Lecture Notes

Statistical Mechanics

Alejandro L. Garcia
San Jose State University

PHYSICS 260 / Spring 2012

Department of Physics, San Jose State University, San Jose CA 95192-0106

Preface

These lecture notes should supplement your notes, especially for those of you (like myself) who are poor note takers. But I don't recommend reading these notes while I lecture; when you go to see a play you don't take the script to read along. Continue taking notes in class but relax and don't worry about catching every little detail. **WARNING: THESE NOTES ARE IN DRAFT FORM AND PROBABLY HAVE NUMEROUS TYPOS; USE AT YOUR OWN RISK.**

The main text for the course is R. K. Pathria and Paul D. Beale, *Statistical Mechanics*, 3rd Ed., Pergamon, Oxford (2011). Other texts that will sometimes be mentioned and that you may want to consult include: K. Huang, *Statistical Mechanics*, 2nd Ed., Wiley, New York (1987); F. Reif, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York (1965); L.D. Landau and E.M. Lifshitz, *Statistical Physics*, Addison-Wesley, Reading Mass. (1969).

Here is the general plan for the course: First review some basic thermodynamics and applications; this is not covered in your text book but any undergraduate thermodynamics book should be suitable.

Given the entropy of a system, all other thermodynamic quantities may be found. The theory of statistical ensembles (Chapters $1-4$ in Pathria) will tell us how to find the entropy given the microscopic dynamics. The theory of ensembles connects mechanics and thermodynamics.

The calculations involved in ensemble theory can be quite hard. We start with the easier non-interacting systems, such as paramagnets, classical ideal gas, Fermion gas, etc., (Chapters $6-8$) and work up to interacting systems, such as condensing vapor and ferromagnets (Chapter $12-13$).

Most of statistical mechanics is concerned with calculating statistical averages but we briefly consider the variation of microscopic and macroscopic quantities about their mean values as predicted by fluctuation theory (Chapter 15). Finally we'll touch on my own specialty, computer simulations in statistical mechanics in Chapter 16.

Alejandro L. Garcia

Chapter 1

Thermodynamics

Equation of State

Lecture 1

We consider thermodynamic systems as physical systems entirely described by a set of thermodynamic parameters. Commonly used parameters include:

- pressure, P, and volume, V
- tension, τ and length, L
- magnetic field, H and magnetization, M

These parameters appear in pairs, the first being a generalized force and the second a generalized displacement, with the product of the pair having the dimensions of energy. All these parameters are well-defined by mechanics and are readily measured mechanically.

An additional thermodynamic parameter, temperature, T, is not a mechanical parameter; we defer its definition for now. Rather than trying to be general, let's say that our thermodynamic parameters are P, V, and T, for example, if our system was a simple gas in a container.

An isolated system (no energy or matter can enter or leave the system) will, in time, attain thermodynamic equilibrium. At thermodynamic equilibrium, the parameters do not change with time. Though thermometers have been in use since the time of Galileo (late 1500's) many concepts regarding temperature were not well understood until much later. In the mid 1700's Joseph Black used thermometers to establish that two substances in thermal contact have the same at thermodynamic equilibrium. This is counterintuitive — if metal and wood are at thermal equilibrium the metal still "feels colder" than the wood. This misleading observation is due to the fact that our sense of touch uses conduction of heat and thus is a poor thermometer.

The *equation of state* is the functional relation between P, V, and T at equilibrium (see Fig. 1.1). An example of an equation of state is the ideal gas law for a dilute gas,

$$PV = NkT \qquad \text{(Ideal gas only)}$$

where N is the number of molecules in the gas and k ($k = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant. Boltzmann's constant does not have any deep meaning. Since the temperature scale was developed long before the relation between heat and energy was understood the value of k simply allows us to retain the old Kelvin and Celsius scales. Later we'll see the theoretical justification for the ideal gas law in the next chapter.

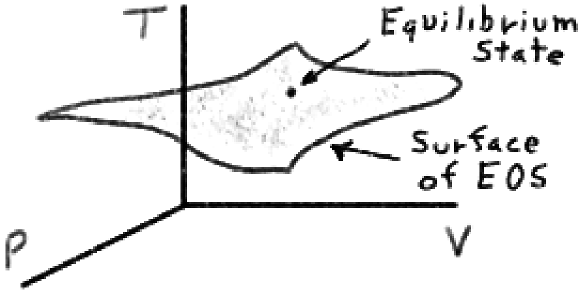

Figure 1.1: Point in state space on the equation of state surface

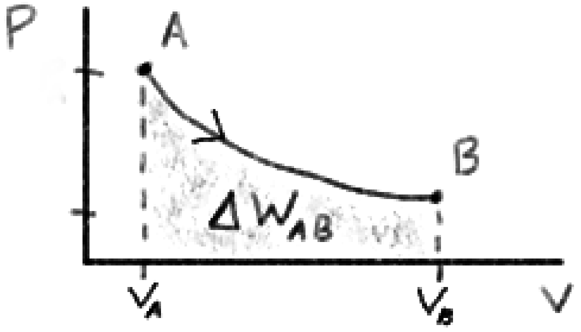

Figure 1.2: Path from points A to B in pressure-volume diagram

You should not get the impression that the equation of state completely specifies all the thermodynamic properties of a system. As an example, both argon and nitrogen molecules have about the same mass and the two gases obey the ideal gas law yet their molar heat capacities are very different (can you guess why Ar and N_2 are different?).

It follows from the equation of state that given V and T, we know P at equilibrium (or in general, given any two we know the third).

First Law of Thermodynamics

Call $U(V,T)$ the internal energy[*] of our system (why don't we write $U(P,V,T)$?). One way that the energy can change is if the system does mechanical work. The work done by a system in going from state A to B is,

$$\Delta W_{AB} = \int_{V_A}^{V_B} P(V)\, dV$$

since Work = Force × Distance = (Force/Area) × (Distance×Area) = (Pressure) × (Volume). Graphically, the work done is the area under the path travelled in going from A to B as shown in Fig. 1.2. The mental picture for this process is that of a piston lifting a weight as the system's volume expands. Notice that if the path between A and B is changed, ΔW_{AB} is different; for this reason $dW = P\,dV$ is *not* an exact differential.

---

[*] NOTATION: Pathria writes $E(V,T)$ instead of $U(V,T)$.

By conservation of energy,

$$U(V_B, T_B) - U(V_A, T_A) = \Delta Q_{AB} - \Delta W_{AB}$$

where ΔQ_{AB} is the non-mechanical energy change in going from A to B. We call ΔQ_{AB} the heat added to the system in going from A to B. If A and B are infinitesimally close,

$$dU = dQ - dW = dQ - PdV \qquad (*)$$

where dU is an exact differential though dQ and dW are not. This is the *first law of thermodynamics*, which simply states that total energy is conserved.

The results so far are very general. If our mechanical variables are different, everything carries over. For example, for a stretched wire the mechanical variables are tension τ and length L instead of P and V. We simply replace $P\,dV$ with $-\tau\,dL$ in equation $(*)$. Note that the minus sign comes from the fact that the wire does work when it contracts (imagine the wire contracting and lifting a weight).

Example: For n moles of an ideal gas held at constant temperature T_0 find the work done in going from volume V_A to V_B.

Solution: The number of molecules in n moles is $N = nN_A$ where $N_A = 6.205 \times 10^{23}$ is Avogadro's number (not to be confused with Avocado's number which is the number of molecules in a guaca-mole).

The ideal gas equation of state is

$$PV = NkT = (nN_A)kT = nRT$$

where $R = kN_A = 8.31$ J/(mol K) is the universal gas constant.

To find the work done, we solve,

$$\Delta W_{AB} = \int_{V_A}^{V_B} P(V)\,dV$$

Since the temperature is fixed at T_0, then $P(V) = nRT_0/V$ or

$$\Delta W_{AB} = nRT_0 \int_{V_A}^{V_B} \frac{dV}{V} = nRT_0 \ln(V_B/V_A)$$

Note that on a different path we would have to know how temperature varied with volume, i.e., be given $T(V)$ along the path.

Heat capacity

We define heat capacity, C, as

$$dQ = CdT$$

Specifically, the heat capacity at constant volume, C_V, is

$$C_V = \left(\frac{\partial Q}{\partial T}\right)_V$$

and the heat capacity at constant pressure

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

The heat capacity gives the amount of heat energy required to change the temperature of an object while holding certain mechanical variables fixed. We also define the heat capacity per unit mass, which is often called the specific heat, and the heat capacity per particle (e.g., $c_V = C_V/N$).

For liquids and solids, since their expansion at constant pressure is often negligible, one finds $C_P \approx C_V$; this is certainly not the case for gases.

For any function $f(x, y)$ of two variables, the differential df is

$$df = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy \qquad \text{(Math identity)}$$

where the subscripts on the parentheses remind us that a variable is held fixed. Applying this identity to internal energy

$$dU(T, V) = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV \qquad (**)$$

If we hold volume fixed (i.e., set $dV = 0$) in this expression and in (*) then we obtain the result

$$dQ = \left(\frac{\partial U}{\partial T}\right)_V dT \qquad \text{(constant } V\text{)}$$

From the above,

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V$$

This result is easy to understand since the change in internal energy equals the heat added when volume is fixed since no work can be done if volume is fixed.

The heat capacity is important because given the equation of state and the heat capacity of a system, *all* other thermodynamic properties can be obtained. This is proved in most thermodynamics textbooks.

Joule's Free Expansion Experiment

Let's apply some of these results to a historically important experiment. Take a gas in an insulated container as shown in Fig. 1.3.

Experimentally, one finds that $T_B = T_A$, that is, the temperature remains constant after the gas expands to fill both sides of the container. If you were thinking that the gas would cool when it expanded, I'll explain the source of your confusion in a moment.

In the free-expansion, the gas does no work so, $\Delta W_{AB} = 0$. The system is insulated so no heat enters or leaves, thus $\Delta Q_{AB} = 0$. From first law, the internal energy U must remain constant.

From $(**)$ for constant U,

$$0 = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV \qquad (dU = 0)$$

Since the experiment shows that the temperature does not change on expansion, then $dT = 0$ and thus for an ideal gas,

$$\left(\frac{\partial U}{\partial V}\right)_T dV = 0 \qquad \text{(Ideal Gas)}$$

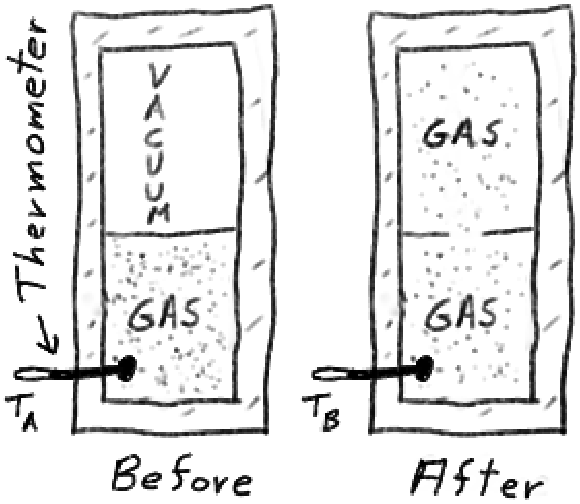

Figure 1.3: Joule free expansion experiment

This means that U does not depend on V so $U(V,T) = U(T)$ for a dilute gas. This defines what one means by an *ideal* gas.

The physical meaning of this result is that when the particles in an ideal gas move farther apart, the energy of the particles does not change. This means that there is zero potential energy associated with any interaction between the particles. In a real gas this is not exactly true since there is a weak interaction between the molecules. In an ideal gas we assume that the potential energy is negligible so all the energy is kinetic energy.

Since the heat capacity at constant volume is defined as

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V$$

for an ideal gas,

$$C_V = \frac{dU}{dT} \qquad \text{(Ideal gas)}$$

or

$$dU = C_V dT \qquad \text{(Ideal gas)}$$

If C_V constant then,

$$U(T) = C_V T + \text{constant} \qquad \text{(Ideal gas)}$$

If we say $U(T = 0) = 0$ then for an ideal gas $U(T) = C_V T$. Later we derive these results using statistical mechanics.

Note: When a gas expands in a piston, it does work and thus if the system is insulated, the gas cools (see Fig. 1.4). In this case $\Delta Q_{AB} = 0$ as before but $\Delta W_{AB} > 0$ so $U_B < U_A$ and thus $T_B < T_A$.

Example: For n moles of an ideal gas held at constant temperature T_0 find the heat added in going from volume V_A to V_B.

Solution: Since $U(T)$ for an ideal gas, if T is fixed then so is the internal energy. The first law says,

$$U_B - U_A = \Delta W_{AB} - \Delta Q_{AB}$$

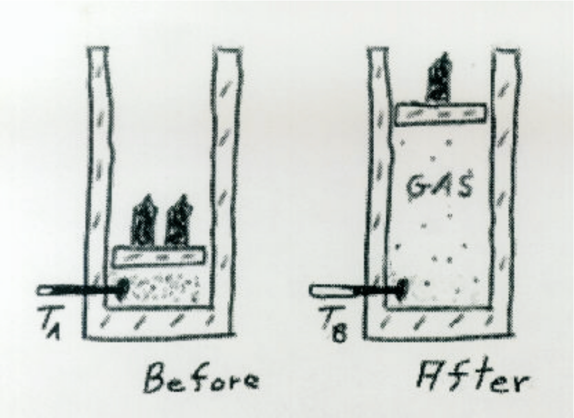

Figure 1.4: Illustration of an ideal gas cooling in an adiabatic expansion under pressure.

since $U_B = U_A$, then $\Delta Q_{AB} = \Delta W_{AB}$, so the heat added equals the work done for an isothermal path (see the previous example).

