\mathrm{N\,/\,m}^2)\,/\,\mathrm{K}$. We must therefore multiply (dS/dV) by a temperature to get dimension pressure. Since T is the same in both systems in equilibrium we can multiply with T, getting to a definition of pressure:

Definition of pressure in an isolated system:

$$p = T \left(\frac{\partial S}{\partial V} \right)_{E,N} , \tag{6.107}$$

This is not the only way to define pressure. We could for example also have included a multiplicative or an additive constant, but we will see below that this definition of pressure does reproduce the correct ideal gas law and other know features of the systems we are interested in.

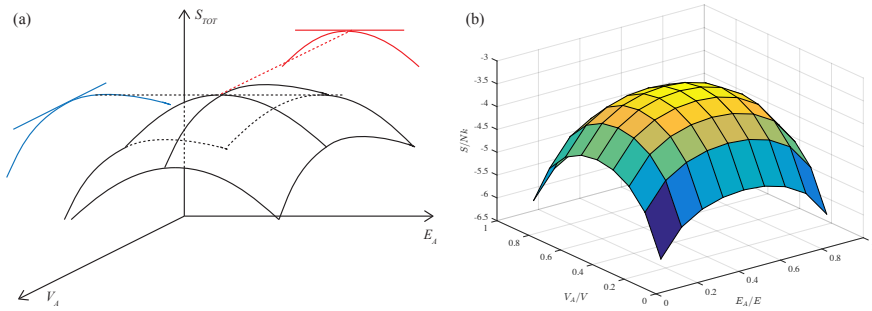


Fig. 6.19 **a** Illustration of the entropy as a function of E_A and V_A. The maximum of the entropy occurs where the derivatives with respect to both E_A and to V_A are zero. **b** Plot of the entropy, $S = S_A + S_B$, as a function of E_A and V_A for two ideal gases in thermal and mechanical contact.

Clickers: Thermodynamics 04

6.7.1 Example: Pressure of ideal gas

We are now in a position to calculate the pressure of a system, given that we know the entropy. For an ideal gas, we found that the entropy was given by Sackur-Tetrode's equation:

$$S = Nk \ln \left(\frac{V}{N} \left(\frac{4\pi m E}{3Nh^2} \right)^{3/2} + \frac{5}{2} \right) . \tag{6.108}$$

However, we only need the volume dependence to find the pressure:

$$S = Nk \ln V + f(N, E) . \tag{6.109}$$

From (6.107) the pressure in the ideal gas is:

$$p = T\left(\frac{\partial S}{\partial V}\right)_{E,N} = T\frac{Nk}{V} \tag{6.110}$$

which corresponds to

$$pV = NkT\,, \tag{6.111}$$

which indeed is the ideal gas law. This also shows we should not include any other factors in the definition of the pressure in (6.107) — it reproduces the correct ideal gas law at it is.

Clickers: Thermal 03

6.8 The thermodynamic identity

We now know enough to develop a differential for the entropy: We would like to see how the entropy S changes with a change in E and V. We do this by writing out the differential for the change dS:

$$dS = \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN\,. \tag{6.112}$$

We have not looked at changes in particle number yet. Let us therefore for now assume that $dN = 0$. The other two expressions we now know

$$\left(\frac{\partial S}{\partial E}\right)_{V,N} = \frac{1}{T}\,,\ \left(\frac{\partial S}{\partial V}\right)_{E,N} = \frac{1}{T}p\,. \tag{6.113}$$

We insert these into the differential in (6.112):

$$dS = \frac{1}{T}dE + \frac{p}{T}dV \ \Rightarrow\ TdS = dE + p\,dV\,. \tag{6.114}$$

This equation is true for any infitesimal change in any system, as long as p and T are well defined and N is constant. It is a fundamental equation in thermodynamics, which we will use many times throughout this book. It is called

The **thermodynamic identity:**

$$T\,dS = dE + p\,dV - \mu\,dN\,. \tag{6.115}$$

(We have not yet allowed N to vary or introduced the chemical potential μ, but we write the equation in this form for completeness. When N is constant, $dN = 0$, and the last term disappears).

6.8.1 The thermodynamic identity and partial derivatives

This equation, in combination with the first and second law of thermodynamics will be our main tool to address thermodynamic processes, as we will see in examples and problems. However, this equation is also a very compact way to remember the various derivatives of the entropy. From (6.115) we can easily find the various partial derivatives by applying simple algebra to the differentials. Let us demonstrate by examples.

First, what is the derivative of S with respect to E when V and N are constant? We can see this directly from the thermodynamic identity, using that $dV = 0$ and $dN = 0$:

$$T\,dS = dE + p\,\underbrace{dV}_{dV=0} - \mu\,\underbrace{dN}_{dN=0} \tag{6.116}$$

$$\frac{dS}{dE} = \frac{1}{T}\,. \tag{6.117}$$

which gives us:

$$\left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{1}{T}\,. \tag{6.118}$$

Similarly, we find the derivative of S with respect to V when E and N are constants? In this case, $dE = 0$ and $dN = 0$, which we insert into the thermodynamic identity in (6.115), getting

$$T\,dS = \underbrace{dE}_{dE=0} + p\,dV - \mu\,\underbrace{dN}_{dN=0} \tag{6.119}$$

$$\frac{dS}{dV} = \frac{p}{T}\,. \tag{6.120}$$

which gives us:

$$\left(\frac{\partial S}{\partial V}\right)_{N,E} = \frac{p}{T}\,. \tag{6.121}$$

This is a very useful technique, which we will use again and again in various disguised throughout this text. Learn and practice it, and you will only have to remember a very few definitions and laws in order to master all of thermodynamics.

Clickers: Thermal 04

6.8.2 Entropy, Heat and Work

If we compare the thermodynamic identity

$$T\,dS = dE + p\,dV - \mu\,dN\ , \tag{6.122}$$

which when $dN = 0$ can be rewritten as

$$dE = T\,dS - p\,dV\ , \tag{6.123}$$

with the first law of thermodynamics

$$dE = Q + W\ , \tag{6.124}$$

We recongnize that the work done on a system, such as a piston, with a pressure p, when the piston is moved so that the volume changes by a small amount dV, is $W = p\,dV$. Notice the sign! The work W in the first law is the work done *on* the system. If the pressure inside the system is p and the pistons expands so the $dV > 0$, the system inside the piston does a positive work pdV on the environment, whereas the work done on the system is the negative of this, $W = -p\,dV$.

Similarly, if there is no mechanical work done on the system, then the change in energy corresponds to the heat, Q, transferred into the system. Therefore, we may conclude that $Q = T\,dS$ in this case.

However, these results are only correct when the processes we consider are sufficiently close to the equilibrium, that is, when the processes are *quasistatic*. Let us now briefly review how we can describe various thermodynamic processes.

6.8.3 Process description

We describe thermodynamic systems by the macrostate variables of the system, such as by S, E, V, N, T, and p. Various thermodynamic

processes are described by how we change these variables and by how we change the system.

We discern between **reversible** and **irreversible** processes. A reversible process is a process that may go both ways — backward and forward. That is, the process has time symmetry. This means that the entropy change for a reversible process in an isolated system must be zero. If it was positive, then the reverse process cannot occur since it would lead to a reduction in entropy.

If we always keep a system very close to its equilibrium state, we call the process **quastistatic**. This is an ideal case, where the process occurs infinitely slowly so that the system always has time to reach thermodynamic equilibrium between each time the system is changed.

In addition, we will introduce various words to describe processes where some of the thermodynamic variables are constant. Usually, we will use the prefix *iso* for the constant variable. For example, a process at constant temperature is called *isothermal*, whereas a process at constant pressure is called *isobaric*. The various processes are

- **quasistatic processes** where the system is changed so slowly that it is at equilibrium at each step
- **adiabatic processes** where there is no thermal energy (heat) transferred to the system
- **isentropic processes** = adiabatic + quasistatic processes
- **isobaric processes** = processes at constant pressure
- **isochoric processes** = processes at constant volume
- **isothermal processes** = processes at constant temperature
- **isoenergetic processes** = processes at constant energy

Most of the theory we develop will be for quasistatic processes. Just be careful about what happens when a process is not quasistatic, then we may have a change in entropy that is larger than what you would expect from work or heat alone.

6.9 Diffusive equilibrium — Chemical potential

We have now seen that two systems that are in *thermal contact*, that is they are allowed to exchange energy, are in equilibrium when the temperatures are the same. If the two systems also are in *mechanical contact*, that is they are allowed to exchange volume, they are in equilibrium

when also the pressures are the same. What if we also allow the two systems to exchange particles — what is the equilibrium condition then?

Two systems that are allowed to exchange particles are said to be in *diffusive contact*. We divide an isolated system into two parts, A and B, so that the total energy, volume and particle number is conserved:

$$E_A + E_B = E \Rightarrow E_A = E - E_B \quad dE_A = -dE_B \tag{6.125}$$

$$V_A + V_B = V \Rightarrow V_A = V - V_B \quad dV_A = -dV_B \tag{6.126}$$

$$N_A + N_B = N \Rightarrow N_A = N - N_B \quad dN_A = -dN_B \; . \tag{6.127}$$

This system is in equilibrium for the macrostate (E_A, V_A, N_A) that makes the total entropy maximum, which occurs when $dS_{TOT} = 0$, where the differential is

$$dS = \left(\frac{\partial S_A}{\partial E_A}\right)_{N_A,V_A} dE_A + \left(\frac{\partial S_A}{\partial V_A}\right)_{N_A,E_A} dV_A + \left(\frac{\partial S_A}{\partial N_A}\right)_{E_A,V_A} dN_A \tag{6.128}$$

$$+ \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B,V_B} dE_B + \left(\frac{\partial S_B}{\partial V_B}\right)_{N_B,E_B} dV_B + \left(\frac{\partial S_B}{\partial N_B}\right)_{E_B,V_B} dN_B = 0 \; , \tag{6.129}$$

which gives

$$dS = \left(\left(\frac{\partial S_A}{\partial E_A}\right)_{N_A,V_A} - \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B,V_B}\right) dE_A \tag{6.130}$$

$$+ \left(\left(\frac{\partial S_A}{\partial V_A}\right)_{N_A,E_A} - \left(\frac{\partial S_B}{\partial V_B}\right)_{N_B,E_B}\right) dV_A \tag{6.131}$$

$$+ \left(\left(\frac{\partial S_A}{\partial N_A}\right)_{E_A,V_A} - \left(\frac{\partial S_B}{\partial N_B}\right)_{E_B,V_B}\right) dN_A = 0 \; , \tag{6.132}$$

which again gives the following conditions:

$$\left(\frac{\partial S_A}{\partial E_A}\right)_{N_A,V_A} = \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B,V_B} \; , \tag{6.133}$$

$$\left(\frac{\partial S_A}{\partial V_A}\right)_{N_A,E_A} = \left(\frac{\partial S_B}{\partial V_B}\right)_{N_B,E_B} \; , \tag{6.134}$$

$$\left(\frac{\partial S_A}{\partial N_A}\right)_{V_A,E_A} = \left(\frac{\partial S_B}{\partial N_B}\right)_{V_B,E_B} \; . \tag{6.135}$$

We introduce this new condition as the chemical potential, μ, which has units of energy and must be the same in both systems when the two systems are in equilibrium:

$$\mu = -T \left(\frac{\partial S}{\partial N} \right)_{E,V} . \tag{6.136}$$

For two systems in diffusive equilibrium, the chemical potential is the same in both systems:

$$\mu_A = \mu_B . \tag{6.137}$$

We will build more intuition for the chemical potential in Chap. 9. Here we have simply introduced it for consistency.

Notice that the minus sign ensures that particles flow from systems with high chemical potential to systems with low potential. We can see this from $\partial S/\partial N$. The system with the largest value for this will gain the most entropy by adding particles. Hence, when we use minus this value we ensure that particles flow from high to low chemical potential.

6.9.1 Thermodynamical identity

With this new definition, all the quantities in the thermodynamic identity are now fully introduced. The thermodynamic identify therefore becomes:

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E} \right)_{V,N} dE + \left(\frac{\partial S}{\partial V} \right)_{E,N} dV + \left(\frac{\partial S}{\partial N} \right)_{E,V} dN \\ &= \frac{1}{T} dE + \frac{p}{T} dV - \frac{\mu}{T} dN , \end{aligned} \tag{6.138}$$

which gives:

$$T\, dS = dE + p\, dV - \mu\, dN , \tag{6.139}$$

and

$$dE = T\, dS - p\, dV + \mu\, dN . \tag{6.140}$$

This expression allows us to gain new intuition into the interpretation of the chemical potential. For a process where S and V is fixed, we see that

$$dE = \mu\, dN , \tag{6.141}$$

which tells us that

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S,V} . \tag{6.142}$$

The chemical potential therefore is change in energy per particle added to the system, when the particle is added so that the volume and the entropy of the system does not change. We will return to better, and more intuitive definitions of the chemical potential in Chap. 9.

6.9.2 Example: Chemical potential of the Einstein crystal

When we know the entropy of a system, $S = S(N, V, E)$, we can use this to find the chemical potential using (6.136). For the Einstein crystal, we found that the entropy was

$$S = Nk \ln E - Nk \ln(\epsilon N) + Nk \; . \tag{6.143}$$

The chemical potential is therefore

$$\begin{aligned} \mu &= -T \left(\frac{\partial S}{\partial N} \right)_{E,V} \\ &= -T \left(k \ln E + k - k \ln(\epsilon N) - k \right) \\ &= -T \left(k \ln E - k \ln \epsilon N \right) \\ &= -kT \ln \frac{E}{N} \epsilon \; . \end{aligned} \tag{6.144}$$

6.9.3 Example: Chemical potential of the ideal gas

Similarly, we can find the chemical potential for the ideal gas from the Sackur-Tetrode equation for the entropy:

$$\begin{aligned} S &= Nk \left(\frac{V}{N} \left(\frac{4\pi m E}{3h^2 N} \right)^{3/2} + \frac{5}{2} \right) \\ &= Nk \left(\ln V \left(\frac{4\pi m E}{3h^2} \right)^{3/2} - \ln N^{5/2} + \frac{5}{2} \right) \; . \end{aligned} \tag{6.145}$$

We take the derivative with respect to N to find the chemical potential:

$$\begin{aligned}\mu &= -T\left(\frac{\partial S}{\partial N}\right)_{E,V} \\ &= -Tk\left(\ln V\left(\frac{4\pi mE}{3h^2}\right)^{3/2} - \ln N^{5/2} + \frac{5}{2}\right) - TNk\frac{5}{2}\frac{1}{N} \\ &= -Tk\left(\ln\frac{V}{N}\left(\frac{4\pi mE}{3Nh^2}\right)^{3/2}\right) \\ &= -Tk\left(\ln\frac{V}{N}\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\right)\end{aligned} \tag{6.146}$$

where we have inserted that $E = (3/2)NkT$ for the ideal gas. We see that the chemical potential depends on the density: Increasing the density (N/V) while keeping T constant would mean that (V/N) becomes smaller, so that the chemical potential becomes larger — the system becomes more willing to give away particles.

Question

Clickers: Heat capacities

6.10 Simulations of the microcanonical system

We have already seen several approaches to simulation of a microcanonical system: The molecular dynamics approach, where energy conservation is ensured by Newton's laws of motion and conservative forces; Stochastic models, such as for the gas; and Monte Carlo methods, such as for the development of the Einstein crystal. Here, we will discuss these simulation methods systematically. Typically, we use simulation methods when we cannot find a theoretical solution for the system we are interested in. In that case, we need several approaches that can be adapted to the problem at hand.

6.10.1 Molecular dynamics simulations

The default molecular dynamics model is a simulation of a system in the microcanonical system: The number of particles, N, the volume, V, and the energy, E, is conserved throughout the simulation. We have seen

examples of simulations of a crystal and of a gas — and we can simulate any system you can think of using this method.

A molecular dynamics simulation produces a sequence of positions and velocities of all the particles in the system, a sequence of microstates as a function of time. However, since the time step in a molecular dynamics simulation is small, the practical upper limit of a simulation is currently around one microsecond. Now, we do not always know if the simulation is able to span over a reasonable set of microstates during this time span. For a liquid at high temperatures, a simulation will typically cover a wide enough area of the possible space of microstates, but for water in ice form or for a glass, only very small changes in the microstates will occur during the whole simulation. We call a system *ergodic* if the time sequence of microstates gives a representative view of the microstates of a system. In principle, we expect all systems to be ergodic, but in practice this may take too long for a simulation to be realized. We must therefore always critically evaluate if a molecular dynamics simulation is sufficiently close to equilibrium to provide relevant statistics.

What can we then use a molecular dynamics simulation for? We have already seen that we can use it to predict fluctuations in variables inside a system: We can use it to measure fluctuations in energy or the number of particles in two parts of a gas, or we can use it to measure fluctuations in energy in a solid. We can also measure temperature or pressure directly from the simulation. Typically, we will not find the temperature from the entropy, but instead we will measure the temperature directly from the average kinetic energy per particle (or degree of freedom) in the system. We will provide many examples of the use of molecular dynamics simulations throughout this book — hopefully these examples will provide you with ideas and inspiration for your own use of similar methods if you need them in your future work.

6.10.2 Stochastic simulations

Ideally, we would like to count all the microstates of a system, and use this to find the multiplicity, and from the multiplicity the entropy. However, this may not always be possible. We may then apply molecular dynamics simulations, which would provide a realistic model of the system, but only a small portion of the set of microstates of the system is probed. A different strategy would be to choose a random sequence of microstates, where each of the microstates are independent of each other, and then

use this set of microstates to estimate the probability and therefore also the multiplicity of a macrostate.

Stochastic model for the sum of dice. We have already applied this strategy to model statistical experiments. For example, if we were interested in describing the probabilty distribution for the sum of 1000 six-sided dice.

This particular problem can of course be solved analytically, but let us see how we would approach it using stochastic simulations. The idea is to generate many microstates, where each microstate corresponds to the result for each of the 1000 dice, and then use the set of microstates to find the probabilities for the macrostates. If the set is representative, we expect good results.

In this case, a microstate is a set $(n_1, n_2, \ldots, n_N)$ if N independent dice throws, so that each n_i is a random number from 1 to 6. The total number of microstates is 6^N, which is a very large number when $N = 1000$. How can we expect to sample this space reasonably? The number of macrostates is much smaller. A macrostate is the sum $Z = \sum_i n_i$, which can vary from N to $6N$. We perform this statistical experiment in python by generating 10000 experiments, each consiting of $N = 1000$ random numbers:

```
import numpy as np
import matplotlib.pyplot as plt
n = np.random.randint(0,6,(1000,10000))
Z = np.sum(n,axis=0)
plt.hist(Z)
```

and plot the resulting distribution of probabilties found from the histogram in Fig. 6.20. Notice that the histogram counts the number of results, $N(z_i)$, in intervals from z_i to $z_i + \Delta z_i$. If the bin size Δz_i is non-zero, we need to divide by the bin size as well as the number of samples to estimate the probability $P(z_i)$ (see Sect. 4.4.4 for a discussion of bin sizes and probabilities). Fig. 6.20 shows that the distributions $P(z)$ approaches a continuous distribution as the number of samples M increases. Indeed, when $M = 10^4$ the distribution already appears to give a good representation of the distribution, even though only a very small fraction of the number of microstates have been probed. However, since most of the randomly selected microstates will be in the region with the most probably macrostates, we still get good estimates for the probabilities of the most probable macrostates. On the other hand, the results for the very unlikely macrostates will be poor, since there will be few if any data points in this range of Z values.

Stochastic model for the Einstein crystal. We can apply a similar approach to sample the microstates of the Einstein crystal, and then use these sampled microstates to calculate the energy q_A of part A of the crystal and q_B of part B, similar to what we did previously. Let us assume that there are N_A oscillators in part A and N_B in part B, so that there are N oscillators in total, and that the total energy of the system is q. How can we generate a random microstate? We have q units of energy and N oscillators to distribute this energy randomly onto. We do this by placing each unit of energy in a randomly selected oscillator. That is, for each unit of energy, we select an oscillator at random, and increase the energy of this oscillator by one unit. When this has been done for all q energy units, we have generated a microstate, $(n_1, n_2, \ldots, n_N)$, where n_i is the number of energy units in oscillator i. The total energy is then $q = \sum_i n_i$. We choose part A to be the first N_A oscillators, that is, part A corresponds to $i = 1, \ldots, N_A$, and part B corresponds to the remaining N_B oscillators, $i = N_A + 1, \ldots, N_A + N_B$. We can therefore calculate q_A and q_B as

$$q_A = \sum_{i=1}^{N_A} n_i \ , \ q_B = \sum_{i=N_A+1}^{N_A+N_B} n_i \ . \tag{6.147}$$

We generate M such microstates and calculate the energy of part A, $q_A(j)$ for each such microstate j. From this sample of microstates we can estimate the probability distribution for q_A. The following program generates states and collects statistics following this stochastic modeling approach:

```
import numpy as np
import matplotlib.pyplot as plt
N = 200; q = 200; M = 1000
NA = int(N/2); NB = int(N/2);
EA = np.zeros(M); EB = np.zeros(M)
for istep in range (M):
    state = np.zeros(N)
    # Generate random state with given energy
    placeA = np.random.randint(0,high=N,size=q)
    for ip in range(len(placeA)):
        i = placeA[ip]; state[i] = state[i] + 1
    # Collect statistics for this state
    qA = np.sum(state[:NA]); qB = q-qA
    EA[istep] = qA; EB[istep] = qB
plt.hist(EA[:istep])
plt.xlabel('q_A'); plt.ylabel('N(q_A)')
```

The resulting distribution for $P(q_A)$ is shown in Fig. 6.20 for various values of M. We see that the distribution approaches a smooth function for

$P(q_A)$ when M is very small compared to the total number of microstates (which is the multiplicity of the Einstein crystal — a very huge number as we know).

Notice that for this particular case, we know the exact solution to the problem. However, in other situations you may not know the exact solution, and this form for Stochastic sampling may be an efficient way to learn about the probability for a macrostate in the system. Notice that it may not always be this simple to generate a state with a given total energy, and we may instead have to depend on a Monte Carlo type model for the system.

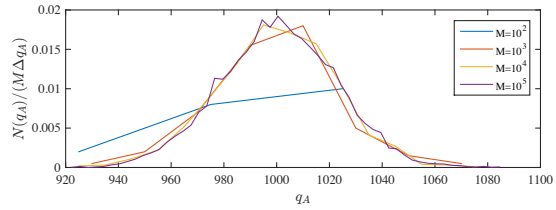


Fig. 6.20 Plot of the estimated probability density $P(q_A)$ for the energy q_A in the left half of an Einstein crystal for various sample sizes M. System with $N = 200$, $q = 2000$, $N_A = N_B = N/2$.

Stochastic model for the Ideal gas. We have already develop a stochastic model for the number of particles in a part of an ideal gas. A microstate of the gas may be the position of all the particles, or simply whether particle i is on the left side ($n_i = 1$) or not ($n_i = 0$). The microstate of the system would then be a list of the positions of all the particles, $(n_1, n_2, \ldots, n_N)$, and the number of particles on the left side is $n = \sum_i n_i$. Since all microstates are equally probable, we simply select each n_i to be 1 or 0 with probability 1/2 each, and generate a set of M microstates. This is what we did in Sect. 4.4.4, where we also studied the resulting distribution of n.

6.10.3 Monte Carlo simulations

In the stochastic model we generate a sequence of uncorrelated microstates: For each sample we generaet a new microstate randomly, ensuring no correlation with the previous state. However, in many cases it may be difficult to generate a microstate from scratch, but it may be simple to make a small perturbation to the microstate, or we may be interested in studying a slow change from an initial microstate of the system. For example, we may want to study how heat is redistributed

in an Einstein crystal, either locally or globally. This means that we start from a microstate $s = \{n_i\}$, which specifies the state, n_i, of each particle/degree of freedom i in the system, and then make small changes to the microstate, while ensuring that the total energy, particle number, or volume are not changed.

This suggested procedure will result in a sequence $s(j)$ of microstates. Every such microstate is indeed a possible microstate, and hence equally probable, but two subsequent microstates are not independently selected from the set of all microstates. Instead, they are correlated. This means that we need to sample many such states — a long sequence — before we make statistical predictions of the probability of a macrostate.

We call this method a Monte Carlo method for the microcanonical system. We have already introduced possible Monte Carlo method for the Einstein crystal, and we will extend this to also include a Monte Carlo method the introduces locations for all the oscillators and rules for possible transitions between microstates simulating local energy transport.

Global dynamics of the Einstein crystal. We already introduced a simple model for global redistribution of energy in an Einstein crystal in Sect. 6.3.3:

1. Generate initial, random microstate.
2. Select a random oscillator to send energy, i_1.
3. Select a random oscillator to receive energy, i_2.
4. If $n_{i,1} > 0$ then $n_{i,1} \to n_{i,1} - 1$ and $n_{i,2} \to n_{i,2} + 1$.
5. Loop to step 2.

As we saw in Fig. 6.9 this model generates and intersting dynamics and it allows us to study how the system approaches equilibrium and how it fluctuates around equilibrium. However, we may add further complexity (and realism) to the model, by insisting that each oscillator is located in a position in space, and that an oscillator only can transfer energy to one of its neighboring oscillator, thus opening for energy (heat) transport.

Local dynamics of the Einstein crystal. We add another level of complexity to the model, by introducing local interactions between the oscillators. All oscillators are placed on a square, two-dimensional grid. Oscillator i will be at position x_i, y_i. We follow the same algorithm as above: We select an oscillator at random, and transfer energy to another oscillator – if possible. However, we now only allow the oscillator to transfer energy to one of its neighbors. We select the neighbor at random

from the four nearest neighbors x_i+1, y_i, x_i-1, y_i, x_i, y_i+1, and x_i, y_i-1. This results in local energy transport as illustrated in Fig. 6.21. You can use the following program to study heat conduction in this simple model of thermal behavior in a solid.

```
# einsteinmclocal01.py
import numpy as np
import matplotlib.pyplot as plt
# Energy transport in an Einstein crystal
Ly = 40 # System size y-direction
LA = 40 # Length of system A in x direction
LB = 40 # Length of system B in x direction
Lx = LA + LB
NA = LA*Ly
NB = LB*Ly
qA = 3000 # Initial energy in system A
qB = 0    # Initial energy in system B
q = qA + qB # Total energy - conserved
N = NA + NB # Total number of oscillators
state = np.zeros((Lx,Ly),float) # 2d state matrix
# Generate initial, random states for A and B
for ia in range(qA):
    ix = np.random.randint(0,LA) # Rnd position from 1 to LA
    iy = np.random.randint(0,Ly) # Rnd position from 1 to Ly
    state[ix,iy] = state[ix,iy] + 1 # Add energy to this site
for ib in range(qB):
    ix = np.random.randint(0,LB)+LA # Rnd pos from LA+1 to LA+LB
    iy = np.random.randint(0,Ly)    # Rnd pos from 1 to Ly
    state[ix,iy] = state[ix,iy] + 1 # Add energy to this site
# Setup dynamic graphics
%matplotlib notebook
fig = plt.figure()
ax = fig.add_subplot(111)
# Simulate state development
nstep = 10000000 # nr of simulation steps
EA = np.zeros(nstep,float) # Energy per oscillator in system A
EB = np.zeros(nstep,float) # Energy per oscillator in system B
for istep in range(nstep):
    # Select an oscillator at random
    ix1 = np.random.randint(0,Lx)
    iy1 = np.random.randint(0,Ly)
    # Check if this oscillator has non-zero energy
    if (state[ix1,iy1]>0):
        # Find a random neighbor
        dx = 2*np.random.randint(0,2)-1    # +/-1 with equal prob
        ix2 = np.mod(ix1 + dx-1,Lx) # mod for periodic boundaries
        dy = 2*np.random.randint(0,2)-1    # +/-1 with equal prob
        iy2 = np.mod(iy1 + dy-1,Ly) # mod for periodic boundaries
        # Transfer energy from (ix1,iy1) to (ix2,iy2)
        state[ix2,iy2] = state[ix2,iy2] + 1
        state[ix1,iy1] = state[ix1,iy1] - 1
    if (np.mod(istep,1000)==0): # Display system at regular intervals
```

```
        ax.clear()
        ax.imshow(state.T)
        fig.canvas.draw()
```

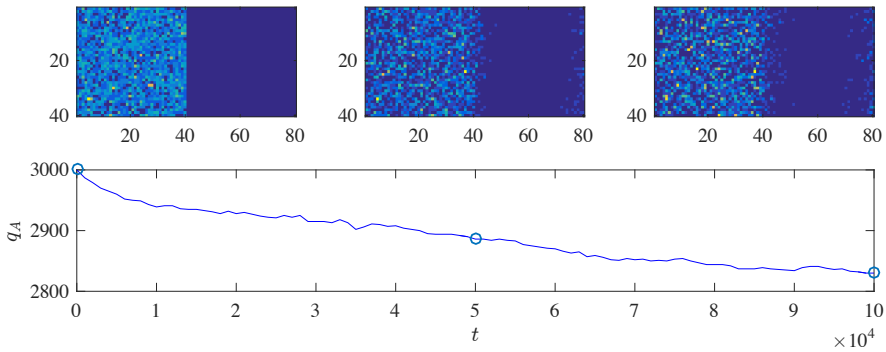


Fig. 6.21 Illustration of the time development of two two-dimensional Einstein crystals in thermal contact.

6.11 Summary

- The **multiplicity** of a macrostate is the number of microstates in the macrostate
- The **Einstein crystal** is a simple model for a solid where the atoms do not interact and where each atom acts as a harmonic oscillator in three dimensions
- For two Einstein crystals in contact we describe the systems by the energy distribution between the two systems, described by the energy q in one of the systems.
- The macrostates for the two Einstein crystals do not have the same number of microstates — some of the macrostates have vastly more number of microstates that others.
- A system develops towards the most probably macrostate
- The number of microstates for a given macrostate is called the multiplicity of the macrostate
- The multiplicity of an isolated system is $\Omega(N, V, E)$.
- The entropy of an isolated system is defined as $S = k \ln \Omega(N, V, E)$
- The temperature of an isolated system is $(1/T) = (\partial S/\partial E)_{N,V}$
- The pressure of an isolated system is $p = T\,(\partial S/\partial V)_{E,N}$

- The chemical potential of an isolated system is $\mu = -T\,(\partial S/\partial N)_{E,V}$
- The first law of thermodynamics states that $\Delta E = Q + W$
- The second law of thermodynamics states that the entropy of an isolated system must increase of be constant in any process. In equilibrium the entropy is at a maximum.
- The thermodynamic identity is $T\,dS = dE + p\,dV - \mu\,dN$
- The heat capacity of a system at constant volume is $C_V = (\partial E/\partial T)_V$

6.12 Exercises

Exercise 6.1: Microstates and macrostates of a random walk

The behavior of an atom or a molecule in a fluid with many atoms or molecules can be considered to behave like a random walker. We will here address the micro- and macrostates of a one-dimensional random walker as a model for atomic (or molecular) diffusion.

The position Z_N of a random walker after N steps can be found by adding each of the steps, X_i, $i = 1, \ldots, N$:

$$Z = \sum_{i=1}^{N} X_i \; . \tag{6.148}$$

Here, we will assume that each step can be either to the left, $X_i = -1$, or to the right, $X_i = +1$, with equal probability.

The microstate for such a walker will correspond to specifying each of the steps, X_i, whereas a macrostate will refer to the position Z_N after N steps.

a) Write down all the microstates for a $N = 4$ steps random walk.

b) Find the probabilities for each of the possible macrostates based on the list of microstates.

c) Compute the multiplicity of each of the macrostates using the bionomial formula and compare with your results from counting all possible states.

Exercise 6.2: Flipping 30 coins

You flip 30 fair coins. A possible macrostate could be defined as the number, n, of heads.

a) How many possible outcomes (microstates) are there of this experiment?

b) How many ways can you get $n = 15$?

c) What is the probability of $n = 15$?

d) What is the probability of $n = 20$?

e) What is the probability of $n = 25$?

f) What is the probability of $n = 30$?

g) Make a program to plot the probability, $P(n)$, to get n.

Exercise 6.3: Stirling's approximation

a) Use python to check the accuracy of Stirling's approximation for $N = 50$ and $N = 100$.

b) Make a plot to illustrate the accuracy of $\ln N! = N \ln N - N$ as a function of N.

Exercise 6.4: Entropy

In a generic system with N particles the multiplicity is given as $\Omega(E, N) = CE^{3N/2}$, where C is a constant.

a) Show that $E = (3/2)NkT$.

b) Show that

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_N < 0 \,, \tag{6.149}$$

and discuss what this means.

Exercise 6.5: Gas in one- and two-dimensions

a) Find an expression for the multiplicity of an ideal gas of N indistinguishable atoms, where each atom only can move along a one-dimensional tube, such as along a narrow nanotube.

b) Find an expression for the multiplicity of an ideal gas of N indistinguishable atoms, where each atom only can move along a two-dimensional flat sheet, such as for electrons in a graphene sheet.

The canonical system

7

We have so far studied isolated systems, systems with constant N, V, and E, such as the systems we studied using molecular dynamics. Such systems are called *microcanonical* systems. Microcanonical systems are theoretically useful, but they are not equally practical from an experimental perspective. If you do an experiment inside a box in the lecture hall, you would typically expect the system inside the box to be in thermal equilibrium with the lecture hall — you would expect the system to be at the same temperature as the lecture hall. Thus, we are often interested in systems that are kept at constant temperature and not at constant energy. This requires contact with a large thermal reservoir that can exhange energy with the system so that the system can keep an (approximately) constant temperature without changing the temperature of the reservoir significantly. For this system N, V, and T are constant, and we call this a *canonical* systems.

Here, we want to develop the statistical mechanics and the thermodynamics of *canonical systems*, systems with given N, V, and T. We will assume that the system may be in any of its states, i, and we will first find the probability for the system to be in a state i when it is in equilibrium with a large reservoir. We will do this using arguments from the microcanonical theory. We will demonstrate how we can link the microscopic scale to the macroscopic scale for canonical systems using the partition function. We will demonstrate that the second law of thermodynamics — the the entropy of an isolated system must increase — is replaced by a similar principle for the canonical system, that Helmholtz free energy will decrease and be minimal in equilibrium.

7.1 Boltzmann statistics in the Einstein crystal

Let us start from a simple example system that we master well — the harmonic oscillator. We will address the motion of a single oscillator in contact with a large reservoir so that the temperature of the oscillator is constant. The reservoir will consist of a huge amount of oscillators in contact with the single oscillator. This is a system we now are proficient in discussing and solving — we can therefore use our experience with the harmonic oscillator system to gain insight into the canonical system.

7.1.1 Einstein crystal – numerical example

We already know how to study a system consiting of one harmonic oscillator in contact with a system of may oscillators as illustrated in Fig. 7.1, because this system corresponds to two Einstein crystals in contact, a system for which we already have found an exact solution!

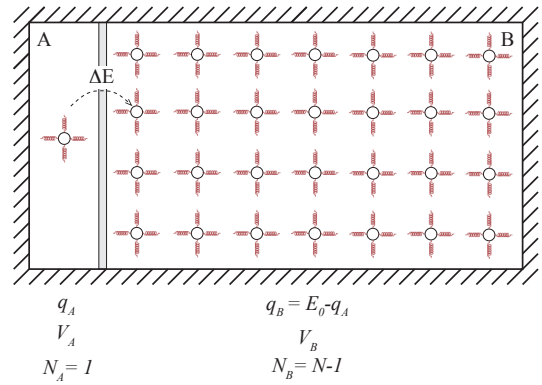


Fig. 7.1 Illustration of a single harmoic oscillator (part A) in contact with a huge reservoir consisting of harmonic oscillators (part B). The total system is isolated, so that for the total system N, V, E are given, whereas for part A only V_A and N_A are given, whereas q_A may vary.

Model system. Our previously developed model of an Einstein crystal divided into two parts, A and B, can be used to represent the canonical system in contact with a large reservoir. Part A is the canonical system and part B is the reservoir, where $N_A = 1$ and $N_B = N - N_A = N - 1 \gg N_A$. The two parts A and B can exchange energy, but the total energy is conserved, $q_A + q_B = q$, as is the number of particles, N_A, N_B, and the volumes, V_A and V_B. The total system, consisting of system A and system B, is an *isolated* system.

The probability for system A to have energy q_A, which is the same as the probability for system A to be in the state q_A, since for this particular

system there is one state per energy. The probability to find system A in state q_A is

$$P(q_A) = \frac{\Omega(q_A, N_A)\,\Omega(q_B, N_B)}{\Omega_T}\,, \tag{7.1}$$

where q_A can range from 0 to q, and $\Omega(q_A, N_A)$ is the multiplicity of an Einstein crystal with N_A oscillators and q_A energy, and is given by

$$\Omega(q_A, N_A) = \begin{pmatrix} q_A + N_A - 1 \\ q_A \end{pmatrix}\,. \tag{7.2}$$

The total multiplicity Ω_T is the sum of all the multiplicities for different values of q_A.

Calculation of probabilities. The probabilities $P(q_A)$ are plotted using the following program. The resulting plot in Fig. 7.2a shows that the probability falls off very rapidly as q_A increases.

```
# Einstein solid NA - NB probability
import numpy as np
import matplotlib.pyplot as plt
from scipy.special import comb
NA = 1
NB = 199
q = 200
N = NA + NB
# Find multiplicity of all macrostates
omegaA = np.zeros(q+1)
qAvalue = np.zeros(q+1)
omegaB = np.zeros(q+1)
omegaTOT = np.zeros(q+1)
# Loop through all macrostates and find multiplicity
for istate in range(0,q):
    qA = istate
    qAvalue[istate] = qA
    omegaA[istate] = comb(qA+NA-1,qA)
    qB = q - qA
    omegaB[istate] = comb(qB+NB-1,qB)
    omegaTOT[istate] = omegaA[istate]*omegaB[istate]
PA = omegaTOT/sum(omegaTOT)
plt.plot(qAvalue,PA,'-o')
plt.xlabel('q_A'), plt.ylabel('P(q_A)')
```

Characterizing the plot of $P(q_A)$. How can we characterize the functional form of $P(q_A)$? There are several tricks from experimental and observational physics on how to determine the underlying functional form of an observed probability distribution. Based on ideas from statistics we could expect it to be of an exponential or a power-law form. An

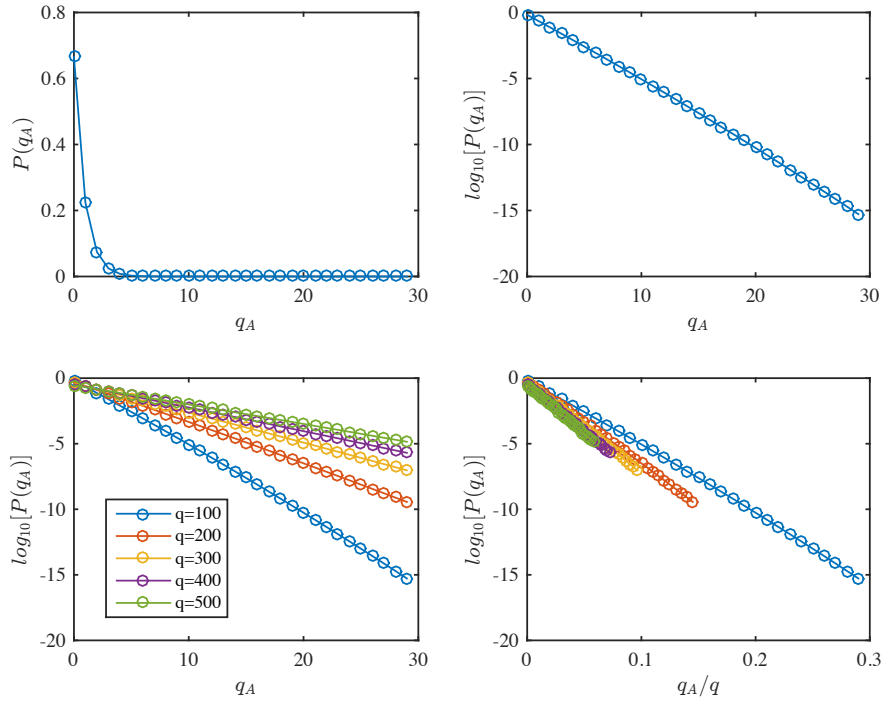


Fig. 7.2 Probability $P(q_A, N_A = 1)$ for the Einstein oscillator with $q = 200$, $N = 200$.

exponential form, means that the function $P(q_A)$ will be on the form

$$P(q_A) = Ce^{-q_A/q^*} \,, \tag{7.3}$$

where C and q^* are constants. How can we test if the measured distribution has this form? We notice (and this is a common trick that it is useful for you to learn) that if we take the logarithm of (7.1.1), we get:

$$\ln P(q_A) = \ln C - \frac{q_A}{q^*} \,, \tag{7.4}$$

which is a linear function of q_A. This means that if we plot $\ln P$ along the y-axis and q_A along the x-axis, we expect the plot to be linear if the functional form in (7.1.1) is correct. Fig. 7.2b shows the resulting plot. Indeed, the curve is approximately linear — our theory formulated in (7.1.1) seems to be reasonable!

Effect of temperature. Maybe our theory becomes better if we increase q and therefore the number of possible values for q_A? We try this by

plotting curves for $P(q_A)$ for several values of q using the following program:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.special import comb
NA = 1
NB = 199
qq = [150,200,250,300,400,450,500]
nqq = len(qq)
for iqq in range(0,nqq):
    q = qq[iqq]
    N = NA + NB
    # Find multiplicity of all macrostates
    nstate = q+1
    omegaA = np.zeros(nstate)
    qAvalue = np.zeros(nstate)
    omegaB = np.zeros(nstate)
    omegaTOT = np.zeros(nstate)
    # Loop through all macrostates and find multiplicity
    for istate in range(0,nstate):
        qA = istate-1
        qAvalue[istate] = qA
        omegaA[istate] = comb(qA+NA-1,qA)
        qB = q - qA
        omegaB[istate] = comb(qB+NB-1,qB)
        omegaTOT[istate] = omegaA[istate]*omegaB[istate]
    PA = omegaTOT/sum(omegaTOT)
    j = np.where(qAvalue<40)
    plt.subplot(2,1,1)
    plt.plot(qAvalue[j],np.log10(PA[j]),'-o',label="q =" +str(q))
    plt.xlabel('q_A'), plt.ylabel('log_{10}[P(q_A)]')
    plt.subplot(2,1,2)
    plt.plot(qAvalue[j]/q,np.log10(PA[j]),'-o',label="q =" +str(q))
    plt.xlabel('q_A/q'), plt.ylabel('log_{10}[P(q_A)]')
plt.subplot(2,1,1); plt.legend()
plt.subplot(2,1,2); plt.legend()
```

The resulting curves for $\ln P(q_A)$ are shown in Fig. 7.2c. Hmmm. Interesting. The slope of the curves, and therefore the q^* value, apparently depends on the total energy q in the system. How is this dependence? We could measure the slope and plot it as a function of q to find out that it appears that q^* appears to be proportional to q. And we can again check this hypothesis by plotting $\ln P(q_A)$ as a function of q_A/q. If q^* is proportional to q we would then expect all the curves for various q-values to fall onto a common curve. The resulting plot in Fig. 7.2d shows that this is approximately the case. You can convince yourself that this is indeed the case, and that the approximation becomes better for larger N- and q-values, by redoing the plot for larger values of N and q.

Now, we have shown how $P(q_A)$ depends on the total energy q in the system. But if we think of the system as a single oscillator – oscillator A – in contact with a large system of oscillators, we can think of the rest of the oscillators as a large reservoir with a given temperature T_B describing system B. Previously, we found that that the temperature of an Einstein crystal in the limit when q and N is large is given by $E = q\Delta\epsilon = NkT$, so that the temperature in system B is proportional to q, that is, $q = cT$, where c is a constant.

We can therefore sum up the results of our numerical study and the approximate theoretical model we have introduced to describe the results:

$$P(q_A) = Ce^{-\frac{q_A}{cT}} , \tag{7.5}$$

where C and c are constants. Now, let us see if we can make a theory for this behavior.

7.1.2 Einstein crystal in heat bath — Theory

We have a complete methodology from the microcanonical system which we can use to address such a problem: What is the probability for a single oscillator to be in a particular state i [1] with energy ϵ_i?

We consider a big system, the reservoir or heat bath R, for example an Einstein crystal consisting of many oscillators, in contact with the system S, here consisting of a single oscillator as illustrated in Fig. 7.3. (We called the system A and the reservoir B in the numerical example above). We will assume that the reservoir is so large that it is not affected by the system: If we place a hot cup of water in the auditorium, it will cool down to the temperature of the auditorium without increasing the temperature of the auditorium significantly. The air and other materials in the auditorium is an example of a large heat bath.

System and reservoir. The composite system consisting of the system and reservoir is isolated and is characterized by constant E, V, and N. The composite system is therefore a microcanonical system and all microstates in this system are equally probable. We can therefore find the probability for the composite system to be in a particular microstate and we can use this to find the probability for the system S to be in a state i. If the system S is in a state with energy ϵ_i the reservoir has

[1] We use the term state here in the quantum mechanical meaning. A state i corresponds to a quantum mechanical state, $|i\rangle$ or Ψ_i, and has an associated energy ϵ_i. Notice that several states may have the same energy.

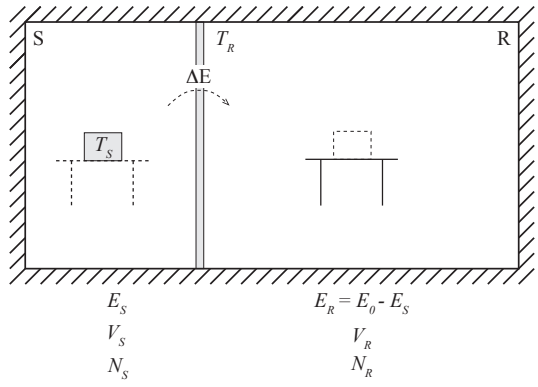


Fig. 7.3 Illustration of a system S in thermal contact with a large reservoir R. The composite system consisting of S and R is isolated.

energy $E_0 - \epsilon_i$ so that the total energy is $E = \epsilon_i + (E_0 - \epsilon_i) = E_0$. The probability for the system to be in state i is given by the multiplicity of the whole system, consisting of both system S and reservoir R:

$$P(i) = \frac{\Omega_S\, \Omega_R}{\sum_i \Omega_S \Omega_R} \ , \tag{7.6}$$

where the sum is for normalization of the probability (to ensure the sum of probabilities of all possible states is 1 — the system must be in one of the possible states).

Probability of state i. Since the system S is in state i the multiplicity of this state is 1: We have specified the state so it is only one way it can be in this state. (Note that state here is not the same as the energy, because several states can have the same energy and still be separate states. It is therefore important to remember that we here consider states and not only the energy.)

What is the multiplicity of the reservoir? For a given microstate i of the system S, the reservoir R can be in many microstates. The multiplicity of the macrostate where system R has energy $E_0 - \epsilon_i$ is $\Omega(E_0 - \epsilon_i)$, where E_0 is the total energy of the system. The probability of the reservoir is therefore

$$P(i) = \frac{\Omega_S \Omega_R}{\sum_i \Omega_S \Omega_R} = C\ (1\, \Omega_R) \ , \tag{7.7}$$

To simplify the algebra we work with the logarithm of the probability:

$$\ln P(i) = \ln C + \ln \Omega_R(E_0 - \epsilon_i) \ . \tag{7.8}$$

We can expand $\ln \Omega_R$ around E_0 since $\epsilon_i \ll E_0$ (the heat bath is large):

$$\ln \Omega_R(E_0 - \epsilon_i) = \ln \Omega_R(E_0) + \left(\frac{\partial \ln \Omega_R(E)}{\partial E}\right)_{V,N} (-\epsilon_i) + \mathcal{O}(\epsilon_i^2) \ , \quad (7.9)$$

Where we can now use that $S = k \ln \Omega$ for the reservoir:

$$\left(\frac{\partial \ln \Omega_R(E)}{\partial E}\right)_{V,N} (-\epsilon_i) = -\left(\frac{(1/k)\partial S_R}{\partial E}\right)_{V,N} \epsilon_i \ . \quad (7.10)$$

Now, we use that the temperature T_R of the reservoir is defined as $1/T_R = (\partial S_R/\partial)_{N,V}$ when the system is in equilibrium. In equilibrium the temperature in the reservoir and the temperature in the system will be the same, and we can simply use $T = T_R = T_S$ for the temperature in both systems. The multiplicity is therefore

$$\ln \Omega_R(E_0 - \epsilon_i) = \ln \Omega_R(E_0) - \frac{\epsilon_i}{kT} + \mathcal{O}(\epsilon_i^2) \ , \quad (7.11)$$

The probability is therefore

$$P(i) = \frac{\Omega_S \, \Omega_R}{\sum \Omega_S \, \Omega_R} \, e^{-\epsilon_i/kT} = \frac{\Omega_R}{\sum \Omega_R} \, e^{-\epsilon_i/kT} = \frac{1}{Z} e^{-\epsilon_i/kT} \ . \quad (7.12)$$

Where Z is determined by the normalization rule for probabilities:

$$\sum_i P(i) = 1 = \frac{1}{Z} \sum_i e^{-\epsilon_i/kT} \ . \quad (7.13)$$

and therefore

$$Z = Z(N, V, T) = \sum_i e^{-\epsilon_i/kT} \ , \quad (7.14)$$

The function Z is called the *partition function.*

7.1.3 Comparison with numerical model

This theory is only valid for large reservoirs. We can therefore compare the theory with our numerical results for the Einstein crystal from above, but we will only expect the match to be good when N and q are large. We notice that the function we found to represent the numerical data was

$$P(q_A) = Ce^{-q_A/cT} \ . \quad (7.15)$$

The state in a system with a single harmonic oscillator is described by the state i, so that the energy levels are $\epsilon_i = \Delta\epsilon i$, where we see that the states can be enumerated by i or by q_A. The theoretical results for the

probabilities for Boltzmann statistics is therefore:

$$P(q_A) = \frac{1}{Z} e^{-(q_A\,\Delta\epsilon)/(kT)} \,. \tag{7.16}$$

Where we see that the constant $c = k/\epsilon$ and that the prefactor C is Z (which also is a function of T, but it does not depend on q_A). The theory therefore both explains the observed numerical results, and provides us with a better understanding for the behavior of the system.

7.2 The partition function and Boltzmann statistics

The preceeding discussion is completely general and not limited to an Einstein crystal. For a system S in thermal contact with a huge reservoir R, we can assume that in equilibrium the temperature, T_S, in system S is given as the temperature, T_R, of the reservoir, $T_S = T_R = T$. The system S is therefore a canonical system with given N, V, T. The system S may be in a set of possible states i with energies ϵ_i, corresponding to the quantum mechanical states of the system. For such as system, the probability for the system to be in a state i is given by the Boltzman factor:

Boltzman probabilities: The probability for a system S described by N, V, T, to be in a state i with energy ϵ_i is given as

$$P(\epsilon_i) = \frac{1}{Z} e^{-\epsilon_i/kT} \,, \tag{7.17}$$

where the system is *in equilibrium*.

Partition function: The sum $Z = Z(N, V, T)$ is called the partition function. The sum is over *all the states* i of the system:

$$Z(N, V, T) = \sum_i e^{-\epsilon_i/kT} \,. \tag{7.18}$$

The symbol Z comes from the German "Zustandsumme" – sum of all *states*. A common mistake is to assume that the sum is over all the

energies of a system, and not over all the states. However, energies may be degenerated — there can be several states that have the same energy — and in this case you need to include one term in the partition function sum for each of the states in the system. Alternatively, you can introduce the degeneracy $g(\epsilon)$ of a system, the number of states of energy ϵ, and then sum over all possible energies instead

$$Z(N,V,T) = \sum_{\epsilon} g(\epsilon)\, e^{-\epsilon/kT} \ . \tag{7.19}$$

7.2.1 Average values and fluctuations

How can we use this new theory to connect the microscopic description of the system given in the probabilities for the states with the macroscopic description of the system given by thermodynamics? We can do this by calculating averages of the thermodynamic quantities.

We call a system with constant N, V, E a **microcanonocal** system, and a system with constant N, V, T – that is a system in contact with a large heat reservoir of temperature T – a **canonical** system. In the microcanonical system the total energy is given, and we can measure the temperature T. The temperature T will fluctuate: If we measure temperature in a molecular dynamics simulation, we will get a temperature that varies in time, $T(t)$. However, we also found that for realistic values of N – which typically are around 10^{20} – the values for T are very sharp: Deviations beyond 10^{-10} of its average value are very unlikely. For a canonical system, the temperature T is constant, but the energy E will fluctuate. We can therefore find the average value (and later the fluctuations) of the energy and use this to connect the microscopic theory of Boltzmann statistics with the macroscopic theory of thermodynamics.

The average value (or expectation value) of a quantity Q_i that depends on the state i follows the definition of averages in Chap. 4.

The **average** of a quantity Q_i, which depends on the state i of a canonical system with given N, V, T, is:

$$\overline{Q}_i = \sum_i P(i)\, Q_i = (1/Z) \sum_i Q_i\, e^{-\epsilon_i/kT} \tag{7.20}$$

The average of the energy E is therefore

$$\bar{E} = \langle E \rangle = \sum_i \epsilon_i P(\epsilon_i) \ , \tag{7.21}$$

where the sum is over all states i of the system S. Since we have an expression for $P(i)$ we can simplify this expression to:

$$\bar{E} = \sum_i \epsilon_i \frac{1}{Z} e^{-\epsilon_i/kT} \ . \tag{7.22}$$

Here, we will employ a common trick, which is to introduce $\beta = 1/(kT)$. The equation then becomes

$$\bar{E} = \frac{1}{Z} \sum_i \epsilon_i e^{-\beta\epsilon_i} \ . \tag{7.23}$$

The trick is now to recognize that the sum is the derivative of a sum which is simpler to calculate:

$$\sum_i \epsilon_i e^{-\beta\epsilon_i} = -\frac{d}{d\beta} \sum_i e^{-\beta\epsilon_i} = -\frac{dZ}{d\beta} \ . \tag{7.24}$$

This provides us with a coupling between the microscopic and the macroscopic descriptions of the system:

The average energy of a canonical system with states i with energies ϵ_i is

$$\bar{E} = -\frac{1}{Z}\frac{d}{d\beta} Z = -\frac{d \ln Z}{d\beta} \ , \tag{7.25}$$

where $Z = \sum_i e^{-\beta\epsilon_i}$ and $\beta = 1/(kT)$.

This expression allows us to calculate the energy E directly from the partition function. Now we have the tools to connect the microscopic and the macroscopic also in systems with constant N, V, and T. Let us look at some example applications.

7.2.2 Example: Two-state system

Let us start with a very simple case: a two state system in contact with a thermal reservoir with temperature T as illustrated in Fig. 7.4. We have an atom/molecule in contact with a thermal bath with temperature T.

The system may be in two states with energies ϵ_0 and ϵ_1. What is the internal energy and the heat capacity of this system?

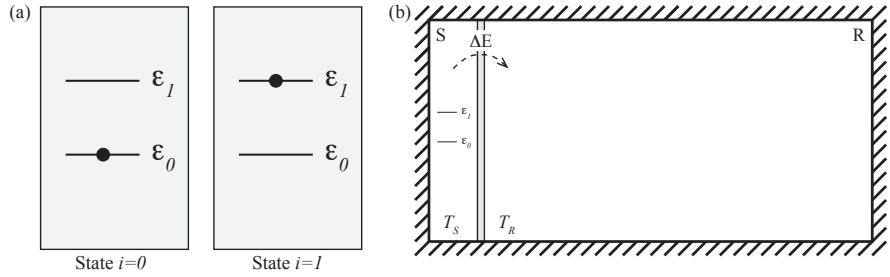


Fig. 7.4 **a** Illustration of two-state system. **b** Two-state system in contact with a large thermal reservoir with temperature T_R.

Finding the partition function. To find the behavior of such a system, our first step is always to find the partition function, Z. We may choose our base energy level as we like, so let us choose $\epsilon_0 = 0$ to simplify the calculations.

$$Z = \sum_i e^{-\epsilon_i} kT \ , \tag{7.26}$$

where the sum is over all the *states*. In this case, all the states corresponds to all the energies.

$$Z = e^{-0/kT} + e^{-\epsilon/kT} = 1 + e^{-\epsilon/kT} \ . \tag{7.27}$$

Finding the average energy. The average energy is

$$\bar{E} = \frac{1}{Z} \sum_i e^{-\epsilon_i/kT} \, \epsilon_i = \frac{1}{Z} \left(0 \, e^{-0/kT} + \epsilon \, e^{-\epsilon/kT} \right) = \epsilon \frac{e^{-\epsilon/kT}}{1 + e^{-\epsilon/kT}} \ . \tag{7.28}$$

Plotting the average energy. The energy as a function of temperature is shown in Fig. 7.5. The results were plotted by the following program:

```
import numpy as np
import matplotlib.pyplot as plt
x = np.linspace(0,3,1000)
E = np.exp(-1/x)/(1+np.exp(-1/x))
plt.subplot(2,1,1)
plt.plot(x,E)
plt.xlabel('kT/\epsilon'), plt.ylabel('U/\epsilon')
plt.subplot(2,1,2)
CV = np.diff(E)/np.diff(x)
xmid = 0.5*(x[1:]+x[0:-1])
```

