

Thermodynamics

William F. Barnes

June 13, 2014

Contents

1	Introduction	6
1.1	Temperature	6
1.2	Thermal Expansion	6
1.3	Thermal Stress	7
1.4	Thermal Energy	7
1.5	Specific Heat and Heat Capacity	8
1.6	Avogadro's Number and Mass Units	8
1.7	Phases of Matter, Phase Diagrams	9
1.8	PT Diagram	9
1.9	PV Diagram	10
1.10	Quasistatic Approximation	10
1.11	Phase Change and Latent Heat	10
2	Heat Transfer	12
2.1	Conduction	12
2.2	Convection	13
2.3	Electromagnetic Radiation	13
	Wandering in the Desert	14
3	Ideal Gas Theory	15
3.1	Ideal Gas Law	15
3.2	Boyle's Law	15
3.3	Charles's Law	15
3.4	Lord William Kelvin	15
3.5	Avogadro's Law	15
3.6	Ideal Gas Model	16
3.7	Mixtures and Dalton's Law	16

3.8	Vapor Pressure, Humidity, Clouds	16
4	Kinetics of Ideal Gas	18
4.1	Pressure	18
4.2	Internal Energy, RMS Speed	19
4.3	Maxwell Speed Distribution	20
4.4	Three Speeds	20
4.5	Equipartition Theorem	22
4.6	Diatomic Molecules	22
4.7	Intermolecular Collisions	23
4.8	The van der Waals Model	23
5	Particle Effusion	24
5.1	Statistical Calculation of Pressure	24
5.2	Effusion Rate	25
5.3	Effusion of Ideal Gas Particles	26
6	Heat Capacity of Gases and Solids	27
6.1	Ideal Gas at Constant Volume	27
6.2	Ideal Gas at Constant Pressure	27
6.3	Elemental Solids	28
7	First Law of Thermodynamics	28
7.1	Isothermal Process	28
7.2	Isochoric Process	29
7.3	Isobaric Process	29
7.4	Adiabatic Process	29
7.5	Processes in Ideal Gas	29
	Isothermal	29
	Isochoric	29
	Isobaric	29
	Adiabatic	30
8	Thermodynamic Cycles	31
8.1	Otto Cycle	32
8.2	Diesel Cycle	34
8.3	Carnot Cycle	35
8.4	Refrigerators	36
9	Reversible Processes	37
9.1	Free Expansion of Ideal Gas	38
9.2	Carnot Cycle Revisited	38
10	Entropy, Second Law of Thermodynamics	39
10.1	Entropy Calculations	39
	Free Expansion of Ideal Gas	39
	Isothermal Expansion of Ideal Gas	39
	Adiabatic Process in Ideal Gas	40
	Entropy of Light	40

10.2 Second Law of Thermodynamics	40
Entropy statement	40
Engine statement	41
Refrigerator statement	41
11 Systems in Contact, Equilibrium	41
11.1 Thermal Equilibrium	41
11.2 Mechanical Equilibrium	42
11.3 Diffusive Equilibrium	42
12 Thermodynamic Potentials	43
12.1 Enthalpy	43
12.2 Helmholtz Free Energy	43
12.3 Gibbs Free Energy	43
12.4 Grand Free Energy	43
12.5 Thermochemistry Tables	44
Electrolysis of Water	44
13 Thermodynamic Identities	45
13.1 First Thermodynamic Identity	45
13.2 Second Thermodynamic Identities	45
13.3 Thermodynamic Variables as Partial Derivatives	45
13.4 Maxwell Relations	46
13.5 Extensive and Intensive Quantities	46
13.6 Linearity between G and N	47
14 Equilibrium and Energy Minimization	47
14.1 Helmholtz Minimum	47
14.2 Gibbs Minimum	48
14.3 Grand Minimum	48
15 Introduction to Statistical Mechanics	48
15.1 Microstates and Macrostates and Multiplicity (oh my!)	49
16 Combinatorics	49
16.1 Party of Strangers	49
16.2 Four-Digit Passwords	50
16.3 Arrangements, Permutations, and Combinations	52
16.4 Two-State Coin	52
17 Probability Theory	54
17.1 Definitions	54
17.2 Questions in Probability Theory	55
The Birthday Problem	55
Laptop Repair Shop	56
18 Statistical Systems	57
18.1 Discrete and Continuous Distributions	58
18.2 Random Variables	59
18.3 Large Numbers and Stirling's Approximation	60

18.4 Gaussian Distribution	61
18.5 Binomial Distribution	62
18.6 Radioactive Decay	63
18.7 Poisson Distribution	64
19 Microcanonical Ensemble	65
19.1 Multiplicity and Entropy	65
19.2 Microcanonical Systems in Contact	65
19.3 Ergodic Hypothesis	67
More Mathematics	68
20 Paramagnetism	69
20.1 Thermodynamics of a Paramagnet	69
20.2 Negative Temperature	70
20.3 Temperature Reparameterization	71
21 Einstein Solid	71
21.1 Multiplicity of Einstein Solid	72
21.2 Thermodynamics of Einstein Solid	72
21.3 Combined Einstein Solids	73
22 Quantum Ideal Gas	73
22.1 Many Particles in a Box	74
22.2 Volume of a Hypersphere	74
22.3 Multiplicity of Ideal Gas	75
22.4 Entropy of Ideal Gas, Gibbs Paradox	76
22.5 Recovering Classical Results	76
22.6 Chemical Potential of Ideal Gas	77
22.7 Fugacity	77
22.8 Generalization and Other Dimensions	78
23 Canonical Ensemble	78
23.1 Minimization of Helmholtz Free Energy	78
23.2 Legendre Transform	79
23.3 Boltzmann Factor and Partition Function	79
23.4 Observables	80
23.5 Fluctuations	81
23.6 Non-Interacting Particles	82
Distinguishable Particles	82
Indistinguishable Particles	82
23.7 Equipartition Theorem Derived	82
24 Canonical Ideal Gas	83
24.1 Partition Function and Thermodynamic Potentials	83
24.2 Maxwell Speed Distribution Derived	84
24.3 Van der Waals Equation Derived	85
25 Simple Harmonic Oscillator System	85
25.1 Classical Simple Harmonic Oscillator	86
25.2 Quantum Simple Harmonic Oscillator	87

26 Two-State Systems	88
26.1 Quantum Two-State System	88
26.2 Rabi Oscillations	89
26.3 Statistical Two-State System	90
26.4 Entropic Rubber Band	91
Microcanonical Analysis	92
Canonical Analysis	92
27 Paramagnetism as Canonical Ensemble	93
27.1 Classical Paramagnet	93
27.2 Two-State Quantum Paramagnet	94
27.3 General Quantum Paramagnet	94
28 Grand Canonical Ensemble	95
28.1 Minimization of Grand Free Energy	95
28.2 Gibbs Factor and Grand Partition Function	96
28.3 Observables	97
28.4 Fluctuations	98
28.5 Ideal Gas as Grand Canonical Ensemble	98
29 Imaginary Time and Density Operator	99
29.1 Temperature as Imaginary Time	99
29.2 Density Operator	99
29.3 Spin 1/2 System	101
Diagonal Hamiltonian	101
Non - Diagonal Hamiltonian	102
29.4 Position Representation	104
29.5 Free Particle in a Periodic Box	104
29.6 Simple Harmonic Oscillator	105
29.7 Grand Canonical Ensemble	107
30 Quantum Statistics	107

1 Introduction

1.1 Temperature

Temperature is a scale used for judging the amount of ‘hotness’ of something. A precise scientific definition of temperature is hard to pin down, but going from intuition we can say:

Temperature is a quantity that eventually becomes the same for two objects after sustained contact.

You might know already that temperature relates to the vibration speed of atoms in matter - this is correct but doesn’t lend to a fundamental understanding of temperature. We won’t have a better definition of temperature for a while, but the point can be restated as the *zeroth law of thermodynamics*, which reads

If system A and system B are each in equilibrium with system C, then A and B are in equilibrium with each other.

By ‘equilibrium’ we mean the objects concerned have been in thermal contact for a sufficiently long time.

Problem 1: At atmospheric pressure (also known as *standard pressure*), liquid water boils at $212^{\circ}F$ and freezes at $32^{\circ}F$. On the centigrade scale, these temperatures are $100^{\circ}C$ and $0^{\circ}C$, respectively. Calculate the one temperature at which the Fahrenheit and centigrade scales are in agreement, assuming each scale is linear. Answer: $-40^{\circ}F = -40^{\circ}C$

1.2 Thermal Expansion

Objects tend to change their size (volume) in response to a change in temperature, a phenomenon named *thermal expansion*. Taking a one-dimensional example, suppose a thin metal rod of length L_0 has initial temperature T_0 . If an external heat source changes the rod’s temperature by ΔT , the length is observed to change by ΔL . Restoring the temperature to T_0 , the length restores to L_0 . It follows that a linear ‘law’ for thermal expansion can be written as

$$L = L_0 + \alpha L_0 \Delta T, \quad (1)$$

where the parameter α is the *linear thermal expansion coefficient*, measuring typically around 10^{-5} per degree centigrade. The temperature change ΔT may be positive or negative.

Thermal expansion also occurs in two- and three-dimensional objects. Taking a 2D example, consider a plate of dimensions $L_0 \times W_0$. With a change in temperature, the area of the plate becomes

$$A = L_0 W_0 (1 + \alpha \Delta T)^2 \approx L_0 W_0 (1 + 2\alpha \Delta T). \quad (2)$$

For small temperature displacements, $\alpha^2 \Delta T^2$ is a reasonably small contribution to thermal expansion and is ignored to first-order approximation. Thus the area scales linearly with temperature, with effective expansion factor 2α .

For three dimensions, the effective expansion constant is $\beta = 3\alpha$. If the volume of the object is V_0 , this means

$$V \approx V_0 (1 + \beta \Delta T). \quad (3)$$

Problem 2: A circular piece with diameter d is removed from a room-temperature sheet of aluminum. If the sheet is heated in an oven, how does the area of the hole change with temperature? Does the hole increase or decrease in area? Answer: $\Delta A = \pi d^2 \alpha \Delta T / 2$, increasing.

Problem 3: On a cool 4°C morning, a driver fills his aluminum fuel tank to the full capacity of 106.0 L. That evening, he checks the fuel level and finds 103.4 L remaining in the tank. The coefficient of volume expansion for gasoline is $9.5 \times 10^{-4}/^\circ\text{C}$, and for aluminum is $7.25 \times 10^{-5}/^\circ\text{C}$. What was the maximum temperature reached by the fuel tank during the day? Answer: 32.70°C

Problem 4: Suppose some linear material has a thermal expansion coefficient that varies with temperature such that $\alpha(T) = A + BT + CT^2$. Derive a formula for the length of the material as a function of temperature change. Answer:

$$L = L_0 (A\Delta T + B\Delta T^2/2 + C\Delta T^3/3)$$

1.3 Thermal Stress

An object confined in space may not be able to physically undergo thermal expansion (or contraction), in which case the object gains *thermal stress*. Suppose a thin metal rod of fixed length L_0 is held between two clamps at temperature T_0 with no initial forces on the ends of the rod. If the temperature changes, the rod ‘wants’ to obey equation (1), written here as

$$\frac{\Delta L}{L_0} = \alpha \Delta T,$$

where α is the linear thermal expansion coefficient.

Meanwhile, Young’s modulus of the rod is defined as $Y = (F/A)/(\Delta L/L_0)$, so we write

$$\frac{\Delta L}{L_0} = \frac{F}{AY},$$

where F is the linear tension in the rod (positive or negative) and A is the cross-sectional area. Since ΔL must be zero, we therefore have

$$\frac{F}{A} = -Y\alpha\Delta T. \quad (4)$$

1.4 Thermal Energy

The phenomenon of ‘heat’ readily manifests as a consequence of some other type of energy expense, whether it be mechanical, chemical, electrical, or nuclear. We define *thermal energy*, denoted Q , as the broad subset of energies and potentials that couple strongly to temperature. Said another way, heat is the energy that flows between systems that are not in equilibrium. Like all other energies in physics, thermal energy must be conserved.

Sir James Joule (1818-1889) found experimentally that the amount of work required to raise the temperature of water is directly proportional to the temperature rise. One *calorie* is the amount of energy required to raise the temperature of one gram of water by one degree centigrade. One food calorie, denoted with a capital C , is 1000 ordinary calories. One calorie equals 4.186 Joules. One *BTU*, or British Thermal Unit, is $252\text{cal} = 1055\text{ J}$.

1.5 Specific Heat and Heat Capacity

A thermal energy quantity dQ transferred to a substance of mass m (without changing the phase) is linearly proportional to a small change in temperature by

$$c = \frac{1}{m} \frac{dQ}{dT}, \quad (5)$$

where the proportionality constant c is the *specific heat* of the substance. Equation (5) can often be integrated in the variable T , giving

$$Q = mc\Delta T \quad (6)$$

(an abuse of notation but common in thermodynamics culture). A typical solid metal has c around 300 J per kilogram of material per degree centigrade. In the same units, liquid water has $c = 4187$, frozen water has $c = 2108$, and water vapor has $c = 1996$.

The amount of thermal energy per degree centigrade that a substance can hold equals the specific heat times the total mass m , written as

$$mc = nC. \quad (7)$$

The constant C is called the *heat capacity*. The dimensionless constant n is the number of *moles* of substance considered, where one mole or *mol* is defined as 6.022×10^{23} particles. We'll see subsequently that heat capacity comes in two flavors, denoted 'constant volume' and 'constant pressure'. Values of C_P and C_V , respectively, are often reported in units of J per degree centigrade per mole.

Another term encountered in chemical physics is the *molar mass*, denoted M , and defined as

$$M = \frac{m}{n}. \quad (8)$$

1.6 Avogadro's Number and Mass Units

The conversion factor 6.022×10^{23} between the number of particles and the number of moles (of any substance) is called *Avogadro's Number*, denoted N_A .

By definition, one mole of Carbon-12 contains N_A atoms and weighs precisely 12 g . From this we define the *unified atomic mass unit*, which is a single number used to approximate both the proton mass and the neutron mass. The unified atomic mass unit is $1.660539 \times 10^{-24} g$. In 'ordinary' or *physical units*, the unified *amu* is roughly 0.9928 proton masses or 0.9914 neutron masses.

The *molecular mass* is defined as the number of *atomic mass units* (masses of protons and neutrons) in an element or molecule. As written in the periodic table, molecular mass is never a whole number for a given element - this is *not* an artifact of unified atomic mass units, nor are the electrons responsible. Rather, 'pure' elements occur in varieties called *isotopes*, where atoms of the same element have a different neutron count and different mass. Reported molecular masses are a weighted statistical average across the isotopes of an element.

For example, 99.2% of hydrogen atoms observed in nature contain one proton and zero neutrons, denoted 1_1H . Meanwhile, 0.8% of hydrogen occurs with one proton and one neutron, denoted 2_1H . Statistically 0% of hydrogen contains two neutrons per atom, meaning 3_1H is exceedingly rare. The weighted average of the respective masses per isotope is therefore

$$0.992 \times (1 \text{ amu}) + 0.008 \times (2 \text{ amu}) + 0 = 1.008 \text{ amu},$$

matching the reported molecular mass of hydrogen. Other elements are reported in the same fashion.

Problem 5: An H_2O molecule consists of 10 protons, 8 neutrons, and 10 electrons. Use the periodic table values $m_H = 1.008$ and $m_O = 15.999$ to compute the mass of a single water molecule in grams. Answer: $2.99 \times 10^{-23} g$.

1.7 Phases of Matter, Phase Diagrams

Phase is a term synonymous with ‘state of matter’ used for classifying the overall spatial arrangement and dynamics of the particles in a substance. Common phases of matter are solid, liquid, and gas - but the list doesn’t stop there. More exotic phases such as plasmas and liquid crystals occur in nature and the laboratory. It turns out that most (if not all) matter has different phases of existence; an extreme example being the duality between volcanic rock (solid state) and lava (liquid state).

The phase of a given substance is chiefly determined by (i) the surrounding temperature, and (ii) the surrounding pressure. This is visualized by using a *phase diagram*, which comes in two flavors: PT and PV , standing for pressure-temperature or pressure-volume, respectively.

1.8 PT Diagram

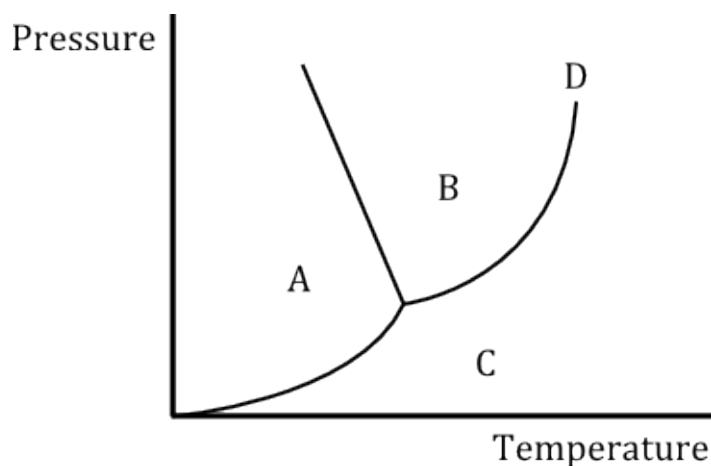


Figure 1: Typical PT phase diagram.

Figure 1 illustrates a typical PT diagram for a substance (water for example). A small enough ‘parcel’ of a substance occupies just one point in the phase diagram at a given time, and the parcel’s ‘trajectory’ in a phase diagram is continuous over time, but not necessarily smooth. Boundaries between regions in the phase diagram are associated with abrupt changes in the phase.

- Region *A* has low temperature and high pressure, corresponding to the solid state.
- Region *B* has moderate temperature and moderate pressure, corresponding to the liquid state.

- Region *C* has high temperature and low pressure, corresponding to the gaseous state.
- Region *D* indicates the supercritical state, a phase mixture of liquid and gas.
- Solid lines are boundaries that separate two phases.
- The junction of regions *A*, *B*, *C* is called the *triple point*.

1.9 PV Diagram

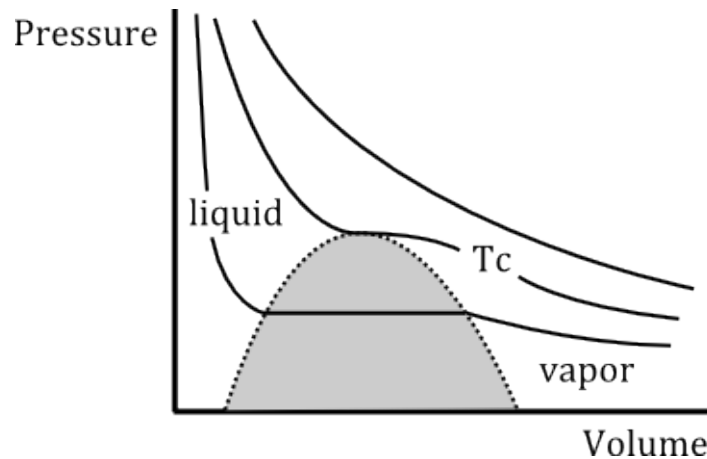


Figure 2: Typical *PV* phase diagram.

A variation of the *PT* diagram is the *PV* diagram, where the volume replaces temperature on the horizontal axis. Figure 2 illustrates a typical *PV* diagram for a substance (water for example, excluding the frozen state). The solid curves are called *isotherms*, which are lines of constant temperature. The shaded area under the dotted arch represents the liquid-vapor equilibrium phase. The isotherm T_C is the *critical temperature*.

1.10 Quasistatic Approximation

The *quasistatic approximation* is the limit when thermodynamic processes occur infinitely slowly. In such a case, *the whole substance occupies one point in the phase diagram at a given time*.

1.11 Phase Change and Latent Heat

Changing the temperature or pressure of a substance causes some shift in the phase diagram. Crossing the boundary between any two regions (see Figure 1) corresponds to a *phase change*. At such a boundary, molecular order of the substance changes radically, which involves an energy change of *chemical* origin.

When a substance is *on* a phase boundary, all energy added or removed from the substance contributes to phase change with the temperature remaining constant. To illustrate, consider a glass containing an ice + water mixture with initial temperature precisely at the freezing point of water, 0°C . When the glass is placed in a room-temperature environment, the mixture steadily absorbs heat from the surroundings, but

the temperature does not go above 0°C ! Rather, all energy contributes to melting the ice. Only when the ice is melted will the system's temperature begin to increase. The same experiment works in reverse: if instead the ice + water mixture were placed in a frigid -100°C environment, the temperature would remain 0°C until all liquid water phase changes to ice.

The thermal energy of a substance is not governed by (5) during phase change because ΔT is zero by construction. Instead, the change in energy is called *latent heat*, which is proportional to the mass of material being changed, or

$$Q = mL_\alpha. \quad (9)$$

The amount of mass that changes phase per unit energy is a constant denoted L_α , where subscript α denotes which phase boundary is being crossed.

Subscript $\alpha = v$ stands for ‘vaporization’, corresponding to the liquid-gas boundary, and $\alpha = f$ stands for ‘fusion’, corresponding to the solid-liquid boundary. Water at atmospheric pressure has $L_v = 2.256 \times 10^6 \text{ J/kg}$ and $L_f = 3.34 \times 10^5 \text{ J/kg}$.

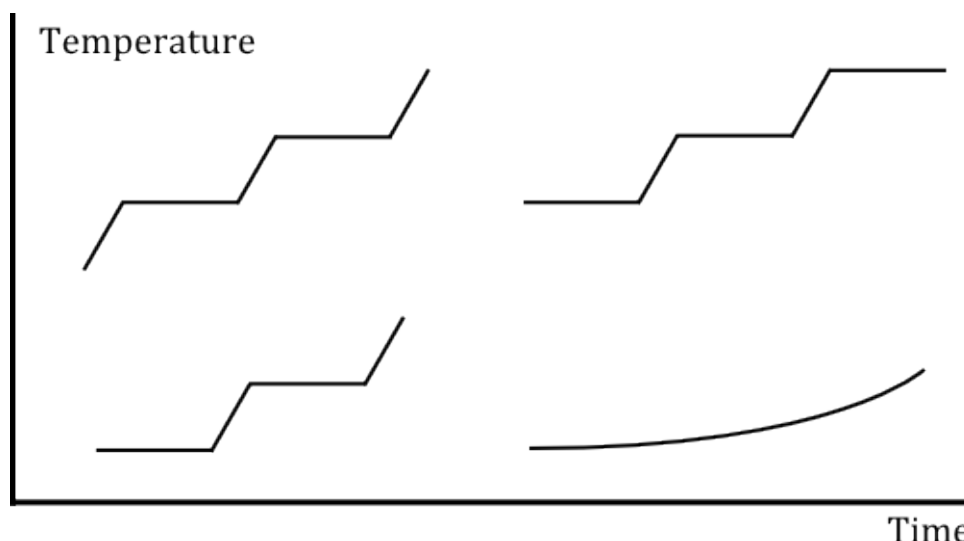


Figure 3: Temperature curves for tea kettle problem.

Problem 6: A tea kettle is filled with ice at -10°C and placed on a burner that is steadily getting hotter. Over a long enough time, the kettle empties due to evaporation. Of the four curves drawn in Figure 3, which best represents the temperature inside the kettle as a function of time? Use the quasistatic approximation. Answer: Top left.

Problem 7: How much ice at 0°C must be added to a liter of water at 80°C so as to end up with all liquid at 20°C ? Answer:

$$\frac{m_{\text{ice}}}{m_0} = \frac{80^\circ\text{C} - 20^\circ\text{C}}{20^\circ\text{C} - (L_f/c)_{\text{water}}}$$

Problem 8: A 15 g ice cube at 0°C is mixed with 2 g of steam at 100°C in an isolated container. What is the final state of the system? Answer: (i) Since $m_{\text{steam}}L_v - m_{\text{ice}}L_f < 0$, all steam condenses, releasing 4512 J , which melts 13.51 g of ice, leaving 1.49 g in the solid

state. (ii) Next, the 2 g of hot water must come to equilibrium with the melted ice, giving $0 = 2(x - 100^\circ\text{C}) + 13.51x$, where x solves to 12.9°C . (iii) Finally, let 15.51 g of water at 12.9°C come to equilibrium with the remaining ice, giving $0 = 15.51(y - 12.9^\circ\text{C}) + 1.49L_f/c$. Solving for y , the final system has 17 g of water at about 5°C .

Problem 9: A 20 g ice cube at -10°C is mixed with 2 g of steam at 100°C in an isolated container. What is the final state of the system? Answer: (i) Energy required to warm ice to 0°C : 422 J. (ii) Amount of steam condensed to accomplish this: 0.187 g. (iii) Energy required to melt ice: 6680 J. (iv) Energy released by remaining steam condensation: 4090 J. (v) Energy released from cooling 2 g of water from 100°C to 0°C : 837 J. (vi) Energy available to melt ice: 4927 J. (vii) Mass of ice melted: 14.8 g. (viii) Final state has about 5 g of ice and 17 g of water at 0°C .

2 Heat Transfer

2.1 Conduction

Conduction is heat transfer due to microscopic motions of particles. Rapidly-moving atoms in the heated part of an object exchange kinetic energy with ‘colder’ atoms. A material’s affinity to conduct heat in this way is the *thermal conductivity*, k , measured in Watts per meter per degree centigrade. Incidentally, most materials that easily conduct electric currents also have high thermal conductivity - a feature attributable crystalline atomic structure.

Suppose a hot object H is connected to a cold object C by a conductor having cross-sectional area A and thermal conductivity k . Objects H and C are maintained at constant temperatures T_H and T_C , respectively. The rate of energy flow dQ/dt across the conductor is:

$$\frac{dQ}{dt} = -\frac{d}{dt} \int \nabla U \cdot d\vec{x} = - \int \nabla U \cdot \frac{d\vec{x}}{dt}$$

The ∇U term is proportional to (i) the temperature difference $T_H - T_C$ and (ii) the conductor cross section A . The $d\vec{x}/dt$ term characterizes the speed of energy flow, characterized by thermal conductivity. The differential equation for heat conduction is therefore

$$\frac{dQ}{dt} = -kA\nabla T, \quad (10)$$

where ∇T is known as the *temperature gradient*. (The negative sign makes sure that heat flows from hot to cold.) In the special case that the conduction takes place in one dimension along length L , the rate of heat flow toward the cooler region is

$$\frac{dQ}{dt} = kA \left(\frac{T_H - T_C}{L} \right).$$

The R -value of a material is equal to the thickness (along the direction of heat flow) divided by the conductivity:

$$R = \frac{L}{k}$$

One inch of wood has $R \approx 1$. Typical New England houses have $R \approx 19$ for the walls and $R \approx 30$ for the roof. The R -value is an additive quantity, with $R_{eff} = R_1 + R_2 + \dots$

Problem 10: The walls of a house are insulated with 2.0 cm of styrofoam ($k_{\text{SF}} = 0.01 \text{ W/mK}$), and 15.0 cm of fiberglass ($k_{\text{FG}} = 0.04 \text{ W/mK}$). The styrofoam is on the exterior of the house. The outside temperature is 0°C , and the interior of the house is maintained at 20°C . (i) Calculate the temperature on the SF-FG interface. (ii) What is the rate of heat transfer per m^2 through both layers of insulation? (iii) Which provides greater net insulation, an additional 1.0 cm of styrofoam, or an additional 5.0 cm of fiberglass? Answer: 6.96°C , 3.5 W , fiberglass.

Problem 11: Inside a conductive cylinder of length L and radius R , heat flows radially from the axis $r = 0$ toward the wall at R with $T(0) > T(R)$. Calculate thermal energy flow rate between any two radii r_1 and r_2 in cylinder. Answer: $dQ/dt = k2\pi RL(T_1 - T_2) / \ln(r_2/r_1)$.

Problem 12: If a tank of water is left in very cold conditions, a sheet of ice forms on the water surface and grows downward over time. Assuming all heat transfer takes place through the ice, and not through the container walls: (i) Determine the thickness Z of the ice as a function of time t . (ii) Calculate the thickness of the ice sheet that will form in one day. (iii) If the tank is 50 cm deep, how many days does it take to freeze of the water? (iv) If the tank is 10 m deep, how many days does it take to freeze of the water? Answer: (i) $Z = (\Delta T 2k / (\rho L_f))^{1/2} \sqrt{t}$, (ii) 9.49 cm, (iii) $2.40 \times 10^6 \text{ sec}$, (iv) $6.40 \times 10^6 \text{ days}$

2.2 Convection

Consider a kettle of water initially at 10°C that sits on a kitchen stove burner. The water will transfer heat in a number of ways:

1. Conduction: Heat is introduced to the liquid by contact with its metal enclosure, where water molecules conduct the heat further into the volume by near-neighbor collisions.
2. Convection cells: Hexagonally-packed fluid parcels nudge their way to the surface, deposit heat into the atmosphere, and sink down again.
3. Turbulence: The bottom of the fluid becomes much warmer than the surface, and the convection cells burst into mushroom-cloud shapes.
4. Boiling: Discernable patterns are foregone, and the fluid expels its heat by ejecting the most energetic particles.
5. Steam: Steam rapidly expands (if not constrained), doing work on the atmosphere thus losing energy until its temperature reaches equilibrium with the surroundings.

Steps (2) through (5) each involve convection. *Convection* is heat transfer due to macroscopic displacement of a fluid. The displacement can be spontaneous, as when steam carries heat away from a hot cup of tea - or forced, as done inside a refrigerator compressor.

2.3 Electromagnetic Radiation

Electromagnetic radiation is energy transfer that uses photons as a carrier. All objects continuously absorb and expel energy through photon exchange with the environment.

The net rate of heat flow is given by Stefan's law,

$$\frac{dQ}{dt} = Ae\sigma (T_{obj}^4 - T_{env}^4) \quad , \quad (11)$$

where A is the object's exposed surface area, e is the emissivity ($0 \leq e \leq 1$) of the material, and σ is the *Stefan-Boltzmann* constant, $5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$. Note the temperature is counted in Kelvin units, which scales proportionally to the centigrade scale but equals zero at $T = -273.15^\circ \text{C}$. (See Lord William Kelvin below.)

A hot electric stove coil emits intense infrared radiation with wavelengths ranging from 10^3 to 10^6 nm . (Your skin easily detects these rays.) With increasing power, the range of emitted wavelengths widens, and eventually the coil visibly glows red with $\lambda = 700 \text{ nm}$. (Your eyes easily detect these.) With even more increasing power, the stove coil becomes 'white hot', emitting a very wide spectrum of wavelengths.

A hypothetical object useful for discussing radiation is the *blackbody*, which has $e = 1$, absorbing all incident radiation. Contrastly, a *reflector* absorbs no radiation.

Problem 13: Earth's upper atmosphere receives roughly $1.50 \times 10^3 \text{ W/m}^2$ of energy from the sun by thermal radiation. The distance from the earth to the sun is $1.50 \times 10^{11} \text{ m}$, and the radius of the sun $6.96 \times 10^8 \text{ m}$. Use conservation of flux to calculate the surface temperature of the sun. Answer: About 6000 Kelvin.

Wandering in the Desert

Problem 14: Consider an unfortunate person walking at 5 km/h on a hot day in the desert, wearing only a bathing suit. The person's skin temperature tends to rise due to four mechanisms:

- Energy is generated by metabolic reactions in the body at a rate of 280 W , and almost all of this energy is converted to heat that flows to the skin.
- Heat is delivered to the skin by convection from the outside air at a rate equal to $k' A_{\text{skin}} (T_{\text{air}} - T_{\text{skin}})$, where k' is $54 \text{ J/}^\circ \text{C m}^2 \text{ h}$, the exposed skin area A_{skin} is 1.5 m^2 , and the air temperature T_{air} is 47°C .
- The skin absorbs radiant energy from the sun at a rate of 1400 W/m^2 .
- The skin absorbs radiant energy from the environment, which has temperature 47°C .

Assume the emissivity of the skin is $e = 1$ and the skin temperature is 36°C . (i) Determine the net heat flow into the person due to the four mechanisms. (ii) At what rate (in liters per hour) must perspiration evaporate from the person's skin to maintain a constant skin temperature? (Assume sweat is made of water which has $L_v = 2.42 \times 10^6 \text{ J/kg}$.) (iii) Suppose instead that the person is protected by light colored clothing with $e = 0$ such that the skin exposed is 0.45 m^2 . (The convective heat exchange is not affected by the clothing.) What is the rate of perspiration now required? Answer: (i) 280 W , 0.248 W , $2.10 \times 10^3 \text{ W}$, 893 W , (ii) 4.87 L/h , (iii) 1.75 L/h

3 Ideal Gas Theory

3.1 Ideal Gas Law

We now compile a few historical achievements into an equation of state called the *ideal gas law*:

$$PV = NK_B T \quad (12)$$

The constant K_B is the *Boltzmann constant*, and has the experimental value $1.381 \times 10^{-23} \text{ J/K}$.