Entropy

Lecture 2

The first law of thermodynamics may be written as:

$$dU = dQ - dW$$

We have an expression for work in terms of thermodynamic parameters:

$$dW = PdV \qquad \text{or} \qquad \Delta W_{AB} = \int_{V_A}^{V_B} PdV$$

where ΔW_{AB} is the work done by the system in going from state (T_A, V_A) to (T_B, V_B) along some given path.

We now introduce a similar result for heat:

$$dQ = TdS \qquad \text{or} \qquad \Delta Q_{AB} = \int_{S_A}^{S_B} TdS$$

where ΔQ_{AB} is the heat added to the system. Unfortunately, the entropy, S, is not as easily measured as other thermodynamic variables. We can write:

$$dS = \frac{dQ}{T} \qquad \text{or} \qquad S_B - S_A = \int_{Q_A}^{Q_B} \frac{1}{T} dQ$$

The differential dS is exact (like dU and dV) so if we select a path between A and B such that $T = T_0$ is fixed, then:

$$S_B - S_A = \frac{1}{T_0} \int_A^B dQ = \frac{1}{T_0} \Delta Q_{AB}$$

By experimentally measuring ΔQ_{AB}, we obtain S.

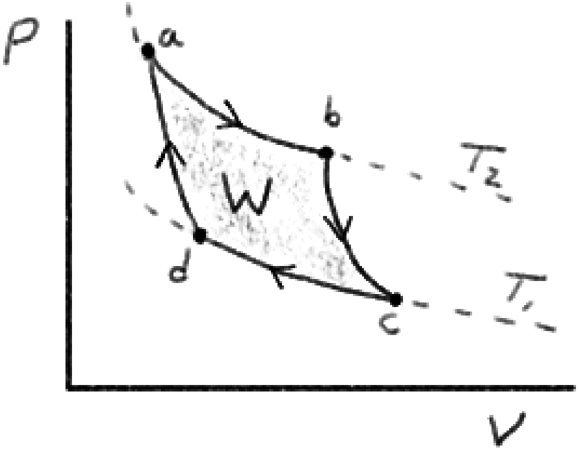

Figure 1.5: Carnot engine cycle in $P - V$ diagram

But what if we cannot get from A to B along an isothermal path? Along an adiabatic path (no heat added or removed) $dQ = 0$ so $dS = 0$. In general, we can build a path between any two points A and B by combining isothermal and adiabatic paths and thus calculate $S_B - S_A$.

The entropy is important in that given $S(U, V)$ for a physical system, *all* other thermodynamic properties may be computed. We derive and use this result when we develop statistical mechanics.

Third Law of Thermodynamics

The entropy is a function of the thermodynamic parameters so we may write it as $S(T, V)$. The third law says:

$$S(T = 0, V) = 0$$

that is, at zero Kelvin the entropy is zero. Thus, we can find the entropy for arbitrary T and V by computing:

$$S(T, V) = \int_{(T=0,V)}^{(T,V)} \frac{dQ}{T}$$

along any convenient path. There are many deep meanings to the third law and I encourage you to plunge in and read about them on your own.

Engines and Refrigerators

A heat engine is described by a closed cycle in a $P - V$ diagram for which the enclosed area (i.e., the work done) is positive (contour loop is clockwise). If the loop is counterclockwise we have a refrigerator.

Call the heat added in a cycle Q_+ and the work done W; we define engine efficiency as $\eta = W/Q_+$, that is, the fraction of the heat added that is converted into work.

An important example is a *Carnot engine*. In a Carnot engine the cycle consists of a pair of isothermal paths and a pair of adiabatic paths (see Fig. 1.5).

Notice that the Carnot cycle is a rectangle in the $T - S$ diagram (see Fig. 1.6). The area enclosed in the $P - V$ diagram is the work done. Since the cycle is a closed loop, our final and initial internal

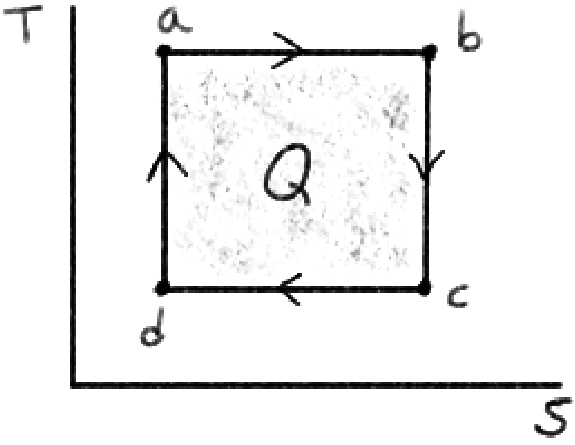

Figure 1.6: Carnot engine cycle in $T-S$ diagram

energies U are equal. By first law this means $W = Q$ where Q, the net heat added, is the area enclosed in the $T-S$ diagram.

On leg ab an amount of heat $Q_+ = T_+(S_b - S_a)$ is added to the system; this is the area under this leg in the $T-S$ diagram. On the leg, cd, an amount of heat $Q_- = T_-(S_b - S_a)$ is removed from the system. Since $Q = Q_+ - Q_- = W$ then

$$\eta = \frac{Q_+ - Q_-}{Q_+} = 1 - \frac{Q_-}{Q_+} = 1 - \frac{T_-}{T_+} \qquad \text{(Carnot Engine)}$$

The efficiency of a Carnot engine is given by this simple relation involving the ratio of temperatures. Using the fact that a Carnot engine is reversible you can prove that no engine can be more efficient than a Carnot engine.

The Carnot engine has important theoretical significance. Using the fact that $Q_-/Q_+ = T_-/T_+$ for a reversible path, Clausius established that dQ/T was an exact differential, which led to the definition of entropy. Lord Kelvin established the absolute temperature scale using the fact that the efficiency of a Carnot engine only depends on the temperatures of the reservoirs and not on the properties of the materials.

Second Law of Thermodynamics

Now that we have the connection between heat and entropy, namely $dQ = TdS$, we may write $dU = dQ - dW$ as

$$dU = TdS - PdV$$

where, as usual, we'll just consider the mechanical variables to be P and V. From this we can get a few useful relations, such as,

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

Note that since $(\frac{\partial S}{\partial U})_V = 1/T > 0$ the entropy increases with increasing energy, keeping volume fixed. But let's formulate a more fundamental result, namely the second law of thermodynamics.

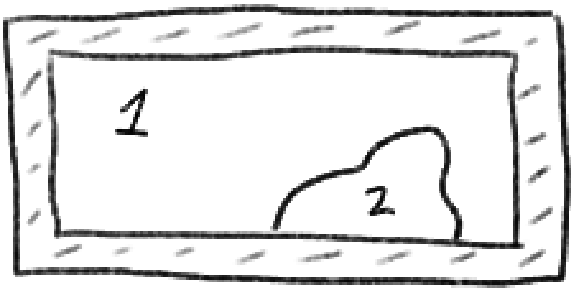

Figure 1.7: Isolated system divided into two parts

Consider an isolated system (no heat or work in or out) that we arbitrarily divide into two parts (see Fig. 1.7). Call the total volume $V = V_1 + V_2$. Similarly, the total internal energy is $U = U_1 + U_2$. These thermodynamic variables are said to be *extensive*. At thermal equilibrium $T_1 = T_2$, at mechanical equilibrium $P_1 = P_2$; temperature and pressure are *intrinsic* parameters.

Each subsystem has an entropy that depends on temperature and pressure; entropy is an extensive variable like energy. For example, given $S(E, V)$ then $S(cE, cV) = cS(E, V)$ where c is a constant; notice that if instead we had $S(E, P)$ then $S(cE, cP) \neq cS(E, P)$ since P is an intrinsic variable. Another extensive variable is N, the number of molecules in the system, so if we allow it to vary then we'd write $S(cE, cV, cN) = cS(E, V, N)$.

The second law states that the total entropy $S_1(T_1, P_1) + S_2(T_2, P_2)$ is maximum if and only if $T_1 = T_2$ and $P_1 = P_2$, that is, at equilibrium.

Proof: Start from:

$$dU = TdS - PdV$$

or

$$dS = \frac{1}{T}dU + \frac{P}{T}dV$$

Since total entropy is $S = S_1 + S_2$, then, $dS = dS_1 + dS_2$ and,

$$dS = \frac{1}{T_1}dU_1 + \frac{P_1}{V_1}dV_1 + \frac{1}{T_2}dU_2 + \frac{P_2}{T_2}dV_2$$

Total energy and volume are fixed so $dU_1 = -dU_2$ and $dV_1 = -dV_2$. Thus,

$$dS = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)dU_1 + \left(\frac{P_1}{T_1} - \frac{P_2}{T_2}\right)dV_1$$

Total entropy is extremum if and only if $dS = 0$ yet $dS = 0$ if and only if $T_1 = T_2$ and $P_1 = P_2$. With a little extra work, we may show that this extremum is a maximum.

We have two forms for the second law: Verbal form: The total entropy of an isolated system is maximum at equilibrium; Mathematical form: $dU = TdS - PdV$. Can derive one from the other.

Thermodynamic Manipulations

We often want to manipulate partial derivatives in thermodynamics calculations. Use the partial derivative rules:

$$\left(\frac{\partial x}{\partial y}\right)_w \left(\frac{\partial y}{\partial z}\right)_w = \left(\frac{\partial x}{\partial z}\right)_w \qquad \text{(D1)}$$

$$\left(\frac{\partial x}{\partial y}\right)_z = \frac{1}{\left(\frac{\partial y}{\partial x}\right)_z} \qquad \text{(D2)}$$

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1 \qquad \text{(D3)}$$

with D3 known as the chain relation; sometimes it is written as

$$\left(\frac{\partial x}{\partial z}\right)_y = -\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x$$

Finally, if $f(x, y)$,

$$\left(\frac{\partial}{\partial x}\right)_y \left(\frac{\partial f}{\partial y}\right)_x = \left(\frac{\partial}{\partial y}\right)_x \left(\frac{\partial f}{\partial x}\right)_y = \frac{\partial^2 f}{\partial x \partial y} \qquad \text{(D4)}$$

These are most of the math identities we need to manipulate partial derivatives.

We define the following useful thermodynamic quantities:

$$\alpha \equiv \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P \qquad \text{(coefficient of thermal expansion)}$$

$$K_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \qquad \text{(isothermal compressibility)}$$

Various relations among different quantities may be obtained, for example,

$$C_P - C_V = \frac{TV\alpha^2}{K_T}$$

which many thermodynamics texts derive.

Example: Determine $C_P - C_V$ for an ideal gas.
Solution: For an ideal gas, $PV = NkT$ or $V = \frac{NkT}{P}$ so:

$$\alpha = \frac{1}{V}\left(\frac{\partial}{\partial T}\right)_P \left(\frac{NkT}{P}\right) = \frac{1}{V}\frac{Nk}{P} = \frac{1}{T}$$

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

and thus,

$$C_P - C_V = \frac{TV(1/T)^2}{(1/P)} = \frac{PV}{T} = Nk$$

Notice that $C_P > C_V$ since heating a gas at constant pressure will cause the gas to expand and do work thus more heat energy is thus required for a given change in temperature.

Maxwell Relations

Lecture 3

Consider our expression for the second law:

$$dU = TdS - PdV$$

Since dU is an exact differential, we can use the mathematical identity for the differential of a function of two variables,

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

Comparing the two expressions above,

$$T = \left(\frac{\partial U}{\partial S}\right)_V$$

and

$$P = -\left(\frac{\partial U}{\partial V}\right)_S$$

which gives the formal definitions of temperature and pressure for a thermodynamic system.

Using these results along with the partial derivative identity (D4),

$$\frac{\partial^2 U}{\partial S \partial V} = \left(\frac{\partial}{\partial V}\right)_S \left(\frac{\partial U}{\partial S}\right)_V = \left(\frac{\partial}{\partial S}\right)_V \left(\frac{\partial U}{\partial V}\right)_S$$

so

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

or

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V \qquad \text{(M1)}$$

This identity is called a *Maxwell relation*; it is a relation between T, V, P, and S arising from the fact that dU is an exact differential.

There are 3 other Maxwell relations:

$$\begin{aligned} \left(\frac{\partial T}{\partial P}\right)_S &= \left(\frac{\partial V}{\partial S}\right)_P \qquad \text{(M2)} \\ \left(\frac{\partial S}{\partial V}\right)_T &= \left(\frac{\partial P}{\partial T}\right)_V \qquad \text{(M3)} \\ \left(\frac{\partial S}{\partial P}\right)_T &= -\left(\frac{\partial V}{\partial T}\right)_P \qquad \text{(M4)} \end{aligned}$$

Example: Express the rate of change of temperature when volume changes adiabatically in terms of α, K_T, and C_V.

Solution: The quantity we want to find is:

$$\begin{aligned} \left(\frac{\partial T}{\partial V}\right)_S &= -\left(\frac{\partial P}{\partial S}\right)_V \qquad \text{(using M1)} \\ &= -\left(\frac{\partial P}{\partial T}\right)_V \Big/ \left(\frac{\partial S}{\partial T}\right)_V \qquad \text{(using D1)} \end{aligned}$$

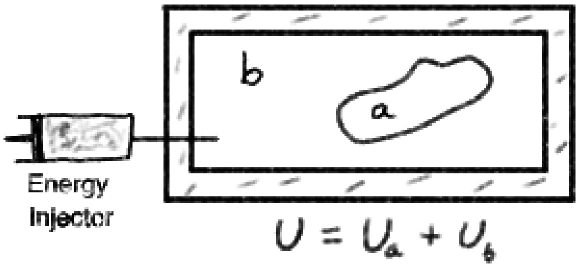

Figure 1.8: Isolated system into which energy is injected.

$$
\begin{aligned}
&= \frac{\left(\frac{\partial P}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_P}{\frac{1}{T}\left(T\left(\frac{\partial S}{\partial T}\right)_V\right)} && \text{(using D2 and D3)} \\
&= \frac{\left(\frac{-1}{VK_T}\right)(\alpha V)}{\frac{1}{T}C_V} && \text{(defn. of } \alpha, K_T\text{, and } C_V\text{)} \\
&= \frac{-\alpha T}{K_T C_V}
\end{aligned}
$$

Thermodynamics Potentials

From the second law, we know that for an isolated system (U and V fixed) the entropy S is maximum at equilibrium. We now establish a related result: for a system with fixed S and V, the internal energy is *minimum* at equilibrium.