```
plt.plot(xmid,CV)
plt.xlabel('kT/\epsilon'), plt.ylabel('C_V/\epsilon')
```

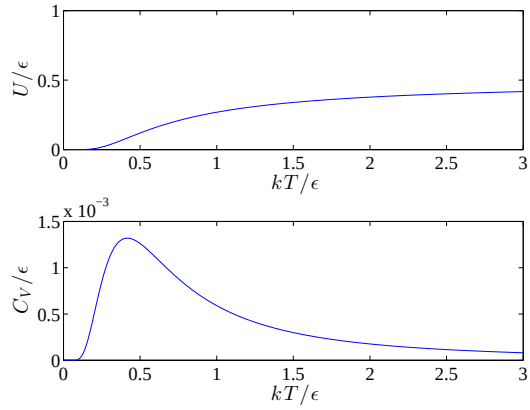


Fig. 7.5 Two-state system. **a** Plot of energy as a function of temperature. **b** Plot of heat capacity as a function of temperature.

Limiting behavior. We check the two limits analytically:

- What happens when $T \to \infty$? In this case, the term $e^{-\epsilon/kT} \to 0$. The probabilities for both state are therefore equal and equal to $1/2$, and and the average energy is $\epsilon/2$.
- At low temperatures, the only possible state is the state with the lowest energy, $P(1) = 1$ and $P(2) = 0$, and the average energy is also zero.

Heat capacity. Finally, we can find the heat capacity, C_V, at constant V from

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{N,V} , \tag{7.29}$$

either by numerical derivation of the $E(T)$ curve, of by analytical derivation of

$$C_V = \epsilon \frac{\partial}{\partial T} \frac{1}{e^{\epsilon/kT} + 1} \tag{7.30}$$

$$= \epsilon \, (-1) \, \frac{1}{\left(e^{\epsilon/kT} + 1\right)^2} \, e^{\epsilon/kT} \left(-\frac{1}{kT^2}\right) \tag{7.31}$$

$$= k \left(\frac{\epsilon}{kT}\right)^2 \frac{e^{\epsilon/kT}}{\left(e^{\epsilon/kT} + 1\right)^2} \, . \tag{7.32}$$

7.2.3 Example: Fluctuations and heat capacity

There is a general result, which is interesting from a computational perspective, which allows us to calculate the heat capacity from the fluctuations in the energy.

Average energy. We found above in (7.25) that the average of the energy is found as:

$$\bar{E} = -\frac{\partial}{\partial\beta} \ln Z \ , \tag{7.33}$$

where $Z = \sum_i = \sum_i e^{-\epsilon_i/kT}$ and $\beta = 1/(kT)$.

Heat capacity. Similarly, we can express the heat capacity as the derivative of $\bar{E}$ with respect to β:

$$C_V = \left(\frac{\partial \bar{E}}{\partial T}\right) = \frac{\partial \bar{E}}{\partial \beta}\frac{\partial \beta}{\partial T} = \frac{1}{kT^2}\left[\frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} - \frac{1}{Z^2}\left(\frac{\partial Z}{\partial \beta}\right)^2\right] \ , \tag{7.34}$$

where we used that

$$\frac{\partial \beta}{\partial T} = \frac{\partial}{\partial T}\frac{1}{kT} = -\frac{1}{kT^2} \ . \tag{7.35}$$

Fluctuations in energy. Now, we see that

$$\bar{E^2} = \sum_i \epsilon_i^2 e^{-\beta\epsilon_i} = \frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} \ , \tag{7.36}$$

Heat capacity and energy fluctuations. This gives us the relation:

$$C_V = \frac{1}{kT}\left(\bar{E^2} - \bar{E}^2\right) \ . \tag{7.37}$$

This means that if we can measure the fluctuations in E we can use this to find the heat capacity. Nice! This is particularly useful if we study a molecular/atomic system!

7.2.4 Sharpness of the canonical system

Based on this result from the fluctuations, we can also determine the sharpness of the thermodynamic averages for a canonical system.

From (7.37) we know that the variation in the energy of a system, ΔE, is proportional to the square root of the heat capacity:

$$\Delta E = \sqrt{\bar{E^2} - \left(\bar{E}\right)^2} = \sqrt{C_V kT}\,, \tag{7.38}$$

Now, we expect both the energy and the heat capacity (which is the derivative of energy with temperature) to be **extensive** properties — properties that are proportional to the number of particles N. This means that the relative fluctuation in E is given as

$$\frac{\Delta E}{\bar{E}} \propto \frac{\sqrt{N}}{N} \propto (N)^{-1/2}\ . \tag{7.39}$$

In the limit of realistic N, typically on the order of 10^{20} – we see that the relative error is 10^{-10}, which is generally not measurable. The thermodynamic properties are therefore sharp.

7.2.5 Example: Einstein crystal in the canonical system

We can use the new formalism to address the behavior of an Einstein crystal in contact with a large, thermal reservoir, for example, a vibrating diatom in thermal contact with a large gas as illustrated in Fig. 7.6. First, we find the probability for a single oscillator to be in a given state, and then we find the average energy of a single oscillator.

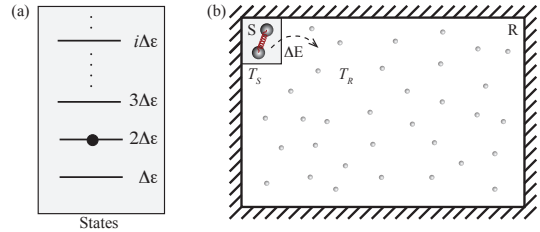


Fig. 7.6 Illustration of a single harmonic oscillator in contact with a large reservoir. **a** Illustration of the states of the diatomic harmonic oscillator. **b** Illustration of the oscillator in contact with a large reservoir.

States of the Einstein crystal. The states of an harmonic oscillator are

$$\epsilon_i = \epsilon_0 + i\,\Delta\epsilon\,, \tag{7.40}$$

where $\epsilon_0 = \hbar\omega/2$ and $\Delta\epsilon = \hbar\omega$. We choose the energy level so that $\epsilon_0 = 0$, which is simply a small shift in energy, so that the energy levels are

$$\epsilon_i = i\,\Delta\epsilon\,, i = 0, 1, 2, \ldots\ . \tag{7.41}$$

Partition function for the Einstein crystal. The partition function is then given as the sum over all states:

$$Z = \sum_{i=0}^{\infty} e^{-\epsilon_i/kT} \ . \tag{7.42}$$

We can simplify the notation by introducing the notation $\beta = 1/(kT)$. The partition function is then:

$$Z = \sum_{i} e^{-\beta\,\epsilon_i} = \sum_{n=0}^{\infty} e^{-\beta\,\Delta\epsilon\,i} \ , \tag{7.43}$$

which is a geometric series, with sum

$$Z = \frac{1}{1 - e^{-\beta\,\Delta\epsilon}} \ . \tag{7.44}$$

Energy of the Einstein crystal. We can find the average energy either directly from the sum or from the expression in (7.25):

$$\bar{E} = -\frac{\partial}{\partial\beta} \ln Z = \frac{\partial}{\partial\beta} \ln\left(1 - e^{-\beta\,\Delta\epsilon}\right) \tag{7.45}$$

$$= \frac{1}{1 - e^{-\beta\,\Delta\epsilon}}\, e^{-\beta\,\Delta\epsilon}\, \Delta\epsilon = \Delta\epsilon \frac{1}{e^{\beta\,\Delta\epsilon} - 1} \ . \tag{7.46}$$

We insert $\beta = kT$, getting:

$$\bar{E} = \Delta\epsilon \frac{1}{e^{\Delta\epsilon/kT} - 1} \ . \tag{7.47}$$

we can introduce the characteristic temperature $\theta_E = \Delta\epsilon/k$ — the Einstein temperature for the oscillator, which is a constant for a given $\Delta\epsilon$ — getting:

$$\bar{E} = \Delta\epsilon \frac{1}{e^{\theta_E/T} - 1} \ . \tag{7.48}$$

Plot of the energy and the heat capacity of an Einstein crystal. It is illustrative to plot this expression as a function of T/θ_E using:

```
import numpy as np
import matplotlib.pyplot as plt
x = np.linspace(0,1.5,1000)
E = 1.0/(np.exp(1.0/x)-1.0)
plt.subplot(2,1,1)
plt.plot(x,E)
plt.xlabel('T/\theta_E'), plt.ylabel('E/\epsilon')
```

```
plt.subplot(2,1,2)
CV = np.diff(E)/np.diff(x)
xx = 0.5*(x[1:]+x[0:-1])
plt.plot(xx,CV)
plt.xlabel('T/\theta_E'), plt.ylabel('C_V/\epsilon')
```

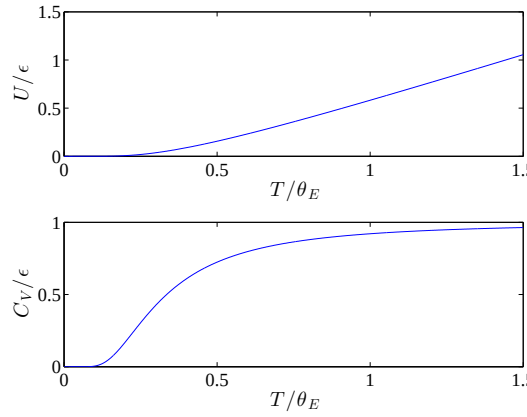


Fig. 7.7 A single Einstein crystal in contact with a reservoir with temperature T. **a** Plot of the energy E as as function of T. **b** Plot of the heat capacity as a function of temperature.

(I leave it for you to find an analytic expression for the heat capacity.)

Harmonic oscillators in diatomic gases. It is interesting to notice the behavior of the heat capacity for a single harmonic oscillator. This would be the contribution to the heat capacity for the vibration of a diatomic molecule. Notice that when $T \ll T_E$ there is no contribution from the oscillations, whereas when $T \gg T_E$ (or about at T_E), there is a contibution from the oscillations. This allows us to sketch the expected effects of oscillations on the behavior of a diatomic ideal gas — where the behavior goes from $3/2NkT$ to $5/2NkT$ at around T_E for the corresponding oscillation. A nice prediction.

7.2.6 Example: Single spin system

We can now apply these methods to address the behavior of a paramagnetic system. For a single dipole in a magnetic field, there are two possible states, one state with spin up and energy $-\mu B$ and one state with spin down and energy μB.

Partition function for a single spin. The partition function for a single spin is therefore:

$$Z = \sum_i e^{-\beta\,\epsilon_i} = e^{-\beta\,\mu B} + e^{\beta\,\mu B} = 2\cosh(\beta\,\mu B)\ . \tag{7.49}$$

That was easy!

Energy for a single spin. We can find the average energy, $\bar{E}$, from:

$$\bar{E} = -\frac{\partial}{\partial\beta}\ln Z = -\frac{1}{\cosh(\beta\,\mu B)}(\sinh(\beta\,\mu B)\,\mu B) = -\mu B\,\tanh(\beta\,\mu B)\ , \tag{7.50}$$

and, as you will see further down, to get the result for N independent spins we simply multiply by N.

This was very simple compared with the long calculation you had to do in the microcanonical system!

7.2.7 Example: N-spin system

How can we find the behavior of a system of N non-interacting spins? If there is no interaction between the spins, then the energy of spin j depends only on the direction of spin j relative to the external magnetic field. If we write the direction of spin j as $S_j = \pm 1$, then the energy associated with spin j is $\epsilon(S_j) = -S_j\,\mu B$. The total energy of the system is therefore

$$\epsilon(S_1, S_2, \ldots, S_N) = -S_1\,\mu B - S_2\,\mu B - \ldots - S_N\,\mu B = \sum_j -S_j\,\mu B\ . \tag{7.51}$$

Partition function. The possible states of the system are then possible values for $(S_1, S_2, \ldots, S_N)$, to find the partition function we must therefore sum over all possible values for S_1, all possible values for S_2, and so on:

$$Z_N = \sum_{(S_1,S_2,\ldots,S_N)} e^{-\epsilon(S_1,S_2,\ldots,S_N)/kT} \tag{7.52}$$

$$= \sum_{S_1=\pm1} \sum_{S_2=\pm1} \cdots \sum_{S_N=\pm1} e^{-\epsilon(S_1)/kT-\epsilon(S_2)/kT-\ldots-\epsilon(S_N)/kT} \tag{7.53}$$

$$= \sum_{S_1=\pm1} \sum_{S_2=\pm1} \cdots \sum_{S_N=\pm1} e^{-\epsilon(S_1)/kT}\, e^{-\epsilon(S_2)/kT} \cdots e^{-\epsilon(S_N)/kT} \tag{7.54}$$

$$= \underbrace{\sum_{S_1=\pm1} e^{-\epsilon(S_1)/kT}}_{=Z_1} \underbrace{\sum_{S_2=\pm1} e^{-\epsilon(S_2)/kT}}_{=Z_1} \cdots \underbrace{\sum_{S_N=\pm1} e^{-\epsilon(S_N)/kT}}_{=Z_1} \tag{7.55}$$

$$= Z_1\, Z_1 \ldots Z_1 \tag{7.56}$$

$$= Z_1^N \,. \tag{7.57}$$

Where we have introduced Z_1 as the partition function for one spin, which we found in (7.49), and Z_N as the partition function for N non-interacting spins.

Energy of N-spin system. This means that the energy, E, which depends on $\ln Z$, simply will have a factor N in front:

$$\bar{E}_N = -\frac{\partial}{\partial\beta} \ln Z_N = -\frac{\partial}{\partial\beta} \ln Z_1^N == -N \frac{\partial}{\partial\beta} \ln Z_1 = N\bar{E}_1 \,. \tag{7.58}$$

However, the use of this argument depends critically on the assumption that the spins do not interact. What happens if they do? We will address this in the next example.

7.2.8 Example: Ising spin system

Ideal paramagnet. We have made many approximations in our model for a paramagnet. A paramagnet consists of many magnetic dipoles. These dipols may interact with an external magnetic field. However, each such dipole also sets up a magnetic field, which all the other dipoles may interact with. So far we have addressed the behavior of an *ideal paramagnet*, a paramagnet where there are no interactions among the spins. They only interact with the external field. But this is only an approximation to a real system. How can we make such a model more realistic?

Realistic paramagnet. To make a more realistic model, we need to include the interaction between spins. Every spin sets up a magnetic

field related to the spin direction, $B_i = cS_i$, where S_i is the direction of spin i. However, this field is not constant in space, but will decay with the distance from spin i. The interaction between spin i and spin j will therefore also depend on the distance r_{ij} between the spins.

As a first approximation, we will assume that the magnetic field set up by a spin decays rapidly in space, so that we only need to include the interactions between the nearest neighbors. If the spins are organized in a regular lattice, such as along a line in one dimension, in a square lattice in two dimensions or in a cubic lattice in three dimensions, then a spin S_i, will set up a magnetic field B_i, which will interact with the neighboring spins S_j, giving rise to energies $\epsilon_{ij} = -JS_iS_j$, where J is a coupling constant the includes the effect of the magnetic field B_i set up by spin i, the decay of the field with distance, and the coupling between the field B_i and spin S_j.

Energy of a state. The total energy of the system will therefore depend on the state $(S_1, S_2, \ldots, S_N)$ of the spin system, and the energy as a function of the state will also include terms of the form $S_i\, S_j$ for all the neighboring spins (i, j).

$$\epsilon(S_1, S_2, \ldots, S_N) = -J \sum_{(i,j)} S_i\, S_j + \sum_i (-\mu B)\, S_i \; . \tag{7.59}$$

where the sum (i, j) is over all i and for each i the sum is also over all the neighbors, j, of i. The neighbors of a spin i is illustrated in Fig. 7.8. This model is called the **Ising model** for a paramagnet. Notice that the energy now contains cross-terms $S_1\, S_2$, so we cannot use the exact same technique as before to find the partition function.

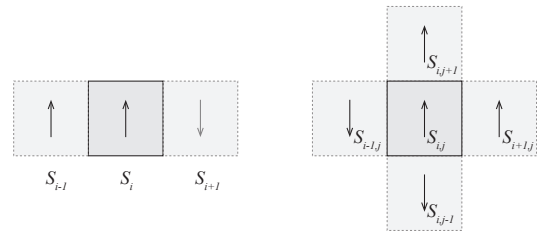


Fig. 7.8 Neighbors in the Ising model in one- and two-dimensions. **a** The neighbors for a spin S_i in one dimension are S_{i-1} and S_{i+1}. **b** The neighbors for a spin $S_{i,j}$ in two dimensions are $S_{i-1,j}$, $S_{i+1,j}$, $S_{i,j-1}$, and $S_{i,j+1}$ as illustrated.

Let us look at the energy of the Ising system. We see that the energy contribution from the spin-spin interaction is low $(-J)$ if the two spins are pointing in the same direction, and it is high $(+J)$ if the two spins are pointing in opposite directions. This spin interaction will therefore lead

to a situation where all the spins point in the same direction — given by the extenal field — in the lowest energy configuration. However, as temperature increases, fluctuations will arise, and it is not so clear what happens.

How can we find the partition function for this system? It turns out it is possible to find the partition function for a one-dimensional system.

Solution in one dimension. For a one-dimensional Ising model with no external magnetic field, the energy of a state $(S_1, S_2, \ldots, S_N)$ is

$$\epsilon(S_1, S_2, \ldots, S_N) = -J\left(S_1S_2 + S_2S_3 + S_3S_4 + \ldots + S_{N-1}S_N\right) . \tag{7.60}$$

The partition function can therefore be written as

$$Z = \sum_{S_1=\pm1} \sum_{S_2=\pm1} \cdots \sum_{S_N=\pm1} e^{J\beta S_1S_2} e^{J\beta S_2S_3} \cdots e^{J\beta S_{N-1}S_N} . \tag{7.61}$$

There is an old trick to calculate this particular sum. We notice that the sum over S_N is the same if $S_{N-1} = 1$ or is $S_{N-1} = -1$:

$$\sum_{S_N=\pm1} e^{J\beta S_{N-1}S_N} = e^{J\beta} + e^{-J\beta} = 2\cosh J\beta , \tag{7.62}$$

Now, when this sum is calculated, we can repeat this argument with the sum over S_{N-1}, and so on, until we reach S_2. The final sum over S_1 will then only yield a factor 2. The partition function is therefore

$$Z = 2^N \left(\cosh J\beta\right)^{N-1} \simeq \left(2\cosh J\beta\right)^N , \tag{7.63}$$

where the approximation is good when N is large. This is exactly the same as for the system of N independent spins, but we have replaced μB with J. From the partition function we can find the average energy:

$$\bar{E} = -\frac{\partial}{\partial\beta} \ln Z = -NJ \tanh J\beta . \tag{7.64}$$

We see that when $T \to 0$, $E \to -NJ$, that is, all the spins are parallel. When $T \to \infty$, $E \to 0$, which means that the spins become random, so that on average the sum of the interactions between the spins are zero.

ams 6: Comment on transition to ordered state? Not a phase transition in this case.

Approximate solutions in higher dimensions. It is also possible to find the partition function for a two-dimensional system. Onsager found this solution already in the 1940ies, and his solution was later confirmed by

Yang [2]. However, the mathematical methods needed are non-trivial and will not be introduced here. Instead, we will approximate the partition function and the equilibrium behavior of the system using approximate solutions. We will address methods to find approximate, statistical solutions in the canonical system in Sect. 7.10, and we will also address the Ising model in detail there.

7.2.9 Degenerate states

A common mistake is to sum over the energies in a system and not over the states when you calculate the partition function. What is the difference? Several states may have the same energy!

We can group microstates with the same energies together. Each state with an energy ϵ_s, will contribute a term $\exp(-\beta\epsilon_s)$ to the sum. Instead of summing over all the states i, we can sum over all the energies, but then we need to include how many states there are for a particular energy ϵ_s, the degeneracy $g(\epsilon_s)$:

$$Z = \sum_{\text{states } i} e^{-\beta\epsilon_i} = \sum_{\text{energy levels } s} g(\epsilon_s) e^{-\beta\epsilon_s} \ . \tag{7.65}$$

We will use both methods, choosing the method that is most adapted to the problem.

7.2.10 Example: System with degenerate states

For example, we could have studied a system where the number of states with the same energy increases with the energy. Let us introduce a system with the following energy states:

$$\epsilon_i = \epsilon(1, 2, 2, 3, 3, 3, 4, 4, 4, 4) \tag{7.66}$$

In this case the multiplicity of energy ϵ is 1, of 2ϵ is 2, of 3ϵ is 3 and so on.

The partition function is:

$$Z = \sum_i e^{-\beta\epsilon_i} = \sum_s g(\epsilon_s) e^{-\beta\epsilon_s} \ , \tag{7.67}$$

where $g(1) = 1$, $g(2) = 2$, $g(3) = 3$ and $g(4) = 4$. The sum is therefore:

$$Z = 1\,e^{-\beta\epsilon} + 2\,e^{-2\beta\epsilon} + 3\,e^{-3\beta\epsilon} + 4\,e^{-4\beta\epsilon}\ . \tag{7.68}$$

We find the average energy in the same way:

$$\bar{E} = \sum_i \epsilon_i e^{-\beta\epsilon_i} = \sum_s g(s)\,\epsilon_s\,e^{-\beta\epsilon_s}\ , \tag{7.69}$$

where the sum over i is over all the states, and the sum over s is over all the energies.

The average energy is therefore:

$$\bar{E} = \epsilon\frac{1\,e^{-\beta\epsilon} + 2\times 2e^{-2\beta\epsilon} + 3\times 3e^{-3\beta\epsilon} + 4\times 4e^{-4\beta\epsilon}}{Z}\ . \tag{7.70}$$

7.3 Thermodynamics in the canonical ensemble

In the microcanonocal system, all microstates have the same probability, $P(i) = 1/\Omega$, whereas in the canonical system, the probability of a microstate is weighted with its energy, ϵ_i, according to the Boltzmann factor $P(i) = (1/Z)\,e^{-\epsilon_i/kT}$. In the microcanonical system, the coupling between the microscopic models and the macroscopic thermodynamics, occurs through $S(E,V,N) = k\ln\Omega(E,V,N)$. Can we find a similar simple relation between the microscopic and the macroscopic for the canonical system where (T,V,N) are constants? In addition, we have discovered the second law of thermodynamics, which is valid for an isolated system. Can we find a similar law for a canonical system? Here, we will address these questions, starting from the thermodynamic identity and its extension to the canonical system.

We have already found a connection between the micro-scopic and the macro-scopic representation of a canonical system: We found that we can calculate the (average) energy $\bar{E}$ in the system directly from the partition function:

$$\bar{E} = \sum_i \epsilon_i P(i) = -\frac{\partial}{\partial\beta}\ln Z\ , \tag{7.71}$$

where $\beta = 1/(kT)$. We will now use this relation as a basis for the coupling between the microscopic and the macroscopic.

7.3.1 Helmholtz free energy

For the microcanonical ensemble, entropy is a function of E, V, N, $S = S(E, V, N)$, and the thermodynamic identity is

$$TdS = dE + pdV - \mu dN \ , \tag{7.72}$$

Thermodynamic identity for a (S, V, N) system. We can change from $S = S(E, V, N)$ to the energy as a function of S,V,and N:

$$E = E(S, V, N) \ , \tag{7.73}$$

and we can find the corresponding differential relations by rearranging the thermodynamic identity:

$$dE = TdS - pdV + \mu dN \ . \tag{7.74}$$

Indeed, from this rearrangement we see that when E is a function of S, V, N, we can find the T, p and μ from various derivatives of $E = E(S, V, N)$.

Thermodynamic identity for a (T, V, N) system. Now, to describe the canonical system, we would like to have a similar differential to dE, but with the differentials dT, dV, and dN. We would like to have a relation like:

$$dA = ()dT + ()dV + ()dN \ , \tag{7.75}$$

where we need to find the unknown functions/variables shown as () in the differential. Also, we need to figure out how we can construct A to have these properties. There is a formal way to do this, using Legendre transforms, which is robust (meaning that it always works) and we will introduce it in Sect. 7.6. For now, we will motivate the Legendre transform by observing that the differential for E almost got us there, we only need to replace the dS term with some dT term. We need to subtract the TdS term and then get a new XdT term with the T dependence. We see that if we subtract $d(TS)$ on both sides of the thermodynamic identity, we get

$$\begin{aligned} dE - d(TS) &= TdS - d(TS) - pdV + \mu dN \\ &= TdS - TdS - SdT - pdV + \mu dN \\ &= -SdT - pdV + \mu dN \ , \end{aligned} \tag{7.76}$$

which has exactly the right form! This is the quantity we were looking for. It is an energy, since it has units energy, and it is a function of T, V, and N. We call this quantity F, Helmholtz free energy:

Helmholtz free energy: $F(T,V,N) = E - TS$

The differential for F, Helmholtz free energy, is:

$$dF = -SdT - pdV + \mu dN \ . \tag{7.77}$$

Application of the thermodynamic identity for the canonical system. This identity allows us to find the entropy, S, the pressure p, and the chemical potential μ if we have an expression for Helmholtz free energy:

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \ , \ p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \ , \ \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \ . \tag{7.78}$$

This means that if we know Helmholtz free energy, we can calculate the remaining thermodynamical properties of the system. However, to do this we need an expression for $F(V,T,N)$ either from experiments, simulations, or theory. We have already found how we can connect the energy E to the microscopic theory through the use of the partition function $Z(T,V,N)$. How can we relate F to the partition function and thus complete the coupling between the microscopic and the macroscopic also for the canonical system?

7.3.2 A microscopic expression for Helmholtz free energy

We have already found that the energy E can be related to the microscopic description of the system through the partition function:

$$\bar{E} = -\frac{\partial \ln Z(T,V,N)}{\partial \beta} \ , \tag{7.79}$$

We can therefore find the energy if we know the partition function. However, for the canonical system, we would also like to be able to calculate quantities such as the entropy, the pressure or the chemical potential. We would therefore like to have an expression for Helmholtz

free energy, because, as we saw above, we can then find all quantities of interest from partial derivatives of F. We know that Helmholtz free energy, F, is

$$F = E - TS \,, \tag{7.80}$$

We already know E, but we do not know S — yet. However, we know that $dF = -SdT - pdV + \mu dN$ and therefore we know that

$$S = \left(\frac{\partial F}{\partial T}\right)_{V,N} \,, \tag{7.81}$$

which we can insert back into (7.80), giving

$$F = E + T\left(\frac{\partial F}{\partial T}\right)_{V,N} \,. \tag{7.82}$$

We recognize this as a differential equation for F.

$$F = -\frac{\partial \ln Z(V,T,N)}{\partial \beta} + T\left(\frac{\partial F}{\partial T}\right)_{V,N} \,. \tag{7.83}$$

We can find the solutions to this differential equation, by starting from a good guess — which of course is correct. A solution to this differential equation is

$$F = -kT \ln Z(V,T,N) \,. \tag{7.84}$$

We can demonstate that this is a solution by substitution. First, we find the derivative of F with respect to T:

$$\begin{aligned}\left(\frac{\partial F}{\partial T}\right)_V &= \frac{\partial}{\partial T} - kT \ln Z \\ &= -k \ln Z - kT \frac{\partial \ln Z}{\partial T} \\ &= -k \ln Z - kT \frac{E}{kT^2} \,.\end{aligned} \tag{7.85}$$

Then we insert this into the differential equation for F in (7.83) and find that

$$E + T\left(-k \ln Z - (E/T)\right) = -k \ln Z \,, \tag{7.86}$$

which proves that $F = -kT \ln Z$ is one possible solution to the differential equation in (7.83). However, there may in addition be a constant term $F = -kT \ln Z + aT$. To address the aT term, we need to invoke the *third law of thermodynamics*, which states that the entropy of a system

approaches a small number or zero as the temperature approaches zero. This is because as the temperature approaches zero, the system will be frozen into the lowest possible energy configurations, and the entropy will only be the possible degeration of the lowest energy, which will typically be a small number. Since the entropy is the derivative of F with respect to T, we see that the entropy only will approach zero or a small number if a is zero or a small number, that is, is $a \simeq 0$.

We have therefore proven the fundamental relationship between thermodynamics and the microscopic models for the canonical system:

Relation between Helmholtz free energy and the microscopic description of a canonical system:

$$F(T,V,N) = -kT \ln Z(T,V,N) \; , \tag{7.87}$$

where $Z(T,V,N) = \sum_i e^{-\epsilon_i/kT}$ is the partition function for the system.

I find this equation beautiful. It is simple. It connects the microscopic and the macroscopic worlds. It allows all types of calculations of S, p and μ from the partition function. We can now apply this to find all thermodynamic properties of systems described in the canonical ensemble.

7.3.3 Helmholz free energy decreases in approach to equilbrium

For the micro-canonical ensemble we have seen how the system develops towards equilibrium and what characterizes the equilibrium state: The system evolves towards a the most likely macrostate – that is, towards the macrostate with most microstates – equivalently the system evolves towards maximum entropy:

$$\Delta S \geq 0 \; . \tag{7.88}$$

This condition also characterizes the equilibrium state: The entropy is a maximum for the equilbrium state.

For example, for the Einstein crystal, divided the system into two parts, A and B. We saw that the total entropy of the system $S = S_A + S_B$

increases with time, and that the total entropy is maximal at equilibrium. This is indeed what characterizes the equilibrium state in the system.

Do we have similar principles for the canonical system? Does the entropy increase for a canonical system? Does the energy increase? Or is there some other principle that we do not yet know? We can study this in the Einstein crystal. We make a system consisting of two parts, a small part A, which is the system, and a large part B, which is the reservoir. We can make a system with $N_A = 10$ and $N_B = 990$. This should ensure that system B is so large that the temperature in this system does not change – it is constant and equal to the reservoir temperature $T_B = T_R$. The total energy is $q = q_A + q_B$. We can simulate the dynamics of this system using the monte-carlo method we developed previously. This would allow us to simulate the time dynamics of the system as it develops through a sequence of microstates. This is done by the following program. (Notice that if you use this program in a Jupyter notebook, you may need to split the code into two cells, so that the `for`-loop is in the second cell. In this way, the figure will be created with correct size in the first cell before it is updated in the second cell).

```
# MC for a two part Einstein crystal
import numpy as np
import matplotlib.pyplot as plt
NA = 10
NB = 990
qA = 300
qB = 9700
q = qA + qB # Total energy
N = NA + NB
nstep = 1000000
nbetween = 1000
state = np.zeros(N,float)
# Generate initial, random state
placeA = np.random.randint(0,NA,qA)
for ip in range(len(placeA)):
    i = placeA[ip]
    state[i] = state[i] + 1
placeB = np.random.randint(0,NB,qB)+NA
for ip in range(len(placeB)):
    i = placeB[ip]
    state[i] = state[i] + 1
# Simulate state development
EA = np.zeros(nstep,float)
EB = np.zeros(nstep,float)
TBSA = np.zeros(nstep,float)
TB = np.zeros(nstep,float)
# Setup dynamic graphics
%matplotlib notebook
```