3.2 Boyle's Law

In the seventeenth century, Robert Boyle (1627 - 1691) performed numerous studies on the properties of gases using the materials of his day (simple glassware and metals). The most famous fruit of his effort, called *Boyle's Law*, states that *the pressure of a fixed amount of gas maintained at constant temperature is inversely proportional to the volume of the gas*:

$$P \propto 1/V$$

3.3 Charles's Law

A similar discovery about gases that came out of France in the 1780s, now known as *Charles's Law*, states that *the volume of a fixed amount of gas maintained at constant pressure is directly proportional to the absolute temperature of the gas*:

$$V \propto T$$

3.4 Lord William Kelvin

The linear relationship between volume V and temperature T found by Charles has an interesting feature that perhaps he didn't notice. All experimental data was indicating that any plot that displays $V(T)$ for *any* substance at *any* pressure has an x -intercept at -273.15°C . Scottish physicist Kelvin realized the significance of this in 1848, and seized the opportunity to name a new temperature scale after himself.

The *Kelvin temperature scale* resembles the centigrade system in every way with one exception: the bottom of the scale, 0 *Kelvin*, corresponds to -273.15°C , also known as *absolute zero*. As it turns out, absolute zero is the limit low temperature for any object in the universe. *Unless otherwise specified, all temperatures in this document are expressed in Kelvin units.*

3.5 Avogadro's Law

Italian scientist Amedeo Avogadro made his contribution in 1811, hypothesizing quite correctly that *the volume of a gas at constant temperature and pressure is proportional to the number of particles N composing the gas*:

$$V \propto N$$

Sometimes in physics (and more often in chemistry) the quantity NK_B is written instead as nR , where n is the number of moles of the gas, and R is the *ideal gas constant*, measured as $8.314 \text{ JK}^{-1}\text{mol}^{-1}$.

3.6 Ideal Gas Model

The ideal gas law is only an approximate description of a realistic gas. In a practical setting, equation (12) holds if the following assumptions can be safely made:

- The separation between gas particles is much greater than the diameter of a given particle.
- The volume of the container holding the gas is much greater than the total volume of the constituent gas particles.
- Particles do not interact among themselves, with the only exception being perfectly elastic collisions.
- Particles exchange momentum with the container walls, but walls don't deform.
- At a given instant, P , V , N , and T are uniform throughout the gas (quasistatic approximation).

In common thermodynamics culture, the subscript B is typically dropped from the Boltzmann constant; thus we simply write K for now on.

Problem 15: A rigid box of capacity 0.5 m^3 is initially open, but is then sealed, trapping air inside at 20°C at atmospheric pressure. The box is then heated until the pressure inside becomes 3 times that of the atmosphere. Calculate the temperature of the trapped air (in Kelvin) after heating. How many particles are in the box? Answer: $T = 879.45\text{ K}$ and $N = 1.25 \times 10^{25}$.

3.7 Mixtures and Dalton's Law

A mixture of ideal gases is itself an ideal gas and still obeys $PV = NKT$. For instance, if a pressurized vessel holds a mixture of helium and nitrogen, the gases are fully mixed, sharing the same V and T . In the variable N , we really mean $N_{He} + N_{Ni}$. Each element also makes its own contribution to P , called a *partial pressure*, where $P = P_{He} + P_{Ni}$ for the example at hand. This is known as Dalton's law. In general, the law of partial pressures reads

$$P = \sum_j P_j, \quad (13)$$

where index j sums over the ideal gas species in the mixture.

3.8 Vapor Pressure, Humidity, Clouds

The *vapor pressure* is the pressure at which the vapor phase of a substance is in equilibrium with the solid or liquid phase. For example, if water is sealed in a container with air above, liquid molecules evaporate into the air. This occurs until the partial pressure P_W of water *in the air* is equal to the vapor pressure P_0 .

The vapor pressure of a substance varies with the temperature, and for water the relationship is reproduced in the table below:

Temperature (°C)	Vapor Pressure (Pa × 10 ³)
10	1.23
12	1.40
14	1.60
16	1.81
18	2.06
20	2.34
22	2.65
24	2.99
26	3.36
28	3.78
30	4.25
40	7.34

Note one Pascal (Pa) is equal to one Newton per square meter.

The *relative humidity* is defined as the ratio

$$H = \frac{P_W}{P_0},$$

always expressed as a percentage and never exceeding 100%. Because Earth's atmosphere is considered open instead of closed, H is typically much less than unity. On a rainy day, H becomes close to unity. In either case, P_0 is much less than the atmospheric pressure P_A , which is near $10^5 Pa$.

Problem 16: At $24^\circ C$ and at 50.0% relative humidity, what is the partial pressure of water vapor in the atmosphere? Answer: $P_W = 1.50 \times 10^3 Pa$

Problem 17: At $24^\circ C$ and at 50.0% relative humidity, what is the mass density of the water vapor in the air? (Hint: $M = 18 \times 10^{-3} kg/mol$.) Answer: $\rho = 0.0109 kg/m^3$

Consider an ice-cold glass sitting in a room. After a short time, water droplets form on the outside of the glass. To explain this, observe first that the air surrounding the glass, which itself contains water molecules, is cooled by conduction. The vapor pressure of the surrounding molecules therefore decreases (see table above), causing the relative humidity surrounding the glass to increase. When the relative humidity reaches 100%, the corresponding temperature is called the *dew point*, denoted T_{dew} .

At the dew point, further cooling cannot result in increased humidity, which is already at unity, so the density of water molecules surrounding the glass must decrease. The excess molecules stick to the glass to form droplets. Note that the dew point is always less than the surrounding room temperature. The higher the relative humidity, the closer the dew point is to the temperature.

Problem 18: Suppose the air temperature in a laboratory is $24^\circ C$, and a physics student cools a half-filled drink by adding ice. He finds that water begins to condense on the glass when its temperature reaches $12^\circ C$. What is the relative humidity in the room?

Answer: 46.8%

Clouds and fog are easily explained in the terms defined above. Simply put, condensation (100% relative humidity) occurs at the altitude where the air temperature equals the dew point.

Problem 19: Consider a spring day when the air temperature on the ground is 24°C and the relative humidity is 50%. Assume that the partial pressure of water does not change with elevation, and that the air temperature decreases with altitude at a rate of $0.06^\circ\text{C}/100\text{ m}$. (i) At what altitude will clouds form? (ii) If the relative humidity on the ground is higher, will the clouds form at a higher or lower altitude? (iii) If the temperature gradient is less, will the clouds form at higher or lower altitude? Answer: (i) 1833 m, (ii) lower, (iii) higher

4 Kinetics of Ideal Gas

We will take a strictly classical approach to the ideal gas problem, leaving a proper quantum-level analysis to subsequent chapters. An ideal gas is assumed to be isotropically distributed throughout its container. This means that a given particle's direction of travel (not its speed) is completely random. It follows that the container walls exchange momentum with the gas particles isotropically, and at a rate determined by *some* kind of average speed of the gas particles.

4.1 Pressure

Consider a sealed vessel of volume V containing N gas particles of mass m - not necessarily ideal gas. Next, consider a cylindrical Gaussian membrane of length L inside the vessel, held so its axis (the x -direction) is perpendicular to the inner surface of the vessel. One end of the Gaussian touching the container's inner wall with common area A , assumed circular.

Question: How many of the particles ΔN enclosed in the membrane will collide with the container wall in time interval Δt ? Answer: Statistically half of the enclosed particles collide with the wall (the other half are heading away). A particle will collide with the wall in a time interval Δt if it is heading toward the wall and within a distance $L = |v_x| \cdot \Delta t$ of the wall, where v_x is the x -component particle's absolute speed, $\sqrt{\langle v^2 \rangle}$. The collision rate is also proportional to the overall density of the gas, N/V . Therefore we may write

$$\Delta N = (1/2) (A |v_x| \Delta t) (N/V) .$$

Upon each collision with the container wall, there are two contributions to the momentum exchange: (i) the wall first absorbs momentum p_x to momentarily stop the gas particle, and (ii) the wall then imparts momentum $-p_x$ to the particle, with a total exchange of $\Delta p = 2mv_x$ per collision.

The pressure of the gas is, by definition,

$$P = \Delta N \frac{1}{A} \frac{\Delta p}{\Delta t} ,$$

and after inserting what we know about ΔN and Δp , this is

$$P = \frac{N}{V} m \langle v_x \rangle^2 .$$

Assuming the gas to be isotropic and living in three dimensions, it follows that

$$\langle v^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle = 3 \langle v_x^2 \rangle ,$$

bringing us to the result

$$P = \frac{1}{3} \frac{N}{V} m \langle v^2 \rangle . \quad (14)$$

Note that the applicability of (14) extends beyond ideal gases, as we have not yet required or assumed the ideal gas law.

Problem 20: In a 60 sec interval, 700 hailstones strike a 0.5 m^2 glass window at an angle of 45° with respect to the normal to the window surface. Each hailstone has mass of 4.0 g and speed 8.0 m/s. If the collisions are purely elastic, Calculate the average force and pressure on the window.

4.2 Internal Energy, RMS Speed

If we assume gas particles do not interact through any attractive or repulsive forces, all thermal is purely kinetic - a quantity called *internal energy*. The internal energy is therefore easy to write down: multiply the number of particles by some kind of average kinetic energy of one particle.

$$U = N \left(\frac{1}{2} m \langle v^2 \rangle \right) \quad (15)$$

Eliminating the $\langle v^2 \rangle$ term between (14) and (15) yields another strong and general statement about gases,

$$U = \frac{3}{2} PV . \quad (16)$$

Finally narrowing the analysis down to ideal gases, replace PV with NKT by (12), giving

$$U = \frac{3}{2} NKT . \quad (17)$$

Eliminating U between (15) and (17) gives a direct relation between the temperature and internal speed. This particular result is the *root mean square speed*:

$$\sqrt{\langle v^2 \rangle} = v_{rms} = \sqrt{\frac{3KT}{m}} \quad (18)$$

A comment on mixtures: If we consider a mixture of gases (rather than a pure sample of identical particles), variables m , $\langle v^2 \rangle$, and N would each need a subscript to track each species. Variables P , N , and U add as $P = \Sigma P_j$, $N = \Sigma N_j$, and $U = \Sigma U_j$. Without any loss of generality, we focus on single-species gases for the remainder of this section. All results are easily applied to mixtures.

Problem 21: The temperature of 7.00 mol of helium gas is increased by 2.00 K. What is the change in internal energy?

4.3 Maxwell Speed Distribution

The natural question now is, what's the precise distribution of molecular speeds in the ideal gas? Do *all* ideal gas particles of the same mass travel at the same speed? Are 50% of the particles above v_{rms} with the other 50% below? The answer lies in a complicated formula called the *Maxwell(-Boltzmann) speed distribution*

$$D(v) dv = \left(\frac{m}{2\pi KT} \right)^{3/2} 4\pi v^2 e^{-mv^2/2KT} dv, \quad (19)$$

which shall be derived when I introduce Boltzmann factors. The curve 'opens up' parabolically near speed $v = 0$, has a peak featuring *three* important speeds, and dies off exponentially for very large v .

Using the tricks of standard probability theory, the chance of finding a molecule with speed within the interval $[v_1, v_2]$ equals the area under the *probability density* $D(v)$ spanning that interval. This means

$$P(v_1 < v < v_2) = \int_{v_1}^{v_2} D(v) dv. \quad (20)$$

The probability of finding a molecule at any one exact speed is zero because the area under an infinitely skinny spike in $D(v)$ is negligible. Therefore we always talk about speed intervals, the smallest being Δv , as in

$$P(v \pm \Delta v/2) = D(v) \Delta v.$$

The factor $4\pi (m/2\pi KT)^{3/2}$ comes from the normalization requirement

$$1 = \int_0^\infty D(v) dv,$$

where inserting (19) reduces the problem to a standard *Gaussian integral*. Normalization asserts the notion that any given particle is going to have *some* speed between 0 and ∞ . Note that the integral is over the one-dimensional variable v . The fact that velocities are 3D vectors is accounted for by the $4\pi v^2$ factor in $D(v)$.

4.4 Three Speeds

We were able to weasel the formula $v_{rms} = \sqrt{3KT/m}$ out of simplistic arguments, but the Maxwell speed distribution is needed to carry the analysis forward. The easiest trick is to take set derivative of (19) to zero to find v at its maximum, yielding

$$v_{max} = \sqrt{\frac{2KT}{m}}. \quad (21)$$

By definition, more particles are moving at speed v_{max} than any other. Note v_{rms} is about 22% larger than v_{max} , and we'll soon see that the *actual* average speed $\langle v \rangle$ lies somewhere in between.

It may have been tempting to assume v_{rms} or v_{max} as the average speed $\langle v \rangle$, but the statistical definition of $\langle v \rangle$ yields a different result entirely. Since we already have the probability density $D(v)$ available, the calculation amounts to solving the integral

$$\langle v \rangle = \int_0^\infty v \cdot D(v) dv.$$

Explicitly, this is

$$\langle v \rangle = 4\pi \left(\frac{m}{2\pi KT} \right)^{3/2} \int_0^\infty v^3 e^{-mv^2/2KT} dv = \sqrt{\frac{8KT}{\pi m}} \int_0^\infty u e^{-u} du ,$$

where the substitutions $q = v\sqrt{m/2KT}$ and $\sqrt{u} = q$ have been made. Lucky enough, the calculation of $\langle v \rangle$ doesn't involve any Gaussian integrals. The remaining u integral evaluates to unity, yielding the average velocity at last:

$$\langle v \rangle = v_{ave} = \sqrt{\frac{8KT}{\pi m}} \quad (22)$$

Writing each result side-by-side, we learn:

$$v_{max} < v_{ave} < v_{rms}$$

$$\sqrt{\frac{2KT}{m}} < \sqrt{\frac{8KT}{\pi m}} < \sqrt{\frac{3KT}{m}}$$

The numerical coefficients are 1.414, 1.596, and 1.732 respectively. Stated in terms of probability density,

$$D(v_{max}) > D(v_{ave}) > D(v_{rms}) .$$

It should also be noted that $v_{rms} = \sqrt{\langle v^2 \rangle}$ is correctly predicted by the Maxwell speed distribution by replacing v^3 by v^4 in the integral above.

Problem 22: The escape speed for a particle to leave the gravitational influence of a massive body is given by $(2GM/R)^{1/2}$. The temperature near the top of Jupiter's multicolored cloud layer is about 140 K. The temperature of Earth's atmosphere at 20 km is 220 K. Calculate the RMS speed of H_2 molecules in each of those environments. Give the answer as a fraction of the escape speed from the respective planet. Answer: 14.7% for Earth and 2.16% for Jupiter, explaining the rarity of H_2 in Earth's atmosphere. Jupiter's atmosphere consists of 89% H_2 plus other gases.

Problem 23: Ceres is an asteroid with a mass equal to 0.014 times the mass of the Moon, has density 2400 kg/m^3 , and surface temperature around 200 K. Can this object support an O_2 atmosphere? Answer: $v_{rms}/v_{esc} = 72.9\%$ thus O_2 will largely leave the asteroid.

Problem 24: Gas molecules in a container have the following measured speed distribution:

Speed (m/s)	Percentage
220	10%
250	10%
500	15%
650	30%
900	20%
1300	15%

Calculate the average speed, the rms speed, and the most probable speed. If the system is ideal gas of molecular mass $M = 50 \times 10^{-3} \text{ kg/mol}$, what is the temperature? Answer: 690 m/s, 768 m/s, 650 m/s, 1180 K.

4.5 Equipartition Theorem

Our perspective of the ideal gas has taken all particles to be pointlike spheres. Question: Can we extend the discussion to include small molecules with definite structure? Answer: Yes, so long as we find a way of keeping track of two new contributions to the internal energy:

1. Non-spherical molecules are able transfer kinetic energy of *rotation* to other molecules by collision.
2. Molecules may self-vibrate, as atomic structures are held together with flexible bonds, quadratic in stength.

The internal energy of a ‘generalized ideal gas’ is generally split between translations, rotations, and vibrations. Such a gas still obeys the equation of state $PV = NKT$, as molecular rotations and vibrations contribute nothing to the pressure and volume.

The internal energy is not generally governed by (17). In order to accurately describe the internal energy of the generalized ideal gas, the *equipartition theorem* must be borrowed from from statistical mechanics. As we will *prove* in subsequent chapters, theorem reads:

The internal energy per molecule gains a factor of $KT/2$ for each independent quadratic degree of freedom.

To illustrate, consider again the ideal gas made of pointlike spheres. There are three degrees of mechanical freedom, namely translations in the x , y , z directions, where the kinetic energy of translations goes as the *square* of each velocity component. Applying the equipartition theorem gives $U = 3 \times KT/2$, verifying (17). In the general case, the kinetic energy of an ideal gas is

$$U = \frac{f}{2}NKT, \quad (23)$$

where f is the integer number of quadratic degrees of freedom per molecule. Note f is not necessarily constant for a given substance; we’ll discuss shortly that f can vary with temperature.

4.6 Diatomic Molecules

Consider a generalized ideal gas made of N identical molecules, each consisting of two hydrogen atoms bound by a quadratic potential. This is a system of dumbbell-shaped H_2 molecules. In accordance with (23), we count the quadratic degrees of freedom per molecule as follows:

- Three translational modes, so f is at least 3.
- Two degrees of rotational freedom. The diatomic molecule may spin like a baton in two orthogonal modes, however rotations about the axis joining the atoms do not change during collision. So far, $f = 3 + 2$.
- Two contributions from the single vibrational mode (a combination of both kinetic and potential energies), increasing f again by 2.

We expect therefore that the energy of a diatomic gas is

$$U_{dia} = (7/2) NKT .$$

However, $f = 7$ is only observed at high temperatures. Experiment shows that a sample of H_2 gas will exhibit only translational behavior until the temperature reaches about $200K$. Below this temperature, the rotational and vibrational energies are ‘frozen out’ and don’t contribute to the energy. By $400K$, the two rotational modes are fully active. Vibrational modes finally show themselves around $4000K$. Conclude that diatomic molecules have f appearing as 3, 5, or 7, depending on T .

Problem 25: Use the equipartition theorem to find the total rotational energy of the molecules in 4.00 mol of diatomic gas at $310 K$.

4.7 Intermolecular Collisions

Now I introduce two new ideas in the kinetic theory of gas: the *mean free path* and the *mean free time*. That is, we want to know (i) the distance a molecule travels before colliding with another molecule, and (ii) the time between such collisions.

Suppose an ideal gas consists of N identical molecules of radius r . At a given instant, each particle is moving in a straight trajectory through its container, and a collision occurs if any particle attempts to occupy the same space as one of its neighbors, with $2r$ being the minimal separation between any two particles.

The picture is equivalently framed and more easily analyzed if we focus one ‘test’ particle in the gas, taking along two awkward but correct simplifications: (i) all other particles in the gas are considered stationary while the test particle goes, and (ii) all other particles in the ideal gas are considered as points, but our test particle swells to radius $2r$.

In time interval Δt , the sphere of radius $2r$ carves out a Gaussian cylinder of length L and volume $\pi (2r)^2 L$. When the cylinder’s volume approaches the average volume per molecule in the gas, a collision is very likely. Solving for L gives the mean free path,

$$L \approx \frac{1}{4\pi r^2} \frac{V}{N} , \quad (24)$$

where V is the container volume and N is the total number of particles.

The mean free time is a piece of cake, since $L = v\Delta t$, right? But just *which* velocity should be used to make this approximation? I argue in favor of using v_{rms} , because we used a similar Gaussian cylinder in establishing equation (14), where v appears as v_{rms} . Therefore, the mean free time is approximately

$$\Delta t \approx \frac{L}{v_{rms}} . \quad (25)$$

4.8 The van der Waals Model

The *van der Waals* model is an extension of the ideal gas law that accounts for the particle size and the interactions between particles. To write a generalized ideal gas law, two experimental constants enter the mix - blandly named a and b .

The starting point is the ideal gas model, with equation of state $PV = NKT$. Supposing each of N particles in a sample (gas or liquid) has volume b , we make the replacement

$$V \rightarrow V - Nb ,$$

where the variable V is understood as the volume available for the sample to occupy. The particles themselves do not contribute to V .

Next we assume particles interact through the attractive *Leonnard Jones potential*. Without increasing temperature, particle interactions lead to an increase in number density N/V and pressure P , as evident by considering a particle on the surface layer of the sample. On the surface layer, the net force on a particle is inward, prorortional in strength to N/V . The total force inward is proportional to the surface area of V , contributing another factor N/V . It follows that the correction to pressure is

$$P \rightarrow P + a (N/V)^2 ,$$

where a is a constant of dimension kgm^5s^{-2} .

The van der Waals equation is

$$\left(P + a \frac{N^2}{V^2} \right) (V - Nb) = NKT , \quad (26)$$

and is extremely accurate for many types of gases and fluids. Water vapor has $a = .5507 Jm^3mol^{-2}$ and $b = 3.04 \times 10^{-5} m^3/mol$. In the same units, N_2 gas has $a = .1361$ and $b = 3.85 \times 10^{-5}$, and H_2 gas has $a = .0247$ and $b = 2.65 \times 10^{-5}$.

5 Particle Effusion

In this section we study *particle effusion*, which is the problem of gas particles rushing out of a small aperture into open space. If you get lost in the math in this section, the results should be useful to you all the same.

Consider a sealed vessel of volume V containing N gas particles of mass m - not necessarily ideal gas. If a small hole is made in the side of the vessel, we want to know (i) do certain particles have a quicker tendency to leave through the hole than do other particles, and (ii) what is the rate of effusion through the hole?

We first argue that inside the vessel, a given particle has any speed between zero and infinity in any direction. Introducing $f(\vec{v})$ as a generalized Maxwell distribution (19), we write the normalization condition as

$$\int_{all \vec{v}} f(\vec{v}) d^d v = 1 ,$$

where d is the number of dimensions in velocity space.

5.1 Statistical Calculation of Pressure

For a warm-up calculation, examine the pressure incident on a small patch dA of the sealed vessel's inner wall. Imagine a Gaussian cylindrical membrane of length L inside the vessel whose end is touching the inner z -wall with common area element $-dA\hat{z}$, where $dA = dxdy$. For particles with speed v inside the Gaussian cylinder, the length is represented by the vector $\vec{L} = (dt)\vec{v}$, and the volume is $d\vec{A} \cdot \vec{L} = dA dt v_z$. The number of particles ΔN destined to collide with the wall is proportional to (at least) the macroscopic density N/V and the volume of the Gaussian cylinder. Upon collision, the particle reverses the z -component of its momentum, with $dp_z = 2p_z$. Particles with negative p_z are heading away from the wall and don't collide.

Collecting these observations, write the pressure ‘operator’ (in the spirit of probability theory) as

$$[P_v] = \Delta N \frac{dF}{dA} = \frac{N}{V} v_z 2p_z,$$

and the total pressure is given by

$$P = \int_{\vec{v}} [P_v] f(\vec{v}) d^3v .$$

Explicitly, this is

$$P = \int_{v_x=-\infty}^{v_x=\infty} \int_{v_y=-\infty}^{v_y=\infty} \int_{v_z=0}^{v_z=\infty} \frac{N}{V} v_z 2p_z f(\vec{v}) dv_x dv_y dv_z .$$

Note the careful choice of limits on the v_z variable.

Since $f(\vec{v})$ is an even function, it does no harm to integrate over the interval $-\infty < v_z < \infty$ and divide by a factor of 2. The integral simplifies to

$$P = \frac{N}{V} \int_{\vec{v}} (u_z p_z) f(\vec{v}) d^3v ,$$

and by definition of the statistical average, we have

$$P = \frac{N}{V} \langle u_z p_z \rangle ,$$

where average quantities are contained within brackets $\langle \rangle$. Finally, observe that

$$\langle u_z p_z \rangle = \frac{1}{3} \langle up \rangle ,$$

and the pressure has a form reminiscent of equation (14):

$$P = \frac{1}{3} \frac{N}{V} \langle up \rangle$$

5.2 Effusion Rate

Now we approach the problem of effusion, which first entails finding the rate R (measured in $m^{-2}s^{-1}$) at which particles that strike area dA . In analogy to the pressure calculation, we write

$$R = \int_{\vec{v}} [R_v] f(\vec{v}) d^3v ,$$

with the corresponding ‘operator’

$$[R_v] = \frac{\Delta N}{\Delta A} = \frac{N}{V} v_z .$$

Expressed in 3D spherical coordinates, the rate is

$$R = \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi/2} \int_{v=0}^{v=\infty} \frac{N}{V} v \cos \theta f(\vec{v}) v^2 dv \sin \theta d\theta d\phi ,$$

where the z -axis is perpendicular to the container wall on the patch dA . The gas exists only above the $z = 0$ plane, as indicated by the choice of limits on the variable θ . Integrating over angles ϕ and θ , the calculation reduces to

$$R = \frac{N}{V} \int_0^\infty \pi v f(\vec{v}) v^2 dv ,$$

and the normalization condition

$$\int_0^\infty f(\vec{v}) 4\pi v^2 dv = 1$$

helps us recognize the remaining integral as the average velocity $\langle v \rangle$, up to a constant.

We therefore find

$$R = \frac{1}{4} \frac{N}{V} \langle v \rangle , \quad (27)$$

known as the instantaneous *effusion rate*. That is, if the element dA is replaced by a small hole in the vessel, particles at any given instant are leaving the vessel at a rate equal to (27). Notably, this result is attained without assuming any form for $f(\vec{v})$.

5.3 Effusion of Ideal Gas Particles

Now further expand the problem of ideal gas particles effusing through a small z -facing perforation in the container wall. The rate at which particles leave is given by (27), but we want to know the average z -direction kinetic energy \bar{E} carried per particle.

The result is given calculated from

$$\bar{E} = \int_{\vec{v}} [E_v] f(\vec{v}) d^3v ,$$

where $[E_v]$ is given by

$$[E_v] = \frac{N}{V} v_z \times \frac{1}{R} \frac{m v_z^2}{2} ,$$

and the quantity $f(\vec{v}) d^3v$ may be replaced by the Maxwell distribution (19), leaving out the factor of 4π .

Expressing the integral in 3D spherical coordinates and performing the trivial integration over ϕ , we have

$$\bar{E} = \frac{2\pi m N}{2R V} \left(\frac{m}{2\pi K T} \right)^{3/2} \int_0^{\pi/2} \cos^3 \theta \sin \theta d\theta \int_0^\infty v^5 e^{-m v^2 / 2 K T} dv .$$

The integrals evaluate to $1/4$ and $8(KT)^3/m^3$, respectively. In the variable R , we replace $\langle v \rangle$ by equation (22). After all the dust settles, arrive at

$$\bar{E} = K T .$$

This result is rather interesting because it's twice what we might expect. An ideal gas that is free to move in three dimensions has energy $(1 + 1 + 1)KT/2$ per particle in accordance with (17). In our problem, energy \bar{E} is associated with motion in the z -direction only, yet the energy comes out to $(1 + 1)KT/2$ per particle. It follows that the total energy of the effused particles is $2KT$ per particle.

Evidently, the fastest-moving particles in a vessel are more likely to effuse than the slow-moving particles. In a large system that undergoes sustained effusion, the temperature inside the vessel decreases significantly due to this effect.

6 Heat Capacity of Gases and Solids

Recall that the heat capacity is the factor that couples energy change to temperature change for a given substance:

$$C = \frac{dE}{dT} \quad (28)$$

Suppose a known quantity of energy ΔQ is introduced to a sample of gas, raising its temperature by some amount. It is tempting but generally wrong to assume that *all* of ΔQ contributes to the kinetic energy of the gas. Instead, some added energy is usually ‘wasted’ on volume expansion of the system.

When the system is constrained to a fixed volume, the heat capacity is represented by the variable C_V . A similar variable C_P applies when pressure is instead held constant.

6.1 Ideal Gas at Constant Volume

Let’s momentarily squelch the volume expansion problem by considering a sample of ideal gas inside a sealed, infinitely rigid container of volume V . It follows that the gas maintains constant volume if the system is heated or cooled - only the pressure and temperature are variable, so all added heat contributes to the kinetic energy.

The constant - volume heat capacity of the ideal gas is calculated from a T -derivative of (23), giving

$$C_V = \frac{f}{2}NK. \quad (29)$$

Due to complications that are later solved by quantum mechanics, (29) is only accurate for some gases. For one mole of monatomic gas with point-like constituents with $f = 3$ (such as helium or argon), the prediction for C_V evaluates to

$$C_{V_{mon}} = \frac{3}{2}R = 12.47 \frac{J}{mol \cdot K},$$

in very good agreement with experiments.

6.2 Ideal Gas at Constant Pressure

Now suppose an ideal gas is heated while maintaining constant pressure. The energy ΔQ added contributes not only to the internal motions of the gas, but toward volume expansion - doing mechanical work on the surroundings. By energy conservation, this means

$$\Delta Q = \Delta U_{int} + P\Delta V,$$

where ΔQ is heat energy added to the system externally, ΔU is the increase in translational, rotational, and vibrational motions of the molecules, and $P\Delta V$ is the energy needed to shove the environment out of the way as the gas expands. It’s readily checked that $P\Delta V$ has units of energy, and is the generalization of work $F \cdot dx$.

Taking a T -derivative of the above yields the heat capacity at constant pressure on the left. Replacing ΔU using (23) and rewriting $P\Delta V$ with the ideal gas law, this is

$$C_P = C_V + NK, \quad (30)$$

a widely useful result experimentally, allowing C_P and C_V to need not both be measured for a given sample.

More insight into the ideal gas is gained from the ratio C_P/C_V , which comes out to

$$\gamma = \frac{C_P}{C_V} = 1 + \frac{2}{f}. \quad (31)$$

Problem 26: Methane (CH_4) is a 3-dimensional molecule - a tetrahedron with carbon in the center. Determine the best approximations for C_V and γ . Answer: $24.9 \text{ J mol}^{-1} \text{ K}^{-1}$ and 1.33

6.3 Elemental Solids

The equipartition theorem allows easy calculation of the heat capacity of many solids. For our purposes, a solid is regarded as a semi-frozen ideal gas; a periodic arrangement of molecules with no contributions to thermal energy from translational or rotational motions. The structure is held together by quadratic potentials, thus such solids have six degrees of freedom per molecule: three components of kinetic energy due to vibration and three corresponding quadratic potentials.

According to (29), one mole of ‘frozen gas’ with $f = 6$ has

$$C_{V_{\text{solids}}} = \frac{6}{2}R = 24.9 \frac{\text{J}}{\text{mol} \cdot \text{K}},$$

which is obeyed by *all* elemental solids at high enough temperature, known as the *rule of Dulong and Petit*.

7 First Law of Thermodynamics

The first law of thermodynamics is a statement of energy conservation. If U is the internal energy of a substance with pressure P and volume V , thermal energy dQ added (or removed) affects the system by

$$dU = dQ - PdV, \quad (32)$$

known as the *first law of thermodynamics*. To make sense of the sign of the PdV term, imagine adding heat $+dQ$ to a gas sample using a candle. If the experiment is prepared such that the temperature does not change, making $dU = 0$, then the gas necessarily expands its volume. The PdV term must appear negative to balance out dQ .

It’s important that volume change dV in a thermodynamic system must happen ‘slow’ enough for T and P to respond - a circumstance we previously called *quasistatic*.

Thermodynamic Processes

A *thermodynamic system* is any collection of parts that permits heat transfer between its members. A closed thermodynamic system allows no energy or material transfer across its boundary, where an open system does allow such transfer. Heat transfer occurs via some thermodynamic process, described as follows:

7.1 Isothermal Process

An *isothermal* process is one that leaves the system’s temperature unchanged, which generally means there is no change in the system’s internal energy. The first law with $dU = 0$ reads $0 = dQ - PdV$.

7.2 Isochoric Process

A process that leaves the system's volume unchanged (easier said than done) is classified as *isochoric*. The dV term in the first law is zero by construction, so all heat added to the system contributes to internal energy via $dU = dQ - 0$.

7.3 Isobaric Process

A process that leaves the system's pressure unchanged (easily said and easily done) is classified as *isobaric*. For example, the macroscopic work done on a gas is not an integral $\int PdV$, but simply the product $P\Delta V$.

7.4 Adiabatic Process

Adiabatic processes are those that have zero net heat transfer into or out of the system. The first law of thermodynamics quantifies this by $dU = 0 - PdV$, where evidently any internal energy change dU is balanced out by response in pressure and volume.

7.5 Processes in Ideal Gas

We now examine the consequences of thermodynamic processes for a controlled sample of ideal gas.

Isothermal

Suppose an ideal gas sample is heated isothermally and is allowed to expand. According to equation (17), the internal energy change $\Delta U = (f/2) NK\Delta T$ is zero because $\Delta T = 0$, and therefore all added heat contributes to volume expansion according to $dQ = PdV$. The total energy entering the gas equals the total work done on the environment, given by the integral

$$W = \int PdV = \int_{V_i}^{V_f} \frac{NKT}{V} dV = NKT \ln \left(\frac{V_f}{V_i} \right). \quad (33)$$

Isochoric

An isochoric process permits no volume change, thus all added energy contributes to molecular kinetics according to $dQ = dU$. The internal energy change is given by (23), namely $dU = (f/2) NKdT$. We therefore write

$$\Delta Q = \frac{f}{2} NK\Delta T, \quad (34)$$

where the factor $fNK/2$ should be familiar as equation (29), the constant-volume heat capacity C_V .

Isobaric

An isobaric process leaves the pressure unchanged, but the temperature and volume are prone to change. Solving for dQ in the first law and differentiating gives

$$\frac{dQ}{dT} = \frac{dU}{dT} + P \frac{dV}{dT}.$$

For an ideal gas, we see, for the isobaric case,

$$\frac{dQ}{dT} = C_P \quad \frac{dU}{dT} = C_V \quad P \frac{dV}{dT} = NK, \quad (35)$$

which is identical to the statement that gave us equation (30).