Consider the closed system, arbitrarily divided into two parts, a and b; the total energy is $U = U_a + U_b$ (see Fig. 1.8).

Consider the diagram shown in Fig. 1.9 illustrating the possible values of energy for the two subsystems. The three diagonal lines indicate three values of total energy, $U_1 < U_2 < U_3$. In each case, at equilibrium the total energy is divided between the two subsystems a and b. At equilibrium, for a total system energy of U_1, the total system entropy is S_1 (and similarly for S_2 and S_3).

As U increases, S must increase since:

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T} > 0$$

so $S_1 < S_2 < S_3$.

Now follow the contour of constant entropy S_1. From the diagram, if the system has entropy S_1, (and V is fixed) then equilibrium occurs at the minimum energy U_1.

To summarize: For fixed U and V thermodynamic equilibrium implies maximum S. For fixed S and V thermodynamic equilibrium implies minimum U.

The internal energy is called a *thermodynamic potential* because it is minimum at equilibrium (for fixed S and V).

There are 3 other thermodynamic potentials:

$$
\begin{aligned}
H &= U + PV & dH &= TdS + VdP && \text{(Enthalpy)} \\
A &= U - TS & dA &= -SdT - PdV && \text{(Helmholtz free energy)} \\
G &= U + PV - TS & dG &= -SdT + VdP && \text{(Gibbs free energy)}
\end{aligned}
$$

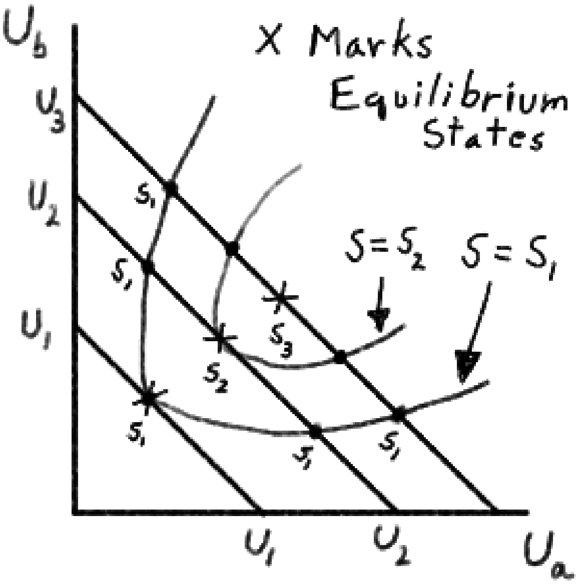

Figure 1.9: Energy graph with entropy contours ($S_2 > S_1$) and equilibrium points (marked by $\times$).

For fixed S and P, thermodynamic equilibrium implies minimum H. For fixed T and V, thermodynamic equilibrium implies minimum A. For fixed T and P, thermodynamic equilibrium implies minimum G. The Gibbs free energy is important to experimentalists since they often do their work at constant temperature and pressure.

Gibbs-Duhem Equation

One last general result that should be mentioned is Gibbs-Duhem equation,

$$SdT - VdP = 0 \qquad \text{(Gibbs-Duhem)}$$

We arrive at this result using the Euler theorem for extensive functions, which states that if f is an extensive function, that is, if $f(cx_1, \ldots, cx_n) = cf(x_1, \ldots, x_n)$ then

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)_{x_j} x_i$$

Since internal energy is an extensive function, $U(S, V)$, then

$$U = \left(\frac{\partial U}{\partial S} \right)_V S + \left(\frac{\partial U}{\partial V} \right)_S V = TS - PV$$

Finally, the differential of this expression is,

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

so by our result for the second law the term in the curly braces is zero.

But wait a minute! In the discussion of thermodynamic potentials we said that the Gibbs free energy was,

$$G = U + PV - TS \qquad dG = -SdT + VdP$$

Does the Gibbs-Duhem equation imply that $dG = 0$? No, because we've been neglecting the dependance of the energy on number of particles; instead of writing $U(S, V)$ it's more complete to write $U(S, V, N)$. The second law is then generalized to,

$$dU = TdS - PdV + \mu dN$$

where μ is the *chemical potential.* The full form of the Gibbs-Duhem equation is then

$$SdT - VdP + Nd\mu = 0 \qquad \text{(Gibbs-Duhem)}$$

so $dG = Nd\mu$, that is the Gibbs free energy is also the chemical potential per particle.

Van der Waals Equation of State (Pathria §12.2)

Real substances have complicated equations of state. Van der Waals introduced a simple model that extends the ideal gas law to dense gases and liquids.

The Van der Waals equation of state is:

$$\left(P + \frac{an^2}{V^2}\right)(V - bn) = nRT$$

where n is the number of moles and the positive constants a and b depend on the material. To understand the origin of these constants, let's consider them separately.

Take $a = 0$, then:

$$P = \frac{nRT}{(V - bn)}$$

Notice that $P \to \infty$ if $V \to bn$. There is a minimum volume to which we can compress the gas; for one mole, this volume is b so b/N_A is the volume of a single molecule. This term simulates the strong repulsion between atoms when they are brought close together (Coulomb repulsion of electron shells).

To understand the other constant, a, write the equation of state as,

$$P = \frac{nRT}{V - nb} - \frac{an^2}{V^2}$$

The larger the value of a, the smaller the pressure. This term represents the weak, attractive force between atoms in the gas. This binding force reduces the pressure required to contain the gas in a volume V. The smaller the volume, the closer the atoms get and the stronger this binding force becomes. However, if we further decrease V, the repulsive term, $(V - nb)$, kicks in and keeps the gas from collapsing.

Fig. 1.10 shows the isotherms on the $P-V$ diagram. The inflection point is called the *critical point.* Using the fact that $(\partial P/\partial V)_T = (\partial^2 P/\partial V^2)_T = 0$ at the critical point you can work out the critical values,

$$T_c = \frac{8a}{27bR} \qquad P_c = \frac{a}{27b^2} \qquad V_c = 3bn \qquad \frac{P_c V_c}{nRT_c} = \frac{3}{8}$$

For $T < T_c$ the equation of state has an S-shape.

The region where the slope $(\partial P/\partial V)_T$ is positive is thermodynamically impossible; this region is called the *spinoidal region.* The isothermal compressibility

$$K_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$$

of a substance must be positive or else second law is violated. To understand why, consider Fig. 1.11; weight is added to the piston and if $K_T < 0$ then the piston rises. Heat energy is spontaneously removed from the reservoir and converted into work, a violation of the second law of thermodynamics.

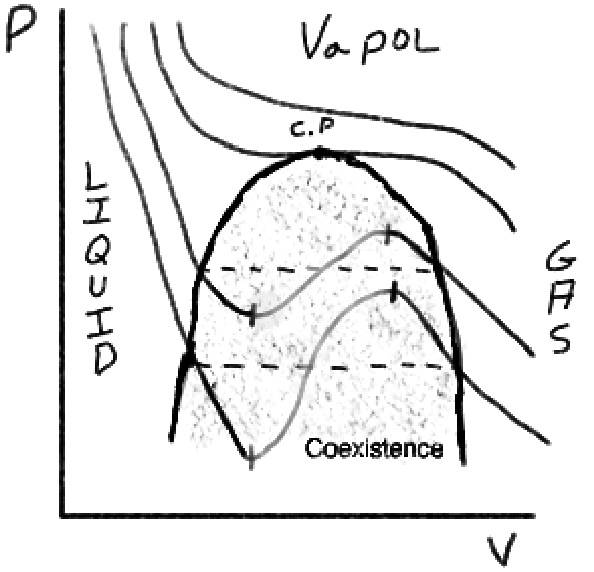

Figure 1.10: Phase diagram in $P-V$ plane for Van der Waals gas.

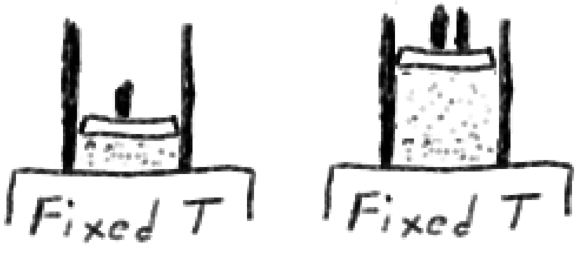

Figure 1.11: Illustration of work done when weight added if $K_T < 0$.

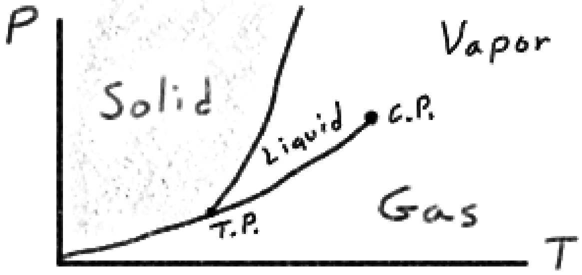

Figure 1.12: Phase diagram in $P-T$ plane.

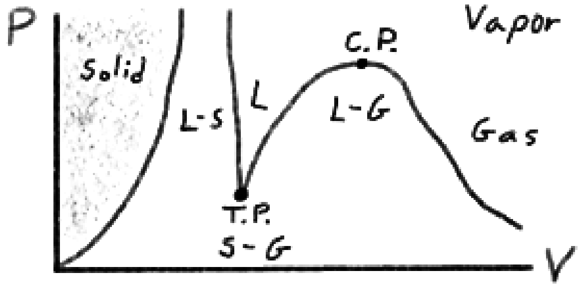

Figure 1.13: Phase diagram in $P-V$ plane.

Thermodynamics of Phase Transitions

The equation of state for a real substance is a complicated function, especially since we have phase transitions. Since the EOS is difficult to draw in 3D, we often look at 2D projections (see Fig. 1.12 and 1.13).

What makes the van der Waals model so interesting is that the critical point in the model is liquid-gas critical point, the point with the highest temperature at which liquid and gas can coexist.

Given that the spinoidal bend in the van der Waals equation of state is forbidden, you should be curious as to what replaces it. Specifically, if we travel along an isotherm by increasing the pressure to take the system from the gas phase (large V) to the liquid phase (small V), what does this isotherm look like?

To answer this question, consider the Gibbs free energy,

$$dG = -SdT + VdP$$

Say that starting from P_A, we move along the isotherm of temperature $T_0 < T_c$. Along an isotherm $dT = 0$ so

$$G(P, T_0) = \int_{P_A}^{P} VdP + C$$

where the constant of integration C can be a function of T_0 and P_A.

By rotating the $P-V$ diagram to a $V-P$ diagram (see Fig. 1.15, the qualitative shape of the integral $\int V\,dP$ can be sketched to give the Gibbs free energy shown in Fig. 1.16. Finally, the Gibbs free energy is a thermodynamic potential that is minimum at equilibrium for fixed P and T. This means that in traveling from point A to point H the pressure at points B through F must be equal.

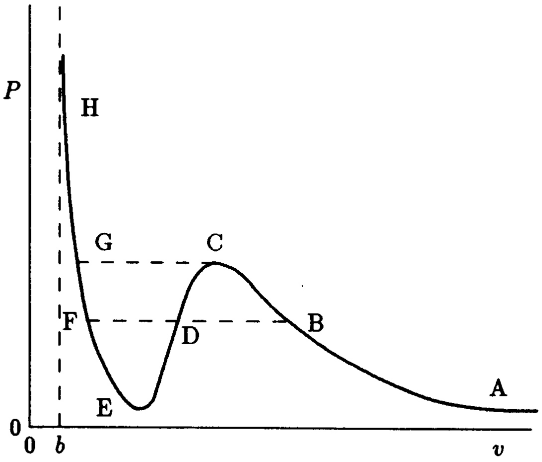

Figure 1.14: Isotherm in $P - V$ plane for Van der Waals fluid.

The gas-liquid coexistence lines are indicated as dashed lines in Fig. 1.17. The requirement that the Gibbs free energy be minimum can be shown to give the *Maxwell construction*, which states that the S-shaped areas above and below the coexistence line must be equal.

As the pressure is dropped, we pass from liquid to gas. If we do the experiment very carefully, we can "miss" the coexistence curve and travel on the metastable branch (see Fig. 1.18). While this alternative branch is experimentally attainable the system is unstable. Kick the lab table and it will spontaneously jump to the $L - G$ branch.