```
fig = plt.figure()
ax1 = fig.add_subplot(411)
ax2 = fig.add_subplot(412)
ax3 = fig.add_subplot(413)
ax4 = fig.add_subplot(414)
for istep in range(nstep):
    i1 = np.random.randint(0,N) # Select oscillator at random
    if (state[i1]>0):  # Check if it has energy
        i2 = np.random.randint(0,N) # Then find other oscillator
        state[i2] = state[i2] + 1
        state[i1] = state[i1] - 1
    # Calculate T_B S_A
    EA[istep] = np.sum(state[:NA-1])
    EB[istep] = np.sum(state[NA:])
    qA = EA[istep]
    qB = EB[istep]
    omegaA = comb(NA+qA-1,qA)
    TB[istep] = qB/NB
    TBSA[istep] = TB[istep]*np.log(omegaA)
    if (np.mod(istep,nbetween)==0):
        # Plot state
        ax1.clear()
        ax1.bar(np.arange(NA),state[:NA],color='b')
        ax1.bar(np.arange(NA+1,N),state[NA+1:],color='r')
        #a = ax1.axis(); a(2) = N; ax1.axis(a);
        #ax1.xlabel('i'); ax1.ylabel('n_i');
        # Plot avg energy in each system
        ax2.clear()
        ax2.plot(EA[:istep]/NA,'-r',EB[:istep]/NB,'-b');
        #ax2.xlabel('t'); ax2.ylabel('q_A/N_A , q_B/N_B')
        # Plot T_B
        ax3.clear()
        ax3.plot(EB[:istep]/NB,'-b')
        #ax3.xlabel('t'); ax3.ylabel('T_B')
        # Plot
        ax4.clear()
        ax4.plot(EA[:istep],'-r',TBSA[:istep],'-b',EA[:istep]-TBSA[:istep],'-k')
        # Draw
        fig.canvas.draw()
```

We need to decide on the initial distribution of energy in the system. We try to alternatives: First, we try the case where system A has a large amount of the energy than an even distribution, we choose $q_A = 300$, and $q_B = 9700$, giving $q = 10000$. In this case, the average energy per oscillator in system A is 30 and the average energy per oscillator in system B is 9.8. We measure the entropy, $S_A(t)$ and energy $q_A(t)$ of system A, $S_A(t)$, as a function of time t. Fig. 7.9 shows the development of $E_A(t)$ and $E_B(t)$. We see that the energy in system A decreases with time and reaches an approximately constant level – corresponding to the equilibrium state of the system. We see that in the equilibrium state

– after a long time – the energy is evenly distributed throughout the system, and the average energy per oscillator in system A and system B, E_A/N_A and E_B/N_B become the same value, close to the average of the total energy per oscillator, q/N, in the system. This is not surprising. We expect the average energy per oscillator to be related to the temperature, and we know that in equilibrium the temperatures in the two systems will be the same. We also see that system B is indeed large. In the plot that shows both E_A/N_A and E_B/N_B we see that the change in E_B/N_B is very small compared with the change in system A. This is what we mean when we call a reservoir *large* – it is so large that the temperature does not change significantly. There are, however, changes in E_B/N_B as well, as evident from the plot of E_B/N_B alone, but these fluctuations are small.

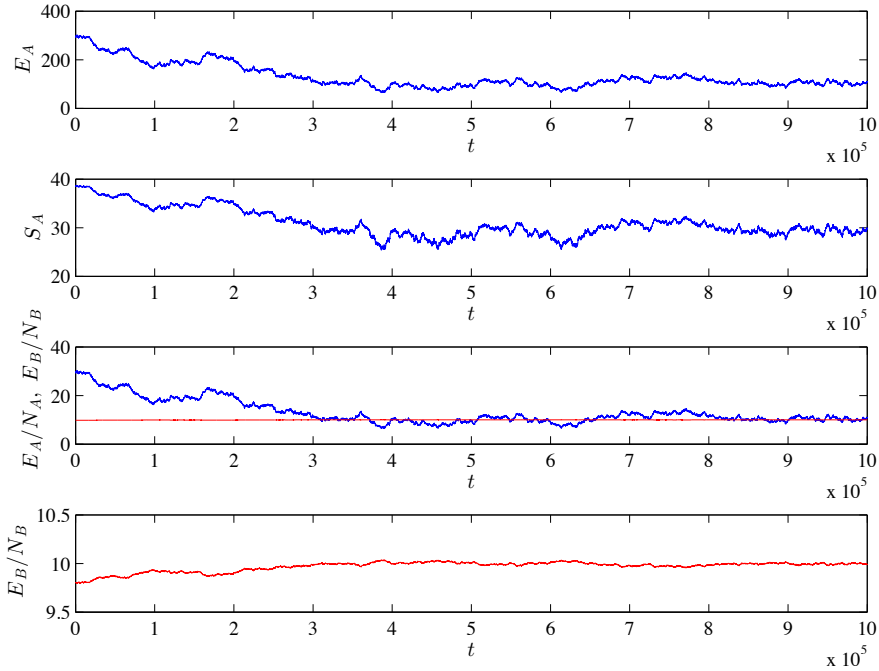


Fig. 7.9 Time development of the energies in two systems A with $N_A = 10$ and B with $N_B = 990$ in thermal contact.

The energy in system A clearly decreases with time in this system. What about the entropy? We can measure the entropy in system A through the formula $S_A = k \ln \Omega_A(q_A, N_A)$, which we have developed previously. Now, q_A will change with time, and therefore as will S_A.

This measurement was implemented in the program above. Fig. 7.10 shows both the energy and the entropy for system A. We see that the entropy in system A decreases in time an approaches an approximately stationary value. We expect the fluctuations to become smaller for larger systems and the entropy to become sharp. These observations suggest that in equilibrium – which is what we reach after a long time – the system has reached a minimum in entropy and a minimum in energy. Hmmm. That was surprising – the entropy is decreasing in this system. (Convince yourself that this is not unphysical and that the second law of thermodynamics still is valid).

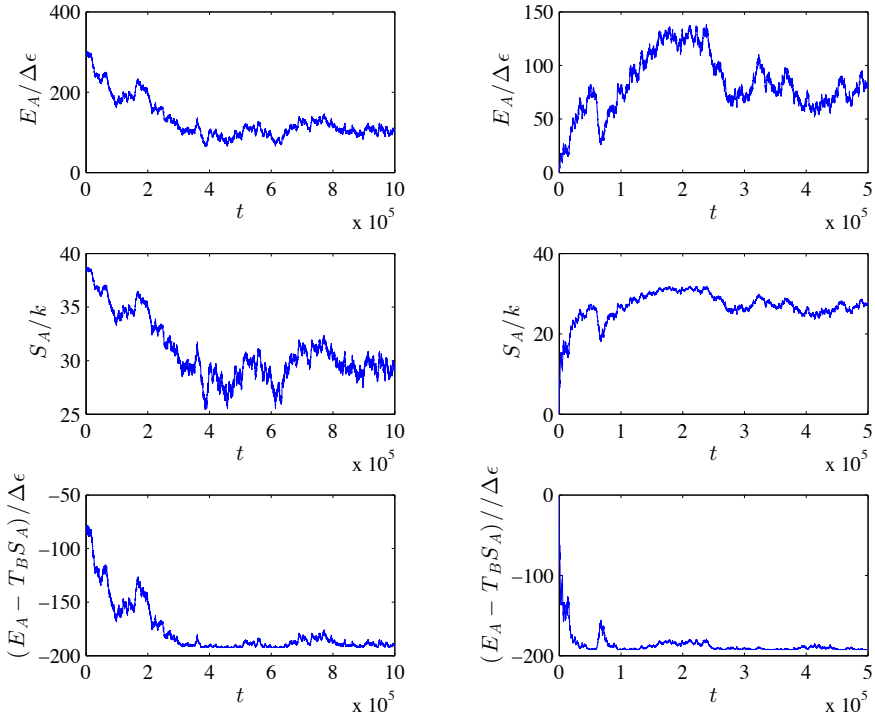


Fig. 7.10 Time development of the energies and entropies for an Einstein system with $N_A = 10$ and $N_B = 990$ for two different initial distributions of energy. In the left figures the system start with $q_A = 300$, $q_B = 9700$ and in the right figures the system start with $q_A = 0$ and $q_B = 10000$.

Let us try the same simulation, but from a different starting point – the case where system A initially has less energy than the average energy, $q_A = 0$ and $q_B = 10000$. We expect the same equilibrium state in this system as in the previous since the number of oscillators is the same and

the total energy in the system is the same. How does it develop in time? Fig. 7.10 shows that the entropy, $S_A(t)$, and energy, $q_A(t)$, of system A in this case both increase, and then reach a stationary value which is a maximum of both entropy and energy? Hmmm. Now it is not easy to see what kind of rule would determine the behavior of the system: Should the entropy be a maximum or a minimum? Should the entropy increase or decrease?

It is time to revisit the theory and develop a better theoretical understanding of this process. We know that the whole system, system A and system B, is thermally isolated and therefore described as a microcanonical system with E, V, and N constant. For this system, the energy is constant $E_A + E_B = E$ and the total entropy must be increasing:

$$\Delta S = \Delta S_A + \Delta S_B \geq 0 \, . \tag{7.89}$$

The change in entropy in system A is indeed what we observed in the simulations – it could be both positive or negative. This is what we want to develop a theory for. What about the development of system B? Since system B is very large, we will now assume that system B is very close to equilibrium – so close that we can assume that system B is in equilibrium. In this case, we know that we can describe the behavior of system B with the thermodynamic identity:

$$TdS = dE + pdV - \mu dN \, , \tag{7.90}$$

where dV and dN is zero for system B – there is no change in volume or the number of oscillators. We therefore get $T_B \Delta S_B = Q_B = \Delta E_B$, where the change in energy in system B comes from the thermal energy Q_B transferred from system A to system B, which again corresponds to $\Delta E_B = -\Delta E_A$, and therefore we have $\Delta S_B = -\Delta E_A / T_B$, which we insert back into (7.89), getting:

$$\Delta S = \Delta S_A - \frac{\Delta E_A}{T_B} \geq 0 \, . \tag{7.91}$$

We rearrange, getting

$$T_B \Delta S_A - \Delta E_A \geq 0 \, , \tag{7.92}$$

where we can now assume that T_B does not change – it is (approximately) constant during this process, and we can write this as the change in $(E_A - T_B S_A)$:

$$\Delta(T_B S_A - E_A) \geq 0 \,, \tag{7.93}$$

or

$$-\Delta(E_A - T_B S_A) \geq 0 \,, \tag{7.94}$$

giving, finally:

$$\Delta(E_A - T_B S_A) \leq 0 \,. \tag{7.95}$$

We have now found a criterion for how the system should develop in time, and a criterion for the equilibrium. In equilibrium, we expect $E_A - T_B S_A$ to be minimal and we expect $E_A - T_B S_A$ to decrease with time and approach a minimal value.

In equilibrium, $T_A = T_B$, and we therefore know that in equilibrium

$$\Delta(E_A - T_A S_A) \geq 0 \,, \tag{7.96}$$

where we recognize this as Helmholtz free energy, $F_A = E_A - T_A S_A$ for system A. We have therefore found a new principle, valid for a canonical system with T, V, N constant: For this system, Helmholtz free energy will always decrease $\Delta F \leq 0$, and the equilibrium state is characterize as a minimum for Helmholtz free energy, F.

First, let us check if this indeed is the case in the simulations we performed above. We have already found E_A and S_A as part of the simulations. How can we find T_B from the simulations? The simplest approach would be to notice that system B is large, so that the high temperature and large system size approximation for the energy of an Einstein crytal can be used, $E_N \simeq N_B k T_B$, which we can use to find the temperature through $T_N = E_N/(N_B k)$. We are now able to calculate $E_A - T_B S_A$, but we should now check that the units we use are ok. In the program, we calculate q_A, q_B and $\ln \Omega$ for S_A. We know that $E_A = q_A \epsilon$ and $E_B = q_B \epsilon$, where ϵ is the characteristic energy (the energy level spacing) for the harmonic oscillator in the Einstein crystal. We therefore get:

$$\begin{aligned} E_A - T_B S_A &= \epsilon q_A - (\epsilon q_B)/(N_B k) k \ln \Omega(q_A, N_A) \\ &= \epsilon \left(q_A - (q_B/N_B) \ln \Omega(q_A, N_A)\right) \,, \end{aligned} \tag{7.97}$$

which shows that we can use $q_A(q_B/N_B) \ln \Omega$ to characterize F_A in the simulations. Fig. 7.10 shows $F_A(t)$ for the two simulated models. We see that even if S_A and E_A are increasing for one system and decreasing in the other system, Helmholtz free energy F_A is decreasing in both cases, and approaches a stationary value as time increases, corresponding to the

equilibrium value of F_A. Indeed, the theory we have developed describes the system we have studied, and provides us with additional insight: Helmholtz free energy is minimal in equilibrium.

This means that Helmholtz free energy plays a similar role in canonocal systems (T, V, N constant) as the entropy did for the microcanonical system (E, V, N constant). We call the principle that the free energy is minimal in equilibrium a minimum free energy principle, and we will see later that such principles are very useful and will allow us to calculate equilibrium states in for example fluid or solid mixtures, and for both biological, geological and meteorological systems of interest.

Where does the word **free** energy come from? We see that if T and N is constant, then the work that is performed by a canonical system is

$$(\Delta F)_T = -p\Delta V = -\Delta W \ , \tag{7.98}$$

so that the work ΔW done on the environment by the system at constant temperature and N is equal to the loss in free energy for the system: Hence the name *free energy*. (This is only true for reversible processes).

7.3.4 Helmholtz free energy is a minimum in equilibrium

We can also show that Helmholtz free energy is not only an extremum value, but it is actually a minimum:

The total energy is $E = E_S + E_R$. The total entropy is

$$\begin{aligned} S &= S_R + S_S \\ &= S_R\left(E - E_S\right) + S_S(E_S) \\ &\simeq S_R(E) - E_S \left(\frac{\partial S_R}{\partial U_R}\right)_{V,N} + S_S(E_S) \ . \end{aligned} \tag{7.99}$$

Where we also know that

$$\left(\frac{\partial S_R}{\partial E_R}\right)_{V,N} = \frac{1}{T} \ , \tag{7.100}$$

and therefore we get

$$S = S_R(E) - \frac{1}{T} F_S \ , \tag{7.101}$$

where $F_S = E_S - TS_S$ is the free energy of the system. Now $S_R(E)$ is a constant. We also know that $S = S_R + S_S$ is maximum with respect to E_S in equilibrium. There F_S must be a minimum.

7.3.5 Maxwell relations

We can now find a new group of relationsm, thermodynamic relations, based on the second derivative of Helmholtz free energy: If the energy is continuous in its variables, we know that the order of derivatives for the second derivative is not important, therefore:

$$\left(\frac{\partial}{\partial T}\left(\frac{\partial F}{\partial V}\right)_{T,N}\right)_{V,N} = \frac{\partial^2 F}{\partial V \partial T} = \frac{\partial^2 F}{\partial T \partial V} = \left(\frac{\partial}{\partial V}\left(\frac{\partial F}{\partial T}\right)_{V,N}\right)_{T,N} . \tag{7.102}$$

We call this type of relation a **Maxwell relation**. We can use Maxwell relations to find non-trivial relations between thermodynamic variables. For example, we can now show that

$$\left(\frac{\partial}{\partial T}\left(\frac{\partial F}{\partial V}\right)_{T,N}\right)_{V,N} = \left(\frac{\partial p}{\partial T}\right)_{V} = \left(\frac{\partial}{\partial V}\left(\frac{\partial F}{\partial T}\right)_{V,N}\right)_{T,N} = \left(\frac{\partial S}{\partial V}\right)_{T} . \tag{7.103}$$

We will find more such relations later as we introduce more variables. (There are three possible permutations in this case, but the other two involve the chemical potential, which we are not familiar with yet).

7.4 Composite systems

7.4.1 Aspects of motion

The energy of a realistic system has many contributions. For example, for a diatomic gas, the energy of a single molecule consists of terms relating to the translational motion of the molecule, the rotational motion of the molecule, and the vibration of the molecule. To specify the state of the molecule we therefore need to describe the state of each of these *aspects* of the molecule, for example by listing the states i for translation, the states j for rotation, and the states n for vibration: (i, j, n). Typically, the contribution to the total energy from each of these aspects are independent: The total energy is the sum of the energies from each of the aspects of motion (or from each of the degrees of freedom, if you prefer that term). The total energy is therefore

$$\epsilon(i, j, n) = \epsilon_i + \epsilon_j + \epsilon_n , \tag{7.104}$$

In order to find the partition function for the diatomic molecule, we therefore need to sum over all the states of translation, rotation, and vibration. If these contributions are independent, we need to sum over all values of i, and for each value of i, we need to sum over all values of j, and for each value of j we must sum over all values of n. We write this as

$$Z = \sum_i \sum_j \sum_n e^{-(\epsilon_i+\epsilon_j+\epsilon_n)/kT} = \sum_i \sum_j \sum_n e^{-\epsilon_i/kT}\, e^{-\epsilon_j/kT}\, e^{-\epsilon_n/kT} \,. \tag{7.105}$$

This can be further simplified by realizing that the i-terms do not vary when we sum over j and n, and they can therefore be placed outside those sums. Similarly, the j-terms do not vary when we sum over n, and they can therefore be placed outside the sum over n, giving:

$$Z = \underbrace{\sum_i e^{-\epsilon_i/kT}}_{Z_{\text{trans}}} \underbrace{\sum_j e^{-\epsilon_j/kT}}_{Z_{\text{rot}}} \underbrace{\sum_n e^{-\epsilon_n/kT}}_{Z_{\text{vib}}} = Z_{\text{trans}}\, Z_{\text{rot}}\, Z_{\text{vib}} \,. \tag{7.106}$$

The total partition function is therefore the product of the partition functions for each of the aspects of motion — as long as the states of the various aspects are independent.

Partition function for independent aspects of a single particle:

$$Z = Z_{\text{trans}}\, Z_{\text{rot}}\, Z_{\text{vib}} \,. \tag{7.107}$$

This argument can be extended to any new degrees of motion we add to such a system, for example, the possibility for the atom or molecule of the gas to be in a ground state or an excited state. However, the same technique can also be used to find the partition function for a system of several particles, if the states of each of the particles are independent. Let us see how.

7.4.2 System of two particles

Let us start by a system consisting of two different particles, each described independently by the quantum states $\epsilon_{1,i}$ and $\epsilon_{2,j}$. The partition function is

$$\begin{aligned} Z &= \sum_i \sum_j e^{-\beta\epsilon_{1,i}}\, e^{-\beta\epsilon_{2,j}} \\ &= \underbrace{\sum_i e^{-\beta\epsilon_{1,i}}}_{Z_1}\, \underbrace{\sum_j e^{-\beta\epsilon_{2,j}}}_{Z_2} \\ &= Z_1\, Z_2 \ . \end{aligned} \tag{7.108}$$

For two non-interacting systems, with *distinguishable particles*, the partition function is the product of the partition functions!

Now, what if the two particles are not distinguishable? In that case, we have included too many elements in the partition function sum, because we have included both the term for the state

$$e^{-\beta(\epsilon_{2,j}+\epsilon_{1,i})} \tag{7.109}$$

and the term for the state

$$e^{-\beta(\epsilon_{2,i}+\epsilon_{1,j})} \ , \tag{7.110}$$

but if the particles are indistinguishable, these two states are the same, and should only be included once in the sum. We should therefore divide the sum by 2:

$$Z = \frac{1}{2} Z_1\, Z_2 \ , \tag{7.111}$$

This is almost correct. But the sum also includes terms where the state of both particles are the same, but only one such term is included in the sum:

$$e^{-\beta(\epsilon_{2,i}+\epsilon_{1,i})} \ , \tag{7.112}$$

and I have also divided this part of the sum by two. Since these terms not have been double counted, we should not divide these parts of the sum by two. However, the error we make by dividing the whole sum by two will be small as long as there are few of such states — and there are very few of these states compared with the other states. We therefore use the approximation:

$$Z \simeq \frac{1}{2} Z_1\, Z_2 \text{ (indistinguishable/identical)} \tag{7.113}$$

when the two particles are indistinguishable, and the exact result

$$Z = Z_1\, Z_2 \text{ (distinguishable/different)} \tag{7.114}$$

when the particles are distinguishable.

7.4.3 System of many particles

This result can be generalized to N particles. If the particles are distinguishable, we find the partition function as the product of the partition functions for each of the particles:

$$Z = Z_1 Z_2 \dots Z_N \ . \tag{7.115}$$

However, for indistinguishable particles, we have included the terms where all the states of all the particles are different too many times. How many times have we included such terms? The number of ways that we can rearrange the N different states among the N indistinguishable particles, which is $N!$ ways. We must therefore divide by $N!$ in order to get the correct result for these terms. In addition, there will be terms where some of the states are the same, and we will make a small error by also dividing these terms by $N!$, but this error is small as long as there are few of cases where the states are the same compared to the cases where all the states are different. We therefore have the following approximate result for indistinguishable particle and exact result for distinguishable particles:

The partition function for a system of N particles with partition functions, Z_1, for each particle is:

$$Z_N = Z_1^N \text{ (distinguishable particles)} \tag{7.116}$$

$$Z_N = \frac{Z_1^N}{N!} \text{ (indistinguishable particles)} \tag{7.117}$$

This result now makes it easy to find

- the partition function for a system of N identical harmonic oscillators: The oscillators are distinguishable (because they are localized in various positions in space), and hence the partition function of N oscillators is $Z_N = Z_1^N$, where Z_1 is the partition function for one oscillator
- the partition function for a system of N identical magnetic spins in an external magnetic field: The spins are distinguishable (because they

are localized in various positions in space), and hence the partition function of N spins is $Z_N = Z_1^N$, where Z_1 is the partition function for one spin
- the partition function for a system of N identical particles in an ideal gas: The particles are indistinguishable (since they are identical and only limited to being inside a box), and hence the partition function of N particles s $Z_N = Z_1^N/N!$, where Z_1 is the partition function for one particle

We notice that since Helmholtz free energy depends on the logarithm of the partition function, we see that the Helmholtz free energy for N *distinguishable* particles is

$$F_N = -kT \ln Z_1^N = -NkT \ln Z_1 = N F_1 \ . \tag{7.118}$$

This corresponds to an addition principle for the system: Helmholtz free energy is the sum of the free energies for each of the particles, as long as the particles are distinguishable.

However, for N *indistinguishable* particles, Helmholtz free energy is

$$F_N \simeq -kT \ln Z_1^N/N! \simeq -NkT \ln Z_1 + NkT(\ln N + 1) = NF_1 + NkT(\ln N + 1) \ . \tag{7.119}$$

7.4.4 Example: Ideal gas in the canonical ensemble

We can use these new results to find the partition function for an ideal gas. First we find Z_1 for one particle in a box, and then we find the partition function, Z_N for an ideal gas of N particles in a box.

Single particle in a box. For a single particle, we may also have to include various aspects of motion, such as the rotational and vibrational degrees of freedom for a diatom. However, let us postpone these differences, and start by finding the partition function for translational motion of a single particle in a box. The box has sides L and volume V. From quantum mechanics we know that the possible states of a particle in a box are described by the quantum numbers n_x, n_y, n_z, representing the states of motion in the x-, y-, and z-directions. The energy for the state (n_x, n_y, n_z) is

$$\epsilon_n = \frac{h^2}{8mL^2}\left(n_x^2 + n_y^2 + n_z^2\right) = a\left(n_x^2 + n_y^2 + n_z^2\right) \ , \tag{7.120}$$

where $a = h^2/(8mL^2)$. The partition function is given by the sum over all the states n_x, n_y, n_z:

$$Z_{1,\text{trans}} = \sum_{n_x}\sum_{n_y}\sum_{n_z} e^{-a\left(n_x^2+n_y^2+n_z^2\right)} = \sum_{n_x}\sum_{n_y}\sum_{n_z} e^{-an_x^2}\, e^{-an_y^2}\, e^{-an_z^2} \;, \tag{7.121}$$

We recognize this as the sum over three identical degrees of freedom for motion along the x-, y-, and z-direction, and we can separate these sums just as we did for composite systems:

$$Z_{1,\text{trans}} = \sum_{n_x}\sum_{n_y}\sum_{n_z} e^{-an_x^2}\, e^{-an_y^2}\, e^{-an_z^2} = \underbrace{\sum_{n_x} e^{-an_x^2}}_{Z_{1,x}} \underbrace{\sum_{n_y} e^{-an_y^2}}_{Z_{1,y}} \underbrace{\sum_{n_z} e^{-an_z^2}}_{Z_{1,z}} \;. \tag{7.122}$$

where each of the sums are identical and equal to ξ

$$Z_{1,\text{trans}} = \xi^3 \;,\; \xi = \sum_{n_x=0}^{\infty} e^{-an_x^2} \;. \tag{7.123}$$

For small values of a we can approximate the sum by the integral:

$$\xi \simeq \int_0^\infty \exp(-an^2)\,, dn = \frac{\sqrt{\pi}}{2\sqrt{a}} \;. \tag{7.124}$$

The partition function is therefore given by:

$$Z_{1,\text{trans}} \simeq \frac{\pi^{3/2}}{8}\left(\frac{2mL^2}{\beta\hbar^2\pi^2}\right)^{3/2} = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V \;. \tag{7.125}$$

To simplify this expression furthre, we introduce the *quantum density* $n_Q(T)$ defined as:

$$n_Q = \left(\frac{m}{2\pi\hbar^2\beta}\right)^{3/2} = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} \;. \tag{7.126}$$

The partition function for the translational motion of a single particle in a three-dimensional box of size $V = L^3$ is then:

$$Z_{1,\text{trans}} = V\, n_Q = \frac{n_Q}{n} \;, \tag{7.127}$$

where $n = 1/V$ is the density of particles.

The average particle energy is given by:

$$\bar{E}_{1,\text{trans}} = -\frac{\partial \ln Z_{1,\text{trans}}}{\partial \beta} = -\frac{\partial}{\partial \beta} \ln V \left(\frac{m}{2\pi\hbar^2 \beta} \right)^{3/2} = \frac{3}{2} kT \ . \quad (7.128)$$

This is in correspondence with the equipartition principle, because the particle has only one degree of freedom for each direction of motion, hence three degrees of freedom in total. The energy per particle due to the translational motion is therefore $(3/2)\, kT$ according to the equipartition principle.

If we want to include more aspects of the gas, this would appear as additional terms in the partition function for the individual particle. For example, for a diatomic gas we would then also include partition function for rotation and vibration:

$$Z_1 = Z_{1,\text{trans}}\, Z_{1,\text{rot}}\, Z_{1,\text{vib}} = \xi^3\, Z_{1,\text{rot}}\, Z_{1,\text{vib}} = \ . \quad (7.129)$$

These terms will also be included in the energy of a single particle and therefore also in the heat capacity of a single particle in the gas. Let us now use this result to find the partition function for a gas of N particles.

Ideal gas of N particles in a box. The particles in an ideal gas are indistinguishable, hence the partition function for the N-particle gas is

$$Z_N = \frac{Z_1^N}{N!} = \frac{1}{N!} \left(Z_{1,\text{trans}}\, Z_{1,\text{rot}}\, Z_{1,\text{vib}} \right) \ . \quad (7.130)$$

From this we find either the energy directly, or Helmholtz free energi, which we then use to find the entropy and the pressure of the gas. We start with the energy:

$$\bar{E}_N = \frac{\partial \ln Z_N}{\partial \beta} = N \frac{\partial \ln Z_1}{\partial \beta} \quad (7.131)$$

$$= N \frac{\partial}{\partial \beta} \left(\ln Z_{1,\text{trans}} + \ln Z_{1,\text{vib}} + \ln Z_{1,\text{rot}} \right) \quad (7.132)$$

$$= N\, E_{1,\text{trans}} + N\, E_{1,\text{vib}} + N\, E_{1,\text{rot}} \ . \quad (7.133)$$

Where we have already in (7.128) found that $E_{1,\text{trans}} = (3/2)kT$. The energy of an ideal monoatomic gas (with only translational degrees of freedom) is therefore

$$E_N = \frac{3}{2} NkT \ . \quad (7.134)$$

Notice that we simply sum the various energy terms to find the total energy. Similarly, we find Helmholtz free energy from the logarithm of the partition function

$$F_N = -kT \ln Z_N = -kT \left(N \ln Z_1 - N \ln N - N \right) \tag{7.135}$$

$$= -NkT \left(\ln Z_{1,\text{trans}} + \ln Z_{1,\text{vib}} + \ln Z_{1,\text{rot}} - \ln N + 1 \right) . \tag{7.136}$$

Let us find the behavior for a monatomic gas with only translational degrees of freedom. We insert $Z_{1,\text{trans}} = n_Q/n$, getting

$$F_N = -NkT \left(\ln(V\, n_Q) - \ln N + 1 \right) = NkT \left(\ln((V/N) n_Q + 1 \right) . \tag{7.137}$$

From this we can find both the pressure and the entropy of the ideal gas. The pressure is found from the derivative of F with respect to V:

$$p = -\left(\frac{\partial F}{\partial V} \right)_{N,T} = \frac{NkT}{V} \;\Rightarrow\; pV = NkT \, . \tag{7.138}$$

Which we recognize as the equation of state for the ideal gas. Similarly, the entropy is

$$S = -\left(\frac{\partial F}{\partial T} \right)_{V,N} = -\frac{\partial}{\partial T} NkT \left(\ln \frac{V}{N} n_Q(T) + 1 \right) \tag{7.139}$$

$$= Nk \left(\ln \frac{V}{N} n_Q(T) + 1 \right) + NkT \frac{\partial}{\partial T} \ln n_Q(T) \tag{7.140}$$

$$= Nk \left(\ln \frac{V}{N} n_Q(T) + 1 \right) + NkT \frac{\partial}{\partial T} \ln T^{3/2} \tag{7.141}$$

$$= Nk \left(\ln \frac{V}{N} n_Q(T) + \frac{5}{2} \right) \tag{7.142}$$

$$= Nk \left(\ln \frac{V}{N} \left(\frac{2\pi mkT}{h^2} \right)^{3/2} + \frac{5}{2} \right) . \tag{7.143}$$

These results are, of course, the same as we found in the microcanonical system, but now expressed in terms of T, V, and N.

7.4.5 Example: Comparing microcanonical and canonical analysis

We now have all the tools needed to address complex and realistic problems using both the microcanonical and the canonical system. How do the analysis and the results differ for the two cases? We will address this in a model for vacancies.

Real crystals are more complicated than the simple Einstein crystals. Real crystals

Microcanonical approach. How can we adress this system using a microcanonical approach? In the microcanonical system the total energy of the crystal is given. In a simplified description of the system we assume that the energy associated with moving a single atom from the interior to the surface does not depend on the number of atoms n that are moved from the interior to the surface: There is a fixed cost for moving an atom from the interior to the surface. This is only valid if the number of atoms moved from the interior to the surface is small compared with the number of atoms N in total in the crystal, so that it is unlikely that two atoms next to each other are removed or that two atoms are placed next to each other on the surface, because in both these cases we would need to consider the number of present neighbors in order to find the energy to remove a particle from the interior and the energy to place a particle on the surface. In this simplified picture, we assume that the energy of an atom in the interior is $-Z\epsilon$, where Z is the number of neighbors in the interior, and that the energy for an atom on the surface is $-z\epsilon$. For example, $Z = 6$ for a cubic crystal and $z = 1$ for a surface atom on a flat, cubic crystal surface if we only include an energy ϵ for each binding to each of the 6 neighbors in a cubic lattice structure as illustrated in Fig. 7.11.

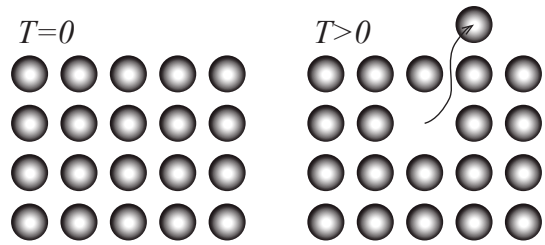


Fig. 7.11 Illustration of energies in a small cubic crystal.

In this case, the energy of the system is given by the number of vacancies n:

$$E = -\left(nz\epsilon + (N-n)Z\epsilon\right) \ . \tag{7.144}$$

In order to develop a thermodynamic description of this model system, we find the multiplicity and the entropy of this model system. The multiplicity is given by the number of possible microstates consistent with the macrostate, that is, consistent with a given number n of vacancies. We can describe a microstate by the positions of the vacancies inside

the crystal and by the position of the vacancies on the surface and find the total multiplicity as the product of the number of ways to place n vacancies in the N interior sites (bulk), Ω_b, and the number of ways to place n vacancies in the approximately $N^(2/3)$ surface sites (surface) Ω_s:

$$\Omega(n,N) = \Omega_b(n,N)\Omega_s(n,N) \ , \tag{7.145}$$

If the number of vacancies n is small compared with the number of atoms, N, in the crystal, then the number of possible ways to arrange n vacancies among N atomic positions is $N!/(n!(N-n)!)$. Since the number of surface sites is much smaller that the number of interior sites ($N^(2/3) \ll N$ when N is large), we ignore Ω_s. (We elaborate on this assumption at the end). The entropy of the system is therefore

$$S = k \ln \Omega = k \ln \frac{N!}{n!(N-n)!} \ . \tag{7.146}$$

We apply Stirling's approximation $\ln(x!) = x \ln x - x$, getting:

$$S \simeq k \left(N \ln N - n \ln n - (N-n) \ln(N-n) \right) \ . \tag{7.147}$$

We use this to find the temperature of the system:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N,V} = \frac{\partial S}{\partial n} \frac{\partial n}{\partial E} \ . \tag{7.148}$$

We find the last term from

$$\frac{\partial E}{\partial n} = \frac{d}{dn} \left(-z\epsilon n - Z\epsilon(N-n) \right) = -z\epsilon + Z\epsilon = (Z-z)\epsilon = \Delta\epsilon \ . \tag{7.149}$$

And we find that

$$\left(\frac{\partial S}{\partial n} \right)_N = k \left(-n/n - \ln n + \ln(N-n) + (N-n)/(N-n) \right) = k \ln \frac{N-n}{n} \ . \tag{7.150}$$

Thus the temperature is

$$\frac{1}{T} = \frac{k}{\Delta\epsilon} \ln \frac{N-n}{n} \ . \tag{7.151}$$

To find the number of vacancies as a function of temperature, we invert the expression for the temperature

$$\frac{1}{T} = -\frac{k}{\Delta\epsilon} \ln \frac{n}{N-n} \ , \tag{7.152}$$

which gives:

$$\frac{n}{N-n} = e^{-\frac{\Delta\epsilon}{kT}} \ , \tag{7.153}$$

which we can solve for n, getting:

$$n = e^{-\frac{\Delta\epsilon}{kT}}(N-n) \ , \tag{7.154}$$

$$n\left(1+e^{-\frac{\Delta\epsilon}{kT}}\right) = Ne^{-\frac{\Delta\epsilon}{kT}} \ , \tag{7.155}$$

$$\frac{n}{N} = \frac{e^{-\frac{\Delta\epsilon}{kT}}}{1+e^{-\frac{\Delta\epsilon}{kT}}} \ . \tag{7.156}$$

Canonical approach. Let us see how we can address this same system as a canonical system. We assume that the crystal of N atoms is in thermal equilibrium with a large reservoir so that it has a constant temperature T. In this case, we do not specify the number of vacancies n for the system, since this would correspond to specifying the total energy of the system. To find the thermodynamic behavior of this system we find the partition function and use this to find the (average) energy of the system. The partition function is given as the sum over all the possible states i of the system. To describe the state we need to decribe the energy ϵ_i for each of the N lattice sites in the system. The energy of the lattice site i is $\epsilon_i = -Z\epsilon$ if an atom occupies the lattice at this site and $\epsilon_i = -z\epsilon$ if the site is empty and the atom is on the surface. (Here, we assume that the number of atoms on the surface is so small, that we can assume that we can describe each site independently of the others. Otherwise, we would need to include the interactions of an atom with all the surrounding atoms in detail). Notice that the number of atoms, N, is not changed, it is only where teh atoms are placed that change. Each lattice site i can therefore be in one of two possible states, and the sum over all possible states must therefore be a sum over all possible lattice sites, $i_1 = 0, 1$, $i_2 = 0, 1$, $i_3 = 0, 1$, ..., $i_N = 0, 1$:

$$Z = \sum_{i_1=0}^{1}\sum_{i_2=0}^{1}\cdots\sum_{i_N=0}^{1} e^{-(\epsilon_{i_1}+\epsilon_{i_2}+\ldots+\epsilon_{i_N})/kT} \tag{7.157}$$

This sum can be split into a product of sums since there are no interactions between the energies in the various sites (this is our assumption):

$$Z = \sum_{i_1=0}^{1} e^{-\epsilon_{i_1}/kT} \sum_{i_2=0}^{1} e^{-\epsilon_{i_2}/kT} \dots \sum_{i_N=0}^{1} e^{-\epsilon_{i_N}/kT} \tag{7.158}$$

$$= \left(\sum_{i_1=0}^{1} e^{-\epsilon_{i_1}/kT} \right)^N \tag{7.159}$$

This is indeed the general result we found above: For a system of independent components, the partition function is the product of the partition functions for each of the components. Here, the total partition function is the partition function for each of the N lattice positions. We find the single-site partition function

$$Z_1 = e^{z\epsilon/kT} + e^{Z\epsilon/kT} = e^{Z/kT} \left(1 + e^{-\Delta\epsilon/kT}\right) \tag{7.160}$$

The energy of the system is then

$$E = -\frac{d \ln Z_N}{d\beta} = N \frac{d}{d\beta} \left(\beta Z + \ln\left(1 + e^{-\Delta\epsilon\beta}\right)\right) = -NZ + \frac{\Delta\epsilon e^{-\Delta\epsilon/kT}}{1 + e^{-\Delta\epsilon/kT}} \tag{7.161}$$

We can use this to find the number n of vacancies, through $E = -NZ + n\Delta\epsilon$, giving

$$n = \frac{e^{-\Delta\epsilon/kT}}{1 + e^{-\Delta\epsilon/kT}} \ . \tag{7.162}$$

Which is exactly the same result as we found for the microcanonical system.

Full model.

7.4.6 Example: Extension of a long polymer

Motivation. A long polymer such as the DNA molecule consists of sequence of connected monomers. The length of a monomer is Δl. We will here address a one-dimensional model for a polymer, consisting of N monomers folded along the x-axis. Each monomer can be folded either to the right or to the left. The number of monomers folded to the right is N_R and the number folded to the left is $N_L = N - N_R$. If one end of the polymer is placed at $x = 0$, the other end will be at a position $x = L$:

$$L = \Delta l \left(N_R - N_L\right) = \Delta l \left(N_R - (N - N_R)\right) = \Delta l \left(2N_R - N\right) \ . \tag{7.163}$$

Thermodynamic identity. This system is therefore not described with a volume, but with a length L instead. How do we need to modify the thermodynamic identity for this? For a quasi-static, reversible process, the first law of thermodynamics gives

$$dE = Q + W \ , \tag{7.164}$$

where $Q = TdS$. The work done on a polymer when it is extended from L to $L + dL$ is $W = fdL$, where f is the force pulling on the polymer. (Notice that this force has the opposite direction to a pressure acting on a piston in the case of an ideal gas). Thus, the thermodynamic identity is

$$dE = TdS + fdL \ \Rightarrow \ TdS = dE - fdL \ , \tag{7.165}$$

Helmholtz free energy. How do we use this to modify the expression for Helmholtz free energy, so that we can find a way to calculate f from F and hence from the partition function for the system?

Helmholtz free energy is given as $F = E - TS$, therefore

$$dF = dE - TdS - SdT = TdS + fdL - TdS - SdT = -SdT + fdL \ . \tag{7.166}$$

This means that we can find the entropy and the force from Helmholtz free energy using:

$$S = -\left(\frac{\partial F}{\partial T}\right)_{f,N} \ , \ f = \left(\frac{\partial F}{\partial L}\right)_{T,N} \ . \tag{7.167}$$

We therefore have all the tools needed to address the behavior of a polymer system — if we only have an expression for Helmholtz free energy.

Partition function. The partition function can be found by summing over all the possible states of the system. The state is described by giving the direction of each of monomer element. We call the state of element i, s_i, which may point in the positive direction, $s_i = +1$, or in the negative direction, $s_i = -1$. The number of steps to the right is then the number of times $s_i = +1$ and the number of steps to the left is the number of times $s_i = -1$.

In this simplest version of a model for a polymer, we assume that the energy of the polymer does not depend on how it is folded or whether it folds onto itself. The energy is just a constant, $N\Delta\epsilon$, where $\Delta\epsilon$ is the energy of a single monomer.

The partition function is the sum over all states consistent with the macroscopic parameters T, L, N. This means that we need to sum the energies of all the states that results in the length L:

$$Z(T, L, N) = \sum_i e^{-\epsilon_i / kT} , \tag{7.168}$$

where $\epsilon_i = N\Delta\epsilon$ is the energy of the state, which is independent on the configuration of the polymer. The sum is over all states that given a total length L of the polymers. This corresponds to a sum over all the states with $N_R = (1/2)\,(N + (L/\Delta l))$ and $N_L = N - N_R$. How many ways can we arrange the monomer so that we have N_R of the states s_i being $+1$ and $N_L = N - N_R$ begin -1? Since the order in which the monomers are folded does not matter for N_R, this is the same as the number N_R of ways that we can select N_R states from N. That is, the partition function sum consists of $\Omega(N_R, N)$ identical elements:

$$Z(T, L, N) = \sum_i e^{-N\Delta\epsilon/kT} = \binom{N}{N_R} e^{-N\Delta\epsilon/kT} . \tag{7.169}$$

Helmholtz free energy. We can now find Helmholtz free energy from

$$F = -kT \ln Z = -kT \left(\ln \frac{N!}{N_R!(N - N_R)!} - \frac{\epsilon}{kT} \right) = N\Delta\epsilon - -kT \ln \frac{N!}{N_R!(N - N_R)!} \tag{7.170}$$

We apply Stirling's approximation, $\ln(x!) = x \ln x - x$, to find an expression we can take the derivative of

$$F = N\Delta\epsilon - kT \left(N \ln N - N_R \ln N_R - (N - N_R) \ln(N - N_R) \right) . \tag{7.171}$$

The force f on the polymer. We find the force f from Helmholtz free energy

$$f = \left(\frac{\partial F}{\partial L} \right)_{T,N} = \left(\frac{\partial F}{\partial N_R} \right)_{T,N} \left(\frac{\partial N_R}{\partial L} \right) , \tag{7.172}$$

where

$$\left(\frac{\partial N_R}{\partial L} \right) = \frac{d}{d_L} \left((1/2)\,(N + (L/\Delta l)) \right) = \frac{1}{2\Delta l} . \tag{7.173}$$

and

$$\left(\frac{\partial F}{\partial N_R}\right)_{T,N} = \frac{\partial}{\partial N_R}\left(N\Delta\epsilon - kT\left(N\ln N - N_R\ln N_R - (N-N_R)\ln(N-(N_R)\right)\right) \tag{7.174}$$

$$= -kT\left(-(N_R/N_R) - \ln N_R + \ln(N-N_R) + 1\right) \tag{7.175}$$

$$= +kT\ln\frac{N_R}{N-N_R}\,. \tag{7.176}$$

We can now replace N_R with $N_R = (1/2)\,(N + (L/\Delta l)$, giving

$$f = \frac{kT}{2\Delta l}\,\ln\frac{N_R}{N-N_R} \tag{7.177}$$

$$= \frac{kT}{2\Delta l}\,\ln\frac{(N+L/\Delta l)/2}{(N-L/\Delta l)/2} \tag{7.178}$$

$$= \frac{kT}{2\Delta l}\,\ln\frac{(N+L/\Delta l)}{(N-L/\Delta l)} \tag{7.179}$$

$$= \frac{kT}{2\Delta l}\,\ln\frac{(1+L/N\Delta l)}{(1-L/N\Delta l)} \tag{7.180}$$

This force is a purely entropic force. The force is due to the forced uncurling of the polymer.

7.4.7 Example: Ising model in 1d and 2d

Motivate the introduction of Monte Carlo methods in a system where it is not simple to find the analytical solutions to the partition function.

7.4.8 Example: Grain boundaries

Make a simple model for grain formation and grain boundaries. Develop a 1d model. Make a 2d model. Simulate using MC methods. Show coalesence and grain boundary dynamics in this model. How to describe these types of processes?

7.5 On the great importance of kT

Discuss how we now can interpret the meaning of kT and how we can use this to analyze the behavior of systems qualitatively.

7.6 Transformation between ensembles

We introduced Helmholtz free energy, F, as a function with the natural variables T, V, N. The method we used to develop F from E is a general method, that we can use to gain a very useful insight into the relation between the various *free energies* that we will use in thermodynamics. The various free energies may often seem unmotivated – you may remain with the impression that they were dreamed up or that we arrived at them in some unspecified way. Not so – they are all derived in a very natural way.

Let us start from the microcanonical system. This system is thermally isolated and described by the state variables E, V, N. These three state variables provide the complete state of the system. From these state variables – and from the relation $S = S(E, V, N)$, we can calculate other state variables such as T, p, and μ for the system.

From a thermodynamic perspective – where $S(E, V, N)$ is a sharp function of E, V, and N – we could instead describe the system with the variables S, V, and N, and then find the energy E from a relation $E = E(S, V, N)$. We can use the thermodynamic identity to see how we can gain all the thermodynamic properties from $E(S, V, N)$ in this case:

$$TdS = dE + PdV - \mu dN \;\Rightarrow\; dE = TdS - PdV + \mu dN \; . \qquad (7.181)$$

We use these differential forms to find what are the variables describing the system. Here is it S, V, and N, since these are the varibles for dE. And we can also use this identity to read out the partial derivatives of E:

$$\left(\frac{\partial E}{\partial S}\right)_{V,N} = T \; , \; \left(\frac{\partial E}{\partial V}\right)_{S,N} = -P \; , \; \left(\frac{\partial E}{\partial N}\right)_{S,V} = \mu \; . \qquad (7.182)$$

Now, we would like to describe the system with the variables T, V, and N instead of S, V, and N, just as we did for the canonical ensemble. What is then the quantity corresponding to E – the quantity A that when we take the differential of A we end up with T, V, and N as the free variables. We want to develop a function $A = A(T, V, N)$ so that

$$dA = \left(\frac{\partial A}{\partial T}\right)_{V,N} dT + \left(\frac{\partial A}{\partial V}\right)_{T,N} dV + \left(\frac{\partial A}{\partial N}\right)_{T,V} dN \; . \qquad (7.183)$$

How can we create such a function and what would the partial derivatives be? Fortunately, there is a general theory in the form of Legendre transforms that allow us to transform a function $E(S, V, N)$ into a new

function $A(T,V,N)$, where T is the partial derivative of E with respect to S. This transform is called the Legendre transform. In general, if we have a function $f = f(x,y,z)$ and we want to replace x with its *partial derivative conjugate*, t:

$$t = \left(\frac{\partial f}{\partial x}\right)_{y,z} . \tag{7.184}$$

This is achieved by the Legendre transform:

$$\mathcal{L}f(x,y,z) = g(t,y,z) = f(x(t,y,z),y,z) - xt . \tag{7.185}$$

This is what we do for the transformation between the microcanonical and the canonical ensemble: We move from a description of the system in terms of $E(S,V,N)$ to $A(T,V,N)$, by introducing

$$A(T,V,N) = E(S(T,V,N),V,N) - S(T,V,N)\left(\frac{\partial E}{\partial S}\right)_{V,N} = E - TS , \tag{7.186}$$

and this is what we called the Helmholtz free energy. Let us check that this free energy indeed has the correct dependent variables by finding the differential for F using the thermodynamic identity:

$$\begin{aligned} dF &= d(E-TS) \\ &= dE - TdS - SdT = dE - (dE + pdV - \mu dN) - SdT \\ &= -SdT - pdV + \mu dN , \end{aligned} \tag{7.187}$$

which indeed shows that the dependent variables – the variables occuring in the differentials on the right hand side, dT, dV, and dN – are indeed T, V, and N as we wanted.

Now, we are ready to examine other possible combinations of state variables. What if we want to study a system where the volume is no longer constant – so that we want to replace V with its partial derivative conjugate. What is the conjugate of V? It is $-p$:

$$-p = \left(\frac{\partial E}{\partial V}\right)_{V,N} . \tag{7.188}$$

We can now use the Legendre transform to generate the potential

$$H = E - V\left(\frac{\partial E}{\partial V}\right)_{S,N} = E - V(-p) = E + pV . \tag{7.189}$$

This potential is called the **enthalpy** of the system. Let us check that it has the correct dependent variables by finding the differential for H:

$$\begin{aligned} dH &= d(E+pV) \\ &= dE + pdV + Vdp \\ &= (TdS - pdV + \mu dN) + pdV + Vdp \\ &= TdS + Vdp + \mu dN \ . \end{aligned} \tag{7.190}$$

Indeed, the dependent variables are S, p, and N.

Let us look at two further systems that we will introduce later on. What if we want to study systems at constant T, p, N – that is, we want to replace V with p in F? In that case, we apply the trasform to F, getting

$$G = F + pV = E - TS + pV \ , \tag{7.191}$$

which is called **Gibbs free energy**.

Similarly, we will want to study chemical reactions in open systems with constant T, V, and μ, but where N may vary. This corresponds to changing from N to μ in F, again applying the transform we get

$$\Phi = F - \mu N \ , \tag{7.192}$$

which is called the **Grand potential**.

It is also useful to develop an interpretation of *free energy as the energy needed to generate a system out of nothing*. For the microcanonical ensemble, the energy E is the energy we need to generate the system.

For the canonical system, we also need to generate the energy E of the system, but now we do this while in equilibrium with a large reservoir with temperature T – we therefore get some energy for free from the reservoir, in the form of heat transfer from the reservoir. This corresponds to TS. The energy needed to create a canonical system with T, V, and N is therefore given by Helmholtz free energy, $F = E - TS$. This is also the energy we would be able to get out of the system if we annhilated the system – again we would gain the energy E, but we would now lose the energy TS to the reservoir.

For a system at constant E, p, N – corresponding to the microcanonical system, but at constant pressure instead of constant volume, we need to generate the energy E, but in addition we need to make room for the volume V of the system, which requires work done by the system on the environment, and this work is pV. The energy needed to create the system is therefore $H = E + pV$, the **enthalpy** of the system.

For a system in equilibrium with a heat bath with temperature T and in an atmosphere/environment with pressure p, we need to generate the energy E and we need to do a work pV to make room for the system, but since we are in equilibrium with a heat bath, we get a thermal energy of TS from the reservoir. The total energy needed is therefore $G = E - TS + pV$, which is the **Gibbs free energy** of the system.

We therefore see that the free energies are the energies needed to generate a system in various conditions. Tables of free energies are therefore useful tools to determine the energy needed to create various systems or to go from one system to another system through chemical reactions or physical transformations, as demonstrated by the following examples.

7.6.1 Example: Fuel cell

Fuel cells are used in cars, such as the BMW 750Hl. In a fuel cell hydrogen and oxygen is combined to form water:

$$\mathrm{H_2} + \frac{1}{2}\mathrm{O_2} \rightarrow \mathrm{H_2O}\ . \tag{7.193}$$

What is the energy produced in this reaction, if it occurs at constant temperature $T = 298\mathrm{K}$ and constant pressure?

In this case, we know that the energy we get from creating 1 mole of water, while removing 1 mole of hydrogen and half a mole of oxygen, is given by the Gibbs free energy. The energy from the reaction will be

$$\Delta G = \Delta G(H_2O) - \Delta G(H_2) - \frac{1}{2}\Delta G(O_2)\ . \tag{7.194}$$

In the table for the Gibbs free energy for water in Schroder, it is the total ΔG which is listed. We find that for water $\Delta G = -237kJ$ per mole. It is negative, because to generate water from its basic compounds, we gain energy.

Since this process occurs at constant temperature corresponding to room temperature, we know that Gibbs free energy has two components, one component from the change in enthalpy and one component from the change in entropy:

$$\Delta G = \Delta H - T\Delta S\ , \tag{7.195}$$

We therefore lose some of the energy to heat. How much energy is converted to heat? We can estimate this in two ways: (i) We can simply use the change in enthalpy tabulated in Schroder – this gives us the total energy produced in a system that is not in thermal contact with a heat bath; or (ii) We can calculate the changes in entropy also based on entropy values from Schroder.

The change in enthalpy for the reaction is

$$\Delta H = -286\text{kJ} \, . \tag{7.196}$$

This means that the entropy change, due to heat, is $\Delta G = \Delta H - T\Delta S$, giving $T\Delta S = Q = \Delta H - \Delta G = -286\text{kJ} + 237\text{kJ} = -49\text{kJ}$, where the negative sign means that the heat is transferred from the system to the environment/reservoir in the process. This means that the efficiency of this process is 237/286, which is 0.83, which is very high!

7.6.2 Example: Muscle cell

We can think of a muscle cell as a fuel cell, but using different chemicals than the hydrogen fuel cell, of course. The general form of reaction in the muscle cell is to use glucose to generate water and CO_2:

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O \, . \tag{7.197}$$

If we assume that the reaction occurs at room temperature and at one atm pressure, what is the enthalpy and Gibbs free energy for this reaction?

How can we find ΔH and ΔG for such compound processes? In the tables in Schroder, and in other sources, the values for ΔH and ΔG for each compound, such as for glucose, is the enthalpy or Gibbs free energy when the compound is made from each component in its most elemental, stable form. For glucose, we assume that it is made from C (graphite), O_2 and H_2. We can therefore find the ΔH or ΔG for the reaction, by first converting each compound into their elemental forms and the converting from the elemental forms and into the new compound – the result of the chemical reaction. We can therefore find what it would take to turn the left side into its basic components, which is the negative of what it would take to create the left side from its basic compents. We call the enthalpy and Gibbs free energy of the left side ΔH_L and ΔG_L respectively – the energies needed to create the left side. Similarly, ΔH_R and ΔG_R are the energies needed to create the right side from its basic

components:

$$\Delta H_L = \Delta H_{C_6H_{12}O_6} + \Delta H_{6O2} = -1273\text{kJ} + 6\,0\text{kJ} = -1272\text{kJ}\ , \quad (7.198)$$

$$\Delta G_L = \Delta G_{C_6H_{12}O_6} + \Delta G_{6O2} = -910\text{kJ} + 6\,0\text{kJ} = -910\text{kJ}\ . \quad (7.199)$$

Similarly, for the right hand side:

$$\Delta H_R = \Delta H_{6CO_2} + \Delta H_{6H_2O} = 6\ -393\text{kJ} + 6\ -286\text{kJ} = -4074\text{kJ}\ , \quad (7.200)$$

$$\Delta G_R = \Delta G_{6CO_2} + \Delta G_{6H_2O} = 6\ -394\text{kJ} + 6\ -237\text{kJ} = -3786\text{kJ}\ . \quad (7.201)$$

The change in enthalpy occurs by creating the right side and removing the left side, thus

$$\Delta H = \Delta H_R - \Delta H_L = -4074\text{kJ} + 1272\text{kJ} = -2802\text{kJ}\ , \quad (7.202)$$

and the change in Gibbs free energy is similar:

$$\Delta G = \Delta G_R - \Delta G_L = -3786\text{kJ} + 910\text{kJ} = -2876\text{kJ}\ . \quad (7.203)$$

Gibbs free energy is negative, meaning that this is the energy available for mechanical work done by the muscle on the environment.

We see that the magnitude of Gibbs free energy is larger than the magnitude of the enthalpy. What does this imply? Since $\Delta G = \Delta H - TS = \Delta H - Q$, where Q is the heat transferred into the muscle from the environment due to the energy in the thermal bath, we see that in this case:

$$Q = \Delta H - \Delta G = -2802\text{kJ} + 2876\text{kJ} = 74\text{kJ}\ . \quad (7.204)$$

which is the heat dumped from the environment to the muscle (to help it do work at the same temperature).

7.7 Equipartition and velocity distributions

7.7.1 Equipartition theorem

We can demonstrate how we can use the results from the canonical system to prove the *equipartition theorem*. We address terms in the energy of a system that are quadratic in shape:

$$\epsilon(q) = cq^2\ , \quad (7.205)$$

where q is a degree of freedom, such as the position for a harmonic oscillator, the velocity for the kinetic energy, the angular velocity for the rotational kinetic energy, or any other quadratic term.

We will assume that the states are densely, but discretely placed in q-space (but that they are not degenerate), so that there are states every Δq. The partition function is therefore

$$Z = \sum_q e - \beta c q^2 \ , \tag{7.206}$$

which we will approximate by the integral over all q-values:

$$Z = \frac{1}{\Delta q} \sum_q e^{-\beta c q^2} \Delta q = \frac{1}{\Delta q} \int e^{-\beta c q^2} dq \ , \tag{7.207}$$

we change variable to $x = \sqrt{\beta c} q$, so that $dx = \sqrt{\beta c} dq$, and

$$Z = \frac{1}{\Delta q} \frac{1}{\sqrt{\beta c}} \int_{-\infty}^{\infty} e^{-x^2} dx \ , \tag{7.208}$$

where the integral is simply a numerical value (which we do not need to know) – we call it b.

$$Z = \frac{1}{\Delta q \sqrt{\beta c}} b = C \beta^{-1/2} \ . \tag{7.209}$$

(Notice that we are here only conserned with the scaling with β of this function – not the exact functional form.) The energy is then found from:

$$\bar{E} = -\frac{1}{Z} \left(\frac{\partial Z}{\partial \beta} \right) = \frac{1}{C \beta^{-1/2}} (-1/2) C \beta^{-3/2} = \frac{1}{2} \beta = \frac{1}{2} kT \ , \tag{7.210}$$

and we have proved the equipartition theorem – which is true as long as the states are spaced in q-space with no degeneracy. Notice that this assumption is only reasonable as long as the spacing between levels is much smaller than kT, which is not true for quantum systems at low energies.

7.7.2 Maxwell Velocity Distribution

For a single particle in an ideal gas, we know that the probability for the particle to be in the state (n_x, n_y, n_z) is given by the corresponding Boltzmann term

$$P(n_x, n_y, n_z) = \frac{1}{Z} e^{-a(n_x^2+n_y^2+n_z^2)/kT} \ . \tag{7.211}$$

But what is the probability for the particle to have a particular velocity? What is the velocity distribution of an ideal gas.

Velocity distribution in one dimension. First, let us specify the question more. The translational motion of a particle in an ideal gas is independent in the x-, y-, and z-directions. Hence the state is specified by three, independent, numbers n_x, n_y, and n_z. These numbers are related to the momentum of the particle. From quantum mechanics, we recall that the momentum is $p_x = \hbar(\pi/L)\, n_x$, and similarly for the other directions.

(This can be seen from the expression for the energy. The energy for motion along the x-axis is $\epsilon(n_x) = an_x^2$. This corresponds to the kinetic energy of the particle, and therefore $mv_x^2/2 = p_x^2/2 = \epsilon(n_x) = an_x^2 = \hbar^2/(2m)\,(\pi/L)n_x^2$, which gives $p_x = \hbar\,(\pi/L)\, n_x$ — up to the sign.)

We can therefore relate the velocity to the momentum, $v_x = p_x/m = \hbar(\pi/mL)\, n_x$. What is then the probability for a particular velocity, $P(v_x)$? We know the answer to this. If v_x is exactly equal to a value $v_x = \hbar(\pi/mL)n_x$ for some n_x, then the probability is $P(n_x)$, otherwise the probability is zero. Hmmmm. This is not that practical. Instead, we may ask what the probability is for the velocity to be in some interval from v_x to $v_x + dv_x$, which corresponds to the n_x value to be in the interval from n_x to $n_x + dn_x$. This corresponds to what we called the *probability density* in Sect. 4.5.7:

$$P(v_x \text{ is in range } (v_x, v_x + dv_x)\,) = P(v_x) d_v x \ , \tag{7.212}$$

where we call $P(v_x)$ the probability density for v_x. Fortunately, the probability density along the n_x axis is simple. There is one state at $n_x = 1$, then one at $n_x = 2$, then one at $n_x = 3$. This means that the number of states per unit n_x is one. We can therefore estimate the probility to find a state in the range from n_x to $n_x + dn_x$ as the probility to be in one of the states times the number of states in the interval

$$P(n_x) dn_x = \frac{1}{Z} e^{-an_x^2/kT} dn_x \ , \tag{7.213}$$

The probability to be in one of the states is the Boltzmann factor for n_x and the number of states is simply dn_x — since the density of states is 1 along the n_x axis. The probability for the velocity to be in the range v_x to $v_x + dv_x$ is the probability for n_x to be in the corresponding range from $n_x(v_x)$ to $n_x(v_x) + dn_x$, which is

$$P(n_x)dn_x = P(v_x(n_x))\frac{dn_x}{dv_x}dv_x = \frac{1}{Z}e^{-mv_x^2/2kT}\frac{mL}{\hbar\pi}dv_x = Ce^{-mv_x^2/2kT} \ . \tag{7.214}$$

The probability density for the velocity along the x-axis is therefore a Gaussian function in v_x. The constant C can be determined from the normalization of the probability density.

Velocity distribution in three dimensions. We can use this argument for each of the three dimensions, finding that the probability density for the velocity in three dimensions is

$$P(v_x, v_y, v_z)\, dv_x\, dv_y\, dv_z = C^3 e^{-m(v_x^2+v_y^2+v_z^2)/2kT}\, dv_x\, dv_y\, dv_z \tag{7.215}$$

$$= C^3 e^{-mv^2/2kT}\, dv_x\, dv_y\, dv_z \ , \tag{7.216}$$

where $v = (v_x^2 + v_y^2 + v_z^2)^{1/2}$ is the speed. This distribution is given in the v_x, v_y, v_z-space. What if we instead wanted to know the probability for v to be in some interval from v to $v + dv$. Then we need to know what volume in $\mathbf{v}$-space this corresponds to: It corresponds to a spherical shell from v to $v + dv$, which has a volume $4\pi v^2 dv$ when dv is small. Hence the probability to find the velocity in the range from v to $v + dv$ is

$$P(v)dv = C^3 e^{-mv^2/2kT}\, 4\pi v^2\, dv \ , \tag{7.217}$$

where the factor C is determined by normalization:

$$\int_0^\infty C^3 e^{-mv^2/2kT} 4\pi v^2 dv = 1 \ \Rightarrow \ C^3 = (m/2\pi kT)^{3/2} \ . \tag{7.218}$$

This distribution of speeds is called the **Maxwell distribution** of the speeds of particles in an ideal gas:

The **Maxwell distribution** of the speeds v of a particle in a three-dimensional ideal gas is

$$P(v)dv = \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi v^2 e^{-mv^2/2kT} dv \ . \tag{7.219}$$

Fig. 7.12 illustrates both the distribution of v_x, the velocity along the x-axis, and the distribution of v, the speed of the particle.

Average velocity. Notice that the average velocity, $\mathbf{v}$, is zero, while the average speed, v, is non-zero. This is because the distributions for v_x,

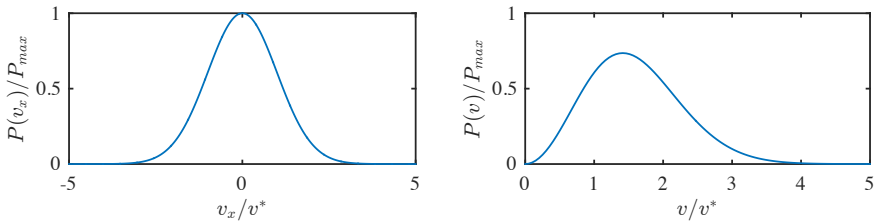


Fig. 7.12 **a** Probability density for the velocity v_x along the x axis for a particle in an ideal gas. **b** Probability density for the speed v of a particle in an ideal gas. Here, $v^* = \sqrt{kT/m}$.

v_y, and v_z are symmetric around zero, while the distribution for the magnitude, v, of the velocity of course is zero when $v < 0$. The average velocity is

$$\langle v_x \rangle = \int_{-\infty}^{\infty} C e^{-mv_x^2/2kT} v_x dv_x = 0 \,, \tag{7.220}$$

and the average speed is (after simplifying the integral):

$$\langle v \rangle = \int_{-\infty}^{\infty} \left(\frac{m}{2\pi kT}\right)^{3/2} e^{-mv^2/2kT} v^2 dv = \sqrt{\frac{8kT}{\pi m}} \,. \tag{7.221}$$

The velocity of N_2 in the air. This means that for air, which mostly consists of nitrogen, the average speed of the nitrogen molecules at $T = 300\,\mathrm{K}$ is:

$$\langle v \rangle = \sqrt{\frac{8kT}{\pi\, m}} = \sqrt{\frac{8 \times 1.38 \times 10^{-23}\,\mathrm{m^2\,kg\,s^{-2}\,K^{-1}} \times 300\,\mathrm{K}}{\pi \times 28 \times 1.66 \times 10^{-27}\,\mathrm{kg}}} \simeq 476\,\mathrm{m/s} \,. \tag{7.222}$$

We could also ask for the probability for a molecule to move with at least twice the average velocity, that is $P(v \geq 2v_a)$, where we have written $v_a = \langle v \rangle$ to simplify the notation. This is estimated by summing up all the probabilities for $v \geq 2v_a$, that is, by integrating $P(v)$ from $2v_a$ to infinity:

$$P(v \geq 2v_a) = \int_{v_a}^{\infty} P(v)\, dv = \int_{v_a}^{\infty} 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} \int_{v_a}^{\infty} v^2 e^{-mv^2/2kT}\, dv \,. \tag{7.223}$$

We change variable to simplify the exponent, introducing $u = v\sqrt{m/2kT}$, $du = \sqrt{m/2kT}dv$, and

$$u_a = v_a\sqrt{m/2kT} = 2\sqrt{8kT/\pi m}\,\sqrt{m/2kT} = 2\sqrt{4/\pi} \simeq 2.26 \,. \tag{7.224}$$

$$P(v \geq 2v_a) = 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} \left(\frac{2kT}{m}\right)^{3/2} \int_{u_a}^{\infty} x^2 e^{-x^2}\, dx \ . \tag{7.225}$$

This integral is solved numerically in python:

```
from scipy.integrate import quad
import numpy as np
def fun(x):
    return x**2*np.exp(-x**2)
quad(fun,2.26,np.inf)
```

```
Out[12]: (0.007454366605316702, 1.0002475807012429e-11)
```

The probability is therefore 0.0075 to have a speed of more that twice the average speed, which corresponds to the probability for the speed of a dinitrogen molecule to be higher than 952 m/s.

7.8 Statistical entropy

From the average energy we get:

$$\bar{E} = \sum_s \epsilon_s P(s) \ , \tag{7.226}$$

Now, the total change in energy is therefore:

$$dE = \sum_s \epsilon_s dP(s) + \sum_s P(s) d\epsilon_s \ , \tag{7.227}$$

Where we can write the second term as:

$$\sum_s P(s) d\epsilon_s = \sum_s P(s) \frac{d\epsilon_s}{dV} dV = -\bar{P} dV \ , \tag{7.228}$$

and therefore the change in energy is:

$$d\bar{E} = \sum_s \epsilon_s dP(s) - \bar{P} dV \ . \tag{7.229}$$

If we compare with the thermodynamic identiy, we see that we can interpret:

$$TdS = \sum_s \epsilon_s dP(s) \ , \tag{7.230}$$

This expressions can (with some tricks) be related to another expression for the entropy – which we could instead have started from:

$$TdS = \sum_s \epsilon_s dP(s) = -kT \sum_s \ln Z dP(s) - kT \sum_s \ln P(s) dP(s) \tag{7.231}$$

where we have used that

$$P(s) = (1/Z) \exp(-\epsilon_s/kT) \ , \tag{7.232}$$

and therefore that

$$\epsilon_s = -kT(\ln Z + \ln P(s)) \ . \tag{7.233}$$

Notice that $\sum_s dP(s) = d \sum_s P(s) = d(1) = 0$. We therefore get:

$$dS = -k \sum_s \ln P(s) dP(s) = -k \sum_s d(P(s) \ln P(s)) \ , \tag{7.234}$$

and

$$S = -k \sum_s P(s) \ln P(s) \ , \tag{7.235}$$

which is a general expression for the entropy. This relation also holds in the microcanonical ensemble, since then $P_s = 1/\Omega$ for all states and therefore we find $S = k \ln \Omega$ after summing.

7.9 General Monte-Carlo framework

Introduction to a general monte-carlo framework that will be used throughout the book to study for Microcanonical systems, Polymers, Random walks, Vacancies, Ising model, Einstein crystals, XY model type systems, Binary mixtures, and Lattice gases.

7.10 Simulations of a canonical system

While molecular dynamics simulations were simple to perform in the microcanonical ensemble, they are not similarly simple in a canonical system because energy is no longer conserved and we need to introduce the effect of a large heat bath. There are several ways this can be done, and we discuss these briefly below. The connection to the heat bath means that the canonical system lends itself more naturally to stochastic simulations methods such as Monte Carlo methods. We will therefore

start with a more systematic introduction to Monte Carlo simulations methods and their application to the canonical system.

Here, we will first demonstrate how we can use Monte Carlo methods for more general purposes, such as for calculating an integral, before we address how we can apply Monte Carlo methods to collect statistical about a canonical system.

7.10.1 Monte Carlo estimates of integrals

How can we find the integral of a function $f(x)$?

$$y = \int_{x_0}^{x_1} f(x)\, dx \ . \tag{7.236}$$

This integral can be interpreted as the integral under the curve $f(x)$ from x_0 to x_1. We know how to solve it numerically by various integration methods, but we can also use a stochastic (Monte Carlo) algorithm to estimate the area.

Hit or miss algorithm. For example, we can use a "hit or miss" algorithm: To estimate the integral we choose a box from $y = 0$ to some $y_1 > f(x)$ for all x on the interval. We then select N points randomly in the interval x_0 to x_1 and $y_0 = 0$ to y_1. We can then estimate the integral by the fraction of points falling below the curve $f(x)$ multiplied by the area of the box:

$$y \simeq \frac{N_{\text{below}}}{N} (y_1 - y_0)\ (x_1 - x_0)\ . \tag{7.237}$$

We expect this to become a better and better estimate as N goes to infinity. Indeed, it can be shown that the error is proportional to $1/\sqrt{N}$ for this method.

Hit and miss integration examples. How does this method work for a simple example? Let us try to integrate the function

$$f(x) = (\sin(10\,x) + \cos(50\,x))^2\ , \tag{7.238}$$

This function, and $N = 1000$ random points, are illustrated in Fig. 7.13. In this figure we also show the estimates integral I as a function of the number of steps i used to estimate the integral. This was found using the following simple implementation of the hit or miss algorithm. Notice that we first use the symbolic package to find the exact solution for comparison.

```
# Symbolic solution
import numpy as np
import matplotlib.pyplot as plt
import sympy as sy
u = sy.symbols('u')
f = (sy.sin(10*u)+sy.cos(50*u))**2
g = sy.integrate(f)
# Find exact integral from 0 to 1
fint = float(g.subs(u,1))
# Convert to Python function
fun = sy.lambdify(u,f,"numpy")
# Plot f(x)
plt.subplot(2,1,1)
xx = np.linspace(0,1,1000)
yy = fun(xx);
plt.plot(xx,yy);
plt.subplot(2,1,2)
# Calculate integral using hit and miss
fmax = 4.0 # Upper limit of f(x) for x (0,1)
N = 1000
hit = 0.
intest = np.zeros((N,1),float)
for i in range(1,N):
    x = np.random.rand();
    y = np.random.rand()*fmax
    if (y<fun(x)):
        hit = hit + 1
    intest[i] = hit/i*fmax
plt.plot(np.arange(0,N),intest,'-r')
plt.plot(np.arange(N),fint*np.ones(N),'k:')
plt.xlabel('i');plt.ylabel('I');
```

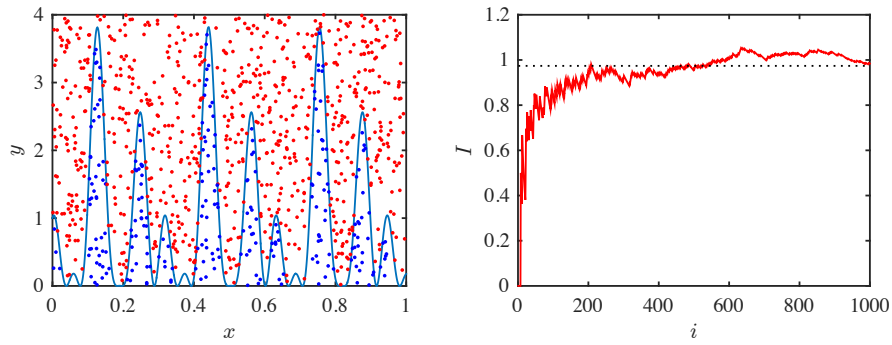


Fig. 7.13 **a** Plot of the function $f(x)$ and the points used to estimate the integral. The color indicates if the point is below $f(x)$ or not. **b** Plot of the estimated integral I as a function of the number of iterations i used for the iteration.

Hit and miss in high dimensions. Why would we calculate the integral in this way? It seems very inefficient. This is true for a one dimensional function. However, for integrals in high dimensions, ordinary integration method does not scale well as the number of dimensions increase, but the stochastic method introduced here is also efficient in high dimensions.

Notice that we do not use this very simple algorithm for real integration. Instead, there are more sophisticated methods that give a more rapid approach to the correct answer. We will not introduce these here for integration, but instead demonstrate examples of more advanced methods, such as importance sampling, for simulations of a canonical system.

7.10.2 Monte Carlo method for a canonical system

For a canonical system, we want to simulate the behavior of a system that is in equilibrium with a heat reservoir of temperature T. In the microcanonical ensemble, we know that all microstates are equally probable, but in the canonical ensemble, the probabilty of a microstate is given by the Boltzmann factors. How can we then sample from the microstates and ensure that we satisfy this particular distribution?

Sampling microstates. We would like to generate a sequence of microstates of the canonical system: A sequence $\{i_1, i_2, \ldots\}$ of states i. We may do this by changing the microstates at random, for example by making a small change to one particle: Flipping a spin, changing the energy of an oscillator, or moving a particle in a gas. However, since the microstates are not equally probable, each of the microstates in the sequence do not weigh equally when we calculate averages from this sequence of microstates. A microstate i of the system has energy E_i and should be weighted according to the probability, $\exp(-E_i/kT)/Z$, of this state. This would for example allow us to estimate the average energy from the sequence $\{i_1, i_2, \ldots, i_M\}$ of (random) states:

$$\bar{E} \simeq \frac{1}{Z} \sum_{j=1}^{M} E_{i_j} e^{-E_{i_j/kT}} \ . \tag{7.239}$$

Unfortunately, we do not know Z, but we can estimate the normalization condition by the measured sum of all states, so that the sum of probabilities for all the states are normalized:

$$Z \simeq \sum_{j=1}^{M} e^{-E_{i_j}/kT} \ . \tag{7.240}$$

The estimate for the energy becomes:

$$\bar{E} \simeq \frac{\sum_{j=1}^{M} E_{i_j} e^{-E_{i_j}/kT}}{\sum_{j=1}^{M} e^{-E_{i_j}/kT}} \ . \tag{7.241}$$

This gives us a reasonable method to perform a statistical sample. We still need the sequence of states to be representative in order for this sequence to give reasonable results, just as we assumed in the microcanonical system. However, this method is in general very slow! Why? Because may of the M states will typically have high energy, and thus have very small probabilities to occur, and therefore contribute very little to both the sum and the average.

Importance sampling. We call the method we have introduced so far a *naive* Monte Carlo method, because we sample completely at random, and hence, we sample many states with low probability. It would be better if we could use the Boltzmann factors to guide our choice of random states, and instead sample mostly the important states. We call methods to do this *importance sampling* methods. The idea is to generate a sequence of microstates that are probable, while still representing a sufficiently random sample of the the microstates.

Metropolis algorithm. An example of an importance sampling method is the Metropolis algorithm, which was introduced by Nicholas Metropolis in 1953 [1]. The Metropolis algorithm generates a sequence of states, $\{i_1, i_2, \ldots\}$. We may start from any random state, i_1. To generate the next state, i_2, we first make a small perturbation of the state i_1, $E'_{i_1} = E_{i_1} + \Delta E$, for example by flipping a spin in a system of magnetic spins, by changing the energy of a harmonic oscillator, or by moving a particle in an ideal gas. If the energy of the new is lower, $\Delta E < 0$ then we accept the change since this will move the system to a more probable state. If the energy is higher in the new state, $\Delta E > 0$, then we accept the change with a probability given as $e^{-\Delta E/kT}$. We then repeat this process to generate a sequence of states.

Subset of states. The Metropolis algorithm generates a subset of states that have high probabilities to occur: States with low energies will occur frequently in the sequence, whereas states with higher energies will be less likely, but not completely improbable. This means that averages

using this sequence may be better than for a completely random sequence — although we cannot really here quantify how much better.

The states in the generated sequence are Boltzmann distributed. However, the Metropolis algorithm not only generates a sequence of states that are better suited for averages. In addition, the states in the sequence occur with a frequency given by the Boltzmann distribution! This means that a state i with energy ϵ_i, will occur with a frequency proportional to $e^{-\epsilon_i/kT}$ in the sequence of states. (The Metropolis algorithm can actually be formulated generally for any distribution, not only the Boltzmann distribution, and it can be used to generate a sequence of numbers that are distributed according to the prescribed distribution).

How can we demonstrate that the frequency of occurence in the sequence is Boltzmann distributed? This is usually demonstrated using the concepts of Markov chains, which are typically introduced in introductory courses in statistical mechanics. The idea is that for two states i and j, where $\epsilon_i \leq \epsilon_j$, the probability for a transition from $j \to i$ according to the Metropolis algorithm is $P(j \to i) = (1/M)$, because this transition will occur, but the probability to choose state j among M possible trial states is $(1/M)$. The probability for a transition from $i \to j$ is $P(i \to j) = (1/M)e^{-(\epsilon_j - \epsilon_i)/kT}$. Thus the ratio of the two transition probabilities are

$$\frac{P(i \to j)}{P(j \to i)} = \frac{(1/M)e^{-(\epsilon_j - \epsilon_j)/kT}}{(1/M)} = \frac{e^{-\epsilon_j/kT}}{e^{-\epsilon_i/kT}} \ . \tag{7.242}$$

This is the ratio of the Boltzmann factors for the two states. However, there are may other states in the system, and there are therefore also many other ways to go from state i to state j. But even if we include any intermediate sequence of states between i and j, the ratio of the probability of a sequence linking state i and j, and its reverse sequence, would still correspond to the Boltzmann factor. Hence, the frequencies at which they occur will also be this exact ratio.

We have therefore argued that the sequence of states generated by the Metropolis algorithm will sample the states with frequencies proportional to their frequency of occurrence — a strong statement which means that we can also use this sequence to gather information about the distribution of states as well as use the sequence to estimate the partition function and average values.

ams 7: The argument as presented here is weak.

Disadvantages of the Metropolis algorithm. It is important to remember that the Metropolis algorithm typically only makes small changes to the states in each step, therefore the states in the sequence will be *correlated*. We would expect that if we only use every Nth state in the sequence, the states would be sufficiently uncorrelated, but we do not yet have a good estimate as to what value of N would be sufficient to generate sufficiently uncorrelated statistics. In addition, we may have started the sequence from an initial state which is unlikely, and we would therefore also need to wait an initial time until the system has reached a stationary and representative state, where the states in the sequence represent the variations in the system in a good way. Again, we do not really know how long we need to wait or what is meant by the word representative here — it may take a very long time to reach reasonable states. However, we will keep these limitations in mind, but still use the Metropolis algorithm to study systems where our analytical methods are insufficient.

Implementation of Metropolis algorithm. We can now implement an algorithm to generate a sequence of states using the following steps of the Metropolis algorithm:

1. Choose an initial microstate, for example, random spins in the Ising magnet or random energies for each oscillator in the Einstein crystal.
2. Make a random change in the microstate, a trial change. Calculate the change of energy going from the old state to the new state, ΔE. For example, for the Ising system we may choose a spin at random and see what happens if we flip it; for the Einstein crystal choose an oscillator at random and change its energy by ± 1.
3. Generate a random number, r, between 0 and 1
4. If $r \leq \exp(-\Delta E/kT)$ accept the change. (This formulation means that we will always accept a change where $\Delta E > 0$.)
5. Repeat

This generates a sequence of states that you can now make measurements on. Let us see how this is implemented for an Ising magnet system.

Application to Ising magnet. An Ising system consists of a set of spins, S_i, at a site i, which interacts with neighboring spins as well as with an external magnetic field, H, so that the total energy of an Ising system with spins $(S_1, S_2, \ldots, S_N)$ is

$$\epsilon = \sum_i \left(S_i \sum_{j:i} -JS_j - H \right) , \tag{7.243}$$

where the sum over $j : i$ are over all the neighbors of site i. We can then implement the Metropolis algorithm to generate a sequence states for the Ising system by selecting a spin, and then find the probability for flipping this spin to be either up or down, and the flip the spin according to these probabilities:

- Choose a random site i
- Find the energy of this site if $S_i = 1$ and if $S_i = -1$, E_+ and E_-.
- Set the spin up with probability $\exp(-\beta E_+)/C$ and down with probability $\exp(-\beta E_-)/C$ with $C = \exp(-\beta E_+) + \exp(-\beta E_-)$.
- Repeat.

ams 8: Illustrate this algorithm with a figure.

How do we find the energy of the site? We calculate it from the energy function. We therefore need to know $m_i = \sum_j S_j$ for the neighbors of site i and calculate $E_+ = -Jm_i - H$ and $E_- = Jm_i + H$, where $H = \mu B$. We can do this for one site i at a time, as is done in the following script. Resulting simulation patterns are shown in Fig. 7.14.