Adiabatic

An adiabatic process usually involves a rapid expansion or compression of a gas that is ‘too fast’ for external heat transfer, making $dQ = 0$. Combining the first law with equation (23) tells us

$$\frac{f}{2} dT = -T \frac{dV}{V} \quad \rightarrow \quad TV^{\gamma-1} = \text{constant}. \quad (36)$$

Replacing T using the ideal gas law, we get a second useful identity,

$$PV^{\gamma} = \text{constant}. \quad (37)$$

Note the two constants in (36) and (37) are *not* the same, hence the soviet spelling.

We now calculate the work entering an ideal gas during an adiabatic process by the definition $W = \int P dV$, resulting in

$$W = \frac{1}{1-\gamma} (V_f P_f - V_i P_i). \quad (38)$$

Notice the variable γ is always > 1 by construction, thus (38) has a buried negative sign. If the overall sign of W is positive, the work is done by the gas. If negative, the work is done on the gas.

One more way to calculate the work entering an ideal gas during an adiabatic process begins with the first law, giving $W = -\Delta U$, which simplifies, after also using (29), to

$$W = -NK \frac{f}{2} (T_f - T_i) = -\frac{PV}{T} \frac{f}{2} (T_f - T_i), \quad (39)$$

which is identical to (38). Note the term (PV/T) may be evaluated at *any* point in the gas’s phase space during the adiabatic process, so long as the combination of P , V , and T are known simultaneously.

Problem 27: Show that ideal gas undergoing an adiabatic process obeys

$$\frac{dT}{dP} = \frac{2}{f+2} \frac{T}{P}. \quad (40)$$

Problem 28: Consider an ideal gas of N particles with initial temperature T_i and initial volume V_i that is compressed to final volume V_f . Which thermodynamic process, adiabatic or isothermal, requires the least energy to compress the gas?

Problem 29: A scuba diver is swimming a depth of 25 m, where the pressure is 3.5 atm. The air she exhales forms bubbles 8.0 mm in radius, which rise to the surface where the pressure is 1 atm. Assume the bubbles remain at the uniform 300 K temperature of the surrounding water. How much work is done by each bubble as it expands to the rising

surface? Answer: $0.953 J$

Problem 30: When a quantity of monatomic ideal gas expands at a constant pressure of $4.00 \times 10^4 Pa$, the volume of the gas increases from $2.00 \times 10^{-3} m^3$ to $8.00 \times 10^{-3} m^3$. (i) What is the change in internal energy of the gas? (ii) Does heat flow into out of the gas? (iii) What is the magnitude of the heat flow? Answer: $360 J$, $600 J$ into the gas.

Problem 31: At height z in the atmosphere, (i) show that a horizontal thin layer of air with density $\rho(z)$ and pressure $P(z)$ obeys the equation $dP/dz = -mg\rho(z)$, where m is the average mass of molecules in the air and $g = 9.8 m/s^2$. (ii) Assuming the air is ideal gas, show that $P(z) = P(0)e^{-mgz/KT(z)}$. (iii) Consider a bubble of gas that moves upward fast by expanding or shrinking adiabatically. Show that the temperature gradient experienced by the bubble is given by

$$\left| \frac{dT}{dz} \right| = \left| \frac{2}{f+2} \frac{mg\rho(z)}{P(z)} \right|.$$

8 Thermodynamic Cycles

A *thermodynamic cycle* is a sequence of processes that bring an open system's state variables to a previous configuration, and is depicted as a closed loop plotted in a PV diagram. It follows that the net internal energy change ΔU is zero for a complete thermodynamic cycle, meaning that any heat added to a system must be removed somewhere in the same cycle.

During any infinitesimal volume change, work done is equal to PdV . Integrated over a thermodynamic cycle, the total work done equals the area enclosed by the loop on the PV diagram. If the loop traces out in a clockwise direction, the work output of the system is positive. When counter - clockwise, the system absorbs work over a complete cycle.

A *heat engine* is a device that utilizes a thermodynamic cycle in repetition to perform 'useful' work. The *efficiency* of a heat engine is defined as the ratio of output work over input heat energy. Using the first law we write

$$e = \frac{W}{Q_H} = \frac{Q - \Delta U}{Q_H} = \frac{Q}{Q_H}. \quad (41)$$

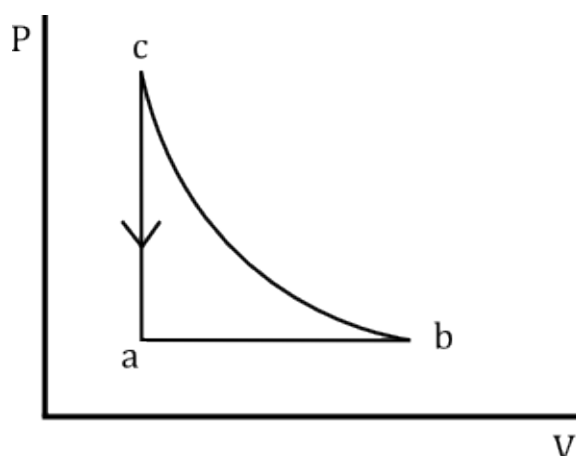
Through a complete thermodynamic cycle, the net change in internal energy ΔU is zero, and the net heat flow Q equals $Q_H + Q_C$. The efficiency is therefore

$$e = \frac{Q_H + Q_C}{Q_H} = 1 - \left| \frac{Q_C}{Q_H} \right|,$$

using the fact that Q_C is negative for a typical heat engine.

Problem 32: Two moles of an ideal diatomic gas are taken around the cycle abc as shown in Figure 4. Data: $P_a = 1.00 \times 10^5 Pa$, $P_c = 1.40 \times 10^5 Pa$, $V_a = 0.049 m^3$. If the path from b to c is isothermal, calculate the work done by the gas. Answer: $-2300 J$

Problem 33: Two moles of an ideal diatomic gas are taken around the cycle abc as shown in Figure 4. Data: $P_a = 1.00 \times 10^5 Pa$, $P_c = 1.40 \times 10^5 Pa$, $V_a = 0.100 m^3$. If the path

Figure 4: A thermodynamic cycle's PV diagram.

from b to c is adiabatic, calculate the volume of the gas at point c . Answer: 0.13 m^3

Problem 34: A cylinder with a piston contains 0.150 mol of nitrogen at $1.80 \times 10^5 \text{ Pa}$ at 300 K . The nitrogen may be treated as an ideal diatomic gas whose molecules can rotate, but not vibrate. The gas is first compressed isobarically to half its original volume. It then expands adiabatically back to its original volume, and is finally heated isochorically to its original pressure. (i) Do the arrows on the corresponding PV diagram point clockwise or counterclockwise? Draw it. (ii) Compute the temperatures at the beginning and end of the adiabatic expansion. (iii) Compute the minimum pressure. (iv) Compute the work done by nitrogen during one thermodynamic cycle. Answer: (i) counterclockwise, (ii) 150 K , 114 K , (iii) $6.82 \times 10^4 \text{ Pa}$, (iv) -75 J

8.1 Otto Cycle

Most small automobiles and nearly all motorized recreational vehicles (excluding diesel and hybrid drives) operate on the *Otto cycle*. Inside Otto engines, a fuel-air mixture (approximately ideal gas) is compressed inside a piston and then ignited with a spark, resulting in rapid volume expansion - ‘blasting’ the piston outward. The spent fuel-air mixture is ejected from the piston and the process repeats. Engines in general have a different number of steps required to complete their thermodynamic cycles, with popular conventions being ‘two-stroke’ and ‘four-stroke’.

Let’s analyze an Otto engine just when a piston is fully expanded and filled with a fuel-air mixture (approximately ideal gas) ready to ignite. In this ‘cool’ state the volume V is a maximum, V_{max} . The temperature is T_0 , and we’ll label the pressure P_0 . the Otto cycle consists of four thermodynamic processes that follow, also illustrated in Figure 5.

1. Compression stroke: An external force (work left over from a previous cycle) rapidly compresses the piston to volume V_{min} . The pressure of the fuel-air mixture thereby increases to P_1 , and the temperature to T_1 . Because no significant heat is added to the cylinder during the compression stroke, it is an adiabatic process.
2. Fuel ignition: With the piston fully compressed, a spark ignites the fuel-air mixture, quickly releasing a great deal of kinetic energy Q_H into the gas. The pressure jumps

quickly to P_2 , and likewise the temperature to T_2 . The volume remains constant at V_{min} .

3. Power stroke: The cylinder's volume expands to V_{max} again, and whatever device is attached to the moving part of the piston meanwhile receives a great deal of organized work W (driving the car forward, for instance). The pressure drops during this process, but not quite down to P_0 ; call it P_+ . The spent fuel-air mixture is still 'hot' at this point, so the temperature becomes T_+ , slightly greater than T_0 .
4. Exhaust stroke: The cylinder expels the spent fuel-air mixture of energy Q_H to the environment and takes in a new fuel-air packet, returning the pressure to P_0 . The piston effectively remains at V_{max} (after intake of new fuel), and is ready again for a compression stroke. (Back to step 1.)

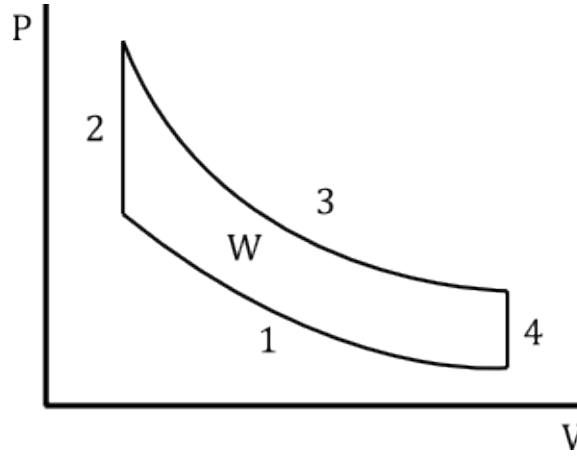


Figure 5: The Otto cycle.

The efficiency of the Otto engine is straightforward to calculate by (41). Heat input comes solely from igniting the fuel-air mixture, quantified as

$$Q_H = nC_V \Delta T_{ign} ,$$

where n is the number of moles of fuel-air mixture in the piston, and ΔT_{ign} is the instant temperature increase when the gas is ignited. After the power stroke, the spent fuel-air mixture returns to V_{max} but is still hot, and is ejected from the piston as exhaust. The heat ejected from the piston (per cycle) is

$$|Q_C| = nC_V |\Delta T_{rej}| ,$$

where ΔT_{rej} is temperature difference $T_0 - T_+$ between fresh fuel and spent fuel, respectively. Notice that Q_C 'wants' to be negative because ΔT_{rej} is negative. The proper statement really quantifies the heat entering, not leaving the piston. Thus,

$$Q_C = nC_V \Delta T_{rej} ,$$

which is indeed negative. The work output is unambiguously $Q = Q_H + Q_C$.

The efficiency then, by (41), is

$$e = 1 + \frac{Q_C}{Q_H} = 1 + \frac{\Delta T_{rej}}{\Delta T_{ign}} = 1 + \frac{(T_0 - T_+)}{(T_2 - T_1)},$$

but we can do better. The adiabatic transitions $(P, T)_0 \rightarrow (P, T)_1$ and $(P, T)_2 \rightarrow (P, T)_+$ obey the identities (36) and (37), which eliminate half of the T variables. While we're at it, define the *compression ratio* $r = V_{max}/V_{min}$. The formula for e boils down to

$$e = 1 - \frac{1}{r^{\gamma-1}}. \quad (42)$$

In the end we see the efficiency depends only on the compression ratio r and on the quality of the working substance through γ , the ratio C_P/C_V by equation (31).

Problem 35: How much of Q_C may be used to bolster to the power stroke of an engine? Answer: None, not even theoretically.

Problem 36: A gasoline engine takes in $8.00 \times 10^3 J$ of heat and delivers $1.60 \times 10^3 J$ of work per cycle. The heat is obtained from gasoline with a heat combustion of $4.6 \times 10^4 J/g$. (i) What is thermal efficiency? (ii) How much heat is discarded in each cycle? (iii) What mass of fuel is burned in each cycle? (iv) If the engine goes through 70.0 cycles per second, what is the power output in watts? (v) What is the power in horsepower? Answer: .212, $6.70 \times 10^3 J$, $1.85 \times 10^{-4} kg$, $1.62 \times 10^5 W$, 217 hp.

Problem 37: A Toyota Celica GT has a four cylinder Otto-cycle engine with a compression ratio of $r = 9.50$. The diameter of each cylinder, called the *bore* of the engine is 87.1 mm. The distance that the piston moves during the compression, called the *stroke* of the engine is 90.9 mm. The initial (and minimal) pressure of the air - fuel mixture (when the volume is maximal) is $8.50 \times 10^4 Pa$, and the initial temperature is 300 K (same as the environment). Assume that 200 J of heat is added to each cylinder in each cycle by burning the fuel (gasoline), which has $C_V = 20.5 Jmol^{-1}K^{-1}$ and $\gamma = 1.40$. (i) Calculate the work done in each cycle. (ii) Calculate the thermal energy released when the gas is cooled to the temperature of the outside air. (iii) Calculate the minimal volume of the air - fuel mixture. (iv) Calculate P , V , and T just before ignition. Answer: 119 J, 81.3 J, $6.05 \times 10^{-4} m^3$, $1.99 \times 10^6 Pa$, $6.37 \times 10^{-5} m^3$, 738 K.

8.2 Diesel Cycle

Some small automobiles and nearly all road shipping vehicles (excluding hybrid drives) operate on the *diesel cycle*. Like the Otto cycle, the diesel cycle entails four thermodynamic processes. Inside a diesel engine piston however, fuel-air mixture (approximated as ideal gas) is compressed until it self - ignites without the help of a spark, causing the power stroke. The ignition takes place in a constant - pressure condition, and the volume jumps near - instantly, much unlike the Otto cycle's ignition scenario. All other aspects of the two cycles are similar.

Problem 38: Figure 6 is the PV diagram for the diesel cycle. The compression stroke, the instant of fuel ignition, power stroke, and exhaust stroke are labeled 1, 2, 3, 4, respectively. Assuming steps 1 and 3 are adiabatic, show that the efficiency of the diesel

engine is given by

$$e = 1 - \frac{1}{\gamma} \frac{(c/b)^{-\gamma} - (c/a)^{-\gamma}}{(c/b)^{-1} - (c/a)^{-1}},$$

where volumes a , b , and c are indicated on the horizontal axis and γ is the ratio C_P/C_V .

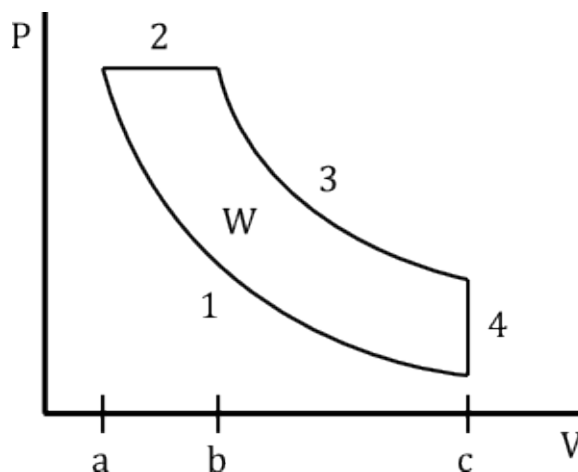


Figure 6: The Diesel cycle.

8.3 Carnot Cycle

Sadi Carnot (1796-1832) designed a hypothetical engine with maximal efficiency by using only isothermal and adiabatic processes. All heat transfer during the Carnot cycle occurs on isotherms; there is no heat transfer across finite temperature differences. No machine using the Carnot cycle would be ‘fast’ enough for industrial standards, but Carnot engines are still the most efficient type.

The gaseous working substance within a Carnot engine piston never leaves the piston, and undergoes no chemical change. Carnot engines rely on a literal ‘hot’ reservoir from which to draw energy, and a similar ‘cold’ reservoir to absorb energy. These two reservoirs may in practice be as simple as two containers of water - one hot and the other cold. The four stages of the Carnot cycle go as follows:

1. Slow process: A piston in the ‘compressed’ state enclosing a gas of volume V_{min} is attached to a reservoir at high temperature T_H and absorbs heat Q_H isothermally. The volume swells to V_1 and the reservoir is removed.
2. Fast process: The gas expands adiabatically, lowering its temperature from T_H until it reaches T_C . The volume swells once more to reach a maximum V_{max} .
3. Slow process: The system now contacts the cold reservoir of temperature T_C where isothermal compression takes the volume to V_2 (with $V_2 < V_1$). Reservoir removed.
4. Fast Process: The gas is compressed adiabatically, raising its temperature from T_C until it reaches T_H . The volume returns to V_{min} . (Back to step 1.)

We now calculate the efficiency of the Carnot engine. In one cycle, heat input contributes to volume expansion on an isotherm, governed by (33), or

$$Q_H = W_{12} = NKT_H \ln \left(\frac{V_1}{V_{min}} \right) .$$

The heat used in one cycle takes place during isothermal compression, or

$$Q_C = W_{34} = NKT_C \ln \left(\frac{V_2}{V_{max}} \right) = -NKT_C \ln \left(\frac{V_{max}}{V_2} \right) .$$

The two adiabatic paths obey identities (36) and (37), from which we can show

$$\frac{Q_C}{Q_H} = -\frac{T_C}{T_H} . \quad (43)$$

The efficiency by (41) is therefore

$$e = 1 - \frac{T_C}{T_H} . \quad (44)$$

It looks like Sadi Carnot was on to something. The efficiency of his cycle depends on neither the construction of an engine nor the properties of the working substance; only the reservoir temperatures are important.

Problem 39: Consider a Carnot engine operating between temperatures T_H and T_C , where T_C is above the ambient temperature T_0 . A second engine operates between the temperatures T_C and T_0 . Show that the overall efficiency of this system is equal to the efficiency of a single Carnot engine operating between T_H and T_0 .

8.4 Refrigerators

A refrigerator operates as a heat engine in reverse - work is required to displace heat from the cold reservoir to the hot reservoir. The working substance that circulates inside a refrigerator is a fixed amount of water, freon, or ammonia (no chemical exhaust). The operation is outlined as follows:

1. Starting in the gaseous phase, the working substance is forced through a compressor, emerging with reduced volume and increased temperature and pressure.
2. The substance goes through a long heat - conducting tube called the *condenser*, which is in thermal contact with the environment (hence the refrigerator needs to ventilate). In the condenser, the working substance loses its heat and changes phase to liquid.
3. The working substance then encounters a very small 'bottleneck' called the *throttle* that hinders most of the circulation. On the input side of the throttle, the temperature and pressure of the working substance are relatively high. The output side is at much lower temperature and pressure.
4. As a cold liquid, the working substance passes through the refrigerator body, warming again to the gaseous phase, and is pulled into the compressor. (Back to step 1.)

Like the heat engine, the refrigerator obeys the first law

$$Q_{in} = \Delta U + W ,$$

where Q_{in} is the sum $Q_H + Q_C$, the heat lost and gained through steps (2) and (4) above. Moreover, for cyclic processes the working substance obeys $\Delta U = 0$. Also note that $Q_H < 0$, $Q_C > 0$, and $W < 0$.

The merit of a refrigerator is not its efficiency, but instead the *coefficient of performance*, defined as the ratio of heat removed from the refrigerator body over the work required to do so. That is,

$$K = \frac{|Q_C|}{|W|} = \frac{|Q_C|}{|Q_H| - |Q_C|} . \quad (45)$$

Problem 40: A Carnot engine operating between two thermal reservoirs has an efficiency of e . When it is run in reverse, it becomes a Carnot refrigerator. Calculate the coefficient of performance. Answer: $(1 - e)/e$

Problem 41: An ideal air conditioner takes heat from a room at $68.5^\circ F$ and transfers it to the environment, which is at $93.1^\circ F$. For each joule of electrical energy required to operate the air conditioner, how many joules of thermal energy are removed from the room?

Problem 42: Liquid refrigerant at a pressure of $1.34 \times 10^5 \text{ Pa}$ leaves the expansion valve of a refrigerator at $-22^\circ C$. It then flows through the vaporization coils inside the refrigerator and leaves as vapor at the same pressure at a temperature of $-18^\circ C$, the same temperature as the inside of the refrigerator. The boiling point of the refrigerant at this pressure is $-18^\circ C$, the heat of vaporization is $1.60 \times 10^5 \text{ J/kg}$, and the specific heat capacity of the vapor is $485 \text{ J/kg}^{-1} K^{-1}$. The coefficient of performance of the refrigerator is $K = 2.52$. If 8.00 kg of refrigerant flows through the refrigerator each hour, find the electric power that must be supplied to the refrigerator. Answer: 123 W

Problem 43: A heat pump designed for southern climates extracts heat from the outside air, and delivers air at $20^\circ C$ to the inside of the house. Compute the coefficient of performance of the heat pump (i) in the south, where the outside temperature is $5^\circ C$; (ii) in the north, where the outside temperature is $-10^\circ C$. (iii) Two identical houses, one in the north and one in the south, are heated by this pump, and maintain temperatures of $20^\circ C$. Considering heat loss through the walls, windows, and roof, what is the ratio of the electrical power required to heat the two houses? Answer: (i) 18.5, (ii) 8.77, (iii) 4.0

9 Reversible Processes

Many processes in classical mechanics are *reversible*, where the system may backtrack along its trajectory in configuration space without violating any laws of physics or probabilistic notions. Harmonic oscillators and central force orbits are example systems that exhibit time - reversal symmetry (fundamentally unchanged if $t \rightarrow -t$), and are thus reversible.

In thermodynamics, only some types of processes are reversible, and only if conditions are right. One example of a reversible thermodynamic process is the isothermal phase change between water and ice at $0^\circ C$ in a sealed container.

Irreversible processes in thermodynamics are many - which is any process that takes a system into a permanent higher state of ‘disorder’. For instance, ignition in an engine piston sharply disorders the fuel-air mixture, and the probability of the exhaust molecules spontaneously reassembling into an unignited gas is immeasurably small. The previous configuration is essentially unattainable.

9.1 Free Expansion of Ideal Gas

Imagine a sealed, perfectly - insulating vessel with ideal gas contained within half of its total volume. The other half of the vessel is vacuum, and a thin membrane separates the two compartments. The membrane is then punctured and the gas is allowed to freely expand into the vacuum until the two compartments are in equilibrium.

The free expansion is an adiabatic process - no heat enters or leaves the vessel through the insulated walls. Also, the system does zero work while the gas expands into vacuum - no force is required. What happens to the temperature might not be directly intuitive, but the free expansion is indeed isothermal by (39). The internal energy change is therefore also zero. Free expansion of ideal gas is not quasistatic.

In summary, we see

$$Q = 0 \qquad W = 0 \qquad \Delta T = 0 \qquad \Delta U = 0 ,$$

clearly showing the energy state of the gas is unchanged. Yet, if we wanted to compress the system back to its original configuration, we *would* have to do work to do so, necessarily adding energy to the system. Free expansion of ideal gas is therefore irreversible.

We may say the ‘quality’ of the energy stored in the gas is higher before the free expansion. To illustrate, suppose a small windmill - like ‘turbine’ was placed on the vacuum side of the vessel. When punctured, gas rushing through the membrane can be used for spinning the turbine, exporting energy as useful work to whatever is connected to the turbine. This setup would be useless if the turbine were introduced after the expansion took place. The same energy would be in the vessel, but the energy would need to be ‘organized’ to spin the turbine.

9.2 Carnot Cycle Revisited

Looking more closely at the adiabatic and isothermal processes that make up the Carnot cycle, we are now equipped to notice that each step in the Carnot cycle is reversible. It follows that *any* reversible path in phase space may be approximated as a sequence of adiabatic and isothermal processes - a stack of Carnot cycles.

Recall the Carnot cycle obeys equation (43), namely

$$\frac{Q_C}{T_C} + \frac{Q_H}{T_H} = 0 .$$

For a closed path made of many adiabatic and isothermal segments, this generalizes to

$$\sum \frac{Q}{T} = 0 ,$$

where taking the infinitesimal limit smooths the sum into a line integral:

$$\oint \frac{dQ}{T} = 0 \tag{46}$$

The question now is, what is so important about the quantity dQ/T ?

10 Entropy, Second Law of Thermodynamics

We now quantify the amount of ‘disorder’ in a thermodynamic system. The new player is called the *entropy*, represented by S , and has several definitions. A suitable definition for our purposes applies to quasistatic processes only, and is contained in the differential relation

$$dS = \frac{dQ}{T} . \quad (47)$$

A stronger definition is $S = K \cdot \ln(\Omega)$, with Ω being the *multiplicity*, which is the number of microscopic configurations available to the system. (A discussion left for statistical mechanics.) In any case, S is proportional to the amount of energy unavailable to do work. Entropy is a state variable, like pressure or volume.

Problem 44: When 1.00 kg of water at 0°C is frozen to make ice, what is the entropy change? Answer: -1200 J/K

Problem 45: An object of mass m_1 , specific heat c_1 , and temperature T_1 is placed in contact with a second object mass m_2 , specific heat c_2 , and temperature $T_2 > T_1$. As a result, object 1 heats to temperature T and object 2 cools to T' . (i) Show that energy conservation requires that $m_1 c_1 (T - T_1) = m_2 c_2 (T_2 - T')$. (ii) Show that the entropy increase of the system is

$$\Delta S = m_1 c_1 \ln(T/T_1) + m_2 c_2 \ln(T'/T_2) .$$

10.1 Entropy Calculations

For reversible processes, the entropy change ΔS can be found by direct integration of (47). To compute ΔS for an irreversible process between two states, it suffices to find a path consisting of reversible process that would connect those two states and compute the change in entropy for that path. Why? It follows from (46) that a thermodynamic processes that takes a system from state a to state b over some path in phase space causes the same change in entropy ΔS_{ab} as any other choice of path.

Any process that is both adiabatic and quasistatic is *isentropic*, having zero entropy change.

Free Expansion of Ideal Gas

The entropy change ΔS for free expansion of ideal gas is irreversible and non - quasistatic with $Q = W = 0$ and $\Delta U = \Delta T = 0$. The entropy may calculated by the technique described above; finding a reversible process - namely isothermal expansion - that brings the system to the same final state as does the irreversible process.

For this problem however, the correct answer can be blindly (yet correctly) calculated: start with (47) and deploy the ideal gas law $PV = NKT$ and the first law $dQ = dU + PdV$, giving

$$\Delta S = \int \frac{dQ}{T} = NK \ln \left(\frac{V_f}{V_i} \right) .$$

Isothermal Expansion of Ideal Gas

Isothermal processes have $\Delta T = \Delta U = 0$, and the first law equates added heat with added work, $Q = W$. By (33), we know $Q = W = NKT \ln(V_f/V_i)$, and dividing by the

constant T gives the entropy increase,

$$\Delta S = NK \ln \left(\frac{V_f}{V_i} \right) ,$$

identical to the ΔS calculation for free expansion of ideal gas.

Adiabatic Process in Ideal Gas

Not all adiabatic processes are reversible, but those involving ideal gas typically are. Starting with (47), replace dQ with $dU + PdV$ according to the first law, using the ideal gas law and equation (23) to simplify. After a bit of algebra, arrive at

$$\Delta S = \int \frac{dQ}{T} = NK \ln \left(\frac{T_f^{(f/2)} V_f}{T_i^{(f/2)} V_i} \right) = 0 .$$

By equations (31) and (36), the argument inside the logarithm is unity, and we verify the isentropic claim: $\Delta S = 0$ for reversible adiabatic processes.

Entropy of Light

A ‘gas’ of photons is referred to as blackbody radiation. Recall that Stefan’s law, equation (11), reminds us that the rate of electromagnetic energy transfer is given by

$$\frac{dQ}{dt} = A\sigma (T_{obj}^4 - T_{env}^4) ,$$

where $e = 1$ and for simplicity we’ll take the T_{env} to be constant. The total energy in a volume V of photon gas is

$$Q = A\sigma (T_{obj}^4 - T_{env}^4) \int \frac{dx}{c} = \frac{V\sigma}{c} (T_{obj}^4 - T_{env}^4) ,$$

and it follows that $dQ = (V\sigma/c)4T_{obj}^3 dT$. Using equation (47), the entropy change ΔS as a function of ΔT_{obj} is

$$\Delta S = \int \frac{dQ}{T} = \frac{4V\sigma}{c} \int T^2 dT = \frac{4V\sigma}{3c} (T_{obj}^3 - T_0^3) .$$

For simplicity, take $T_{env} = 0$ and $T_0 = 0$. As a function of Q and V , the entropy ΔS becomes

$$S = \frac{4}{3} \frac{Q}{T} .$$

10.2 Second Law of Thermodynamics

The *second law of thermodynamics* is a restriction on the entropy of the universe. The statements that follow are equivalent statements of the second law.

Entropy statement

The entropy of a closed system can never decrease. The entropy of an open system can be forced to decrease, but the net entropy of the system + surroundings necessarily increases.

Engine statement

It is impossible for any system to undergo a cyclic process in which it absorbs heat at a single temperature and converts the heat completely into mechanical work. The efficiency of a heat engine is never unity.

Refrigerator statement

It is impossible for any process to have as its sole result the transfer of heat from a cooler to a hotter body. That is, heat never spontaneously flows from cold to hot.

11 Systems in Contact, Equilibrium

11.1 Thermal Equilibrium

Consider two thermodynamic systems A and B that are in separate, sealed containers that don't allow any fluctuations in pressure, volume, or number of particles contained. The systems are in thermal contact - allowing changes (by conduction or radiation) in temperature and hence the internal energy of A and B , but the total $U = U_A + U_B$ remains constant.

If the systems are initially at different temperatures $T_A > T_B$, sustained thermal contact will result in *thermal equilibrium* with a common temperature T_C for each system. This is an irreversible process with $T_A > T_C > T_B$. The total entropy $S = S_A + S_B$ necessarily increases as thermal equilibrium is reached, which we prove using the second law.

The heat transfer between the two systems is equal and opposite, and at any instant relates to temperature and entropy by (47), which means

$$T_A dS_A = -|dQ| \quad T_B dS_B = +|dQ| \quad .$$

Since the warmer system A has higher temperature than the cooler system B , conclude that dS_A is smaller than dS_B . That is, the entropy of the lower - temperature system is more sensitive to heat exchange, thus the sum $\Delta S_A + \Delta S_B$ is nonzero as we set to prove.

An equivalent statement comes from considering the derivative $\partial S / \partial U_A$, which equals zero at thermal equilibrium by the second law. Looking more closely, this means

$$\frac{\partial S_A}{\partial U_A} + \frac{\partial S_B}{\partial U_A} = 0 \quad .$$

Now, because $U = U_A + U_B$, it follows that $\partial U_B = -\partial U_A$, and hence

$$\frac{\partial S_A}{\partial U_A} = \frac{\partial S_B}{\partial U_B} \quad .$$

We already agree that T should be the same for conjoined systems in thermal equilibrium, and above we saw $\partial S / \partial U$ is the same for each. Let's seize the opportunity and actually *define* temperature:

$$T = \left(\frac{\partial S}{\partial U} \right)^{-1}_{V,N} \quad (48)$$

In this definition, the system's volume V and number of particles N are held constant.

11.2 Mechanical Equilibrium

Consider two thermodynamic systems A and B that are in separate containers with fixed number of particles at common temperature T . The containers are then joined by a moveable but impenetrable membrane separating A from B . The membrane finds a position where it feels zero net force, corresponding to equal pressures in each system. This condition is called *mechanical equilibrium*.

The second law says $S = S_A + S_B$ is maximum at mechanical equilibrium, which we may quantify by $\partial S / \partial V_A = 0$. Since $V = V_A + V_B$ is constant, infer $\partial V_B = -\partial V_A$, and conclude that

$$\frac{\partial S_A}{\partial V_A} = \frac{\partial S_B}{\partial V_B}$$

at equilibrium. We have seen the quantity $\partial S / \partial V$ has *something* to do with pressure, but the dimensionality is off. Conveniently, the dimension of $\partial S / \partial V$ is short by one factor of Kelvin units. It does no harm to multiply both sides by one and the same factor T , giving

$$T \frac{\partial S_A}{\partial V_A} = T \frac{\partial S_B}{\partial V_B} .$$

Alas, we're in position to relate pressure to entropy as we did the temperature. The equation

$$P = T \left(\frac{\partial S}{\partial V} \right)_{U,N} \quad (49)$$

does the trick for a system with fixed internal energy and fixed number of particles. Note that (49) is merely a tool for calculation, not a refined definition of pressure.

11.3 Diffusive Equilibrium

We now let systems A and B exchange both energy and particles through a porous membrane, but the volume of each individual system remains fixed. Assume initially that system A has larger P , V , N , and T than does system B . Placed in contact, the systems achieve *diffusive equilibrium* in addition to thermal equilibrium.