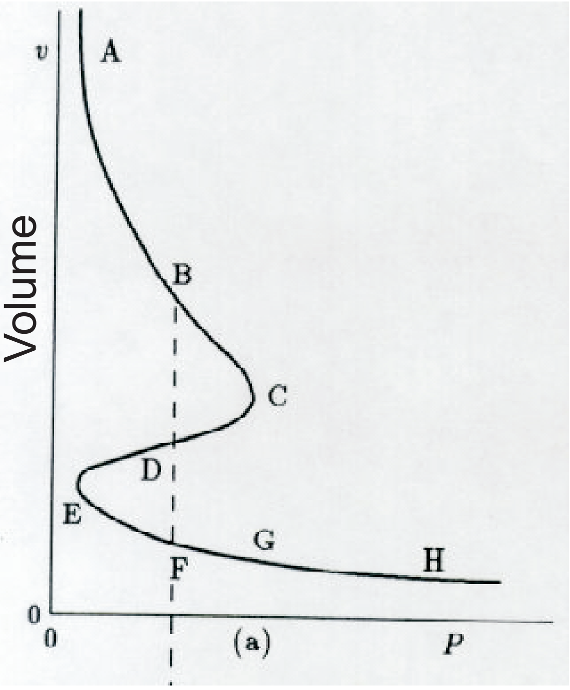

Figure 1.15: Isotherm in $V-P$ plane for Van der Waals fluid. Compare with Fig. 1.14

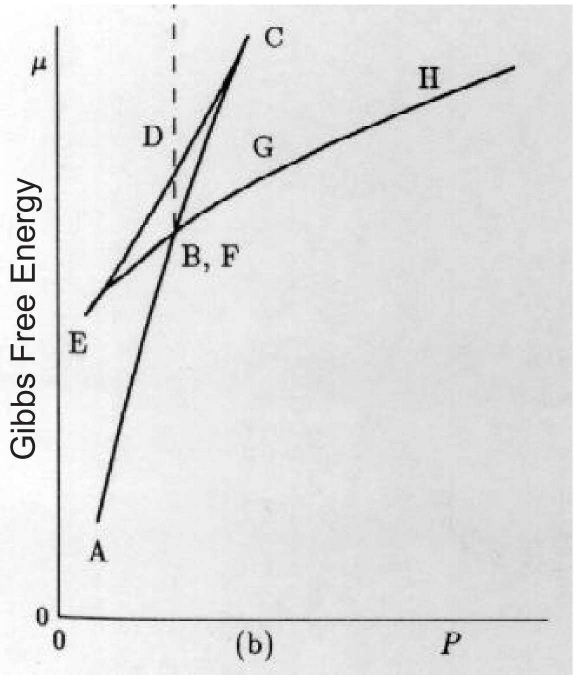

Figure 1.16: Gibbs free energy versus pressure for the isotherm shown in Fig. 1.15.

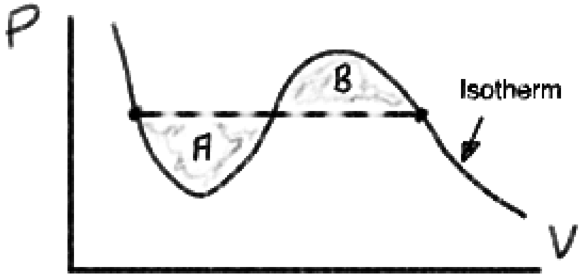

Figure 1.17: Maxwell construction for coexistence isotherm of Van der Waals gas.

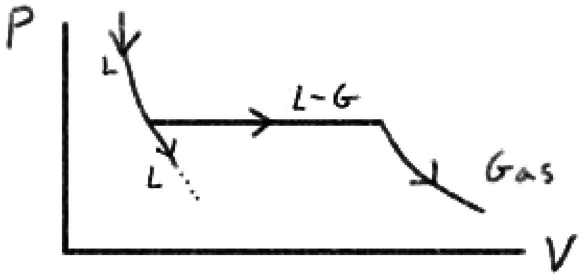

Figure 1.18: Superheated liquid or supercooled gas.

Chapter 2

Ensemble Theory I

Ergodic Hypothesis (§2.1 – 2.2)

Lecture 4

Since this is a statistical mechanics class, it is time that we end our brief review of thermodynamics and move on to statistical mechanics. Pathria and Beale give an introduction to ensemble theory in their Chapter 1 then give a more formal exposition in Chapter 2. In these notes the coverage is slightly different, just to give you an alternative way to understand how it all fits together.

Given certain information about a substance (e.g., equation of state, heat capacity), thermodynamics allows us to derive various other quantities. What is missing is a way to bridge from the mechanical (classical or quantum) description of atomic interactions to the macroscopic level of thermodynamics. Statistical mechanics serves as the bridge.

We first consider the formulation of statistical mechanics using classical mechanics. Since Gibbs, Maxwell, and Boltzmann established the theory with no knowledge of quantum physics we retrace their steps. Some systems (e.g., dilute gas) are well described by classical mechanics. Surprisingly we later find that the quantum systems are not much harder to do. In fact Pathria and Beale start with quantum systems, where states are easier to count, and then formulate ensemble theory for classical system in their Chapter 2.

From Hamiltonian dynamics, we know that a system is completely described by its canonical coordinates and momenta. For an N-particle system, we write the coordinates $q_1, q_2, ..., q_{3N}$ (in 3D each particle has 3 coordinates, such as x, y, and z). Similarly the momenta are $p_1, p_2, ..., p_{3N}$.

The instantaneous state of the system is given by:

$$(q, p) \equiv (q_1, ..., q_{3N}, p_1, ..., p_{3N})$$

The system is not frozen in this state; particles move and interact and thus the q's and p's change with time. The Hamiltonian may be used to compute the dynamics.

At any instant in time, the values (q, p) may be viewed as a point in $6N$ dimensional space (hard to draw). In time, this point moves around in this *phase space*, the trajectory completely describes the evolution of the system (see Fig. 2.1).

It is neither possible nor desirable to compute this trajectory for real systems (in which $N \approx 10^{23}$).

The trajectory cannot wander arbitrarily through phase space. For example, if our system is isolated (so U and V are fixed) then only points that satisfy these constraints can be reached.

In Fig. 2.2 suppose that only points within the shaded region are permitted by the constraints on the system.

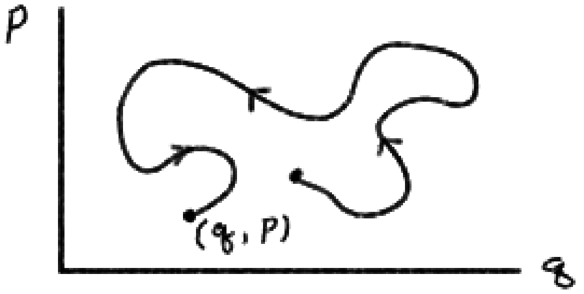

Figure 2.1: Point in phase space evolving in time.

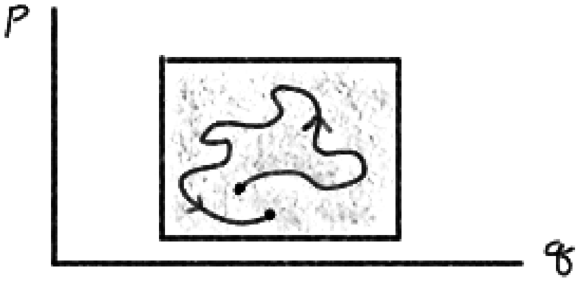

Figure 2.2: Point in phase space constrained to remain within restricted region.

We now introduce the fundamental postulate of equilibrium statistical mechanics:

> An isolated system at equilibrium is *equally likely* to be found in any accessible state, that is, all points in the allowed region of phase space are equally probable. This is called the *ergodic hypothesis.*

This is a *big* assumption. It is very difficult to prove that a system will even *visit* every allowed point much less prove that all are equally probable. In the absence of a mathematical proof, we use the postulate of equal probabilities as a working hypothesis (that happens to be well supported by experiments and numerical simulations).

Ensemble Formulation (§2.1 – 2.3)

Mathematically, there are various ways to proceed. Pathria and Beale define a phase space density $\rho(q, p)$ for points. I prefer to think in terms of probability, call

$$\mathcal{P}(q, p) \equiv \text{ Probability that the system is in state } (q, p)$$

Since all states are equally probable,

$$\mathcal{P}(q, p) = \frac{1}{\text{Total number of allowed states}} = \frac{1}{\Gamma}$$

Mathematically, $\mathcal{P}(q, p)$ and $\rho(q, p)$ differ only by a constant of normalization.*

---

*NOTATION: Pathria and Beale (pg. 2) calls the total number of allowed states Ω instead of Γ.

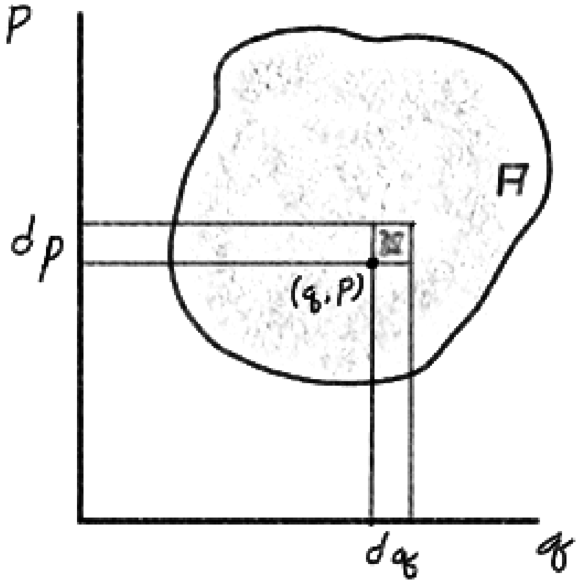

Figure 2.3: Infinitesimal element in phase space.

One technical problem, a continuous space has an infinity of points so “total number” is an ill-defined concept.

Two equally valid solutions to this technical difficulty:

1) Continuous Probability Distribution

Define $\mathcal{P}(q,p)\,dqdp$ to be the probability that the system is in a state within the infinitesimal volume between (q,p) and $(q+dq,p+dp)$ (see the Fig. 2.3).

Our equal probability hypothesis gives:

$$\mathcal{P}(q,p)\,dqdp = \frac{dqdp}{\int_A dqdp}$$

where the integral is over the allowed region in phase space (given the constraints that energy, U, and volume, V, are fixed).

Since all points in the allowed region are equally probable, the probability of the system being a given sector of phase space is just the fractional volume of phase space occupied by that sector.

2) Coarse Graining

Suppose we partition the real estate in phase space into finite lots of volume h^{3N}, as shown in Fig. 2.4.

The function $\mathcal{P}(q,p)$ is the probability that the system is in lot (q,p). The equal probability hypothesis now becomes:

$$\begin{aligned}\mathcal{P}(p,q) &= \frac{1}{\text{Number of lots inside } A}\\ &= \frac{1}{\Gamma}\end{aligned}$$

While this coarse graining of phase space into parcels might seem ad hoc, it is a useful construction. Specifically, it is closer to what we find in quantum mechanics—there is a discrete set of allowed states.

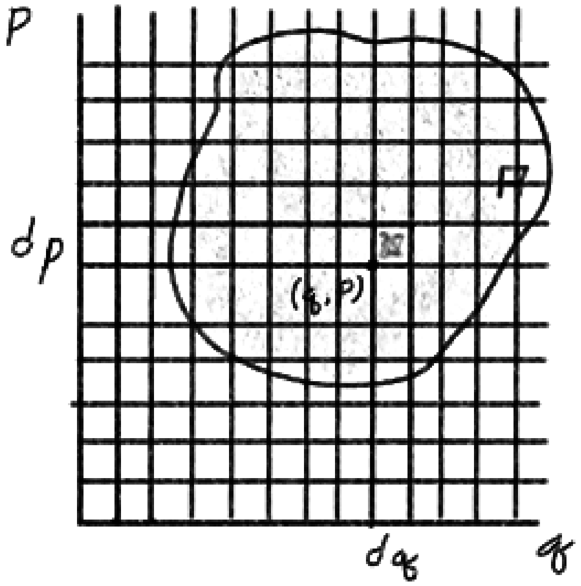

Figure 2.4: Coarse graining of phase space into discrete cells.

By the uncertainty principle:

$$(\Delta q_1 \Delta p_1)(\Delta q_2 \Delta p_2)\ldots(\Delta q_{3N} \Delta p_{3N}) \approx h^{3N}$$

so the factor h is taken to be Planck's constant.

Given the allowed slop in the above formulation, you should realize that the details won't matter in the end.

Microcanonical Ensemble (§1.1, 1.2, 2.3)

Now that we know the probability for a system to be in a given state, we could work out average values, such as average pressure, etc.. Since all allowed states are equally probable, we could just (just, HA!) find the pressure for every possible state and average all the values together.

A more efficient approach is to compute entropy and then use thermodynamics to find any other quantity of interest. Since the entropy is intimately connected to the number of states, this approach will not be as difficult as it sounds.

We first consider isolated systems with N particles for which internal energy U and volume V are fixed. We want to get an expression for $S(U, V, N)$.

Various ways we can define entropy; any definition that satisfies the laws of thermodynamics is valid. The most useful definition is,

$$S = -k \int_A dq dp \, \mathcal{P}(q, p) \ln \mathcal{P}(q, p)$$

where the integrals are over all allowed states and k is Boltzmann's constant. If we think of dividing phase space into a discrete set of states then we define entropy as,

$$S = -k \sum^{\Gamma} \mathcal{P}(q, p) \ln \mathcal{P}(q, p)$$

Probability theory tells us this expression for S gives the uncertainty (or disorder) in a system. You can prove that it possesses all the desired features for entropy, for example, it is maximum when all states are equally probable thus S is maximum at equilibrium.

Example: Suppose that a system has only 2 states, a and b. Show that the entropy is maximum when $\mathcal{P}_a = \mathcal{P}_b = 1/2$.