```
# Heat bath Monte Carlo Ising model one spin at a time
import numpy as np
import matplotlib.pyplot as plt

nstep = 100     # Number of MC steps
N = 100         # System size
Jdivk = 2.0/3.0 # Interaction
Hdivk = 0.0     # External field
T = 0.1         # Dimensionless temperature
JdivkT = Jdivk/T
HdivkT = Hdivk/T

# Make initial random spin configuration
spins = np.random.randint(0,2,(N,N))*2-1
N2 = N*N

# Function to evolve system one time step
from numba import jit
@jit
def evolve_spins():
    Nx = spins.shape[0]
    Ny = spins.shape[1]
    N2 = Nx*Ny
    for jstep in range(N2):
        # Select a spin at random
```

```
        posx = np.random.randint(0,Nx)
        posy = np.random.randint(0,Ny)
        # Make list of neighbors
        neighbors = (np.array([np.mod(posx+1,Nx),np.mod(posx-1,Nx),posx,posx]),np.array([posy,posy
        sumneigh = sum(spins[neighbors])
        thisspin = spins[posx,posy]
        DeltaE = -JdivkT*thisspin*sumneigh - HdivkT*thisspin
        change = (np.random.rand()<np.exp(DeltaE))*(-2)+1
        spins[posx,posy] = spins[posx,posy]*change