The second law is satisfied when $\partial S / \partial N_A = 0$ where $S = S_A + S_B$ and $N = N_A + N_B$, implying $\partial N_B = -\partial N_A$. We may write the differential entropy per particle as

$$\frac{\partial S_A}{\partial N_A} = \frac{\partial S_B}{\partial N_B} .$$

Sticking a factor of the equilibrium temperature $T = T_A = T_B$ on each side shifts the dimension to energy units. While we're at it, make that factor $-T$ to follow textbook convention. The conclusion is

$$-T \frac{\partial S_A}{\partial N_A} = -T \frac{\partial S_B}{\partial N_B} .$$

The quantity $-T \cdot \partial S / \partial N$ has a special name called the *chemical potential*, denoted by the greek μ as:

$$\mu = -T \left(\frac{\partial S}{\partial N} \right)_{U,V} \quad (50)$$

By construction, μ is the energy required to raise N by one unit in a system with constant T , U , and V . It follows from (50) that particles flow from a system with higher μ to a system with lower μ .

When there are multiple particle species present in a system, μ adopts a subscript i for each species, and the quantity μdN becomes $\sum \mu_i dN_i$.

12 Thermodynamic Potentials

12.1 Enthalpy

In order to create a system of volume V that has internal energy U in an environment at pressure P , the total energy required is called the *enthalpy*, defined as

$$H = U + PV . \quad (51)$$

Enthalpy is the internal energy of a system plus the work required to shove the environment out of the way in order to make room the system. Conversely, H is the total energy recovered if you could annihilate the system.

12.2 Helmholtz Free Energy

A system assembled in isothermal conditions may borrow energy from the environment (easing the job of the assembler) by an amount equal to TS , where T is the temperature and S is the entropy of the system. The net work required is therefore

$$A = U - TS , \quad (52)$$

the *Helmholtz Free Energy*. If the system is annihilated in constant - temperature and constant - volume conditions, A is the energy recoverable.

12.3 Gibbs Free Energy

In conditions where both temperature and pressure are constant, creating a system requires the *Gibbs free energy*

$$G = U + PV - TS , \quad (53)$$

which means you supply the internal energy U and the work PV to shove the environment out of the way. The environment supplies TS for you, hence the minus sign. In constant T and P conditions, G is the energy recovered if the system is annihilated.

Bringing chemical potential into the discussion, you may suspect that μ is the proportionality factor between the particle count N and the energy of the system, but *which* energy? It will turn out that we're looking for the Gibbs free energy with the proper formula being, at fixed temperature and pressure,

$$G = \mu N . \quad (54)$$

12.4 Grand Free Energy

One more grouping of terms heavily used in describing irreversible processes in open systems is the *grand free energy* or *grand potential*, defined as

$$\Phi = U - TS - \mu N . \quad (55)$$

The quantity (55) is the energy recovered by reducing the system to zero size - putting all particles and energy back into the reservoir.

Using equation (54), it's straightforwardly shown that

$$\Phi = -PV . \quad (56)$$

12.5 Thermochemistry Tables

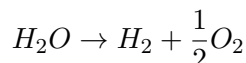
When a thermodynamic system is changed, the response in enthalpy H , entropy S , and Gibbs free energy G varies per constituent species and per surrounding conditions. Such state variables are used heavily in the sciences, particularly chemistry, and have been measured for many known substances. Data is condensed into *thermochemistry tables*, also called entropy tables.

At ‘standard’ room temperature (298 K) and at ‘standard’ atmospheric pressure (1 atm), also known as STP conditions, the state variables ΔH , S , and ΔG of common substances behave as listed in the following table:

Species	Name	Enthalpy ΔH [kJmol ⁻¹]	Entropy S [Jmol ⁻¹ K ⁻¹]	Gibbs free energy ΔG [kJmol ⁻¹]
H_2O	Liquid Water	- 285.83	69.95	- 237.15
H_2O	Water Vapor	- 241.83	188.84	- 228.59
H_2	Hydrogen Gas	0	130.7	0
O_2	Oxygen Gas	0	205.07	0
O	Monatomic Oxygen	249.170	161.055	231.731
N	Nitrogen Gas	472.704	153.298	455.563
NH_3	Ammonia	- 45.90	192.77	- 16.37

Electrolysis of Water

Consider a 1 mol sample of liquid water at STP conditions. When electric current is applied through the sample, the reaction



takes place. We will use the ΔH - and S - data from the table above to calculate the change in Gibbs free energy ΔG for this system. That is, what energy must the current - driving apparatus supply?

The enthalpy ΔH of the reaction is needed first, which is, for one mole,

$$\Delta H = -286 \text{ kJ}$$

by the above table. Of the 286 kJ produced, an amount of work equal to $1 \text{ atm} \times \Delta V = 4 \text{ kJ}$ must go into the atmosphere to make room for the new gases, with 282 kJ remaining in the system. The net entropy change for the reaction is

$$S = \left(131 + \frac{205}{2} - 70 \right) \text{ JK}^{-1} = 163 \text{ JK}^{-1},$$

which is evidently positive. Thus, the environment contributes energy equal to $TS = 298 \times 163 \text{ J} = 49 \text{ kJ}$. The change in Gibbs free energy is therefore

$$\Delta G = -286 \text{ kJ} + 49 \text{ kJ} = -237 \text{ kJ},$$

verifying the top row of the above table, and answering the question at hand.

13 Thermodynamic Identities

13.1 First Thermodynamic Identity

The first law of thermodynamics (32) is cited frequently throughout this study. In light of the chemical potential, the first law should be modified to include an extra term, yielding *the first thermodynamic identity*:

$$dU = TdS - PdV + \mu dN \quad (57)$$

Note especially that dQ has been replaced with TdS , and it's understood that we're working under the quasistatic approximation.

13.2 Second Thermodynamic Identities

In analog to (57), it would be nice to come up with similar differential relations for enthalpy, Helmholtz free energy, and Gibbs free energy.

Beginning with enthalpy $H = U + PV$, the differential form dH is

$$dH = dU + PdV + VdP .$$

Replacing dU with (57) and simplifying gives the dH - identity

$$dH = TdS + VdP + \mu dN . \quad (58)$$

A similar exercise starting with Helmholtz free energy $A = U - TS$ yields the dA - identity

$$dA = -SdT - PdV + \mu dN . \quad (59)$$

Cranking the Gibbs free energy $G = U + PV - TS$ through the same ritual delivers the dG - identity

$$dG = -SdT + VdP + \mu dN . \quad (60)$$

Note that each of these results may be referenced as the second thermodynamic identity.

13.3 Thermodynamic Variables as Partial Derivatives

Various derivatives of thermodynamic identities (57) - (60) yield extremely useful formulae for thermodynamic variables that have the appearance of (48), (49), and (50). Following are only several of the common relations:

$$\begin{aligned} S &= - \left(\frac{\partial A}{\partial T} \right)_{V,N} & S &= - \left(\frac{\partial G}{\partial T} \right)_{P,N} \\ P &= - \left(\frac{\partial A}{\partial V} \right)_{T,N} & V &= \left(\frac{\partial G}{\partial P} \right)_{T,N} \\ \mu &= \left(\frac{\partial A}{\partial N} \right)_{T,V} & \mu &= \left(\frac{\partial G}{\partial N} \right)_{T,P} \end{aligned}$$

Defined in terms of partial derivatives, the coefficient of thermal expansion, namely $\beta = 3\alpha$ from equation (3) may be written as

$$\beta = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_P .$$

A similar property of matter is the isothermal compressibility κ_T , defined as

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T .$$

Problem 46: Assume some equation of state (not necessarily the ideal gas equation) connects a system's pressure P , temperature T , and volume V . For the infinitesimal change $P \rightarrow P + dP$, $T \rightarrow T + dT$ with V fixed, (i) show that $(\partial V / \partial P)_T dP + (\partial V / \partial T)_P dT = 0$. (ii) Next show that

$$\frac{dP}{dT} = \left(\frac{\partial P}{\partial T} \right)_V = -\frac{(\partial V / \partial T)_P}{(\partial V / \partial P)_T} .$$

Problem 47: Around room temperature, liquid water has a thermal expansion coefficient $\beta = 2.5 \times 10^{-4} K^{-1}$ and isothermal compressibility $\kappa_T = 4.5 \times 10^{-10} Pa^{-1}$. Estimate the pressure that must be exerted on a drop of water in order to prevent it from expanding while increasing from $20^\circ C$ to $30^\circ C$. Use this result to explain why it is easier to measure C_P rather than C_V . Answer: $55.8 atm$

Problem 48: For a van der Waals fluid obeying equation (26) at fixed T and N , use the second thermodynamic identity (60) to show that the Gibbs free energy is

$$G = -NKT \ln(V - Nb) + \frac{N^2 b K T}{V - Nb} - \frac{2aN^2}{V} + C(T) . \quad (61)$$

13.4 Maxwell Relations

The first derivative of a thermodynamic potential yields some special quantity. It turns out that the second derivative of a thermodynamic potential is also significant. From calculus, we know that successive derivatives of some general function Φ with respect to variables x, y yields a result independent of the order of x, y . That is,

$$\frac{\partial}{\partial x} \frac{\partial \phi}{\partial y} = \frac{\partial}{\partial y} \frac{\partial \phi}{\partial x} .$$

wfb i think this whole paragraph is bullshit

13.5 Extensive and Intensive Quantities

All variables in thermodynamics fall into one of two classifications. A quantity may be *extensive*, where the scale of the variable goes with the size of the system, or it may be *intensive*, where the scale of the variable doesn't care about the scale of the system. To illustrate, consider a thermodynamic system with the laundry list of state variables. If the system is duplicated and attached to itself, which variables duplicate, and which don't? You should conclude:

Extensive: M, V, N, U, A, G, S

Intensive: P, T, μ, ρ

Properties of intensive and extensive variables:

- An extensive quantity multiplied by an intensive quantity yields an extensive quantity.

- The sum of two extensive quantities is also extensive.
- A ratio of extensive quantities yields an intensive quantity.
- The product of two extensive quantities is rare in calculations - check twice if you encounter this.
- The sum of an extensive quantity and an intensive quantity should never occur.

13.6 Linearity between G and N

Consider the dG - identity (60) when T and P are constant. We write

$$\mu = \left(\frac{\partial G}{\partial N} \right)_{T,P} ,$$

meaning that adding one particle increases a system's energy by μ . As many particles are added, μ does not change so long as T and P do not change. Integration of the above therefore delivers equation (54), namely $G = \mu N$.

You might wish that a similar relation connects A to N , by starting with $\mu = (\partial A / \partial N)_{T,V}$. However, adding particles while maintaining fixed T and V forces the density ρ to slowly increase, which directly changes μ .

14 Equilibrium and Energy Minimization

Consider a thermodynamic system with entropy S_{sys} in an environment having entropy S_{env} . The entropy of the universe obeys

$$S_{uni} = S_{sys} + S_{env},$$

with the differential version reading $dS_{uni} = dS_{sys} + dS_{env}$. Proceed by replacing dS_{env} using the first thermodynamic identity, giving

$$dS_{uni} = dS_{sys} + \left[\frac{1}{T} (dU + PdV - \mu dN) \right]_{env} . \quad (62)$$

By the second law, S_{uni} is always increasing. If the system comes to equilibrium with the environment, it follows that dS_{uni} approaches zero. (Drop the subscript $_{sys}$ for now on.)

14.1 Helmholtz Minimum

Suppose a system with fixed V and N is placed in an environment of constant T . By (62) we have

$$dS_{uni} = \frac{1}{T} (TdS) + \frac{1}{T} (dU_{env}) ,$$

and notice that $dU_{env} = -dU$ by energy conservation. Looking at the differential version of the Helmholtz free energy (52), we may replace $TdS - dU$ by $-dA$, or

$$dS_{uni} = -\frac{1}{T} (dA)_{T,V,N} . \quad (63)$$

The lesson: A system with fixed V and fixed N in an environment at constant T will do whatever it can to minimize the Helmholtz free energy.

14.2 Gibbs Minimum

In an environment of constant T and P , suppose a system is designed such that N is fixed, but V is allowed to change. The combined entropy of the system plus its environment equals the entropy of the universe, S_{uni} . Starting again with (62) leads to

$$dS_{uni} = \frac{1}{T} (TdS) + \frac{1}{T} (dU_{env} + PdV_{env}) .$$

From energy and volume conservation, it follows that $dU_{env} = -dU$ and $dV_{env} = -dV$. Looking at the differential version of the Gibbs free energy (53), we may replace $TdS - dU - PdV$ by $-dG$, or

$$dS_{uni} = -\frac{1}{T} (dG)_{P,T,N} . \quad (64)$$

The lesson: A system with fixed N in an environment at constant T and constant P will do whatever it can to minimize the Gibbs free energy. It follows that if T is allowed to vary, the system instead minimizes the Enthalpy.

14.3 Grand Minimum

Suppose a thermodynamic system of fixed V is allowed to exchange both U and N with its environment. Assume the chemical potential μ of the system is already in equilibrium with the environment, as is the temperature T . That is, we must interpret equation (62) as written.

When the conditions T , V , and μ are each constant, the differential version of the grand free energy (55) reads

$$d\Phi = dU - TdS - \mu dN ,$$

which is precisely bracketed term in (62). It immediately follows that

$$dS_{uni} = -\frac{1}{T} (d\Phi)_{T,V,\mu} . \quad (65)$$

The lesson: A system with fixed V in thermal and diffusive equilibrium with the environment will do whatever it can to minimize the grand free energy.

15 Introduction to Statistical Mechanics

In order to gain more significant insight into thermodynamics we must restart the entire subject from a highly magnified point of view. At small scales (and for cold systems) the effects of quantum mechanics are essential. For starters, we dismiss the notion that matter is made of continuous substance. The famous ‘gold foil’ experiment by Rutherford around 1910 demonstrated that matter is mostly vacuum, sparsely populated with discrete atomic nuclei, each separated from its neighbors by a length much greater than the diameter of the nucleus itself.

To throw the notion of true exactness out the window, the Heisenberg uncertainty relation $\Delta x \Delta p > \hbar/2$ forbids simultaneous knowledge of the position and the momentum of any particle. It makes sense therefore to discuss large thermodynamic systems in terms of their statistics. Observables such as pressure, temperature, volume, etc. should emerge as statistical averages of many-particle behavior.

15.1 Microstates and Macrostates and Multiplicity (oh my!)

Consider a system of N molecules contained in volume V . Each particle follows some path $\vec{r}(t)$ and carries momentum $\vec{p}(t)$. For each time t , a ‘mental snapshot’ of all \vec{r} and \vec{p} constitutes a *microstate*. A microstate is a precise list of each particle’s location, momentum, spin, or any other thermodynamically relevant information. One single microstate may contain a vast amount of information, and a system may skip through microstates very rapidly.

To illustrate, suppose our box has uniform pressure P on all of the walls. During any pressure measurement the system goes through many microstates, yet all of them correspond to P . The pressure is thus classified as a *macrostate*. It would be extremely unlikely for the pressure to momentarily read zero, perhaps corresponding to all molecules occupying one half of the box and not the other. Such a microstate is allowed, but would practically never occur. It seems that the *observed thermodynamic average is the macrostate corresponding to the greatest number of microstates*.

The system also possesses ‘global’ variables such as internal energy U , temperature T , entropy S , and so on. While a great multitude of microstates are sampled, the system’s ‘occupation’ of U , T , and S may change slowly, if at all. We also classify these as macrostates, which are some kind of average that depends on the system’s occupation of microstates.

The number of microstates that correspond to a given macrostate is called the *multiplicity*, denoted by Ω , the greek symbol ‘Omega’. In thermodynamic systems, there are generally far more microstates available to a system than there are macrostates. Just *how many* microstates correspond to a given macrostate has a deep consequence in thermodynamics.

We’ll see that the name of the game in statistical physics is to calculate the number of microstates Ω that correspond to some macrostate.

16 Combinatorics

The goal is to attain familiarity with counting the number of discrete microstates available to a general system, an application of a branch of mathematics called *combinatorics*.

16.1 Party of Strangers

Suppose there are N people occupying a room, and no two people are yet acquainted. When any two people shake hands, those two become mutually acquainted. Question: For all N occupants to become mutually acquainted, how many handshakes $\Omega(N)$ must take place?

Answer: One given person in the crowded room must shake $N - 1$ hands in order to meet everyone else, so for starters $\Omega \propto (N - 1)$. All N people must shake $N - 1$ hands, so $\Omega \propto N$ also. Incidentally each pair of people will have become acquainted twice, so the final answer needs to be divided by two. Conclude that

$$\Omega = \frac{1}{2}N(N - 1) \ .$$

16.2 Four-Digit Passwords

A typical combination lock has two macrostates: open and closed. The essence of a combination lock is the majority of microstates correspond to the ‘closed’ position, and just one microstate that corresponds to ‘open’.

Imagine you are tasked with guessing the combinations that open an assortment of locks. I won’t present here an algorithm for opening locks in general (ask me in person about that), but I argue it’s useful to know the total number of microstates Ω that correspond to the closed position possible for a given lock. The fewer the possibilities, the sooner you might guess the combination.

In the examples that follow, each lock A , B , C , and D is unlocked by a four-digit combination.

Lock A

Suppose lock A allows each of the four digits in its combination to be occupied by any integer between 0 and 9, inclusive (like a short phone number). The surest way to attack this kind of problem is to start listing all allowed combinations in some kind of sequence, and hopefully pick out a pattern that helps uncover the total.

Writing the possible combinations for lock A is equivalent to writing the sequence:

$$0000, 0001, 0002, 0003, \dots, 9997, 9998, 9999$$

The total number of combinations obviously amounts to

$$\Omega_A = 10,000 = 10^4 .$$

That is, a random guess at lock A ’s password has a $1/10,000$ chance of being correct.

Lock B

Next consider lock B , designed similarly to lock A but the allowed integers span only 1 to 4, inclusive. The multiplicity of this lock must be smaller than that of lock A because integers 5-9 are excluded. However, the previously-used sequence technique

$$1111, 1112, 1113, \dots, 4442, 4443, 4444$$

avails no shortcuts for counting the total number of possible combinations.

To proceed we first solve a simplified version of the problem: suppose the combination is instead a two-digit number with the same integers 1 through 4 available. It’s easy to write out the sixteen available combinations for the two-digit lock:

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

Now observe that the list of four-digit combinations is every possible pairing of the sixteen elements of the two-digit combination. Therefore,

$$\Omega_B = 16^2 = 4^4 .$$

Corollary

Locks A and B exhibit the property where any given digit in the combination is independent of any other digit - the numbers aren't consumed as they are used. If n equals the number of digits in a combination and m is the number of integers available for a given digit, then the total number of combinations for locks of type A and B is

$$\Omega = m^n .$$

Lock C

Lock C has a four-digit combination of integers 1-4, but each integer can only be used once in the combination. (The buttons stay down when pressed.) Since no two digits in the combination are allowed to be equal, we expect $\Omega_C < 4^4$, as the formula $\Omega = m^n$ produces a number too large. In sequence form, the available combinations are:

$$1234, 1243, 1324, \dots, 4231, 4312, 4321$$

Each combination has the integers 1-4 'consumed' in a different order. For choosing the first digit, all four integers are available, so we know at least $\Omega_C \propto 4$. The second digit in the combination will be one of 3 remaining integers, indicating $\Omega_C \propto 4 \cdot 3$. The third digit has 2 possibilities (gain a factor of two), and there is no freedom in the last digit (gain a factor of one).

Evidently, the total number of combinations possible for lock C is:

$$\Omega_C = 4 \times 3 \times 2 \times 1 = 4!$$

The aspect of lock C making the multiplicity so simple is that the number of digits in the combination happens to equal the number of integers available. Generalizing to n -digit combinations with precisely n integers available per digit is straightforward:

$$\Omega = n!$$

Lock D

Lock D has a four-digit combination consisting of integers between and including 0 and 9, but no two integers in the combination are equal.

Choosing the first digit in lock D 's combination, there are 10 integers available. As no repeated integers are allowed, there are 9 available integers for the second digit. For the third and fourth digits, there are 8 and 7 available integers, respectively. Each possible combination leaves six of the integers unused.

The number of allowed combinations must be:

$$\Omega_D = 10 \times 9 \times 8 \times 7 = \frac{10!}{6!} ,$$

which generalizes to

$$\Omega = \frac{m!}{(m-n)!} ,$$

where m is the number of integers available to the lock type whose combination consists of n nonrepeated digits.

16.3 Arrangements, Permutations, and Combinations

The number A_m of all arrangements of m distinguishable objects is:

$$A_m = m! \quad (66)$$

Any number p of identical objects in the set of m objects induces a factor of $p!$ in the denominator on the right side of (66). This is understood by casting the repeated elements so they are not unique (with say, prime or tilde notation), and then the double-counting is corrected with factors of $1/p!$.

The number P_m^n denotes the number of permutations in where n objects can be chosen in a specific order out of m objects, and is equal to:

$$P_m^n = \frac{m!}{(m-n)!} \quad (67)$$

Finally, we'll soon prove that the number C_m^n denotes the number of combinations in where n objects can be chosen out of m objects with order being unimportant. This is:

$$C_m^n = \frac{m!}{n!(m-n)!} \quad (68)$$

Note that the 'party of strangers' result $N(N-1)/2$ is a special case of (68). If there are $N = m$ unacquainted people in a system, there are C_m^n necessary 'handshakes' required for total friendship. A handshake is defined as any pairing of $n = 2$ people.

16.4 Two-State Coin

Let's devise a game that involves tossing a balanced (a.k.a. *fair*) coin for many trials - an example of a *two-state system*. Each time the coin lands on 'heads' (denoted H) your score increases by 1 (starting from zero). If the coin lands on tails (denoted T), your score decreases by 1 (negative score is allowed).

If this game runs for N hypothetical trials, the net score is some integer k . For example, a game with for consecutive 'heads' has $k = 4$, and a game with four consecutive 'tails' has $k = -4$. For a net score of $k = 0$ after four tosses, there are several ways to do this: $HHTT$, $HTHT$, $TTHH$, and so on.

The main question: What is the number of microstates $\Omega(k, N)$ that correspond to the macrostate k when the coin has been tossed N times? I'll show that the counting trick is to visualize the entire output of the game: if there were N tosses, assume there were N coins, each tossed once.

Consider the outcome where all N coins have landed on 'tails', corresponding to $k = -N$. There is only $\Omega = 1$ way this result can happen - rearranging the trials changes nothing. If instead the final score were $-N + 1$, this means any one of the coin tosses resulted in 'heads'. Since there were N tosses in total, we deduce that $\Omega(-N+1, N) = N$.

If any two of the tosses show 'heads' with the rest being 'tails', the multiplicity is proportional to $N(N-1)$. The result must be divided by 2 because the two 'heads' outcomes could occur in any order (much unlike the four-digit passwords of the previous example):

$$\Omega(-N+2, N) = \frac{N(N-1)}{2}$$

Next we bring the total score to $k = -N + 3$, corresponding to three 'heads' results among $N - 3$ 'tails'. This has obvious consequences for Ω , namely the numerator gains

a factor of $N - 2$, corresponding to the number of available coins (that are not already ‘heads’), and the denominator gains a factor of 3 by identical arguments that gave us the 2 previously.

Generalizing the pattern, we see that adding m points to the original score of $k = -N$ gives multiplicity

$$\Omega(-N + m, N) = \frac{N(N-1)(N-2)(N-3)\cdots(N-m+1)}{2 \cdot 3 \cdot 4 \cdots (m)} .$$

In tighter notation, the denominator is simply m - factorial and the numerator is N - factorial divided by $(N - m)$ - factorial, which is

$$\Omega(-N + m, N) = \frac{N!}{m!(N-m)!} .$$

Sometimes the above is written as

$$\Omega(m) = \frac{N!}{m!(N-m)!} = \binom{N}{m} , \quad (69)$$

and spoken ‘ N choose m ’ in statistics lingo. As a reality check, it’s readily verified that $\Omega = 1$ for both extreme cases $m = 0$ (all ‘tails’) and $m = N$ (all ‘heads’). The maximum in Ω should occur when there are equal numbers of ‘heads’ - up coins and ‘tails’ - up coins, in which case $m = N/2$. A plot of Ω versus m is sharply peaked and symmetric about $m = N/2$.

We now ask a bigger question: For a game of N trials, what is the *total* multiplicity Ω_T across all possible scores k ? The total multiplicity of the system is equal to the sum of each individual multiplicity:

$$\Omega_T = \sum_{m=0}^N \Omega(-N + m, N) = \sum_{m=0}^N \frac{N!}{m!(N-m)!}$$

Evaluating the sum looks time - consuming at best, and avoiding it would be ideal. We are in luck for this type of system: each of the N coin tosses can go one of two ways, so the total multiplicity is just the number 2 multiplied by N instances of itself, as

$$\Omega_T = 2^N$$

which reveals an interesting identity relating 2^N to a sum of factorials:

$$\sum_{m=0}^N \frac{N!}{m!(N-m)!} = 2^N \quad (70)$$

The multiplicity $\Omega(k, N)$ can be written more neatly for export. First note the variable m is equivalent to the number of ‘heads’ - up coins, N_{\uparrow} . Observe next that the total coins N must equal $N_{\uparrow} + N_{\downarrow}$, where N_{\downarrow} is the number of ‘tails’ - up coins. So far then, Ω looks like

$$\Omega(N_{\uparrow}, N) = \frac{N!}{N_{\uparrow}!(N - N_{\uparrow})!} = \frac{N!}{N_{\uparrow}!N_{\downarrow}!} , \quad (71)$$

but we can do better. By substituting $N/2 = N_{\uparrow} - k$ and $N/2 = N_{\downarrow} + k$, we have a clear expression of the multiplicity of a two-state system with equilibrium value $k = 0$:

$$\Omega(k, N) = \frac{N!}{(N/2 - k)!(N/2 + k)!} \quad (72)$$

17 Probability Theory

Probability theory is a tool for understanding systems whose characteristics are generally (i) not precisely known, and/or (ii) too numerous to discuss on an element-by-element basis. The probability P for the occurrence of some event E is written as $P(E)$, defined as some number between 0 and 1, inclusive. If the total multiplicity is Ω_T for the entire system, the probability of finding the system in state E is

$$P(E) = \frac{\Omega(E)}{\Omega_T}, \quad (73)$$

where $\Omega(E)$ is the multiplicity of state E .

17.1 Definitions

- If two events A and B are independent, the joint probability is

$$P(A \text{ and } B) = P(A \cap B) = P(A)P(B).$$

For example, if two coins are flipped the chance of both showing ‘tails’ is $P(T \cap T) = (1/2)(1/2) = 1/4$.

- If events A and B are mutually exclusive, the probability of either event occurring is

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

For example, the chance of rolling a 3 or 4 on a six-sided die is $(1/6) + (1/6) = 1/3$.

- When there is a risk of double-counting the microstates available to a given macrostate, events are not mutually exclusive, and the term $P(A \text{ or } B)$ is corrected:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

For example, when drawing a single card at random from a 52-card deck, the chance of getting a heart or a face card J, Q, K (or one that is both) is $(13/52) + (12/52) - (3/52) = 11/26$.

- The conditional probability of event A occurring under given condition B is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

- Baye’s theorem states that the inversion of $P(A|B)$ may be written as

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}.$$

- The probability of event Q occurring under the set of conditions $\{A_j\}$ may be expressed as a probabilistic expansion:

$$P(Q) = \sum_j P(Q|A_j)P(A_j)$$

17.2 Questions in Probability Theory

Problem 49: Consider the library of letters $B\$ = ACEFGILMNTUX$. We construct a ‘word’ by removing at random a letter from $B\$$, and placing it in line with the next letter removed from $B\$$ until the library is empty. Compute the probability of arranging the word $W\$ = MAGNETICFLUX$. Answer: $\Omega = 12! \rightarrow P(W\$) = 1/12!$

Problem 50: Consider the library of letters $C\$ = FLUXELECTRIC$. We construct a ‘word’ $X\$$ by removing at random a letter from $C\$$, and placing it in line with the next letter removed from $C\$$ until the library is empty. Compute the probability of arranging the word $X\$ = ELECTRICFLUX$. Answer: $\Omega = 12!/(2!2!2!) \rightarrow P(X\$) = 8/12!$

Problem 51: From a 52-card deck, how many 5-card poker hands are possible? What is the probability of randomly drawing a ‘royal flush’ 10-J-Q-K-A in any one suit? Answer:

$$\Omega = \frac{52!}{5!(52-5)!} \rightarrow P = \frac{4}{\Omega} = \frac{1}{649740} \approx 0.00000154$$

Problem 52: A fair coin is tossed N times. Calculate the probability of getting precisely N_{\uparrow} heads. Answer:

$$P(N_{\uparrow}) = \frac{N!}{2^N N_{\uparrow}! (N - N_{\uparrow})!}$$

Problem 53: From a 52-card deck, you draw three cards. Compute the probability that these cards read King - Queen - Jack, in that order (mixed suits allowed). Solve this by two methods: permutations and conditional probability. Answer:

$$\Omega = \frac{52!}{(52-3)! 4^3} \rightarrow P = \frac{4^3}{52 \times 51 \times 50}$$

$$P = P(K) P(Q|K) P(J|(Q|K)) = \frac{4}{52} \times \frac{4}{51} \times \frac{4}{50}$$

Problem 54: Compute the probability of drawing three Kings in a row from a 52-card deck. Solve this by two methods: permutations and conditional probability. Answer:

$$\Omega = \frac{52!}{3!(52-3)!} \frac{3!(4-3)!}{4!} \rightarrow P = \frac{24}{52 \times 51 \times 50}$$

$$P = P(A) P(A|A) P(A|(A|A)) = \frac{4}{52} \times \frac{3}{51} \times \frac{2}{50}$$

Problem 55: From a 52-card deck, compute the probability of getting any ‘pair’ (QQ, 77, AA, etc.) drawn in a 7-card hand. Answer:

$$P = 1 - \prod_{n=1}^6 \frac{52-4n}{52-n} \approx 0.79$$

The Birthday Problem

Problem 56: Consider a classroom of total population N . Question: What is the probability that any two people were born on the same day? Answer:

This problem shall be solved heuristically for a general audience. Begin with the trivial case $N = 2$, in where there is a $1/365$ chance of a common birthday:

$$P(2) = \frac{1}{365} = 1 - \frac{364}{365}$$

The result is written in the form $1 - X$ so we may focus on X , the probability of *no* common birthday.

A third person entering the system, making $N = 3$, has $365 - 2 = 363$ available days to avoid a common birthday. The probability becomes

$$P(3) = 1 - \frac{364}{365} \cdot \frac{363}{365},$$

and the pattern is now obvious. For total population N , the probability that some pair of people share a birthday is:

$$P(N) = 1 - \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{(365 - N + 1)}{365} = 1 - \frac{365!}{365^N (365 - N)!}$$

Note that X has been expressed as a recursion of conditional probabilities

$$X(n|n-1) = \frac{365 - (n-1)}{365} \quad \rightarrow \quad X(N) = \prod_{n=2}^N X(n|n-1),$$

which could also have been written directly by the permutation formula

$$X(N) = \frac{P_{365}^N}{365^N} = \frac{365!}{365^N (365 - N)!}.$$

Following is a list of various populations N with their corresponding $P(N)$:

N	P(N)
5	2.71%
10	11.7%
20	41.1%
23	50.7%
30	70.6%
50	97.0%

Laptop Repair Shop

(Taken from CMPSCI240 UMass Amherst 2013.)

Problem 57: You work in a laptop repair shop. 80% of laptops brought in have been dropped, 15% of laptops have had a drink spilled on them, and 5% of laptops have a variety of other problems. A customer drops off a laptop and doesn't tell you what happened to it. You notice the laptop is emitting a slight coffee - like smell. Based on your knowledge of broken laptops, you estimate that 20% of dropped laptops have a slight coffee - like smell, 65% of laptops that have had something spilled on them have a slight coffee - like smell, and 5% of laptops that have some other problem have a slight coffee - like smell.

Provide labels for the events described in the problem. Find the probability that the laptop had something spilled on it given that it has a slight coffee - like smell. Answer:

Denote D for ‘drop’, S for ‘spill’, and O for ‘other’. Let the letter F denote ‘slight coffee - like smell’. The information provided in the problem may be written:

$$\begin{aligned} P(D) &= .80 & P(S) &= .15 & P(O) &= .05 \\ P(F|D) &= .20 & P(F|S) &= .65 & P(F|O) &= .05 \end{aligned}$$

We further deduce:

$$P(F) = P(F|D)P(D) + P(F|S)P(S) + P(F|O)P(O) = .26$$

To answer the question, we need to compute $P(S|F)$, which is the inversion of $P(F|S)$. Thus:

$$P(S|F) = \frac{P(S)P(F|S)}{P(F)} = .375$$

Problem 58: After closer inspection, you note that the laptop has no cracks on the case. Based on your knowledge of broken laptops, you estimate that 80% of dropped laptops have cracked cases, 11% of laptops that have had something spilled on them have cracked cases, and 9% of laptops that have some other problem have cracked cases. If the probability that a laptop smells like coffee and the probability that a laptop has a cracked case are conditionally independent of each other given the case of the damage (drop, spill, or other), what is the probability that the laptop had something spilled on it if it has a slight coffee - like smell and no cracks in the case? Answer:

The problem asks us to evaluate $P(S|F \cap Z^C)$, where Z denotes ‘crack’ and Z^C denotes ‘no crack’. Due to the independence between F and Z^C , the term $P(F \cap Z^C|S)$ decouples into $P(F|S)P(Z^C|S)$. Using Baye’s theorem and the expansion formula we write

$$P(S|F \cap Z^C) = \frac{P(S)P(F|S)P(Z^C|S)}{\sum_{i=D,S,O} P(F \cap Z^C|X_i)P(X_i)},$$

and meanwhile the problem has stated

$$P(Z|D) = .80 \quad P(Z|S) = .11 \quad P(Z|O) = .09,$$

or equivalently:

$$P(Z^C|D) = 1 - .80 \quad P(Z^C|S) = 1 - .11 \quad P(Z^C|O) = 1 - .09$$

Evaluating $P(S|F \cap Z^C)$ is now straightforward:

$$P(S|F \cap Z^C) = \frac{.15 \times .65 \times .89}{.20 \times .20 \times .80 + .65 \times .90 \times .15 + .05 \times .91 \times .05} = .7169$$

18 Statistical Systems

Any system that contains very much ‘information’ is suitable for statistical analysis. For instance, the temperature T , pressure P , and density ρ of a gas are observable macrostates of the constituent molecules. During a single measurement of an observable (think of a pressure gauge), the system skips *randomly* through a great multitude of microstates. The sum of such measurements are a kind of system average.