Solution: Our definition for entropy is:

$$\begin{aligned} S &= -k\sum^{\Gamma} \mathcal{P}\ln\mathcal{P} \\ &= -k(\mathcal{P}_a\ln\mathcal{P}_a + \mathcal{P}_b\ln\mathcal{P}_b) \end{aligned}$$

But probabilities must sum to one so $\mathcal{P}_a + \mathcal{P}_b = 1$ or $\mathcal{P}_b = 1 - \mathcal{P}_a$ so:

$$S = -k(\mathcal{P}_a\ln\mathcal{P}_a + (1-\mathcal{P}_a)\ln(1-\mathcal{P}_a))$$

To find the maximum, we take the derivative

$$\frac{\partial S}{\partial \mathcal{P}_a} = -k\left(\ln\mathcal{P}_a + \mathcal{P}_a\left(\frac{1}{\mathcal{P}_a}\right) - \ln\left(1-\mathcal{P}_a\right) + \left(1-\mathcal{P}_a\right)\left(\frac{-1}{1-\mathcal{P}_a}\right)\right)$$

so

$$0 = -k(\ln\mathcal{P}_a - \ln(1-\mathcal{P}_a))$$

or

$$\mathcal{P}_a = 1 - \mathcal{P}_a$$

so $\mathcal{P}_a = 1/2$. You can check that this extremum is a maximum.

For an isolated system, the equal probability hypothesis gives the probability $\mathcal{P}(q,p) = 1/\Gamma$ so:

$$\begin{aligned} S &= -k\sum^{\Gamma} \mathcal{P}(q,p)\ln\mathcal{P}(q,p) \\ &= -k\sum^{\Gamma}\left(\frac{1}{\Gamma}\right)\ln\left(\frac{1}{\Gamma}\right) \\ &= +k\left(\frac{1}{\Gamma}\right)\ln\left(\Gamma\right)\left(\sum^{\Gamma}1\right) \\ &= k\left(\frac{1}{\Gamma}\right)\ln\left(\Gamma\right)\left(\Gamma\right) \\ &= k\ln\Gamma \end{aligned}$$

Aside from the constant k, the entropy is the logarithm of the number of accessible states.

Entropy and Thermodynamics (§1.3, 2.3)

We just discussed how, for a system with internal energy, U, and volume, V, the equal probability hypothesis gave us an entropy

$$S = k\ln\Gamma$$

where Γ is the number of states that the system can be in. The set of allowed states for fixed U and V is called the *microcanonical ensemble.*

Pathria and Beale point out that it doesn't matter if we use

$$\text{Number of states such that} \begin{cases} U = E \\ U \leq E \leq U + \Delta U \\ E \leq U \end{cases}$$

where E is the energy of an individual state. That is, we can count just states that have exactly energy U, or count all states with energy in the range U and $U + \Delta U$ or even counting all states with energy less than or equal to U. This will be useful because sometimes one set of states is easier to count that another, as you will see in later examples.

Given the number of states in the microcanonical ensemble $\Gamma(U,V)$ we have the following recipe:

$$\begin{array}{ll} \text{1) Evaluate} & S(U,V) = k \ln \Gamma(U,V) \\ \text{2) Given} & S(U,V) \text{ solve for } U(S,V) \\ \text{3) Evaluate} & \\ & T = \left(\frac{\partial}{\partial S}\right)_V U(S,V) \\ & P = -\left(\frac{\partial}{\partial V}\right)_S U(S,V) \end{array}$$

The third step comes from using

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

Finally, use conventional thermodynamics to find any other desired quantities (e.g., heat capacity, compressibility).

Notice that temperature now has a mechanical definition since

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V = \left(\frac{\partial}{\partial U}\right)_V (k \ln \Gamma(U,V))$$

so

$$\frac{1}{T} = k \frac{1}{\Gamma} \left(\frac{\partial \Gamma}{\partial U}\right)_V$$

or

$$T = \frac{1}{k} \frac{\Gamma(U,V)}{\left(\frac{\partial \Gamma}{\partial U}\right)_V}$$

The temperature is inversely proportional to the rate at which the number of states increases as internal energy increases. Of course it is difficult find a "number of allowed states" meter in the lab. Simpler to calibrate thermometers that use properties, such as electrical conductivity, that vary with temperature and are easy to measure mechanically.

For most physical systems, the number of accessible states increases as energy increases so $(\partial\Gamma/\partial U) > 0$. There are exceptional systems for which the energy that can be added to the system has some maximum saturation value. For these systems, one can have $(\partial\Gamma/\partial U) < 0$ giving *negative temperatures.* Note that negative temperatures occur not when a system has very little energy but rather when it has so much energy that it is running out of available states. You'll do an example of such a system as a homework problem.

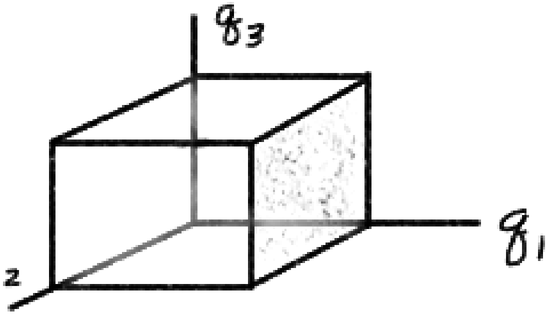

Figure 2.5: Restricted region of phase space in coordinates.

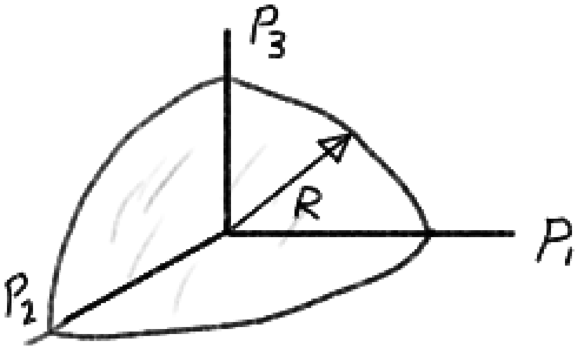

Figure 2.6: Restricted region of phase space in momentum.

Classical Ideal Gas – Microcanonical Ensemble (§2.4)

Lecture 5

We now calculate a specific example, an ideal monatomic gas using the microcanonical ensemble. Take N particles (of mass m) in a volume, V. The total energy is just the sum of the individual kinetic energies.

$$\begin{aligned} U &= \frac{1}{2}m\sum_{i=1}^{N}(v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2) \\ &= \frac{1}{2m}\sum_{j=1}^{3N}p_j^2 \end{aligned}$$

For just one particle (with position q_1, q_2, and q_3 in the x,y, and z directions) we can sketch the allowed region of coordinate phase space (see Fig. 2.5).
The particle can be anywhere *inside* this rectangle of volume $L_xL_yL_z = V$.

For just one particle (with momenta p_1, p_2, and p_3) we can sketch the accessible area in momentum phase space (see Fig. 2.6).

The allowed states are all points on the *surface* of this sphere with radius $R = \sqrt{2mU}$. The surface area of a 3D sphere is $4\pi R^2$.

If we have N particles instead of just one, the box in coordinate phase space is a $3N$ dimensional box with volume

$$\underbrace{(L_xL_yL_z) \times (L_xL_yL_z) \times \ldots \times (L_xL_yL_z)}_{3N \text{ terms}} = V^N$$

Similarly, for N particles instead of one, the accessible states in momentum space are the points on the surface of a $3N$ dimensional sphere. The surface area of a $3N$ dimensional sphere is $B_N R^{3N-1}$ where B_N is a geometric constant (see Appendix C in Pathria and Beale). The number of accessible states is thus,

$$B_N R^{3N-1} = B_N \left(\sqrt{2mU}\right)^{3N-1} \approx B_N \left(2mU\right)^{3N/2}$$

Since $N \gg 1$, we can set $3N - 1 \approx 3N$.

Collecting the above results, the total number of accessible states is the product of the number of states in coordinate space times the number of states in momentum space, so

$$\begin{aligned} \Gamma(U,V) &= \left(V^N\right) \times \left(B_N (2mU)^{3N/2}\right) \\ &= B_N (V(2mU)^{3/2})^N \end{aligned}$$

We now follow the steps in our recipe; first get the entropy,

$$\begin{aligned} S &= k \ln \Gamma(U,V) \\ &= kN \ln(V(2mU)^{3/2}) + k \ln B_N \end{aligned}$$

The temperature in an ideal gas is found by solving

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial U}\right)_V \\ &= kN \frac{1}{V(2mU)^{3/2}} \frac{3}{2} V (2m)^{3/2} U^{1/2} \\ &= \frac{\frac{3}{2}kN}{U} \end{aligned}$$

This relation between energy and temperature is usually written as

$$U = \frac{3}{2} kNT \qquad \text{(Monatomic ideal gas)}$$

This confirms the result found in the Joule free expansion experiment—the internal energy depends only on the temperature and not on volume.

To get the pressure we could use $P = -\left(\frac{\partial U}{\partial V}\right)_S$. Instead let's use

$$\begin{aligned} \frac{P}{T} &= \left(\frac{\partial S}{\partial V}\right)_U \\ &= \left(\frac{\partial}{\partial V}\right)_U \left[kN \ln(V(2mU)^{3/2}) + k \ln B_N\right] \\ &= kN \frac{1}{V(2mU)^{3/2}} \left(2mU\right)^{3/2} \\ &= \frac{kN}{V} \end{aligned}$$

or $PV = NkT$, the ideal gas law.

Our derivation was for a monatomic dilute gas such as helium or argon. Most dilute gases are *not* monatomic, for example nitrogen and oxygen form diatomic molecules.

However, for dilute gases, the energy of the molecules is independent of the molecules' positions. This means that, in general, for a gas,

$$\Gamma(U,V) = V^N f(U)$$

where f is some undetermined function. Simply put, the accessible area in coordinate phase space is always proportional to V^N where N is the number of molecules.

The pressure, in general, is then

$$\begin{aligned}\frac{P}{T} &= \left(\frac{\partial S}{\partial V}\right)_U = \left(\frac{\partial}{\partial V}\right)_U \left(k\ln(V^N f(U))\right) \\ &= \frac{kN}{V}\end{aligned}$$

Thus the ideal gas law does not depend on the molecular composition of the gas.

On the other hand, for a diatomic dilute gas,

$$U = \frac{5}{2}NkT \qquad \text{(Diatonic ideal gas)}$$

We derive this result later (can get it using microcanonical ensemble but derivation is clumsy).

Microcanonical Ensemble in Quantum Mechanics (§1.1, 1.2)

We will study three types of ensembles, the microcanonical (fixed U,V, and N), the canonical ensemble (fixed T,V, and N) and the grand canonical ensemble (fixed T,V, and Gibbs free energy G). For each ensemble we consider both classical and quantum mechanical formulations. As Ronald Reagan once said, “You’ve seen one redwood you’ve seen them all” and it’s almost the same with ensembles. Once you get the hang of it, the various types of ensembles are not so difficult to use.

So far we’ve considered the microcanonical ensemble for classical systems; let’s consider the quantum mechanical case.

Consider a quantum mechanical system that can be in any one of a set of discrete states. By the postulate of equal probabilities, we say that all states for which the energy is U and the volume is V are equally probable. The entropy is

$$S(U, V) = k \ln \Gamma(U, V)$$

where $\Gamma(U, V)$ is the number of allowed states. In a sense, the quantum mechanical case simplifies matters since we have discrete states to count.

Combinatorics

Counting can be tricky; here’s a quick review: The number of ways of making M selections from a population of N individuals is

Distinguishable Selections

$${}_NP_M = \frac{N!}{(N-M)!} \qquad \text{(Permutations)}$$

Indistinguishable Selections

$${}_NC_M = \frac{N!}{(N-M)!M!} = \binom{N}{M} \qquad \text{(Combinations)}$$

where the factorial is defined as $N! = N \times (N-1) \times \ldots \times 3 \times 2 \times 1$.

Example: Consider a class with $N = 10$ persons (labelled $A, B, ..., J$). In a beauty contest, find the number of ways to award 1st, 2nd, and 3rd prize.

Solution: Since the selections are distinguishable (1st prize is different from 2nd prize) making the list of possible selections we have,

$$\begin{array}{c} ABC \\ ACB \\ ABD \\ \vdots \end{array}$$

The number of permutations is

$$\frac{10!}{(10-3)!} = \frac{10!}{7!} = 10 \cdot 9 \cdot 8 = 720$$

Example (cont.): If in a class of 10 students three students flunk the class, find the number of ways of selecting them.

Solution: In this case the selections are indistinguishable so the list of possible selections is

$$\begin{array}{c} ABC \\ ABD \\ \vdots \end{array}$$

The number of combinations is $\frac{10!}{7!3!} = 120$.

Finally, the number of ways of putting M indistinguishable objects into N distinguishable boxes is

$$\binom{N+M-1}{M} = \binom{N+M-1}{N-1} = \frac{(N+M-1)!}{(N-1)!M!}$$

Quantum Harmonic Oscillators (§3.8)

Lecture 6

Our first quantum mechanics example is a system composed of a collection of N uncoupled harmonic oscillators. The total energy U is fixed; this system has no volume (or pressure) so the number of states Γ only depends on U. The oscillators are distinguishable.

The energy levels for an individual oscillator are $\epsilon(n) = \hbar\omega(n+1/2)$, $n = 0, 1, 2, \ldots$. Call n_i the energy level for oscillator i and

$$M = \sum_{i=1}^{N} n_i$$

the total number of quanta of energy in the system. Figure 2.7 shows some states for $M = 3$, $N = 4$.

The total energy in the system is

$$U = \sum_{i=1}^{N} \epsilon_i(n_i) = \sum_{i=1}^{N} \hbar\omega\,(n_i + 1/2) = \hbar\omega\,(M + N/2)$$

Notice that fixing U is equivalent to fixing M.