# Initialize graphics
%matplotlib notebook
fig = plt.figure()
ax = fig.add_subplot(111)

# Evolve the spins nstep times and visualize the dynamics
for istep in range(nstep):
        evolve_spins()
        ax.clear()
        ax.imshow(spins.T)
        fig.canvas.draw()
```

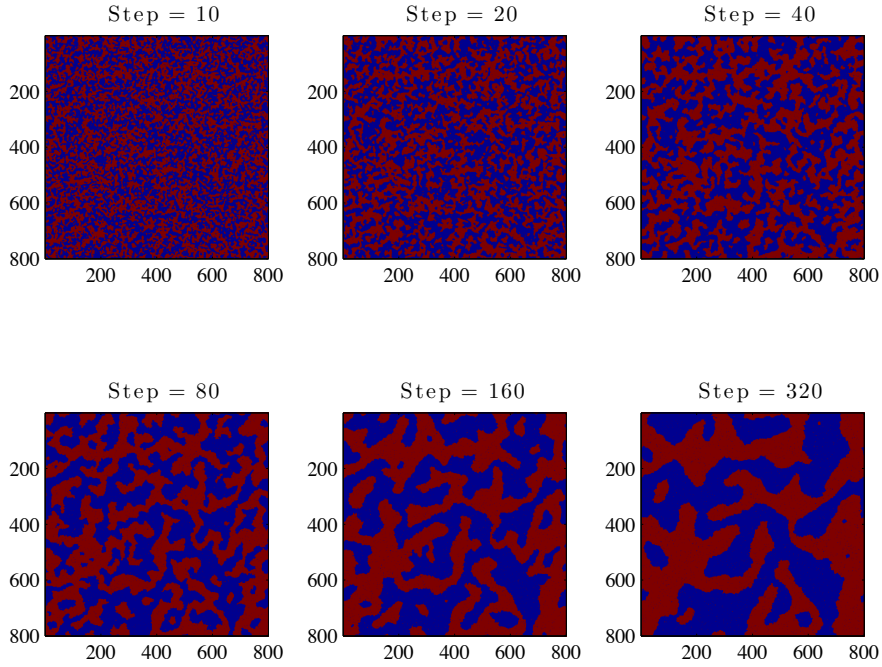


Fig. 7.14 Illustration of spins in the Ising model for $T = 0.8$.

7.10.3 Detailed balance

Are there any limits to what kinds of rules we can choose for the transitions between states like we did in the Metropolis algorithm? Could we invent any rule we like?

No! The transition rates must satisfy a principle called *detailed balance*. What is this?

What do we want to model? We want to model a sequence of microstates for the system – like a time development of microstates. Let us write the states as j: The possible states of the system is 1, 2, 3, etc. If the system is in state i at time t we can write this as $i(t)$.

We want to find the next state, $j(t+\Delta t)$. Then, we need the probability to get to state j for all possible configurations at time t. And not only that, in general we may consider that case that the probability to be in state j at time $t + \Delta t$ depends not only on the state the system was in at a time t, but also all the previous states at earlier time. Hmmm. That was a lot of information. We can simplify the problem if we remove the history dependence – if we assume that the probability to be in a state $j(t + \Delta t)$ only depends on the state $i(t)$ it was in at time t and not on what happened before that. We can think of this as implying that the state $i(t)$ contains all the relevant information about the system and determines what new states can occur. If the state only depends on the previous state and not on the history we call the sequence of states a *Markov chain*.

What we need is the probability $P(j, t + \Delta t)$ to be in state j for each possible state i at t. And we need this probability for all states j and i. We call the probablity to go from state i at t to state j at $t + \Delta t$ the transition probability:

$$W_{ij} = W\left(i \to j\right) = P(j(t + \Delta t)|i(t)) \ , \tag{7.244}$$

where the symbol | means "given that" — it is a conditional probability.

We know that $W_{ij} \geq 0$ and that $\sum_j W_{ij} = 1$: Starting from state i at time t we know that it must end up in some state j at time $t + \Delta t$.

The probability $P(j, t + \Delta t)$ is therefore:

$$P(j, t + \Delta t) = \sum_i P(j, t + \Delta t|i, t) = \sum_i W_{ij} P(i, t) \ . \tag{7.245}$$

What determines how this system develops? The master equation – a cool name – which is given as

$$\frac{dP(j,t)}{dt} = -\sum_i W_{ji} P(j,t) + \sum_i W_{ij} P(i,t) \ . \tag{7.246}$$

ams 9: Remove or finish this discussion.

7.11 Summary

7.12 Exercises

Exercise 7.1: Numerical values for the partition function

Consider an atom with two possible states, a ground state with energy 0 and an excited state with an energy of 1 eV.

a) Find the partition function, $Z(T)$, for this system.

b) Plot the partition function as a function of temperature. What are the numerical values for $T = 300\,\mathrm{K}$, $T = 3000\,\mathrm{K}$, and $T = 30,000\,\mathrm{K}$?

c) Plot the probability for the system to be in its ground state as a function of T.

Exercise 7.2: Changing the energy scale

An atom can be in three states with energies $-\epsilon$, 0, and ϵ.

a) Find the partition function for this atom. Plot the probability for the system to be in each of the three states as a function of kT/ϵ.

We are free to choose the energy scale. If we shift the energy scale by ϵ, the atom may instead be in states with energy 0, ϵ, and 2ϵ.

b) Find the partition function for this atom now and plot the probability for the system to be in each of the three states as a function of kT/ϵ.

c) What aspects of this problem change when we change the energy scale, and what aspects do not change?

Exercise 7.3: Thermal excitations in the Hydrogen atom in the Sun

The transition energies between levels in the Hydrogen atom are given by the Rydberg formula

$$\frac{1}{\lambda} = Rz^2 \left(n_1^{-2} - n_2^{-2} \right) \ , \tag{7.247}$$

where z is the atomic number and $R = 1.097373\,10^7\,\mathrm{m}^{-1}$. The first few energy levels are $\epsilon_1 = -13.6\,\mathrm{eV}$, $\epsilon_2 = -3.4\mathrm{eV}$, $\epsilon_3 = -1.51\,\mathrm{eV}$, and $\epsilon_4 = -0.85\,\mathrm{eV}$. The degree of degeneracy is $g(n) = n^2$ — and double that if spin degeneracy is included. The temperature in the Sun's atmosphere is typically $T = 5800\,\mathrm{K}$.

a) What is the ratio of probabilities for a Hydrogen atom to be in its first excited states compared to its ground state? For every billion atoms in the ground state, how many are approximately in the first excited state?

Exercise 7.4: Thermal excitations of the water molecule

A water molecule can vibrate in many ways. Typical models for water include both the vibration of the positions of each of the hydrogen atoms relative to the oxygen atom as well as the flexing of the angle between the two OH bonds — this angle vibrates around the average angle of 104° with a typical frequency of $\omega = 3.0\,10^{12}\,\mathrm{s}^{-1}$. You may assume that this vibration can be modelled as a harmonic oscillator with energy levels $\epsilon(n) = \hbar\omega(n + 1/2)$.

a) Find the partition function Z for $T = 300\,\mathrm{K}$. (It is sufficient only to include the first three terms in the sum).

b) What is the probability for the water molecule to be in flexing states $n = 0, 1, 2$?

c) Compare this with the vibrational states for the OH bond, for which $\omega \simeq 3.4\,10^{13}\mathrm{s}^{-1}$. Comment on the results.

Exercise 7.5: Three-level system

We will study a system consisting of a single particle that may be in three possible states with energies $-\epsilon$, 0, and ϵ.

a) Find the partition function, $Z(T, V, N)$, for the system.

b) Find Helmholtz free energy for the system, $F = F(T, V, N)$.

c) Find the entropy from Z. What is the limiting value when $kT/\epsilon \to \infty$, and when $kT/\epsilon \to 0$. Comment on the results?

d) Find the energy, E, for the system from a sum over the probabilities for each state and show that the result is in accordance with $F = E - TS$, with the expressions for F and S you found above.

Exercise 7.6: Partition function of two systems

Consider a system consisting of two parts, A and B. Show the under certain conditions, we can write the partition function of the system as a product of the partition functions for each of the parts, $Z = Z_A\, Z_B$.

Exercise 7.7: Arrhenius processes

Many processes in chemistry, biology and physics are considered to be thermally activated processes. In this case, the rate of the process is controlled by an energy barrier that the partners in the process must overcome in order for the reaction/process to happen. The energy barrier is called the activation energy, $\Delta\epsilon$. The probability for two partners in a process to attain the energy $\Delta\epsilon$ is given by the Boltzmann factor, $e^{-\Delta\epsilon/kT}$.

a) Show that if the process can occur at an energy of $\Delta\epsilon$ or higher, the rate of the process will be proportianal to the Boltzmann factor, $e^{-\Delta\epsilon/kT}$.

Experimental results are often analyzed in an Arrhenius diagram. You draw $1/T$ along the x-axis, and the logarithm of the measured quantity along the y axis.

b) Fig. 7.15 shows experimental measurements of the etching velocity of silicon in a chemical etch ($HF:HNO_3:H_2O$). Find the activation energy for the process.

Exercise 7.8: Energy fluctuations

Here, we will address energy fluctuations in a system held at constant (T, V, N) — that is, for a canonical system.

a) Show that the energy fluctuations can be expressed as

$$\langle(\epsilon - \langle\epsilon\rangle)^2\rangle = (kT)^2 \left(\frac{\partial E}{\partial T}\right)_V . \tag{7.248}$$

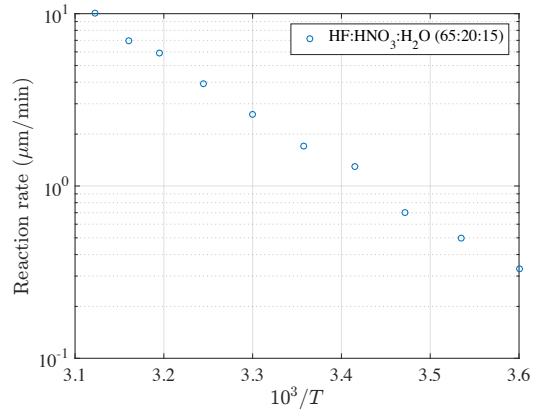


Fig. 7.15 Measurement of the reaction rate as a function of temperature for Silicon in a chemical etch of ($HF:HNO_3:H_2O$).

b) Estimate the typical relative fluctuation in the energy of a monatomic ideal gas with a volume of 1 liter at room temperature and atmospheric pressure.

Exercise 7.9: Hydrogen gas

The H_2 molecule has vibrational states with energies

$$\epsilon_i \simeq \epsilon \left(1.03i - 0.03i^2\right) \ , \ i = 0, 1, 2, \ldots, 15 \tag{7.249}$$

(There are no more discrete levels beyond $i = 15$ because the molecule disassociates — it breaks up.)

a) Plot ϵ_i/ϵ as a function of i. How does the spacing between energy levels change with i?

b) Write a computer program to calculate the partition function as a function of kT/ϵ.

c) Calculate the average energy and the heat capacity, C_V, for the Hydrogen molecule and plot the results as functions of kT/ϵ.

d) Compare the results to that of an harmonic oscillator with the same energy difference between the ground state and the first excited state.

Exercise 7.10: One-dimensional ideal gas

We study a one-dimensional ideal gas with N particles. The particles can move freely along the x-axis over a distance L. Ignore the y- and z-directions.

a) Find the partition function of the gas.

b) Find the energy of the gas at the temperature T.

c) Find the entropy for the gas.

Assume that the N particles can move freely in the x- and the y-directions over an area L^2.

d) Find the partition function, the energy and the entropy for the gas.

If we limit the dimensions of the gas, we can also limit the degrees of freedom of the gas we are studying. Quantum mechanically, we assume that the wave function and the Hamilton operator can be separated into x-, y-, and z-components, $\Psi(\mathbf{r}) = \Psi(x)\,\Psi(y)\,\Psi(z)$. The eigenvalues for the energy can then also be written as $E = E_{xy} + E_z$. The two-dimensional case we could therefore consider to be a result of shrinking on of the edges, L_z, of a cube with edges L. Then we know that (i) the rest energy in the z-direction is large, (ii) the distance between the energy levels in the z-direction is large. We can therefore assume that all the particles will be in the ground state in the z-direction.

e) Show that a good partition function for the two-dimensional case will be

$$Z = Z_{2D}\ \left(\exp\left(-\epsilon_{z,0}/kT\right)\right)^N\ , \tag{7.250}$$

where Z_{2D} is the partition function found for the two-dimensional system above, and $\epsilon_{z,0}$ is the ground state for motion in the z-direction.

f) Find the energy using the partition function in (7.250).

Exercise 7.11: Heat capacities of an ideal gas

The heat capacity at constant pressure and volume are defined as

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V\ ,\ C_p = \left(\frac{\partial E}{\partial T}\right)_p\ , \tag{7.251}$$

respectively.

a) Show that for an ideal gas $C_p - C_V = Nk$.

Exercise 7.12: Compressibility of an ideal gas

The compressibility κ_x for an process where x is constant is defined as:

$$\kappa_x = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_x . \tag{7.252}$$

a) Find an expression for the isothermal compressibility for an ideal gas.

b) May the comressibility be negative for a system in equilibrium?

Exercise 7.13: Equipartion for a linear degree of freedom

We consider a system where the energy depends on the absolute value of the degree of freedom, x, and not on the square of the degree of freedom: $\epsilon(x) = cx$, where c is a constant. (For example, for a relativistic particle, the kinetic energy is proportional to its momentum). Show that the equipartition theorem for this system gives an average energy of kT per such a linear degree of freedom.

Exercise 7.14: Speed of an oxygen molecule

a) Show that the most likely speed of a molecule is $v = \sqrt{2kT/m}$ according to the Maxwell speed distribution.

b) Find the average speed, the most probable speed, and the standard deviation of the speed of an oxygen molecule at room temperature.

c) What is the fraction of oxygen molecules travelling at less than 300 m/s at room temperature?

Exercise 7.15: Escape velocities

We will here address the escape velocity of particles in the Earth's upper atmosphere, at a distance of about $R = 6\,400\,textkm$ from the center of the Earth. The temperature at this height is approximately $T = 1000\,\mathrm{K}$.

a) What is the espace velocity of a particle at this distance from the Earth's center?

b) What is the probability for a nitrogen molecule, a hydrogen molecule and a helium atom to have a velocity that is higher than the escape velocity? Discuss the consequences.

c) Compare with the escape velocity on the moon's surface ($R = 1\,737\,\mathrm{km}$) and discuss the consequences of your results.

Exercise 7.16: Informational entropy

In information theory, the entropy of a signal or the entropy of a discrete random variable with possible states (outcomes) s_i is defined as

$$H = -\sum_i P(s_i) \ln P(s_i) \ , \tag{7.253}$$

where the sum is over all the states i of the system and $P(s_i)$ is the probability for state i with outcome s_i.

a) Show that this formula corresponds to Boltzmann's formula $S/k = \ln \Omega$ for the microcanonical system where the probability for all accessible states are the same, $P = 1/\Omega$.

b) Show that this formula also corresponds to the entropy for a Canonical system in thermal equilibrium at temperature T, where the probability for a state s with energy ϵ_s is $P(s) = (1/Z)e^{-\epsilon_s/kT}$.

7.13 Projects

Thermodynamics and Heat engines 8

We have now developed the microscopic foundation of thermodynamics. We have demonstrated the arrow of time and basic concepts such as entropy and free energy and how to calculate them based on a microscopic theory. We can, for example, calculate the entropy of a microcanonical system or the Helmholtz free energy for a canonical system, and then find other state variables through derivatives of these quantities. However, there are many ways we can determine the state variables and the equations of state for macroscopic systems: It can indeed be through a microscopic theory, it can be through simulations, or it can be through direct, experimental measurements. Thermodynamics addresses how we can describe macroscopic systems when we assume that the basic laws of thermodynamics are known as well as equations of state for all the systems involved.

In this chapter, we will review thermodynamics, and then see how we can apply theromodynamics to discuss heat machines, such as steam engines, diesel engines, stirling engines, and heat exchangers. We will place particular emphasis on the stirling engine, which is elegant in design and frequently used for solar energy applications. Indeed, the best solar energy stirling machines have efficiencies around 30 percent, which is quite amazing given the primitivity of a mechanical machine compared with high technology approaches such as solar cells.

8.1 Thermodynamics

In thermodynamics we discuss the behavior of a **system** , which is a part of the universe, which is separated from the rest of the universe, which we call **the environment**. The system may be **thermal contact** with the environment, if it is allows to exchange energy with the environment, or it may be **thermally insulated**, if there is no energy exchange with the environment. An isolated system has interactions with the enviroment, either through changes in shape of the system; through changes in the number of particles in the system, or through changes in the energy in the system. For an isolated system the energy, E, volume, V, and number of particles, N, are constant.

Thermal equilibrium. Two systems are in thermal equilibrium with each other if they are in thermal contact with each other and we have waited for a long time. What characterizes thermal equilibrium? In thermal equilibrium (or thermodynamic equilibrium), the two systems have the same temperature. For two systems A and B in thermal equilibrium, we know that $T_A = T_B$. The **zeroth law of thermodynamics** states that if two systems are in equilibrium with a third system, they are also in thermal equilibrium with each other. If the two systems A and B both are in thermal equilibrium with system C, this means that $T_A = T_C$ and $T_B = T_C$, and therefore $T_A = T_B$. They are therefore also in thermal equilibrium with each other.

State variables. We describe thermodynamic system with **state variables** , which describe the macroscopic state (macro-state) of a system. State variables can be variables such as energy, E, pressure, p, volume, V, entropy, S, number of particles, N, chemical potential, μ etc. There are two types of state variables: **extensive** and **intensive** variables. Extensive variables depend on (are proportional to) the number of particles, N, such as energy, E, entropy, S, and, of course, the number of particles, N. Intensive varialbes do not depend on the number of particles, such as the density, ρ, the pressure, p, the temperature, T, and the chemical potential, μ.

Equations of state. We describe the behavior of a thermodynamic system though **equations of state**, which are equations that relate various state variables. For example, for the ideal gas, we know that there is an equation relating pressure, P, volume, V, the number of particles, N, and temperature, T. In general, we can write this equation of state as

$$f(P, V, N, T) = 0 \ , \tag{8.1}$$

where $f()$ is some function. For the ideal gas, we know this function — we know the equation of state:

$$PV - NkT = 0 \ . \tag{8.2}$$

We will develop many similar equations of state for other systems throughout this text.

State variables are path independent. It is important to notice that any state variable depends on the state of the system, but not on how the system got to this state. If we say that the energy $E(S, V, N)$ is a state variable, we know that it is always the same for the same set of values for S, V, and N, independent of how the system got there. This is different for the work performed on a system or the thermal enegy transfered to a system. These depend on how the system moves from one state to another state, and are therefore not state variables.

Processes. Changes in the state of the system occur through **processes**, which correspond to changes in the state variables of the system. We can sketch such changes in various ways, as illustrated in Fig. 8.1. We can describe the state of a system in P, V-space, as we will often do for machines. The state of the system is then given by a pair P, V of coordinates. (We have then typically assumed that N is constant). Or we can describe the system in S, T space. A process is then a path from one state to anther state in the system, as shown as lines in the sketch. The two processes R_1 and R_2 which take the system from state 1 (P_1, V_1) to state 2 (P_2, V_2) are shown in the figure. If we want to determine the work done on the system in each of these processes we need to calculate the integrals over the curves of PdV, so that the work for the two processes are:

$$W_{R-1} = \int_{R_1} PdV \ , \ W_{R-2} = \int_{R_2} PdV \ , \tag{8.3}$$

and these two work integrals are generally not equal. (For the given figure they are clearly different, since the integrals correspond to the areas under the $P(V)$ curves for the two processes). Also notice that a process does not need to be described as function $P(V)$, but is more generally described as a curve that may not have a one-to-one correspondence between P and V or other state variables.

Reversible and irreversible processes. Processes may be **reversible** or **irreversible**. A reverisible processes is just that: A process that can

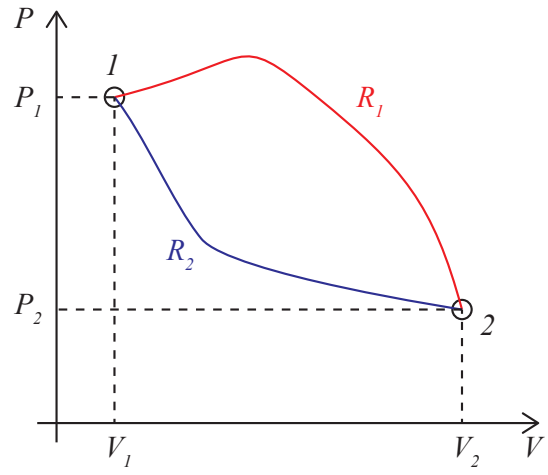


Fig. 8.1 Sketch of two processes R_1 and R_2 in a P-V digram. Each point in the P, V space corresponds to a state of the system.

be reversed. This must, for example, mean that the entropy change in the process must be zero. For a reversible process there cannot be any friction forces, since such forces would ensure that we do not return to the same state if we reverse a process.

Quasi-static processes. Often, we want to specify processes where the system is always very — infinitesimally — close to equilibrium. We call such processes **quasi-static** processes. All reversible processes are necessarily quasi-static. We must always be very close to equilibrium for a process to be reversible. However, there can be quasi-static processes that are not reversible, for example, processes where the internal friction in the system is non-negligible.

8.2 Engines

It is now time to address some of the consequences of thermodynamics. Thermodynamics was developed before we had a good microscopic understanding. Instead, it was developed partly out of necessity - because we needed the tools to develop and understand **engines**.

What is an engine? A **heat engine** is an engine that absorbs energy and converts (some of) it into mechanical work. Only part of the thermal energy can be converted into mechanical energy.

In general, heat and work are to different ways to transfer energy – and to change the internal state of a system: $\Delta E = W + Q$. Heat is the

thermal transfer of energy between systems in thermal contact. Work is the mechanical transfer of energy to a system by changes in the external parameters (volume, pressure, magnetic field, gravitational potential).

Indeed, we can argue that the conversion of heat into energy is one of the most important processes in our society. It was the basis of the industrial revolution where in particular the steam engine was important. Still, steam engines and steam turbines are important and used to convert heat from burning gas or coal into electric energy – in todays power plant. Indeed, the use of hot water from solar heating or geological systems is one possible way to produce renewable energy from our environment.

There is a fundamental difference between heat and work related to the transfer of entropy. For reversible, quasi-static processes we know that $TdS = dE + pdV + \mu dN$ and $dE = Q + W$, therefore when $W = 0$ we know that $Q = TdS$. (Notice that the entropy does not change in a quasistatic, adiabatic process).

How can we use this relation to understand engines? The idea of an engine is to convert thermal energy transfered into a system into mechanical energy (work) done by the system, and then returns to its initial state. (Otherwise we could not continue to do this process).

Now, if we transfer thermal energy $Q = TdS$ into the system, the source of the thermal energy must decrease its entropy. However, if all the thermal energy was converted into work - which may be done adiabatically - we do not have any other change in entropy - and we would have a net decrease in entropy. This is not possible - from the second law of thermodynamics. We therefore also need to transfer entropy in addition to doing work - and this is done through the transfer of some of the energy that was not converted into work.

8.2.1 A simple piston

Let us now see how we can create a *heat engine* – a machine that converts heat to mechanical work. The simplest machine we can imagine consists of a single piston filled with a gas, as illustrated in fig. 8.2. We are going to heat up the gas in the piston. What happens to the system? How does the system behave?

What do we expect to happen if all we know is that we add a thermal energy Q – heat – to the system. We expect the temperature in the gas to increase and the pressure in the gas to increase. As the pressure increases, we will expect the pressure to push on and move the piston,

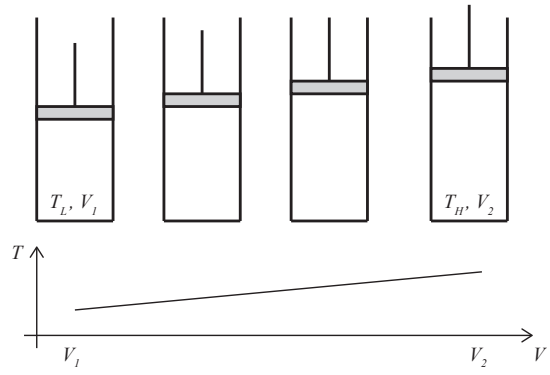


Fig. 8.2 Illustration of a piston which is supplied a heat Q so that the piston increases its temperature and expands, doing work on the environment in the process.

so that the system expands – which subsequently will lead to a cooling of the gas and a lowering of the pressure.

This process is a bit too unspecified. Let us be more precise in how we describe the process. Let us assume that the gas starts at some temperature T_L (the Low temperature) and is then suddenly put in contact with a reservoir with temperature T_H (the High temperature). The temperature in the gas will then gradually increase, until it reaches the temperature T_H, and the piston will move, changing the volume of the system from some initial volume V_1 to a final volume V_2.

We can be even more specific. Let us divide this process into two parts. First, we place the gas in thermal contact with a reservoir with temperature T_H so that the gas increases its temperature to T_H, but we keep the volume constant. Then we allow the volume to change, while the gas still is in thermal contact with the reservoir, so that the temperature is constant in the subsequent process, where the volume increases from V_1 to V_2. Now, the piston is expanded, and the temperature inside the piston is high. The machine has performed work with the motion of the piston, but only once. If we want the machine to continue to produce work, we need to be able to repeat the process – we need the machine to return to its initial state, so that we can add heat yet again and produce more work. We need to have a **cyclic process**, corresponding to a loop in the P, V diagram for the process. We therefore need to compress the piston back to a volume V_1 and reduce the temperature to T_L. However, if we first return the volume to V_1 while the system is at the high temperature, the work done will be exact the same as was done to expand the gas – and the net work done will be zero. However, if we instead first cool the gas to T_L, and then push the piston back at the lower temperature and

pressure, we will do less work on the piston to return it to its original position and the machine will be able to perform a net work.

Let us look at the various steps in this process in detail. The system starts in state 1, where the temperature is T_L and the volume is V_1. This state is illustrated in fig. 8.3. The system then comes in contact with a reservoir with temperature T_H so that the temperature increases from T_L to T_H without any change in volume in the system. This process takes us from a point 1 at (P_1, V_1) in the diagram to a point 2 at (P_2, V_1). We call such a process an **isochor** process, since it is a process without any change in volume.

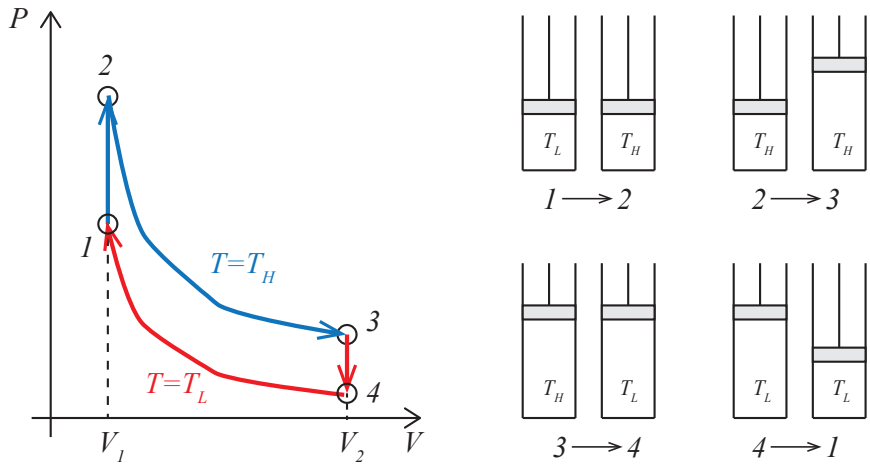


Fig. 8.3 Illustration of an engine as a sequence of processes in a P-V-diagram.

$1 \to 2$**: an isochor heating.** The process is illustrated in fig. 8.3. The work done by the system on the environment is zero, since the volume does not change. This means that the change in energy is just due to the thermal energy transferred. From the first law, we have $\Delta E_{1,2} = -W_{1,2} + Q_{1,2} = Q_{1,2}$, when $W_{1,2} = 0$. (Notice that we have changed the sign of work when we discuss engines – we prefer to work with positive work when the machine does work on the environment). We can therefore find the thermal energy transferred (the heat) from the change in energy. For an ideal monoatomic gas, the energy is $E = (3/2)NkT$, and the change in energy is therefore $\Delta E_{1,2} = (3/2)Nk(T_H - T_L)$, which is also the heat transferred to the system in this process.

$2 \to 3$**: an isothermal expansion.** In the next step of the process, the gas expands while keeping the temperature constant. We call processes

at constant temperature **isothermal processes**. How does this process look in the P, V diagram? We can find how P depends on V from the equation of state for the ideal gas: $PV = NkT$, and since this process occurs at $T = T_H$, we have $P = (NkT_H)/V$. We can therefore find the work done by the system on the environment in this step:

$$W_{2,3} = \int_{V_1}^{V_2} PdV = \int_{V_1}^{V_2} \frac{NkT_H}{V} dV = NkT_h \ln \frac{V_2}{V_1} \,. \tag{8.4}$$

If we perform work on the enviroment, but keep the temperature in the gas constant, thermal energy must be transferred from the reservoir. This is also evident from the first law, because $\Delta E_{2,3} = -W_{2,3} + Q_{2,3}$. For an ideal (monoatomic) gas, we know that $E = (3/2)NkT$. For a process with constant T there is therefore no change in energy, and $\Delta E_{2,3} = 0$, hence $Q_{2,3} = W_{2,3}$: all the energy used to perform the work is taken from the reservoir.

$3 \to 4$**: an isochor cooling.** Now, we cool the system to the initial temperature T_L by putting the system in contact with another reservoir at a lower temperature T_L. The discussion for this process is similar to the isochor heating: The heat $Q_{3,4}$ corresponds to the change in energy because the work is zero (since there is no change in volume). We therefore have $Q_{3,4} = \Delta E_{3,4} = (3/2)Nk(T_L - T_H)$. Now, this is a negative value, indicating that this heat is tranferred out of the system and not into the system.

$4 \to 1$**: an isothermal compression.** Finally, the system is brought back to the beginning by compressing the system back to the starting volume V_1. This is again an isothermal process at $T = T_L$ and we can find the work by integrating $P(V) = NkT_L/V$:

$$W_{4,1} = \int_{V_2}^{V_1} \frac{NkT_L}{V} dV = NkT_L \ln \frac{V_1}{V_2} \,, \tag{8.5}$$

and this work corresponds to the heat $Q_{4,1} = W_{4,1}$. Both the work and the heat are negative, indicating that heat is tranferred out of the system and work is done on the system in this process.

8.2.2 Net work and efficiency of the machine

We have now described a complete cycle of the machine. The cycle is illustrated in fig. 8.3.

We see that the work performed from 2 to 3 is the integral of $P(V)$ with respect to V, and we can interpret this as the area under the curve of the $P(V)$ curve from 2 to 3. Similarly, the work performed from 4 to 1 is (minus) the area under the $P(V)$ curve from 4 to 1. The net work done in the whole cycle is therefore the area enclosed by the $P(V)$ curve for the entire, cyclic process.

The **net work** is the sum of all the works done in each subprocess. This corresponds to the area enclosed by the curve in the P-V diagram. You can see this by realizing that the system does positive work when the system moves to the right along the $P(V)$ curve. The area under this part of the curve is therefore the positive work done. The system does negative work when the system moves to the left along the $P(V)$ curve. If we subtract these two areas – as illustrated in fig. 8.4 – we get the net work, which corresponds to the enclosed area.

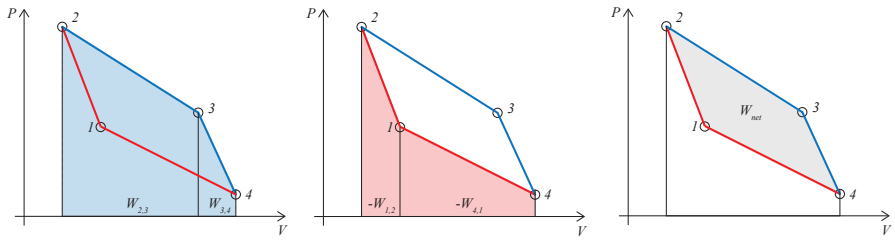


Fig. 8.4 Illustration of work done by the system on the environment for a complete cycle of a machine.

The **net heat** is the sum of all the heats transferred in each subprocess. The **added heat** corresponds to the sum of the positive heats ($Q > 0$). We can apply the first law to a complete cycle. In this case, the change in energy is zero, since the energy at the beginning and end of the cycle must be the same since the energy is a state variable. Hence, the net work must correspond to the net het added to the system: $W_{net} = Q_{net}$ and $\Delta E_{net} = 0$.

What is the efficiency, e, of this machine? It is reasonable to call the efficiency how much work you get out for each unit of heat you put into the system:

$$e = \frac{W_{net}}{Q_{add}} \,. \tag{8.6}$$

We can calculate the efficiency of the machine we have described, by inserting the values we have found for work and heat. We notice that the added heat corresponds to the heat added in process $1 \to 2$ and $2 \to 3$,

that is $Q_{net} = Q_{1,2} + Q_{2,3}$. The efficiency is therefore

$$e = \frac{W_{net}}{Q_{net}} = \frac{NkT_H \ln V_1/V_0 - NkT_L \ln V_1/V_0}{NkT_h \ln V_1/V_0 + Nk(T_H - T_L)} \,. \tag{8.7}$$

This is a nice result. And as we shall soon see – the machine we have described here corresponds to a Stirling machine, which is used for example for solar array systems.

8.2.3 Limits on efficiency

We have found a particular efficiency in this case. But how good is this machine? To answer that we need to know of efficient a machine can be – is there a theoretical maximum in efficiency for a heat engine? The answer is yes – and we can find this maximum.

Let us look at this more systematically. We can sketch an engine as in figure 8.5.

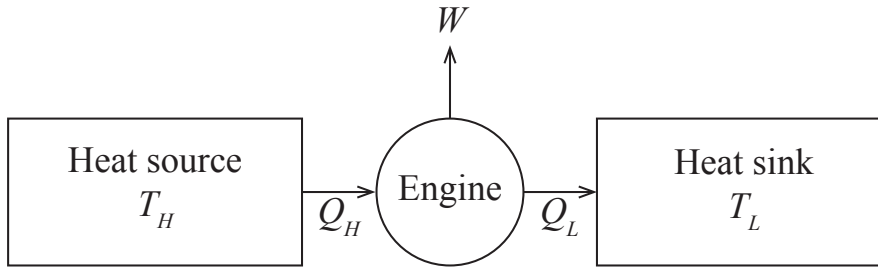


Fig. 8.5 Sketch of a heat engine.

This engine works between a heat source at T_H and a heat sink at T_L.

- The engine performs a cycle that should end in the engine coming back to the state is started from (not microstate, of cource, but the same macrostate).
- In this cycle a thermal energy Q_H is transferred into the engine.
- The engine does work W.
- And the engine transfers a thermal energy Q_L out of the engine and into the heat sink.

- The heat source and heat sink are so large that their temperatures do not change in this process.

We have so far defined the work done *on the system* as positive, but for an engine we are usually more interested in the work done *by the system*. We will therefore use the common convention Q_H, Q_L and W all are positive. In this notation, the first law of thermodynamics is

$$\Delta E = Q_H - Q_L - W \ , \tag{8.8}$$

where W is the net work done in a complete cycle, Q_H are all the heats that are added to the system – this may occur in more than one of the processs – and Q_L are all the heats that are removed from the system – again this may occur in more than one processes. Notice that we have changed the sign in front of the work, because the work W is now the work done by the system on the environment. For a complete cycle, the change in E must be zero, because E is a state variable and therefore will be the same in the first and the last point in the cycle, because these are the same states. The first law of thermodynamics for this system is therefore

$$W = Q_H - Q_L \ , \tag{8.9}$$

the net work W is equal to the net thermal energy (heat) added during a complete cycle.

What is the efficiency of such as machine: It is what we get out (net work) divided by what we put in (heat in Q_H):

$$e = \frac{W}{Q_H} \ . \tag{8.10}$$

Ok. But what is the maximum efficiency we can get? It is limited by the laws of thermodynamics. From the first law of thermodynamics, we see that $Q_H = W + Q_L$. The efficiency is therefore:

$$e = \frac{Q_H - Q_L}{Q_H} = 1 - \frac{Q_L}{Q_H} \ . \tag{8.11}$$

We can use the second law to relate Q_L and Q_H. From the second law we know that

$$dS = dS_H + dS_L = -\frac{Q_H}{T_H} + \frac{Q_L}{T_L} \geq 0 \ , \tag{8.12}$$

which gives:

$$\frac{Q_L}{Q_H} \geq \frac{T_L}{T_H} \,. \tag{8.13}$$

and therefore

$$e = 1 - \frac{Q_L}{Q_H} \leq 1 - \frac{T_L}{T_H} \,. \tag{8.14}$$

(Notice that we have here assumed that the heat is deposited to the low temperature and taken from the high temperature). When the cycle is reversible, we know that the equality holds, and in this case the efficiency is:

$$e = 1 - \frac{T_L}{T_H} \,. \tag{8.15}$$

This is the maximum efficiency of an engine. A reversible engine of this type is called a *Carnot engine*. Notice that this means that all engines that operate between the same two temperatures (and that are reversible) have the same efficiency – independent of how they are constructed.

However, in practice, no real engine reaches the theoretical efficiency, and differences in the actual implementation is therefore important - typical engines reach about 30-40% of the theoretical limit, often even less.

Clicker example of cycles

8.3 The Carnot cycle

Let us show an example of a cycle – a heat engine – that achieves the theoretical limit in an idealized setting. The carnot cycle consists of four steps. It starts from a state 1 with a temperature T_H. (It is in contact with a heat bath with this temperature).

Step 1 ($1 \to 2$) – Isothermal expansion: First, it executes an isothermal expansion from $1 \to 2$. The system expands while in contact with the heat source. This means that heat is transfered into the system. For a gas this corresponds to the gas pushing on a piston turning a crank to produce work. Since the system stays at the same temperature, the engine must get thermal energy from the heat bath. (We assume that heat is flowing into the system, even though the system and the bath are at the same temperatures - for example by expanding the system in very small steps, so that we are always infinitesimally close to equilibrium).

We could then compress the system isothermally to return to state 1. This would complete the cycle, but exactly the same same work would be needed to push the piston back - we would not have done any net work! We therefore need a few more steps - we could reduce the pressure in the system, so that we could return the piston at a lower pressure, which would require less work. But then we would need to increase the pressure again before returning to the initial state. We could also lower the pressure by doing an adiabatic expansion, and this is what we do:

Step 2 ($2 \to 3$) – adiabatic expansion. We now isolate the system (remove the thermal contact). The volume contines to increase, but no heat is added. Then both the pressure and the temperature decreases. The temperature decreases all the way to T_L.

We also perform work in this part of the cycle - but now the energy that goes into the work comes from the internal energy of the system, $\Delta E_{2,3} = W_{2,3}$, since the system is thermally isolated.

Now, we can restore the system to its initial condition, starting at a lower pressure. First, we perform an isothermal compression at T_L.

Step 3 ($3 \to 4$) – Isothermal compression. Work is done on the system by pressing the piston back in. This would increase the temperature of the system, but instead thermal energy is flowing out of the system and into the heat sink (T_L).

Finally, we need to increase the temperature of the system back to its original state.

Step 4 ($4 \to 1$) – Adiabatic compression. We now need to increase the temperature (and pressure) of the system back to its original state. We remove the contact with the heat sink, and compress the system without any thermal energy flowing into it - this is done until the temperature is back at T_H (and the pressure is back to its initial state).

8.3.1 Carnot cycle for ideal gas

Let us look at the Carnot cycle for an ideal gas in detail. The cyclic process is sketched in the $P(V)$ diagram in fig. 8.6.

Isothermal expansion ($1 \to 2$). The gas starts at $T_1 = T_H$ and V_1. The gas expands isothermally to V_2. First, how does this curve look like in the diagram?

If the temperature is constant, we know that $pV = NkT$, and hence we know that

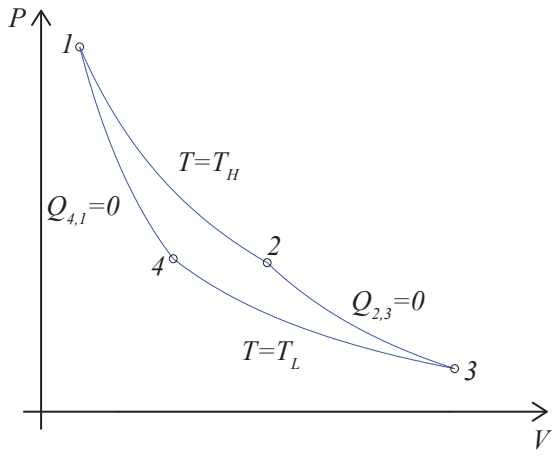


Fig. 8.6 Illustration of the Carnot cycle for an ideal gas.

$$p = \frac{NkT}{V} \tag{8.16}$$

for this curve.

Since the temperature is constant, and the energy U for an ideal gas is linear in temperature, we know that $dU = 0$ for this process, hence $Q = W$. The heat transferred is therefore given by the work:

$$W = \int_{V_1}^{V^2} pdV = NkT \int_{V_1}^{V_2} \frac{1}{V} dV = NkT \ln \frac{V_2}{V_1} \ , \tag{8.17}$$

(Notice, that you would have found the same by using that the heat is $Q = TdS$ for this process and that the entropy is $S = Nk \ln V + f(N, U)$, which would give the same result for the heat).

We can draw the curve from 1 to 2 in a pV diagram, and notice that the work is the area under the curve from 1 to 2.

Adiabatic expansion ($2 \to 3$). Now, the gas is disconnected from the bath at T_H and thermally isolated - so that the work is given as the change in internal energy of the system. The work done in this process is simple to find from $dU = Q + W$, where now $Q = 0$, and therefore $dU = W$, where $U = (3/2)NkT$ for a monatomic ideal gas and $((f + 3)/2)NkT$ for a general ideal gas. The work is therefore simply

$$W_{23} = \frac{3}{2} Nk \left(T_H - T_L\right) \ . \tag{8.18}$$

What does the curve $P(V)$ look like for this process?

For a quasistatic, adiabatic process the entropy is constant. The entropy for an ideal gas is given by the Sackur-Tetrode equation:

$$S = Nk\left(\ln\frac{V}{N} + \frac{3}{2}\ln\frac{4\pi mU}{3Nh^2} + \frac{5}{2}\right) , \tag{8.19}$$

where

$$U = \frac{3}{2}NkT , \tag{8.20}$$

and therefore

$$S = Nk\left(\ln\frac{V}{N} + \frac{3}{2}\ln\frac{2\pi mkT}{h^2} + \frac{5}{2}\right) . \tag{8.21}$$

Constant S therefore means that $\ln VT^{3/2}$ is constant, and therefore that $VT^{3/2}$ is constant. We can use this to find the volume at the end of the expansion:

$$T_L V_3^{2/3} = T_H V_2^{2/3} , \tag{8.22}$$

that is

$$V_3/V_2 = (T_H/T_L)^{3/2} . \tag{8.23}$$

And we can also find the shape of the $P(V)$ curve - which we often simply call an adiabat, by inserting $T = pV/Nk$ into $VT^{3/2} = \text{const.}$, which gives $V^{5/2}p^{3/2}$ is constant, or $PV^{5/3}$ is constant, and the curve is $P = C/V^{5/3}$, which we can sketch.

Isothermal compression ($3 \to 4$). The gas is now in contact with the bath at T_L and compressed isothermally to the volume V_4. How should we choose V_4? We must choose V_4 so that we can go along an adiabat from 4 to 1, and therefore V_1, T_H, and V_4, T_L must be connected by an adiabat, therefore

$$V_4 T_4^{3/2} = V_1 T_H^{3/2} \Rightarrow V_4/V_1 = (T_H/T_L)^{3/2} , \tag{8.24}$$

where we again have that $(T_H/T_L)^{3/2} = V_3/V_2 = V_4/V_1$.

What work is done on the gas? We find the work by integrating

$$W_{34} = \int_3^4 pdV = NkT_L \ln\frac{V_3}{V_4} = NkT_L \ln\frac{V_2}{V_1} . \tag{8.25}$$

A corresponding energy is transferred into the system as heat, since the internal energy does not change for an ideal gas during an isothermal process. The heat is therefore

$$Q_{34} = Q_L = W_{34} \ . \tag{8.26}$$

Adiabatic compression ($4 \to 1$). Finally, the gas is isolated thermally again, and the gas is compressed along an adiabat from 4 to 1. The work done in this process is:

$$W = \frac{3}{2} Nk \left(T_H - T_l\right) \ , \tag{8.27}$$

which cancels the work done during the isentropic expansion from 2 to 3.

The net work done by the whole cycle can be read from the pV-diagram. Look at the area below the curve from 1 to 2, 2 to 3, 3 to 4 and 4 to 1, and see that the net work corresponds to the area enclosed by the curve. This net work is

$$W_{\text{net}} = Q_H - Q_L = Nk \left(T_H - T_L\right) \ln \frac{V_2}{V_1} \ . \tag{8.28}$$

Efficiency: We can put this back into the efficiency, and find that the efficiency is

$$e = \frac{W_{\text{net}}}{Q_H} = \frac{T_H - T_L}{T_H} = 1 - \frac{T_H}{T_L} \ . \tag{8.29}$$

8.4 Carnot refrigerator

A refrigerator is a heat engine that works in reverse - it uses work to cool a system down. It takes heat from the low temperature region – inside the refrigerator – and combines it with work, and dumps heat at a high temperature bath - outside the refrigerator.

In this case the efficiency depends on Q_L and W, but we call it the coefficient of performance (COP):

$$\text{COP} = \frac{Q_L}{W} \ . \tag{8.30}$$

Again, we can use the first law to relate the three energies:

$$Q_H = Q_L + W \ , \tag{8.31}$$

which gives

$$\text{COP} = \frac{Q_L}{Q_H - Q_L} = \frac{1}{Q_H/Q_L - 1} \ . \tag{8.32}$$

What are the limits for the efficiency of a refridgerator? The second law gives us:

$$\frac{Q_H}{T_H} \geq \frac{Q_L}{T_L} \Rightarrow \frac{Q_H}{Q_L} \geq \frac{T_H}{T_L} \ . \tag{8.33}$$

(Notice that the inequality is in reverse compare to before because the entropy is flowing in the opposite direction). This gives:

$$\text{COP} \leq \frac{1}{T_H/T_L - 1} = \frac{T_L}{T_H - T_L} \ . \tag{8.34}$$

The coefficient may be higher than 1 - indeed it will usually be. For example, for a kitchen fridge we may have $T_H = 298\text{K}$ and $T_L = 255\text{K}$, which gives COP = 5.9. Notice that COP goes to infinity as the temperature difference goes to zero! Also notice that we can make a Carnot cycle refrigerator - but this is not how things are done in practice since the cycle would take too long time. We will address a few more realistic processes in the following.

8.5 Realistic model for Stirling Engine

We can gain more insight into how real engines work by developing a realistic model for a Stirling engine. Here, we will work on a model engine used in teh 2.670 class at MIT, inspired by a report by M. Byl in 2002. The Stirling engine we will model is sketched in fig. 8.7. This engine consist of two connected chambers. One chamber on the left, where the displacement piston is working. The volume of the gas in this chamber remains constant throughout the process, but the portion that is in contact with the hot temperature, T_H, and the portion that is in contact with the cold region, T_L, varies in time. We will assume that the whole gas has the same temperature $T(t)$. This means that the temperature will vary in time, but we assume that the temperature in the gas is the same everywhere along the machine.

Let us start with a quantitative description of this machine. We will assume that there is a constant transport of heat from the surrounding reservoirs into the machine. There is a transfer of heat into the machine in the hot region and a transfer of heat out of the machine in the cold region. We will apply a simple law for heat conduction to estimate the *rate* at which thermal energy is transferred to the gas. The heat transfer from outside to inside the piston is proportional to the difference in

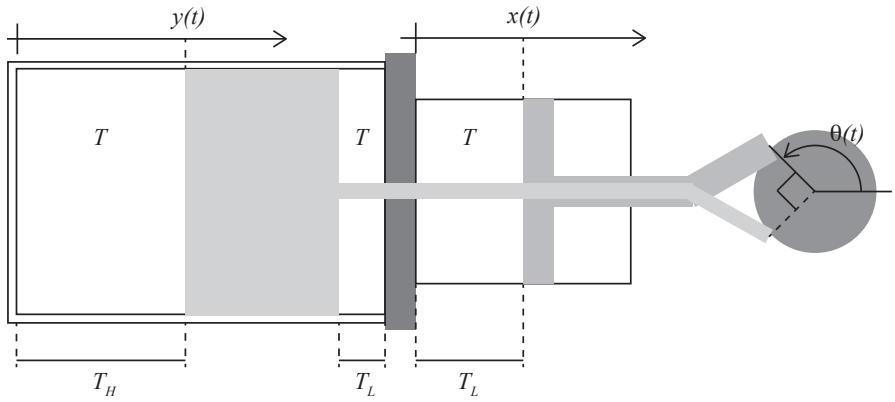


Fig. 8.7 Illustration of a realistic Stirling engine using air as the gas medium.

temperature, $T_{out} - T_{in}$, the area of contact, A, and a constant, μ, that depends on the material properties of the cylinder wall. In the hot region, the only contact is through the large cylinder on the left. The heat transported through the wall and into the gas in a time interval Δt is then

$$Q_H = A_H \mu \, (T_H - T) \, \Delta t \; , \tag{8.35}$$

where $\mu = 100000\mathrm{W/m}^2$ (calculated from the thermal conductance of steel and the thickness of the cylinder wall), and A_H depends on the position of the displacer piston.

How do we describe the positions of the displacer piston and the work piston? We notice that the two pistons are attached to a flywheel – through some mechanism that is not properly sketched in the figure – with the net effect that the pistons move a distance $2R$ back and forth along their axis, where R is the radius of the flywheel at the point of attachment. The position $x(t)$ of the work piston can then be written

$$x(t) = R \, (1 + \sin \theta(t)) \; , \tag{8.36}$$

with a proper choice of origin for θ. The position, $y(t)$, of the displacer cylinder should move with a phase difference of $\pi/2$, so that

$$y(t) = y_{max} \, (1 + \cos \theta(t)) \; , \tag{8.37}$$

where y_{max} is the maximum extension of the piston. If we know $x(t)$ and $y(t)$, we can find the surface areas A_H and A_L for the thermal conduction. At the hot end, the area A_H only depends on the maximum possible area, corresponding to the y_{max} multiplied with the circumference of the

cylinder. We therefore find

$$A_H(t) = A_{H,max}\,(1+\cos\theta(t))\ . \tag{8.38}$$

where $A_{H,max} = A_0$ is the total surface area of the part of the large cylinder that contains gas.

For the cold end we need to include the heat transfer from the large cylinder and the transfer for the small cylinder:

$$A_L(t) = A_0\,(1-\cos\theta(t)) + A_1(x)\ , \tag{8.39}$$

where $A_1(x) = A_{1,max}x(t)/x_{max}$ is the area of contact for the small cylinder to the right in the figure.

This means that in a short time interval Δt, the change in (internal) energy E of the gas in the system is

$$\Delta E = Q + W\ , \tag{8.40}$$

where $Q = Q_H - Q_L$, with $Q_H = A_H\mu(T_H - T)$ and $Q_L = A_L\mu(T - T_L)$. In addition, we need to include the work, W, done on the system. This work is done by the work piston, which has moved a distance $\Delta x = (\Delta x/\Delta t)\Delta t$ with a force $F = (P - P_0)A_p$ acting on the piston from the gas, where P is the pressure in the gas, P_0 is the outside pressure, and A_p is the area of the piston in contact with the gas. The work is therefore

$$W = -F\Delta x = -(P - P_0)A_p\frac{\Delta x}{\Delta t}\Delta t\ . \tag{8.41}$$

If we describe the position $x(t)$ through the angle $\theta(t)$, we see that

$$\frac{\Delta x}{\Delta t} = \frac{dR(1+\sin\theta)}{dt} = R\frac{d\theta}{dt}\cos\theta = R\omega\cos\theta\ . \tag{8.42}$$

Now, we have a complete description of ΔE:

$$\Delta E = Q_H - Q_L + W = A_H\mu(T_H - T)\Delta t - A_L(T - T_L)\Delta t - PA_pR\omega\cos\theta\Delta t\ , \tag{8.43}$$

which gives a differential equation for $E(t)$:

$$\frac{dE}{dt} = A_H(t)\mu(T_H - T(t)) - A_L(t)(T(t) - T_L) - (P(t) - P_0)A_pR\omega(t)\cos\theta(t)\ . \tag{8.44}$$

In addition, we need to determine the motion of the flywheel. The flywheel has a moment of inertial I and is affected by a torque, τ, due

to the force on the work piston. In addition, we will include a frictional effect due to friction in the flywheel as well as viscous damping due to gas motion in the displacement cylinder. The total torque on the flywheel is then

$$\tau = FR\cos\theta - b\omega = (P - P_0)A_pR\cos\theta - b\omega \ , \tag{8.45}$$

where b is a viscous coefficient describing the effective damping in the system. This gives us the equation of motion for θ:

$$\frac{d^2\theta}{dt^2} = FR\cos\theta - b\frac{d\theta}{dt} \ . \tag{8.46}$$

We now have two equations, describing θ and E. In addition, we need to be able to calculate T, V and P in the system. We can find V from geometric considerations. The volume in the displacement cylinder is constant, V_c, but the volume in the work cylinder depends on x:

$$V(t) = V_c + A_px \ , \tag{8.47}$$

where A_px is the volume of the gas in the work piston. We can now find the temperature through

$$E = C_VT \ \Rightarrow \ T = \frac{E}{C_V} \ , \tag{8.48}$$

where $C_V = c_Vm$ is the heat capacity for the gas in the cylinder. And we find the pressure using the ideal gas law:

$$P = \frac{NkT}{V} = \frac{2}{5}\frac{E}{V} \ , \tag{8.49}$$

where we have used that $E = (5/2)NkT$ for the gas at realistic temperatures.

This completes the set of equations needed to determine the behavior of the machine. We can now model the machine by integrating the equations of motion for θ and the time development of E to find the behavior of the machine in a $P(V)$ diagram.

First, we need to know the values of the parameters for a realistic machine. We will use $T_H = 600\text{K}$, $T_L = 300\text{K}$, $R = 1.25\text{cm}$ (radius of flywheel), $A_0 = 40\text{cm}^2$ (surface area of active area of displacement cylinder), $A_p = 1.9\text{cm}^2$ (surface area of piston), $A_{1,max}/x_{max} = 4.9\text{cm}$ (perimeter of the work piston cylinder), $V_c = 40\text{cm}^3$ (volume of gas in displacement cylinder), $I = 4\text{kg}\cdot\text{cm}^2$, $b = 0.7{\cdot}10^{-3}\text{N/s}$ (damping constant representing viscous processes), $P_0 = 10^5\text{Pa}$ (trykk i omgivelsene), $C_V =$

717J/kg/K (heat capacity for the gas), $\mu = 100000\text{W/m}^2$ (heat transfer coefficient for the cylinder wall), and $\rho =$ (mass density for the gas).

The motion of the cylinder is solved by the following implementation, where we have used Euler-Cromer for the time integration with a time step of $\Delta t =$.

The resulting behavior is illustrated in fig. 8.8. The simulation is for 5 seconds, but only the first second is shown for $x(t)$ and $T(t)$, because the behavior is completely regular after 1 second. We see that both the position and the temperature has a short transient behavior before they start having a periodic behavior. We can illustrate the behavior in the periodic part of the motion by plotting $P(T)$ and $P(V)$ for the last 2.5 seconds of the simulation. We see that the behavior of the machine indeed is a closed loop in the $P - V$ diagram. However, the actual path in the $P - V$ diagram looks very different from the idealized case we used to model the behavior of the system.

We can estimate the efficiency of this – more realistic – Stirling engine: $e = W_{net}/Q_{in}$, by measuring the heat flow into the engine during a cycle and the net work done by the machine during a cycle. The net work corresponds to the area enclosed by the curve in the $P - V$ diagram, but in the model it is easier to calculate the net work done by the piston and the sum of the heat, Q_H, transferred into the system. Unfortunately, this results in a very low efficiency. (Something may be wrong with the calculation – can you find it?).

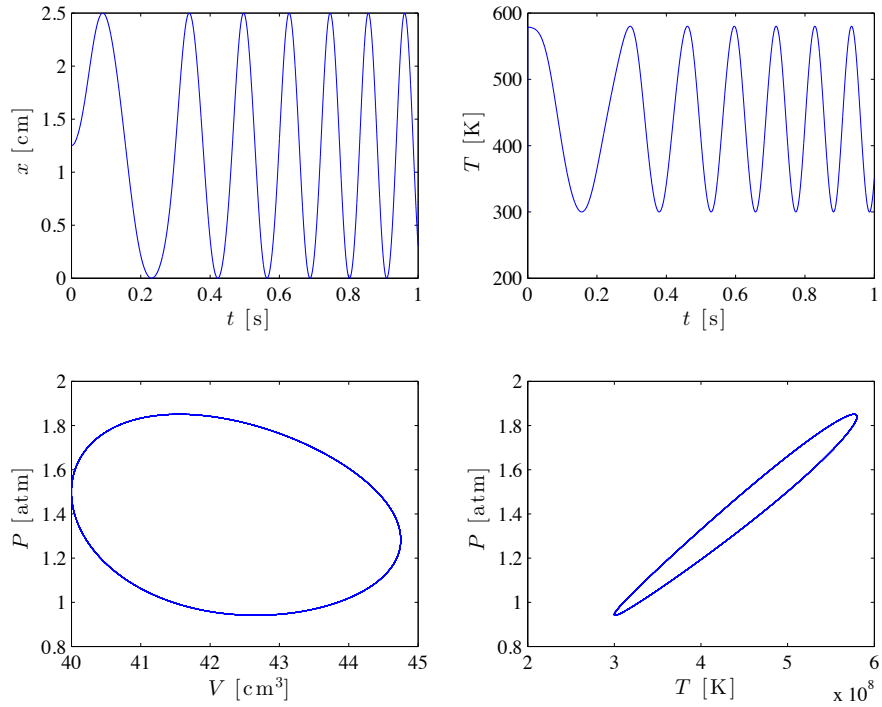


Fig. 8.8 Plots of the temporal behavior of the model Stirling engine.

8.6 Additional aspects of heat engines and thermodynamics (not releavnt for exam)

8.6.1 Path dependence of Heat and Work

We need to be careful with how we write differences and differentials.

We have been careful not to write dW and dQ, since W and Q are not state functions.

S, E, etc are state functions and therefore form pure differentials.

If we go a complete loop, we will come back to the same value.

The integral over a closed loop is therefore zero.

But not so for W and Q. The integral around a complete loop of W is the net work done. And the integral around a complete loop of Q is the same - since they are related by $dE = W + Q$, and the integral of dE around a compete loop is zero.

8.6.2 Heat and work at constant temperature and constant pressure

Isothermal work. In a reversible, isothermal process T is constant and $TdS = d(TS)$, therefore $Q = TdS = d(TS)$ and

$$W = dE - Q = dE - d(TS) = dF \tag{8.50}$$

In this case Helmholtz free energy is a proper energy.

Isobaric work. Many processes occur at constant pressure - such as open processes in the atmosphere.

If we boil a liquid at constant pressure, the piston above a boiling system may move a distance dV, and the work due to this movement would be $-pdV = -d(pV)$. This work would be included in the total work done by the system. But if we compress the system, we get this work for "free" from the atmosphere, and if we expand the system we have to perform this work against the atmosphere, but this work cannot be used for anything else. It is therefore often useful to subtract $-d(pV)$ from the work, and only get the effective work performed on the system. We define this effective work as

$$W' = W + d(pV) = dE + d(pV) - Q = dH - Q \ , \tag{8.51}$$

where we have introduced the entalpy $H = U + pV$. This plays the same role at constant P as U plays in processes at constant V.

The term pV is required to displace the surrounding atmosphere to get room for the system.

(We assume that there are other types of work in addition to this pressure work).

Two classes of constant pressure processes:

Processes with no effective work $Q = dH$ (no other work than the work needed to expand the system against the given pressure). Evaporation of a liquid - the heat of evaporation is the enthalpy difference between the liquid and gas phase.

Processes at constant T and constant p. Then $Q = TdS = d(TS)$ and the effective work is

$$W' = dF + d(pV) = dG \ , \tag{8.52}$$

where we have introduced $G = F + pV = E + pV - TS$, the Gibbs free energy. The effective work performed at constant pressure and temperature is equal to the change in Gibbs free energy. Useful for

chemical reactions where the volume changes at constant pressure and temperature.

8.6.3 Chemical Work

(Kittel, p 250)

8.6.4 Joule-Thompson effect

(May use for exercise instead - process on van der Waals gas).

8.7 Summary

8.8 Exercises

Exercise 8.1: Adiabatic process

The change of state of a system is called *adiabatic* if no heat is exchanged with the environment along the reversible path of the process from state 1 to state 2. Let us address the behavior of 1 mol of an ideal gas. The molar heat capacity at constant volume is $c_{V,m}$. Assume that the gas is going through an adiabatic change ΔT and ΔV.

a) Show that for the adiabatic change

$$c_{V,m}\frac{\Delta T}{T} + R\frac{\Delta V}{V} = 0 \,, \tag{8.53}$$

where $R = N_A k$ is the molar gass-constant, where N_A is Avogardro's number.

b) Integrate (8.53) of demonstrate that the following expression is valid for any pair of states 1 and 2 connected by an adiabatic path:

$$T_1 V_1^{\alpha} = T_2 V_2^{\alpha} \,, \text{ where } \alpha = R/c_{V,m} \,. \tag{8.54}$$

c) Show that the result in (8.54) alternatively may be expressed as

$$p_1 V_1^{\gamma} = p_2 V_2^{\gamma} \text{ or } T_1 p_1^{\beta} = T_2 p_2^{\beta} \,. \tag{8.55}$$

where $\beta = (1-\gamma)/\gamma$ and $\gamma = c_{p,m}/c_{V,m} = 1 + R/c_{V,m}$.

d) Draw adiabatic curves ($pV^\gamma = \text{const.}$) and isotherms in a p-V-diagram. Assume that the gas is diatomic with $c_{V,m} = (5/2)R$.

Exercise 8.2: Adiabatic compressibility

The compressibility of a system is defined as

$$\kappa_x = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_x , \tag{8.56}$$

for a process where x is a constant. For an ideal gas we know that $pV^\gamma = \text{const.}$ for an adiabatic process, where $\gamma = c_{p,m}/c_{V,m}$.

a) We often use the term isentropic to describe an adiabatic process. Why?

b) Show that for an ideal gas the isentropic compressibility, κ_S, is smaller than the isothermal compressibility, κ_T, by a factor γ.

The speed of sound in a gas is given as $v = (\kappa_S \rho)^{-1/2}$, where ρ is the mass density of the gas.

c) Using the results from above, show that the speed is sound can be written as $v = \sqrt{\gamma RT/M}$, where M is the average molecular weight of the gas.

d) Calculate the numerical value for the speed of sound in air at 300 K by using $M \simeq 29\,\text{g/mol}$.

e) The speed of sound is closely connected to the average velocity, v_{avg}, of molecules in a gas. Try to express v_{avg} by R, T, and M and compare with the speed of sound.

f) Discuss the difference between a sound wave that (hypothetically) propagated as an isothermal process and one that propagated as an isentropic process. How can you change the parameters of gas in order to get a more isothermal process during sound propagation?

Exercise 8.3: Crossing adiabatic curves

a) Prove that two different adiabatic curves for an ideal gas cannot cross each other in a p-V-diagram.

Hint. Assume that they cross each other, complete a cycle with an isotherm, and show that such a process is in violation of the second law.

Exercise 8.4: Work of a gas

We study a He-gas in a cylinder with a piston. Initially, the system is in equilibrium at $T = 2\,\mathrm{K}$, $p = (1/36)\,\mathrm{atm}$, $V = 600\,\mathrm{cm}^3$. You may assume that the internal energy of the gas is zero at this point.

a) The temperature is increased to $T = 288\,\mathrm{K}$ at constant volume. If you assume that He behaves as a monatomic ideal gas, how much heat is adsorbed? And what is the internal energy of the gas? May all this energy be extracted in the form of heat? May all this energy be used to perform mechanical work?

b) The gas expands isentropically until the temperature is 2 K. How much work is performed and what is the internal energy of the gas? Check that this process does not violate the various formulations of the second law of thermodynamics.

c) The gas is compressed isothermally back to its original volume. How much work is done in this cycle og how much energy in the form of heat is adsorbed? What is the efficiency of the cycle?

Exercise 8.5: String engine

We know that a rubber band can be described by $S(x) = b - k_B a x^2$, and $K = -2ak_BTx$ (Hooke's law), where x is the elongation of the band, a and b are constants, T is the temperature, K is the tension in the band (the force), and S is the entropy.

a) Draw a Carnot-cycle for the rubber-band-engine in a K-x-diagram. (The cycle consists of two isothermal and two isentropic processes).

b) Find the work done in the process $1 \to 2$ and $3 \to 4$.

c) Find the heat adsorbed in process $1 \to 2$ and $3 \to 4$.

d) Show that the efficiency of the machine is given as the Carnot efficiency.

Exercise 8.6: Three-point cycle

Fig. 8.9 shows a cycle for an imaginary machine that operates on an ideal gas. You can assume that the heat capacities C_V and C_p are constant throughout the process.

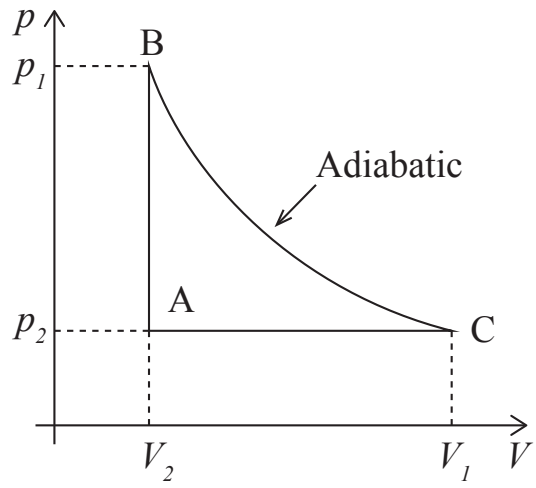


Fig. 8.9 Illustration of a cyclic process in the $p - V$ plane. The process from B to C is adiabatic.

a) Show that the efficiency of this machine is

$$e = 1 - \gamma \frac{(V_1/V_2) - 1}{(p_1/p_2) - 1} \,, \tag{8.57}$$

where $\gamma = C_p/C_V$.

8.9 Projects

Chemical potential and Gibbs Distribution

9

We have now developed a microscopic understanding and the thermodynamic theory for systems that are closed to all transport: isolated systems with constant E, V, and N, and for systems that are only open to thermal transport, systems with constant T, V, and N. We started with an isolated, closed system, and then opened the possibility for energy transport.

For the microcanonical system, we found that the link between the microscopic and the macroscopic follows from $S = k \ln \Omega(E, V, N)$ and we found that the multiplicity, and hence entropy, of such a system must increase and equilibrium is obtained when the multiplicity, and hence entropy, is at its maximum.

For the canonical system, we found that the microscopic decription is given by the Boltzmann factors and the partition function, $Z = Z(T, V, N)$, and that the macroscopic state is found through $F = -kT \ln Z$. The thermodynamic principle that detemines the behavior of this system is that the system develops towards a minimum in Helmholtz free energy, F, and that the equilibrium is characterized by a minimum in F.

There are, however, many systems of interest that are open to particle transport. For example, how can we determine the equilibrium state of an ideal gas in the atmosphere — where the particles are free to flow from one region to another; the equilibrium state of electrons in a voltage field; or the equilibrium state of a gas-water mixture? If we pour water into an empty container of a given volume and keep the temperature constant, how much of the water will be water, how much will be ice and how much

will be vapor? And what principle determines this equilibrium? Similar questions could be asked about chemical reaction, such as the reaction between hydrogen and oxygen to form water — what is the equilibrium concentrations of each component?

To answer such questions, we need to extend our microscopic and thermodynamic framework to also address systems that are open to transport. In this chapter we will introduce the basic principles needed to address the statistical mechanics and thermodynamics of *open* systems — systems where the particle number may change.

9.1 Chemical potential

We say that system is open to transport if particles can be transported in or out through the boundaries of the system. In that case, the number of particles is no longer a constant in the system, and we call such systems *diffusional open* or *open to diffusion*.

In order to address the thermodynamics of systems open to diffusion, we need thermodynamic principles similar to what we have found previously for the microcanonical and the canonical systems. We have seen that:

- If two systems have the same temperature there is no net energy flow between them.
- If two systems have the same pressure, there is no net change of volumes
- If two systems have the same X, there is no net flow of particles between them

We need to identify this X. Let us do this by expanding gradually from systems we already know how to treat.

9.1.1 Chemical potential in the canonical system

Let us start with a system S that is in thermal contact with a large reservoir with temperature T, but otherwise of constant particle number and constant volume: This system is a canonical system. We divide this system into two parts A and B that are in diffusional contact. For example, if the system is a gas, we simply divide the container into two parts and call one part A and the other part B as illustrated in Fig. 9.1.

Particles can pass from A to B and the other way. This is similar to the molecular dynamics model for the ideal gas where half the system was empty. (Similar, but not identical, since our molecular dynamics model was for constant energy and not constant temperature!). We can characterize this sysytem by N the total number of particles in the system, where $N_A + N_B = N$ is constant, but N_A and N_B may vary.

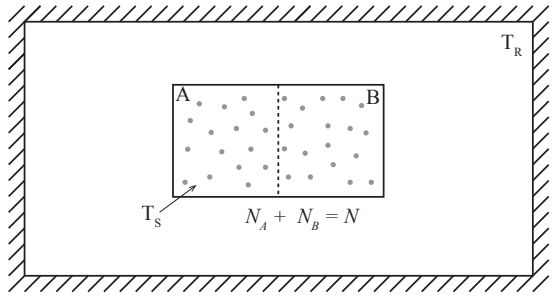


Fig. 9.1 Illustration of a system S in thermal contact with a large reservoir so that the temperatures $T_R = T_S$ are the same in the two systems. The system is divided into two parts, A and B, that are in diffusive contact so that the total number of particles is conserved, $N_A + N_B = N$.

Helmholtz free energy for combined system. This system is a canonical system, and we know that it will develop towards minium Helmholtz free energy and that the equilibrium is characterized by a minimum of the Helmholtz free energy. What is the Helmholtz free energy for this system? It is the sum of the energy for the components: $F = F_A + F_B$. What happens if we look at a small change in this system: a small transfer of dN_A particles from B to A:

$$N_A \to N_A + dN_A \tag{9.1}$$

$$N_B \to N_B + dN_B = N_B - dN_A \tag{9.2}$$

In this case, there is no change in the temperature or volume of systems A and B, only in the number of particles. Since the change in the number of particles is small, we can find the change in Helmholtz free energy by a Taylor expansion:

$$dF = d(F_A + F_B) = dF_A + dF_B = \left(\frac{\partial F_A}{\partial N_A}\right)_{V,T} dN_A + \left(\frac{\partial F_B}{\partial N_B}\right)_{V,T} dN_B \ . \tag{9.3}$$

Since $N = N_A + N_B$ is constant, we know that $dN_A = -dN_B$:

$$\begin{aligned} dF &= \left(\frac{\partial F_A}{\partial N_A}\right)_{V,T} dN_A - \left(\frac{\partial F_B}{\partial N_B}\right)_{V,T} dN_A \\ &= \left(\left(\frac{\partial F_A}{\partial N_A}\right)_{V,T} - \left(\frac{\partial F_B}{\partial N_B}\right)_{V,T}\right) dN_A \end{aligned} . \tag{9.4}$$

Helmholtz free energy is minimal in equilibrium. In equilibrium, F is minimal, and this change must be zero:

$$\left(\frac{\partial F_A}{\partial N_A}\right)_{V,T} = \left(\frac{\partial F_B}{\partial N_B}\right)_{V,T} . \tag{9.5}$$

This looks very similar to what we found for the canonical system. This is the quantity related to particle transport which must be the same for the two systems to be in *diffusive equilibrium*. This is the X we were looking for!

Definition of chemical potential. We call this quantity the chemical potential. Indeed, we will use this as the definition of the chemical potential:

Chemical potential:

$$\mu(T,V,N) = \left(\frac{\partial F}{\partial N}\right)_{V,T} , \tag{9.6}$$

or

$$\mu(T,V,N) = F(V,T,N+1) - F(V,T,N) . \tag{9.7}$$

This provides us with a condition for diffusive equilibrium:

A condition for **diffusive equilibrium** between two systems A and B (at constant T and V) is the the chemical potential in the two systems are the same: then

$$\mu_A = \mu_B . \tag{9.8}$$

Chemical potential as a discrete difference. It may seem strange to define a derivative in N, which is essentially a discrete quantity: the number of particles. We could therefore instead have used the definition

$$\mu = F(N+1,V,T) - F(N,V,T) , \tag{9.9}$$

which is the same as the definition above for large N. This definition also emphasizes that:

> The chemical potential can be interpreted as the change in Helmholtz free energy when we add one particle to the system, keeping the temperature and volume constant.

Transport occurs from high to low chemical potential. We also know how a system that it out of equilibrium develops: Helmholtz free energy must decrease. The change in Helmholtz free energy is we transfer an amount dN_A from part A to part B is

$$dF = \left(\frac{\partial F_A}{\partial N_A}\right)_{V,T} dN_A - \left(\frac{\partial F_B}{\partial N_B}\right)_{V,T} dN_A \tag{9.10}$$

$$= \mu_A dN_A - \mu_B dN_A = (\mu_A - \mu_B)\, dN_A \,. \tag{9.11}$$

If the system is out of (diffusive) equilibrium, we know that the system will change in such a way that the change in N_A will result in a lower value of F as the system approaches equilibrium, hence $dF \leq 0$. We therefore see that if $\mu_B > \mu_A$ then $dN_A > 0$ for dF to be negative, hence the flux is from B to A when μ_B is larger than μ_A. We therefore conclude that

> **Flux and chemical potential**: Particles flow from large to small chemical potential.

This will occur until the chemical potential in the two systems are equal and a further change in particles does not change Helmholtz free energy. This effects makes intuitive sense since taking a particle from a part with large chemical potential and transferring it to a part with lower chemical potential will reduce the total Helmholtz free energy of the system and therefore move the system towards equilibrium.

9.1.2 Thermodynamic identity for diffusive systems

We can now extend the thermodynamic identity to also include changes in N. We already found that $dF = -SdT - pdV + XdN$ for Helmholtz free energy. Previously we did not have a good intuition for X, but

now we have defined this X as the chemical potential. The potential for Helmholtz free energy is therefore

$$dF = -SdT - pdV + \mu dN \ , \tag{9.12}$$

and we can use this to find the thermodynamic identity

$$TdS = dE + pdV - \mu dN \ . \tag{9.13}$$

9.1.3 Chemical potential for multi-particle systems

If several chemical species are present, then each species will have its own chemical potential:

Chemical potential of species j in a multi-species system:

$$\mu_j(T, V, N) = \left(\frac{\partial F}{\partial N_j}\right)_{T,V,N_k} \ , \tag{9.14}$$

where all the other N_k, $k \neq j$, are kept constant in the derivative.

9.1.4 Example: Chemical potential of the Einstein crystal

We found the Helmholtz free energy for the Einstein crystal with energy level spacing $\Delta\epsilon$ to be:

$$Z_N = \left(\frac{1}{1 - e^{-\Delta\epsilon/kT}}\right)^N \ , \tag{9.15}$$

and

$$F = -kT \ln Z_N = NkT \ln\left(1 - e^{-\Delta\epsilon/kT}\right) \ , \tag{9.16}$$

which gives

$$\mu = \left(\frac{\partial F}{\partial N}\right) = kT \ln\left(1 - e^{-\Delta\epsilon/kT}\right) \ . \tag{9.17}$$

9.1.5 Example: Chemical potential of the Ideal gas

For an ideal gas, the partition function can be written as:

$$Z_N = \frac{1}{N!} Z_1^N \ , \ Z = Z_{\text{trans}}\, Z_{\text{vib}}\, Z_{\text{rot}} \ , \tag{9.18}$$

where $Z_{\text{trans}} = n_Q(T)\, V$ and $n_Q(T) = \left(2\pi mkT/h^2\right)^{3/2}$ for a three-dimensional ideal gas. For a monatomic gas $Z_{\text{vib}} = Z_{\text{rot}} = 1$, and $Z_1 = Z_{\text{trans}}$. Helmholtz free energy is therefore:

$$F = -kT \ln Z_N = -kT \left(N \ln Z_1 - \ln N!\right) \ . \tag{9.19}$$

We use Stirling's approximation $\ln N! \simeq N \ln N - N$, getting:

$$\frac{d}{dN} \ln N! = \frac{d}{dN} \left(N \ln N - N\right) = \ln N + 1 - 1 = \ln N \ , \tag{9.20}$$

which gives

$$F = -kT \left(\ln Z_1 - \frac{d}{dN} \ln N!\right) = -kT \left(\ln Z_1 - \ln N\right) = kT \ln(N/Z_1) \ . \tag{9.21}$$

We insert $Z_1 = Z_{\text{trans}} = n_Q(T)\, V$, getting:

$$\mu = kT \ln \left(\frac{N}{V\, n_Q(T)}\right) = kT \ln \left(\frac{n}{n_Q(T)}\right) \ . \tag{9.22}$$

(We get the same result if we use $\mu = F(N) - F(N-1)$ as the definition for the chemical potential, and then we do not even have to use Stirling's approximation!).

We see that the chemical potential increases with the number of particles per volume, n, which is consistent with our general observation above: Particles flow from parts with high n to systems with low n. (Notice that particle flow therefore tends to smooth out the particle density). We notice that for classical concentrations, that is when $n/n_Q \ll 1$, the chemical potential of an ideal gas is always negative.

9.1.6 Potential energy and the chemical potential

We can better understand the chemical potential by looking at a system with a difference in potential energy. Let us look at two systems A and B at the same temperature that may exchange particles, but the two systems are not yet in diffusive equilibrium. The two systems may for example be two ideal gases with different densitites, $n_B > n_A$, so that

$\mu_B > \mu_A$. If the two systems are placed in contact, particles will flow from B to A.

How can we stop that flow? By changing the chemical potential, for example by lifting one of the systems up in the gravitational field (if the particles have masses). How does a change in potential energy change the chemical potential? We can find out be addressing Helmholtz free energy and then taking the derivative to find the chemical potential. Helmholtz free energy consists of two parts, $F = E - TS$. We expect the energy E of the system to depend on the position in a gravitational field. If we lift an ideal gas from height $z = 0$ to a height z, this corresponds to adding a potential energy $u = mgz$ to all the particles. The energy therefore depends on the height, $E(z) = E(0) + mgz$. However, if we add a potential energy to the system, we only add a constant to all the energy levels for all the particles, we do not change the distribtion of energy levels. Therefore, we do not expect the entropy of the system to change, $S(0) = S(z)$. Helmholtz free energy is thererfore

$$F(z) = E(z) - TS(z) = E(0) + Nmgz - TS(0) = F(0) + Nmgz = F(0) + Nu \,, \tag{9.23}$$

where $u = mgz$ is the potential energy per particle. The chemical potential is therefore

$$\mu(z) = \left(\frac{\partial F}{\partial N}\right)_{V,T} = \left(\frac{\partial F(0)}{\partial N}\right)_{V,T} + u = \mu(0) + u \,. \tag{9.24}$$

The effect of adding a potential energy to system A — such as by lifting it in a graviational field — is to increase the chemical potential by the potential energy per particle, u.

It is common to divide the chemical potential into two parts: The *external chemical potential* is the potential energy per particle in an external field, and the *internal chemical potential* energy is the chemical potential that would be present without the external field.

Internal and **external** chemical potential:

$$\mu_{\mathrm{TOT}} = \mu_{\mathrm{int}} + \mu_{\mathrm{ext}} \,. \tag{9.25}$$

Where the external chemical potential, μ_{ext}, is the potential energy per particle in an external field, and the internal chemical potential, μ_{int}, is the chemical potential without the external field.

If $\mu_B > \mu_A$, we can therefore stop the flux of particles from system B to system A by increasing the potential energy per particle in system A by $u = \mu_B - \mu_A$. Then the chemical potential in the two systems are equal, and there is no net flow of particles. We could use this to measure the chemical potential: We apply a potential difference and determine at what potential difference net particle flow stops.

9.1.7 Example: Barometric pressure formula

Fig. 9.2 illustrates a situation where a part B of an ideal gas was lifted a height z in the gravity field — how must the density change for part A to be in equilibrium with part B after this displacement? How can we use the theory we haved developed for the chemical potential to predict the distribution of densities in the atmosphere?

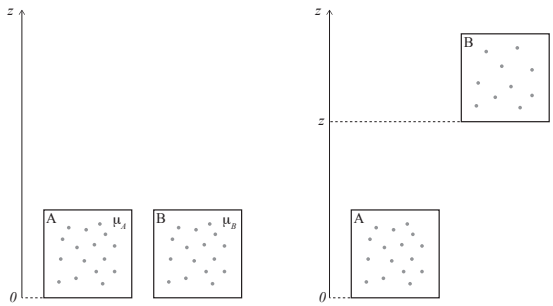


Fig. 9.2 Illustration of two parts of an ideal gas, where part B has been displaced a distance z in the vertical direction.

Let us for simplicity assume that the atmosphere has the same temperature T at all heights. (This is generally not true). The chemical potential of gas at a height z is then

$$\mu(z) = \mu_{\text{gas}} + \mu_{\text{grav}} = kT \ln(n(z)/n_Q(T)) + mgz \ , \tag{9.26}$$

where mgz is the potential energy of a gas particle at height z above the surface. In equilibrium, the chemical potential must be the same everywhere — otherwise there would be a flux from regions with high chemical potential to regions with low chemical potential. The chemical potential at $z = 0$ must therefore be the same as the chemical potential at z:

$$\mu(z) = kT \ln(n(z)/n_Q) + mgz = kT \ln(n(0)/n_Q) = \mu(0) \ , \tag{9.27}$$

which gives

$$n(z) = n(0)e^{-mgz/kT} \ . \tag{9.28}$$

What does that mean for the pressure of an ideal gas? If the temperature does not depend on height, the pressure is

$$pV = NkT \ \Rightarrow \ p = \frac{N}{V}kT = nkT \ , \tag{9.29}$$

and

$$p(z) = p(0)\, e^{-mgz/kT} = p(0)\, e^{-z/z_c} \ , \tag{9.30}$$

where $z_c = kT/mg$ is a characteristic height for the decay of the density.

Notice that the result depends on the mass of the gas molecules! For N_2 the mass of a molecule is 28 which gives $z_c = 8.5\,\text{km}$. Lighter molecules will extend further up - and will mostly have escaped from the atmosphere. Notice that T is not really constant, and $n(z)$ is generally more complicated. And also notice that various gases have different m, and will therefore have different values of z_c, which means that the composition of the air will change with distance z as well.

9.1.8 Example: Batteries

ams 10: Add this.

9.1.9 Example: Helmholtz free energy for system in potential field

We study a system with where each particle is described by the states i with energies ϵ_i. We know that we can describe the thermodynamics of this system by first finding the partition function and then Helmholtz free energy for a system of N particles. How does the partition function and Helmholtz free energy change if we add a potential energy ϵ_0 to each particle?

The states, i', of the system after we add the potential energy can be numbered the same way as the states i of the original system, but their energies are changes to $\epsilon_i + N\epsilon_0$, where the factor N is included because there are N particles, each with a potential energy, ϵ_0, so that the total potential energy added to the system is $E_0 = N\epsilon_0$. The partition function, Z', for the system after adding the potential energy is therefore:

$$Z' = \sum_i e^{-(\epsilon_i + N\epsilon_0)/kT} = e^{-N\epsilon_0/kT} \sum_i e^{-\epsilon_i/kT} = e^{-N\epsilon_0/kT} Z \ , \qquad (9.31)$$

where Z is the partition function for the system without the potential energy. We find Helmholtz free energy from the partition function

$$F' = -kT \ln Z' = -kT \ln \left(e^{-N\epsilon_0/kT} \, Z \right) \qquad (9.32)$$

$$= -kT \ln e^{-N\epsilon_0/kT} - kT \ln Z = N\epsilon_0 + F \ , \qquad (9.33)$$

where F is Helmholtz free energy for the system without the potential energy. The chemical potential of the system is then

$$\mu' = \left(\frac{\partial F'}{\partial N} \right)_{V,T} = \epsilon_0 + \mu \ , \qquad (9.34)$$

where μ is the chemical potential without the potential energy. Adding a potential energy per particle therefore results in adding the potential energy to the chemical potential.

9.1.10 Example: Modeling potential gradients

Fig. 9.3 shows a system with a variation in the chemical potential. The system stretches from $x = 0$ to $x = L$, and within the system the chemical potential is $\mu(x)$. Is this system in equilibrium? What potential energy must you add in order to make the system be in equilibrium?

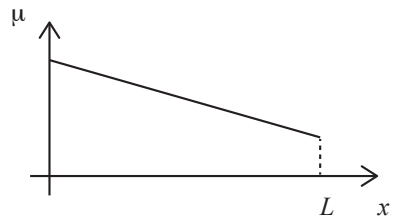


Fig. 9.3 Plot of the chemical potential $\mu(x)$ as a function of x in a system.

Clearly, the system is not in equilibrium. Because if it were in equilibrium the chemical potential would be the same everywhere. Since the chemical potential is not the same, there will be a flow of particles from regions with high chemical potential to regions with small chemical potential.

We can make this argument more quantitative. Let us assume that the whole system is at a constant temperature T, and that diffusion is possible inside the system, but not across the boundaries at $x = 0$ and $x = L$. The whole system therefore has T, V, and N constant, and its equilibrium is determined from the minimum of Helmholtz free energy.

We can find Helmholtz free energy of the whole system by dividing the system into small parts from x to $x + \Delta x$, where Helmholtz free energy of this small part is $F(x)$. If each part is independent, that is, if the energy levels in each part are independent, then each part will have its separate sum over states in the partition function and Helmholtz free energy will be the sum of the free energies for each part. The total free energy is then

$$F_{tot} = \sum_x F(x) \ . \tag{9.35}$$

We can use this to discuss how the chemical potential must develop as the system approaches equilibrium. If the chemical potential is not constant, as sketched in the figure, we can redure the total Helmholtz free energy by moving a particle from a position with high chemical potential to a position with low chemical potential. When we remove a particle from a point, the decrease in F is the chemical potential at that point, and, similarly, when we add the same particle to another point, the increase in F is the chemical potential. If we remove a particle from a region with high chemical potential and place it in a region with low chemical potential, we therefore remove more from Helmholtz free energy than we put back, hence the free energy is reduced. This process will continue until the chemical potential is equal everywhere.

Another alternative to obtain equilibrium is to add an external potential, so that the total chemical potential is constant. For example, we can add the external potential $u(x)$ given as $u(x) = u_0 - \mu_0(x)$, where u_0 is a constant and $\mu_0(x)$ is the initial chemical potential. The total chemical potential is then:

$$\mu(x) = \mu_0(x) + u(x) = \mu_0(x) + u_0 - \mu_0(x) = u_0 \ , \tag{9.36}$$

which is a constant. In practice, we may not add any potential, but only simple potentials such as a linear potential in the case of an electrical potential difference for charged particles or a gravitional field for particles with non-negligible masses.

Clicker question on $\mu(x)$ as introduction.

9.2 Phase equilibrium at constant volume

We started out by wanting to study diffusive equilibrium because we were interested in processes such as the equilibrium between different phases of a material and chemical reactions that transform one material to another. Now, we have introduced a new concept: the chemical potential. Let us see how we can use this concept to address the equilibrium between phases of a material.

The basic phases of matter are solid, liquid and gas (and plasma). A phase equilibrium is an equilibrium between two or several phases of a material. If I hold a cup of water in the lecture room, and place a lid on the cup, I have a system with water molecules in a given volume, at a given temperature (the temperature in the room is the temperature of the heat reservoir), and a given number of molecules, particles, of water. Inside the cup, there will be a given number of water molecules. What determines how much of the water is a solid, how much is a liquid and how much is a gas? If the cup contained water at room temperature, we expect the closed cup to contain some water and some water vapor. How can we determine the number of particles in each phase and how can we determine the state of the phases, such as the pressure in the vapor?

9.2.1 Equilibrium between an ideal gas and an ideal crystal

To address this situation, we will discuss phase equilibrium in a simplified model in which we can understand all the processes in detail. We will address equilibrium between a solid, a crytal of a given particle type, and a gas consisting of the same particles, such as an equilibrium between ice and water vapor at sub-zero temperatures. We will model this system, a cup with water molecules, as a system with a given temperature, T, volume, V, and number of particles, N. The system consists of two subsystems: a gas system (subindex g) of N_g gas particles in the gas phase; and a solid system (subindex s) of N_s solid particles. The gas will be modelled as an ideal gas and the solid as an Einstein crystal. The system is illustrated in fig. 9.4 and the symbols used to describe the system are listed in Table **??**.

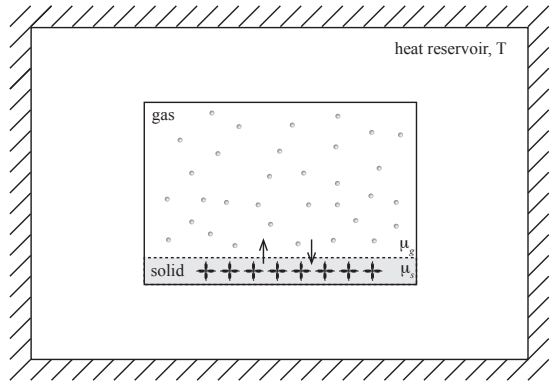


Fig. 9.4 Illustration of a system with phase equilibrium between a solid and a gas phase of the same particles.

Table 9.1 Symbols used to describe phase equilibrium between an Einstein crystal and an ideal gas of the same particles.

Symbol	Description
N_g	Number of gas particles
N_s	Number of solid particles
V_g	Volume of gas
V_s	Volume of solid
n_g	Density of gas
n_s	Density of solid

Approach. Our plan is the find the chemical potential for each of the two subsystems. In equilibrium, the chemical potentials must be equal, otherwise there would be a flow of particles from the part of the system with high chemical potential towards the part with a low chemical potential. We can calculate the chemical potential from Helmholtz free energy, which is turn can be found from the partition function for each of the two systems.

Helmholtz free energy for an ideal gas. We have previously found that Helmholtz free energy for a gas in a volume V_g is

$$F_g = -N_g kT \left(\ln \frac{n_Q}{n_g} + 1 \right) \ , \ n_Q(T) = \left(\frac{2\pi mkT}{h^2} \right)^2 \ , \tag{9.37}$$

where $n_g = N_g/V_g$ is the density of the gas.

Helmholtz free energy for an Einstein crystal. The Helmholtz free energy for the Einstein crystal is

$$F_s = N_{osc} kT \ln \left(1 - e^{-\Delta\epsilon/kT} \right) \ . \tag{9.38}$$

Here, we have written N_{osc} because this is the energy is for a given number of *oscillators* and not a given number of *particles*. For a three-dimensional system, each particle can oscillate in three different, independent directions (x, y, and z), and the number of oscillators is therefore $N_{osc} = 3N_s$, where N_s is the number of particles in the solid. Helmholtz free energy for an Einstein crystal with N_s particles is therefore:

$$F_s = 3N_s kT \ln\left(1 - e^{-\Delta\epsilon/kT}\right) . \tag{9.39}$$

Comparing energy levels for Helmholtz free energy. To compare the energy for the gas and the crystal, the reference level of the energies must be the same. For a particle in a gas, we have assumed that the particle has zero energy when at rest. Similarly, for the crystal, we have assumed that the lowest energy level is at zero energy. However, we know that the particles in a crystal are bound to the crystal, and that we need to add an energy to such a particle to pull the particle from the crystal and make it part of the gas. The energy added corresponds to the binding energy, ϵ_0, of a particle in the crystal. This means that a particle in the lowest energy level in the crystal has an energy $-\epsilon_0$ compared with an energy at rest in the gas, we must therefore add a term $-N_s\epsilon_0$ to Helmholtz free energy for the crystal to account for the binding energies.

$$F_s = 3N_s kT \ln\left(1 - e^{-\Delta\epsilon/kT}\right) - N_s\epsilon_0 . \tag{9.40}$$

Finding the chemical potentials. The chemical potentials for both the gas and the solid can now be found from F_g and F_c:

$$\mu_g(V_g, T, N_g) = \left(\frac{\partial F_g}{\partial N_g}\right)_{V_g,T} = kT \ln \frac{n_g}{n_Q(T)} , \tag{9.41}$$

and

$$\mu_s(V_s, T, N_s) = \left(\frac{\partial F_s}{\partial N_s}\right)_{V_s,T} = 3kT \ln\left(1 - e^{-\Delta\epsilon/kT}\right) - \epsilon_0 . \tag{9.42}$$

In equilibrium we assume that the two chemical potentials are the same.

$$\mu_g(V_g, T, N_g) = kT \ln \frac{n_g}{n_Q(T)} = \mu_s(V_s, T, N_s) = 3kT \ln\left(1 - e^{-\Delta\epsilon/kT}\right) - \epsilon_0 . \tag{9.43}$$

We solve for n_g, getting:

$$n_g = n_Q(T)e^3\left(1 - e^{-\Delta\epsilon/kT}\right)e^{-\epsilon_0/kT} \ . \tag{9.44}$$

From the ideal gas law for the gas, we know that $pV_g = N_g kT$ and therefore $n_g = N_g/V_g = p/kT$. We insert this to get an equation for the pressure in the gas:

$$p(T) = kTn_Q(T)e^3\left(1 - e^{-\Delta\epsilon/kT}\right)e^{-\epsilon_0/kT} \ . \tag{9.45}$$

9.2.2 Vapor pressure

This result tells us that inside the closed cup, there will be both a solid and a gas, and the pressure of the gas will be given by $p(T)$. We call this the *vapor pressure* of the system and the curve $p(T)$ is called the *coexistence curve*. We have illustrated $p(T)$ in Fig. 9.5. What does this curve tell us? In this form, it describes the state inside the closed system.

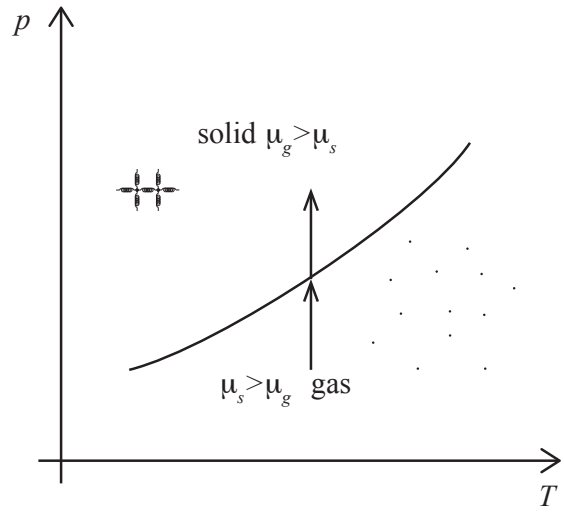


Fig. 9.5 Plot of the phase coexistence curve $p(T)$ for the ideal gas - Einstein crystal model system.

Let us see if we can use the model and the chemical potentials to understand what would happen if we instead of doing the experiment at constant volume, did the experiment at constant pressure: Instead of doing the experiment with a closed cup, we do the experiment inside a piston, where the outside of the piston has a pressure corresponding to the pressure in the lecture room. What would happen if we start the

system on the $p(T)$ curve described by the model, and then increase the pressure?

In the model, it is only the chemical potential of the gas that depends on the pressure. And the chemical potential of the gas will increase if we increase the pressure and decrease if we decrease the pressure. At $p(T)$ the chemical potential in the gas and the solid is the same. If we increase the pressure from this state, the chemical potential in the gas will increase while the chemical potential in the solid remains constant (this is in our model, not necessarily in the actual, physical system). As a result particles will diffuse from the gas and into the solid – until all the particles are in the solid. This means that the region above the $p(T)$ curve in the $p - T$ diagram corresponds to a region where the whole system is a solid. If we decrease the pressure, the chemical potential in the gas will decrease while the chemical potential in the solid will remain constant, hence all the particles will diffuse to the gas. The region below $p(T)$ in the $p - T$ diagram hence corresponds to a gas. The phase coexistence curve is therefore the border between the solid and the gas phases!

Now, there are a few problems with this argument. Primarily, the problem is that we have described the system in terms of T, V, and N, but the system we have sketched with the piston is not a system with constant V, but rather a system with constant p. We therefore need to introduce methods to address systems with constant T, p, N, including methods to find the equilibrium states in such systems. We will therefore introduce the (T, p, N) system and the corresponding free energy — Gibbs free energy — before we return to the model system to complete the discussion of the phase coexistence curve.

9.3 Gibbs Free Energy

We need to develop a theory to address systems where T, p, N are constant, corresponding to a system in a piston in a large reservoir with a given temperature and pressure as illustrated in Fig. 9.6. This corresponds to a laboratory situation or a natural situation where systems are in a surrounding gas or liquid and therefore in equilibrium with an external pressure. How can we develop such a theory?

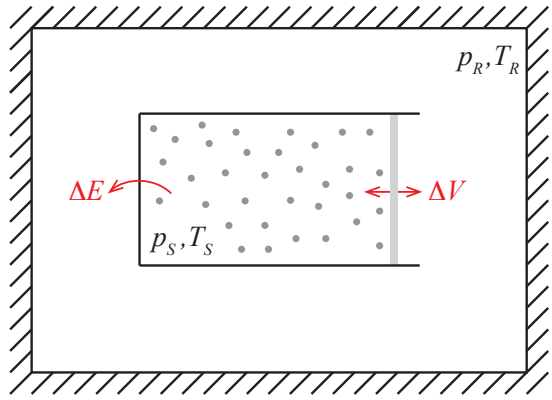


Fig. 9.6 Illustration of a system in thermal and mechanical contact with a huge reservoir with pressure p_R and temperature T_R.

9.3.1 Definition of Gibbs free energy

Changing variables from V to p. We have already found that for the micro-canonical system, we characterize the energy by E, the internal energy, and the development of the system toward equilibrium is determined by the second law of thermodynamics, $\Delta S \geq 0$. Based on this we developed Helmholtz free energy through a Legendre transformation: We exchanged the variable S with the conjugate variable T (conjugate because $(\partial E/\partial S)_{V,N} = T$). This was done by introducing

$$F = E - \left(\frac{\partial E}{\partial S}\right)_{V,N} = E - TS \; . \tag{9.46}$$

Now, we want to introduce a new free energy with the natural variables T, p, N. We know that F has the variables T, V, and N and the differential

$$dF = -SdT - pdV + \mu dN \; . \tag{9.47}$$

We would now like to replace pdV by some term which depends on dp. We see that we can do this by adding pV to F, because the new differential of $d(F + pV)$ will then not contain pdV, but instead a term Vdp. If we introduce $G = F + pV$ then the differential of G is

$$\begin{aligned} dG &= d\,(F + pV) \\ &= dF + pdV + Vdp \\ &= -SdT - pdV + \mu dN + pdV + Vdp \\ &= -SdT + Vdp + \mu dN \; . \end{aligned} \tag{9.48}$$

We see that G indeed has T, p, and N as natural variables.

Gibbs free energy is defined as

$$G(T,p,N) = E - TS + pV \ , \tag{9.49}$$

and its differential is

$$dG = -S\,dT + V\,dp + \mu\,dN \ . \tag{9.50}$$

We can therefore find the temperature, the volume and the chemical potential from Gibbs free energy:

$$S = -\left(\frac{\partial G}{\partial T}\right)_{p,N} , V = \left(\frac{\partial G}{\partial p}\right)_{T,N} , \mu = \left(\frac{\partial G}{\partial N}\right)_{T,p} \tag{9.51}$$

We leave it as an exercise for the reader to show what new Maxwell relations can be derived using Gibbs free energy.

Legendre transform method. We could also have found this by applying the Legendre transform. We could define G by replacing V by p in $F(T,V,N)$:

$$G = F - V\left(\frac{\partial F}{\partial V}\right)_{T,N} = F - V(-p) = F + pV \ , \tag{9.52}$$

which is the same as we found above.

9.3.2 Minimum Free Energy principle

We can also demonstrate a minimization principle for Gibbs free energy similar to what we found for Helmholtz free energy. We assume that our system is contained in a large reservoir with pressure p_R and temperature T_R. The total system, consisting of the system and the reservoir, is isolated, and therefore a microcanonical system. We therefore know that the entropy of the total system must increase and be maximum in equilibrium.

The entropy of the total system consists of the entropies of the system, S_S, and the entropy of the reservoir, S_R, $S_{TOT} = S_R + S_S$. Let us now look at a small change of the energy and the volume of the system, ΔE_S and ΔV_S. Since the total energy and volume is conserved, $E_{TOT} = E_S + E_R$

and $V_{TOT} = V_S + V_R$, we know that $\Delta E_S = -\Delta E_R$, and $\Delta V_S = -\Delta V_R$: We change the energy in the system by transferring that energy to the reservoir and if we increase the volume in the system, the volume of the reservoir decreases by the same amount. For any such change in the internal distribution of energy or volume, the total entropy must increase:

$$\Delta S_{TOT} = \Delta S_R + \Delta S_S \geq 0 \; . \tag{9.53}$$

How can we express the change in entropy of the reservoir in terms of changes in the system? First, we notice that the change in energy of the reservoir must be due to the heat transferred to the reservoir or the work done on the reservoir (the first law of thermodynamics):

$$\Delta E_R = Q_R + W_R = T_R \Delta S_R - p_R \Delta V_R \; . \tag{9.54}$$

where we have used that the heat transferred is $Q_R = T_R \Delta S_R$. Notice the sign of the work done on the reservoir. This means that the change in entropy of the reservoir is

$$T_R \Delta S_R = \Delta E_R + p_R \Delta V_R \; . \tag{9.55}$$

Since the total energy and the total volume is conserved, we have that

$$\Delta E_R = -\Delta E_S \; , \; \Delta V_R = -\Delta V_S \; , \tag{9.56}$$

which we insert into (9.55), getting

$$T_R \Delta S_R = -\Delta E_S - p_R \Delta V_S \; . \tag{9.57}$$

We insert this into (9.53), getting the condition

$$T_R \Delta S_{TOT} = T_R \, \Delta S_R + T_R \, \Delta S_S = T_R \Delta S_R - \Delta E_S - p_R \Delta V_S = \Delta \left(T_R S_S - E_S - p_R V_S \right) \geq 0 \tag{9.58}$$

In equilibrium $T_R = T_S$, and we recognize this as Gibbs free energy, $G_S = E_S - T_S S_S + p_S V_S$:

$$T_R \Delta S_{TOT} = -\Delta \underbrace{\left(E_S - T_S S_S + p_S V_S \right)}_{G_S} \geq 0 \; \Rightarrow \; \Delta G_S \leq 0 \; . \tag{9.59}$$

This demonstrates that Gibbs free energy always decreases and is minimum in equilbrium.

For a (T, p, N) system Gibbs free energy is $G = F + pV = E - TS + pV$. Gibbs free energy decreases as the system approaches equilibrium:

$$\Delta G \leq 0 \tag{9.60}$$

and is minimum in equilibrium.

This principle will turn out to be a vary powerful theoretical and practical tool as we will see below.

9.3.3 Gibbs free energy and the chemical potential

Gibbs free energy is an extensive quantity, so that Gibbs free energy for a system of N particles is

$$G(T, p, N) = N g(T, p) \, , \tag{9.61}$$

where $g(T, p)$ is the Gibbs free energy per particlem, which only depends on the intensive variables T and p.

Notice that since T and p are intensive, this means that

$$g(T, p) = \frac{G(T, p, N)}{N} = \left(\frac{\partial G}{\partial N} \right)_{T,p} = \mu(T, p) \, . \tag{9.62}$$

Notice that this is true for G, but it is not true for F!

9.3.4 Gibbs free energy for a system of several species

If a system consists of k different species of particles with chemical potentials $\mu_j(p, T)$, $j = 1, 2, \ldots, k$, then Gibbs free energy for the system is

$$G = \sum_j N_j \, \mu_j(p, T) \, . \tag{9.63}$$

The thermodynamic identity for a k-species system becomes

$$T \, dS = dE + p \, dV - \sum_{j=1}^{k} \mu_j \, dN_j \tag{9.64}$$

and the differential for G becomes

$$dG = -T\,dS + V\,dp + \sum_{j=1}^{k} \mu_j\,dN_j\ . \tag{9.65}$$

We can also for this expression for Gibbs free energy demonstrate that Gibbs free energy for a system of k species is also minimum in equilibrium.

Notice that the species can be molecules or atoms of different types, but it can also be atoms or molecules of the same type, but in different phases, where the particles will have different chemical potentials.

9.4 Phase coexistence at constant pressure

We have now developed the basic tools needed to discuss the behavior of a system at constant T, p, and N. This corresponds to the case where we have a system of a given number of particles inside a box, but where the volume of the box is varying while the pressure outside (and inside) the box is constant. This is what happens in a system with a piston. Now, we can address the same question as we did above for the T, V, and N system: What is the state of a system of particles, which may be in several phases, inside a piston for a given set of T, p, and N values?

9.4.1 Phase coexistence in a simple model

We address the same simplified system as we did before. In our simplified model the system consists of N particles, of which N_g are in a gas phase and N_s are in a crystal phase. If the system is in a large reservoir with constant temperature and pressure and the number of particles does not change in the system, the equilibrium state of the system is determined by the minimum of the total Gibbs free energy of the entire system, consisting of both gas and crystal particles.

Equilibrium condition. The equilibrium of this system can be found either as a minimum of the Gibbs free energy, or as the state when the chemical potential in the two phases are the same. The total Gibbs free energy for this system can be written as

$$G_{TOT} = N_s g_s(p,T) + N_g g_g(p,T)\ , \tag{9.66}$$

where $g_s(p,T) = \mu_s(p,T)$ is the Gibbs free energy per particle in the solid phase, that is for a particle in the crystal, and $g_g(p,T) = \mu_g(p,T)$

is Gibbs free energy per particle in the gas phase. We can therefore look for that state with the minimum of G_{TOT}, as this will be the equilibrium state. Alternatively, we realize that the minimum of G_{TOT} occurs when $dG_{TOT} = 0$, and the differential of Gibbs free energy for a multi-species system is

$$dG = -SdT + Vdp + \sum_j \mu_j dN_j \ . \tag{9.67}$$

For a system at constant T and p, we know that $dT = dp = 0$. We are then left with

$$dG = \sum_j \mu_j dN_j = \mu_g dN_g + \mu_s dN_s \ . \tag{9.68}$$

Since the total number of particles is conserved, $N = N_s + N_c = \text{const.}$, we know that $dN_s = -dN_g$, and therefore

$$dG = (\mu_g - \mu_s)\, dN_g = 0 \ \Rightarrow \ \mu_g(p,T) = \mu_s(p,T) \ . \tag{9.69}$$

This means that for a system with two species, where the total number of particles is conserved, the chemical potential of the two species must be identical in equilibrium. That is, we can use the same principle as we found for a system at constant (T, V, N).

Finding the equilibrium. Our plan is therefore to find Gibbs free energy for each of the species, and use the expressions for Gibbs free energy to find the chemical potentials, and from this find the equilibrium concentration of gas and solid.

Gibbs free energy for the solid. For the crystal, we find Gibbs free energy from

$$G_s(T,p,N_s) = F_s(T,V_s,N_s) + pV_s \ , \tag{9.70}$$

where V_s is the volume of the crystal. The volume does not appear in F_s, but we will need an expression for V_s for the pV_s term. We therefore make a small addition to the ideal crystal (Einstein crystal) model: We assume that there is a volume v_0 associated with each crystal particle, so that the volume of the crystal is $V_s = N_s v_0$. We insert the result

$$F_s(T,V_s,N_s) = 3N_s kT \ln\left(1 - e^{-\Delta\epsilon}\right) - N_s\epsilon_0 \tag{9.71}$$

from our calculations above and find

$$G_s(T,p,N_s) = 3N_s kT \ln\left(1 - e^{-\Delta\epsilon}\right) - N_s\epsilon_0 + pN_s v_0 \ . \tag{9.72}$$

We can then calculate $g_s(T,p) = \mu_s(T,p)$:

$$g_s(T,p) = \mu_s(T,p) = \frac{G_s}{N_s} = 3kT \ln\left(1 - e^{-\Delta\epsilon}\right) - \epsilon_0 + pv_0 \ . \qquad (9.73)$$

We see that the chemical potential for the crystal is modified by the term pv_0 compared with the expression we have above for the canonical T, V, N system.

Gibbs free energy for the gas. Similarly, we find Gibbs free energy for the gas:

$$G_g(T,p,N_g) = F_g(T,V_g,N_g) + pV_g = -N_g kT \left(\ln \frac{n_Q}{n_g} + 1\right) + pV_g \ . \qquad (9.74)$$

We apply the ideal gas law, $pV_g = N_g kT$ and $p = (N_g/V_g)kT = n_g kT$, which gives $n_g = p/(kT)$:

$$G_g(T,p,N_g) = -N_g kT \left(\ln \frac{n_Q kT}{p} + 1\right) + N_g kT = N_g kT \ln \frac{p}{n_Q\, kT} \ . \qquad (9.75)$$

which gives

$$g_g(p,T) = \mu_g(p,T) = kT \ln \frac{p}{n_Q\, kT} \ . \qquad (9.76)$$

(This is exactly the same result for μ_g as we would get if we inserted $n_g = p/(kT)$ into the result for $\mu_g(T,V)$ from our previous discussion.)

Condition for phase equilibrium. Phase equilibrium occurs for the p_e,T_e pair for which $\mu_c(p_e,T_e) = \mu_g(p_e,T_e)$, which is given by

$$\mu_g = kT_e \ln \frac{p}{n_Q(T_e)\, kT_e} = \mu_c = 3kT_e \left(1 - e^{-\Delta\epsilon/kT_e}\right) - \epsilon_0 + p_e v_0 \ . \qquad (9.77)$$

Unfortunately, this equation is not simple to solve, since p_e occurs both in the logarithm and directly in the equation. However, if $p_e v_0$ is much smaller than kT, we may neglect the $p_e v_0$ term and the resulting $p_e(T_e)$ curve will be the same as we found above.

Phase equilibrium from Gibbs free energy. Represeting the phase equilibrium problem using Gibbs free energy provides us with a richer description and understanding of the phase equilibrium. The system consists of N_s of the solid species and N_g of the gas species. The total Gibbs free energy is therefore

$$G_{TOT} = G_s + G_g = N_s g_s(p,T) + N_g g_g(p,T) \ , \qquad (9.78)$$

Fig. 9.7 illustrates g_s and g_g as a function of T for a given value of p. We see that since g_g is proportional to kT (except for a logarithmic dependency on T which is weaker), and g_s approximately is proportional to $3kT$, the slope of g_s with T is higher than that of g_g. This implies that $g_s < g_g$ for small T and $g_s > g_g$ for large T. We find that $g_s = g_g$ at $T = T_e$, that is at the phase co-existence curve.

From (9.78) we therefore see that when $g_s < g_g$, we get a minimum in the total Gibbs free energy is all the particles are in the solid phase, whereas when $g_g s < g_g$, we get a minimum in the total Gibbs free energy is all the particles are in the gas phase. This means that

$$N_s = \begin{cases} N \text{ when } T < T_e \ (g_s < g_g) \\ 0 \text{ when } T > T_e (g_s > g_g) \end{cases} , \ N_g = \begin{cases} 0 \text{ when } T < T_e \ (g_s < g_g) \\ N \text{ when } T > T_e \ (g_s > g_g) \end{cases} \tag{9.79}$$

as illustrated in Fig. 9.7.

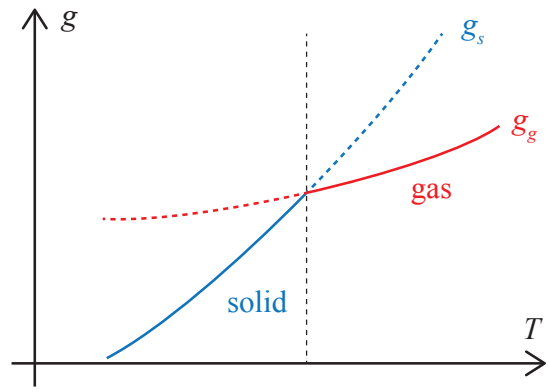


Fig. 9.7 Plot of the Gibbs free energies per particles for the gas and the crystal. The dashed regions indicate where the other phase has a lower Gibbs free energy, the solid lines show the minimum of the total Gibbs free energy and hence the actual behavior of the system.

Phase coexistence and Gibbs free energy. For this system, the $p_e(T_e)$ represents a singular region. It is only when the pressure and temperature are exactly on this curve that there is co-existence between the two phases in the system. If the system starts at the co-existence curve, and we increase the temperature in the system, then the chemical potential in the gas will be lower than the chemical potential of the crystal (or, similarly, the Gibbs free energy per particle will be lower in the gas than in the crystal), and the system will only contain gas particles. If we lower the temperature from $p_e(T_e)$, all the particles will be in the crystal phase.

Behavior at the phase coexistence curve. Let us look at what happens as we cross the coexistence curve. We start from a temperature that is

lower that T_e (for a given p), and then gradually (and slowly) increase the temperature. As long as $T < T_e$, the Gibbs free energy of the crystal will be lower than the Gibbs free energy of the solid, and all the particles will be in the solid phase until we reach $T = T_e$. At this point, the number of particles in the crystal and the gas phases are not known — they are not determined by the minimum in Gibbs free energy! Any choise of N_s and N_g (consistent with the condition that $N_g + N_s = N$) will give the same Gibbs free energy

$$G_{tot} = N_g g_g + N_s g_s \ , \tag{9.80}$$

because $g_g = g_s$ when $p = p_e(T_e)$. So what determines N_s and N_g in this case? The history! How we have come to this state. If we came to the state from $T < T_e$ then $N_g = N$ and $N_s = 0$.

Changing the relative amount of gas and solid. How can we change how much of the system is in the gas state and how much is in the solid state (that is, how can we change N_g and N_s) while the system is at $T = T_e$? We can do this by traferring thermal energy — heat — into the system. This will not lead to a change in temperature in the system, since the system will remain on the phase coexistence curve as long as there are particles in both the crystal and gas phases in the system. Instead, the heat will take particles from the crystal phase and into the gas phase. We realize that this does require energy, because the particles in the crystal are bound with a binding energy ϵ_0. But there are other contributions as well.

Energy for phase transformation. What is the heat needed to transfer one particle (atom or molecule) from the solid to the gas — as illustratrated in Fig. 9.8? We can analyze this using the first law of thermodynamics. The change in the energy of the system when $N_g \to N_g + 1$ and $N_s \to N_s - 1$ is

$$\Delta E = Q + W \ , \tag{9.81}$$

where the work done on the system by the environment is $W = -p\Delta V$. The heat transferred into the system therefore goes into work done by the system on the environment because the volume changes and a change in internal energy in the system:

$$Q = \Delta E + p\Delta V \ . \tag{9.82}$$

Change in internal energy per particle. We use the high-temperature approximation for the internal energy of the solid and the gas: $E_g \simeq$

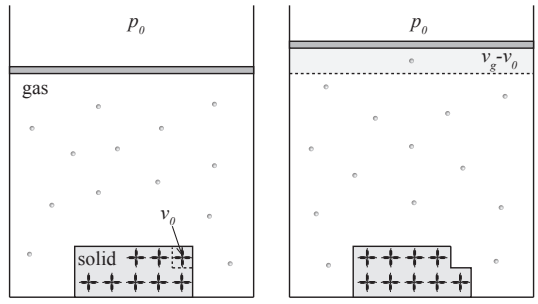


Fig. 9.8 Illustration of the change in the system as one particle moves from the solid to the gas phase.

$(3/2)N_gkT$, for a monatomic gass, and $E_s \simeq 3N_skT - N_s\epsilon_0$. The energy per particle is therefore $e_g = (3/2)kT$ and $e_s = 3kT - \epsilon_0$. We notice that without the binding energy, we would decrease the energy if we transferred one particle from the solid to the gas, since in thermal equilibrium each particle has $kT/2$ per degree of freedom of motion, and there are more degrees of freedom for a particle in a solid than in a (monatomic) gass. This difference becomes smaller for a diatomic gas where there are additional degrees of freedom for rotation and vibration. The change in energy from transferring one particle from the solid to the gas is therefore

$$\Delta e = e_g - e_s = \frac{3}{2}kT - (3kT - \epsilon) = \epsilon_0 - \frac{3}{2}kT \; . \tag{9.83}$$

This shows that if the binding energy is small ($\epsilon_0 \simeq 0$), then the internal energy decreseas when a particle is transferred from the solid to the gas.

Work done per particle. What is the work done by the system $-w = p\Delta v$ per particle that is transferred from the solid to the gas? The volume of the gas is V_g and the volume of the solid is $V_s = N_sv_0$, where v_0 is the volume per particle in the solid. The work per particle is therefore

$$w = p\left(\frac{V_g}{N_g} - \frac{V_s}{N_s}\right) \; . \tag{9.84}$$

We insert the ideal gas law for the ideal gas: $pV_g = N_gkT$, which gives $V_g/N_g = kT/p$. We insert this and $V_s = N_sv_0$, getting:

$$w = p\left(\frac{N_g\,kT}{p\,N_g} - \frac{N_s\,v_0}{N_s}\right) = kT - pv_0 \; . \tag{9.85}$$

Heat per particle for phase transformation. The heat added in order to transfer a single particle from the solid to the gas phase is therefore

$$q = e + w = \epsilon_0 - \frac{3}{2}kT + kT - pv_0 = \epsilon_0 - \frac{1}{2}kT - pv_0 \ . \tag{9.86}$$

Usually, we will assume that $pv_0 \ll kT$ and we can therefore neglect the last term.

This means that in order to transfer one particle from the solid to the gas, we need to add a thermal energy $q \simeq \epsilon_0 - kT/2$ to the system. This seems reasonable, since we need to add the binding energy ϵ_0, we need to perform work against the external pressure in order to increase the volume of the gas, and we gain energy due to the difference in degrees of freedom for a solid and a gas particle.

Latent heat in simplified model. We call the heat added to the system in order to transform one particle from solid to gas, the *latent heat* of the system. The latent heat is usually given as the heat needed to transfer one mole of the substance from one phase to another, and we use the symbol L for the latent heat of the system or the symbol ℓ for the latent heat per particle:

The **latent heat** in the simplified phase coexistence model is

$$\ell \simeq \epsilon_0 - kT/2 \ . \tag{9.87}$$

9.4.2 General phase coexistence

For the simplified model, we only address two phases: a solid and an gas. In general systems, there are typically three phases — gas, liquid and solid — and the phase diagram becomes more complicated. Fig. 9.9 illustrates the more general situation. (Let us for now assume that this plot is found from an experimental or a computational study). Phase coexistence occurs along the lines separating the various phases. How does the system go from one phase to another? Let us look at an isotherm in detail: The red dashed line in Fig. 9.9 corresponds to the line $T = T_e$. At high pressures, $p > p_e$, the system is in the liquid phase, and the whole system is a liquid. If we gradually decrease the pressure in the system, keeping the temperature constant, we move along the $T = T_e$ line. Eventually, the system will reach the pressure $p = p_e$ corresponding

to phase coexistence between liquid and gas. As we reach $p = p_e$ the whole system will be in liquid form. Based on our experience from the simplified system, we expect that we must add energy in the form of heat in order to bring particles from the liquid and into the gas state. This occurs at constant pressure, but the volume of the system will increase as the system changes from pure liquid, through liquid and gas coexistence, to pure gas.

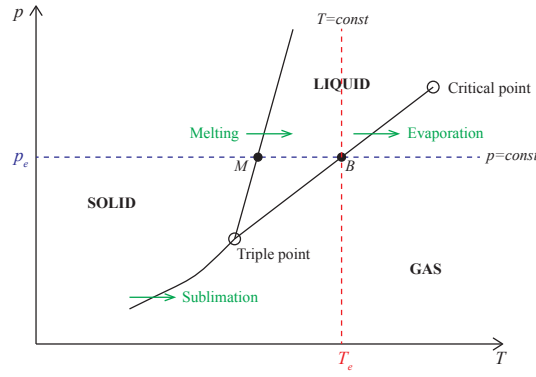


Fig. 9.9 Phase plot in P-T space.

Isothermal process in a p-V-diagram. How do we expect this isothermal process to look like in a p-V-diagram? We know that for an ideal gas, an isotherm has the shape $p = (NkT)/V$. We also expect $p(V)$ to fall rapidly with volume in the liquid state, but we do not yet know exactly how the functional shape is. We have sketched the behavior $p(V)$ along an isotherm in Fig. 9.10. If we start from a high pressure and slowly decrease the pressure at constant temperature, we argued above that when the system reaches $p = p_e$, the whole system will be a liquid, but as we add heat to the system to transfer particles from the liquid to the gas state, the volume will increase, while the pressure remains the same. This corresponds to a horisontal line in the p-V-diagram as shown in Fig. 9.10. Along this horisontal line there is both liquid and gas in the system, but the amount of gas increases (and the amount of liquid decreases) as the volume increases, until all the liquid has been transformed into a gas, and the system will behave as a gas, following the isothermal behavior, $p = (NkT)/V$ behavior of an ideal gas.

Alternative description. We could also think of this process as a process where we gradually changed the volume: We pull on the piston slowly and observe the internal state of the system. This will result in the behavior

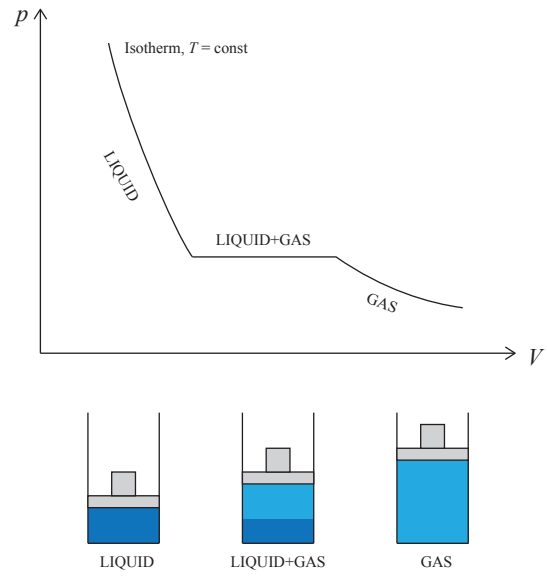


Fig. 9.10 Illustration of an isothermal process crossing the liquid-gas phase boundary plotted in p-V space.

from Fig. 9.10. Heat will also in this case flow into the system from the thermal bath in order to keep the system at the same temperature while the system is expanding and particles are transferred from liquid to solid. Notice that the pressure is a function of the volume, but that the volume is not a function of pressure: There are many possible volumes for the pressure p_e.

Melting, evaporation, and sublimation. It is common to call the diagram in Fig. 9.9 a *phase diagram*, since it divides the p-T space into regions that belong to different phases.

The various phase-transformation processes illustrated in Fig. 9.9 have different names. We call the transformation from a solid to a liquid, *melting*. We call the tranformation from a solid directly to a gas, *sublimation*, and we call the transformation from a liquid to a gas *evaporation*.

Triple-point and critial point. We have also marked two other special points in the $p(T)$ phase diagram in Fig. 9.9. The *triple-point* is the temperature and pressure at which we have coexistence between all three phases. We will return to methods to address the equilibrium near this point below. The triple-point for water at a pressure of one atmosphere is very close to 0°C, and this point is indeed used to define the kelvin scale.

The *critial point* is the point in p, V, T space beyond which gases and liquids are called supercritical, and there is no longer any phase transition between a liquid and a gas. Various values for the critical temperatures are listed in Table **??**.

Table 9.2 Critical temperatures for various substances.

Material	Critical T
O_2	154.3 K
H_2O	647.1 K
CO_2	304.2 K
N_2	126.0 K
H_2	033.2 K

9.4.3 General phase coexistence and Gibbs free energy

The equilibrium state of a system at constant T, p, and N is determined by the minimum of the Gibbs free energy of the system. For a system with three phases: gas (g), liquid (l), and solid (s), the total Gibbs free energy can be written as

$$\begin{aligned} G_{TOT} &= G_g(N_g, p, T) + G_l(N_l, p, T) + G_s(N_s, p, T) \\ &= N_g\, g_g(p, T) + N_l\, g_l(p, T) + N_s\, g_s(p, T)\ , \end{aligned} \tag{9.88}$$

where g_g, g_l, and g_s are the Gibbs free energy per particles for the gas, liquid and solid phases, and $N_g + N_l + N_s = N$ are the particles in the gas, liquid and solid phases respectively.

In general, the three free energies will have different dependencies of p and T. In Fig. 9.11 we have sketched the behavior of the Gibbs free energies for a sample system. Just like for the simplified model, we know that the system will be in the phase with the minimum Gibbs free energy per particle. Phase coexistence occurs where the Gibbs free energies per particle are equal: The evaporation transition (e) from gas to liquid occurs for the p_e, T_e values where $g_g(p_e, T_e) = g_l(p_e, T_e)$; and the melting ($m$) transition from liquid to solid occurs where $g_l(p_m, T_m) = g_s(p_m, T_m)$. This provides us with a method to calculate the phase coexistence curves if we know the Gibbs free energies for the various phases in the system. In the following we use this to find the shape of the $p(T)$ curve without knowing all the details of the underlying system.

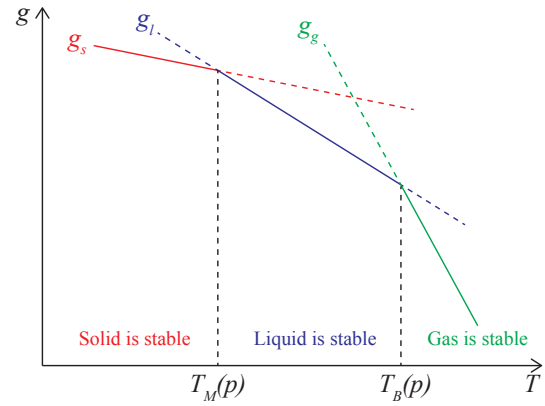


Fig. 9.11 Illustration of the gibbs free energies per particle as a function of temperature at constant pressure for the solid (s), liquid (l), and gas (g) phases. The melting and boiling points, T_M and T_B, are illustrated.

9.4.4 Clausius-Clapeyron's equation

Let us find a general expression for the coexistence curve, $p_e(T_e)$, using the liquid-gas transition as an example. Coexistence occur when the Gibbs free energy per particle in the liquid and gas phases are the same:

$$g_l(p,T) = g_g(p,T)\ . \tag{9.89}$$

This means that if we start at a point p_0, T_0 on the coexistence curve, and make a small change dT to $T_0 + dT$ we must also make a change dp in the pressure to still remain on the curve, so that (9.89) is satisfied for the original point, (p_0, T_0), and for the new point, $(p_0 + dp, T_0 + dT)$:

$$g_l(p_0, T_0) = g_g(p_0, T_0) \tag{9.90}$$

$$g_l(p_0 + dp, T_0 + dT) = g_g(p_0 + dp, T_0 + dT) \tag{9.91}$$

We subtract the two equations, getting

$$\underbrace{g_l(p_0 + dp, T_0 + dT) - g_l(p_0, T_0)}_{=dg_l} = \underbrace{g_g(p_0 + dp, T_0 + dT) - g_g(p_0, T_0)}_{=dg_g}\ . \tag{9.92}$$

For small changes dp and dT we can approximate $g_l(p_0 + dp, T_0 + dT)$ by its Taylor expansion around p_0, T_0:

$$g_l(p_0 + dp, T_0 + dT) \simeq g_l(p_0, T_0) + \left(\frac{\partial g_l}{\partial p}\right)_T dp + \left(\frac{\partial g_l}{\partial T}\right)_p dT\ , \tag{9.93}$$

and similarly for $g_g(p,T)$. We recognize the partial derivatives as

$$\left(\frac{\partial g_l}{\partial p}\right)_T = v_l \ , \ \left(\frac{\partial g_l}{\partial T}\right)_p = -s_l \ . \tag{9.94}$$

where the small symbols indicate that the volume and entropy are *per particle*. We insert this into (9.92), getting

$$dg_l = v_l\, dp - s_l\, dT = dg_g = v_g\, dp - s_g\, dT \ . \tag{9.95}$$

We rearrange to find an equation for dp/dT:

$$(v_l - v_g)dp = (s_l - s_g)dT \ , \tag{9.96}$$

and the coexistence curve $p(T)$ is therefore determined by the equation:

$$\left(\frac{dp}{dT}\right)_{\text{along curve}} = \frac{s_g - s_l}{v_g - v_l} \ . \tag{9.97}$$

Interpretation of Δs and Δv. How can we interpret the two differences $\Delta s = s_g - s_l$ and $\Delta v = v_g - v_l$? We see that Δs is the increase in entropy when we transfer one molecule from the liquid to the gas phase. Similarly, Δv is the change in volume when we transfer one molecule from the liquid to the gas. We can relate Δs to the heat added to the system:

$$Q = T\Delta s = T\,(s_g - s_l) = \ell \ , \tag{9.98}$$

which is called the *latent heat* of evaporation.

The **latent heat** of evaporation is

$$\ell = T\,(s_g - s_l) \ , \tag{9.99}$$

where s_g is the entropy per particle in the gas and s_l is the entropy per particle in the liquid phase.

Latent heat. It is common to describe the latent heat by the latent heat for the evaporation of one mole and not one particle. The molar latent heat is $L = T(s_g - s_l)$, where s_g and s_l are the entropies of one mole of gas and liquid phase of the molecules. (If we use molar values for the entropy or latent heat, we also need to use molar values for the volumes v.) Notice that the latent heat is generally dependent on temperature: $L = L(T)$. We can now rewrite (9.97) as

Clausius-Clapeyron's equation:

$$\frac{dp}{dT} = \frac{\Delta s}{\Delta v} = \frac{T\Delta s}{T\Delta v} = \frac{L(T)}{T\Delta v} \,, \tag{9.100}$$

where $L(T)$ is the (molar) latent heat for the transition and Δv is the (molar) volume change.

Simplification for liquid-gas transition. For the liquid-gas transition, we can simplify the expression because $v_g \gg v_s$, so that $\Delta v = v_g - v_s \simeq v_g$. In addition, the gas phase can be described approximately by the ideal gas law, so that $v_g = p/kT$ (or if we use molar values, $v_g = p/RT$, where $R = N_A k$, where N_A is Avogadro's number). The equation for the coexistence curve then becomes

$$\frac{dp}{dT} = \frac{L(T)}{T\Delta v} \simeq \frac{L(T)p}{kT^2} \,, \tag{9.101}$$

and

$$\frac{dp}{p} = \frac{L(t)dT}{kT^2} \,. \tag{9.102}$$

Given an expression for $L(T)$ we can integrate this equation to find $p(T)$.

Coexistence curve for a constant latent heat. If the latent heat $L(T) \simeq L_0$ is approximately constant over a range of temperatures T, we can find an analytical solution to the phase coexistence curve:

$$\int_{p_0}^{p} \frac{dp}{p} = \int_{T_0}^{T} \frac{L(T)dT}{kT^2} \simeq L_0 \int_{T_0}^{T} \frac{dT}{kT^2} \,, \tag{9.103}$$

which gives

$$\ln p - \ln p_0 = -\frac{L_0}{kT} + \frac{L_0}{kT_0} \,, \tag{9.104}$$

and

$$p(T) = p_0' e^{-\frac{L_0}{kT}} \,. \tag{9.105}$$

For many systems of practical importance this equation is sufficiently precise. For example, Fig. XXX shows that $\ln p$ as a function of $1/T$ for water is approximately a straight line over a temperature range corresponding to the range from ice to the critical point of water.

9.4.5 Example: Transition from graphite to diamond

9.4.6 Latent heat and enthalpy

We can use themodynamic tables that list enthalpies to find the latent heat for a transition. The latent heat correspond to the difference of the enthalpy, $H = E + pV$, between two phases when the process occurs at constant pressure p.

This is simple to demonstrate, because along the coexistence curve we know that $\mu_l = \mu_g$. The theromdynamic identity is therefore

$$TdS = dE + pdV - (\mu_g - \mu_l)\, dN \ , \tag{9.106}$$

but since $\mu_g = \mu_l$ the last term is zero. At constant pressure, the latent heat is the heat TdS transferred, which is

$$L = TdS = dE + pdV = dH = dE + pdV + \underbrace{Vdp}_{=0} = H_g - H_l \ . \tag{9.107}$$

At constant pressure the latent heat for a transition from phase g to phase l can be related to the difference in enthalpy for the two phases:

$$L = H_g - H_l \ . \tag{9.108}$$

Experimentally, we can find the values of H by integrating the heat capacity at constant pressure:

$$C_p = T\left(\frac{dS}{dT}\right)_p = \left(\frac{\partial U}{\partial T}\right)_p + p\left(\frac{\partial V}{\partial T}\right)_p = \left(\frac{\partial H}{\partial T}\right)_p \ . \tag{9.109}$$

and

$$H = \int C_p dT \ . \tag{9.110}$$

9.5 Van der Waals fluid model

We have so far used the ideal gas model as our canonical model for a gas. But the model is limited — it does not contain any interactions between the gas particles. How well does this model really describe a real gas or a real liquid? We can test this by studying the behavior of a

real, although simulated, fluid using the molecular dynamics simulation method we introduced previously.

9.5.1 Molecular dynamics simulation of Lennard-Jones gas

In a molecular simulation we can include any type of interactions between the molecules in the system. We could include a model of real water (see App. XX), but most of the properties of gases and fluids are well captured by a simpler model, such as the Lennard-Jones model used to describe the Argon system.

Lennard-Jones model system. In the Lennard-Jones model, the interatomic potential is

$$U_{LJ}(r) = 4\epsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right) . \tag{9.111}$$

This potential has a strong repulsive interaction at short ranges, described by a potential with a r^{-12} dependency, and a weak attractive interaction at long distances, described by a dipole-dipole interaction with a r^{-6} dependency. The Lennard-Jones system therefore contain both repulsive and attractive interactions. The Lennard-Jones system can describe all the three phases of the Argon system: a solid, a liquid, and a gas, depending on the pressure, density and temperature in the system.

Molecular simulation of phase equilibrium. We can use molecular simulations to study a phase equilibrium directly. We can model a microcanonical system with constant E, V, and N. In this system, if the energy is low enough, there will always be an equilibrium between a gas and a liquid phase. If we start the simulation from a regular pattern of atoms with random velocities, the system will quickly coalesce and coarsen: Droplets of liquid will spontaneously nucleate and grow, and the regions between the droplets will be filled with a gas with much lower density than the solid. Fig. 9.12 shows a sequence of snap-shots from a simulation with the following initialization file for the LAMMPS simulator:

```
# 3d Lennard-Jones coalesence
units       lj
atom_style  atomic
lattice     fcc 0.2
region      box block 0 20 0 20 0 20
create_box  1 box
```

```
create_atoms    1 box
mass         1 1.0
velocity     all create 0.1 87287
pair_style   lj/cut 2.5
pair_coeff   1 1 1.0 1.0 2.5
neighbor     0.3 bin
neigh_modify    every 20 delay 0 check no
fix 1 all nvt temp 0.6 0.6 1.0
dump         id all atom 100 vdw-coalesce01.lammpstrj
thermo       100
run      20000
```

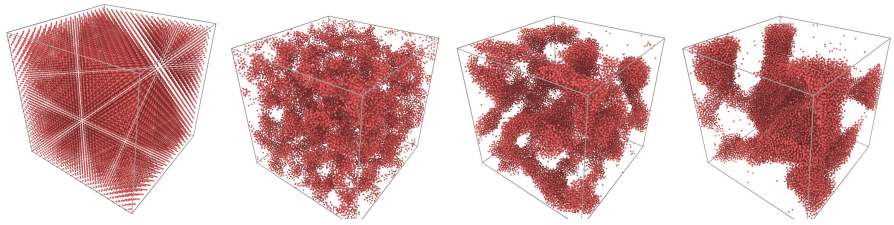


Fig. 9.12 Snapshots from simulations of a Lennard-Jones system at a low temperature, $T = 0.1T^*$, show coalescence and phase coexistence between a liquid and a gas phase.

In this case, the simulated system is at a constant temperature. This is specified by the `fix` command. Here, we provide a fix of the (N, V, T) type, specified by the option `nvt`. The remaining parts of the line

```
fix 1 all nvt temp 0.6 0.6 1.0
```

specifies what region the fix is applied to. The three numbers 0.6, 0.6, 1.0 tells that the fix should change the temperature from $T_0 = 0.6$ to $T_1 = 0.6$ over a simulation time corresponding to $\Delta t = 1.0$, and then continue with a fix at $T_1 = 0.6$ from that time onward. This simulation is therefore of a canonical system. Notice that it is also possible to model systems at a constant pressure using special techniques that change the volume of the system to keep the pressure (approximately) constant.

The resulting snapshots demonstrate that the model is able to model the phase coexistence between a liquid and a gas — and indeed also the non-equilibrium transition towards a stable equilibrium with a given density of both fluid and gas.

Measuring the equation of state in a molecular model. We can use the molecular dynamics model also to calculate the equation of state — the relation between the pressure, volume, particle number and temperature for a real gas with attractive as well as repulsive interactions. To do this, we need to vary the temperature, the number of particles, and the volume

and measure the pressure in the system. Here, we choose to model a microcanonical system, hence the temperature will fluctuate somewhat during the simulation. We keep the number of particles constant, but vary the temperature and the system size. We write a short script to generate input scripts for the molecular simulation program, and run the simulator from within the script. The resulting data is analyzed for each simulation, and then data from all simulations are collected using the follow script:

```
from pylab import *
import re
pressarr = array([]) # Store pressures
volarr = array([]) # Store volumes
Tarr = array([]) # Store temperatures
myvelocities = array([3.0,3.5,4.0,4.5,5.0,5.5,6.0])
myvolumes = array([0.010, 0.020, 0.040, 0.080])
for ivel in range(0,size(myvelocities)):
    for ivol in range(0,size(myvolumes)):
        # Change the word mydensity to myvolumes[ivol]
        infile = open("in.vdwstat30",'r')
        intext = infile.read()
        infile.close()
        replacestring = "%f" % (myvolumes[ivol])
        intext2=intext.replace('mydensity',replacestring)
        # Change the word myvelocity to myvelocities[ivel]
        replacestring = "%f" % (myvelocities[ivel])
        intext3=intext2.replace('myvelocity',replacestring)
        infile = open("in.tmp",'w')
        infile.write(intext3)
        infile.close()
        # Run the simulator
        print "Executing lammps < in.tmp"
        os.system("lammps < in.tmp") # Run lammps
        # Extract data from trajectory of simulation
        d = dump("tmp.lammpstrj") # Read sim states
        tmp_time,simbox,atoms,bonds,tris,lines = d.viz(0)
        dx = simbox[3]-simbox[0]
        dy = simbox[4]-simbox[1]
        vol = dx*dy # Volume of box
#        t = d.time(), n = size(t)
        # Read statistics calculated in simulation
        l = logfile("log.lammps")
        # Find pressure averaged over all timesteps
        press = average(l.get("Press"))
        temp = average(l.get("Temp"))
        # Store calculated values in arrays
        pressarr = append(pressarr,press)
        volarr = append(volarr,vol)
        Tarr = append(Tarr,temp)
# Plot the results
plot(Tarr,pressarr,'o'),xlabel("T"),ylabel("P"),show()
```

and the file used by the program is

```
# 3d Lennard-Jones gas
units        lj
dimension    3
atom_style   atomic
lattice      fcc mydensity
region       box block 0 10 0 10 0 10
create_box   1 box
create_atoms 1 box
mass         1 1.0
velocity     all create myvelocity 87287
pair_style   lj/cut 2.5
pair_coeff   1 1 1.0 1.0 2.5
neighbor     0.3 bin
neigh_modify every 20 delay 0 check no
compute         newke all ke
fix          1 all nve
dump         1 all custom 100 tmp.lammpstrj id type x y z vx vy vz
thermo_style custom step temp press vol c_newke
thermo       100
run          5000
```

Comparison with ideal gas law. These molecular simulations allow us to test the precision of the ideal gas law, and maybe we can also learn how we can modify the ideal gas law to better represent the material in the simulations. The ideal gas law states that $pV = NkT$, which gives us a functional form for $p = NkT/V$. We can measure and plot p as a function of T for a given number N of atoms and for a given volume V and test the accuracy of the model. Fig. 9.13 shows the measured $p(T)$ [1] from molecular simulations and the corresponding curve $p = NkT/V$. The ideal gas law is clearly reasonable, but not a very good approximation. The law seems to be wrong in two ways: There is an additional constant in the actual data: $p = p_0 + NkT/V$, and in addition, the slope of the $p(T)$ curve is not given by NkT/V, but is somewhat adjusted, $p = p_0 + c(V)\,NkT$. Let us now address what the origin of these differences may be.

9.5.2 Modifications to the ideal gas law

What can be the origin of these differences between the observed equation of state and the equation of state for the ideal gas? What effects in the Lennard-Jones model have not been included in the ideal gas law? There

[1] The pressure p is measured inside the simulation program. We are here simply reading the measured values from the simulation log file.

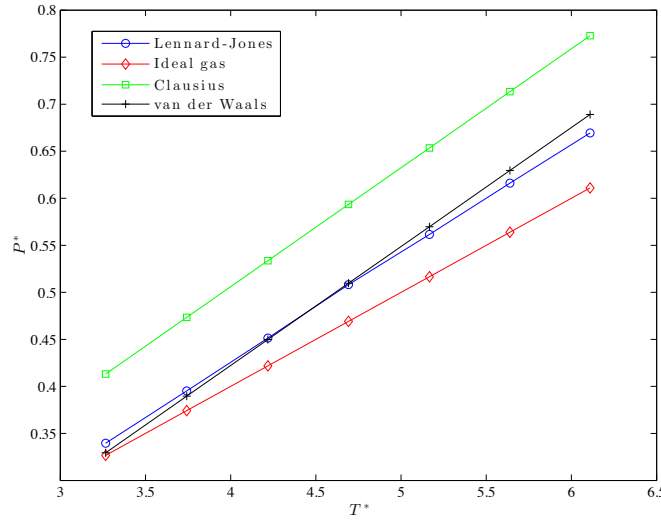


Fig. 9.13 Plot of $p(T)$ as measured in molecular dynamics simulations using the Lennard-Jones potential, for the ideal gas law, and for various modifications of the ideal gas law.

are two main interaction in the Lennard-Jones potential: A *repulsion* at short distances and an *attraction* at long distances. Neither are included in the ideal gas law. How can we include these effect in a simple modification?

Including the effect of a repulsive interaction. First, how can we include the repulsion at short distances? The effect of a repulsion is that atoms will not be very close to each other. When the atoms are far apart, that is for low densities, this effect does not matter much. But as the atoms are pressed closer together, we will notice that a significant part of the volume is not available for motion because it is too close to other atoms: We have a region of *excluded volume*.

We can make a simple model for the effect of an excluded volume by assuming that each atom has a small volume b associated with it, so that the excluded volume is Nb. The volume that is actually available for the atoms are then $V - Nb$ and not V. This will lead to a modification of the ideal gas law. However, we do not only want to modify the ideal gas law, we want to develop a new theory for a more realistic gas. This means that we would need expressions not only for the pressure, but also for the entropy and the chemical potential — which can be found from Helmholtz free energy. Let us therefore see how we can modify the free energy. For an ideal gas (IG), the Helmholtz free energy is

$$F_{IG} = -NkT\left(\ln\frac{n_Q(T)}{n}+1\right) = -NkT\left(\ln\frac{Vn_Q(t)}{N}+1\right) . \quad (9.112)$$

In this equation, the volume only occurs once, and we can include the excluded volume by replacing $V \to V - Nb$:

$$F_C = -NkT\left(\ln\frac{(V-Nb)n_Q(t)}{N}+1\right) . \quad (9.113)$$

This model for a gas is called the Clausius model, and the corresponding equation of state is found by calculating p as a derivative of F:

$$p_C = -\left(\frac{\partial F_C}{\partial V}\right)_{T,N} = \frac{NkT}{V-Nb} . \quad (9.114)$$

Clausius' law:

$$p_C = (NkT)/(V-Nb) . \quad (9.115)$$

We have plotted Clausius' law in the same plot as the ideal gas law and the observed behavior for $p(T)$. Clausius law is an improvement over the ideal gas law, but there are still clear discrepancies. Let us now see how we also can include the effect of attractive interactions.

Including the effect of attractive interactions. While Clausius law is an improvement of the ideal gas model, it does not include the correct shift, p_0, in pressure. We therefore need an additional term. Based on the plot in Fig. 9.13 we see that the additional term must reduce the effective pressure: The actual pressure must be lower than the pressure predicted by Clausius law. What physical effect may cause this? Attraction between the particles! We therefore need to include the effect of the attractive part of the Lennard-Jones model. This will give rise to a binding energy — a reduction in the energy and in Helmholtz free energy — when the density is high.

Let us construct a simplified model for the attractive interaction between the atoms. An atom in the gas will experience attractive interactions from all the other atoms. The potential energy of this interaction is the sum of the potential energies for all the other atoms. The potential energy of atom i is

$$u_i = \sum_{j\neq i} U_{LJ}(r_{ij}) , \quad (9.116)$$

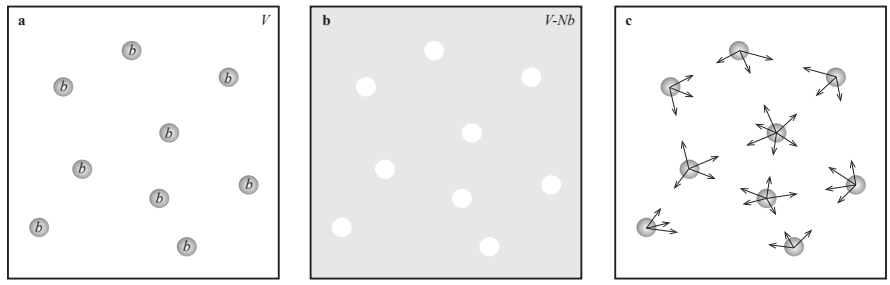


Fig. 9.14 Illustration of attractive and repulsive interactions in a van der Waals gas. **a** and **b** Illustration of the excluded volume Nb , corresponding to a volume b for each atom. **c** Illustration of the attractive interactions between the atoms. This gives rise to an average binding energy.

where r_{ij} is the distance from atom i to atom j. However, since the potential energy only depends on the distance, we can instead sum over distances, but then we need to include the number of atoms at each distance, which corresponds to the density of atoms measured around atom i, $n_i(r)$ times the volume element dV:

$$u_i = \int_V U_{LJ}(r) n_i(r) dV \ . \tag{9.117}$$

For each atom, the actual $n_i(r)$ will vary in time as all the other atoms move about. Instead of calculating this integral for each atom, we will assume that all atoms experience the same, avarage density $n(r)$ as a function of distance to the atom, and that this average density can be written as $n(r) = n\phi(r)$, where n is the average density of the system and $\phi(r)$ gives the variations relative to the average density. The potential energy of atom i is then:

$$u_i = \int_V U_{LJ}(r) n\phi(r) dV = n \underbrace{\int_V U_{LJ}\phi(r) dV}_{=-2a} = -2an \ . \tag{9.118}$$

The integral does not depend on N or V. We have introduced the notation $-2a$ for the integral. (You will see below why we have introduced the factor 2). The theory can be made even simpler by assuming that $\phi(r)$ is constant in space, but this is not necessary for our argument.

The potential energy of one atom is u_i. What is the potential energy of the whole gas? The gas consists of N atoms, and it may therefore be tempting to simply multiply u_i by N. But this is not correct, because we should only include the potential energy for each pair of particles once,

and the number of pairs is $N(N-1)/2$ and not N^2. Since u_i already includes the interactions from N atoms, we must multiply u_i by $N/2$ to find the total potential energy: $u_{tot} = (N/2)(-2an) = -anN$. This potential energy must be added to the Helmholtz free energy without interactions to find the (approximative) Helmholtz free energy with attractive interactions:

$$F_{vdw} = -NkT\left(\ln\frac{n_Q(T)}{n}+1\right) - anN \tag{9.119}$$

$$= -NkT\left(\ln\frac{Vn_Q(T)}{N}+1\right) - a\frac{N^2}{V}\ . \tag{9.120}$$

This model is called the *van der Waals model* of a gas. The model was introduced in the doctoral dissertation of Johannes Diderik van der Waals (1837-1923; Nobel 1910). The equation of state for the van der Waals gas is found from the pressure:

$$p = -\left(\frac{\partial F_{vdw}}{\partial V}\right)_{T,N} = \frac{NkT}{V-Nb} - \frac{N^2a}{V^2}\ , \tag{9.121}$$

which also can be written as

$$\left(p+\frac{N^2a}{V^2}\right)(V-Nb) = NkT\ , \tag{9.122}$$

which is called the *van der Waals equation of state*.

The **van der Waals model** for a gas:

$$F_{vdw} = -NkT\left(\ln\frac{(V-Nb)n_Q(T)}{N}+1\right) - a\frac{N^2}{V} \tag{9.123}$$

Equation of state for the van der Waals system:

$$\left(p+\frac{N^2a}{V^2}\right)(V-Nb) = NkT\ , \tag{9.124}$$

The van der Waals model is described by two parameters, the excluded volume b and the attractive interaction integral, a. First, let us see how these parameters are determined, then let us see how we can describe the

behavior of the system in a general way, where the details of the particular parameters values only are included as dimensionless parameters.

Parameters for the Lennard-Jones system. How can we determine the values of b and a for an actual gas? We can find these either from experimental data, by simulations of more complicated fluids such as water, or by theoretical considerations for simple fluids such as a Lennard-Jones system. The actual values of a and b will be different for different substances, and they may even vary for the same substance.

For the Lennard-Jones system we may estimate the values of a and b. The excluded volume corresponds to a volume that is not available for motion. This must be related to σ, the characteristic length of the potential. We could assume that the excluded volume starts where the potential is zero, that is for a volume of radius σ in the Lennard-Jones potential. However, this is too large, since the atoms often may be compressed into this zone. Instead, we will guess that the exluced volume is one half of the volume corresponding to σ, that is

$$b = v_{ex} = (1/2)(4\pi/3)\sigma^3 = (2\pi/3)\sigma^3 \ . \tag{9.125}$$

We can estimate a from the integral of the potential, assuming that the density is constant, $n(r) = n$ ($\phi(r) = 1$):

$$a \simeq \int_{\sigma/2^{1/3}}^{\infty} -\frac{\sigma^6}{r^6} 4\pi r^2 dr = \frac{8\pi}{3}\sigma^3 \ . \tag{9.126}$$

This $p(T)$ curve for this set of parameter values are plotted in Fig. 9.13, and the resulting model fits excellently to the data measured in molecular dynamics simulations of the Lennard-Jones model.

9.5.3 Properties of the van der Waals system

The van der Waals gas describes a gas with interactions — both attraction and repulsion. How does the resulting gas behave? And does this system display the behavior of both a gas and a fluid, just as we saw the simulation system did? In order to answer these questions we need to analyze the behavior of the van der Waals gas. We start by studing the behavior of isotherms, which shows that the system indeed displays a phase equilibrium, then we use the Gibbs free energy to discuss the phase equilibrium of the system, and discover that phase equilibrium only occurs at low temperatures and pressures, at sufficiently high temperatures and

pressures — beyond what we call the critical point — there is no longer a difference in phase.

Non-dimensional equation of state. The equation of state provides the relation between p and V for a given temperature, T, and number of particles (atoms/molecules) in the gas. However, the equation of state contains two parameters, a, and b, that are material specific — the values of a and b depens on the particular gas we are studying. But the general form of the equation is state is general. Could we make it even more general by rescaling the pressure, volume and temperature of the gas? We may say that the values of a and b can be included as characteristic values for the pressure, volume and temperature, and then we can measure all the quantities in these quantities. For example, it seems reasonable to measure the volume in terms of the excluded volume, Nb. We can rewrite the equation of state in a form where this is clear:

$$\left(p + \frac{(Nb)^2}{V^2}\frac{a}{b^2}\right)\left(\frac{V - Nb}{Nb}\right) = \frac{NkT}{Nb} \,, \tag{9.127}$$

and we see that we may measure pressure in terms of a/b^2, giving:

$$\frac{a}{b^2}\left(\frac{p}{a/b^2} + \frac{(Nb)^2}{V^2}\right)\left(\frac{V - Nb}{Nb}\right) = \frac{NkT}{Nb} \,. \tag{9.128}$$

We introduce the characteristic volume, $V^* = Nb$, and the characteristic pressure, $p^* = a/b^2$, and divide by a/b^2 on both sides, getting

$$\left(\frac{p}{p^*} + \frac{(V^*)^2}{V^2}\right)(V^* - 1) = \frac{kTb^2}{ab} \,. \tag{9.129}$$

Finally, we introduce the characteristic temperature, $T^* = a/kb$, leaving us with a very simple equation of state:

$$\left(\frac{p}{p^*} + \left(\frac{V^*}{V}\right)^2\right)(V^* - 1) = \frac{T}{T^*} \,. \tag{9.130}$$

Why did we introduce these reduced coordinates? Now, all the material-specific details are in the characteristic values, whereas the equation of state is general. We can therefore study general properties of the van der Waals gas — properties that are valid for all van der Waals gases.

While the characteristic values p^*, V^*, and T^* we have introduced are very simple and lead to a simple equation of state in the rescaled

coordinates, it is customary to use a slightly different set of characteristic values, which is only a numerical rescaling: $p_c = p^*/27 = a/27b^2$, $V_c = 3V^* = 3Nb$, and $kT_c = 8/27\,kT^* = 8/27\,ka/b$. We can rewrite the equation of state using these variables (see exercise), getting:

$$\left(p' + \frac{3}{V'^2}\right)\left(V' - \frac{1}{3}\right) = \frac{8}{3}T' \,. \tag{9.131}$$

where $p' = p/p_c$, $V' = V/V_c$, and $T' = T/T_c$. The only reason to do this is that the behavior of the van der Waals gas with these rescaled (non-dimensional) parameters, have simpler numerical values for important points in the phase diagram. We will therefore use this rescaling.

Law of corresponding states. This equation contains no remnants of a or b! There are only numbers and the non-dimensional pressure, volume and temperature. We can therefore study the behavior of this equation of state and learn about all possible van der Waals gases! This way of writing the equation of state is called the *law of corresponding states*:

The **law of corresponding states**:

$$\left(p' + \frac{3}{V'^2}\right)\left(V' - \frac{1}{3}\right) = \frac{8}{3}T' \,. \tag{9.132}$$

In terms of the dimensionless variables $T' = T/T_c$, $V' = V/V_c$, and $T' = T/T_c$, where p_c, T_c, and V_c correspond to the pressure, temperature and volume at the critical point of the gas.

Notice that the behavior of real gases is not to a high precision in correspondence with this equation.

Plotting the behavior of a van der Waals gas. How can we now characterize the behavior of the van der Waals gas? We could start to look at isotherms. We know that for an ideal gas the isotherms behave as $p = (NkT)/V$ in a p-V-diagram. How does the isotherms of a van der Waals gas look like? We plot isotherms for various values of T' using the following script:

```
import numpy as np
import matplotlib.pyplot as plt
That = [0.9,0.95,1.0,1.05]
for i in range(len(That)):
    T = That[i]
```