18.1 Discrete and Continuous Distributions

In any statistical system, the probability of observing macrostate k is equal to the multiplicity of that macrostate divided by the total multiplicity of the system:

$$P(k) = \frac{\Omega(k)}{\Omega_T}$$

To remove any ambiguity, Ω_T is understood to be one of

$$\Omega_T = \sum_k \Omega(k) \quad \text{or} \quad \Omega_T = \int_{\text{all } k} d\Omega(k) ,$$

depending on whether $\Omega(k)$ is discrete or continuous.

It follows from (73) that the sum of probabilities across the whole range of macrostates k must equal unity. That is,

$$\sum_{\text{all } k} P(k) = \frac{\Omega(k_1) + \Omega(k_2) + \Omega(k_3) + \dots}{\Omega_T} = \frac{\Omega_T}{\Omega_T} = 1 ,$$

or in integral form,

$$\int_{\text{all } p} dp = \int_{\text{all } k} \left(\frac{dp}{dk} \right) dk = 1 .$$

The quantity (dp/dk) is called the *probability density*, represented by $p(k)$. The property that all probabilities add to unity is called *normalization*, and is enforced by either

$$\sum_k P(k) = 1 \quad \text{or} \quad \int_{\text{all } k} p(k) dk = 1 . \quad (74)$$

The probability density has uses beyond writing the normalization condition. By definition, the probability of measuring macrostate k between k_a and k_b is the integral of the probability density function over the interval. That is,

$$P(k_a < k < k_b) = \int_{k_a}^{k_b} p(k) dk . \quad (75)$$

For a given probability distribution, the average value of observable q , written as $\langle q \rangle$, is defined as

$$\langle q \rangle = \sum_q q \cdot P(q) \quad \text{or} \quad \langle q \rangle = \int_{\text{all } q} q \cdot p(q) dq , \quad (76)$$

for discrete and continuous distributions, respectively. In a similar manner, the mean square of an observable can be calculated by replacing q with q^2 , giving

$$\langle q^2 \rangle = \sum_q q^2 \cdot P(q) \quad \text{or} \quad \langle q^2 \rangle = \int_{\text{all } q} q^2 \cdot p(q) dq . \quad (77)$$

The degree of ‘smearing’ in observations of q is formally called the *variance*, defined as

$$\text{Var}(q) = \langle q^2 \rangle - \langle q \rangle^2 . \quad (78)$$

A more popular quantity that keeps track of smearing around a central average is the *standard deviation*, which has both discrete and continuous definitions

$$\sigma = \sqrt{\sum_{i=1}^N (q_i - \langle q \rangle)^2 \cdot P(q_i)} \quad , \quad \sigma = \sqrt{\int_{\text{all } q} (q - \langle q \rangle)^2 \cdot p(q) dq} , \quad (79)$$

and is simply the square root of the variance (proof left to the reader).

18.2 Random Variables

Two random variables x, y with probability distributions $p_x(x), p_y(y)$ are *independent* if $p(x, y)dxdy = p_x(x)p_y(y)$, where $p(x_0, y_0)dxdy$ is the probability to find the variables x, y at intervals dx, dy , respectively around x_0, y_0 .

Consider a series of random variables $\{x_k\}$ where $1 \leq k \leq N$. The relation

$$X = \sum_k x_k$$

also satisfies

$$\langle X \rangle = \sum_k \langle x_k \rangle, \quad (80)$$

which is proven as follows: Assume two random variables x, y (special case $k = 2$). The probability to find x in the vicinity x_0 in the interval dx (independent of y) is $p_x(x_0)dx$. Let us call $p_{x,y}(x_0, y_0)dxdy$ that probability that the two variables simultaneously are within a window of dx, dy of x_0, y_0 . Proceed by writing

$$p(x) = \int_{-\infty}^{\infty} p_{x,y}(x, y) dy.$$

Then, using $\langle X \rangle = \langle x + y \rangle$, we may write

$$\langle X \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} (x + y) p_{x,y}(x, y) dy,$$

which becomes

$$\langle X \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} (x) p_{x,y}(x, y) dy + \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} (y) p_{x,y}(x, y) dy,$$

simplifying to $\langle X \rangle = \langle x \rangle + \langle y \rangle$. Generalizing to k variables, verifying the above relation is trivial.

Problem 59: Consider the relation

$$Var(X) = \langle X^2 \rangle - \langle X \rangle^2,$$

where $X = \sum_k x_k$. Show that if the variables $\{x_k\}$ are independent, then

$$Var(X) = \sum_k [Var(x_k)]. \quad (81)$$

Answer: Assume two random variables x, y (special case $k = 2$):

$$Var(X) = \langle (x + y)^2 \rangle - \langle x + y \rangle^2$$

Using equation (80),

$$Var(X) = \langle x^2 \rangle + \langle y^2 \rangle + 2\langle xy \rangle - \langle x \rangle^2 - \langle y \rangle^2 - 2\langle x \rangle \langle y \rangle,$$

where if x, y are independent then the cross terms cancel, meaning

$$Var(x) = Var(x) + Var(y).$$

Generalizing to k variables, verifying the above relation is trivial.

18.3 Large Numbers and Stirling's Approximation

The preceding examples in combinatorics probability theory have introduced a variety of counting techniques for the multiplicity Ω of a system. All is well until we try manipulations of Ω , due largely to the common occurrence of factorial ($X!$ -like) terms. For instance, try to puzzle out a simple answer to:

$$\frac{d}{dx} (x!) = ?$$

If a simple derivative is so ugly, we should avoid being freighted with factorial terms as we confront the subject of statistical mechanics.

The goal is to find a new way to write $N!$ for the large- N approximation, and the formula that does this job is the *Stirling approximation*. There are many derivations of this formula, and the one I present here is not the simplest. Begin with the definition of the *gamma function*,

$$\Gamma(z) = (z-1)! = \int_0^\infty t^{z-1} e^{-t} dt, \quad (82)$$

from which we may write

$$n! = \int_0^\infty e^{-x} x^n dx.$$

Looking at out the $e^{-x} x^n$ term, notice one derivative reveals this is a very sharply peaked function at $x = n$. The function is so sharp in fact, that we use this to approximate $x \approx n$ under the integral. That is, suppose that $x = n + \epsilon$, where ϵ is very small compared to n . The quantity $e^{-x} x^n$ becomes $e^{-(x+\epsilon)} (x+\epsilon)^n$. Next take the natural logarithm of our working quantity

$$\ln(e^{-x} x^n) = \ln(e^{-(x+\epsilon)} (x+\epsilon)^n),$$

and let the algebra bloom on the right side. Along the way, use the expansion

$$\ln(1+y) \approx y - \frac{y^2}{2} + \frac{y^3}{3} - \dots \quad y \ll 1 \quad (83)$$

to land at

$$\ln(e^{-x} x^n) \approx n \ln n - n - \frac{\epsilon^2}{2n} + \frac{\epsilon^3}{3n^2}.$$

With no predilection to poetry, exponentiate and integrate.

$$\int_0^\infty e^{-x} x^n dx \approx \int_0^\infty e^{n \ln n - n - \frac{\epsilon^2}{2n} + \frac{\epsilon^3}{3n^2}} d\epsilon$$

The integral on the left is simply N -factorial, which follows from the gamma function. The right hand side needs a few adjustments before we proceed. The integration variable is no longer x (held $\approx n$), but is now ϵ . Since the curve under the integral is so narrow, it does no harm to extend the lower limit from 0 to $-\infty$. The problem is reduced to

$$n! \approx \int_{-\infty}^\infty e^{n \ln n - n - \frac{\epsilon^2}{2n} + \frac{\epsilon^3}{3n^2}} d\epsilon = e^{n \ln n - n} \int_{-\infty}^\infty e^{-\frac{\epsilon^2}{2n} + \frac{\epsilon^3}{3n^2}} d\epsilon.$$

Finally, let's agree that the ϵ^3 term is negligible. The remaining (Gaussian) integral evaluates to $\sqrt{2\pi n}$, yielding the anticipated Stirling's approximation,

$$n! \approx e^{n \ln n - n} \sqrt{2\pi n} = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}. \quad (84)$$

The log of (84) is another common form of Stirling's formula:

$$\ln(n!) \approx n \ln n - n + \ln(\sqrt{2\pi n}) . \quad (85)$$

In this form, it's evident that the $\sqrt{2\pi n}$ term can be much smaller than the leading terms, and sometimes the shortened versions will often suffice in problem solving:

$$n! \approx e^{n \ln n - n} \quad \ln(n!) \approx n \ln n - n \quad (86)$$

18.4 Gaussian Distribution

The most common and powerful tool in statistics is known as the *Gaussian distribution*, informally called the *Bell curve*.

To derive the mathematical apparatus we begin, strangely enough, by considering is the multiplicity of the two-state system,

$$\Omega(k, N) = \frac{N!}{\left(\frac{N}{2} - k\right)! \left(\frac{N}{2} + k\right)!} .$$

This function is continuous for large N and is symmetric about its very sharp peak at $k = 0$. Away from $k = 0$, the multiplicity tends to zero quickly. Next, Stirling's approximation (85) lets us replace each factorial term in Ω , giving

$$\ln \Omega \approx N \ln N - \left(\frac{N}{2} - k\right) \ln \left(\frac{N}{2} - k\right) - \left(\frac{N}{2} + k\right) \ln \left(\frac{N}{2} + k\right) + \frac{\ln(2/\pi N)}{2} ,$$

where the $\ln()$ terms have been stacked and the simple N and k terms have canceled out. After a bit of algebra with help from (83), the answer boils down to

$$\ln(\Omega(k, N)) \approx N \ln 2 - \frac{2k^2}{N} + \frac{1}{2} \ln \left(\frac{2}{\pi N} \right) .$$

Exponentiating both sides delivers the beautiful answer for the multiplicity of the two-state system with large N :

$$\Omega(k, N) \approx 2^N e^{-2k^2/N} \sqrt{\frac{2}{\pi N}} \quad (87)$$

Equation (87) is composed of three significant factors. Recall first that 2^N is the count of total microstates available to a two-state system. Next we have the anticipated 'Gaussian' exponential term $e^{-2k^2/N}$ containing all of the k - dependence, which is responsible for the overall shape of Ω . The final term is $\sqrt{2/\pi N}$, which at first appears as a pesky constant, but is soon to become a very lucky constant.

Dividing $\Omega(k, N)$ by $\Omega_T = 2^N$ yields, by definition, the probability density function of a two-state system,

$$p(k) = \frac{\Omega(k, N)}{\Omega_T} = \left(\frac{2}{\pi N} \right)^{1/2} e^{-2k^2/N} . \quad (88)$$

We see that the 'pesky' constant $\sqrt{2/\pi N}$ is precisely what's needed to satisfy the normalization condition (74), namely $\int p(k) dk = 1$. The details of normalization entail the fact that

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} , \quad (89)$$

which can be proved by squaring the integral and switching to polar coordinates.

Let us continue with a slightly more generalized version of $p(k)$, supposing the maximum in $p(k)$ occurs not at $k = 0$, but at the constant $k = a$ such that

$$p(k) = \left(\frac{2}{\pi N}\right)^{1/2} e^{-2(k-a)^2/N}. \quad (90)$$

With $p(k)$ on - hand, it's naturally interesting to calculate $\langle k \rangle$ and $\langle k^2 \rangle$:

$$\langle k \rangle = \int_{-\infty}^{\infty} k \cdot p(k) dk = \sqrt{\frac{2}{\pi N}} \int_{-\infty}^{\infty} k \cdot e^{-2(k-a)^2/N} dk = a$$

$$\langle k^2 \rangle = \int_{-\infty}^{\infty} k^2 \cdot p(k) dk = \sqrt{\frac{2}{\pi N}} \int_{-\infty}^{\infty} k^2 \cdot e^{-2(k-a)^2/N} dk = a^2 + \frac{N}{4}$$

The $\langle k \rangle$ integral is solved by making the substitution $k - a = x$ which splits the problem into two terms: one term is the integral of an odd function over a symmetric interval and contributes nothing, and the surviving term is just a . The $\langle k^2 \rangle$ integral involves the same substitution $k - a = x$, but one more mathematical insight is needed. Look again at (89) and differentiate both sides with respect to a (not x), resulting in

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2} dx = \frac{\sqrt{\pi}}{2a^{3/2}}. \quad (91)$$

The standard deviation σ is readily calculated, coming out as:

$$\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2 \quad \rightarrow \quad \sigma = \sqrt{\frac{N}{4}} \quad (92)$$

With this, we may write a cleaner definition of the normalized Gaussian distribution in terms of its mean and standard deviation:

$$p(k) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} e^{-(k-\langle k \rangle)^2/2\sigma^2} \quad (93)$$

Problem 60: The multiplicity of the two-state system is given by

$$\Omega(x, N) = A(N) e^{-2x^2/N},$$

where N is a very large number. Determine the constant $A(N)$ by two different methods: (i) use your knowledge of the true multiplicity Ω , and (ii) treat the system as a Gaussian distribution. Answer: $A(N) = 2^N \sqrt{2/\pi N}$

18.5 Binomial Distribution

We established previously that the multiplicity of the two-state coin is

$$\Omega(k, N) = \frac{N!}{k!(N-k)!},$$

where N is the number of trials and k is the number of outcomes corresponding to 'heads'. Now we generalize the above to accomodate the 'unfair' coin. Suppose the probability

of the 'heads' outcome is p , and correspondingly for tails, $1 - p$. The multiplicity Ω becomes

$$\Omega(k, N, p) = p^k (1 - p)^{N-k} \frac{N!}{k! (N - k)!} . \quad (94)$$

The above result applies to more than coins. The probability of having k successful outcomes of any event with weighting p after N trials is given by (94).

One unfortunate aspect of (94) is that the large- N behavior is ambiguous at first sight. Proceed by the following method: define a two-state random variable z_j that is equal to one if event A with weight p occurs in the j -th trial, and is equal to zero otherwise. The average value of z_j is

$$\langle z_j \rangle = P(A) \hat{z}_j^A + P(\text{not } A) \hat{z}_j^{\text{not } A} = p \cdot 1 + (1 - p) \cdot 0 = p ,$$

and furthermore the average of z_j^2 is

$$\langle z_j^2 \rangle = p \cdot 1^2 + (1 - p) \cdot 0^2 = p .$$

Next, the total number of occurrences k_A of event A involves a sum over all N trials as

$$k_A = \sum_{j=1}^N z_j ,$$

and it follows that

$$\langle k_A \rangle = N \langle z_j \rangle = Np \quad \quad \langle k_A^2 \rangle = Np^2 .$$

In accordance with equation (78), the variance in k_A shall be denoted σ^2 and is given by

$$\sigma = \sqrt{Np(1 - p)} , \quad (95)$$

which is equivalent to (92) for $p = 1/2$. Note the overall form $\sigma \propto \sqrt{N}$, a common feature of large- N systems.

18.6 Radioactive Decay

An unstable atom is one that undergoes a decay, such as by ejecting a nucleon, positron, or gamma ray. Having no internal time - keeping mechanism, an unstable atom is entirely 'unaware' of its absolute age. Suppose we begin observation of an unstable atom at $t = 0$, and that we need to know the probability that the atom 'survives' until time t .

Begin by writing the conditional probability of the atom decaying in a small time window Δt after time $t > 0$

$$P_{\Delta t/t}^{\text{decay}} = \tau^{-1} \Delta t ,$$

where τ^{-1} is an 'average lifetime' constant related to (but not precisely equal to) the half - life of the element, defined such that $\Delta t \ll \tau$. The probability of the atom being 'still alive' in the interval Δt is correspondingly

$$P_{\Delta t/t}^{\text{alive}} = 1 - \Delta t/\tau .$$

Next decompose the 'alive' case into a product of conditional probabilities, namely by slicing the time t into n copies of the short interval Δt as

$$P^{\text{alive}}(t) = P_{\Delta t/t_1}^{\text{alive}} \cdot P_{\Delta t/t_2}^{\text{alive}} \cdots P_{\Delta t/t_n}^{\text{alive}} = \left(1 - \frac{t}{n\tau}\right)^n .$$

Letting $n \rightarrow \infty$, and also using the identity

$$\lim_{n \rightarrow \infty} \left(1 + \frac{B}{n}\right)^n = e^B, \quad (96)$$

we find the probability that a single unstable atom will have an observed lifetime t starting from $t = 0$:

$$P(t) = e^{-t/\tau} \quad (97)$$

18.7 Poisson Distribution

Imagine counting the number of water molecules that pass a point in a river, the flow of which being completely characterized by the average speed v . Over time interval t , measured molecule count ‘should’ be vt , but this exact value should rarely be attained. Instead, due to the nature of fluctuations in the river, one would measure a spread of values about vt . Naturally we wonder, what is the probability $P_k(t)$ that k molecules are measured over the interval t ?

Proceed by the same method as done in the ‘Radioactive Decay’ above. First slice the interval t into $N \rightarrow \infty$ copies of the very short interval Δt . It follows that the number of particles per ‘bin’ in Δt decreases and limits at zero, implying

$$P_0(\Delta t) \gg P_1(\Delta t) \gg P_2(\Delta t) \gg P_3(\Delta t) \gg \dots$$

Moreover, each k - bin in Δt is its own weighted two-state system, with the probability of the ‘filled’ state being $v\Delta t$. This is an application of the binomial distribution (94), so immediately write

$$P_k(t) = (v\Delta t)^k (1 - v\Delta t)^{N-k} \frac{N!}{k!(N-k)!},$$

and the variables N and $\Delta t = t/N$ must be eliminated. In the limit $N \rightarrow \infty$, we have

$$\lim_{N \rightarrow \infty} \left(1 - \frac{vt}{N}\right)^{N-k} = e^{-vt},$$

and because

$$\frac{N!}{k!(N-k)!} = N(N-1)(N-2)\dots = N^k + \dots,$$

we also have

$$\lim_{N \rightarrow \infty} \left(\frac{vt}{N}\right)^k \frac{N!}{k!(N-k)!} = (vt)^k.$$

Redfining the product vt as a ‘dimensionless’ parameter q , we assemble the remarkably common *Poisson distribution*:

$$P_k(q) = e^{-q} \frac{q^k}{k!} \quad (98)$$

Going back to the original example, $P_k(t)$ is the probability that exactly k of a river’s molecules have passed a fixed point in elapsed time t if the overall river speed is v . What remains to be found is the average count $\langle k \rangle$, and also the spread in that number, given by $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$. By brute force, we have:

$$\langle k \rangle = e^{-q} \sum_{k=0}^{\infty} k \frac{q^k}{k!} = e^{-q} q \frac{d}{dq} \sum_{k=0}^{\infty} \frac{q^k}{k!} = e^{-q} q \frac{d}{dq} e^q = q$$

$$\begin{aligned}\langle k^2 \rangle &= e^{-q} \sum_{k=0}^{\infty} k^2 \frac{q^k}{k!} = e^{-q} q \frac{d}{dq} q \frac{d}{dq} \sum_{k=0}^{\infty} \frac{q^k}{k!} = e^{-q} q \frac{d}{dq} q \frac{d}{dq} e^q = q(1+q) \\ \Delta k = \sigma &= \sqrt{q^2 + q} - q = \sqrt{q}\end{aligned}\tag{99}$$

19 Microcanonical Ensemble

The *microcanonical ensemble* is a ‘stage’ for describing an isolated thermodynamic system - one that is restricted to having fixed energy, fixed volume, and fixed number of particles. The fundamental assumption of the microcanonical ensemble (a corollary of the preceding sections) is that *all microstates are equally accessible*.

19.1 Multiplicity and Entropy

Now we address how the multiplicity Ω is related in thermodynamics. We actually *define* the entropy as

$$S = K \ln(\Omega) , \tag{100}$$

where the constant K is the Boltzmann constant; the same that appears in the ideal gas equation (12), carrying dimensions *Joule/Kelvin*.

We’ll see subsequently that definition (100) is completely consistent with all notions of entropy established previously.

19.2 Microcanonical Systems in Contact

Why we define S to be the natural log of Ω , as opposed to say, $S \propto \Omega$, becomes evident when two thermodynamic systems $A(\Omega_A, S_A)$, $B(\Omega_B, S_B)$ are placed in contact. When joining A and B to create system C , the entropy and multiplicity must satisfy (i) $S_C = S_A + S_B$ and (ii) $\Omega_C = \Omega_A \Omega_B$, respectively. Any multiplicity Ω is therefore called a *super - extensive* quantity. In order to satisfy (i) and (ii), the only reasonable definition is $S = K \ln(\Omega)$.

Two thermodynamic systems placed in contact eventually reach some kind of equilibrium: thermal, mechanical, chemical, etc. It’s appropriate to check that two microcanonical systems also adhere to such notions.

Consider two systems A and B , with energy, multiplicity, and entropy represented by U_i , $\Omega_i(U_i)$, $S_i(U_i)$ with $i = (A, B)$. Placed in contact, A and B will exchange energy while maintaining the total $U_C = U_A + U_B$. It follows that U_A and U_B are not independent, and it makes sense to express all quantities in terms of one favored energy variable. We choose U_A , thereby replacing all instances of U_B , namely

$$U_B = U_C - U_A \quad , \quad dU_B = -dU_A .$$

The combined multiplicity is the product

$$\Omega_C(U_A) = \Omega_A(U_A) \Omega_B(U_C - U_A) ,$$

and the total entropy is

$$S(U_A) = K \ln(\Omega_C(U_A)) .$$

To proceed we expand the total entropy near thermal equilibrium by Taylor series, giving

$$S(U_A) \approx S(\bar{U}_A) + \left(\frac{\partial S}{\partial U_A} \right)_{\bar{U}_A} (U_A - \bar{U}_A) + \frac{1}{2!} \left(\frac{\partial^2 S}{\partial U_A^2} \right)_{\bar{U}_A} (U_A - \bar{U}_A)^2 + \dots,$$

where the energy U_A is near its equilibrium value \bar{U}_A . The term that is linear in U_A drops out, as

$$\left(\frac{\partial S}{\partial U_A} \right)_{\bar{U}_A} = K \left(\frac{\partial}{\partial U_A} \ln \Omega_A + \frac{\partial}{\partial U_B} \ln \Omega_A \right) = \frac{\partial S_A}{\partial U_A} + \frac{\partial S_A}{\partial U_B} = \frac{\partial S_A}{\partial U_A} - \frac{\partial S_B}{\partial U_B},$$

which conceals a statement about thermal equilibrium. Recall equation (48), namely $T = (\partial S / \partial U)^{-1}$, and we have

$$\left(\frac{\partial S}{\partial U_A} \right)_{\bar{U}_A} = \frac{1}{T_A} - \frac{1}{T_B} = \frac{1}{T} - \frac{1}{T} = 0,$$

which tells us the temperature tends to become the same for systems A and B .

The quadratic term in $S(U_A)$ does not suffer a similar cancelation. In fact, one more U_A derivative shows us, at thermal equilibrium,

$$\left(\frac{\partial^2 S}{\partial U_A^2} \right)_{\bar{U}_A} = -\frac{1}{T_A^2} \frac{\partial T_A}{\partial U_A} + \frac{1}{T_B^2} \frac{\partial T_B}{\partial U_A} = -\frac{1}{T^2} \left(\frac{1}{(C_V)_A} + \frac{1}{(C_V)_B} \right),$$

where $(C_V)_i$ is the respective system's heat capacity. Now define the variable σ such that

$$\frac{1}{\sigma^2} = - \left(\frac{\partial^2 S}{\partial U_A^2} \right)_{\bar{U}_A},$$

so that the entropy simplifies to

$$S(U_A) \approx S(\bar{U}_A) - \frac{1}{2\sigma^2} (U_A - \bar{U}_A)^2.$$

The combined multiplicity is actually a Gaussian distribution

$$\Omega_C(U_A) = e^{S(U_A)/K} \approx e^{S(\bar{U}_A)/K} e^{-(U_A - \bar{U}_A)^2 / 2\sigma^2}, \quad (101)$$

with probability density and normalization condition, respectively,

$$p(U_A) = \left(\frac{1}{2\pi\sigma^2} \right)^{1/2} e^{-(U_A - \bar{U}_A)^2 / 2\sigma^2}, \quad \int_{\text{all } U_A} p(U_A) dU_A = 1. \quad (102)$$

It's often convenient to deal in dimensionless integration variables, accomplished by relabeling the energy $U_A = \epsilon u_A$, where ϵ has units of energy and u_A is a dimensionless number. Equations (102) now take the form

$$p(u_A) = \left(\frac{\epsilon^2}{2\pi\sigma^2} \right)^{1/2} e^{-\epsilon^2(u_A - \bar{u}_A)^2 / 2\sigma^2}, \quad \int_{\text{all } u_A} p(u_A) du_A = 1.$$

With the energy represented by a dimensionless number, look once more at the combined multiplicity,

$$\Omega_C(u_A) \approx e^{S(\bar{u}_A)/K} e^{-\epsilon^2(u_A - \bar{u}_A)^2 / 2\sigma^2}.$$

Now establish the tautological connection between the magnitude $e^{S(\bar{u}_A)/K}$ and the width σ . To get from $\Omega(u_A)$ to $p(u_A)$, recall the definition of a probability, (73), from which we obtain

$$p(u_A) = \frac{\Omega(u_A)}{\sum_{u_A} \Omega_C(u_A)} = \frac{e^{S(\bar{u}_A)/K}}{\sum_{u_A} \Omega_C(u_A)} e^{-\epsilon^2(u_A - \bar{u}_A)^2/2\sigma^2} .$$

Evidently,

$$\frac{e^{S(\bar{u}_A)/K}}{\sum_{u_A} \Omega_C(u_A)} = \left(\frac{\epsilon^2}{2\pi\sigma^2} \right)^{1/2} ,$$

which may or may not be a useful identity in problem solving.

19.3 Ergodic Hypothesis

A fundamental assumption of microcanonical ensemble theory is that all microstates are equally available to a system. We've seen so far that a system is most likely to exhibit the macrostate that has the greatest number of corresponding microstates. Such comments refer to the phase space of a given system - any observable is a kind of 'ensemble average'.

Meanwhile, the concept of *time* has been implied but not specified. We would hope that *for any given observable, the ensemble average and the time average are in agreement*, known as the *ergodic hypothesis*.

Suppose an observable f is some real - valued function on the phase space Γ that is integrable with respect to the microcanonical ensemble measure μ . The time average of f is defined as

$$\bar{f} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(x(t)) dt , \quad (103)$$

which must equal the ensemble average

$$\langle f \rangle = \int_{\Gamma} f(x) d\mu(x) \quad (104)$$

in order to satisfy the ergodic hypothesis.

To demonstrate the point, consider a single one-dimensional harmonic oscillator that has hamiltonian

$$H(q, p) = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2 .$$

The equation of motion of the oscillator is

$$q(t) = A \cos(\omega t) ,$$

where if the total energy of the system is E , then $A = \sqrt{2E/m\omega^2}$ (proof left to the reader). The first order of business is to determine \bar{q}^2 , the time-average mean-square position of the oscillator. (We calculate the square of q because $\bar{q} = 0$ for the harmonic oscillator.) This amounts to solving the integral

$$\bar{q}^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T q(t)^2 dt .$$

Substituting $q(t)$ and evaluating delivers

$$\bar{q}^2 = \left(\frac{A^2}{2} \right) \lim_{T \rightarrow \infty} \left(1 + \frac{\sin(2\omega T)}{2\omega T} \right) = \frac{A^2}{2} = \frac{E}{m\omega^2} .$$

Now we must check that the time average $\bar{q^2}$ is equal to the ensemble average $\langle q^2 \rangle$, which is calculated using (104). Unpacking the definition, start with

$$\langle f(q, p) \rangle = \frac{\int \rho(q, p) f(q, p) dq dp}{\int \rho(q, p) dq dp}, \quad (105)$$

where $f(q, p) = q$ and the distribution function $\rho(q, p)$ is given by

$$\rho(q, p) \propto \delta(H(q, p) - E).$$

Next we exploit the fact that when $\phi(y_0) = 0$,

$$\int \delta(\phi(y)) dy = \frac{1}{\phi'(y_0)}. \quad (106)$$

Pick out the integral in the denominator of (105) for a moment - using (106) as a hint, this can be solved while treating q as a constant. That is,

$$\int \rho(q, p) dq dp \propto \int \delta(H(q, p) - E) dq dp = \int \left(\frac{1}{p/m} \right) \Big|_{\frac{p^2}{2m} + \frac{m\omega^2}{2} q^2 - E = 0} dq.$$

The variable p is gone after evaluation, and the remaining q integral is

$$\sqrt{\frac{2}{mE}} \int \frac{dq}{\sqrt{1 - \frac{m\omega^2 q^2}{2E}}}.$$

Before we go wild substituting $q = A \cos(\omega t)$ into the above, remember this integral is to be evaluated over the whole phase space, so for now recast the quantity ωt in the parameter θ . With this in mind, discover that

$$\int \rho(q, p) dq dp \propto \frac{2}{m\omega} \int_0^\pi d\theta = \frac{2\pi}{m\omega}.$$

That does it for the denominator of (105). Evaluating the numerator is no different until the q -integral, where we now have an extra q^2 inside the integrand:

$$\int \rho(q, p) f(q, p) dq dp \propto \sqrt{\frac{2}{mE}} \int \frac{q^2 dq}{\sqrt{1 - \frac{m\omega^2 q^2}{2E}}} = \frac{2}{m\omega} A^2 \int_0^\pi \cos^2(\theta) d\theta = \frac{A^2 \pi}{m\omega}$$

Finally, $\langle q^2 \rangle$ can be assembled from (105), giving

$$\langle q^2 \rangle = \frac{A^2 \pi}{m\omega} \frac{m\omega}{2\pi} = \frac{A^2}{2},$$

matching the answer for $\bar{q^2}$ as predicted by the ergodic hypothesis.

More Mathematics

There are several other hypotheses and theorems to develop at this stage to ensure we're 'allowed' the use of certain mathematical constructs in subsequent analyses. However, I omit the topics of Liouville's theorem, Poisson brackets, and Hamiltonian formalism from this document - these items fit more nicely in a study devoted to classical mechanics.

20 Paramagnetism

The simplest application of microcanonical ensemble theory is paramagnetism. Magnetism is a phenomenon caused by circulating electric currents in matter or vacuum. Most solids generally fall into one of three categories: paramagnetic, diamagnetic, or ferromagnetic.

Ferromagnetic materials have the strongest coupling to magnetic phenomena, and are most familiar to everyday experience. Inside a ferromagnetic material, atoms are arranged into macroscopic *domains*, wherein electrons orbit their respective nuclei such that the angular momentum vector of each electron points the same way. A strong permanent magnet consists of many domains, where any given cluster of domains has a net magnetic moment that is strong in a direction we decide to call ‘north’.

Paramagnetic materials exhibit weak affinity to magnetism, and neither type carries a permanent magnetic field. For our purposes, a paramagnetic material is one that responds attractively to an external magnetic field. That is, electrons inside the material choose their direction of orbit in order to increase the total magnetic field near the paramagnet. In addition, a paramagnetic material is attracted to an external field and moves toward it.

A diamagnetic material acts in the opposite manner as a paramagnet - it tries to oppose an external magnetic field and is repelled from it. These phenomena only occur when the temperature of the material is low enough - otherwise thermal noise prevents domain alignment.

20.1 Thermodynamics of a Paramagnet

Consider a large- N paramagnetic material made of spin $1/2$ particles, each having magnetic moment $\vec{\mu} = 2\mu_B \vec{S}$, where μ_B is the Bohr magneton and \vec{S} is the spin of a given particle. When an external magnetic field $\vec{B} = +B\hat{z}$ is applied, many particles align their spin along the external field’s direction; a few others go against. The paramagnet in turn acts as a two-state system with overall behavior governed by the strength of \vec{B} . The total number of particles in the system obeys the relation $N = N_{\uparrow} + N_{\downarrow}$, and the internal energy of the paramagnet in terms of N_{\uparrow} and N_{\downarrow} is

$$U = -N_{\uparrow}\mu B + N_{\downarrow}\mu B .$$

Starting from (72), the multiplicity of this two-state paramagnetic system in terms of U and N is

$$\Omega(U, N) = \frac{N!}{N_{\uparrow}!N_{\downarrow}!} = \frac{N!}{\left(\frac{N}{2} - \frac{U}{2\mu B}\right)! \left(\frac{N}{2} + \frac{U}{2\mu B}\right)!} ,$$

and Stirling’s approximation brings Ω to

$$\Omega(U, N) = 2^N e^{-2(U/2\mu B)^2/N} \sqrt{\frac{2}{\pi N}} .$$

Divide by 2^N to attain the probability density of this two-state paramagnet:

$$p(U) = \sqrt{\frac{2}{\pi N}} e^{-2(U/2\mu B)^2/N} \quad (107)$$

Meanwhile, the entropy $S = K \ln(\Omega)$ comes out to

$$S = K \left[\ln 2 + \frac{1}{2} \ln \left(\frac{2}{\pi N} \right) - \frac{U^2}{2N\mu^2 B^2} \right]. \quad (108)$$

You should verify that S decreases as more domains align, and also that S is maximum when $N_{\uparrow} = N_{\downarrow}$.