Now we need to count the number of possible states given M and N. This is equivalent to the number of ways that M indistinguishable objects may be placed into N distinguishable boxes, the number of states is

$$\Gamma(U) = \binom{N+M-1}{M} = \frac{(M+N-1)!}{(N-1)!M!}$$

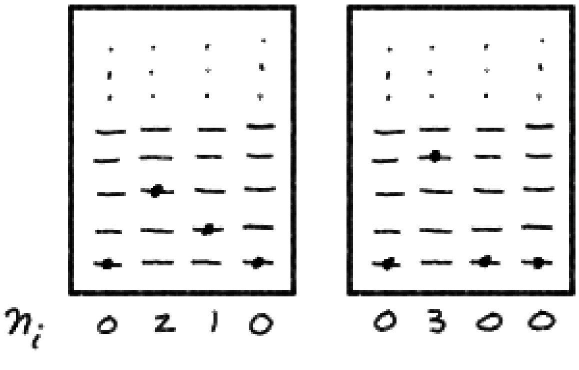

Figure 2.7: Two possible quantum harmonic oscillator system states for $M = 3$, $N = 4$.

where $M = U/\hbar\omega - N/2$.

The entropy in the system is

$$S(U) = k \ln \Gamma \approx k \left[(M + N) \ln (M + N) - M \ln M - N \ln N \right]$$

where we've used the approximation $\ln(x!) \approx x \ln(x) - x$ for $x \gg 1$. We can now proceed to compute other thermodynamic quantities. For example,

$$\frac{1}{T} = \frac{dS}{dU} = \frac{dS}{dM}\frac{dM}{dU} = \left(k \ln(M + N) - k \ln M\right)\left(\frac{1}{\hbar\omega}\right)$$

which gives us the following relation between M and T,

$$M = \frac{N}{e^{\hbar\omega/kT} - 1}$$

so

$$U = \frac{\hbar\omega N}{e^{\hbar\omega/kT} - 1} + \frac{1}{2}\hbar\omega N$$

Notice that $M \to 0$, $U \to \frac{1}{2}\hbar\omega N$ as $T \to 0$ and $M \to \infty$, $U \to \infty$ as $T \to \infty$. This is the Einstein model for a solid. Though it gives the right qualitative behavior and is better than a classical description of a solid this model is superseded by the Debye model.

Example: Find the heat capacity for a system of N quantum mechanical harmonic oscillators.

Solution: Since the system cannot do mechanical work, $C_P = C_V = C$. In other words, the internal energy is only a function of temperature so,

$$\begin{aligned} C = \frac{dU}{dT} &= \frac{dU}{dM}\frac{dM}{dT} \\ &= (\hbar\omega)\frac{-N}{\left(e^{\hbar\omega/kT} - 1\right)^2}\frac{-\hbar\omega}{kT^2} \\ &= \left(\frac{\hbar\omega}{kT}\right)^2 \frac{kN}{\left(e^{\hbar\omega/kT} - 1\right)^2} \end{aligned}$$

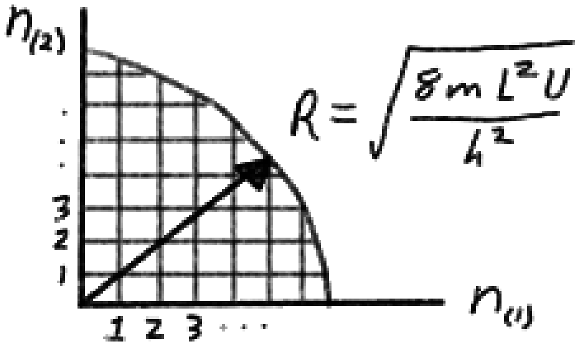

Figure 2.8: Free particle states with energy $E \leq U$.

Monatomic Ideal Gas (§1.4)

We previously studied this system using classical mechanics, we now consider it using quantum mechanics.

First, consider a single particle of mass m in a cubical box of side L, so $V = L^3$. The energy levels for a particle in a 3D box are

$$\epsilon(n_x, n_y, n_z) = \frac{h^2}{8mL^2}\left(n_x^2 + n_y^2 + n_z^2\right)$$

where $n_x, n_y, n_z = 1, 2, \ldots$ are the quantum numbers.

Now suppose we have N identical, but distinguishable, particles; particle i has energy $\epsilon_i(n_{x,(i)}, n_{y,(i)}, n_{z,(i)})$. The particles do not interact (i.e., ideal gas) so the total energy is

$$E = \sum_{i=1}^{N} \epsilon_i = \frac{h^2}{8mL^2} \sum_{i=1}^{N} \left[\left(n_{x,(i)}\right)^2 + \left(n_{y,(i)}\right)^2 + \left(n_{z,(i)}\right)^2 \right]$$

or

$$E = \frac{h^2}{8mL^2} \sum_{r=1}^{3N} n_{(r)}^2$$

where the index r runs over all $3N$ quantum numbers.

Now we have to count the number of possible combinations of quantum numbers that give an energy $E = U$. An easier job is to count the number of combinations that have a total energy $E \leq U$; we'll call Σ the total number of states such that $E \leq U$.

Our constraint may be written as

$$\frac{8mL^2U}{h^2} \geq \sum_{r=1}^{3N} n_{(r)}^2$$

Figure 2.8 illustrates this constraint for just two quantum numbers. The number of combinations for n's that satisfies the constraint is approximately the number of unit squares inside the circular wedge with radius R. But this is nothing more than the volume of the wedge. Notice that we have a wedge instead of a sphere since $n_{(r)} > 0$ so we only consider the wedge in the first quadrant.

The volume of a $3N$ dimensional sphere of radius R is

$$\mathcal{V}_{3N}(R) = \frac{\pi^{3N/2}}{(3N/2)!} R^{3N}$$

The volume of the first quadrant wedge is

$$\left(\frac{1}{2}\right)^{3N}\left(\frac{\pi^{3N/2}}{(3N/2)!}R^{3N}\right)$$

The total number of states with energy $E \leq U$ is thus

$$\Sigma(U,V) = \left(\frac{1}{2}\right)^{3N}\frac{\pi^{3N/2}}{(3N/2)!}\left(\frac{8mL^2U}{h^2}\right)^{3N/2}$$

or

$$\Sigma(U,V) = \left(\frac{V}{h^3}\right)^{N}\frac{(2\pi mU)^{3N/2}}{(3N/2)!}$$

The entropy is $S(U,V) = k\ln\Sigma(U,V)$ so

$$\begin{aligned} S(U,V) &\approx kN\ln\left[\frac{V}{h^3}\left(\frac{4\pi mU}{3N}\right)^{3/2}\right] + \frac{3}{2}N \\ &= kN\ln V + \frac{3}{2}kN\ln U + f(N) \end{aligned}$$

where we used Stirling's approximation $\ln(x!) \approx x\ln x - x$. From this we get

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V = \frac{\frac{3}{2}kN}{U}$$

so $U = \frac{3}{2}kNT$ and

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_U = \frac{kN}{V}$$

so $PV = NkT$.

You can compare this result with our previous expression for $S(U,V)$ obtained from classical mechanics; aside from a few unimportant constants, they are the same. From the above expression, you can derive the ideal gas law, $PV = NkT$ and the energy-temperature relation $U = \frac{3}{2}kT$ as we did in the classical case.

Finally, there is an interesting result that connects Σ, the number of states with energy $E \leq U$ and Γ, the number states with energy $E = U$. For a d-dimensional sphere of radius R, the fraction of the volume which lies within the outer shell of thickness ΔR,

$$\begin{aligned} \frac{\mathcal{V}_d(R) - \mathcal{V}_d(R-\Delta R)}{\mathcal{V}_d(R)} &= 1 - \frac{\mathcal{V}_d(R-\Delta R)}{\mathcal{V}_d(R)} \\ &= 1 - \frac{(R-\Delta R)^d}{R^d} \\ &= 1 - \left(1 - \frac{\Delta R}{R}\right)^d \\ &= 1 - \exp(-d\Delta R/R) \qquad (d >> 1) \end{aligned}$$

If d is very large then virtually all points in the sphere are infinitesimally close to the surface. For example, if $d = 1000$, then 90% of the volume is inside the outer shell of thickness $\Delta R \approx R/400$. Imagine if $d \approx 10^{23}$, how dramatic this effect becomes. For this reason, as $N \to \infty$, the two ways to count states become equivalent since $\Sigma \approx \Gamma$. In most physical systems, the number of states in phase space increases very rapidly with energy (e.g., as U^N).

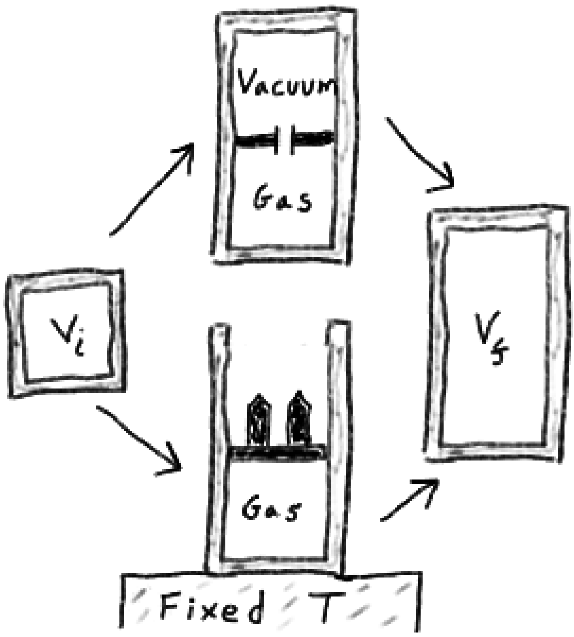

Figure 2.9: Two alternative expansions of a gas in an insulated container leading to the same final state.

Gibbs' Paradox §1.5

Lecture 7

In the previous section, we obtained the following expression for the entropy of a monatomic ideal gas,

$$\begin{aligned} S(U,V,N) &= kN\ln\left[\frac{V}{h^3}\left(\frac{4\pi mU}{3N}\right)^{3/2}\right]+\frac{3}{2}N \\ &= kN\ln\left(Vu^{3/2}\right)+Ns_0 \qquad (*) \end{aligned}$$

where $u = U/N = \frac{3}{2}kT$ is the energy per particle and

$$s_0 = \frac{3k}{2}\left(1+\ln\frac{4\pi m}{3h^2}\right)$$

Now consider the change in entropy when a gas expands from an initial volume V_i to a final volume V_f, holding U (and thus T) fixed. There are two ways to do this, as shown in Fig. 2.9.

For the reversible process, we can compute the change in entropy from the heat added as

$$S_f - S_i = \int_i^f \frac{dQ}{T} = \frac{1}{T}\int_i^f dQ = \frac{1}{T}\Delta Q_{if}$$

The internal energy U is constant (since T is constant) so $\Delta Q_{if} = \Delta W_{if}$. Back in the first lecture we worked out that for the isothermal expansion of a gas

$$\Delta W_{if} = kNT\ln(V_f/V_i)$$

so thermodynamics predicts

$$S_f - S_i = kN\ln(V_f/V_i)$$

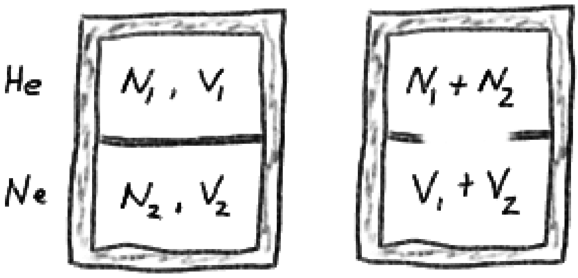

Figure 2.10: Mixing of two different gases (e.g., helium and neon).

Using our statistical mechanics formula for entropy, we have

$$\begin{aligned} S(U,V_f,N)-S(U,V_i,N) &= (kN\ln V_f+\frac{3}{2}kN\ln u+Ns_0) \\ &\quad -(kN\ln V_i+\frac{3}{2}kN\ln u+Ns_0) \\ &= kN\ln(V_f/V_i) \end{aligned}$$

So our results agree; so far, so good.

Now consider the case where we have two *different* gases on each side of a partition (see Fig. 2.10).

We can make the transformation between states irreversibly by punching out the partition. The transformation can be achieved reversibly using semipermeable membranes.

Our statistical mechanics formula predicts a change in entropy of

$$S_f - S_i = kN_1\ln(\frac{V_1+V_2}{V_1}) + kN_2\ln(\frac{V_1+V_2}{V_2})$$

which is the sum of the individual entropy increases. This change in the entropy is called the *entropy of mixing.*

However, suppose that we had the *same* gas (say helium) on each side. If the density on each side of the partition is $N_1/V_1 = N_2/V_2 = (N_1+N_2)\,/\,(V_1+V_2)$ then we expect *no* entropy change in going from the initial to final state. In other words, we can open and close the partition and no change occurs in the system.

Now for the paradox—our statistical mechanics expression for entropy predicts an increase in entropy even when the gas on each side is the same.

There is a second problem with expression $(*)$ for entropy: $S(U,V,N)$ must be an extensive function. If we double the size of a system, the entropy must exactly double (i.e., $S(2U,2V,2N) = 2S(U,V,N)$). However, the function given by equation $(*)$ is *not* extensive.

The resolution of the paradox is to realize that our derivation of $S(U,V,N)$ assumed that the particles were *distinguishable.* However, quantum mechanics tells us that two helium atoms are indistinguishable.

If particles are indistinguishable, the number of states in phase space, Γ or Σ, decreases. The reason is that exchanging particle i and j quantum numbers is not counted as a separate state. For N *indistinguishable* particles, the number of states,Σ, is reduced by a factor of $1/N!$ so

$$\Sigma_{new} = \frac{1}{N!}\Sigma_{old}$$

and

$$S_{new} = k\ln\Sigma_{new} = k\ln\Sigma_{old} - k\ln N!$$

$$
\begin{aligned}
&= S_{old} - k \ln N! \\
&= S_{old} - k \left(N \ln N - N \right) \\
&= kN \ln \left(\frac{V}{N} \left(\frac{U}{N} \right)^{3/2} \right) + NC
\end{aligned}
$$

where C is a constant. You can see that $S(U, V, N)_{new}$ is an extensive function so this resolves Gibbs' paradox. This corrected expression for the entropy of a monatomic ideal gas is called the Sackur-Tetrode equation. Our previous results (before the $1/N!$ correction) are unaffected because the term added to the entropy is independent of U and V.