```
    V = np.linspace(0.5,2.0,1000)
    p = 8.0/3.0*T/(V-1/3)-3.0/(V**2)
    plt.plot(V,p,label="$\\hat{T} = "+str(T)+"$")
plt.xlabel('V/V_c')
plt.ylabel('p/p_c'); plt.legend()
```

Where we have plotted $p(V)$ for values of $T' = T/T_c$ that are below and above 1 in Fig. 9.15.

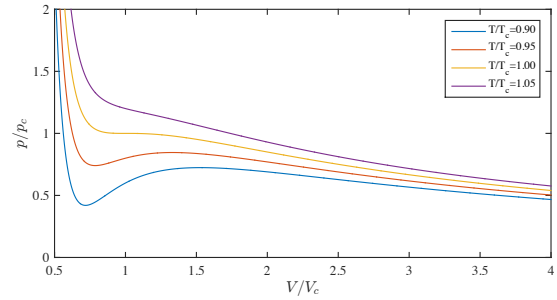


Fig. 9.15 Plot of four isotherms for the van der Waals gas.

Unphysical isotherms. We can now compare these isotherms with the sketch from Fig. 9.10. First, we notice two aspects of the plotted isotherms for the van der Waals gas: (i) When $T < T_c$ there are multiple possible volumes for a given pressure, which means that $V(p)$ is not well defined. This was also the case in Fig. 9.10, where this was interpreted as a two-phase system, where the volume changed as the ratio of gas to liquid was changed. (ii) When $T < T_c$ there are region where $p(V)$ is an increasing function of V. This is clearly unphysical! It would mean that the pressure increases when the volume increases while the temperature is constant. (We will address this behavior using the concept of compressibility in the exercises). This means that there is a region of the $p(V)$ curve that cannot represent an equilibrium situation. How can we determine how this system should behave in equilibrium? We need to find the minimum of the Gibbs free energy of this system.

Gibbs free energy for the van der Waals gas. We find Gibbs free energy from Helmholtz free energy from $G = F + pV$ and then by ensuring that the end results is only a function $G(T, p, N)$ of T, p, and N. We use the expression for Helmholtz free energy from (9.123) to find Gibbs free energy:

$$G = -NkT\left(\ln\frac{(V-Nb)n_Q(T)}{N} + 1\right) - a\frac{N^2}{V} + pV\ , \tag{9.133}$$

We can rewrite Gibbs free energy using the same non-dimensionalization as we did above: $p' = p/p_c$, $V' = V/V_c$, and $T' = T/T_c$, where $p_c = a/(27b^2)$, $V_c = 3Nb$, and $kT_c = 8a/27b$:

$$G = -NkT\left(\ln\left(3\frac{V}{3Nb} - 1\right)\, b\, n_Q(T) + 1\right) - \frac{NaNb}{Vb} + p\frac{3NbV}{3Nb} \tag{9.134}$$

$$\frac{G}{NkT_c} = -\frac{NkT}{NkT_c}\left(\ln\left(3V' - 1\right)\, b\, n_Q(T) + 1\right) - \frac{aN}{3bNkT_cV'} + \frac{3bpV'}{kT_c} \tag{9.135}$$

$$\frac{G}{NkT_c} = -T'\,\ln\left(3V' - 1\right) + f(T') - \frac{aN27b}{3bNV'a} + \frac{3bpV'27b}{8a} \tag{9.136}$$

$$\frac{G}{NkT_c} = -T'\,\ln\left(3V' - 1\right) + f(T') - \frac{9}{8V'} + p'V'\frac{3}{8} \tag{9.137}$$

$$\frac{G}{(3/8)NkT_c} = -\frac{8}{3}T'\,\ln\left(3V' - 1\right) - \frac{3}{V'} + p'V' + f'(T')\,. \tag{9.138}$$

where the function $f'(T')$ only depends on T' and not on p' or V', and where we introduce $g = G/(3/8)NkT_c$ as the dimensionless Gibbs free energy per particle.

Gibbs free energy is supposed to be a function of p, T, and N, and $g = G/N$ should only depend on p and T, but the our result now depends on both T', p', and V'. The challenge is that we found from the plot of $p(V)$ that for $T < T_c$ $V(p)$ is not a unique function. We therefore cannot replace V' by $V'(p')$. Instead, we must use that Gibbs free energy is minimal in equilibrium. Let us therefore plot $g(V'; p', T')$ and find the V' that makes Gibbs free energy minimal — this must be the equilibrium value for V'. When we plot the function, it is more practical to plot g' as a function of $n' = 1/V'$:

$$g'(n'; p', T') = -(8/3)\,T'\ln(3/n' - 1) - 3n' + p'/n' + f'(T')\,. \tag{9.139}$$

Here, we will neglect the last term, since this only depends on T' and will not influence the n' that makes g' minimal.

In Fig. 9.16 we have plotted $g'(n'; p', T')$ near the values $p' = 0.647$ and $T' = 0.9$, which are on the phase coexistence curve. In Fig. 9.16a we have varied the pressure p'. If we start with a pressure below phase coexistence, at $p' = 0.55$, we see that Gibbs free energy has a clear minimum at $n' = 0.315$. This means that in equilibrium, the van der

Waals gas will only contain a single phase with this density — a gas phase.

However, if we increase the pressure, Gibbs free energy will go from a configuration with only one minimum, to a configuration with two (global) minimum values, here occuring for $p' = 0.647$. For pressures just below this value, there is only one minimum in Gibbs free energy, and the whole system is a gas, but when the pressure reaches this value, there are suddenly to minima at $n'_1 = 0.42$ and $n'_2 = 1.65$ so that $g'(n'_1) = g'(n'_2)$. We interpret this to mean that there are now two phases in the system, one phase with a high density — a fluid phase — and one phase with a low density — the gas phase. Notice that at this particular set of p', T' values, all the densities along the stapled line between n'_1 and n'_2 are possible values for the density. When the density of the system is n'_1 it means that all of the system is in the low density gas phase, whereas when the density of the system is n'_2 all of the system is in the high density liquid phase. If half of the system is in the gas phase, and half is in the liquid phase, the density of the system is $n' = 0.5\, n'_1 + 0.5\, n'_2$. This means that Gibbs free energy is the same value, and minimal, for all densities n' in the interval from n'_1 to n'_2, and that the density is determined by how much of the system is in the gas or in the liquid phase. Gibbs free energy therefore has the form $g'_e(n'; p', T,')$

$$g'_e(n'; p', T') \begin{cases} g'(n'; p', T') & n' < n'_1 \\ g'(n_1; p', T') & n'_1 < n' < n'_2 \\ g'(n'; p', T') & n'_2 < n' \end{cases}, \tag{9.140}$$

when p', T' are on the phase coexistence curve.

If we increase the pressure only slightly away from the phase coexistence curve, we see that there is again only one density that gives a minimal Gibbs free energy, and the whole system is in the liquid phase, as illustrated for $p' = 0.75$.

We can present a similar picture if we keep the pressure constant and change the temperature, as illustrated in Fig. 9.16b.

Toward the critical point. What happens if we now increase both the pressure and the temperature of the system, keeping the system on the coexistence curve? We have plotted Gibbs free energy for increaseing p', T' in Fig. 9.17. This illustration gives a very clear picture of what happens as $p' \to 1$ and $T' \to 1$: The densities of the two phases are converging, and Gibbs free energy goes from having two minima to only have a single minima at $p' = 1$, $T' = 1$. We call this point the critical

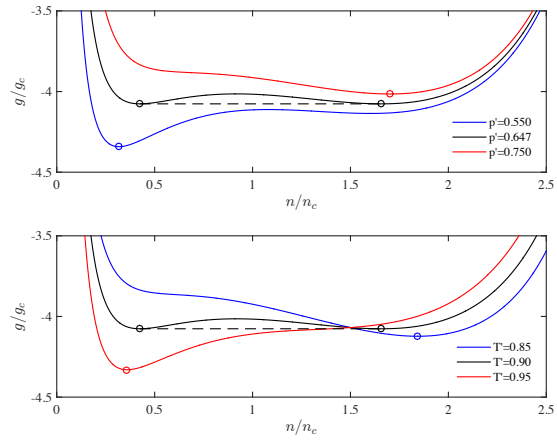


Fig. 9.16 Plot of $g'(n';p',T')$ near the phase coexistence curve.

point. Beyond the critical point, there is only a single phase in the system, and there is no longer any phase coexistence or phase transition. This is the reason for the choice of the numerical prefactors for p_c and T_c, so that the critical point occurs at $(p',T') = (1,1)$.

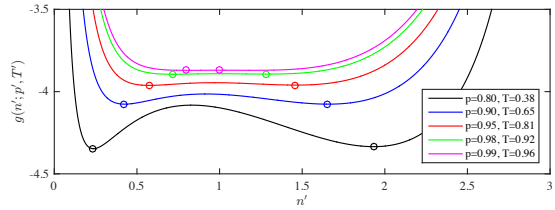


Fig. 9.17 Plot of $g'(n';p',T')$ near the phase coexistence curve.

Behavior of van der Waals gas below T_c. The full behavior of the van der Waals gas system is address in project XX.

$p(T)$ and latent heat for the van der Waals gas.

9.5.4 Surface tension

ams 11: Draft only. Not curriculum in Fys2160.

We introduce $\Delta\mu = \mu_g - \mu_l$ as the difference in chemical potential between the vapor surrounding a liquid droplet and the liquid in bulk, that is for a very large drop (no surface effects).

If $\Delta\mu > 0$ it means that the liquid has lower free energy than the gas/vapor - and liquid droplets may spontaneously form.

However, we must also include the surface free energy, because this tends to increase the energy of the liquid.

For a very small droplet, with a very small radius of curvature, the surface energy will be dominating and the drop can be unstable with respect to the gas/vapor.

Let us study this by addressing Gibbs free energy when a droplet of radius R forms. We introduce

$$\Delta G = G_l - G_g = -(4\pi/3)R^3 n_l \Delta\mu + 4\pi R^2 \gamma \,, \tag{9.141}$$

where n_l is the number density of the liquid (the concentration), and γ is the energy it costs to form a surface - per area.

The liquid drop will grow when $G_l < G_g$. There is a maximum at

$$\frac{d\Delta G}{dR} = 0 = -4\pi R^2 n_l \Delta\mu + 8\pi R\gamma \,, \tag{9.142}$$

which gives

$$R_c = \frac{2\gamma}{n_l \Delta\mu} \,. \tag{9.143}$$

We call this the **critical radius of nucleation**. When the radius is smaller than R the drop will grow smaller - and disappear. When the radius is larger than R the drop will continue to grow larger and larger.

We may (or maybe not?) assume that the gas is a an ideal gas. In that case the chemical potential is

$$\Delta\mu = kT \ln(p/p_{eq}) \,, \tag{9.144}$$

where p_{eq} is the equilibrium vapor pressure of the bult liquid.

We can use realistic values then to estimate R_c for water at 300K and $p = 1.1 p_{eq}$. In addition, we need a value for γ, which we can find from tables, $\gamma = 72\,10^{-3}$ N/m. This gives $R_c \simeq 10^{-8}$ m $= 10$ nm.

9.5.5 Lattice gases

ams 12: lattice gas model: model of general system, mean-field theory for thermodynamics. Gibbs free energy minimization. p(T) curve. Lattice gas automata.

9.5.6 Stability

ams 13: If the gas starts in a local minimum in $g(n)$, it may take some time for the system to reach the global minimum. Metastability and stability, supercooled and superheated water, heterogeneous nucleation, glasses

9.6 Solutions

The systems we have been studying so far mostly only contain one type of particles. The state of the system is described by the number of particles, N, the pressure, p, the temperature, T, and possible also the volume in the case of phase coexistence.

If we introduce another type of particles, we get a *mixture* or a *solution*. There are many types of solutions: a solution may be a mixture of gases; in a liquid we can dissolve gases or solids; in solids we can dissolve gases, liquids or other solids. For example, hydrogen-gas can be dissolved in palladium, or mercury can be dissolved in gold and in many other metals; tin and zink can be dissolved in copper. Solutions are one of the most important classes of systems for chemist, and our understanding of solutions is important for our understanding of many problems of both theoretical and practical importance.

We will here primarily study *dilute solutions*, which are where one of the components are so dilute that they do not affect the energy of the other component. Typical examples are trace gases in air or dilute solutions in water.

Solution definitions:

- the component that is dissolved is called the **solute**
- the substance it is being dissolved in is called the **solvent**
- the resulting product is called the **solution**.

For example, salt water is a *solution*, made from a *solvent*, water, and a *solute*, salt.

While solutions in general may include complex interactions, ideal solutions are simple, since the interactions between the solvent and the element that is dissolved primarily is through the entropy of mixing term. Let us start by looking at mixed ideal gases, often called ideal mixtures.

9.6.1 Ideal mixture

An ideal mixture is a mixture where the species are *non-interacting*:

Ideal mixture: An ideal mixture is a system

- where the species are non-interacting, so that there is no coupling between the energy states of the various species
- with an additional *entropy of mixing* that contributes to the total entropy of the system

Ideal mixtures are good approximations for mixtures in the dilute limit.

Let us develop a thermodynamic theory for the behavior of an ideal mixture of several species. Fig. 9.18 illustrates the mixing of two non-interacting species. First, we consider one of the species in a system of volume V, pressure p and temperature T. Gibbs free energy for species i is then

$$G_i^0 = E_i^0 - TS_i^0 + pV_i^0 = N_i g_i^0 \ , \tag{9.145}$$

where the superscript 0 indicates that this is the values for species i on its own. If we now place several species into the same box (with the same volume $V = V_i^0$ for all the species), then the energies do not change, because the particles do not interact. Similarly, the pV terms do not change even if we add more particles into the same box. But what about the entropy? The total entropy will increase, because there are now more ways to rearrange the various particles. For ideal gases, the entropy for each species would depend on the number of particles, the mass of the particles, the size of the box and the temperature. This internal entropy would not change. In addition, there would be an increase in entropy due to the mixing of the various species. We will therefore assume that the Gibbs free energy for the mixed system will have two contributions: Gibbs free energy for each of the species and a term related to the entropy of mixing:

$$G(p, T, \{N_i\}) = \sum_i N_i g_i^0(p, T, N_i) - TS_{\text{mix}} \ , \tag{9.146}$$

where $g_i^0(p, T, N_i)$ is Gibbs free energy per particle for a system consisting only of type i particles in the same box. (The superscript 0 indicates

that this is the Gibbs free energy without any interactions). What is the entropy of mixing?

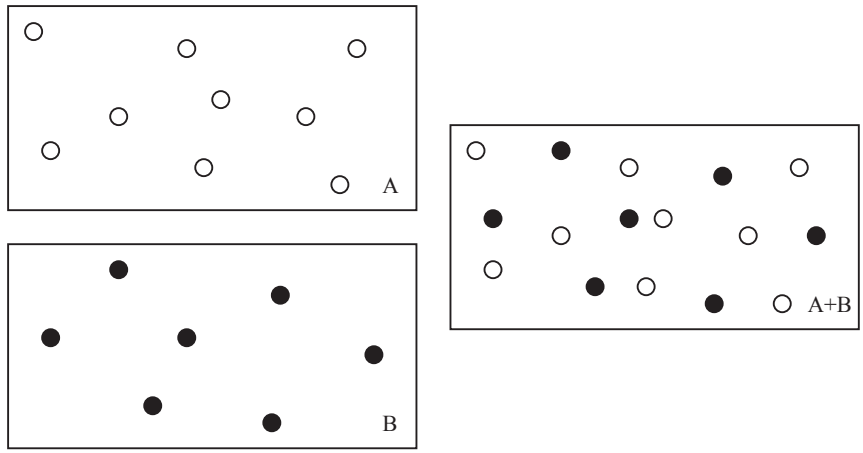


Fig. 9.18 Illustration of a system A of N_A particles of type A and a system B of N_B particles of type B mixed into a common system.

Entropy of mixing. We can find the entropy of mixing from the multiplicity of mixing. For an gas or for a solid of $N = \sum_i N_i$ particles, how many ways can we rearrange the particles? We know that for two species, 1 and 2, the total number of ways that we rearrange N_1 particles of type 1 and N_2 particles of type 2 is

$$\Omega(N_1, N_2) = \begin{pmatrix} N \\ N_1 \end{pmatrix} = \frac{(N_1 + N_2)!}{N_1! N_2!} \tag{9.147}$$

and for k species:

$$\Omega(\{N_i\}) = \frac{N!}{N_1! N_2! \ldots N_k!} \tag{9.148}$$

We find an approximate expression for the mixing entropy, $S_{\text{mix}} = k \ln \Omega$ using Stirling's formula, $\ln y! = y \ln y - y$:

$$S_{\text{mix}}/k = \ln \Omega = N \ln N - N - \sum_{i=1}^{k} (N_i \ln N_i - N_i) \tag{9.149}$$

$$= \underbrace{N}_{=\sum_i N_i} \ln N - N - \sum_{i=1}^{k} N_i \ln N_i + \underbrace{\sum_{i=1}^{k} N_i}_{=N} \tag{9.150}$$

$$= \sum_{i=1}^{k} N_i \ln N - \sum_{i=1}^{k} N_i \ln N_i \tag{9.151}$$

$$= -\sum_{i=1}^{k} N_i \ln \frac{N_i}{N} \, . \tag{9.152}$$

It is common to use the fraction, $x_i = N_i/N$ to describe the relative number of the various components. In terms of fractions, x_i, the entropy of mixing is

$$S_{\text{mix}} = k \sum_i x_i \ln x_i \tag{9.153}$$

Gibbs free energy of an ideal mixture. We insert the entropy of mixing into Gibbs free energy, finding

$$G(p, T, \{N_i\}) = \sum_{i=1}^{k} N_i g_i^0(p, T, N_i) + kT \sum_{i=1}^{k} N_i \ln \frac{N_i}{N} \, . \tag{9.154}$$

We use the term $x_i = N_i/N$ for the mole fraction or the fraction of species i:

$$G(p, T, \{N_i\}) = \sum_{i=1}^{k} G_i^0(p, T, N_i) + NkT \sum_{i=1}^{k} x_i \ln x_i \, . \tag{9.155}$$

(Notice that $\sum_i x_i = \sum_i N_i/N = (\sum_i N_i)/N = 1$).

Chemical potential of individual species in an ideal mixture. The chemical potential of component i can then be found from the partial derivative of Gibbs free energy:

$$\mu_i = \left(\frac{\partial G}{\partial N_i} \right)_{p,T,\{N_j\}} \tag{9.156}$$

where we keep N_j constant for all $j \neq i$. (Notice that $N = \sum_i N_i$ and therefore N is not a constant in this derivative).

$$\mu_i = \left(\frac{\partial G}{\partial N_i}\right)_{p,T,\{N_j\}} \tag{9.157}$$

$$= \left(\frac{\partial G_i^0}{\partial N_i}\right)_{p,T,\{N_j\}} + \left(\frac{\partial}{\partial N_i}\sum_{r=1}^{k} N_r \ln x_r\right)_{p,T,\{N_j\}} \tag{9.158}$$

$$= \mu_i^0(p,T) + kT\ln x_i + \sum_{r=1}^{k} N_r \frac{1}{x_r}\frac{\partial x_r}{\partial N_i} \tag{9.159}$$

$$= \mu_i^0(p,T) + kT\ln x_i + N\frac{\partial x_r}{\partial N_i}\underbrace{\sum_{r=1}^{k} x_r}_{=1} \tag{9.160}$$

$$= \mu_i^0(p,T) + kT\ln x_i \ . \tag{9.161}$$

where $\mu_i^0(p,T) = g_i^0(p,T)$ is the chemical potential of the system when only species i is present.

The chemical potential for component i is therefore only dependent on $x_i = N_i/N = n_i/n = m_i$, and not on the other x_j values. The only dependence on the other atoms is through the N in the denominator of $x_i = N_i/N$. We can check the consistency of the expression by noticing that for $x_1 = 1$ we only have the pure substance, and $\mu_1 = \mu_1^0$, as it should.

The ideal mixture approximation is versatile. This simple expression for the chemical potential is an approximation to case of a general mixtures. We call this approximation the **ideal mixture** approximation. This approximation is versatile. It is clearly applicable to ideal gases, where there are no interactions between the particles, but it is also a reasonable approximation to many other systems. It is well demonstrated the the ideal mixture approximation is:

- almost exact for mixtures of isotopes
- good for mixtures of similar molecules
- very good for the *solvent* in dilute solutions
- reasonably good for use in chemical equilibriums

Let us now see how we can address an ideal mixture of ideal gases.

Solution of ideal gases. Mixtures of ideal gases are ideal mixtures, since there are no interactions between the particles in the gas. For a multi-component gas, with N_i particles (atoms/molecules) of species i, the total number of particles is $N = \sum_i N_i$. The pressure of the gas is

$$p = \frac{NkT}{V} = \sum_{i=1}^{k} \frac{N_i kT}{V} = \sum_{i=1}^{k} p_i \ , \tag{9.162}$$

where $p_i = N_i kT/V$ is called the *partial pressure* of component i. Since the particles are non-interacting, you can think of a multi-species ideal gas as a set of k gases, each in a volume V, and each with a pressure p_i. The total pressure is the sum of the (partial) pressures of each of the gases, $p = \sum_i p_i$.

For an ideal gas the fraction x_i can be expressed in terms of the partial pressure: $x_i = N_i/N = N_i kT/NkT = p_i V/pV = p_i/p$, which gives the following expression for the chemical potential of species i in a mixture of ideal gases:

$$\mu_i(p,T) = \mu_i^0(p,T) + kT \ln x_i = \mu_i^0(p,T) + kT \ln p_i/p \ . \tag{9.163}$$

Where $\mu_i^0(p,T) = kT \ln n_i/n_Q = kT \ln p_i/(n_Q(T)\, kT)$ for an ideal gas.

9.6.2 Binary mixtures

The expressions become simpler when there are only two species, 1 and 2, present. In this case, we can describe the state of the solution only by $x = x_2 = N_2/N$ and the total number of particles, N. We see that $x_1 + x_2 = 1$ gives $x_1 = 1 - x_2 = 1 - x$. In this case we can simplify the expression for Gibbs free energy:

$$G(p,T,N_1,N_2) = N_1 g_1^0(p,T) + N_2 g_2^0(p,T) + NkT \sum_i x_i \ln x_i \tag{9.164}$$

$$= N(1-x) g_1^0(p,T) + Nx g_2^0(p,T) + NkT \left((1-x)\ln(1-x) + x \ln x\right) \ , \tag{9.165}$$

where the last term is due to the entropy of mixing: $S_{\text{mix}} = k\left((1-x)\ln(1-x) + x \ln x\right)$.

Fig. 9.19 shows plots of Gibbs free energy (and the mixing entropy) as a function of x. We illustrate two cases: The case of no mixing and the case of mixing. The case of no mixing corresponds to the case where all of species 1 are on the left side and all of species 2 are on the right hand side of the system. This corresponds to a Gibbs free energy without any additional entropy of mixing, which is simply the linear combination of the Gibbs free energy for each of the two species: $G = (1-x)G_1^0 + xG_2^0$

(illustrated by the red line). The total Gibbs free energy is clearly smaller when the system is mixed due to the effect of the entropy of mixing.

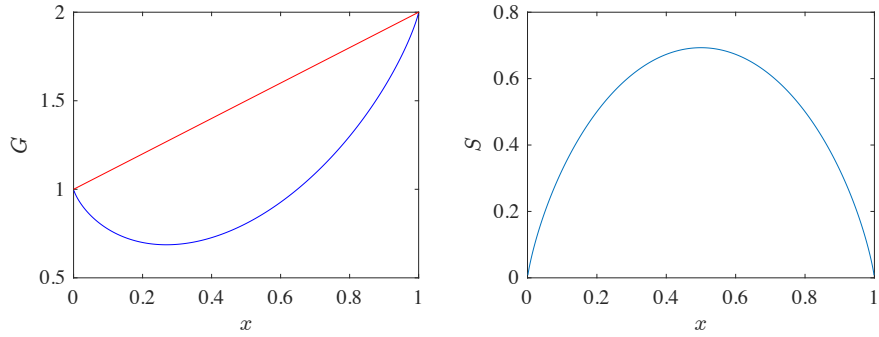


Fig. 9.19 Plots of the Gibbs free energy for a mixed (blue) and an unmixed (red) system.

In non-ideal mixtures interactions between particles of different species will change the internal energies as well, opening also for unmixed states in thermal equilibrium as discussed in Sect. XX

9.6.3 Dilute solutions

The results can be further simplified in the case of a *dilute* solution — a solution where there is few of the solute (2) compared to the solvent (1): $N_2 \ll N_1$. In this case $x = N_2/N \ll 1$, and we see that $x = N_2/N = N_2/(N_1 + N_2) \simeq N_2/N_1$.

Chemical potential for the solvent in a dilute solution. The chemical potential for the solvent (1) in an ideal mixture is

$$\mu_1(p,T) = \mu_1^0(p,T) + kT \ln x_1 = \mu_1^0(p,T) + kT \ln(1-x) \ , \qquad (9.166)$$

when $x = x_2 \ll 1$, we can approximate $\ln(1-x)$ by the first term in its Taylor expansion, $\ln(1-x) \simeq -x$, getting:

$$\mu_1(p,T) = \mu_1^0(p,T) - kTx \ , \qquad (9.167)$$

where $\mu_1^0(p,T)$ was the chemical potential of the solvent before the solute was added. Notice that this means that the chemical potential of the solvent therefore becomes lower after the solute is added.

Chemical potential for the solute in a dilute solution.

ams 14: May add derivation of this?

9.6.4 Example: Vapor pressure above a dilute solution

We have already found an equation for the vapor pressure for a liquid. What happens if we dissolve a substance in the liquid? How will the vapor pressure change?

We can apple Le Chaatlier's principle to gain a quick understanding of what will happen. According to the principle, the system will respond in such a way as to oppose the interaction — the system will therefore respond by increasing the amount of liquid in order to make the solution thinner, at the expense of the vapor. That is, the system will respond by lowering the vapor pressure. But by how much?

The system is illustrated in Fig. 9.20. A closed container at constant pressure, p, and temperature, T, contains N atoms/molecules of a solvent, N_g are in the gas phase and N_l are in the liquid phase. In addition, there are N' dissolved atoms in the liquid. We assume that we can neglect the amount of dissolved atoms in the gas — its vapor pressure is negligible.

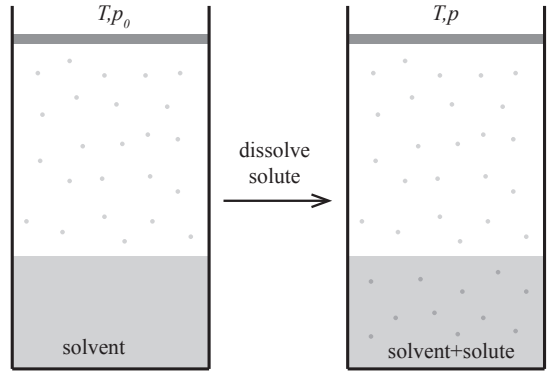


Fig. 9.20 Illustration of a system with a solvent at equilibrium with its vapor at T, p_0. Then a solute is dissolved in the solvent, and the system reaches equilibrium at a new equilibrium vapor pressure p.

The behavior of the system is determined by Gibbs free energy: In equilibrium Gibbs free energy is minimal. Gibbs free energy for the system is

$$G = N_g g_g(T,p) + N_l g_l(T,p) + N' g'(T,p) + N' kT \left(\ln N'/N_l - 1\right) \ , \tag{9.168}$$

where we have used the approximation $N'/(N' + N_l) \simeq N'/N_l \ll 1$. In equilibrium, the chemical potential of the two phases of the solvent must be equal, that is, $\mu_l = \mu_g$. This is true for p_0, T before the substance is dissolved in the solvent, and for p, T after the substance was added. We want to find the change in vapor pressure, that is, $p - p_0 = \Delta p$. We find

the chemical potential by derivation of Gibbs free energy:

$$\mu_g = \left(\frac{\partial G}{\partial N_g}\right)_{T,p,N_l,N'} = g_g(T,p) \ , \tag{9.169}$$

$$\mu_l = \left(\frac{\partial G}{\partial N_l}\right)_{T,p,N_g,N'} = g_v(T,p) - kT\frac{N'}{N_l} \ . \tag{9.170}$$

Here, we now insert $p = p_0 + \Delta p$, and we approximate $g_v(T, p_0 + \Delta p)$ with its first order Taylor expansion, and similarly for $g_l(T, p_0 + \Delta p)$:

$$\mu_g(p_0 + \Delta p, T) = g_g(p_0 + \Delta p, T) = g_g(p_0, T) + \left(\frac{\partial g_g}{\partial p}\right)_T \Delta p \ , \tag{9.171}$$

and similarly

$$\mu_l(p_0+\Delta p, T) = g_l(p_0+\Delta p, T) - kT\frac{N'}{N_l} = g_l(p_0, T) + \left(\frac{\partial g_l}{\partial p}\right)_T \Delta p - kT\frac{N'}{N_l} \ , \tag{9.172}$$

We introduce $x = N'/N_l$, where $x \ll 1$. We notice that

$$\left(\frac{\partial g_l}{\partial p}\right)_T = v_l \ , \ \left(\frac{\partial g_g}{\partial p}\right)_T = v_g \ , \tag{9.173}$$

are the volume of a single atom of gas or liquid.

In equilibrium the chemical potentials must be equal:

$$\mu_l(p,T) = g_l(p_0,T) + v_l\Delta p - kTx = \mu_g(p,T) = g_g(p_0,T) + v_g\Delta p \ . \tag{9.174}$$

$$v_l\Delta p - kTx = v_g\Delta p \ \Rightarrow \ (v_g - v_l)\,\Delta p = -kTx \ , \tag{9.175}$$

which can be simplified further since the volume per particle in the liquid is much smaller than in the gas, $v_l \ll v_g$, and therefore $v_g - v_l \simeq v_g$, and for an ideal gas we know that $v_g = kT/p_0$. We have therefore obtained *Raoult's law*:

$$v_g\Delta p = \frac{\Delta p}{p_0} = x = \frac{N'}{N_l} \ . \tag{9.176}$$

Notice that the change in the vapor pressure is proportional to the amount of dissolved substance. But this change in vapor pressure is independent of what type of substance is dissolved — it is only the number of dissolved particles that matter. We may use this to determine the atomic weight of a substance. If we have a solvent with a known molecular weight, such as for water, we may weigh a given amount, N_l, of

the water (solvent). Then we dissolve a given mass m of a substrance with an unknown molecular weight in the water and measure the reduction in vapor repssure, Δp, and find the molecular weight from

$$M = \frac{m}{N'} = -\frac{mp_0}{N_l \Delta p} \, . \tag{9.177}$$

9.6.5 Example: Osmotic pressure

We can use our newfound theory for dilute solutions to find the osmotic pressure of a solution. Biological membranes ensure that the chemistry inside and outside of the membrane can be different – thus ensuring the machinery of life inside a cell, while living in a potentially hostile environment. Membranes can be made to allow the transport of a solvent such as water while restricting the transport of a solute, such as salt. Fig. 9.21 illustrates a membrane with pure solvent on one side and a solution on the other side. The membrane allows the solvent to move from one side to another, but not the solute.

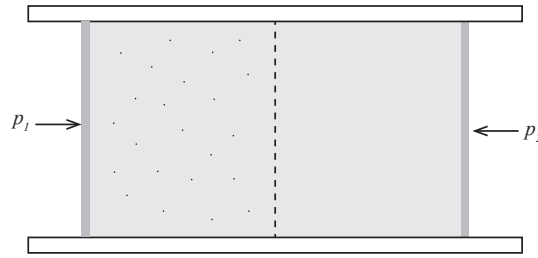


Fig. 9.21 Illustration of a solution (left) and a pure solvent (right) separated by a membrane that is permeable to the solvent, but not to the solute.

If the system starts in this situation, with a pure solvent on one side and a solution on the other, the system will clearly not be in thermodynamic equilibrium, because the chemical potentials are not the same on each side of the membrane. The chemical potential in the solvent is $\mu_2(p,T) = \mu_s^0(p,T)$, while the chemical potential in the (dilute) solution is approximately $\mu_1(p,T) = \mu_s(p,T) = \mu_s^0(p,T) - kT \ln x$. Transport will therefore occur from the side with the solvent to the side with the solution. However, we can stop the transport by increasing the pressure in the solution until the chemical potentials are equal. This occurs when the pressure in the solution is p_1 and the pressure in the solvent is p_2. The chemical potentials are then

$$\mu_2(p_2,T) = \mu_s^0(p_2,T) = \mu_1(p_1,T) = \mu_s^0(p_1,T) - kTx \, . \tag{9.178}$$

If the pressure difference $\Delta p = p_2 - p_1$ is small compared with p_1, we can approximate $\mu_s^0(p_2, T)$ with the Taylor expansion around $p = p_1$:

$$\mu_s^0(p_2) \simeq \mu_s^0(p_1) + (p_2 - p_1) \left(\frac{\partial \mu_s^0}{\partial p} \right)_{T,N,x} . \tag{9.179}$$

We can find the derivative of the chemical potential from the derivative of Gibbs free energy, because $\mu_s^0 = G_s^0/N$, and therefore

$$\left(\frac{\partial \mu_s^0}{\partial p} \right)_{T,N,x} = \frac{1}{N} \left(\frac{\partial G_s^0}{\partial p} \right)_{T,N,x} = \frac{V}{N} = v_s^0 , \tag{9.180}$$

that is, the volume per particle in the pure solute. This gives

$$\mu_s^0(p_2) \simeq \mu_s^0(p_1) + \frac{V}{N} \Delta p \tag{9.181}$$

which we can insert into (9.178), getting

$$\mu_s^0(p_2) \simeq \mu_s^0(p_1) + \frac{V}{N} \Delta p = \mu_s^0(p_1, T) - kTx , \tag{9.182}$$

and therefore

$$\Delta p = -\frac{NkTn}{V} = -\frac{NkTN_B}{VN} = -\frac{N_B kT}{V} , \tag{9.183}$$

where B represent the solute. This pressure difference is called the **osmotic pressure**. And the equation describing the osmotic pressure of a dilute solution is called van't Hoff's formula. This means that the pressure difference needed corresponds to the pressure in an ideal gas with the concentration N_B/V of the solute.

This pressure can be quite significant for realistic, biological systems. For example, for a cell there are about 200 water molecules for every molecule of some other type – this is therefore a dilute solution. Since the atomic weight of water is 18u, the number of water molecules in a liter of water (1 kg) is

$$N_s = \frac{1\text{kg}}{18 \cdot 1.66 \cdot 10^{-27}\text{kg}} , \tag{9.184}$$

and the pressure in a gas of $N_s/200$ particles in 1 liter is

$$p = \frac{NkT}{V} = \frac{(1/200)N_s kT}{0.001\mathrm{m}^3} = \frac{1.38 \cdot 10^{-23}\mathrm{J/K} \cdot 300\mathrm{K}}{200 \cdot 18 \cdot 1.66 \cdot 10^{-27} \cdot 0.001\mathrm{m}^3} = 6.9\mathrm{N/m}^2 \,. \tag{9.185}$$

The pressure difference between the inside and the outside of a cell in equilibrium with pure water is therefore approximately 7 atm – which is high!

If the pressure difference is smaller than the osmotic pressure, molecules from the pure solvent will flow through the membrane and into the solution. However, if the pressure difference is larger than the osmotic pressure, the pressure will drive the solvent from the solution and into the pure solvent. This process is, not surprisingly, called **reverse osmosis** and currently represents one of the best techniques for desalination (removing salt from seawater) to make freshwater. Making good membranes for this process is an area of great current technological and industrial interest.

9.6.6 Example: Change in boiling point above a dilute solution

9.6.7 Example: Henry's law and the solubility of gases

9.6.8 Non-ideal mixtures

9.6.9 Surface growth processes

9.7 Chemical reactions

We now have what we need to address chemical reactions - how some species are transformed into other species without changing the total number of atoms. (The total number of particles may change).

Let us analyze chemical reactions from an example: The formation of water from hydrogen and oxygen.

$$\mathrm{H^+ + OH^- \leftrightarrow H_2O} \,, \tag{9.186}$$

It is common to describe chemical reactions by it stochiometric coefficients, ν_j:

$$\nu_1 A_1 + \nu_2 A_2 + \ldots + \nu_k A_k = 0 \,, \tag{9.187}$$

where A_j describes a chemical species such as O or O_2 or H_2O. The reaction above would then be described as

$$+1\,\mathrm{H}^+ + 1\,\mathrm{OH}^- - 1\,\mathrm{H_2O} = 0\ , \tag{9.188}$$

that is

$$\nu_1 = 1\ ,\ A_1 = \mathrm{H}^+\ ,\ \nu_2 = 1\ ,\ A_2\mathrm{OH}^-\ ,\ \nu_3 = -1\ ,\ A_3 = \mathrm{H_2O}\ . \tag{9.189}$$

What characterizes the equilibrium of a system with hydrogen, oxygen and water present? Typically, we will consider equilibrium at constant pressure and temperature, $dp = dT = 0$, which corresponds to typical laboratory conditions. In this case, thermodynamic equilibrium occurs when Gibbs free energy is minimal, that is, when $dG = 0$. In general, we have that

$$dG = -SdT + Vdp + \sum_{i=1}^{c} \mu_i dN_i = 0\ , \tag{9.190}$$

which for $dT = dp = 0$ simplifies to

$$dG = \sum_{i=1}^{c} \mu_i dN_i = 0\ . \tag{9.191}$$

We have equilibrium when one more chemical reaction – one transformation from the left to the right or opposite – does not change Gibbs free energy. For the chemical reaction at hand, the changes in N_i are not arbitrary – they are linked by the coefficients ν_i of the reaction. For a single reaction, $dN_j = \nu_j$. We therefore have that in equilibrium:

$$dG = \sum_j \mu_j dN_j = \sum_j \mu_j \nu_j = 0\ . \tag{9.192}$$

We can determine this condition for any reaction if we only know the chemical potential μ_j for each of the species. (Notice that this relation is derived for constant p and T, but it also applies to the equilibrium of reactions at constant T and V).

Gibbs-Duhem relation. Notice that it is usual to call the relation

$$dG = -SdT + Vdp + \sum_j \mu_j dN_j = 0\ , \tag{9.193}$$

(which it is in equilibrium) the **Gibbs-Duhem relation** when rewritten as

$$\sum_j \mu_j dN_j = -SdT + Vdp\ . \tag{9.194}$$

9.7.1 Chemical equilibrium for Ideal Gas systems

For a multi-component ideal mixture, such as a system of ideal gases or any other system well described as an ideal mixture, the chemical potential of component i is

$$\mu_i = \mu_i^0 + kT \ln x_i \ , \tag{9.195}$$

where $x_i = N_i/N$ and μ_i^0 is the chemical potential of the pure gas – without any other gases present. The equilibrium condition from (9.192) then gives

$$\sum_i \mu_i \nu_i = \sum_i \mu_i^0 \nu_i + \sum_i kT \ln x_i = 0 \ , \tag{9.196}$$

which gives

$$kT \sum_i \ln x_i = -\sum_i \mu_i^0 \nu_i \ . \tag{9.197}$$

If we now multiply by Avogadro's number, N_A, in each side we get

$$RT \sum_i \nu_i \ln x_i = -N_A \sum_i \mu_i^0 \nu_i = -\Delta G^0 \ , \tag{9.198}$$

where ΔG^0 is the change in Gibbs free energy for the reaction – which is the hypothetical change in G when one mole of H_2 reacts with one mole of OH^-, forming one mole of water at 1 bar. This value of ΔG^0 you can usually find in reference tables. If we take the exponential on both sides, we find

$$\prod_i x_i^{\nu_i} = e^{-\frac{\Delta G^0}{RT}} = K(T) \ , \tag{9.199}$$

where $K(T)$ is only a function of T – and you can often find the value of ΔG^0 in chemical tables. The constant K is usually knows as the equilibrium constant – and we see that we can acually calculate the value of $K(T)$ if we know the chemical potentials μ_i^0 for all the components in the reaction.

For ideal gases, the relation can be rewritten in terms of the concentrations, $n_i = N_i/V$, instead:

$$\begin{aligned} \prod_i x_i^{\nu_i} &= \prod_i \left(\frac{N_i}{N}\right)^{\nu_i} = \prod_i \left(\frac{N_i kT}{NkT}\right)^{\nu_i} \\ &= \prod_i \left(\frac{N_i kT}{pV}\right)^{\nu_i} = \prod_i \left(\frac{N_i}{V}\right)^{\nu_i} \left(\frac{kT}{pV}\right)^{\nu_i} = K(T) \ , \end{aligned} \tag{9.200}$$

which gives

$$\prod_i \left(\frac{N_i}{V}\right)^{\nu_i} = \frac{K(T)}{\prod_i \left(\frac{kT}{pV}\right)^{\nu_i}} = K_1(p,T) \; . \tag{9.201}$$

These equations are called the **law of mass action**

Notice that when we calculate $K(T)$ we must be very careful to choose a consistent value for all the energies - we need to select the zero level in the same way for all the particles. One way to define energies that are internally consistent can be explained through a disassociation reaction, where a molecule A_2 disassociates into 2A. In this case, we should choose the zero level of each composite particle (A_2) to be the energy of the disassociated particles (A) at rest. That is, if the binding energy for A_2 is ϵ (this is the energy needed to place the two consituents of A_2 infinitely far away from each other), we place the ground state of the composite particle (A_2) at $-\epsilon$.

9.7.2 Example: Disassociation of hydrogen

We start with the reaction

$$H_2 \leftrightarrow 2H \; , \tag{9.202}$$

which also can be written as

$$H_2 - 2H = 0 \; , \tag{9.203}$$

The law of mass action gives:

$$\prod_j n_j^{\nu_j} = K(T) \; , \tag{9.204}$$

where $j = 1$ corresponds to H_2, so that $\nu_1 = 1$, and $j = 2$ corresponds to H, so that $\nu_2 = -2$. It is usual to write

$$n_{H_2} as \, [H_2] \tag{9.205}$$

The law of mass action is therefore

$$[H_2]\,[H]^{-2} = \frac{[H_2]}{[H]^2} = K(T) \; , \tag{9.206}$$

This means that

$$\frac{[\mathrm{H}_2]}{[\mathrm{H}]} = \frac{1}{[\mathrm{H}_2]^{1/2}\, K^{1/2}} \,, \tag{9.207}$$

so that the relative concentration of hydrogen is inversely proportional to the concentration of H_2.

9.7.3 Example: pH and the Ionization of water

Water goes through the process

$$\mathrm{H_2O} \leftrightarrow \mathrm{H^+} + \mathrm{OH^-} \,, \tag{9.208}$$

when in liquid form. This process is called the disassociation of water.

The law of mass action gives:

$$\left[\mathrm{H^+}\right] \left[\mathrm{OH^-}\right] = [\mathrm{H_2O}]\, K(T) \,. \tag{9.209}$$

In pure water each of the concentrations are

$$\left[\mathrm{H^+}\right] = \left[\mathrm{OH^-}\right] = 10^{-7} \mathrm{mol\,l^{-1}} \,. \tag{9.210}$$

We can change this concentration by introducing a proton donor. This increases the number of $\mathrm{H^+}$ ions and decreases the number of $\mathrm{OH^-}$ ions to ensure the product of the concentrations is constant.

It is common to introduce the pH through

$$\mathrm{pH} = -\log_{10}\left[\mathrm{H^+}\right] \,. \tag{9.211}$$

The pH of pure water is therefore 7.

Strong acids have low pH values. An apple has pH around 3.

9.7.4 Example: Kinetics, reaction rates, and catalytic processes

What if we study the process

$$\mathrm{A} + \mathrm{B} \leftrightarrow \mathrm{AB} \,, \tag{9.212}$$

Then the rate at which the concentrations changes are related by

$$\frac{dn_{AB}}{dt} = C n_A n_B - D n_{AB} \,, \tag{9.213}$$

where C describes how AB is formed in collisions and D is the reverse process.

In equilibrium the concentrations does not change, and

$$Cn_An_B = Dn_{AB} \ , \tag{9.214}$$

which also are related by the law of mass action:

$$\frac{n_An_B}{n_{AB}} = \frac{D}{C} = K(T) \ . \tag{9.215}$$

Now, what if AB is not formed by collisions between A and B, but in a two step process involving a catalyst E:

$$A + E \leftrightarrow AE \ , \ AE + B \leftrightarrow AB + E \ , \tag{9.216}$$

where E is returned to its original state after the reaction.

What is the point of E? It may increase the rates significantly?

How? The rates are not only determined by the energy of the final configuration, but also by an energy barrier. The rate is determined by the height of the energy barrier and the temperature (Arrhenius processes). However, by introducing E we may lower the energy barrier in each step, increasing the rate of the reaction.

Now, if the process is rapid so that E is short lived, then AE does not form a significant quantity of A. Then the ratio n_An_B/n_{AB} is the same as we found above - it is given by the law of mass action. The route taken by the reaction is not important - the end result is the same.

In equilibrium, the direct and the inverse reaction rates must be the same - what we assumed above - is called the *principle of detailed balance*.

9.7.5 Example: Dissolved Oxygen

Henry's law.

9.7.6 Example: Charge distribution of the Hemoglobin Molecule

Longer analytical and numerical example.

9.8 Old Stuff

9.8.1 Intensive, Extensive Variables, and G

We take two identical systems - for example two system of ideal gas or two systems of the einstein crystal - and put them together, forming a system with double the number of particles.

Some variable will change and some will not in this process.

We call the variable that dot not change **intensive variables**. They are p, T, μ.

Other variables are linear in N: They double when the system doubles. We call these variables **extensive variables**. Examples are U, S, V, N, F, G.

If G is directly proportional to N, we can write $G = Ng(p, T)$, where $g = G/N$. What is g? It is simply

$$\mu = \left(\frac{\partial G}{\partial N}\right)_{T,p} = g \ , \tag{9.217}$$

and

$$G(T, p, N) = N\mu(T, p) \ . \tag{9.218}$$

9.8.2 Multi-component systems

How can we generalize all our results to multi-component systems?

Originally, we introduced the thermodynamic identity from the micro-canonical ensemble, and we found that we could write the entropy, S as a function of U and V, which gave us the differentiale

$$dS = \frac{1}{T}dU - \frac{p}{T}dV \ , \tag{9.219}$$

then we extended to a system with N particles, getting

$$dS = \frac{1}{T}dU - \frac{p}{T}dV - \frac{\mu}{T}dN \ . \tag{9.220}$$

This can now directly be extended to a system with $j = 1, \ldots, k$ different species by introducing a term related to the diffusive equilbrium for each of the species, resulting in (review the original introduction if you are in doubt):

$$dS = \frac{1}{T}dU - \frac{p}{T}dV - \sum_j \frac{\mu_j}{T}dN_j \ . \tag{9.221}$$

Similarly, we can generalize the chemical potential we found from Helmholtz free energy (which is the same as the one we found from the entropy):

$$\mu_j = \left(\frac{\partial F}{\partial N_j}\right)_{T,V,\{N_j\}} . \tag{9.222}$$

And similarly for Gibbs free energy:

$$G(T,p,\{N_j\}) = \sum_j N_j \mu_j . \tag{9.223}$$

The thermodynamic identity then becomes

$$TdS = dU + pdV - \sum_j \mu_j dN_j , \tag{9.224}$$

and the differential for G becomes

$$dG = -SdT + Vdp + \sum_j \mu_j dN_j . \tag{9.225}$$

We will use this to address reactions between different chemical components - by introducing the fundamental laws of chemistry.

9.8.3 Matlab MD script

```
% LJ MD calculation
clear all; clf;
L = 10; % Number of atoms = L^2
N = L*L;
rho = 0.8; % reduced density
Temp = 0.1; % reduced temperature
nsteps = 10000;
dt = 0.02;
printfreq = 1;
% Initial coordinates on cubic grid
r = zeros(N,2);
v = zeros(N,2);
[x y] = meshgrid((0:L-1),(0:L-1));
r(:,1) = x(:); r(:,2) = y(:);
% Rescale to wanted rho
L = L*(1.0/rho^2); r = r*(1.0/rho^2);
% Initialize with wanted T
v = sqrt(Temp)*randn(N,2);
% Internal variables
dt2 = dt*dt;
force = zeros(N,2);
```

```
epot = zeros(N,1); ekin = epot; t = epot;
% Integrate motion
for i = 1:nsteps
    % Velocity-Verlet - part 1
    r = r + v*dt + 0.5*force*dt2;
    r = r + (r<0.0)*L - (r>L)*L; % Periodic
    v = v + 0.5*force*dt;
    % Find forces/accelerations
    [force,energy] = LJforce(r,L);
    % Velocity-Verlet - part 2
    v = v + 0.5*force*dt;
    % Store energies
    epot(i) = energy/N;
    ekin(i) = 0.5*sum(sum(v.*v))/N;
    t(i) = i*dt;
    % Plot
    if (mod(i,printfreq)==0)
        tit = sprintf('Timesteps = %d',i);
        plot(r(:,1),r(:,2),'o');
        title(tit);
        axis equal, axis([0 L 0 L])
        drawnow
    end
end
%%
figure
ii = (1:i-1);
plot(t(ii),ekin(ii),'-r',t(ii),ekin(ii)+epot(ii),':k');
xlabel('t')
ylabel('E');
legend('K','E_{TOT}');
```

```
function [force,energy] = LJforce(r,L);
% Calculate the force on each particle and the
% potential energy of a Lennard-Jones system
% with potential energy for each pair:
% V(r) = 4*U0*((sigma/dr)^12-(sigma/dr)^6))
% Variables in: r(1:N,3) coordinates
%               L system size
% Notice: lengths measured in units of sigma
%         energies measured in units of U0
s = size(r);
npart = s(1); dim = s(2);
L2 = L*0.5; mL2 = -L2;
ff = zeros(npart,dim); % forces
en = 0;
for i = 1:npart
    ri = r(i,:);
    for j = i+1:npart
        rj = r(j,:);
        rij = (ri-rj);
        rij = rij + (rij<mL2)*L - (rij>L2)*L;
        r2 = sum(rij.*rij);
```

```
        ir2 = 1.0/r2;
        ir6 = ir2*ir2*ir2;
        ir12 = ir6*ir6;
        % Calculate force from i-j interaction
        fij = (2*ir12-ir6)*rij*ir2;
        ff(i,:) = ff(i,:) + fij;
        ff(j,:) = ff(j,:) - fij;
        % Calculate energy from i-j interaction
        enij = (ir12-ir6);
        en = en + enij;
    end
end
en = en*4;
energy = en;
ff = ff*24;
force = ff;
return
```

9.9 Summary

9.10 Exercises

Exercise 9.1: Maxwell relations

a) Find the three Maxwell relations for Gibbs free energy.

Exercise 9.2: Compressibility of the van der Waals gas

ams 15: Show problems with vdW gas and compressibility when $T < T_c$.

9.11 Projects

Fermi and Bose gases

10

10.1 Gibbs factor and Gibbs sum

We have already developed a microscopic theory for systems where the temperature, volume and number of particles are constant (T, V, N). Can we develop a similar theory for systems where particles can diffuse, so that N is no longer constant, but where the chemical potential is constant (in equilibrium), that is, for systems described by (T, V, μ). We call such systems grand canonical systems. Such as system is illustrated in Fig. 10.1. The system consists of two parts, the system S and a large reservoir R. The whole system, consisting of parts S and R, is isolated, but the system S may exchange both energy and particles with the reservoir.

This system is similar to the system we introduced when we discussed the canonical system, but we have now opened for particle transport in addition. For the canonical system we found that the probability for the system to be in state i with energy ϵ_i could be expressed as

$$P(i) = \frac{1}{Z} e^{-\epsilon_i/kT} \ , \ Z(T, V, N) = \sum_i e^{-\epsilon_i/kT} \ . \tag{10.1}$$

where the partition function $Z(T, V, N)$ was introduced as a normalization constant. We could use this theoretical prediction for the probabilities of states i to calculate average values for the energy, and we were able to relate the partition function to Helmholtz free energy.

How can we generalize this microscopic theory to a system S in contact with a reservoir R, but where the contact allows both exchange of energy and exchange of particles? Since the whole system in Fig. 10.1 is isolated, the total energy and the total number of particles are conserved, that is

$$E = E_S + E_R = \text{const.}\ ,\quad N = N_S + N_R = \text{const.}\ , \tag{10.2}$$

are both constants. Let us now at a particular case where system S has N_S particles and is in a state i with energy ϵ_i. (Notice that the states and the energies of the states depend on the number of particles in the system). How can we find the probability for this state?

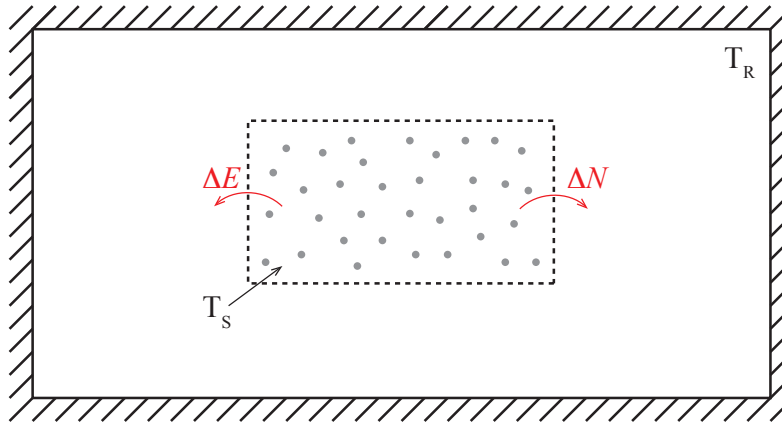


Fig. 10.1 Illustration of a system S in contact with a reservoir R, where the system is allowed to exchange both energy and particles with the reservoir.

We know that in the isolated system, consisting of both the system S and the reservoir R, the probability of every microstate is the same — this is the fundamental assumption of statistical mechanics. We can therefore find the probability of a particular macrostate in this system by counting the number of microstates in the macrostate. We call Ω_S the multiplicity of system S, Ω_R, the multiplicity of the reservoir. The probability for system S to be in a state i is therefore

$$P(N_S, i) = \frac{\Omega_S\,\Omega_R}{\sum \Omega_S\,\Omega_R}\ , \tag{10.3}$$

where the sum is the total number of microstates in the whole system (R+S). If we now address a situation where the number of particles in system S is N_S and the system is in a specific state i, there is only

one way this can be achived, and the multiplicity Ω_S of system S is 1. However, the energy of the reservoir is then $E_R = E - \epsilon_i$, and the number of particles is $N_R = N - N_S$, and there are many states with this combination.

How can we estimate the multiplicity for the reservoir? We use the same trick as we did for the canonical system. First, we look at the logarithm of the multiplicity, $S_R/k = \ln \Omega_R$, and then we realize that the deviations N_S and ϵ_i are small compared with the energy and number of particles in the reservoir, therefore we can Taylor expand the entropy of the reservoir around the value for N, E:

$$k \ln \Omega_R = S_R(N - N_S, E - \epsilon_i) \simeq S_R(N, E) - N_S \left(\frac{\partial S_R}{\partial N_R}\right)_{E,V} - \epsilon_i \left(\frac{\partial S_R}{\partial E}\right)_{E,N} . \tag{10.4}$$

We recognize the partial derivatives from Chap. 6:

$$\left(\frac{\partial S_R}{\partial N_R}\right)_{E_R,V_R} = -\frac{\mu_R}{T_R} \text{ and } \left(\frac{\partial S_R}{\partial E}\right)_{E_R,N_R} = \frac{1}{T_R} . \tag{10.5}$$

We therefore get

$$k \ln \Omega_R = S_R(N - N_S, E - \epsilon_i) \simeq S_R(N, E) + \frac{N_S \mu_R}{T_R} - \frac{\epsilon_i}{T} , \tag{10.6}$$

and therefore the multiplicity is

$$\Omega_R = C\, e^{(N_S \mu_R - \epsilon_i)/kT} , \tag{10.7}$$

where C is a constant. We insert this back into expression in (10.3) for the probability for a state (N_S, i) :

$$P(N_S, i) = \frac{\Omega_S\, \Omega_R}{\sum \Omega_S\, \Omega_R} \simeq \frac{1 \times C\, e^{(N_S \mu_R - \epsilon_i)/kT}}{\sum \Omega_S\, \Omega_R} = C'\, e^{(N_S \mu_R - \epsilon_i)/kT} , \tag{10.8}$$

where the constant C' can be determined by normalization of the probability, where we need to sum over all possible states for the system S: This means summing over all possible values for the number of particles N_S in system S, and for each value of N_S we need to sum over all the states $i(S)$ of the system given that the system has N_S particles:

$$\sum_{N_S} \sum_i C' e^{(N_S \mu_R - \epsilon_i)/kT} = 1 \;\Rightarrow\; (1/C') = \sum_{N_S} \sum_i e^{(N_S \mu_R - \epsilon_i)/kT} . \tag{10.9}$$

We call this normalization constant Gibbs sum or the Grand partition function, denoted by Z_G:

The probability for a system in diffusive and thermal equilibrium with a thermal bath to be in a state (N, i) is given by

$$P(N,i) = \frac{1}{Z_G} e^{(N\mu - \epsilon_i)/kT} \,. \tag{10.10}$$

where Gibb's sum is

$$Z_G(T,V,\mu) = \sum_N \sum_i e^{(N\mu - \epsilon_i)/kT} \,. \tag{10.11}$$

It is important to realize how this sum is performed: over all particle numbers N, and for each particle number, N, over all the possible states for the systems when it has N particles. Its use become clearer through the following examples.

10.1.1 Averages using Gibbs sum

When we know the probability for all the possible states of the system, we can use this to find the average of any macroscopic quantity using the definition of the average:

Average of a macroscopic variable in a grand canonical system described by (T, V, μ):

$$\langle X \rangle = \sum_N \sum_i X(N,i) P(N,i) \,. \tag{10.12}$$

Notice that both the energy, E, and the number of particles, N, now are fluctuating quantities, but for macroscopic systems the values will typically be sharp.

For example, the average number of particles is:

$$\langle N \rangle = \frac{1}{Z_G} \sum_N \sum_i N e^{(N\mu - \epsilon_i)/kT} \,. \tag{10.13}$$

To simplify this expression, we can use the "derivative trick" we used earlier, but now take the derivative with respect to μ:

$$Ne^{(N\mu-\epsilon_s)/kT} = kT\frac{d}{d\mu}e^{(N\mu-\epsilon_s)/kT} , \tag{10.14}$$

and therefore

$$\langle N\rangle = kT\frac{1}{Z_G}\frac{d}{d\mu}Z_G = kT\frac{\partial \ln Z_G}{\partial \mu} . \tag{10.15}$$

This shows that we again find macroscopic quantities directly from the grand partition function, Z_G, just like we did for the partition function for canonical systems.

10.1.2 Gibb's sum for systems with several species of particles

What if we have more that one species of particles in the system? For example a gas like air may consist of both N_2 and O_2 molecules. In this case, the complete system (reservoir and system), is described by $(T, V, N_1, N_2, \ldots, N_k)$ for the k different species. Similarly, we can introduce the chemical potential for each species as

$$\mu_j = \left(\frac{\partial F}{\partial N_j}\right)_{V,T,N_i\neq N_j} . \tag{10.16}$$

Following exactly the same arguments are above, the probability for a state with N_j particles of each of the species and for a state i of this system is

$$P(i, N_1, N_2, \ldots, N_k) = \frac{1}{Z_G}\sum_{(N_1,N_2,\ldots,N_k)} e^{(\sum_j N_j\mu_j-\epsilon_i)/kT} , \tag{10.17}$$

where the Gibb's sum now is

$$Z_G(T, V, (N_1, N_2, \ldots, N_k)) = \sum_{(N_1,N_2,\ldots,N_k)} e^{(\sum_j N_j\mu_j-\epsilon_i)/kT} . \tag{10.18}$$

We demonstrate the use of Gibb's sum for both a single species and for several species through the following example.

10.1.3 Example: CO poisoning

A hemoglobin molecule is illustrated in Fig. 10.2. Each hemoglobin molecule has four independent adsorption sites, each consisting of a Fe^{2+} ion, and each site can couple to one O_2 molecule. Find the probability that the site is occupied with an O_2 molecule, when you assume that the hemoglobin in the blood is in equilibrium with O_2 in the air in your lungs, at a temperature of $T = 310\,\mathrm{K}$, and where the partial pressure of O_2 is $p_O = 0.2\,\mathrm{atm}$.

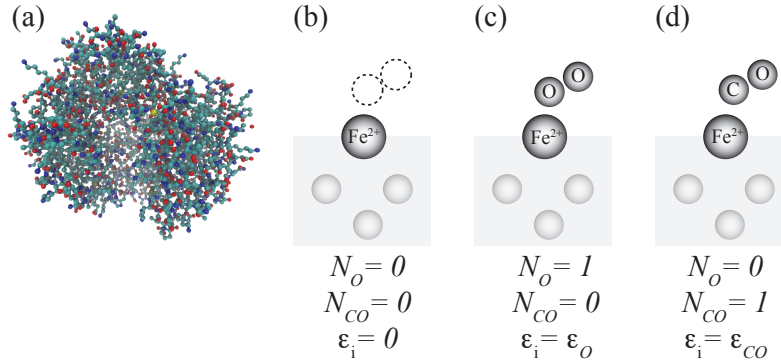


Fig. 10.2 **a** Illustration of a hemoglobin molecule, **b,c,d** Illustration of the three states of the molecule.

Analysis of one species system. We adress a system consisting of a single Fe^{2+} ion in contact with a gas of O_2 molecules. The system therefore has two possible states, as illustrated in Fig. 10.2, either occupied by oxygen ($N = 0$) or not occupied by oxygen ($N = 1$). Here, N, is the number of oxygen molecules attached to the site, and ϵ_i is the energy of the system. The possible energy states are: For $N = 0$ there is only one possible state of the system with energy $\epsilon_i = 0$. For $N = 1$ there is only one possible state of the system with energy $\epsilon = -0.7\,\mathrm{eV}$. (Notice that we could also have studied excited states, which would have added more possible states for the $N = 1$ case, but we disregard this possibility here.)

To find the average number of attached O_2 molecules, we first find the probability, and then calculate the average number of N for the system. To find the probability, we first find the grand partition function:

$$Z_G = \underbrace{e^{(0\,\mu - 0)/kT}}_{(N=0)} + \underbrace{e^{(1\,\mu - \epsilon)/kT}}_{(N=1)} \,. \tag{10.19}$$

We see that this function includes the chemical potential, μ. But we have not been given a value for μ. How can we find this? We assume that in the lungs, the blood is in diffusive equilibrium with the air in the lungs, and we can therefore use the chemical potential for an ideal gas, where we know that the density of O_2 molecules can be related to the partial pressure of O_2 using the ideal gas law: $pV = NkT$, and $N/V = p/kT$, hence

$$\mu = -kT \ln(n_Q/n) = -kT \ln(n_Q(T)\, kT/p) \simeq -0.6\text{eV} \,, \tag{10.20}$$

when $T = 310\text{K}$, which is the temperature in your body, and $p = 0.2\,\text{atm}$, which is the partial pressure for oxygen. This gives

$$e^{-(\epsilon-\mu)/kT} \simeq e^{0.1\,\text{eV}/kT} \simeq 40 \,, \tag{10.21}$$

which means that the probabilty to be occupied is

$$P(N=1) = \frac{40}{1+40} \simeq 0.98 \,. \tag{10.22}$$

Table 10.1 Elements for the partition sum.

N_O	N_{CO}	ϵ_i
0	0	0
1	0	ϵ_O
0	1	ϵ_{CO}

Analysis of two-species system. Now, what happens if CO is also present, which also can be adsorbed at the same site. Now, there are three possible states: unoccupied, occupied by O_2 or occupied by CO. And there are two species, the number of oxygen molecules, N_O, and the number of CO molecules at the site, N_{CO}. We must therefore include both species in the grand partition function. All the possible particle number and the states of the system — all the terms that must be included in the grand partition function — are listed in Table **??**. The partition function is therefore

$$Z_G = e^{(0\,\mu_O + 0\,\mu_{CO} - 0)/kT} + e^{(1\,\mu_O + 0\,\mu_{CO} - \epsilon_O)/kT} + e^{(0\,\mu_O + 1\,\mu_{CO} - \epsilon_{CO})/kT} \,. \tag{10.23}$$

Now we need numbers. CO is more strongly bound, so $\epsilon_{CO} = -0.85\,\text{eV}$. But what is the chemical potential? We could still use the expression for the ideal gas, but with the concentration n_{CO} of CO in air. If CO is x times less abundant that oxygen, we would find that

$$\mu_{CO} = -kT \ln \frac{n_Q}{n_{CO}} = -kT \ln \frac{n_Q}{x\, n_O} = -kT \ln \frac{n_Q}{n_O} + kT \ln x \ , \quad (10.24)$$

where $kT \ln 100 = 0.12\,\text{eV}$, so that $\mu_{CO} = 0.72\,\text{eV}$. This gives for the new Gibbs factor:

$$e^{(1\,\mu_{CO} - \epsilon_{CO})/kT} = 120 \ , \quad (10.25)$$

and therefore

$$P(O) = \frac{40}{1 + 40 + 120} = 0.25 \ . \quad (10.26)$$

So just a small amount of CO is devastating!

10.1.4 Example: Diffusive equilibrium in the Einstein crystal

10.1.5 Example: Vacancies

10.2 Fermi-Dirac and Bose-Einstein distribution

When we discussed the ideal gas we assumed that quantum effects were not important. This was implied when we introduced the term $1/N!$ for the partition function for the ideal gas, because this term only was valid in the limit when the number of states are many compared to the number of particles.

Now, we will address the general case of a quantum gas. The concepts and methods we introduce here will be important for your understanding of applied problems in quantum mechanics, such as electrical conducitivity, solar power cells, thermo-electric materials, neutron stars, and many other processes.

The system we will study is a system of non-interacting particles in a box of size $L \times L \times L = V$. For a single particle in a box, we know that the possible translational states of the particle have energies

$$\begin{aligned} \epsilon(n_x, n_y, n_z) &= \frac{\hbar^2}{2m} \left(\frac{\pi}{L}\right)^2 \left(n_x^2 + n_y^2 + n_z^2\right) \\ &= a \left(n_x^2 + n_y^2 + n_z^2\right) = a n^2 \ , \end{aligned} \quad (10.27)$$

where

$$a = \frac{\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2 = \frac{h^2}{8mV^{2/3}} \ , \tag{10.28}$$

and m is the mass of the particle. The three positive integers n_x, n_y, n_z enumerates all the possible states of the system. That is, the state of one particle in a box is given by (n_x, n_y, n_z), where $n_x = 0, 1, 2, \ldots$, $n_y = 0, 1, 2, \ldots$, $n_z = 0, 1, 2, \ldots$.

This describes a the states of a single particle in a box. However, if we have many particles in the same box, we need to address whether two particles can occupy the same quantum state in the system. This depends on the type of particle:

- For **Fermions** (1/2 spin) Only 0 or 1 fermion can be in a particular state.
- For **Bosons** (integer spin) Any number of bosons can be in a particular state.

Given this condition, we will assume that there are no other interactions between the particles and that they can occupy any of the possible quantum states of the system.

How can we find the thermodynamics of this system using a statistical physics approach? We will apply a simplification that allows us to address each state separately: Let us assume that we study a grand-canonical system, that is, we will assume a system at constant temperature, T, volume, V, and chemical potential μ. We will then use a "trick" – we will look only at a single state in the system, find the probability that this state is occupied, and then sum the occupation numbers of all states to find the total number of particles in the system and the total energy of the system.

Let us first look at a system of Fermions with spin 1/2, such as the electrons in an electron gas. In this case, there can only be one particle in each state. However, for each translational state given by (n_x, n_y, n_z) we will have 2 spin states, $\sigma = \pm 1/2$. Let us pick one such state, (n_x, n_y, n_z, σ) with energy $\epsilon(n_x, n_y, n_z)$ (For example, the state $(1, 0, 1, -1/2)$ with energy $\epsilon = a(1^2 + 0^2 + 1^2) = 2a$). We study the occupancy of this state using the Gibbs sum approach. The Gibbs sum is the sum over the possible number of particles in this state, and for each number of particles, we sum over all the possible states of the system given this number of particles. In this case these sums are simple. The system may have either $N = 0$ particles or $N = 1$ particle. For $N = 0$ the system consisting of the state (n_x, n_y, n_z, σ) can only have

one possible energy – zero. For $N = 1$, the energy of the system is simply $\epsilon(n_x, n_y, n_z, \sigma)$, which we write at ϵ for short. The Gibbs sum is therefore

$$\begin{aligned} Z_G &= \sum_{N=0}^{1} \sum_{s(N)} e^{-(\epsilon_{s(N)} - N\mu)/kT} \\ &= e^{-(0-0\dot{\mu})/kT} + e^{-(\epsilon - 1\dot{\mu})/kT} \\ &= 1 + e^{-(\epsilon-\mu)/kT} \, . \end{aligned} \tag{10.29}$$

In the Gibbs sum formalism, the probability to have N particles in the system and for the system to be in the state $s(N)$ is

$$P(\epsilon_{s(N)}, N) = \frac{1}{Z_G} e^{-\left(\epsilon_{s(N)} - N\mu\right)/kT} \, ., \tag{10.30}$$

The average number of particles in the state $s = (n_x, n_y, n_z, \sigma)$, $\bar{N}_s$ is then

$$\begin{aligned} \bar{N}_{(n_x,n_y,n_z,\sigma)} &= \sum_{N=0}^{1} P(\epsilon_{s(N)}, N) N \\ &= 0 \cdot 1 + 1 \cdot \frac{1}{Z_G} e^{-(\epsilon - \dot{\mu})/kT} \\ &= \frac{e^{-(\epsilon-\mu)/kT}}{e^{-(\epsilon-\mu)/kT} + 1} \\ &= \frac{1}{e^{(\epsilon-\mu)/kT} + 1} \, . \end{aligned} \tag{10.31}$$

This quantity – the average number of particles in a state (n_x, n_y, n_z, σ) – is called the **Fermi-Dirac distribution function**:

$$\bar{N}_{(n_x,n_y,n_z,\sigma)} = f_{FD}(\epsilon; \mu, T) = \frac{1}{e^{(\epsilon-\mu)/kT} + 1} \, . \tag{10.32}$$

Because the average number of particles in a state (n_x, n_y, n_z, σ) only depends on the energy of this state, and not on any other details of the state, we write the average number as a function of the energy ϵ of the state alone.

What does the average number of particles – the Fermi-Dirac distribution function – look like? First, let us look at the functional form. We see that if we introduce $x = (\epsilon - \mu)/kT$, the functional form is

$$f(x) = \frac{1}{e^x + 1} \, , \tag{10.33}$$

This is a function that provides a smooth transition from a value of 1 when $x \ll -1$ to 0 when $x \gg +1$. We have plotted $f_{FD}(\epsilon;\mu,T)$ as a function of $(\epsilon-\mu)/kT$ in Fig. 10.3. The temperature effectively determines the range of the transition region. We can illustrate this by introducing an energy scale in the plot. Let us use the chemical potential μ as the energy scale[1], that is, we plot $f_{FD}(\epsilon;\mu,T)$ as a function of ϵ/μ for different values of T corresponding to $kT = 0.2\mu$, $kT = \mu$ and $kT = 5\mu$. We see that when the temperature goes towards zero, the function approximates a step function, and when the temperature increases the transition becomes broader. These plots were generated by the following short script:

```
import numpy as np
import matplotlib.pyplot as plt
plt.subplot(2,1,1)
x = np.linspace(-10,10);
f = 1.0/(np.exp(x)+1);
plt.plot(x,f)
plt.xlabel('$(\\epsilon-\\mu)/kT$')
plt.ylabel('$f\\epsilon;\\mu,T)$')
plt.subplot(2,1,2)
ktvalues = [5.0,1.0,0.2]
for i,kt in enumerate(ktvalues):
    x = np.linspace(-20,20,100)
    f = 1.0/(np.exp((x-1)/kt)+1)
    plt.plot(x,f,label="$kT = "+str(kt)+"\\epsilon$")
plt.xlabel('$\epsilon/\mu$'); plt.ylabel('$f(\epsilon, \mu,T)$')
plt.legend()
```

What happens when $T = 0$? We see from the plots that the function approaches a step function that goes from 1 to 0 at the chemical potential. We call the chemical potential at $T = 0$ the **Fermi energy**:

$$\epsilon_F = \mu(T = 0) \; . \tag{10.34}$$

At $T = 0$ all the states up to the level ϵ_F are occupied – and none of the levels above ϵ_F are occupied. Later on we will see how we can relate ϵ_F to the number of particles N in the gas.

10.2.1 Bose-Einstein distribution

Bosons are particles of integer spins. For Bosons, more than one particle can be in the same state. Examples of bosons are photons, phonos, and

[1] Notice that here the chemical potential is a given, constant, whereas further on, we will assume that the number of particles is constant instead, and the calculate a chemical potential that will depend on T, V, and N.

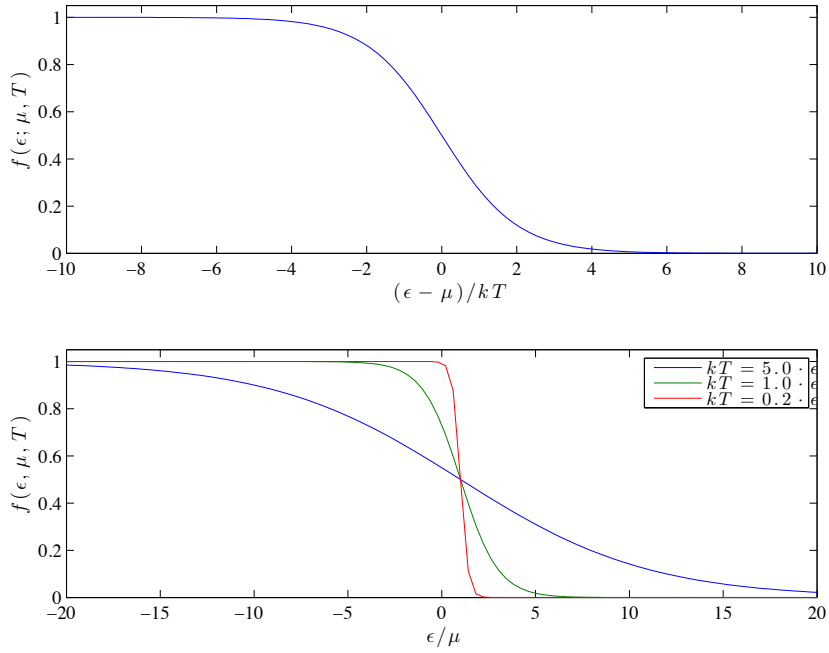


Fig. 10.3 Plot of Fermi-Dirac distribution.

particles such as Helium-4. The Gibbs sum and the average number of particles per state is different for Bosons than for Fermions because there is no limit on the number of particles in a given state for Bosons. This means that the number of particles, N, may run from 0 to ∞. The energy $\epsilon_{s(N)}$ of the system when there are N particles in a state s (where each particle has an energy ϵ_s) is $N\epsilon_s$. The Gibbs sum is therefore

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} \sum_{s(N)} e^{-(\epsilon_{s(N)}-N\mu)/kT} \\ &= \sum_{N=0}^{\infty} e^{-(N\epsilon_s-N\mu)/kT} \\ &= \sum_{N=0}^{\infty} e^{-N(\epsilon_s-\mu)/kT} = \frac{1}{1-e^{-(\epsilon_s-\mu)/kT}} , \end{aligned} \tag{10.35}$$

where the sum is the well-known geometric series. The average number of particles in state s is

$$\begin{aligned}\bar{N}_s &= \sum_N \sum_{s(N)} \frac{N}{Z_G} e^{-(\epsilon_{s(N)}-N\mu)/kT} \\ &= \frac{1}{Z_G} \sum_N N e^{-(N\epsilon_s - N\mu)/kT} \\ &= \frac{1}{Z_G} \sum_N N e^{-N(\epsilon_s-\mu)/kT} \\ &= \frac{1}{Z_G} \sum_N kT \frac{\partial}{\partial \mu} e^{-(\epsilon_s-\mu)N/kT} \\ &= \frac{\partial}{\partial \mu} kT \ln Z_G \ ,\end{aligned} \tag{10.36}$$

where we have used a usual "trick" by introducing the derivative and taking it outside the sum. The result is

$$\bar{N}_s = f_{BE}(\epsilon_s; \mu, T) = \frac{1}{e^{(\epsilon_s-\mu)/kT} - 1} \ . \tag{10.37}$$

Because the distribution only depends on the energy ϵ_s of the state s of the particle-in-box system, it is common to simply write the distribution is a function of ϵ. We have plotted the Bose-Einstein distribution along with the Fermi-Dirac distribution in Fig. 10.4.