With the entropy on hand, use thermodynamic identity (48), namely

$$T = \left(\frac{\partial S}{\partial U} \right)_{V,N}^{-1},$$

to compute the temperature. (It's assumed the volume is fixed.) The temperature T of the paramagnet in terms of its internal energy U is

$$T = -\frac{1}{U} N\mu^2 B^2 / K. \quad (109)$$

The system must have negative internal energy in order to have a positive temperature, which makes sense from an electrodynamics point of view - a dipole that aligns parallel to an external field is bound to its equilibrium orientation.

20.2 Negative Temperature

Curiously, a two-state paramagnet can have negative temperature if conditions are right. If the system is carefully prepared so that most spins oppose the external magnetic field, the temperature is indeed negative, however not colder than absolute zero. The temperature undergoes a major discontinuity as the energy passes through $U = 0$, at which point the temperature is ambiguously either/and $T = -\infty$ or $T = \infty$. If the energy is a small positive number $U \approx \epsilon$, the temperature is around $-\infty$. If the energy is instead negative with $U \approx -\epsilon$, the temperature is approximately ∞ . This is not so far - fetched, because a paramagnet that can still manage $U = 0$ despite the external B - field should be quite hot!

Equation (109) hides some of the more interesting aspects of the two-state paramagnet at extreme nonzero U and T conditions. For instance, the unique state that all N spins up or another with all N spins down each correspond to a 'frozen' condition with $T = 0$. Our formula for T does not reach zero for finite U , which is indeed problematic. I therefore suggest a less hasty calculation of the temperature; this time starting with the multiplicity in the form of (71) and using Stirling's approximation but avoiding expansion (83). This gives a modified formula for S that still involves a few $\ln()$ terms, and definition (48) gives us the refined formula for the temperature of the two-state paramagnet,

$$T = \frac{2\mu B}{K \ln \left(\frac{N\mu B - U}{N\mu B + U} \right)}. \quad (110)$$

This formula equals zero at the boundaries $U = \pm N\mu B$, properly describing the system. (The previous $T \propto U^{-1}$ equation was good at $U = 0$, but not accurate enough at the U extremes.)

20.3 Temperature Reparameterization

Let's revisit the temperature discontinuity that occurs across $U = 0$ in the two-state paramagnet. Since the concept of temperature is rather ambiguous from the start, it does no harm to imagine a new thermodynamic parameter that does T 's job, but has a cleaner mathematical form. A convenient choice would be a parameter proportional to T^{-1} that has units of energy. By including the Boltzmann constant we may write

$$\beta = \frac{1}{KT} . \quad (111)$$

Much of statistical physics is more convenient to write down in terms of β rather than T , particularly because the factor $1/KT$ occurs so frequently. If it happens that the temperature is divergent, β is simply zero. For future reference, the T - derivative of β is

$$\frac{d\beta}{dT} = -\frac{1}{KT^2} . \quad (112)$$

21 Einstein Solid

An accurate and precise description of matter requires full - blown quantum mechanics, which is directly practical only for certain systems such as isolated atoms, harmonic oscillators, and so on. This isn't a comment on the correctness of quantum mechanics, only the usefulness of theory - when many particles are involved, quantum mechanics explodes in terms of its notation and representation.

To analyze a statistical system, physicists often create a model or a 'toy' that captures the system's main behavior but ignores the fine details. I now introduce a toy model called the *Einstein solid*, where each particle in a one-dimensional solid is represented by a distinguished harmonic oscillator, with the energy of any particle being

$$\epsilon_i = \hbar\omega (n_i + 1/2) . \quad (113)$$

The Einstein solid has its basis in quantum mechanics, but all we ever mention is the harmonic oscillator - there is no need for nucleons, electrons, spin, or momentum in the model.

In equation (113), constant ω is the angular vibration frequency of an oscillator, and is considered uniform throughout the solid. Variable n_i is the quantum excitation number of each oscillator, and may occupy any positive integer including zero. The total internal energy E is the sum of all ϵ_i , or

$$E = \sum_{i=1}^N \epsilon_i .$$

Suppose now that all quantum numbers n_i are equal to zero. This unique state corresponds to an energy minimum in the solid, where the remaining bit is called the *zero - point energy*. The zero - point energy cannot be removed from a system by any known means, unless the system is wholly destroyed. For our purposes then, it makes sense to group the zero - point energy into the same category as the rest energy $E = mc^2$, thereby ignoring it thermodynamically. In order to make the internal energy correspond to 0 when all n_i are 0, subtract off the zero - point energy in a definition for U of the Einstein solid:

$$U = \sum_{i=1}^N \epsilon_i - N \frac{\hbar\omega}{2}$$

21.1 Multiplicity of Einstein Solid

Now we determine the multiplicity $\Omega(U, N)$ for a given solid with internal energy macrostate U and number of oscillators N . To begin, visualize an Einstein solid in its ground state with all $n_i = 0$ and $U = 0$. We illustrate this with $N - 1$ slashes, each slash representing the boundary between two oscillators. Taking the example $N = 4$ we write:

$$U = 0 \quad \rightarrow \quad |||$$

If we add one energy unit to the system, any one of the oscillators therefore must raise its quantum number n_i by 1, bringing macrostate U to $\hbar\omega$. Representing the energy unit as a dot, the system is depicted by one of the four microstates:

$$\bullet||| \quad | \bullet || \quad || \bullet | \quad ||| \bullet$$

Adding a second energy unit to the solid brings the internal energy U up to $2\hbar\omega$, and there are 10 ways that 4 oscillators can hold 2 energy units:

$$\begin{array}{ccccc} \bullet\bullet||| & \bullet| \bullet || & \bullet || \bullet | & \bullet ||| \bullet & | \bullet \bullet || \\ | \bullet | \bullet | & | \bullet || \bullet & || \bullet \bullet | & || \bullet | \bullet & ||| \bullet \bullet \end{array}$$

The system is represented this way because we reframe the problem as a two-state analysis.

When a picture of a given microstate is drawn, there are $N - 1$ slashes and $U/\hbar\omega = u$ dots, making of total of $N - 1 + u$ symbols. Of the $N - 1 + u$ symbols, u of them correspond to an energy unit added to the otherwise ground - state system. Therefore, the multiplicity of the Einstein solid equals ‘ $N - 1 + u$ choose u ’, in accordance with equation (69), or

$$\Omega(u, N) = \frac{(N - 1 + u)!}{u! (N - 1)!} \quad (114)$$

21.2 Thermodynamics of Einstein Solid

Using (114), the entropy $S = K \ln(\Omega)$ is calculated with help from Stirling’s approximation (86), and simplifies to

$$S = K \left[N \ln \left(1 + \frac{u}{N} \right) + u \ln \left(1 + \frac{N}{u} \right) \right]. \quad (115)$$

The temperature of an Einstein solid is straightforward to calculate from definition (48), and boils down to

$$T = \frac{\hbar\omega/K}{\ln(1 + N\hbar\omega/U)}. \quad (116)$$

We may as well keep going and solve for the internal energy U in terms of T :

$$U = \frac{N\hbar\omega}{e^{\hbar\omega/KT} - 1} \quad (117)$$

With an explicit form for U in hand, many more calculations are possible, primarily the heat capacity:

$$C_V = \frac{dU}{dT} = \frac{N\hbar^2\omega^2 e^{\hbar\omega/KT}}{KT^2 (e^{\hbar\omega/KT} - 1)^2}. \quad (118)$$

Problem 61: To describe a three-dimensional Einstein solid, what changes must be made to equations (114) - (118)? Answer: $N \rightarrow 3N$.

21.3 Combined Einstein Solids

Consider an Einstein solid with N oscillators and total energy $U \gg \hbar\omega N$. The system is divided into two subsystems A and B such that

$$N_A = N_B = N/2 \quad U_A = U/2 + q \quad U_B = U/2 - q.$$

The temperature of the total solid of N particles must not be sensitive to the arbitrary division into subsystems A and B . Using definition (48), the temperature of subdivision A reads

$$\frac{1}{KT} = \frac{\partial}{\partial U_A} \ln \Omega(U_A, N/2) \big|_{\epsilon=0}.$$

Recalling that $S = K \ln \Omega$, it's straightforward to show, in the limit $U \gg \hbar\omega N$, that

$$KT = \frac{U_A}{N_A},$$

which is the average energy per oscillator in subdivision A . Clearly this result is the same for subdivision B , and is evidently independent of the size(s) of A and B .

Note further that the term KT is really the sum $KT/2 + KT/2$, as the equipartition function gives us one factor of $KT/2$ per quadratic degree of freedom per oscillator. (A linear oscillator has two degrees of freedom.)

Problem 62: Prove that the multiplicity of the combined system is approximately

$$\Omega_{AB} = \Omega(U/2, N/2)^2 e^{-2Nq^2/U^2},$$

where (approximately)

$$\Omega(U/2, N/2)^2 = \left(\frac{Ue}{N\hbar\omega} \right)^N.$$

Problem 63: Treat the energy U_i in each harmonic oscillator ($1 \leq i \leq N$) as a random variable. Each U_i may be replaced by $U_i \rightarrow \langle U \rangle + q_i$, where q_i is the 'excess' (positive or negative) energy of the i^{th} oscillator and $\langle U \rangle = U/N$. Calculate the standard deviation in q_i using definition (78). Answer:

$$\begin{aligned} \langle q_i \rangle &= 0 & \langle q_i^2 \rangle &= \int_{-\infty}^{\infty} q^2 A e^{-2Nq^2/U^2} dq & A &= \sqrt{\frac{2N}{\pi U^2}} \\ \langle q_i^2 \rangle &= A \frac{\sqrt{\pi}}{2} \left(\frac{U^2}{2N} \right)^{3/2} = \frac{U^2}{4N} & \sigma_{q_i} &= \sqrt{\langle q_i^2 \rangle} = \frac{U}{2\sqrt{N}} \end{aligned}$$

22 Quantum Ideal Gas

The classical ideal gas model successfully accounts for the statistical behavior of many gas systems. Our construction of ideal gas theory thus far has been a stitching together of empirical conjectures. Using microcanonical theory along with quantum mechanics, we now construct the ideal gas model in a more quantitative fashion.

22.1 Many Particles in a Box

The ideal gas is ‘defined’ as particles occupying a container, which is related to a standard example in quantum mechanics. Consider a one-dimensional ‘box’ of length L occupied by a particle of mass m . The walls at $x = 0$ and $x = L$ are impenetrable. The behavior of the particle is determined by the Schrodinger equation, namely

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = \left[\frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x, t) \right] \psi(x, t) , \quad (119)$$

where ψ is the wavefunction of the particle, and V is the potential energy of the particle, which is zero for this problem.

The particle-in-a-box solution to (119) is

$$\psi(x, t) = \sqrt{\frac{2}{L}} e^{-i\epsilon t/\hbar} \sin(kx) ,$$

where the energy ϵ and wavenumber k are related by $\epsilon = \hbar^2 k^2 / 2m$. Explicitly, these are:

$$\epsilon = \frac{n^2 \hbar^2}{8mL^2} \quad k = \frac{n\pi}{L}$$

The number n is any positive nonzero integer. (Exclude the negative n as we are dealing with standing-wave solutions.) Generalizing to three dimensions, the energy becomes

$$\epsilon = \frac{\hbar^2}{8mL^2} (n_x^2 + n_y^2 + n_z^2) . \quad (120)$$

That is, a particle trapped in a three-dimensional box of side L has a quantized energy according to (120), which is *not* a continuous function.

Generalizing the particle-in-a-box setup to accommodate N particles, first define the total energy U as the sum

$$U = \sum_{j=1}^N \epsilon_j ,$$

and meanwhile define a ‘dimensionless energy’ N^* such that

$$(N^*)^2 = \sum_{i=1}^{3N} n_i^2 = \frac{8mL^2 U}{\hbar^2} . \quad (121)$$

22.2 Volume of a Hypersphere

Consider again a single particle in a box. The energy (or momentum) of that particle may be represented by one point (n_x, n_y, n_z) in the three-dimensional n -space. Of course, n -space is not smooth and continuous, as all n take integer values only. A particle with given energy ϵ is located at some finite radius in n -space from the origin.

When we consider a gas of N particles instead of one, the number of dimensions of n -space jumps from 3 to $3N$. A gas of energy E has corresponding ‘radius’ equal to N^* , the square root of (121).

Ultimately the multiplicity Ω of the ideal gas must be calculated, which involves finding how many $3N$ -uples of nonnegative integers $(n_1, n_2, n_3, \dots, n_{3N})$ satisfy equation (121).

It follows that we must determine the volume of (the first octant of) the $3N$ -dimensional sphere of radius N^* .

Rather than using the variables N^* and $3N$, momentarily denote the number of dimensions as d and the radius of the hypersphere as R . To find the volume, begin by writing

$$V_d = \int_{\sum_{i=1}^d x_i^2 < R^2} dx_1, \dots, dx_d = \int_0^R r^{d-1} dr \int d\Omega_{d-1} = \frac{R^d}{d} \int d\Omega_{d-1},$$

where the integrals have been stated in spherical coordinates.

Next recall the following Gaussian integral identity

$$(\sqrt{\pi})^d = \int e^{x_i^2} dx_i, \dots \int e^{x_d^2} dx_d = \int_0^\infty r^{d-1} e^{-r^2} dr d\Omega_{d-1} = \frac{\Gamma(\frac{d}{2})}{2} \int d\Omega_{d-1},$$

from which we may deduce

$$V_d = \frac{R^d}{d} \pi^{d/2} \frac{2}{\Gamma(\frac{d}{2})} = \frac{R^d \pi^{d/2}}{\Gamma(\frac{d}{2} + 1)} = \frac{R^d \pi^{d/2}}{(\frac{d}{2})!}.$$

for the full sphere. To single out the octant that has only positive n_i , multiply V by a factor of $(1/2)^d$. The final result is

$$V_d = \frac{R^d \pi^{d/2}}{2^d (\frac{d}{2})!}. \quad (122)$$

For completeness, the surface area of a hypersphere relates to the volume by

$$S_{n+1} = 2\pi V_n,$$

which is proven with recursion relations. Explicitly,

$$S_{d-1} = \frac{R^d \pi^{d/2} d}{\Gamma(\frac{d}{2} + 1)}.$$

22.3 Multiplicity of Ideal Gas

The total number of microstates Ω that correspond to an ideal gas of given energy U and volume V with number of particles N may be written as the integral over all possible position and momentum states as follows:

$$\Omega = \frac{1}{N!} \frac{\int dx_1, \dots, dx_{3N} \int dp_1, \dots, dp_{3N}}{h^{3N}} \quad (123)$$

The factor of $1/N!$ has been inserted to account for the indistinguishability of ideal gas particles. If somehow we could uniquely distinguish each particle, the $1/N!$ - factor would not be present.

The position integral is the easiest to evaluate - we simply get $3N$ copies of the length L , or equivalently, N copies of the volume V of the gas. That is, $\int dx_1, \dots, dx_{3N} = V^N$. The momentum integral is less obvious; first recast the dp_i terms in terms of the integers n_i , and arrive at

$$\Omega = \frac{1}{N!} \frac{V^N}{h^{3N}} \left(\frac{\hbar\pi}{L} \right)^{3N} \int dn_1, \dots, dn_{3N},$$

where the remaining integral is the volume of the $3N$ -dimensional hypersphere of radius N^* . Substituting equations (121) and (122) into the above arrive at the anticipated result

$$\Omega(U, V, N) = \frac{V^N}{N! (3N/2)!} \left(\frac{2\pi m U}{h^2} \right)^{3N/2}. \quad (124)$$

Problem 64: Consider an ideal gas of internal energy U and volume V consisting of N particles. For large N , show that

$$\ln \Omega(2U, 2V, 2N) = \ln \Omega(U, V, N)^2$$

is a good approximation.

Problem 65: A container of volume V is filled with one mole of ideal gas. Using equation (124), compute the probability that a region with volume $V/1000$ will be completely void of particles. Answer:

$$P = \frac{\Omega(U, V \frac{999}{1000}, N)}{\Omega(U, V, N)} = \left(\frac{999}{1000} \right)^{6 \times 10^{23}} \approx 0$$

22.4 Entropy of Ideal Gas, Gibbs Paradox

Using the definition of the entropy, namely $S(U, V, N) = K \ln \Omega$, S may be determined with some help from Stirling's approximation. After the dust settles, we find

$$S = NK \left(\ln \left[\frac{V}{N} \left(\frac{4\pi m U}{3N h^2} \right)^{3/2} \right] + \frac{5}{2} \right), \quad (125)$$

a result known as the *Sackur - Tetrode* equation.

I point out that equation (125) is an extensive quantity, as the entropy should be. The pioneers of this subject weren't able to get the correct answer right away. Indeed, it's not highly obvious that the $1/N!$ - term is needed in equation (123). If that term were not present, the above result would have been

$$\tilde{S} = NK \left(\ln \left[V \left(\frac{4\pi m U}{3N h^2} \right)^{3/2} \right] + \frac{3}{2} \right),$$

which is wrong! This subtle error is known as the Gibbs paradox. The lesson: Don't forget to include a factor of $1/N!$ somewhere when dealing with indistinguishable particles.

22.5 Recovering Classical Results

With the entropy in hand, *all* of the ideal gas properties derived in previous sections can be recovered. The identities $dU = TdS - PdV + \mu dN$ and $H = U + PV$ are all we need to get the good ones. I encourage you to retrace the following calculations:

$$\left(\frac{\partial S}{\partial V} \right)_{T, N} = \frac{P}{T} \quad \rightarrow \quad PV = NKT$$

$$\left(\frac{\partial S}{\partial U} \right)_{V, N} = \frac{1}{T} \quad \rightarrow \quad U = \frac{3}{2} NKT$$

$$\begin{aligned} \left(\frac{\partial U}{\partial T}\right)_{V,N} = C_V & \quad \rightarrow \quad C_V = \frac{3}{2}NK \\ \left(\frac{\partial H}{\partial T}\right)_{P,N} = C_P & \quad \rightarrow \quad C_P = \frac{5}{2}NK \end{aligned}$$

22.6 Chemical Potential of Ideal Gas

The chemical potential of the ideal gas is given by the identity

$$\left(\frac{\partial S}{\partial N}\right)_{U,V} = -\frac{\mu}{T},$$

which has not yet been explicitly determined. Inserting our form for S into the above yields

$$\mu = -KT \ln \left[\left(\frac{V}{N}\right) \left(\frac{4\pi m U}{3N h^2}\right)^{(3/2)} \right].$$

The instance of U may be replaced by $3NKT/2$. Simplifying the result gives

$$\mu = -KT \ln \left(\frac{V}{N \lambda_0^3} \right), \quad (126)$$

where the ‘length’ λ_0 is known as the *thermal de Broglie wavelength*. Specifically,

$$\lambda_0 = \frac{h}{\sqrt{2\pi m K T}}. \quad (127)$$

As a function of pressure, the chemical potential reads

$$\mu = \mu_0 + KT \ln \left(\frac{P}{P_0} \right), \quad (128)$$

where $P_0 = KT/\lambda_0^3$ and the term μ_0 has been added on to be accurate for dilute solutions. The ideal has $\mu_0 = 0$.

22.7 Fugacity

The expression (128) for the chemical potential of the ideal gas can be generalized to accomodate non - ideal gases and dilute solutions at constant temperature T . The ratio P/P_0 is replaced by the ratio f/f_0 such that

$$\mu = \mu_0 + KT \ln \left(\frac{f}{f_0} \right),$$

where f is known as the *fugacity* of the substance. Solving for f , we obtain

$$f = f_0 e^{\beta(\mu - \mu_0)}.$$

Fugacity may be interpreted as the pressure required for a system to become diffuse and behave like ideal gas (in which case $\mu_0 = 0$). Dividing through by the pressure to attain the *fugacity coefficient*:

$$z = e^{\beta(\mu - \mu_0)} \quad (129)$$

Note that the notion of fugacity extends to all of statistical physics - this construction is not limited to the ideal gas.

22.8 Generalization and Other Dimensions

So far, the ideal gas model has been written to only account for 3-dimensional gases, which is of course powerful. It's also useful to consider spaces that are not $3D$. The notion of 'volume' must now be written as L^α , where α is the number of spatial dimensions available to the gas. The multiplicity readily takes the form

$$\Omega(U, L^\alpha, N) = U^{\alpha N/2} L^{\alpha N} f(N), \quad (130)$$

where $f(N)$ is some function of the number of particles N . The entropy is still given by $S = K \ln(\Omega)$. Taking various derivatives of S give a generalized ideal gas law, plus an internal energy relation that is consistent with the equipartition theorem:

$$\begin{aligned} \left(\frac{\partial S}{\partial L^\alpha} \right)_{T, N} &= \frac{P}{T} & \rightarrow & & PL^\alpha &= NKT \\ \left(\frac{\partial S}{\partial U} \right)_{L^\alpha, N} &= \frac{1}{T} & \rightarrow & & U &= \frac{\alpha}{2} NKT \end{aligned}$$

Problem 66: Consider a 3-dimensional ideal gas consisting of N_3 particles occupying a volume L^3 that can exchange energy (but not matter) with a 2-dimensional ideal gas consisting of N_2 particles on a surface of area L^2 . Let the total energy E of the system equal the sum $E_3 + E_2$, the energies of the individual subsystems. Find (i) the total number of microstates of the system as a function of E_2 , and find (ii) the equilibrium value of E_2 and $E_3 = E - E_2$. Answer:

$$\begin{aligned} \Omega &= \left(\sqrt{E - E_2} \right)^{3N_3} \left(\sqrt{E_2} \right)^{2N_2} f(L, N_2, N_3) \\ \left(\frac{\partial S}{\partial E_2} \right)_{L^2, N} &= 0 & \rightarrow & & \bar{E}_2 &= \frac{N_2 E}{N_2 + 3N_3/2}, & \bar{E}_3 &= \frac{3N_3 E/2}{N_2 + 3N_3/2} \end{aligned}$$

23 Canonical Ensemble

So far we have developed the notion of the microcanonical ensemble - that is, statistical systems with fixed energy, fixed volume, and fixed number of particles. Now we shift the argument and introduce the *canonical ensemble*. In this regime, a thermodynamic system of fixed volume and fixed number of particles is in contact with a thermal reservoir of fixed temperature T , where now the system's internal energy U is not constant.

23.1 Minimization of Helmholtz Free Energy

Consider a system with fixed V and N in contact with a large reservoir of temperature T . Denote the system's multiplicity Ω_1 , and denote the reservoir's multiplicity as Ω_2 . The total energy is $U_1 + U_2 = U_T$, and the combined multiplicity is

$$\Omega_T = \Omega_1(U_1) \Omega_2(U_T - U_1) .$$

Take the natural logarithm of both sides and evaluate the result with $U_1 \approx 0$. Taylor-expanding the $\ln \Omega_2$ term, we see

$$\ln \Omega_T = \ln \Omega_1(U_1) - U_1 \left. \frac{\partial \ln \Omega_2}{\partial U_2} \right|_{U_2 \approx U_T}, \quad (131)$$

which if you remember equation (48) and $S = K \ln \Omega$, can be written as:

$$\frac{S_T}{K} = \frac{S_1}{K} - \frac{U_1}{KT} \quad \rightarrow \quad -TS_T = U_1 - TS_1$$

The quantity TS_T is really $KT \ln(\Omega_T)$, and $U_1 - TS_1$ is the Helmholtz free energy A of the system.

Meanwhile, we know that the combined system will tend to occupy the state that maximizes Ω_T , or in other words, minimize the quantity $-KT \ln(\Omega_T)$. That is, the canonical system with fixed V and N tends to minimize the Helmholtz free energy

$$A = -KT \ln(\Omega_T) . \quad (132)$$

This observation is consistent with the previously - established equation (63), namely

$$dS_{uni} = -\frac{1}{T} (dA)_{T,V,N} .$$

(The symbol Ω_T shall be subsequently replaced.)

23.2 Legendre Transform

In the canonical ensemble, the Helmholtz free energy $A(T, V, N)$ is the quantity that is minimized at equilibrium. Here I claim that we cannot frame the problem in terms of the internal energy U . In fact, it's not proper to state $U = U(T, V, N)$ because the temperature T is defined as a derivative of U by equation (48), namely

$$T = \left(\frac{\partial S}{\partial U} \right)_{V,N}^{-1} .$$

It is appropriate however to work with the internal energy in terms of the entropy instead as $U = U(S, V, N)$. The problem is, the temperature definition (48) isn't built into $U(S, V, N)$ - the system should be described in terms of both U and it's derivative with respect to S . We would like to build both ideas into one object.

To proceed, select a given point P on a $U(S)$ curve, and determine the slope $\partial U / \partial S = T$ at that point. Draw a straight line through point P having slope T , and extend the line until it hits the U - axis. The U - intercept is precisely the Helmholtz free energy A . That is, we have drawn the line

$$U(S, V, N) = A(T, V, N) + S \left(\frac{\partial U}{\partial S} \right)_{V,N} ,$$

which is equivalent to the usual formula $A = U - TS$.

The action of subtracting TS from U is known as the *Legendre transform*. Notice that A contains the information that U alone does not.

23.3 Boltzmann Factor and Partition Function

Recall that in any statistical system, the probability, now denoted Pr , of observing macrostate k is equal to the multiplicity of that macrostate divided by the total multiplicity. The microstates in a canonical system are not equally accessible, but are dependent on the internal energy of a given microstate at temperature T .

Returning to equation (131), note that the $\ln \Omega_1$ term is much smaller than the others and may be ignored. Taking the exponential of each side (reversing the logarithm), we evidently have $Pr \propto e^{-\beta U_1}$. It follows that a canonical system's probability of occupying energy state k is:

$$Pr(k) = \frac{1}{Z} e^{-\beta U(k)} \quad (133)$$

The term $\exp(-\beta U)$ is called the *Boltzmann factor*. Note that β is equal to $1/KT$, and U is the internal energy of the system in state k . Since the sum of all probabilities must equal unity, the normalization statement

$$\sum_k Pr(k) = 1$$

tells us the sum of all Boltzmann factors is

$$Z(V, T) = \sum_{\text{states } k} e^{-\beta U_k} , \quad (134)$$

which is called the *partition function*.

Note the sum in (134) is over the discrete set of microstates k . It's often useful to cast the sum over discrete energy levels j , and the partition function becomes

$$Z = \sum_{\text{energies } j} g(U_j) e^{-\beta U_j} , \quad (135)$$

where $g(U_j)$ is called the *degeneracy* of energy U_j . The degeneracy counts how many distinct microstates that correspond to the macrostate U_j . Of course, a system's energy may be continuous instead of discrete, in which case the partition function becomes an integral:

$$Z = \int_{\text{energies}} g(U') e^{-\beta U'} dU' \quad (136)$$

Often, the integration domain is neither state numbers k nor energies j . Instead, a general dynamic system will have a partition function of the form

$$Z = \frac{\int \int e^{-\beta H(x,p)} d^m x d^m p}{h^m} , \quad (137)$$

where m is the number of dimensions in position - momentum phase space, H is the Hamiltonian, and h is Planck's constant.

23.4 Observables

The average value of any observable macrostate O may be defined as a generalization of the normalization condition such that

$$\langle O \rangle = \sum_k \hat{O} Pr(k) , \quad (138)$$

where $\langle O \rangle$ is the average value of O as the operator \hat{O} is the weight factor inside the summation.

The internal energy U (omitting the $\langle \rangle$ symbols) of a canonical system is

$$U = \sum_k U_k \frac{1}{Z} e^{-\beta U_k} ,$$

which after some simplification becomes remarkably simple:

$$U = -\frac{\partial}{\partial \beta} \ln Z \quad (139)$$

Similarly, the entropy S may be calculated with the operator $\hat{S} = -K \ln(Pr(k))$, where K is the Boltzmann constant. Proceed by writing

$$S = -K \sum_k \ln(Pr(k)) \frac{1}{Z} e^{-\beta U_k},$$

which reduces to

$$S = \frac{U}{T} + K \ln Z. \quad (140)$$

By thermodynamic identity $A = U - TS$, the Helmholtz free energy is verified to be

$$A = -KT \ln Z.$$

Note/recall that a canonical thermodynamic system with fixed temperature, fixed volume, and fixed number of particles seek equilibrium by minimizing the Helmholtz free energy A .

Equations (139) - (132) are indeed strange, but are completely consistent with the previously - established thermodynamic identities. For instance, the pressure in a canonical system is given by

$$P = -\left(\frac{\partial A}{\partial V}\right)_{T,N}, \quad (141)$$

where substituting (132) into the above we correctly verify the pressure operator as follows:

$$P = \sum_k \left(-\frac{\partial U_k}{\partial V}\right) \frac{1}{Z} e^{-\beta U_k} \quad (142)$$

23.5 Fluctuations

A canonical system tends to minimize the Helmholtz free energy by adjusting state variables toward their average values. It's interesting to examine the system's behavior *near* (instead of *at*) equilibrium. Suppose the internal energy of a canonical system is given by U , and the average of U is denoted $\langle U \rangle$. The probability of that system having internal energy U is equal to

$$Pr(U) = g(U) e^{-\beta U}.$$

In general, the function $Pr(U)$ is sharply - peaked around the average energy $\langle U \rangle$, and is diminished away from $\langle U \rangle$. We therefore approximate the probability as a Gaussian form:

$$Pr(U) = A e^{-(U - \langle U \rangle)^2 / 2\sigma^2}$$

Equating each version of $Pr(U)$ and taking the natural logarithm of both sides, we get

$$\ln(g(U)) - \beta U = \ln A - \frac{(U - \langle U \rangle)^2}{2\sigma^2}. \quad (143)$$

Take a single U - derivative of (143) and evaluate the result at $U = \langle U \rangle$, yielding the identity

$$\beta = \frac{\partial}{\partial U} \ln g(U) .$$

Next, take the second derivative of (143) to attain information on σ . Recognize that $(\partial U / \partial T)$ is precisely equal to the heat capacity at constant volume, C_V , namely because V and N are constant in the canonical ensemble. Conclude that

$$\sigma = \sqrt{KT^2 C_V} \quad (144)$$

For example, an ideal gas at about $300^\circ C$ has $KT \approx (1/40) eV$ (electron - volts). The fluctuation width σ evaluates to about $2 \times 10^{10} eV$, making $\sigma/U \approx 10^{-12}$.

23.6 Non-Interacting Particles

Consider a system of two particles A and B . The state S of the system may be represented as $S = [s_A, s_B]$, and the total energy E of the system is $E = E(s_A) + E(s_B)$.

Distinguishable Particles

If the particles A and B are distinguishable, it follows that $[s_A, s_B] \neq [s_B, s_A]$. That is, the system will change if the two particles are swapped. The partition function is

$$Z = \sum_{s_A} \sum_{s_B} e^{-\beta[E(s_A) + E(s_B)]} = Z_A Z_B ,$$

which means we simply multiply individual partition functions of each particle. This easily generalizes to a system of N distinguishable particles:

$$Z = Z_1 Z_2 Z_3 \dots Z_N \quad (145)$$

Indistinguishable Particles

If the particles A and B are instead indistinguishable, it follows that $[s_A, s_B] = [s_B, s_A]$. which means the system doesn't notice if two particles are swapped. The partition function must be divided by two:

$$Z = \frac{1}{2} Z_A Z_B .$$

The argument straightforwardly applies to a system of N indistinguishable particles:

$$Z = \frac{1}{N!} Z_1 Z_2 Z_3 \dots Z_N \quad (146)$$

23.7 Equipartition Theorem Derived

The equipartition theorem has so far ensured that the internal energy per particle gains of factor of $KT/2$ for each quadratic degree of freedom. There's more to the story: *the internal energy per particle gains a factor of KT/m for each degree of freedom of power m .*

Consider a system of N particles that are subject to the Hamiltonian

$$H = H_0(p) + bx^m ,$$

where b is a constant and m is the power - law for the potential energy term. (For example, take $m = 2$ for a harmonic oscillator.) The kinetic term $H_0(p)$ is left unspecified without loss of generality. The partition function of a single particle is

$$Z_1 = \frac{\int dx e^{-\beta b x^m} \int dp e^{-\beta H_0(p)}}{h} = f(T, V) \int dx e^{-\beta b x^m} .$$

By standard substitution techniques, the partition function becomes

$$Z_1 = f(T, V) \frac{1}{m} (\beta b)^{-1/m} \int_0^\infty x^{1/m-1} e^{-x} dx ,$$

where the integral above is equal to $\Gamma(1/m)$. For the system of N particles, the partition function Z_1 is raised to the N - th power, and we may or may not require a factor of $1/N!$ to keep track of particle (in)distinguishability, which is inconsequential for this derivation - absorb all N - business into the function f . It follows that:

$$Z = f(T, V)^N \left(\frac{1}{m}\right)^N (\beta b)^{-N/m} \Gamma\left(\frac{1}{m}\right)^N$$

Now, the equipartition theorem concerns the internal energy, which is defined as the minus β - derivative of the natural logarithm of the partition function, in other words $U = -\partial \ln Z / \partial \beta$. Without specifying $H_0(p)$, the exact internal energy isn't available unless f is known - we are only interested in the energy per particle due to the b - term. Ultimately we find

$$\frac{\langle U \rangle_m}{N} = \frac{1}{m} K T , \quad (147)$$

which quantifies (and justifies) the equipartition theorem.