If instead of one type of particle we have M distinguishable species with particle numbers $N_1, \ldots, N_M$ then the Gibbs correction is

$$\Sigma_{new} = \frac{1}{N_1! \ldots N_M!} \Sigma_{old}$$

so

$$S_{new} = S_{old} - k \ln N_1! - \ldots - k \ln N_M!$$

which you can check makes $S(U, V, N_1, \ldots, N_M)_{new}$ an extensive function.

Finally, note that the combinatorial argument which gave us the $1/N!$ factor assumes only one particle per quantum state. This is a good assumption except at low temperatures where particle occupancy of states has to be treated more carefully. We return to do the low temperature scenario when we consider Fermi and Bose quantum ideal gases.

Chapter 3

Ensemble Theory II

Canonical Ensemble (§3.2)

Lecture 8

Our fundamental definition for entropy is

$$S = -k \sum_{i}^{states} \mathcal{P}_i \ln \mathcal{P}_i$$

where the sum is over all allowed states, and $\mathcal{P}_i$ is the probability of each state.

In the microcanonical ensemble, only states that had energy $E_i = U$ were allowed and all allowed states were equally probable, thus

$$\mathcal{P}_i = \frac{1}{\Gamma} \qquad \text{(microcanonical ensemble)}$$

where Γ was the total number of allowed states. We found that sometimes it was more convenient to find Σ, the number of states with energy $E_i \leq U$. Besides, we saw that $\Gamma \approx \Sigma$ which comes from the fact that there are very few states with energy less than U (i.e., the number of states increases astronomically fast with increasing energy).

In the canonical ensemble, we let *all* states be allowed states but demand that the average energy equal U. The average energy is

$$\langle E \rangle = \sum_{i}^{states} E_i \mathcal{P}_i$$

We now allow states with energy $E_i > U$. However, our constraint on $\langle E \rangle$ will assign them low probability.

We want to determine $\mathcal{P}_i$ such that S is maximum. There are two constraints on $\mathcal{P}_i$,

$$\sum_i \mathcal{P}_i = 1 \qquad ; \qquad \sum_i E_i \mathcal{P}_i = U$$

The first is the demand that probabilities must sum to unity; the second is our condition that $\langle E \rangle = U$.

To maximize S with these constraints, we use the method of Lagrange multipliers. Introduce the function

$$W = S - \alpha' \sum_i \mathcal{P}_i - \beta' \sum_i E_i \mathcal{P}_i$$

where α' and β' are Lagrange multipliers.

We seek to maximize W so we compute $\partial W/\partial \mathcal{P}_j$ and set the derivative to zero,

$$\begin{aligned}
\frac{\partial W}{\partial \mathcal{P}_j} &= \frac{\partial}{\partial \mathcal{P}_j}\left(-k\sum_i \mathcal{P}_i \ln \mathcal{P}_i - \alpha' \sum_i \mathcal{P}_i - \beta' \sum_i E_i \mathcal{P}_i\right) \\
&= -\sum_i \frac{\partial}{\partial \mathcal{P}_j}\left(k\mathcal{P}_i \ln \mathcal{P}_i + \alpha' \mathcal{P}_i + \beta' E_i \mathcal{P}_i\right) \\
&= -\sum_i \left(k \ln \mathcal{P}_j + k + \alpha' + \beta' E_j\right)\delta_{ij} \\
&= -k(\ln \mathcal{P}_j + \alpha + \beta E_j)
\end{aligned}$$

where $\alpha \equiv 1 + \alpha'/k$, $\beta \equiv \beta'/k$.

Setting $\partial W/\partial \mathcal{P}_j = 0$ gives

$$\ln \mathcal{P}_j = -\alpha - \beta E_j$$

or

$$\mathcal{P}_j = e^{-\alpha} e^{-\beta E_j}$$

Notice that in the canonical ensemble, states are not equally probable. Instead, the higher the energy of a state, the lower the probability (we'll see $\beta > 0$ so as E_j increases $e^{-\beta E_j}$ decreases so $\mathcal{P}_j$ decreases).

But isn't our fundamental assumption of statistical mechanics that all accessible states are equally probable? Not quite. We stated that for an *isolated* system (no energy in or out), all accessible states were equally probable. The canonical ensemble describes equilibrium systems but not isolated equilibrium systems since energy is not fixed. So what type of equilibrium system does the canonical ensemble describe? We'll see in a just a few pages (look at Fig. 3.2 if you just can't wait).

The two Lagrange multipliers are determined by imposing the constraints

$$\sum_i \mathcal{P}_i = 1; \qquad \sum E_i \mathcal{P}_i = U$$

on the above expression for $\mathcal{P}_j$.

From the constraint that the probabilities sum to unity.

$$\sum_i (e^{-\alpha} e^{-\beta E_i}) = 1$$

or

$$e^{-\alpha} \sum_i e^{-\beta E_i} = 1$$

We define the *partition function* as,

$$Q_N \equiv \sum_i e^{-\beta E_i}$$

so $e^{-\alpha} Q_N = 1$ or

$$\alpha = \ln Q_N$$

This gives us α in terms of β.

The probabilities are thus

$$\mathcal{P}_i = \frac{e^{-\beta E_i}}{\sum\limits_i e^{-\beta E_i}} = \frac{e^{-\beta E_i}}{Q_N}$$

For the constraint demanding that the average energy equal U,

$$\begin{aligned}
U &= \sum_i E_i \mathcal{P}_i = \sum_i E_i \frac{e^{-\beta E_i}}{Q_N} \\
&= \frac{1}{Q_N} \sum_i E_i e^{-\beta E_i} \\
&= -\frac{1}{Q_N} \frac{\partial}{\partial \beta} \sum_i e^{-\beta E_i} \qquad \text{(slick trick)} \\
&= -\frac{1}{Q_N} \frac{\partial}{\partial \beta} Q_N
\end{aligned}$$

so

$$U = -\frac{\partial}{\partial \beta} \ln Q_N$$

While this is not an explicit expression for β, it turns out to be good enough. Notice the slick trick used above to change the sum by introducing a derivative.

Collecting our results, we can write the entropy as

$$\begin{aligned}
S &= -k \sum_i \mathcal{P}_i \ln \mathcal{P}_i \\
&= -k \sum_i \left(\frac{1}{Q_N} e^{-\beta E_i} \right) \ln \left(\frac{1}{Q_N} e^{-\beta E_i} \right) \\
&= -k \sum_i \left(\frac{1}{Q_N} e^{-\beta E_i} \right) [-\beta E_i - \ln Q_N] \\
&= +k\beta \sum_i \frac{1}{Q_N} E_i e^{-\beta E_i} + k \ln Q_N \sum_i \frac{1}{Q_N} e^{-\beta E_i} \\
&= k\beta \sum_i E_i \mathcal{P}_i + k \ln Q_N \sum_i \mathcal{P}_i
\end{aligned}$$

Finally,

$$S = k\beta U + k \ln Q_N$$

Our link to thermodynamics is still shaky since we haven't nailed down β in terms of something familiar. Let's hammer using

$$\left(\frac{\partial S}{\partial U} \right)_V = \frac{1}{T}$$

then

$$\frac{1}{T} = \left(\frac{\partial}{\partial U} \right)_V [k\beta U + k \ln Q_N] = k\beta$$

so

$$\beta = \frac{1}{kT}; \qquad T = \frac{1}{k\beta}$$

which is the link we wanted to find.

Our entropy is

$$S = \frac{1}{T} U + k \ln Q_N$$

or

$$kT \ln Q_N = TS - U = -A$$

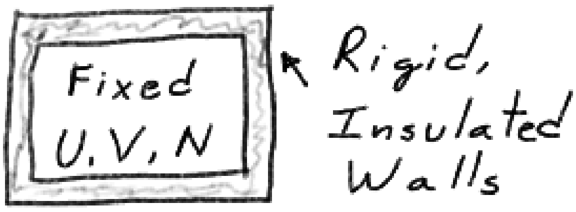

Figure 3.1: Isolated system represented by the microcanonical ensemble.

so

$$Q_N = e^{-A/kT}$$

where A is the Helmholtz free energy.

We have all the results we need. Here is the recipe for using the canonical ensemble:

1. Determine all the possible states that a system can be in given the constraints that volume, V, and particle number, N, are fixed.
2. Determine the energy, E_i, of each state for the system.
3. Compute the partition function by evaluating the sum

   $$Q_N = \sum_i e^{-\beta E_i}$$

   where $\beta = 1/kT$.
4. Given Q_N compute other quantities of interest, for example,

   $$U = -\frac{\partial}{\partial \beta} \ln Q_N$$

   $$S = \frac{1}{T} U + k \ln Q_N$$

   $$A = -kT \ln Q_N$$

   and

   $$P = -\left(\frac{\partial A}{\partial V}\right)_T = kT \left(\frac{\partial}{\partial V}\right)_T \ln Q_N$$

Before going to some examples, I should discuss the physical meaning of the canonical ensemble. We already saw that the microcanonical ensemble represents an isolated system (fixed U, V, and N) such as illustrated in Fig. 3.1. The canonical ensemble is similar except that the system's energy is not strictly fixed to be U, rather on average the energy is U. This turns out to be equivalent to a system at fixed temperature T (see Fig. 3.2).

Important Point: The thermodynamic properties derived from the two ensembles are *identical*. The only reason we use two ensembles is for *computational convenience*; for different problems one is easier to work with than the other. However, all thermodynamic quantities (entropy, equation of state, heat capacities, etc.) are identical (except for small differences that go to zero as $N \to \infty$).

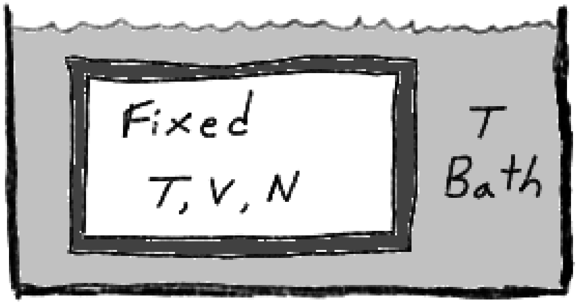

Figure 3.2: Thermal system represented by the canonical ensemble.

Quantum Harmonic Oscillators (§3.8)

Lecture 9

We solved this system using microcanonical ensemble; now let's use canonical ensemble. Again, we have N quantum harmonic oscillators; the energy of each oscillator is

$$\epsilon_j\left(n_j\right) = \hbar\omega\left(n_j + 1/2\right)$$

where $n_j = 0, 1, ...$ is the quantum number specifying the energy level of the oscillator.

The total energy of the system is

$$E = \epsilon_1 + \epsilon_2 + \ldots + \epsilon_N$$

The partition function is

$$\begin{aligned} Q_N &= \sum^{states} e^{-\beta E} \\ &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\ldots\sum_{n_N=0}^{\infty} e^{-\beta(\epsilon_1+\epsilon_2+\ldots+\epsilon_N)} \\ &= \left(\sum_{n_1=0}^{\infty} e^{-\beta\epsilon_1}\right)\left(\sum_{n_2=0}^{\infty} e^{-\beta\epsilon_2}\right)\ldots\left(\sum_{n_N=0}^{\infty} e^{-\beta\epsilon_N}\right) \\ &= \left(Q_1\right)^N \end{aligned}$$

where Q_1 is the partition function for an individual oscillator

$$Q_1 = \sum_{n=0}^{\infty} e^{-\beta\epsilon(n)} = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+1/2)}$$

Before doing this simple sum, two points: 1) Since the individual oscillators may be treated as independent elements and the energy written as the sum of the individual energies, $Q_N = (Q_1)^N$. This is a common (and nice) feature; 2) In this problem the oscillators are taken to be *distinguishable*. For indistinguishable elements, we must include a Gibbs correction to the sum over states, in which case $Q_N = \frac{1}{N!}(Q_1)^N$.

To finish the example, we use the identity

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a} \qquad |a| < 1$$

so

$$Q_1 = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} = \frac{1}{2}\sinh\left(\beta\hbar\omega/2\right)$$

Using

$$U = -\frac{\partial}{\partial\beta}\ln Q_N = -\frac{\partial}{\partial\beta}\left(N\ln Q_1\right)$$

then

$$U = \frac{1}{2}N\hbar\omega\coth(\beta\hbar\omega/2)$$

You can check that this result matches our microcanonical result for U (recall $\beta = 1/kT$).

Monatomic Ideal Gas – Canonical Ensemble

This is the third (and I promise, the last) time we derive the thermodynamics of this system using statistical mechanics. It is worth repeating using the canonical ensemble since our derivation will be useful when we study fermion and boson systems.

Recall that the energy levels of a particle in a cubic box of volume $V = L^3$ are

$$\epsilon\left(n_x, n_y, n_z\right) = \frac{h^2}{8mL^2}(n_x^2 + n_y^2 + n_z^2)$$

Since we have an ideal gas, the interactions between the particles are negligible. The partition function is thus

$$Q_N = \frac{1}{N!}Q_1^N$$

where Q_1 is the partition function for a single particle. Notice we have a factor of $1/N!$ out in front because the particles are indistinguishable.