```
import numpy as np
import matplotlib.pyplot as plt
kt = 1.0
x = np.linspace(-10,10,100)
f = 1.0/(np.exp(x/kt)+1)
plt.plot(x,f,label="FD")
x = np.linspace(0.4,10,100)
f = 1.0/(np.exp(x/kt)-1)
plt.plot(x,f,label="BE")
x = np.linspace(-1.0,10,100)
f = np.exp(-x/kt)
plt.plot(x,f,label="C ")
plt.xlabel('$(\\epsilon-\\mu)/kT$')
plt.ylabel('$f(\\epsilon, \\mu,T)$')
plt.legend()
plt.ylim(top = 2.0)
```

We see that the distribution diverges as ϵ approaches μ, that is, as $\epsilon - \mu$ approaches 0. This means that as the temperature goes towards zero, all the particles will be in the ground state.

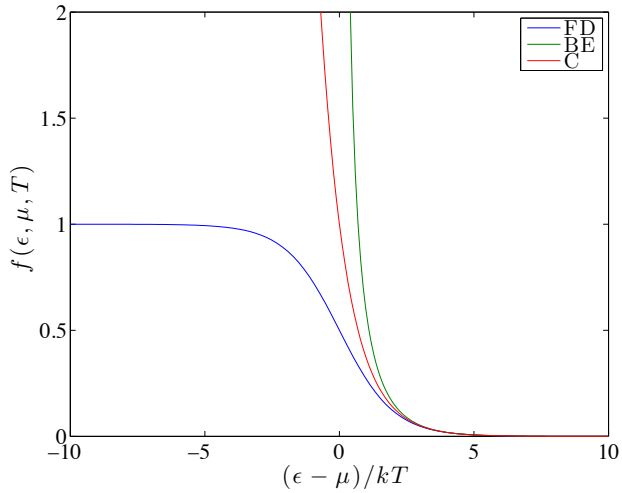


Fig. 10.4 Plot of Bose-Einstein distribution.

10.2.2 Classical limit

A gas is in the classical limit when the number of particles in a given state is much smaller than one. For a gas at room temperature and atmospheric pressure, the typical number of particles in each state is less than 10^{-6}. In this case, the differences between Fermions and Bosons are not important – these differences are only important when the number of particles is one or larger This means that the distribution function $f(\epsilon; \mu, T)$, which gives the number of particles in a state with energy ϵ, also must be very small, $f \ll 1$. This implies that the exponential term must be very large. In that case, we can assume that it is much larger than 1 and the ± 1 in the denominator does not matter. In the classical limit we therefore get that the behavior of a Boson gas and a Fermion gas are identical. The **classical distribution function** is

$$f_{\text{classical}}(\epsilon, \mu, T) \simeq e^{(\epsilon-\mu)/kT} \ , \tag{10.38}$$

We have plotted all the distributions functions in the same plot in Fig. 10.4, where we see that all the distribution functions are indistinguishable for large values of $(\epsilon - \mu)/kT$.

Let us now compare this result with our results for a classical ideal gas in a canonical system, where the chemical potential is a function of T, V, N: $\mu = \mu(T, V, N)$. Since we have now addressed the system using the Gibbs formalism, the (average) number of particles in the system

is a function of T, V, and μ, $\bar{N} = \bar{N}(T, V, \mu)$. How can we compare the two results? We can calculate the average total number of particles, $\bar{N}$, as a function of the chemical potential in the T, V, μ system, and then solve the resulting equation to find the chemical potential as a function of the average total number of particles. (We may not always be able to solve this equation analytically, but we may then resort to numerical methods and still save the day).

For a given chemical potential, μ, the total average number of particles, $\bar{N}$, is the sum of the average number of particles in each state, s, summed over all possible states s:

$$\bar{N} = \sum_s \bar{N}_s = \sum_s f(\epsilon_s, \mu, T) \ , \tag{10.39}$$

In the classical limit, we have argued that the distribution function $f(\epsilon; \mu, T)$ has a simple form, so that the sum is:

$$\bar{N} = \sum_s \exp((\epsilon_s - \mu)/kT = e^{\mu/kT} \sum_s e^{\epsilon_s/kT} \ . \tag{10.40}$$

The last sum is the sum over all states, s, for a particle in a box. We recognize this as the one-particle partition function, Z_1, for the translational motion of one particle in a box:

$$\bar{N} = e^{\mu/kT} Z_1 \ , \tag{10.41}$$

and the chemical potential therefore is:

$$\mu = kT \ln(\bar{N}/Z_1) \ . \tag{10.42}$$

For an ideal gas - that is for a particle in a box - we found that the one-particle partition function is

$$Z_1 = n_Q V \ , \tag{10.43}$$

where

$$n_Q = \left(\frac{mkT}{2\pi h^2}\right)^{3/2} \ , \tag{10.44}$$

was called the quantum concentration. The chemical potential is therefore

$$\mu = kT \ln(\bar{N}/(n_Q V)) = kT \ln(\bar{n}\rangle/n_Q) \ , \tag{10.45}$$

where $\bar{n} = \bar{N}/V$.

10.3 Fermi gases

We can use a similar approach to determine the number of particles in a Fermi gas: We find the average number of particles in each state for a particle in a box, and then find the total average number of particles as the sum of the average number of particles in each state:

$$\bar{N} = \sum_s f(\epsilon_s; \mu, T) \ , \tag{10.46}$$

where the sum is over all the states for a single particle in a box.

However, we are usually instead interested in understanding the behavior in a system with a given number, N, of particles and not a given chemical potential. How can we transform the results we have for a given chemical potential to a system with a given number of particles? Since we know how to calculate the (average) number of particles, $\bar{N}$, in a system, we can simply find the chemical potential, $\mu(V, T, N)$, for which the average number of particles, $\bar{N}$, is equal to the number of particles, N: $\bar{N} = N$. From a more practical point of view, this means that we first calculate $\bar{N} = \bar{N}(T, V, \mu)$ for a given μ, and then find μ from the equation

$$N = \bar{N} = \bar{N}(T, V, \mu) \ . \tag{10.47}$$

10.3.1 Zero temperature limit

Let as start at the simplest case. At very low temperatures, where $T = 0$, the Fermi distribution is a step function. This means that all the states below the chemical potential at $T = 0$ – the Fermi energy – are occupied by one particle, and none of the states above the chemical potential are occupied. (We recall that the chemical potential at $T = 0$ is called the Fermi energy, $\epsilon_F = \mu(T = 0)$.)

In this case, it is simple to find the number of particles in the system. Since all the states below ϵ_F have one particle in them, we simply need to count the number of states with energies below ϵ_F and this will be the (average) number of particles in the gas.

In general, we find the average number of particles in the gas from

$$\bar{N} = \sum_s \bar{N}_s = \sum_s f(\epsilon_s; \mu, T) \ , \tag{10.48}$$

where the sum is over all the states s of the system. The states of a particle in a box are given as (n_x, n_y, n_z, σ). We can sketch the states in the n_x, n_y, n_z space as illustrated in Fig. 10.5. We see that for each set of integers n_x, n_y, n_z we have two states. If we look at a "volume" $\Delta V = \Delta n_x \Delta n_y \Delta n_z$ in this space, we know that $\Delta n_x = \Delta n_y = \Delta n_z = 1$ we will have 2 states in this volume (as long as the volume is in octant where n_x, n_y, and n_z all are positive). We can therefore introduce a **density of states** in the n_x, n_y, n_z-space: $D(n_x, n_y, n_z)$ so that the number of states, $\Delta\mathcal{N}$, in the volume from n_x, n_y, n_z to $n_x + \Delta n_x, n_y + \Delta n_y, n_z + \Delta n_z$ is

$$\Delta\mathcal{N} = D(n_x, n_y, n_z)\Delta n_x \Delta n_y \Delta n_z \ . \tag{10.49}$$

The density of states for this problem – a particle in a box – is uniform in n_x, n_y, n_z-space:

$$D(n_x, n_y, n_z) = \begin{cases} 2 & \text{when } n_x > 0, n_y > 0, n_z > 0 \\ 0 & \text{otherwise} \end{cases} \ . \tag{10.50}$$

The number of states in a volume V is therefore simply $D(n_x, n_y, n_z)V = 2V$ in n_x, n_y, n_z-space.

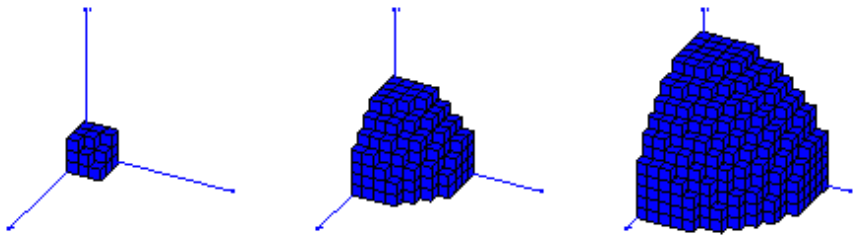


Fig. 10.5 Illustration of the states of a particle in a box. Each vertice of a voxel represents a single state. The three figures illustrates the states that are below n_F for $n_F = 3, 7, 11$.

For $T = 0$ we wanted to count all states with energies below ϵ_F. The energy of a state is

$$\epsilon(n_x, n_y, n_z) = a\left(n_x^2 + n_y^2 + n_z^2\right) = an^2 \ , \tag{10.51}$$

which means that the energy only depends on the distance n from the origin in n_x, n_y, n_z-space. The condition that $\epsilon \leq \epsilon_F$ therefore corresponds to

$$\begin{aligned} \epsilon &\leq \epsilon_F \\ an^2 &\leq \epsilon_F \\ n &\leq \sqrt{\frac{\epsilon_F}{a}} = n_F \ . \end{aligned} \tag{10.52}$$

To find the number of states with $\epsilon \leq \epsilon_F$ we must therefore count the number of states in the n_x, n_y, n_z-space with $n \leq n_F$, that is, we must count all the states inside a sphere with radius n_F.

Let us follow two approaches to find the number of states in this volume. First, we simply use that the denstiy of states in n_x, n_y, n_z-space is uniform, and that the number of states therefore is the volume of the sphere with radius n_F multiplied with the density of states which is 2, taking into account that only positive values of n_x, n_y, n_z. Using this argument, the number of states inside n_F is

$$\bar{N} = 2 \cdot V_F = \underbrace{2}_{\text{2 spins}} \cdot \underbrace{\frac{1}{8}}_{n_x,n_y,n_z>0} \cdot \underbrace{\frac{4\pi}{3} n_F^3}_{\text{volume of sphere}} = \frac{\pi}{3} n_F^3 \ . \tag{10.53}$$

However, it is instructive to calculate $\bar{N}$ following a more formal approach, since this approach can be generalized to the case where $T > 0$. The total average number of particles in the system is the sum of the average number of particles in each state:

$$\bar{N} = \sum_s \bar{N}_s = \sum_s f(\epsilon_s; \mu, T) \ , \tag{10.54}$$

where the states now are given as (n_x, n_y, n_z, σ):

$$\bar{N} = \sum_{n_x} \sum_{n_y} \sum_{n_z} \sum_{\sigma} f(\epsilon(n_x, n_y, n_z); \mu, T) \ . \tag{10.55}$$

If we assume that the sum will be over many states, we can approximate this sum by an integral over (n_x, n_y, n_z), where we also must include the density of states:

$$\bar{N} = \iiint f(\epsilon(n_x, n_y, n_z); \mu, T) D(n_x, n_y, n_z) dn_x dn_y dn_z \ . \tag{10.56}$$

For $T = 0$, $f(\epsilon)$ becomes a step function, and we only need to include the n_x, n_y, n_z values that satisfy $n \leq n_F$:

$$\bar{N} = \iiint_{n \leq n_F} D(n_x, n_y, n_z) dn_x dn_y dn_z \ . \tag{10.57}$$

Now, we can use our knowledge from calculus to spherical coordinates in the n_x, n_y, n_z space. The integral transformation needed to go from Cartesian to spherical coordinates is to replace $dn_x dn_y dn_z = 4\pi n^2 dn$. However, when we integral in spherical coordinates, we must remember that we only include 1/8 of the sphere since we only allow positive coordinates. The integral transform is therefore

$$\bar{N} = \iiint_{n \leq n_F} D(n_x, n_y, n_z) dn_x dn_y dn_z = \int_0^{n_F} D(n_x, n_y, n_z) \frac{4\pi}{8} n^2 dn \ . \tag{10.58}$$

We insert $D(n_x, n_y, n_z) = 2$, getting:

$$\bar{N} = \int_0^{n_F} 2\frac{4\pi}{8} n^2 dn = \int_0^{n_F} \pi n^2 dn = \ . \tag{10.59}$$

From this equation, we see that we can interpret πn^2 as the density of states in n-space (which here is different from n_x, n_y, n_z-space): $D(n) = \pi n^2$.

This integral can be easily solved:

$$\bar{N} = \int_0^{n_F} \pi n^2 dn = \frac{\pi n_F^3}{3} \ . \tag{10.60}$$

This is, of course, the same as we found in (10.53).

Solving for the chemical potential. We insert $n_F = \sqrt{\epsilon_F / a}$ to find the average number of particles in terms of the chemical potential at $T = 0$:

$$\bar{N} = \frac{\pi}{3} n_F^3 = \frac{\pi}{3} \frac{\epsilon_F^{3/2}}{a^{3/2}} \ . \tag{10.61}$$

We can solve for ϵ_F, getting:

$$\left(\frac{3\bar{N}}{\pi} \right) a^{3/2} = \epsilon_F^{3/2} \ , \tag{10.62}$$

and therefore

$$\left(\frac{3\bar{N}}{\pi} \right)^{2/3} a = \epsilon_F \ . \tag{10.63}$$

We recall that a also includes the volume V of the system:

$$a = \frac{h^2}{8mV^{2/3}} \ , \tag{10.64}$$

which gives

$$\epsilon_F = \left(\frac{3\bar{N}}{\pi}\right)^{2/3} \frac{h^2}{8mV^{2/3}} = \frac{h^2}{8m}\left(\frac{3}{\pi}\right)^{2/3}\left(\frac{\bar{N}}{V}\right)^{2/3} . \tag{10.65}$$

We repeat the argument that lead to this conclusion: At $T = 0$ all the states up to the chemical potential are occupied by one particle. We can therefore find the average number of particles, $\bar{N}$, in the gas when the chemical potential is given, by summing up how many staes we have below the chemical potential. We reverse this argument, and find the chemical potential given the number of particles, which gives us $\mu(T = 0) = \epsilon_F$ as a function of $\bar{N}$ (and V) - which is what we found above.

Average energy at $T = 0$. We can also find the average energy of the gas at $T = 0$ using a similar approach. The average energy of a state (n_x, n_y, n_z, σ) is $N(n_x, n_y, n_z, \sigma) \cdot \epsilon(n_x, n_y, n_z, \sigma)$. We can therefore find the total average energy by summing the average energies for all the states:

$$\bar{E} = \sum_{(n_x,n_y,n_z,\sigma)} N(n_x, n_y, n_z, \sigma) \cdot \epsilon(n_x, n_y, n_z, \sigma) . \tag{10.66}$$

The energy only depends on the magnitude of (n_x, n_y, n_z), $\epsilon = an^2$, we can therefore replace the sum with an integral over n:

$$\bar{E} = \int_0^\infty an^2 f(\epsilon; \mu, T) D(n) dn , \tag{10.67}$$

where $D(n)$ is the density of states in the n-space and $D(n)dn$ gives the number of states in the interval from n to $n + dn$.

How can we solve this integral? We will work more with this in general in the following, but for the $T = 0$ case, the integral reduces to a simple form because $f(\epsilon; \mu, T = 0) = 0$ when $n > n_F$ and 1 otherwise. The integral therefore becomes:

$$\bar{E} = \int_0^{n_F} an^2 D(n) dn , \tag{10.68}$$

where $D(n)dn = \pi n^2 dn$:

$$\begin{aligned}\bar{E} &= \int_0^{n_F} an^2 \pi n^2 dn \\ &= \pi a \int_0^{n_F} n^4 dn \\ &= \frac{\pi}{5} a n_F^5 \\ &= \underbrace{\left(a n_F^2\right)}_{=\epsilon_F} \cdot \underbrace{\left(\frac{\pi}{3} n_F^3\right)}_{\bar{N}} \cdot \frac{2}{5} \\ &= \frac{2}{5} \bar{N} \epsilon_F \ .\end{aligned} \tag{10.69}$$

(Notice that we used the result $\bar{N} = (\pi/3)n_F^3$ from (10.61)). This result is surprisingly simple! This simplicity is not that surprising: Since we are at $T = 0$ we have filled up all the states up to the level ϵ_F, starting from the states with lowest energy. Because of the Pauli principle we can only have one Fermion in each state, and the particles will therefore be distributed across a range of states with a range of energies. You can gain some inituition about these states from the illustration of the states in n_x, n_y, n_z-space in Fig. 10.5. However, since the energies of the states increases rapidly with distance from the origin, $\epsilon = an^2$, most of the energy comes from the states far from the origin. The result is that $\bar{E} = (2/5)N\epsilon_F$, which means that many of the states are near ϵ_F. We will gain more intuition into this in the next section.

The result we have found also give us useful physical insights into the quantuum effects on the behavior of a quantum gas. We see that if we keep N constant for the gas, but compress the system (decrease V), then ϵ_F will increase (from (10.65) we know that $\epsilon_F \propto (\bar{N}/V)$), and therefore the energy will increase. This effect gives rise the a repulsion due to Fermi-exclusion.

10.3.2 Density of states

We introduced the density of states in both n_x, n_y, n_z-space and in n-space in order to describe the number of states in a small interval around a point (n_x, n_y, n_z) or n. The density of states in n_x, n_y, n_z-space was $D(n_x, n_y, n_z)$. By this we meant that the number of states between (n_x, n_y, n_z) and $(n_x + dn_x, n_y + dn_y, n_z + dn_z)$ is

$$\mathcal{N} = D(n_x, n_y, n_z) dn_x dn_y dn_z \ , \tag{10.70}$$

where we found that $D(n_x, n_y, n_z) = 2$ when n_x, n_y, and n_z all are positive, and zero otherwise.

We have similarly introduced a density of states in n-space - so that the number of states with magnitude n between n and $n + dn$ is

$$D(n)dn = 2 \cdot \frac{1}{8} \cdot 4\pi n^2 dn = \pi n^2 dn \ , \tag{10.71}$$

where the factor 2 again comes from the two spin states per $\mathbf{n}$ value and the factor of 1/8 is included because only positive values of n_x, n_y, and n_z are used, hence only 1/8 of the volume of a sphere with radius n is in the first octant. We therefore call $D(n) = \pi n^2$ the density of states for the three-dimensional ideal gas. Notice that the density of states would be different for a one- or a two-dimensional gas - the n-dependence would be different!

However, the density of states in n-space is not that useful since the occupation number - $f(\epsilon, \mu, T)$ - is given in terms of ϵ. If we want to calculate average values - such as the average number of particles or the average energy - we perform the "sums" over (n_x, n_y, n_z) which we convert to integrals of n (since all quantities only depend on the length of (n_x, n_y, n_z) - and since the number of states is large so that the integral is a good approximation for the sum):

$$\bar{N} = \sum_{(n_x, n_y, n_z)} 2f(\epsilon(n_x, n_y, n_z); \mu, T) = \int_0^\infty f(\epsilon(n), \mu, T) D(n) dn \ . \tag{10.72}$$

Similarly, we find the average energy from a similar integral:

$$\bar{U} = \sum_{(n_x, n_y, n_z)} 2\epsilon(n_x, n_y, n_z) f(\epsilon(n_x, n_y, n_z); \mu, T) = \int_0^\infty \epsilon(n) f(\epsilon(n); \mu, T) D(n) dn \ , \tag{10.73}$$

Now - we can evaluate this integral by inserting $\epsilon(n) = an^2$. Or we can change integration variable in the integral, and instead integrate over ϵ. That is, we introduce the variable $\epsilon(n)$ and the differential:

$$d\epsilon(n) = \frac{d\epsilon}{dn} dn \ , \tag{10.74}$$

giving the integrals:

$$\bar{N} = \int_0^\infty f(\epsilon; \mu, T) D(n(\epsilon)) \frac{1}{d\epsilon/dn} d\epsilon \ , \tag{10.75}$$

and

$$\bar{U} = \int_0^\infty \epsilon f(\epsilon;\mu,T) D(n(\epsilon)) \frac{1}{d\epsilon/dn} d\epsilon \ . \tag{10.76}$$

From these expressions we see that we can interpret

$$D(n(\epsilon)) \frac{1}{d\epsilon/dn} d\epsilon = D(\epsilon) d\epsilon \ , \tag{10.77}$$

as the density of states in ϵ space. The quantity $D(\epsilon)d\epsilon$ gives the number of states with energies between ϵ and $\epsilon + d\epsilon$.

Let us find the density of states in energy space for the three-dimensional gas. We find it using the integral transformation we introduced above:

$$D(n)dn = D(\epsilon)d\epsilon \ , \tag{10.78}$$

and therefore

$$D(n(\epsilon)) \frac{1}{d\epsilon/dn} = D(\epsilon) \ , \tag{10.79}$$

where we now use that

$$\epsilon(n) = \frac{\hbar}{2m} \left(\frac{\pi}{L}\right)^2 n^2 = an^2 \ , \tag{10.80}$$

and therefore we find:

$$n = (\epsilon/a)^{1/2} \ , \tag{10.81}$$

and that

$$\frac{d\epsilon}{dn} = 2an \ , \tag{10.82}$$

which we insert in (10.79) to get

$$
\begin{aligned}
D(\epsilon) &= D(n)\frac{1}{d\epsilon/dn} \\
&= \pi n^2 \frac{1}{2an} \\
&= \frac{\pi}{2a} n \\
&= \frac{\pi}{2a}\sqrt{\frac{\epsilon}{a}} \\
&= \frac{\pi}{2a^{3/2}}\epsilon^{1/2} \\
&= \frac{\pi(8m)^{3/2}}{2h^3} V\sqrt{\epsilon} \\
&= \frac{3N}{2\epsilon_F^{3/2}}\sqrt{\epsilon}\,.
\end{aligned}
\tag{10.83}
$$

The nice thing about this expression is that it can be interpreted in the same plot - in the same space as we say here - as the distribution function f. This is illustrated in Fig. 10.6, where we have plotted both the density of states $D(\epsilon)$, $f(\epsilon;\mu,T)$ and their product in the same plot.

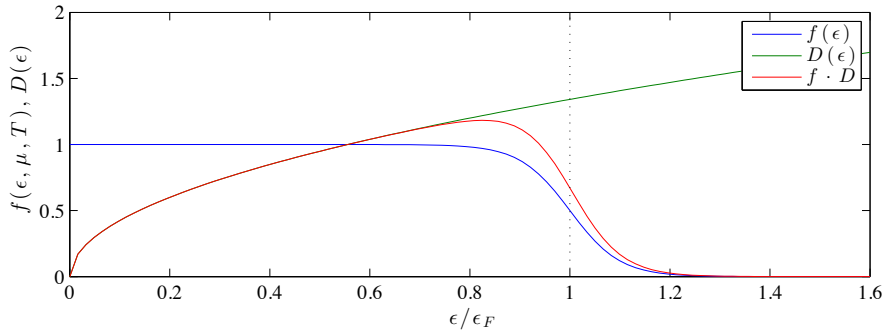


Fig. 10.6 Plot of the density of states, $D(\epsilon)$, the number of particles in a state with energy ϵ, $f(\epsilon;\mu,T)$, and their product for $kT = 0.25 \cdot \epsilon_F$.

10.3.3 Behavior at $T = 0$

We can use Fig. 10.6 to gain more insight into the system at $T = 0$. In this case, the distribution function $f(\epsilon;\mu,T)$ is a step function with the step at ϵ_F. We can therefore think of what happens at $T = 0$ in a slightly different way: We can assume that we have a given number of particles

N. These particles will fill up all the states, starting with the states with the lowest energy first. Since two particles cannot occupy the same state, we fill in the density of states, $D(\epsilon)$, function starting from $\epsilon = 0$ and then up to ϵ_F, as illustrated in Fig. 10.7. In this case, it is simple to find the average number of particles – it is the area under the $D(\epsilon)$ curve from 0 to ϵ_F:

$$\begin{aligned} \bar{N} &= \int_0^\infty f(\epsilon;\mu,T)D(\epsilon)d\epsilon \\ &= \int_0^{\epsilon_F} D(\epsilon)d\epsilon \\ &= \int_0^{\epsilon_F} \frac{\pi}{2a}\sqrt{\frac{\epsilon}{a}}d\epsilon \\ &= \frac{\pi}{2a}\frac{1}{\sqrt{a}}\frac{2}{3}\epsilon_F^{3/2} , \end{aligned} \tag{10.84}$$

where we have inserted $D(\epsilon) = (\pi/2a)\sqrt{\epsilon/a}$. This result corresponds to what we found previously. However, it is simple to see from the $\epsilon^{1/2}$ dependence of $D(\epsilon)$ that the average number of particles will depend on $\epsilon_F^{3/2}$.

10.3.4 Qualitative behavior at low temperatures

We have found the chemical potential, ϵ_F, at $T = 0$. How does the chemical potential depend on temperature at low temperatures? We can gain insights just from the behavior of $D(\epsilon)$ itself. If we look at a gas with a given number N of particles (such as a given number of free electrons in an electron gas), then we can solve the equation

$$N = \bar{N} = \int_0^\infty f(\epsilon;\mu(T),T)D(\epsilon)d\epsilon , \tag{10.85}$$

for $\mu(T)$. What does this integral represent? It represents the area under the curve $f(\epsilon;\mu,T)\cdot D(\epsilon)$ as illustrated in Fig. 10.7. Since the area corresponds to the number of particles, and we assume that the number of particles is conserved, the area must be conserved when the temperature is increased from $T = 0$.

What happens as T is increased? Some of the particles in the states below the chemical potential (area A) are moved into states above the chemical potential (area B) as illustrated in Fig. 10.7. However, the number of particles removed from states below the chemical potential

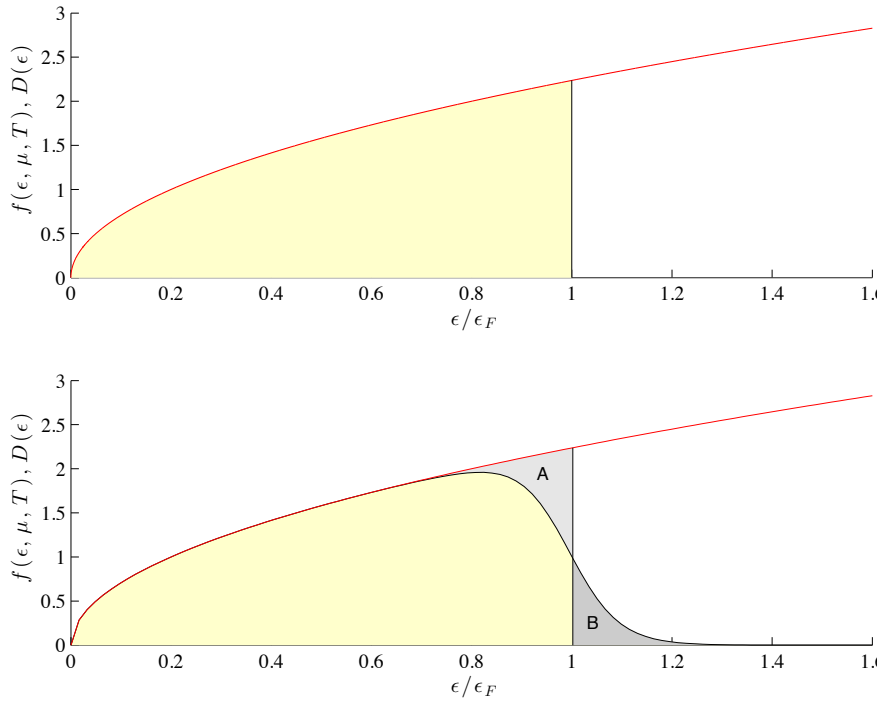


Fig. 10.7 (Top) At $T = 0$ all the states up to ϵ_F are filled. (Bottom) For $T > 0$ the particles in the states in area A have been moved into the the states B with higher energies for $kT = 0.25\epsilon_F$.

(area A), must be equal to the number of particles added to states above the chemical potential (area B) in order for the number of particles to remain constant.

Now, if we assume that the chemical potential μ does not change when T is increased, we see that area B will be larger than area A because $D(\epsilon)$ is an increasing function and $f(\epsilon;\mu,T)$ is symmetric around $\epsilon = \mu$. This is illustrated in the magnification shown in Fig. 10.8. The top figure shows what happens if we assume that the chemical potentia $\mu(T) = \epsilon_F$ also when $T > 0$. In this case the area B is larger than the area A. This is not that easy to see from the figures, since the differences are rather small.

How should $\mu(T)$ depend on T in order to increase the area A and decrease the area B? To get this argument right, it is important to realize what areas we really are to compare. We have illustrated this in the bottom figure in Fig. 10.8. Here we have illustrated the chemical potential by the dotted line. The chemical potential is smaller than ϵ_F in this

figure. We have also illustrated the relevant areas: It is the area below ϵ_F and above ϵ_F we should compare. (Not the area below and above μ – can you see the difference?). At $T = 0$ the area below $D(\epsilon)$ corresponds to the number of particles in the system.

In the top figure in Fig. 10.8 we have argued that area B is slightly larger than area A. In the bottom figure in Fig. 10.8 we see that area A is clearly larger than area B. This means that the correct value for μ lies somewhere between $\mu = \epsilon_F$ in the top figure and $\mu = 0.99 \cdot \epsilon_F$ in the bottom figure. This means that μ must be smaller than ϵ_F, but not as small as illustrated in this figure.

Conclusion: $\mu(T)$ must be a decreasing function of T near $T = 0$.

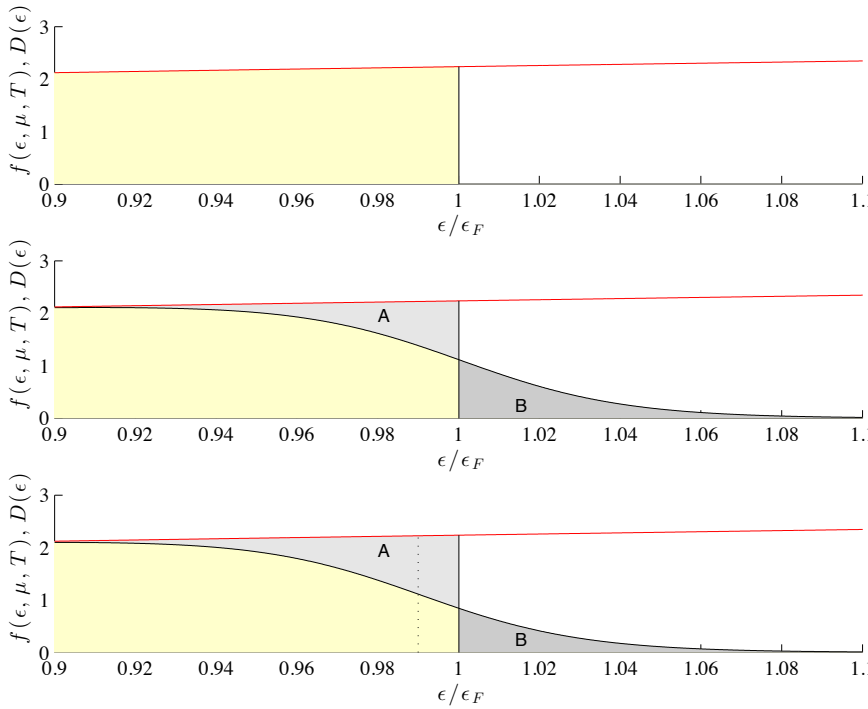


Fig. 10.8 Plot of $f(\epsilon; \mu, T = 0) \cdot D(\epsilon)$ (top), $f(\epsilon; \epsilon_F, T = 0.1T_F) \cdot D(\epsilon)$ (middle), and $f(\epsilon; \mu, T = 0.1T_F) \cdot D(\epsilon)$ (bottom) where $T_F = \epsilon_F/k$.

10.3.5 Quantitative description of low temperature behavior

We can also devise a simplified quantitative argument to estimate how the energy $\bar{E}$ depends on the temperature and then use this argument to find how μ depends on T near $T = 0$. The average energy of the system is

$$\bar{E} = \int_0^\infty \epsilon f(\epsilon; \mu, T) D(\epsilon) d\epsilon \, . \tag{10.86}$$

Let us now assume that the number of particles, N, is given so that $\mu = \mu(T)$. We can then find the heat capacity (at constant volume, since we still work in a T, V system) from

$$C_V = \left(\frac{d\bar{E}}{dT} \right) = \int_0^\infty \epsilon D(\epsilon) \frac{d}{dT} f(\epsilon; \mu(T), T) d\epsilon \, , \tag{10.87}$$

where there are two T-dependencies in f – both a direct dependency and an indirect dependency through $\mu(T)$:

$$\frac{df(\epsilon; \mu(T), T)}{dT} = \frac{\partial f}{\partial T} + \frac{\partial f}{\partial \mu} \frac{d\mu}{dT} \, . \tag{10.88}$$

This gives

$$C_V = \int_0^\infty \epsilon \frac{\partial f}{\partial T} D(\epsilon) d\epsilon + \frac{d\mu}{dT} \int_0^\infty \epsilon \frac{\partial f}{\partial \mu} D(\epsilon) d\epsilon \, . \tag{10.89}$$

We can do the same for $\bar{N}$:

$$\bar{N} = \int_0^\infty f(\epsilon; \mu(T), T) D(\epsilon) d\epsilon \, , \tag{10.90}$$

$$\frac{d\bar{N}}{dT} = \int_0^\infty \frac{\partial f}{\partial T} D(\epsilon) d\epsilon \, , \tag{10.91}$$

$$0 = \int_0^\infty \frac{\partial f}{\partial T} D(\epsilon) d\epsilon + \frac{d\mu}{dT} \int_0^\infty \frac{\partial f}{\partial \mu} D(\epsilon) d\epsilon \, . \tag{10.92}$$

(Where we have used that $(d\bar{N}/dT)$ is constant, since the number of particles in the system does not change when we change the temperature if the number of particles is constant). Up to now we have not done any approximations. Now, we see that f is a rapidly varying function of ϵ around μ:

$$\frac{\partial f}{\partial \mu} = \frac{e^{(\epsilon - \mu)/kT}}{\left(e^{(\epsilon - \mu)/kT} + 1 \right)^2} \, . \tag{10.93}$$

We can therefore approximate the second integral in (10.89) by

$$\int_0^\infty \epsilon D(\epsilon)\frac{\partial f}{\partial \mu}d\epsilon \simeq \mu \int_0^\infty D(\epsilon)\frac{\partial f}{\partial \mu}d\epsilon \ , \tag{10.94}$$

where the right hand side now corresponds to the second integral in (10.92). We can therefore replace

$$\mu int_0^\infty \frac{\partial f}{\partial \mu}D(\epsilon)d\epsilon = -\frac{\mu}{d\mu/dT}\int_0^\infty \frac{\partial f}{\partial T}D(\epsilon)d\epsilon + \ . \tag{10.95}$$

We replace this into the equation for C_V in (10.92) getting

$$C_V \simeq \int_0^\infty \epsilon\frac{\partial f}{\partial T}D(\epsilon)d\epsilon - \mu\int_0^\infty \frac{\partial f}{\partial T}D(\epsilon)d\epsilon = \int_0^\infty (\epsilon-\mu)\frac{\partial f}{\partial T}D(\epsilon)d\epsilon \ . \tag{10.96}$$

We will use the same approximation once more, but this time argue that $\partial f/\partial T$ falls off rapidly with ϵ away from μ, but that $D(\epsilon)$ varies slowly, so that we can assume that $D(\epsilon)$ is approximately constant and equal to $D(\mu)$, simplifying the integral to:

$$C_V \simeq \int_0^\infty (\epsilon-\mu)\frac{\partial f}{\partial T}D(\epsilon)d\epsilon \simeq D(\mu)\int_0^\infty (\epsilon-\mu)\frac{\partial f}{\partial T}d\epsilon \ . \tag{10.97}$$

Here, we need $\partial f/\partial T$:

$$\frac{\partial F}{\partial T} = \frac{(\epsilon-\mu)}{kT^2}\frac{e^{(\epsilon-\mu)/kT}}{\left(e^{(\epsilon-\mu)/kT}+1\right)^2} \ . \tag{10.98}$$

Now, we are near an integral we can solve. We introduce a new integration variable, $x = (\epsilon-\mu)/kT$, so that $d\epsilon = kTdx$, which gives:

$$C_V \simeq D(\mu)k^2T\int_{-\mu/kT}^\infty \frac{x^2e^x}{\left(e^x+1\right)^2}dx \ , \tag{10.99}$$

where again the integral varies rapidly around $x = 0$, and the lower bound therefore can be set to be approximately $-\infty$, and we can solve the integral exactly:

$$C_V \simeq D(\mu)k^2T\int_0^\infty \frac{x^2e^x}{\left(e^x+1\right)^2}dx = k^2TD(\mu)\frac{1}{3}\pi^2 \ . \tag{10.100}$$

We may also approximate $D(\mu) \simeq D(\epsilon_F)$, which gives

$$D(\mu) \simeq D(\epsilon_F) = \frac{3N}{2}\frac{1}{\epsilon_F} = \frac{3}{2}\frac{1}{NkT_F} , \tag{10.101}$$

and therefore

$$C_V \simeq \frac{1}{2}\pi^2 Nk\frac{T}{T_F} = aNT , \tag{10.102}$$

where $kT_F = \epsilon_F$, that is, $T_F = \epsilon_F/k$, and a is a constant that does not depend on N or T. A typical values for T_F for an electron gas is $T_F \simeq 5 \cdot 10^4$K.

This long derivation shows that the heat capacity due to an electron gas (or another Fermi gas) is proportional to T at low temperatures (T small compared with T_F).

10.4 Photon Systems

Let us now use the theory we have developed to address the behavior of a photon gas - a system with electromagnetic waves inside a container – in equilbrium with the container walls. In this case the walls will absorb and emit electromagnetic wave (packets) - so the number of photons is not conserved.

For a container of length L, we assume that the field only consists of standing waves (in equilbrium). These waves can therefore have frequencies $f = c/\lambda$ and corresponding energies $\epsilon = nhf = n\hbar\omega$, where n is an integer corresponding to the state.

(The integer n comes from the number of half-wavelengths that make up the length L: $L = (n/2)\lambda$, where n is an integer.)

This corresponds to the harmonic oscillator we have already studied.

We will now study the occupancy probability for these states. We can assume that the system has constant volume V and temperature T, and that since photons are continually created and destroyed - the chemical potential for the system is therefore zero.

We can then use the Bose-Einstein distribution law:

$$f_{BE}(\epsilon, \mu, T) = \frac{1}{e^{(\epsilon-\mu)/kT} - 1} , \tag{10.103}$$

which we can simplify since $\mu = 0$ to

$$f_{BE}(\epsilon, \mu = 0, T) = \frac{1}{e^{(\epsilon/kT} - 1} , \tag{10.104}$$

which is the same as we found for the harmonic oscillator in the canonical ensemble.

This tells us the occupancy of a given energy level ϵ_s. In addition we need to know the density of states for the photon gas.

For particles (fotons) in a box, the possible (standing wave) solutions are

$$\lambda = \frac{2L}{n} \ , \ \frac{hn}{2L} \ , \tag{10.105}$$

where n is a positive integer. The energy of the photon is

$$\epsilon = pc = \hbar\omega = \frac{hcn}{2L} \ , \tag{10.106}$$

instead of $\epsilon = p^2/2m$ for classical particles.

This means that the energy of a photon generally is proportional to n while the energy of a classical moving particles (in an ideal gas / Fermi gas) is proportional to n^2.

This is also true in three dimensions, where the momentum is independent in the three direction, and equal to $hc\,n/2L$ in each direction:

$$\mathbf{p} = \frac{hc}{2L}(n_x, n_y, n_z) \ , \tag{10.107}$$

and the energy still is $\epsilon = pc$, where p is the magnitude of $\mathbf{p}$:

$$\epsilon = \frac{hc}{2L}\left(n_x^2 + n_y^2 + n_z^2\right)^{1/2} = \frac{hcn}{2L} \ . \tag{10.108}$$

In order to use the distribution function, we need to sum over all possible states (and their corresponding energies) to find the number of photons and the total energy of the photon gas.

Let us first look at the total energy - which we find by summing over all possible n-values - including the effect of two possible polarizations:

$$U = 2\sum_{n_x}\sum_{n_y}\sum_{n_z} \epsilon(n_x, n_y, n_z) f_{BE}(\epsilon, \mu = 0, T) \ , \tag{10.109}$$

where $\epsilon = hcn/L$. Instead of summing, we integrate over n in n-space:

$$U = 2\int_0^\infty \frac{4\pi n^2}{8}\,\frac{hcn}{L}\,\frac{1}{e^{hcn/2LkT} - 1}\,dn \ . \tag{10.110}$$

We can now transform to a variable that simplifies the exponent, choosing $x = hcn/2LkT$, which gives $n = 2LkTx/hc$ and $dn = 2LkT/hc\,dx$:

$$U = \frac{8\pi(kT)^4 L^3}{(hc)^3} \int_0^\infty \frac{x^3}{e^x - 1} dx \ , \tag{10.111}$$

where the integral is $\pi^4/15$. Energy per volume is therefore then

$$\frac{U}{V} = \frac{8\pi^5}{15h^3c^3}(kT)^4 \ . \tag{10.112}$$

This is called the **Stehan-Boltzmann** law of radiation.

We can also find the frequency (or energy, since they are propotional $\epsilon = h\nu$) distribution of the photon gas / blackbody radiation.

What does that mean? We can find for example how many photons are in the range from ν to $\nu + d\nu$, how many photons are in the range from ϵ to $\epsilon + d\epsilon$ - or preferrably – how much energy is in the corresponding range of frequencies / energies: We find this as the number of photons in the range multiplied by the energy per photon, ϵ.

We can read this directly from the integral, realizing that the integral for the total energy can be written as

$$U = \int_0^\infty \epsilon_n D(n) dn = \int_0^\infty \epsilon D(\epsilon) d\epsilon \ , \tag{10.113}$$

or if we are interested in the energy density (U/V) instead:

$$\frac{U}{V} = \int_0^\infty \frac{\epsilon_n}{V} D(n) dn = \int_0^\infty \frac{\epsilon}{V} D(\epsilon) d\epsilon \ , \tag{10.114}$$

which is what we just found above.

Starting from

$$U = 2\int_0^\infty \frac{4\pi n^2}{8} \frac{hcn}{L} \frac{1}{e^{hcn/2LkT} - 1} dn \ . \tag{10.115}$$

we can instead insert $\epsilon = hcn/2L$, getting

$$\frac{U}{V} = \int_0^\infty \frac{8\pi\epsilon}{(hc)^3 e^{\epsilon/kT} - 1} d\epsilon \ , \tag{10.116}$$

and we see that the energy density in the range ϵ to $\epsilon + d\epsilon$ is

$$u(\epsilon) = \frac{8\pi}{(hc)^3} \frac{\epsilon^3}{e^{\epsilon/kT} - 1} \ . \tag{10.117}$$

We can express it instead using ν where $\epsilon = h\nu$:

$$u(\nu) = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/kT}-1}d\nu \ . \tag{10.118}$$

This is called the **Planck radiation law**.

We can plot this distribution in dimensionless form by plotting $x^3/(e^x - 1)$ as a function of $x = \epsilon/kT$.

This function has a maximum at $x = 2.82$, which corresponds to $\epsilon = 2.82\,kT$.

This shows that the maximum energy (or frequency) increases with temperature. This law is called **Wien's law**. (Wien's displacement law).

This law implies that the temperature can been seen from the most prominent color - since the frequency of this this color is proportional to the temperature. (If we can consider the body we examine as blackbody radiation).

10.4.1 Entropy of the photon gas

We can find the entropy, since for constant volume we have

$$TdS = dU \ , \ dS = \frac{dU}{T} = \frac{4aT^3}{T} \ , \tag{10.119}$$

which we integrate from 0 to T, getting:

$$S(T) = 4a\frac{1}{3}T^3 = \frac{32\pi^5}{45}V\left(\frac{kT}{hc}\right)^3 k \ . \tag{10.120}$$

(We can find the total number of photon using the same formula, but with a different prefactor).

10.4.2 Radiation from a black body

What if we have a photon gas and we study the radiation emitted from a small whole in the container.

We say that such a hole radiates as a black body.

The energy flux is the rate of energy emission per unit area.

How much radiates out in a time dt?

The energy in the volume dV which is cdt times the opening area A times some geometric factor g (which is 1/4 - see textbook and problems).

The flux is therefore

$$J = \frac{dE}{Adt} = \frac{udV}{Adt} = \frac{uAcdtg}{Adt} = ucg \ . \tag{10.121}$$

and

$$J = \frac{cU}{4V} = \frac{2\pi^5}{15}\frac{(kT)^4}{(hc)^3} = \sigma T^4 \ , \tag{10.122}$$

where σ is called the **Stefan-Boltmann** constant.

And the law is called **Stefan's law**.

This is from a black body - non-reflecting! - body with surface temperature T.

Applications: Cosmic black body background radiation, Emission and absorption (Kirchoffs law).

10.5 Phonon Systems

We have previously introduced the Einstein model - or the ideal crystal - as our model for a crystal. But in this system there are no interactions between the atoms/particles. Does this have any consequence?

For the Einstein crystal we found that the heat capacity was:

$$C_V = 3Nk\frac{\left(\frac{\epsilon}{kT}\right)^2 e^{\epsilon/kT}}{\left(e^{\epsilon/kT}-1\right)^2} \ . \tag{10.123}$$

where N now is the number of atoms/particles, so that $3N$ is the number of oscillators.

What happens when $\epsilon \ll kT$ for this model? In that range the heat capacity approaches a constant: $3Nk$. This corresponds to the equipartition theorem.

What happens for small temperatures, when $\epsilon \gg kT$? In this range the heat capacity goes to zero exponentially. This is not in correspondence with experiments, which show that the heat capacity goes to zero at T^3 in the low temperature limit.

What went wrong?

We have ignored interactions among the atoms - which probably may be important for lattice vibrations.

Mainly because there are different vibration modes in systems with many atoms - the atoms may vibrate together to form low frequency modes. At low temperatures the high frequencies (high energy) modes are frozen out, but then the low frequency (low energy) modes become more important. By only looking at single atoms we have overlooked

collective modes that have lower energy. Therefore the heat capacity goes to zero slower than predicited by the Einstein model.

The vibration modes in a crystal resembles electromagnetic waves:

They are waves, but they have smaller velocities (much smaller of course). We will here assume that the speed is a constant c_s - even if it acutally depends on the wave length in a crystal.

They have three polarizations. The transverse and one longitudinal. The polarizations really have different velocities. At first we will ignore this effect.

The waves cannot have all possible wavelengths, because the atoms are on a lattice with a given lattice spacing, and we cannot resolve waves with wavelength shorter than (two times) the lattice spacing.

Let us assume that we can otherwise describe a phonon - a lattice vibration mode energy packet - just as we have described photons: with uniformly spaced energy levels:

$$\epsilon_s = h\nu = \frac{hc_s}{\lambda} = \frac{hc_s n}{2L} \, , \tag{10.124}$$

where L is the system size. Again n is the magnitude of the $\mathbf{n}$ vector in n-space.

We also assume that phonons are bosons with $\mu = 0$, so that the distribution function is given by the Bose-Einstein distribution:

$$f_{BE} = \frac{1}{e^{\epsilon/kT} - 1} \, . \tag{10.125}$$

We find the energy and the number of phonons by summing over all possible $\mathbf{n}$ values:

$$U = 3 \sum_{n_x} \sum_{n_y} \sum_{n_z} \epsilon(\mathbf{n}) f(\epsilon, \mu = 0, T) \, . \tag{10.126}$$

The main difference with the crystal lattice is that not all values are possible - we can only have some values for $\mathbf{n}$.

Let us look at the x-direction.

Along this direction we have $N_x = (N)^{1/3}$ atoms as shown in the figure.

This puts at limit on the maximum upper number of n.

This should correspond to a cube in n-space.

The Debye approximation is to assume that we instead include all modes up to a radius n_D in n-space, but so that we ensure that the total number of modes is equal to $3N$:

$$3\sum_{n_x}\sum_{n_y}\sum_{n_z} = \frac{3}{8}\int_0^{n_D} 4\pi n^2 dn = 3N \ , \tag{10.127}$$

This determines n_D. Since the integral is simply the volume of a sphere of radius n_D we find that

$$3N = \frac{3}{8}\frac{4\pi}{3}n_D^3 \ , \tag{10.128}$$

and therefore that

$$n_D = \left(\frac{6N}{\pi}\right)^{1/3} \ . \tag{10.129}$$

Now, we have a theory we can use to find the energy of the phonon gas using the approach we now are used to:

$$\begin{aligned} U &= \sum_{n_x}\sum_{n_y}\sum_{n_z} \epsilon_n f(\epsilon_n, \mu = 0, T) \\ &= \frac{3}{8}\int_0^{n_D} n^2 \frac{h\nu_n}{\exp(h\nu_n/kT) - 1} dn \ , \end{aligned} \tag{10.130}$$

where

$$h\nu_n = \frac{hc_s n}{2L} \ . \tag{10.131}$$

The integral is therefore

$$U = \frac{3\pi}{2}\int_0^{n_D} \frac{hc_s}{2L}\frac{n^3}{\exp(hc_s n/2LkT) - 1} dn \ . \tag{10.132}$$

We introduce the new integration variable $x = hc_s n/2LkT$ getting:

$$x_D = \frac{hc_s n_D}{2LkT} = \frac{hc_s}{2kT}\left(\frac{6N}{\pi V}\right)^{1/3} = \frac{T_D}{T} \ , \tag{10.133}$$

where we call T_D the Debye temperature.

This gives - after some algebra:

$$U = \frac{9NkT^4}{T_D^3}\int_0^{T_D/T} \frac{x^3}{e^x - 1} dx \ . \tag{10.134}$$

This integral cannot be solved analytically, but it is easy to solve numerically.

However, we can also find the high and low temperature limits directly by approximations.

In the high temperature limit the upper bound of the integral is much smaller than 1 and in this limit we can approximate $e^x = 1 + x$, and the integral becomes

$$\begin{aligned} U &= \frac{9NkT^4}{T_D^3}\int_0^{T_D/T}\frac{x^3}{x}dx \\ &= \frac{9NkT^4}{T_D^3}\frac{1}{3}\left(\frac{T_D}{T}\right)^3 \\ &= 3NkT\ . \end{aligned} \tag{10.135}$$

This corresponds to the equipartition principle.

In the lower limit, $T \ll T_D$, the upper limit is so large that we can replace it by infinity (since the e^x will go very quickly to zero, this is not a problem). The integral is therefore the same as we did for the photon gas - with value $\pi^4/15$, and the total energy is

$$U = \frac{3\pi^4}{5}\frac{NkT^4}{T_D^3}\ , \tag{10.136}$$

and the heat capacity in this limit is

$$C_V = \frac{12\pi^4}{5}\left(\frac{T}{T_D}\right)^3 Nk\ , \tag{10.137}$$

which agrees very well with experimental measurements.

We find the intermediate values by numerical integration.

For metals we need to include both the contribution from the phonos and the contributions from the electrons, so that the heat capacity has the behavior:

$$C_V = \gamma T + \frac{12\pi^4 Nk}{5T_D^3}T^3\ , \tag{10.138}$$

when $T \ll T_D$ and $\gamma = \pi^2 Nk^2/2\epsilon_F$.

If we now plot C_V/T as a function of T^2 we can check both constants in the resulting plot, which should be linear in T^2 with an intercept corresponding to γ.

What are typical values for T_D?

For lead 88K

For diamond 1860K

Above T_D you can get away with using the equipartition theorem since the heat capacity by then has reached 95% of its maximum value.

10.6 Boson Gas and Einstein Condensation

We have now looked at photons and phonons. For these particles we do not have a conservation law. What about other bosons, such as Helium-4? In this case the chemical potential is not zero - we will need to determine the chemical potential from the number of particles, just as we did for Fermions.

We start in the limit of small temperatures.

When $T = 0$ all the atoms will be in the lowest energy state.

For atoms in a box of volume $V = L^3$ the energy states are

$$\epsilon = \frac{h^2}{8mL^2}\left(n_x^2 + n_y^2 + n_z^2\right) \ , \tag{10.139}$$

so that the lowest energy state has energy

$$\epsilon_0 = \frac{3h^2}{8mL^2} \ , \tag{10.140}$$

which is a very small value for realistic (macroscopic) values of L.

The Bose-Einstein distribution gives the average number of atoms in this state:

$$N_0 = \frac{1}{\exp((\epsilon_0 - \mu)/kT) - 1} \ , \tag{10.141}$$

As the temperature approaches zero, we know that N_0 will be large, which means that $\exp((\epsilon_0 - \mu)/kT) - 1$ must be close to 1. This means that the exponent is close to zero. In this limit, we can expand the exponent using $\exp(x) = 1 + x$, getting:

$$N_0 = \frac{1}{1 + (\epsilon_0 - \mu)/kT - 1} = \frac{kT}{\epsilon_0 - \mu} \ , \tag{10.142}$$

and therefore that $\mu = \epsilon_0$ when $T = 0$ and then just a bit larger when T is small.

To make life simpler, let us change energy scales, so that $\epsilon_0 = 0$. The result is then

$$N_0 = -\frac{kT}{\mu} \ , \ \mu = -\frac{kT}{N_0} \ . \tag{10.143}$$

10.6.1 Spacing of levels

What is really a realistic value for the spacing of energy levels?

The energy is

$$\epsilon = \frac{h^2}{8mL^2}\left(n_x^2 + n_y^2 + n_z^2\right) \tag{10.144}$$

the energy difference between the lowest and the second lowest is therefore

$$\Delta\epsilon = \frac{3h^2}{8mL^2} \tag{10.145}$$

Now if we look at Helium-4, $m = 6.6 \times 10^{-24}$g and $L = 1$cm, then

$$\frac{\Delta\epsilon}{k} = 1.8 \times 10^{-14}\text{K} \, . \tag{10.146}$$

This is a small splitting!

How could this play an important physical role at temperatures which are Kelvin or at best a thousands of a Kelvin?

10.6.2 Approaching $T = 0$

What happens if we have a constant particle number, N, and we lower the temperature? Then the number of particles in the lowest leverl, N_0, approaches N - all the particles are in the state with the lowest energy.

This means that

$$N_0 = \frac{1}{\exp((\epsilon_0 - \mu)/kT) - 1} \, , \tag{10.147}$$

approaches N (which is very large) when $\mu \simeq \epsilon_0$.

For low temperatures we separate out the number of particles in the lowest state, and write

$$N = N_0 + \sum_{j=1}^{\infty} \frac{1}{\exp((\epsilon_j - \mu)/kT) - 1} \, , \tag{10.148}$$

where we have removed the singularity in the lowest energy level.

We can change this to an integral (since the energy levels are very close), but then we need the density of states, $D(\epsilon)$:

$$N = N_0 + \int_{\epsilon_1}^{\infty} \frac{g(\epsilon)}{\exp((\epsilon - \mu)/kT) - 1} d\epsilon \, , \tag{10.149}$$

Since ϵ_1 also is very small ($\epsilon_1 \simeq 0$) and $g(0) = 0$, we can instead put 0 as the lower bound for the integral. In addition, we still assume that

$\mu = 0$:

$$N = N_0 \int_0^\infty \frac{g(\epsilon)d\epsilon}{e^{\epsilon/kT} - 1} \,, \tag{10.150}$$

where we now insert for $g(\epsilon)$:

$$g(\epsilon) = \frac{2}{\sqrt{\pi}} \left(\frac{2\pi m}{h^2}\right)^{3/2} V\sqrt{\epsilon} \,, \tag{10.151}$$

and we introduce the variable $x = \epsilon/kT$. The integral is then reduced to

$$N = N_0 + \frac{2}{\sqrt{\pi}} \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V \int_0^\infty \frac{x^{1/2}dx}{e^x - 1} \,, \tag{10.152}$$

where the integral is

$$\int_0^\infty \frac{x^{1/2}}{e^x - 1}dx = \frac{\sqrt{\pi}}{2} \cdot 2.612 = 2.315 \,, \tag{10.153}$$

and therefore we find

$$N = N_0 + 2.612 \left(\frac{2\pi mkT}{h^2}\right)^{3/2} V \,. \tag{10.154}$$

This expression is only valid for low temperatures, but the second term increases with temperature. What is the upper limit of validity of this expression? That is when N_0 is zero and all the particles are in higher energy levels, this occurs at a temperature T_C given by:

$$N = 2.612 \left(\frac{2\pi mkT_C}{h^2}\right)^{3/2} V \,, \tag{10.155}$$

which gives

$$kT_C = \frac{1}{\pi(2.612)^{2/3}} \frac{h^2}{2m} \left(\frac{N}{V}\right)^{2/3} \,. \tag{10.156}$$

The critical temperature T_C therefore depends both on details of the particles, through m, and on the density N/V of the system. (for low temperatures).

We can also use T_C to rewrite the expression for N:

$$N = N_0 + N \left(\frac{T}{T_C}\right)^{3/2} \,, \tag{10.157}$$

which gives

$$N_0 = N\left(1 - \left(\frac{T}{T_C}\right)^{3/2}\right) \ . \tag{10.158}$$

In the range $T < T_C$ we have Bose-Einstein condensation. At $T = 0$ all the particles are in the lowest energy state.

What happens when $T < T_C$? In this case our calculation is no longer valid. Instead we must include how the chemical potential varies with temperature. We can do this by solving the equation

$$N = \int_0^\infty \frac{g(\epsilon)}{e^{(\epsilon-\mu)/kT} - 1} d\epsilon \ , \tag{10.159}$$

numerically. We introduce new variables, $x = \epsilon/kT_C$, $t = T/T_C$ and $c = \mu/kT_C$, and get the integral

$$2.315 = \int_0^\infty \frac{x^{1/2} dx}{e^{(x-c)/t} - 1} \ , \tag{10.160}$$

which you now know how to solve.

References

[1] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 1953.

[2] C. N. Yang. The spontaneous magnetization of a two-dimensional Ising model. *Physical Review*, 85, 1952.

Index

addition rule, 67
algorithmic model, 20
arrow of time, 20
`ase`, 43
atomic potential, 15
attractive interactions, 349
average, 68
average energy, 214
average value, 212

Bernoulli process, 85
binomial distribution, 85
Boltzmann statistics, 203

canonical system, 203
Central limit theorem, 106
chemical potential, 312
Clausius' law, 348
Clausius-Clapeyron's equation, 339
complement rule, 66
correlated, 61
critical point, 338
cumulative distribution, 116

data-collapse, 157
detailed balance, 147
diffusive equilibrium, 312
dilute solutions, 360

Einstein crystal, 141, 203
engines, 284
entropy of expansion, 174
entropy of mixing, 174
environment, 281
equation of state, 282
equation of state for an ideal gas, 127
estimator, 63
evaporation, 338
event, 61
excluded volume, 348
exponential distribution, 99
extensive, 282
extensive properties, 216
extensive property, 176
extreme value statistics, 108

face-centered cubic lattice, 34
fair, 61

fcc, 34
fluctuations, 215
free energy, 225
free expansion, 174
frequency, 63
fundamental assumption, 147

Gaussian distribution, 92
Gaussian function, 164
Gibb's paradox, 176
Gibbs free energy, 325

heat capacity, 215
heat engines, 284
Helmholtz free energy, 225

ideal crystal, 141
ideal gas, 124
ideal gas law, 126, 127
ideal mixture, 360
ideal solution, 360
independent, 67
intensive, 282
interacting systems, 222
irreversible, 283
irreversible process, 283
Ising model, 222
isolated system, 203, 281

LAMMPS, 41, 42
lammps trajectory file, 49
latent heat, 336
law of corresponding states, 354
Legendre transform, 327
Lennard-Jones potential, 15, 33
London interaction, 15

macrostate, 26, 79
macrostates, 132
MD, 32
mean-field theory, 349
melting, 338
microcanocical system, 203
microstate, 26, 79
microstates, 132
minimum free energy principle, 327
mixture, 360
Molecular Dynamics, 32
molecular dynamics, 13
Monte Carlo methods, 148
multiplicity, 79, 136
multiplicity function, 154

non-interacting particles, 124
Normal distribution, 106
normalization, 65, 66
normalization condition, 66
normalized probability, 65
numerical statistical experiment, 61

outcome, 61
Ovito, 48

partition function, 209, 210
periodic boundary conditions, 35
phase diagram, 338
probability, 60, 63
probability density, 95
probability distribution, 95
processes, 283

quasi-static, 284
quasi-static processes, 284

random number, 61
random variable, 67
Raoult's law, 366
repulsive interactions, 348
reversible, 283
reversible process, 283

sample space, 61, 67

second law of thermodynamics, 164
sequence of random numbers, 61
sharp multiplicity function, 164
sharpness of canonical system, 216
simple cubic lattice, 34
solute, 360
solution, 360
solvent, 360
span sample space, 67
standard deviation, 74
state variable, 282
statistical experiment, 61
Stirling's approximation, 89
sublimation, 338
system, 281

thermal contact, 281
thermal equilibrium, 282
thermally insulated, 281
thermodynamic limit, 156
thermodynamics, 281
triple-point, 338
two-state system, 213

uncorrelated, 61
uniform distribution, 99
unit cell, 34, 38
universal distributions, 105

van der Waals equation of state, 351
van der Waals force, 33
van der Waals interaction, 15
van der Waals interactions, 33
van der Waals model, 351
Vapor pressure, 324
variance, 74
Velocity Verlet, 36
VMD, 48