24 Canonical Ideal Gas

24.1 Partition Function and Thermodynamic Potentials

An ideal gas in a container of volume V and temperature T has N molecules moving around, each carrying some momentum p_i and kinetic energy $p_i^2/2m$, where m is the mass of each identical particle. The partition function of the ideal gas is a Boltzmann - factor modification of the multiplicity, equation (123). Begin by writing:

$$Z = \frac{1}{N!} (Z_1)^N \quad (148)$$

Note that

$$Z_1 = \frac{\int d^3x \int d^3p e^{-\beta p^2/2m}}{h^3} ,$$

which is the partition function of a single particle. The factor $1/N!$ accounts for the indistinguishability of the gas particles. Evaluating the spatial integral in (148) gives

$$Z = \frac{1}{N!} \frac{V^N}{h^{3N}} \left(4\pi \int_0^\infty p^2 e^{-\beta p^2/2m} dp \right)^N ,$$

where the momentum integral has been expressed in spherical coordinates and the angular integrals have been evaluated. Using the identity

$$\int_0^\infty x^2 e^{-ax^2} dx = \frac{\sqrt{\pi}}{4a^{3/2}} ,$$

the partition function becomes

$$Z = \frac{1}{N!} \left[\frac{V}{h^3} (2\pi m K T)^{3/2} \right]^N = \frac{1}{N!} \left(\frac{V}{\lambda_0^3} \right)^N. \quad (149)$$

Recall that λ_0 is known as the thermal de Broglie wavelength.

The internal energy U is straightforwardly calculated from the identity (139), which readily delivers $U = (3/2)NKT$. The entropy can be calculated from the identity (140). Begin by writing

$$S = -K \sum_k \ln \left[\frac{1}{Z} e^{-\beta p_k^2/2m} \right] \frac{1}{Z} e^{-\beta p_k^2/2m},$$

which simplifies to

$$S = \frac{1}{T} \sum_k \frac{p_k^2}{2m} \frac{1}{Z} e^{-\beta p_k^2/2m} + K \ln Z.$$

The first term is equal to the average internal energy U divided by the temperature, namely $(3/2)NK$ after writing U in terms of T . The second term gives *exactly* what else is needed to recover our previous answer for the entropy, the Sackur - Tetrode equation (125).

Finally, the Helmholtz free energy is easily determined by (132), and for the ideal gas, comes out to

$$A = -NKT \left[1 + \ln \left(\frac{V}{N\lambda_0^3} \right) \right]. \quad (150)$$

Problem 67: Take the appropriate derivative of (150) to attain the ideal gas law.

Problem 68: Consider an ‘ideal’ gas made of N particles that move at relativistic speeds in a three-dimensional volume V . The kinetic energy per particle is no longer given by $E = p^2/2m$, but instead $E = pc$. Show that the partition function is:

$$Z = \frac{1}{N!} \left[8\pi V \left(\frac{KT}{hc} \right)^3 \right]^N$$

Problem 69: Consider an ‘ideal’ gas made of $3N$ particles that move at relativistic speeds in a one-dimensional channel L . The kinetic energy per particle is no longer given by $E = p^2/2m$, but instead $E = pc$. Show that the partition function is:

$$Z = \frac{1}{(3N!)} \left[2L \left(\frac{KT}{hc} \right) \right]^{3N}$$

24.2 Maxwell Speed Distribution Derived

We have previously used the Maxwell speed distribution (19) to gain many insights into the ideal gas:

$$D(v) dv = \left(\frac{m}{2\pi KT} \right)^{3/2} 4\pi v^2 e^{-mv^2/2KT} dv$$

To derive the form of $D(v)$, begin with the notion that a given gas particle must be found *somewhere* in three-dimensional velocity - space according to the normalization of the probability density:

$$1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A e^{-\beta mv^2/2} dv_x dv_y dv_z$$

The probability distribution function inside the integral is simply the Boltzmann factor $\exp(-\beta mv^2/2)$ times a normalization factor A . Converting to spherical coordinates and performing the angular integrals, the above becomes

$$1 = A \int_0^\infty 4\pi v^2 e^{-\beta mv^2/2} dv ,$$

where after solving for A , we pick out $D(v)$ to attain equation (19).

24.3 Van der Waals Equation Derived

The van der Waals equation accounts for ‘ideal gas’ particles that have finite size and are allowed to interact through the attractive Leonnard Jones potential. In an ad hoc fashion we assembled equation (26), namely

$$\left(P + a \frac{N^2}{V^2} \right) (V - Nb) = NKT ,$$

where the constants a and b are specific to the gas or fluid being considered.

Here we use Boltzmann factors and the partition function to derive the van der Waals equation properly. The available volume of the gas is again $V - Nb$. The Leonnard Jones potential imparts an extra average energy $-aN/V$ to each particle. The partition function of the ideal gas therefore generalizes to

$$Z = \frac{1}{N!} \left[\frac{(V - Nb)}{N} \left(\frac{2\pi mKT}{h^2} \right)^{3/2} e^{\beta aN/V} \right]^N . \quad (151)$$

Problem 70: Use thermodynamic identity (139), namely $U = -\partial \ln Z / \partial \beta$, to show that the internal energy of the van der Waals fluid is

$$U = \frac{3}{2} NKT - \frac{aN^2}{V} . \quad (152)$$

Problem 71: Calculate the Helmholtz free energy A of the van der Waals fluid. Answer:

$$A = -NKT \left(1 + \ln \left[\frac{V - Nb}{N} \left(\frac{2\pi mKT}{h^2} \right)^{3/2} \right] \right) - \frac{aN^2}{V} \quad (153)$$

Problem 72: Using identity (141), namely $P = -(\partial A / \partial V)_{T,N}$, attain the van der Waals equation (26).

Problem 73: Use the identity $A = U - TS$ to calculate the entropy of the van der Waals fluid. Answer:

$$S = NK \left(\ln \left[\frac{V - Nb}{N} \left(\frac{4\pi m}{3Nh^2} (U + aN^2/V) \right)^{3/2} \right] + \frac{5}{2} \right) \quad (154)$$

25 Simple Harmonic Oscillator System

The simple harmonic oscillator (abbreviated SHO) is undoubtedly the most frequently encountered physical regime across all theoretical and practical disciplines. Nearly all classical systems exhibit simple harmonic oscillatory behavior when perturbed near a local energy minimum. Analysis of the SHO may be done in both the classical and the quantum viewpoints of matter. Although these approaches are different at face value, they converge in the high - temperature regime.

25.1 Classical Simple Harmonic Oscillator

Consider a collection of N simple harmonic oscillators. Each oscillator is centered on a nonmoving location - there is no net displacement over time. From this, it should follow that the system's volume is constant, and consequentially the pressure is identically zero.

If the oscillators are represented by particles in an attractive potential, we may say the particles *are* distinguishable. Despite the particles being themselves identical, we may easily 'keep track' of the location of each. Suppose each oscillator has position x_i and momentum p_i , all in one dimension, such that the Hamiltonian of the system of oscillators is

$$H = \sum_{j=1}^N \left(\frac{p_j^2}{2m} + \frac{m\omega^2 x_j^2}{2} \right), \quad (155)$$

where m is the mass of each oscillator, and ω is the frequency of each oscillator, presumed constant throughout the system.

The partition function of the system is of the form $Z = (Z_1)^N$, where Z_1 is the partition function of a single particle. Since one particle has both kinetic and spatially - dependent potential energy, we write Z_1 as

$$Z_1 = \frac{1}{h} \int_{-\infty}^{\infty} e^{-\beta m \omega^2 x^2 / 2} dx \int_{-\infty}^{\infty} e^{-\beta p^2 / 2m} dp .$$

Evaluation of Z_1 is a straightforward exercise in Gaussian integrals, and it follows that the system partition function becomes:

$$Z = \left(\frac{1}{\beta \hbar \omega} \right)^N = (\beta \hbar \omega)^{-N} \quad (156)$$

We may immediately evaluate all of the relevant details of the harmonic oscillator. Following we have the internal energy, Helmholtz free energy, enthalpy, entropy, Gibbs free energy, chemical potential, and the grand free energy:

$$U = -\frac{\partial \ln Z}{\partial \beta} = NKT \quad (157)$$

$$A = -KT \ln Z = NKT \ln (\beta \hbar \omega) \quad (158)$$

$$H = U + PV = NKT \quad (159)$$

$$S = -\left(\frac{\partial A}{\partial T} \right)_{V,N} = \frac{U - A}{T} = NK (1 - \ln (\beta \hbar \omega)) \quad (160)$$

$$G = H - TS = NKT \ln (\beta \hbar \omega) \quad (161)$$

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{T,V} = \frac{G}{N} = KT \ln (\beta \hbar \omega) \quad (162)$$

$$\Phi = -PV = U - TS - \mu N = 0 \quad (163)$$

Problem 74: Verify equations (157) - (163).

25.2 Quantum Simple Harmonic Oscillator

The quantum - mechanical view of the simple harmonic oscillator begins with the same Hamiltonian (155) as the classical case. Evolution is determined not by Newton's laws, but instead by the Schrodinger equation (119), namely $i\hbar\partial_t\psi(x,t) = H\psi(x,t)$. It follows from standard quantum mechanics that the quantum SHO energy levels are quantized as

$$\epsilon = \hbar\omega (n + 1/2) ,$$

which was introduced as equation (113) to develop the Einstein solid. The number n is any positive integer including zero.

According to canonical ensemble theory, the partition function for a single oscillator is the sum of all possible Boltzmann factors $\exp(-\beta\epsilon_n)$:

$$Z_1 = \sum_{n=0}^{\infty} e^{-\beta\epsilon_n} = e^{-\beta\hbar\omega/2} \sum_{n=0}^{\infty} \left(e^{-\beta\hbar\omega}\right)^n$$

The summation term in Z_1 is a geometric series, and obeys $\sum_{j=0}^{\infty} x^j = 1/(1-x)$. It follows that the single-particle partition function is equal to

$$Z_1 = \frac{1}{e^{\beta\hbar\omega/2} - e^{-\beta\hbar\omega/2}} = [2 \sinh(\beta\hbar\omega/2)]^{-1} .$$

Consider a system of N non - interacting oscillators. Each particle is 'tied down' to its energy well, so the average displacement of any particle is identically zero. It follows that (i) the particles are distinguishable from each other, and (ii) the pressure of the system is identically zero. The partition function for the whole system is

$$Z = Z_1^N = [2 \sinh(\beta\hbar\omega/2)]^{-N} , \quad (164)$$

which has no dependence on the system's volume, reinforcing statement (ii) above.

As done for the classical SHO, we can derive all of the important thermodynamics by cranking the partition function through the identities and derivatives, etc. Following we have the internal energy, Helmholtz free energy, enthalpy, entropy, Gibbs free energy, chemical potential, and the grand free energy:

$$U = -\frac{\partial \ln Z}{\partial \beta} = N \left(\frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \right) = N\hbar\omega \left(\frac{1}{2} + \langle n \rangle \right) \quad (165)$$

$$A = -KT \ln Z = N \left[\frac{\hbar\omega}{2} + KT \ln \left(1 - e^{-\beta\hbar\omega} \right) \right] \quad (166)$$

$$H = U + PV = N\hbar\omega \left(\frac{1}{2} + \langle n \rangle \right) \quad (167)$$

$$S = \frac{U - A}{T} = N \left[\frac{\hbar\omega}{T} \langle n \rangle - K \ln \left(1 - e^{-\beta\hbar\omega} \right) \right] \quad (168)$$

$$G = H - TS = N \left[\frac{\hbar\omega}{2} + KT \ln \left(1 - e^{-\beta\hbar\omega} \right) \right] \quad (169)$$

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{T,V} = \frac{G}{N} = \left[\frac{\hbar\omega}{2} + KT \ln \left(1 - e^{-\beta\hbar\omega} \right) \right] \quad (170)$$

$$\Phi = -PV = U - TS - \mu N = 0 \quad (171)$$

In the above, the symbol $\langle n \rangle$ has replaced the quantity $1/(\exp(\beta\hbar\omega) - 1)$ for brevity. When we discuss Bose - Einstein statistics, this will be known as the ‘occupation number’.

Problem 75: In the high - temperature limit, show that $\hbar\omega \langle n \rangle$ is equal to $\beta^{-1} = KT$.

Problem 76: In the high - temperature limit, show that equations (165) - (171) reduce to their classical counterparts (157) - (163).

26 Two-State Systems

26.1 Quantum Two-State System

A quantum system in general can have any mixture of continuous and/or quantized energy levels. To make analysis reasonable, we argue that any set of closely - spaced energy levels in the system, so long as they are weakly connected to the rest of the spectrum, may be analyzed as an isolated finite-dimensional system. With this we analyze angular momentum, spin, neutrino flavor, and numerous tunneling problems in quantum mechanics. Also, any pair of energy levels that are mutually affected by near - resonant radiation may be regarded as a two-state system.

Consider a two-state system with discrete energy levels E_1 and E_2 . The state vector at time t is

$$|\psi(t)\rangle = \alpha(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta(t) \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (172)$$

where the column vectors in the above are alternatively written $|e_1\rangle$ and $|e_2\rangle$, respectively. Vector $|\psi\rangle$ is normalized such that $\alpha^2 + \beta^2 = 1$. If the Hamiltonian operator \hat{H} of this system is diagonal, then $|e_1\rangle$ and $|e_2\rangle$ are automatically eigenvectors for the system, whose eigenvalues coincide with E_1 and E_2 .

In general, the Hamiltonian of a two-state system is not diagonal in the basis $(|e_1\rangle, |e_2\rangle)$, but contains mixing terms as

$$\hat{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \begin{bmatrix} \epsilon & -\Delta e^{i\phi} \\ -\Delta e^{-i\phi} & -\epsilon \end{bmatrix}. \quad (173)$$

The energy ϵ is measured from (and thus given by) the average $(E_1 + E_2)/2$. The Δ - term is called a ‘mixing term’, and is responsible for transitions (due to radiation, for instance) between states 1 and 2. (Assume nothing about the ratio ϵ/Δ .)

Time evolution is determined by the time dependent Schrodinger equation (119), $i\hbar\partial_t |\psi\rangle = H |\psi\rangle$. To proceed we strip the time - dependence from $|\psi(t)\rangle$ as

$$|\psi(t)\rangle = e^{-iUt/\hbar} |\psi\rangle,$$

allowing the Schrodinger equation to reduce to

$$\begin{bmatrix} \epsilon - U & -\Delta e^{i\phi} \\ -\Delta e^{-i\phi} & -\epsilon - U \end{bmatrix} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} = 0.$$

The eigenvalues U are found from the determinant condition

$$U_{\pm} = \pm \sqrt{\epsilon^2 + \Delta^2}.$$

The normalized eigenvectors may therefore be written as

$$|\pm\rangle = \frac{1}{\sqrt{2U_{\pm}(U_{\pm} - \epsilon)}} \left[(\epsilon - U_{\pm}) e^{-i\phi} \right],$$

and obey

$$\langle \pm | \mp \rangle = 0 \quad \langle \pm | \pm \rangle = 1.$$

The eigenvectors $|\pm\rangle$ are more conveniently expressed as

$$|+\rangle = \begin{bmatrix} \cos \theta e^{i\phi} \\ -\sin \theta \end{bmatrix} \quad |-\rangle = \begin{bmatrix} \sin \theta e^{i\phi} \\ \cos \theta \end{bmatrix},$$

where the parameter θ relates to the energies by

$$\cos \theta = \sqrt{\frac{1}{2} + \frac{\epsilon}{2U}} \quad \sin \theta = \sqrt{\frac{1}{2} - \frac{\epsilon}{2U}}$$

In the basis $(|+\rangle, |-\rangle)$, the two-state Hamiltonian may be expressed as

$$\hat{H} = U_+ (|+\rangle \langle +|) - U_- (|-\rangle \langle -|).$$

Time evolution of the state vector is easy to write down. Supposing the state of the system at time $t = 0$ is equal to $|\psi_0\rangle$, then for all times t , we have:

$$|\psi(x, t)\rangle = \langle + | \psi_0 \rangle e^{-iU_+ t/\hbar} |+\rangle + \langle - | \psi_0 \rangle e^{-iU_- t/\hbar} |-\rangle \quad (174)$$

In cases where $\epsilon \gg \Delta$, the system tends to get ‘stuck’ in either energy state U_+ or U_- , in the sense that $P(t)$ oscillates about U_{\pm} with amplitude going like Δ^2/ϵ^2 . In the other extreme with $\epsilon = 0$, the probability of observing either energy state, say U_+ , oscillates gradually in time between 1 and 0 with a period proportional to π/Δ .

26.2 Rabi Oscillations

Consider a two - level system with the following time - dependent Hamiltonian

$$\hat{H} = \begin{bmatrix} \epsilon & \delta e^{-i\omega t} \\ \delta e^{i\omega t} & 0 \end{bmatrix}.$$

Using the basis

$$|\psi(t)\rangle = \begin{bmatrix} \alpha(t) e^{-i\omega t} \\ \beta(t) \end{bmatrix},$$

it’s straightforwardly shown that the Schrodinger equation

$$i\hbar \partial_t |\psi(t)\rangle = \hat{H} |\psi(t)\rangle$$

can be written in terms of a time - independent matrix as

$$i\hbar \partial_t \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix} = \begin{bmatrix} \epsilon - \hbar\omega & \delta \\ \delta & 0 \end{bmatrix} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix}.$$

The corresponding energy eigenvalues are

$$\lambda_{\pm} = \frac{1}{2} (\epsilon - \hbar\omega) \pm \frac{1}{2} \sqrt{(\epsilon - \hbar\omega)^2 + 4\delta^2} = \Omega \pm \sqrt{\Omega^2 + \delta^2} = \Omega \pm \kappa,$$

with normalized eigenvectors

$$|\psi_{\pm}\rangle = \frac{1}{\sqrt{\delta^2 + (\Omega \pm \kappa)^2}} \begin{bmatrix} \Omega \pm \kappa \\ \delta \end{bmatrix}.$$

Suppose the two-state system begins in the pure state $|\psi_0\rangle = |e_1\rangle$. The probability amplitude for finding the system in state $|e_2\rangle$ for all times t is given by

$$\langle e_2|\psi(t)\rangle = \langle +|e_1\rangle e^{-i\lambda_+t/\hbar} \langle e_2|+\rangle + \langle -|e_1\rangle e^{-i\lambda_-t/\hbar} \langle e_2|-\rangle,$$

and comes out to

$$A = \langle e_2|\psi(t)\rangle = \frac{-i\delta e^{-i\Omega t/\hbar} \sin(\kappa t/\hbar)}{\kappa}.$$

The square of the norm of the probability amplitude gives the probability itself:

$$P = |\langle e_2|\psi(t)\rangle|^2 = \frac{\delta^2}{\kappa^2} \sin^2(\kappa t/\hbar) = \frac{\delta^2}{\Omega^2 + \delta^2} \sin^2\left(\frac{t}{\hbar} \sqrt{\Omega^2 + \delta^2}\right)$$

The special case of *resonance* occurs when $\hbar\omega = \epsilon$. In this regime we have $\Omega = 0$ and $\kappa = \delta$, in which case the eigenvectors reduce to

$$|\psi_{\pm res}\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} \pm 1 \\ 1 \end{bmatrix}.$$

More interesting is the near - resonance case with $\hbar\omega \approx \epsilon$. The angular oscillation frequency becomes

$$\kappa/\hbar = \frac{\delta}{\hbar} \sqrt{1 + \Omega^2/\delta^2} \approx \delta/\hbar + \Omega^2/2\delta\hbar = \delta/\hbar + \frac{(\epsilon - \hbar\omega)^2}{8\hbar\delta}.$$

26.3 Statistical Two-State System

A pure quantum analysis of the two-state system is indeed precise, but is limited in its practicality for describing many-particle systems, and more importantly, has no obvious way of incorporating the idea of temperature into the predictions. We'll see here that a statistical analysis is simpler to execute *and* accounts for temperature.

The most basic statistical two-state system is a canonical ensemble of N noninteracting distinguishable particles, each allowed to occupy one of two distinct energy levels $U = \pm\epsilon$. The temperature is fixed at T (recall $\beta = 1/KT$). For one particle ($N=1$), two-state partition function is

$$Z_1 = \sum_{j=1,2} e^{-\beta U_j} = 2 \cosh(\beta\epsilon). \quad (175)$$

Adding one distinguishable particle ($N = 2$), the degeneracy of each energy state gains a complication. The maximal energy state $U = 2\epsilon$ is unique, and corresponds to both particles having $U_j = \epsilon$. The minimal energy state $U = -2\epsilon$ is unique for a similar reason.

The middle - energy state $U = 0$ is not unique, because there are $g(U = 0) = 2$ ways that it can occur: each microstate $(+\epsilon, -\epsilon)$ and $(-\epsilon, +\epsilon)$ gives the macrostate $E = 0$. The partition function for the two-particle system is

$$Z_2 = \sum_s g(s) e^{-\beta U_s} = e^{-2\beta\epsilon} + 2 + e^{2\beta\epsilon} = 2^2 \cosh^2(\beta\epsilon) = Z_1^2,$$

which by no surprise is consistent with equation (145). It easily follows that for N distinguishable particles, the partition function is

$$Z_N = 2^N \cosh^N(\beta\epsilon) . \quad (176)$$

With the partition function Z_N known, all of the two-state thermodynamics may be derived. Of particular interest is the internal energy $U = -\partial \ln Z / \partial \beta$, which means

$$U = -N\epsilon \tanh(\beta\epsilon) . \quad (177)$$

Note that the high - temperature limit of (177) yields a formula perfectly analogous to the energy - temperature relation (109) derived for the two-state paramagnet. For small values of β , we have

$$U \approx -\frac{N\epsilon^2}{KT} ,$$

which should remind you that in certain special circumstances, the two-state system can actually have negative temperature.

When the temperature T is very high, β is very small, and the energy U approaches zero. This makes sense because a high - temperature canonical system has all of the microstates more - or - less equally occupied, and the system - wide energy will be near - equal contributions of $+\epsilon$ and $-\epsilon$. Likewise energy is maximized (or minimized) at low temperatures. Particle states aren't as subject to thermal randomization, which allows for greater coherence at the macroscopic level.

26.4 Entropic Rubber Band

Physical systems can sometimes be understood by remarkably simple models; a few successful examples being the ideal gas model, van der Waals model, and the Einstein solid. Here I introduce the so - called *entropic spring* model.

Consider a stretchable material classified as a polymer chain, such as a rubber band or stretchable tube, that is under linear tension. The tension F is analogous to the pressure in an ordinary thermodynamic system, and therefore length replaces volume as the important spatial displacement. The material is also in contact with a thermal reservoir of temperature T . The internal energy U of the material is reasonably defined as zero at the outset - polymers often have much more energy associated with stretching than with bending. The entropic spring system is one with a fixed number of particles at constant temperature and under constant force. It will therefore tend to minimize the Gibbs free energy

$$G = -Fx - TS ,$$

where Fx has units of energy, as it should.

The fine details of the entropic spring model vary per system. Imagine a one-dimensional chain consisting of $N \gg 1$ segments under tension in the x -direction. Each segment may occupy one of the two states: (i) those oriented parallel to the x -direction contribute length a to the variable length L ; (ii) those aligned perpendicular to the x -direction contribute nothing to the length L .

The chain length L is not equal to a times the number of segments unless the material is completely stretched. The length is instead a times the number of x - oriented segments N_{\parallel} . Equivalently, the length is written $L = Nx$, where x is the segment length a multiplied by the ratio N_{\parallel}/N .

Microcanonical Analysis

The multiplicity of this entropic spring, as a two-state system, can be written exactly as

$$\Omega(N_{\parallel}, N) = \frac{N!}{N_{\parallel}! (N - N_{\parallel})!} .$$

Using Stirling's approximation $\ln(B!) \approx B \ln(B) - B$, the entropy $S = K \ln \Omega$ is

$$S = K [N \ln N - N_{\parallel} \ln N_{\parallel} - (N - N_{\parallel}) \ln (N - N_{\parallel})]$$

By inspection of $\Omega(N_{\parallel}, N)$ and/or S , the most probable macrostate of the system, with zero force applied, has $N_{\parallel} = N/2$. Therefore define the 'equilibrium' length as $\langle L_0 \rangle = aN/2$.

Writing N_{\parallel} as a function of x , and minimizing the Gibbs free energy $G = -Fx - TS$ with respect to x , arrive at the equation of state,

$$F = -\frac{KT}{a} \ln \left[\frac{a}{\langle x \rangle} - 1 \right] .$$

Solving for $\langle x \rangle$ immediately leads to $\langle L \rangle$, the length of the entropic chain as a function of temperature T and applied force F :

$$\langle x \rangle = \frac{a}{1 + e^{-\beta Fa}} \quad \rightarrow \quad \langle L \rangle = \frac{aN}{1 + e^{-\beta Fa}}$$

When the polymer chain is near the equilibrium length $aN/2$, the overall behavior of the system behaves as a Hookean spring. To show this, define the displacement length $\Delta = \langle L \rangle - aN/2$. The equation of state reads

$$F = \frac{KT}{a} \left[\ln \left(\frac{1}{2} + \frac{\Delta}{aN} \right) - \ln \left(\frac{1}{2} - \frac{\Delta}{aN} \right) \right] .$$

For small values of Δ , the equation of state is

$$F = \frac{KT}{a} \left[-\ln 2 + \frac{2\Delta}{aN} + \ln 2 + \frac{2\Delta}{aN} \right] = \frac{4KT}{a^2 N} (\langle L \rangle - \langle L_0 \rangle) .$$

Evidently, the effective spring constant is $4KT/a^2 N$. Notice that temperature is in the numerator: the greater the temperature, the more the system resists being stretched. This is opposite the behavior of metals, which lose stiffness at high T .

Canonical Analysis

Now repeat the previous analysis using Boltzmann factors. The $N \gg 1$ polymer chain segments, although connected, are noninteracting. The partition function takes the form $Z = (Z_1)^N$, where the single-link partition function is

$$Z_1 = e^{-\beta E_{\perp}} + e^{-\beta E_{\parallel}} .$$

Segments in the \perp orientation are not contributing to the system's energy, and each segment in the \parallel orientation contributes energy $-Fa$. Thus,

$$Z_1 = 1 + e^{\beta Fa} .$$

The average length per segment $\langle x \rangle$ may be calculated statistically by equation (138). The ‘position operator’ for the \perp orientation is simply $\hat{0}$, and is \hat{a} for the \parallel orientation:

$$\langle x \rangle = \frac{\hat{0} \cdot 1 + \hat{a} \cdot e^{\beta F a}}{Z_1} = \frac{a}{1 + e^{-\beta F a}}$$

Not surprisingly, the canonical and microcanonical calculations of $\langle x \rangle$ are in agreement. So too is the length $\langle L \rangle = N \langle x \rangle$.

At high temperatures, the length $\langle x \rangle$ becomes, after using Taylor expansion on the exponential,

$$\langle x \rangle = \frac{F a^2}{4KT} + \frac{a}{2} .$$

Multiply through by N to (re)(dis)cover the effective Hookean spring constant of the system, $4KT/a^2 N$.

27 Paramagnetism as Canonical Ensemble

27.1 Classical Paramagnet

Consider a ‘solid’ system of distinguishable noninteracting particles at temperature T that are subject to a uniform magnetic field $\vec{B} = B\hat{z}$. In the classical viewpoint, each particle’s magnetic moment $\vec{\mu}$ may be oriented at an arbitrary angle with respect to \hat{z} . A particle with classical magnetic moment $\vec{\mu}$ in a uniform magnetic field $B\hat{z}$ has potential energy

$$U = -\vec{\mu} \cdot \vec{B} = -\mu_z B_z = -\mu B \cos \theta ,$$

where θ is the angle between $\vec{\mu}$ and the z - axis. We are interested in calculating the magnetization $\langle \mu_z \rangle$ of the material, which is the average z - component of $\vec{\mu}$ per particle. With momentum playing no role in the magnetization, the single-particle partition function is

$$Z_1 = \int_0^\pi e^{\beta \mu B \cos \theta} \sin \theta d\theta d\phi = \frac{4\pi}{\beta \mu B} \sinh(\beta \mu B) ,$$

which is an integral over a 4π - shell embedded in three - space. Note that Z_1 would need a momentum - like term if to derive dynamical information about the system.

The magnetization $\langle \mu_z \rangle$ is simply the ensemble average of $\mu \cos \theta$, or

$$\langle \mu_z \rangle = \int_0^\pi \mu \cos \theta e^{\beta \mu B \cos \theta} \sin \theta d\theta d\phi = \mu \left(\coth(\beta \mu B) - \frac{1}{\beta \mu B} \right) , \quad (178)$$

a result known as the *Langevin function*. At high temperatures (or small β), the quantity in parenthesis reduces to $\mu B/3KT$, known as the *Curie law*, which states that the magnetization of a paramagnetic particle is prortional to B and inversely proportional to T :

$$\langle \mu_z \rangle \propto \frac{\mu^2 B}{K T} \quad (179)$$

The prefactor infront of B/T is called the *Curie constant* for a given material.

27.2 Two-State Quantum Paramagnet

Consider a paramagnetic system made of spin - $1/2$ particles. In the two-state viewpoint of paramagnetism, each particle's magnetic moment $\vec{\mu}$ may take one of two alignments: parallel to \hat{z} , or antiparallel to \hat{z} . Parallel and antiparallel vectors carry opposite energies according to

$$U = \pm \mu B .$$

Following 'standard' two-state analysis, the partition function for a single particle is given by (175), namely

$$Z_1 = \sum_{j=1,2} e^{-\beta U_j} = 2 \cosh (\beta \mu B) .$$

The magnetization $\langle \mu_z \rangle = \langle \mu \rangle$ shall be calculated in analogy to the internal energy, equation (177). Specifically,

$$\langle \mu \rangle = \frac{\partial}{\partial (\beta B)} \ln [2 \cosh (\beta \mu B)] = \mu \tanh (\beta \mu B) ,$$

which if the magnetic field strength B is sufficiently low and the temperature T is sufficiently high, reproduces the Curie law with Curie constant $C = \mu^2 / K$.

27.3 General Quantum Paramagnet

In the general viewpoint of paramagnetism, each particle's spin vector has quantized projections on the z - axis. If a given particle's total spin has magnitude J , it follows from standard quantum mechanics that the z - projection of the spin is given by $m_s \hbar$, where m_s is the spin quantum number ranging from $-J$ to J in integer or half - integer steps.

The Hamiltonian of one particle is

$$H = -\mu_z B = -\gamma m_s B ,$$

where γ is the gyromagnetic ratio (a constant) of the particle species. For an isolated electron, the gyromagnetic ratio is $\gamma = g_e \mu_B / \hbar$, where $g_e \approx 2.00$ and μ_B is the Bohr magneton, given by

$$\mu_B = \frac{e \hbar}{2 m_e} , \quad (180)$$

where m_e is the electron mass and e is the electron charge.

The partition function is straightforwardly written as the sum of all possible Boltzmann factors

$$Z_1 = \sum_{m_s=-J}^{m_s=J} \left(e^{-\beta \gamma B} \right)^{m_s} = \frac{1 - e^{-q(J+1)}}{1 - e^{-q}} + \frac{1 - e^{q(J+1)}}{1 - e^q} - 1 ,$$

where the substitution $q = \beta \gamma B$ has been made and the geometric series identity

$$1 + q + q^2 + q^3 + \dots + q^{n-1} = \frac{1 - q^n}{1 - q} \quad (181)$$

has been used. Thus Z_1 simplifies to

$$Z_1 = \frac{\sinh (\beta \gamma B (J + 1/2))}{\sinh (\beta \gamma B / 2)} .$$

By computing the observable $\langle M \rangle \propto \langle \gamma m_s \rangle$, it can be shown that the magnetization of a paramagnet with N particles per unit volume is

$$\langle M \rangle = N\gamma JB_J(x) ,$$

where $x = \beta\gamma JB$, and the function $B_J(x)$ is called the *Brillouin function*, which is a generalization of the Langevin function as follows:

$$B_J(x) = \frac{2J+1}{2J} \coth\left(\frac{2J+1}{2J}x\right) - \frac{1}{2J} \coth\left(\frac{x}{2J}\right) \quad (182)$$

If the magnetic field strength B is sufficiently low and the temperature T is sufficiently high, it can be shown that the Curie constant for any J is

$$C = \frac{\partial \langle M \rangle}{\partial B} = \frac{\mu_B^2}{3K} N g_e^2 J(J+1) .$$

The special case $J = 1/2$ makes the Brillouin function simplify to $\tanh(x)$, which in the small - x limit reduces to x , matching the previously - found result for the two-state paramagnet.

Classical behavior is recovered from the quantum analysis by allowing the spin vector to occupy the whole continuum of solid angle. This can only be accomplished by setting $J \rightarrow \infty$. In this case, the Brillouin function reduces, as expected, to the Langevin function.

28 Grand Canonical Ensemble

We have encountered two regimes so far - the microcanonical and canonical ensembles. Recall that the microcanonical ensemble deals with systems with fixed energy, fixed volume, and fixed number of particles. The canonical ensemble applies to systems in contact with a thermal reservoir of one temperature, but the systems may not exchange matter with the reservoir. Here I introduce the *grand canonical ensemble* where the system exchanges both energy and particles with the environment (the system volume remains fixed).

28.1 Minimization of Grand Free Energy

Consider a system with internal energy U_1 , number of particles N_1 , and volume V_1 that is connected to a reservoir with internal energy U_2 , number of particles N_2 , and volume V_2 . Denote the system's entropy S_1 , and denote the reservoir's multiplicity S_2 . Before writing any equations, we can expect to find that (i) thermal equilibrium implies $T_1 = T_2$, and (ii) diffusive equilibrium implies $\mu_1 = \mu_2$.

The total internal energy is $U_T = U_1 + U_2$, meaning $dU_1 = -dU_2$. Likewise the total number of particles is $N_T = N_1 + N_2$, implying $dN_1 = -dN_2$. The total entropy is

$$S_T = S_1(U_1, N_1) + S_2(U_T - U_1, N_T - N_1) ,$$

and we are interested in the U - and N - derivatives of S . At thermal equilibrium,

$$\frac{\partial S}{\partial U_1} = 0 = \left(\frac{\partial S_1}{\partial U_1} \right)_{V_1, N_1} - \left(\frac{\partial S_2}{\partial U_1} \right)_{V_2, N_2} ,$$

which if you remember equation (48) implies that $T_1 = T_2$, as expected. At diffusive equilibrium,

$$\frac{\partial S_T}{\partial N_1} = 0 = \left(\frac{\partial S_1}{\partial N_1} \right)_{U_1, V_1} - \left(\frac{\partial S_2}{\partial N_1} \right)_{U_2, V_2} ,$$

which if you remember equation (50) implies that $\mu_1 = \mu_2$, as expected.

Now we seek the quantity that is overall minimized for the grand canonical ensemble. Begin by writing the first - order Taylor expansion of the total entropy S :

$$S_T = S_1(U_1, N_1) + S_2(U_2, N_2) - U_1 \left(\frac{\partial S_2}{\partial U_1} \right)_{V_2, N_2} - N_1 \left(\frac{\partial S_2}{\partial N_1} \right)_{U_2, V_2}$$

Note that S_1 is untouched, and that $S_2(U_2, N_2)$ is a (very large) constant of no physical consequence. The remaining terms may be replaced by equations (48) and (50), respectively, to get

$$S_T = S_1(U_1, N_1) - \frac{U_1}{T} + \frac{\mu N_1}{T} . \quad (183)$$

The equilibrium state is the one that maximizes the total entropy S_T , which means we must maximize the quantity $TS - U + \mu N$, dropping the ₁ subscript and multiplying through by T . We are seeking something to minimize, which is therefore

$$\Phi = U - TS - \mu N = -PV ,$$

and you should recognize this result as the grand free energy. Using the definition $S = K \ln \Omega$, it follows that

$$\Phi = -KT \ln \Omega_T . \quad (184)$$

(The symbol Ω_T shall be subsequently replaced.) The equilibrium condition is consistent with the previously - established equation (65), namely

$$dS_{uni} = -\frac{1}{T} (d\Phi)_{T, V, \mu} .$$

28.2 Gibbs Gactor and Grand Partition Function

The grand canonical ensemble is an extension of the (ordinary) canonical ensemble. It follows that the Boltzmann factor must gain a term relating to number of particles and the chemical potential. To see this, return to equation (183) and multiply through by the quantity $K^{-1}T$. Then take the exponential of each side (reversing the logarithm), arrive at

$$\Omega_T = \Omega_1 + e^{-\beta(U_1 - \mu N_1)} .$$

Finally, note that Ω_1 is very small compared to Ω_T and may be ignored.

With the multiplicity known, it follows that the probability of the gand canonical system containing N_s paticles occupying energy state k is:

$$Pr(k, N_s) = \frac{1}{Z_G} e^{-\beta[U(k) - \mu N_s]} \quad (185)$$

The expoential term in (185) is called the *Gibbs factor*, which is a generalization of the Boltzmann factor. The term Z_G is called the *grand partition function*. Since the sum of all probabilities must equal unity, the normalization statement

$$\sum_{N_s} \sum_k Pr(N_s, k) = 1$$

tells us the grand partition function is

$$Z_G = \sum_{N_s} \sum_k e^{-\beta[U(k) - \mu N_s]} . \quad (186)$$

Notice the summation is over both the states k and the number of particles N_s .

Recall that the factor $\exp(\beta\mu)$ is identically the fugacity, equation (129), denoted by the symbol z . Also identify the sum over states k combined with the Boltzmann factor $\exp(-\beta U(k))$ as being equivalent to the canonical partition function Z for N_s particles in the system. We have found:

$$Z_G(z, V, T) = \sum_{N_s} z^{N_s} Z_{N_s}(V, T) \quad (187)$$

28.3 Observables

Recall that the average value of any observable macrostate O is calculated by a summation over all microstates of \hat{O} times the probability density factor. For the grand canonical ensemble, equation (138) generalizes to

$$\langle O \rangle = \frac{1}{Z_G} \sum_{N_s} \sum_k \hat{O} e^{-\beta[U(k) - \mu N_s]} .$$

The pressure in a grand canonical system is easily expressed in terms of T , V , and Z_G without performing a statistical calculation, thus no need for a pressure operator \hat{P} . Recall that the grand free energy Φ is equal to $-PV$, and meanwhile we found that $\Phi = -KT \ln Z_G$ in accordance with equation (184). Evidently then,

$$P = \frac{KT}{V} \ln Z_G . \quad (188)$$

The system's internal energy $U = \langle U \rangle$ is straightforwardly written

$$U = \frac{1}{Z_G} \sum_{N_s} \sum_k U(k) e^{-\beta[U(k) - \mu N_s]} = - \left[\frac{\partial}{\partial \beta} \ln(Z_G) \right]_{z, V} , \quad (189)$$

where the explicit summation has been avoided by the 'derivative trick' first seen in the canonical ensemble analysis. Similarly, the average number of particles $N = \langle N \rangle$ in the system is

$$N = \frac{1}{Z_G} \sum_{N_s} \sum_k N_s e^{-\beta[U(k) - \mu N_s]} = \left[KT \frac{\partial}{\partial \mu} \ln(Z_G) \right]_{V, T} . \quad (190)$$

The Helmholtz free energy may be written as $A = -PV + \mu N$ by well - established identities. Expressing PV in terms of Z_G , and also inverting the chemical potential via $e^{\beta\mu} = z$, we have

$$A = -KT \ln \left(\frac{Z_G}{z^N} \right) . \quad (191)$$

Next, the entropy can be written from the identity $S = (U - A)/T$. Filling in the above results for U and A , we get

$$S = KT \frac{\partial}{\partial T} (\ln Z_G)_{z, V} + K \ln Z_G - NK \ln z . \quad (192)$$

Equivalently, the entropy is given by the statistical calculation

$$S = -K \sum_{N_s} \sum_k Pr(N_s, k) \ln(Pr(N_s, k)) , \quad (193)$$

where $Pr(N_s, k)$ is given by (185).

28.4 Fluctuations

wfb

28.5 Ideal Gas as Grand Canonical Ensemble

The ideal gas regime falls elegantly out of the grand canonical ensemble. Starting with the canonical calculation of the ideal gas partition function Z as (149), the grand partition function Z_G of the ideal gas is

$$Z_G(z, V, T) = \sum_{N_s} z^{N_s} \frac{1}{N_s!} \left(\frac{V}{\lambda_0^3} \right)^{N_s} ,$$

where λ_0 is the thermal de Broglie wavelength $h/\sqrt{2\pi mKT}$. Note the lowercase z is the fugacity.

The formula for Z_G becomes exponential if we take the limits on the N_s -sum as $0 \leq N_s \leq \infty$. To see this, let $y = zV/\lambda_0^3$, and Z_G becomes

$$Z_G = \sum_{N_s=0}^{\infty} \frac{y^{N_s}}{N_s!} = e^y = \exp\left(\frac{zV}{\lambda_0^3}\right) . \quad (194)$$

Taking the natural logarithm of both sides, we have

$$\ln Z_G = \frac{zV}{\lambda_0^3} = \frac{zV (2\pi m)^{3/2}}{h^3 \beta^{3/2}} .$$

The pressure P and internal energy U are given by equations (188) and (189), respectively. Explicitly these are

$$P = \frac{KT}{V} \ln Z_G \quad U = \frac{3}{2} KT \ln Z_G ,$$

implying

$$U = \frac{3}{2} PV .$$

It remains to show that the number of particles N is equivalent to $\ln Z_G$. Recall that the chemical potential of the ideal gas, according to equation (126), is

$$\mu = -KT \ln \left(\frac{V}{N \lambda_0^3} \right) ,$$

where multiplying through by β and taking the exponential of each side, we have N in terms of the fugacity

$$N = \frac{zV}{\lambda_0^3} .$$

Notice next that zV/λ_0^3 is precisely equal to the natural logarithm of Z_G as in equation (194), thus proving that, for the ideal gas, $\ln Z_G = N$. Inserting this into our previous achievements (namely the results P and U), we recover the ideal gas law $PV = NKT$.

29 Imaginary Time and Density Operator

29.1 Temperature as Imaginary Time

Quantum mechanics is built on only few principles mainly centered around the Schrodinger equation. In bra - ket notation, the time dependent Schrodinger equation may be written as

$$\hat{H} |\psi(t)\rangle = -i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle , \quad (195)$$

where $|\psi(0)\rangle$ is the state vector at time t , and the Hamiltonian is assumed to not vary with time. One way to express the solution for $|\psi(t)\rangle$ is

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar} |\psi(0)\rangle ,$$

where the exponential term is taken as a matrix operator. Formally, it is known as the *time evolution operator*, or *propagator* of a closed quantum system.

From standard quantum mechanics, we know that there usually exists a basis of energy eigenfunctions $\{|\phi\rangle_n\}$, in where the Hamiltonian matrix \hat{H} has energy eigenvalues along the diagonal and zero for off - diagonal components. The matrix \hat{H} appears inside the exponent in the time evolution operator, which by Taylor expansion means that the operator is an infinite sum of matrices. If \hat{H} is non - diagonal, the time evolution operator is difficult to write down.

With the Hamiltonian expressed in an orthonormal basis such that \hat{H} is diagonal, it follows that the time evolution operator $\exp(-i\hat{H}t/\hbar)$ is also diagonal. Notice now that the trace of the time evolution operator

$$\text{Tr} \left(e^{-i\hat{H}t/\hbar} \right) = \sum_s e^{-iU_s t/\hbar}$$

looks very much like the canonical partition function Z , except temperature is missing and time is now involved. If we make the interpretation

$$\beta = \frac{it}{\hbar} , \quad (196)$$

then the trace of the imaginary-time evolution operator is *equal* to the partition function:

$$Z = \text{Tr} \left(e^{-\beta\hat{H}} \right) = \sum_s \langle \phi_s | e^{-\beta\hat{H}} | \phi_s \rangle \quad (197)$$

29.2 Density Operator

In the canonical ensemble, observables may be calculated once the partition function for a system is known. In quantum mechanics, we are used to using the probability density $|\phi^*\phi|$. Now that we know these are related by equation (197), define the *density operator* or *density matrix* $\hat{\rho}$ as

$$\hat{\rho} = \frac{1}{Z} e^{-\beta\hat{H}} , \quad (198)$$

or more specifically,

$$\rho_{ss} = \frac{1}{Z} e^{-\beta U_s} |\phi_s\rangle \langle \phi_s| ,$$

with normalization requirement

$$\text{Tr}(\hat{\rho}) = 1 . \quad (199)$$

The density operator is the tool for computing statistical observables in the canonical ensemble. For any observable x corresponding to the operator $\hat{x} = x_{jk}$, the ensemble average of x is

$$\langle x \rangle = \frac{1}{Z} \sum_k e^{-\beta U_k} \langle \phi_k | \hat{x} | \phi_k \rangle = \text{Tr} (\hat{x} \hat{\rho}) . \quad (200)$$

Recall that the Hamiltonian is already assumed to be in a representation with energy eigenvalues on the diagonal and zero elsewhere. If you need to work in a non - energy eigenbasis, perform a *similarity transform* on $\hat{\rho}$, meaning calculate $\hat{\rho}' = S \hat{\rho} S^{-1}$, where S is a matrix whose column entries are the eigenvector components of $\hat{\rho}$. The new matrix $\hat{\rho}'$ is not diagonal. To switch back, use $\hat{\rho} = S^{-1} \hat{\rho}' S$. Changing basis by similarity transform does not affect observables, as

$$\text{Tr} (\hat{x}' \hat{\rho}') = \text{Tr} (S \hat{x} S^{-1} S \hat{\rho} S^{-1}) = \text{Tr} (\hat{x} \hat{\rho}) ,$$

where the quantity $S^{-1} S$ is equal to the identity matrix.

In the general case, the components of the density operator are ρ_{mn} , which evolve in time following the Schrodinger equation. Explicitly,

$$i\hbar \frac{\partial}{\partial t} \rho_{mn} = i\hbar \frac{\partial}{\partial t} (|\psi_m\rangle \langle \psi_n|) = H_{mk} |\psi_k\rangle \langle \psi_n| - |\psi_m\rangle \langle \psi_j| H_{jn} ,$$

where the fact that \hat{H} is Hermitian has been exploited. Simplifying,

$$i\hbar \frac{\partial}{\partial t} \rho_{mn} = H_{mk} \rho_{kn} - \rho_{mj} H_{jn} .$$

Note the right side is the commutation between ρ and H . In basis - free notation, we arrive at

$$i\hbar \frac{\partial}{\partial t} \hat{\rho} = [\hat{H}, \hat{\rho}] ,$$

a close analog to Liouville's theorem from classical mechanics.

Problem 77: Consider the matrix:

$$A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

Find the inverse of A by Gauss - Jordan elimination $AI \rightarrow IA^{-1}$.

Problem 78: Consider the matrix:

$$M = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$$

Determine the eigenvalues, eigenvectors, the similarity matrix S , and its inverse S^{-1} . Verify the relation $M = SDS^{-1}$, where D is the diagonalization of M (nonzero elements are eigenvalues along the diagonal). Answer:

$$\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{bmatrix}$$

29.3 Spin 1/2 System

Consider a quantum paramagnet made of spin - 1/2 electrons subject to external magnetic field \vec{B} . The Hamiltonian operator \hat{H} for one particle is

$$\hat{H} = -\vec{\mu}_s \cdot \vec{B} ,$$

where $\vec{\mu}_s$ is the spin magnetic moment of the particle, where for instance an electron has

$$\vec{\mu}_s = -\frac{g_e e}{2m_e} \vec{S} .$$

The object $\vec{S} = \hat{S}$ is the *spin operator*, which may be written

$$\vec{S} = \frac{\hbar}{2} \vec{\sigma} ,$$

where $\vec{\sigma} = \hat{\sigma}$ are the famous *Pauli spin matrices*:

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Diagonal Hamiltonian

Suppose the magnetic field points only along the z -direction such that $\vec{B} = B\hat{z}$, and also recall that the factor $g_e \approx 2.00$, making the magnitude of the spin magnetic moment is approximately equivalent to the Bohr magneton μ_B . The Hamiltonian operator becomes

$$\hat{H} = -\mu_B B \hat{\sigma}_z ,$$

which has energy eigenvalues

$$U = \pm \mu_B B$$

and eigenvectors

$$|\chi_+\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |\chi_-\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

The density operator $\hat{\rho}$ for this system shall be constructed using the definition:

$$\hat{\rho} = \frac{e^{-\beta \hat{H}}}{Z} = \frac{1}{Z} \left(1 - \beta \hat{H} + \frac{1}{2!} (\beta \hat{H})^2 - \frac{1}{3!} (\beta \hat{H})^3 + \dots \right)$$

The fact that \hat{H} is diagonal is quite useful, as we see

$$\hat{\sigma}_z^{N \text{ even}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \hat{I} \quad \hat{\sigma}_z^{N \text{ odd}} = \hat{\sigma}_z ,$$

resulting in

$$\hat{\rho} = \frac{1}{Z} \left[\hat{I} \cosh(\beta \mu_B B) + \hat{\sigma}_z \sinh(\beta \mu_B B) \right] ,$$

or in matrix form,

$$\hat{\rho} = \frac{2}{Z} \begin{bmatrix} e^{\beta \mu_B B} & 0 \\ 0 & e^{-\beta \mu_B B} \end{bmatrix} ,$$

which in hindsight could have been written without the expansion of the exponential. Indeed when \hat{H} is diagonal, the density matrix is simple to write.

By the normalization condition $\text{Tr}(\hat{\rho}) = 1$, the partition function Z is

$$Z = 2 \cosh(\beta\mu_B B) ,$$

which is precisely what we expect from standard two-state analysis. With the density operator known, we may calculate observable expectation values. For instance, the average z - value of the system spin is

$$\langle S_z \rangle = \frac{\hbar}{2} \text{Tr}(\hat{\sigma}_z \hat{\rho}) = \frac{\hbar}{2} \tanh(\beta\mu_B B) ,$$

and the average x - value of the system spin is

$$\langle S_x \rangle = \frac{\hbar}{2} \text{Tr} \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{bmatrix} \right) = 0 .$$

Non - Diagonal Hamiltonian

Now repeat the analysis in a less convenient basis to derive all the same physics. For eigenvectors we choose those that satisfy $\hat{\sigma}_x |\phi\rangle = \pm |\phi\rangle$, which in terms of the $\{|\chi\rangle\}$ eigenvectors are written:

$$|\phi_+\rangle = \frac{1}{\sqrt{2}} (|\chi_+\rangle + |\chi_-\rangle) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$|\phi_-\rangle = \frac{1}{\sqrt{2}} (|\chi_+\rangle - |\chi_-\rangle) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

The similarity matrix is the juxtaposition of eigenvector components as follows:

$$S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \rightarrow \quad S^{-1} = \sqrt{2} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

Now we may write the new density operator $\hat{\rho}'$ by the similarity transform $\hat{\rho}' = S\hat{\rho}S^{-1}$, giving

$$\hat{\rho}' = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} \cosh(\beta\mu_B B) & \sinh(\beta\mu_B B) \\ \sinh(\beta\mu_B B) & \cosh(\beta\mu_B B) \end{bmatrix} .$$

Meanwhile, the z - and x - spin operators become, respectively:

$$\hat{\sigma}'_z = S\hat{\sigma}_zS^{-1} = \hat{\sigma}_x$$

$$\hat{\sigma}'_x = S\hat{\sigma}_xS^{-1} = \hat{\sigma}_z$$

Calculations of $\langle S_z \rangle$ and $\langle S_x \rangle$ should agree with those above. In this case,

$$\langle S_z \rangle = \frac{\hbar}{2} \text{Tr}(\hat{\sigma}_x \hat{\rho}') = \frac{\hbar}{2} \tanh(\beta\mu_B B)$$

$$\langle S_x \rangle = \frac{\hbar}{2} \text{Tr}(\hat{\sigma}_z \hat{\rho}') = 0 ,$$

as expected.

By the same similarity transform, the Hamiltonian \hat{H}' in the $\{|\phi\rangle\}$ basis comes out to

$$\hat{H}' = S\hat{H}S^{-1} = \begin{bmatrix} H_{11} + H_{22} & H_{11} - H_{22} \\ H_{11} - H_{22} & H_{11} + H_{22} \end{bmatrix},$$

but $H_{11} = -H_{22} = -\mu_B B = E$, so we have

$$\hat{H}' = \begin{bmatrix} 0 & 2E \\ 2E & 0 \end{bmatrix},$$

where the eigenvector $|\phi_+\rangle$ has eigenvalue $2E = -2\mu_B B$, and the eigenvector $|\phi_-\rangle$ has eigenvalue $-2E = 2\mu_B B$.

Problem 79: Consider the Hamiltonian

$$\hat{H}' = \begin{bmatrix} \epsilon & \Delta \\ \Delta & \epsilon \end{bmatrix},$$

which has been represented in the basis $(|1\rangle, |2\rangle)$. Construct the density matrix and the partition function, and then calculate the energy observable $\langle U \rangle$. Answer:

Denote $\hat{\rho}'$ as the non - diagonal density matrix as represented in the basis $(|1\rangle, |2\rangle)$ such that $\hat{\rho}' = S\hat{\rho}S^{-1}$, where $\hat{\rho}$ is the density matrix in some diagonal basis and S is the similarity matrix. Likewise, the non - diagonal Hamiltonian \hat{H}' relates to its diagonal cousin \hat{H} by the similarity transform $\hat{H} = S^{-1}\hat{H}'S$.

The eigenvectors and eigenvalues of \hat{H}' are

$$\begin{aligned} |\psi_+\rangle &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \rightarrow & U_+ = \epsilon + \Delta \\ |\psi_-\rangle &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} & \rightarrow & U_- = \epsilon - \Delta, \end{aligned}$$

which simultaneously reveals the similarity matrix S and the diagonal Hamiltonian \hat{H} . (It turns out that S is identical to that in the spin - 1/2 system solved previously and is not rewritten here.) The density matrix $\hat{\rho}'$ is therefore

$$\hat{\rho}' = S \left(S^{-1} \frac{e^{-\beta\hat{H}}}{Z} S \right) S^{-1} = \frac{e^{-\beta\epsilon}}{Z} \begin{bmatrix} \cosh \Delta & -\sinh \Delta \\ -\sinh \Delta & \cosh \Delta \end{bmatrix},$$

and by normalization $\text{Tr}(\hat{\rho}') = 1$, we have

$$Z = 2e^{-\beta\epsilon} \cosh \Delta.$$

Finally, the energy observable $\langle U \rangle$ is

$$\langle U \rangle = \text{Tr}(\hat{H}'\hat{\rho}') = \epsilon - \Delta \tanh \Delta.$$

29.4 Position Representation

The density operator $\hat{\rho} = \exp(\beta\hat{H})/Z$ works naturally in an orthonormal energy eigenbasis $\{|\phi\rangle\}$ in where \hat{H} is diagonal. It's useful to address which other basis choices are equally natural.

Denote the position eigenstates by basis vectors $\{|r\rangle\}$. The density operator may be written

$$\hat{\rho}_{r'r} = \frac{1}{Z} \langle r' | e^{\beta\hat{H}} | r \rangle = \frac{1}{Z} \sum_{ij} \langle r' | \phi_i \rangle \langle \phi_i | e^{\beta\hat{H}} | \phi_j \rangle \langle \phi_j | r \rangle ,$$

where the inner products between basis sets are analogous to the similarity transform matrices as in $\hat{\rho}' = S\hat{\rho}S^{-1}$. In this case, they are equal to the position representation of the energy eigenstates as follows:

$$\langle r' | \phi_i \rangle = \phi_i^*(r') \quad \langle \phi_j | r \rangle = \phi_j(r)$$

The exponential term is simply the diagonal matrix of energy Boltzmann factors $e^{-\beta U_j}$. The position representation is evidently

$$\hat{\rho}_{r'r} = \frac{1}{Z} \sum_{ij} \delta_{ij} e^{-\beta U_j} \phi_i^*(r') \phi_j(r) . \quad (201)$$

The normalization condition $\text{Tr}(\hat{\rho}_{r'r}) = 1$ must be satisfied. Setting index $r = r'$, we indeed see

$$\text{Tr}(\hat{\rho}_{r'=r}) = \frac{1}{Z} \sum_j e^{-\beta U_j} \int \phi_j^*(r) \phi_j(r) dr = 1 .$$

29.5 Free Particle in a Periodic Box

Consider a single particle of mass m 'confined' to a periodically - repeating cubic volume V . This is the near - equivalent setup used to analyze the quantum ideal gas, with the Hamiltonian being $H = -\hbar^2 \nabla^2 / 2m$. In this scenario the energy of the particle is given by

$$\epsilon = \frac{\hbar^2}{2mL^2} (n_x^2 + n_y^2 + n_z^2) ,$$

where the quantum numbers n_i are nonzero integers with negatives allowed. Hence we use traveling waves rather than standing waves as wavefunctions. The wavefunctions in position representation are

$$\phi_j(\vec{x}) = \frac{1}{\sqrt{V}} e^{i\vec{k}_j \cdot \vec{x}} ,$$

where the momentum \vec{k}_j is equal to $(2\pi/L)\vec{n}$ and relates to the energy by $\epsilon = \hbar^2 k^2 / 2m$.

The density matrix $\hat{\rho}_{x'x}$ in position representation shall be determined by equation (201), which gives

$$\hat{\rho}_{x'x} = \frac{1}{Z} \frac{1}{V} \sum_j e^{-\beta \epsilon_j} e^{i\vec{k}_j \cdot (\vec{x} - \vec{x}')} .$$

Recall that the thermal de Broglie wavelength is defined as $\lambda_0 = h/\sqrt{2\pi mKT}$. The density matrix becomes

$$\hat{\rho}_{x'x} = \frac{1}{Z} \frac{1}{V} e^{-\pi \Delta \vec{x}^2 / \lambda_0^2} \sum_j e^{-(\sqrt{\pi} \lambda_0 / (L) \vec{n}_j - i\sqrt{\pi}/(\lambda_0) \Delta \vec{x})^2} .$$

To a good approximation, the spectrum of n - values may be regarded as continuous, even though the individual n_i are integers. The summation may be replaced by an integral in three-dimensional n - space. We then use the Gaussian integral identity

$$\int_0^\infty q^2 e^{-aq^2} dq = \frac{\sqrt{\pi}}{4a^{3/2}},$$

and the density matrix simplifies to

$$\hat{\rho}_{x'x} = \frac{1}{Z} \frac{1}{\lambda_0^3} e^{-\pi \Delta \vec{x}^2 / \lambda_0^2}.$$

Note that the imaginary exponential has no n - dependence and simply shifts the integral, and has no physical consequence.

The normalization condition $\text{Tr}(\hat{\rho}_{x'=x}) = 1$ pins down the partition function Z . That is,

$$1 = \text{Tr}(\hat{\rho}) = \frac{1}{Z} \frac{1}{\lambda_0^3} \int_V e^0 d^3x \quad \rightarrow \quad Z = \frac{V}{\lambda_0^3},$$

which makes sense - because generalizing to N indistinguishable noninteracting particles, we have precisely the partition function (149) of the ideal gas:

$$Z_N = \frac{1}{N!} \left(\frac{V}{\lambda_0^3} \right)^N$$

The factor $1/N!$ can be formally traced back to the system wavefunction Ψ , which has certain symmetrization requirements when dealing with indistinguishable particles.

The energy of the single particle is the expectation value of the Hamiltonian, namely $H = -\hbar^2 \nabla^2 / 2m$. By equation, (200), we have

$$\langle H \rangle = \text{Tr}(\hat{H} \hat{\rho}) = \int \left(\frac{-\hbar^2}{2mV} \nabla^2 \exp[-\pi \Delta \vec{x}^2 / \lambda_0^2] \right) \Big|_{\Delta \vec{x}=0} d^3x = \frac{3}{2} KT,$$

as expected from the equipartition theorem. Alternatively,

$$\langle H \rangle = \frac{\text{Tr}(\hat{H} e^{-\beta \hat{H}})}{\text{Tr}(e^{-\beta \hat{H}})} = -\frac{\partial}{\partial \beta} \ln \text{Tr}(e^{-\beta \hat{H}}) = -\frac{\partial}{\partial \beta} \ln Z.$$

Problem 80: Evaluate the above integral to prove that $\langle H \rangle = 3KT/2$. See Pathria (1972) Section 5.4.

29.6 Simple Harmonic Oscillator

Density operator formalism can also be applied to the simple harmonic oscillator. For simplicity we consider a one-dimensional example having the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m} \partial_x^2 + \frac{1}{2} m \omega x^2,$$

with energy eigenvalues

$$\epsilon_n = \hbar \omega (n + 1/2),$$

where ω is the angular frequency and n is any positive integer including zero. The corresponding eigenfunctions are

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \frac{H_n(\xi)}{\sqrt{2^n n!}} e^{-\xi^2/2}$$

where

$$\xi = \sqrt{\frac{m\omega}{\hbar}} x,$$

and H_n are the (in)famous Hermite polynomials, obeying

$$H_n(q) = (-1)^n e^{q^2} \frac{d^n}{dq^n} e^{-q^2}.$$

The density matrix $\hat{\rho}_{xx'}$ in position representation shall be determined by equation (201), giving

$$\langle x | e^{-\beta \hat{H}} | x' \rangle = \frac{1}{Z} \sum_{n=1}^{\infty} e^{-\beta \epsilon_n} \psi_n(x) \psi_n(x').$$

Substituting the energy equation and wave functions, along with a hint from Kubo (1965), pp. 175 - 77, the above sum evaluates to

$$\hat{\rho}_{xx'} = \frac{1}{Z} \sqrt{\frac{m\omega}{2\pi\hbar \sinh 2\tau}} \times \exp \left[\frac{-m\omega}{4\hbar} \left((x+x')^2 \tanh \tau + (x-x')^2 \coth \tau \right) \right],$$

where $\tau = \beta\hbar\omega/2$.

The trace of $\hat{\rho}_{xx'}$ must equal unity, meaning

$$1 = \frac{1}{Z} \sqrt{\frac{m\omega}{2\pi\hbar \sinh 2\tau}} \int_{-\infty}^{\infty} \exp \left[\frac{-m\omega}{\hbar} x^2 \tanh \tau \right] dx.$$

Solving for Z , the partition function is

$$Z = \frac{1}{2 \sinh \tau} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}},$$

the single-particle case of equation (164).

The probability density $\hat{\rho}_{x=x'}$ has two very interesting limits. In the high - temperature (a.k.a. classical) limit where $\beta\hbar\omega \ll 1$, the distribution smears out all quantum effects, and is purely thermal:

$$(\hat{\rho}_{x=x'})_{T \rightarrow \infty} \approx \sqrt{\frac{m\omega^2}{2\pi KT}} \times e^{-m\omega^2 x^2 / 2KT}$$

(meaning there is no dependence on \hbar). On the other hand, when the temperature is low, $\beta\hbar\omega \gg 1$, and we find

$$(\hat{\rho}_{x=x'})_{T \rightarrow 0} \approx \sqrt{\frac{m\omega}{\pi\hbar}} \times e^{-m\omega x^2 / \hbar}$$

which is purely a quantum - mechanical result that doesn't mention temperature. In fact, it's proportional to the probability density of the ground - state oscillator.

The average energy $\langle H \rangle$ is easy to compute:

$$\langle H \rangle = -\frac{\partial}{\partial \beta} \ln Z = \frac{\hbar \omega}{2} \coth \left(\frac{\beta \hbar \omega}{2} \right)$$

Problem 81: In the vicinity $x = x'$, eliminate the tanh factor in the probability density $\hat{\rho}_{x=x'}$ to show

$$\langle x | \hat{\rho} | x \rangle = \sqrt{\frac{m\omega^2}{2\pi \langle H \rangle}} \times e^{-m\omega^2 x^2 / 2 \langle H \rangle}.$$

Note the root - mean - square displacement of the oscillator is simply

$$x_{rms} = \sqrt{\frac{\langle H \rangle}{m\omega^2}},$$

consistent with the equipartition theorem.

29.7 Grand Canonical Ensemble

Density operator formalism readily generalizes to situations where the number of particles is not fixed - known already as the grand canonical ensemble. As we wrote in equation (186), the grand partition function $Z_G(\mu, V, T)$ involves the sum over all (energy, momentum, etc.) states k available to a system, nested within the sum over the number of particles N_s in each energy state.

The generalization of the canonical density operator (198) is therefore

$$\hat{\rho} = \frac{1}{Z_G} e^{-\beta[\hat{H} - \mu \hat{N}]}, \quad (202)$$

where μ is the chemical potential of the system. Note that the density operator must not only commute with the Hamiltonian but also with the number operator \hat{N} .

With the normalization requirement $\text{Tr}(\hat{\rho}) = 1$, we find

$$Z_G = \text{Tr} \left(e^{-\beta[\hat{H} - \mu \hat{N}]} \right) = \sum_{N_s} \sum_k e^{-\beta[U(k) - \mu N_s]}, \quad (203)$$

permitting the calculation of observables. For some observable x , the ensemble average $\langle x \rangle$ is

$$\langle x \rangle = \frac{1}{Z_G} \text{Tr} \left(\hat{x} e^{-\beta \hat{H}} e^{\beta \mu \hat{N}} \right).$$

Recall next that, according to equation (129), the fugacity coefficient z is a shorthand for $e^{\beta \mu}$. We use this to replace everything in terms of the canonical partition function Z_N and the canonical ensemble average $\langle x \rangle_N$ for a given N . Thus,

$$\langle x \rangle = \left(\sum_{N=0}^{\infty} z^N Z_N \right)^{-1} \left(\sum_{N=0}^{\infty} z^N \langle x \rangle_N Z_N \right). \quad (204)$$

30 Quantum Statistics