The single particle partition function is

$$Q_1 = \sum_{n_x=1}^{\infty}\sum_{n_y=1}^{\infty}\sum_{n_z=1}^{\infty} e^{-\beta\epsilon(n_x,n_y,n_z)}$$

or

$$Q_1 = \left(\sum_{n_x=1}^{\infty}\exp\left(-\frac{\pi\lambda^2 n_x^2}{4L^2}\right)\right)(\text{same for } n_y)\,(\text{same for } n_z)$$

where

$$\lambda = \frac{h}{\sqrt{2\pi mkT}}$$

is the thermal wavelength (the de Broglie wavelength of a particle with energy $\approx kT$).

In the classical limit the de Broglie wavelength of the particles is much smaller than the distance between them $\left(\lambda \ll \sqrt[3]{V/N} = L/\sqrt[3]{N}\right)$, so we may replace the sums with integrals.

$$\sum_{n_x=1}^{\infty}\exp\left(-\frac{\pi\lambda^2 n_x^2}{4L^2}\right) \approx \int_0^{\infty} dx\,\exp\left(-\frac{\pi\lambda^2 x^2}{4L^2}\right)$$

This Gaussian integral is easy to compute,

$$\int_0^{\infty} dx\,\exp\left(-\frac{\pi\lambda^2 x^2}{4L^2}\right) = \frac{1}{2}\sqrt{\frac{\pi}{\left(\frac{\pi\lambda^2}{4L^2}\right)}} = \frac{L}{\lambda}$$

so

$$Q_1 \approx \left(\frac{L}{\lambda}\right)\left(\frac{L}{\lambda}\right)\left(\frac{L}{\lambda}\right) = \frac{V}{\lambda^3}$$

The N particle partition function is

$$Q_N = \frac{1}{N!}Q_1^N = \frac{1}{N!}\left(\frac{V}{\lambda^3}\right)^N$$

The internal energy is

$$U = -\left(\frac{\partial}{\partial\beta}\right)_V \ln Q_N = 3N\left(\frac{\partial}{\partial\beta}\right)_V \ln\lambda = \frac{3N}{2\beta} = \frac{3}{2}NkT$$

The pressure is

$$P = kT\left(\frac{\partial}{\partial V}\right)_T \ln Q_N = kT\left(\frac{\partial}{\partial V}\right)_T \ln V^N = \frac{NkT}{V}$$

confirming all of our previous results.

Canonical Ensemble in Classical Mechanics (§3.5)

Lecture 10

So far we've developed the canonical ensemble in the framework of physical systems with discrete states. This formulation is appropriate for quantum mechanics but in classical mechanics, our phase space is continuous. However, the extension to continuous states is simply a replacement of sums over states with integrals in phase space.

Consider a physical system whose state is given by the coordinates and momenta $(q,p) = \{q_1,\ldots,q_{3N}, p_1,\ldots,p_{3N}\}$. The energy of this state is $E(q,p)$. In the canonical ensemble, the probability of being in this state is

$$\mathcal{P}(q,p)\,dqdp = \frac{1}{Q_N}e^{-\beta E(q,p)}\,dqdp$$

where the partition function is

$$Q_N = \frac{1}{h^{3N}}\int dq\int dp\, e^{-\beta E(q,p)}$$

where the integral over q is restricted to positions inside the volume V and the integral over p is unrestricted.

Our previous identities linking Q_N to standard thermodynamic quantities (S,U,P,etc.) are unchanged.

Equipartition Theorem (§3.7)

Let's derive an important general result using the classical canonical ensemble. Again, consider a system whose energy is $E(q_1,\ldots,q_{3N},p_1,\ldots,p_{3N}) = E(q,p)$.

Suppose that EITHER:

a1) The total energy splits additively as

$$E(q,p) = \epsilon_i(p_i) + E'(q,p')$$

where the prime indicates that p_i is missing *and* **a2)** The function ϵ_i is quadratic in p_i so

$$\epsilon_i(p_i) = bp_i^2$$

where b is a constant.

OR

b1) The total energy splits additively as

$$E(q,p) = \epsilon_i(q_i) + E'(q'p)$$

where the prime indicates that q_i is missing *and* **b2)** The function ϵ_i is quadratic in q_i as

$$\epsilon_i(q_i) = bq_i^2$$

where b is a constant.

Case a) is common since often the energy of a system is of the form

$$E = \text{kinetic energy} + \text{potential energy}$$

and the kinetic energy of a particle is

$$K = \frac{1}{2}m\left|\vec{v}\right|^2 = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right)$$

so each component satisfies a1) and a2).

Case b) occurs when the potential energy of a particle is well approximated by the harmonic oscillator potential, $\frac{1}{2}kq^2$, where k is the effective spring constant.

Back to our derivation. We want to compute the average value of ϵ_i, that is

$$\begin{aligned}\langle\epsilon_i\rangle &= \frac{1}{h^{3N}}\int dq \int dp\, \epsilon_i \mathcal{P}(q,p) \\ &= \frac{\frac{1}{h^{3N}}\int dq \int dp\, \epsilon_i \exp\left(-\beta E(q,p)\right)}{\frac{1}{h^{3N}}\int dq \int dp\, \exp\left(-\beta E(q,p)\right)}\end{aligned}$$

Consider case a) (the derivation for b) is similar)

$$\begin{aligned}\langle\epsilon_i\rangle &= \frac{\int dp_i \epsilon_i e^{-\beta\epsilon_i}\int dq \int dp' e^{-\beta E'(q,p')}}{\int dp_i e^{-\beta\epsilon_i}\int dq \int dp' e^{-\beta E'(q,p')}} \\ &= \frac{\int dp_i \epsilon_i e^{-\beta\epsilon_i}}{\int dp_i e^{-\beta\epsilon_i}} = \frac{-\frac{\partial}{\partial\beta}\int e^{-\beta\epsilon_i}dp_i}{\int e^{-\beta\epsilon_i}dp_i} \\ &= -\frac{\partial}{\partial\beta}\ln\left(\int e^{-\beta\epsilon_i}dp_i\right)\end{aligned}$$

since $\frac{1}{f(x)}\frac{d}{dx}f(x) = \frac{d}{dx}\ln f(x)$.

Using a2),

$$\begin{aligned}\langle\epsilon_i\rangle &= -\frac{\partial}{\partial\beta}\ln\left(\int e^{-\beta b p_i^2}dp_i\right) \\ &= -\frac{\partial}{\partial\beta}\ln\left(\beta^{-1/2}\int e^{-by^2}dy\right) \qquad \text{(Use } y \equiv \beta^{1/2}p_i\text{)} \\ &= -\frac{\partial}{\partial\beta}\left[-1/2\ln\beta + \ln\int e^{-by^2}dy\right] \qquad \text{(No } \beta \text{ dependence in 2nd term)} \\ &= \frac{1}{2}\frac{1}{\beta} = \frac{1}{2}kT\end{aligned}$$

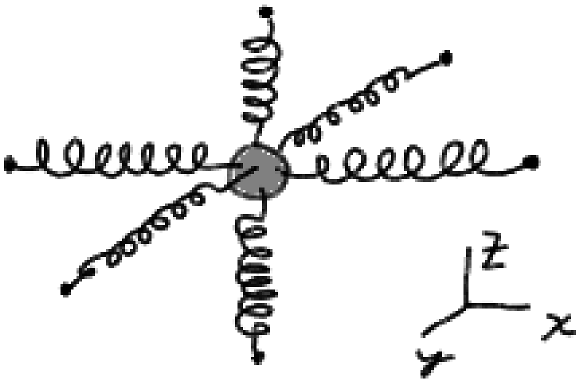

Figure 3.3: Atom in a solid modeled as mass on springs.

Thus the average value of energy for component i is $\frac{1}{2}kT$. The derivation for case b) is identical except we interchange $p \Longleftrightarrow q$.

Whenever the energy of a particle has a component that goes as p^2 or q^2 we call this a *degree of freedom.* From the equipartition theorem, each degree of freedom has an average energy of $\frac{1}{2}kT$. A monatomic gas atom has three degrees of freedom (d.o.f.) so the energy in one mole of the gas is $U = \frac{3}{2}N_AkT = \frac{3}{2}RT$. At low T, a diatomic gas has five degrees of freedom (three translational d.o.f. for the center of mass motion plus two rotational kinetic energy d.o.f.) in the absence of vibrational states. At high temperatures, the molecule may be approximated as a pair of masses coupled by a spring giving seven degrees of freedom (three translational d.o.f. for each atom in the molecule plus one potential d.o.f. from the spring). Thus a diatomic gas has a molar specific heat of $c_V = \frac{5}{2}R$ at low temperatures and $c_V = \frac{7}{2}R$ at high temperatures.

Heat Capacity of a Solid

Consider the following simple model for a solid. Each atom is independent and rattles around in a cage formed by springs (see Fig. 3.3).

The energy of the atom may be written as

$$\begin{aligned} E_{Atom} &= \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) \\ &+ \frac{1}{2}k\left(q_x^2 + q_y^2 + q_z^2\right) \end{aligned}$$

Using equipartition theorem on each term,

$$\langle E_{Atom} \rangle = 6\left(\frac{1}{2}kT\right) = 3kT$$

For one mole of atoms

$$U = N_A \langle E_{Atom} \rangle = 3N_AkT = 3RT$$

Finally, the heat capacity (for 1 mole)

$$c = \frac{dU}{dT} = 3R$$

(for solids, expansion is negligible so $c_P \approx c_V = c$).

Our result says that the molar heat capacity of a solid is $3R$ *independent* of the material. This is called the law of Dulong and Petit. The result is quite accurate, despite the simplicity of the model. At low temperatures, quantum effects become important and a more complex formulation (e.g., Debye theory) is needed.

Chemical Potential

Lecture 11

So far we have implicitly assumed that the number of particles, N, in a system was fixed. We now consider the more general case in which N is allowed to vary.

As we have seen, the entropy, $S(U, V, N)$, and internal energy, $U(S, V, N)$, are functions of N. Specifically we have

$$dU = \left(\frac{\partial U}{\partial S}\right)_{V,N} dS + \left(\frac{\partial U}{\partial V}\right)_{S,N} dV + \left(\frac{\partial U}{\partial N}\right)_{S,V} dN$$

or

$$dU = TdS - PdV + \mu dN$$

where the chemical potential is defined as

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

that is, μ is the change in the internal energy with N, given that S and V are fixed. Notice that μ has the dimensions of energy per particle; in fact it is the Gibbs free energy per particle.

An alternative expression for chemical potential is

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

which may be obtained from $TdS = dU + PdV - \mu dN$.

For a monatomic ideal gas, the Sackur-Tetrode equation gives the entropy,

$$S(U, V, N) = \frac{5}{2}kN + kN\ln\left(C_0\frac{VU^{3/2}}{N^{5/2}}\right)$$

where C_0 is a constant. The chemical potential is

$$\mu = -kT\ln\left(C_0\frac{VU^{3/2}}{N^{5/2}}\right)$$

Notice that even an inert monatomic gas (such as neon) has a chemical potential, i.e., μ has no more to do with chemistry than, say, temperature. Also notice that the chemical potential is negative since increasing the number of particles will increase the number of accessible states (and thus increase S) unless the energy U simultaneously decreases. From this observation and the definition

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

we see that the chemical potential is the amount of energy that must be removed when one particle is added in order to keep S fixed.

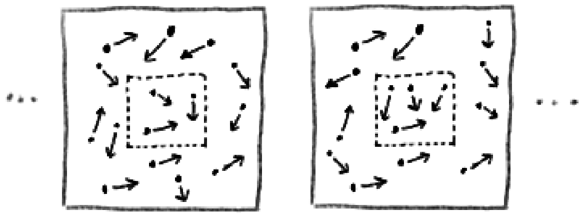

Figure 3.4: Some possible states for the grand canonical ensemble.

Example: Consider a system of distinguishable particles with energy levels $e = 0, \epsilon, 2\epsilon, \ldots$. For a system with $N = 2$ particles and energy $U = 2\epsilon$, find the entropy and the chemical potential.

Solution: For $N = 2$ and $U = 2\epsilon$ you can count that there are $\Gamma = 3$ accessible states, specifically,

$$(2\epsilon, 0) \qquad (0, 2\epsilon) \qquad (\epsilon, \epsilon)$$

So the entropy is $S = k \ln \Gamma = k \ln 3$. If we add one particle keeping the energy fixed then the number of accessible states increases to 6, specifically,

$$\begin{array}{ccc} (2\epsilon, 0, 0) & (0, 2\epsilon, 0) & (0, 0, 2\epsilon) \\ (\epsilon, \epsilon, 0) & (\epsilon, 0, \epsilon) & (0, \epsilon, \epsilon) \end{array}$$

To find the chemical potential, we need to find how much energy we must remove to bring the entropy back down to its original value. You can check that with $N = 3$ and $U = 1$ the number of accessible states returns to $\Gamma = 3$, specifically,

$$(\epsilon, 0, 0) \qquad (0, \epsilon, 0) \qquad (0, 0, \epsilon)$$

Thus

$$\mu = \left(\frac{\Delta U}{\Delta N} \right)_{S,V} = \frac{-\epsilon}{1} = -\epsilon$$

Grand Canonical Ensemble (§4.1–4.3)

In the microcanonical ensemble, our ensemble of systems consisted of all states with a given energy, volume, and N. For the canonical ensemble, we relaxed the condition that energy be fixed and simply demanded that *on average* the energy equaled U. For the grand canonical ensemble, we further relax the condition that the number of particles is fixed and simply demand that *on average* it equal N.

Some allowed states in the grand canonical ensemble are illustrated in Fig. 3.4. The physical system is the dashed rectangle, that is, a volume V, is delineated in a large reservoir. The energy and number of particles in the volume V may vary yet their averages, U and N, are fixed.

Some time back, we saw how the canonical ensemble was derived from the condition that the entropy

$$S = -k \sum_i^{states} \mathcal{P}_i \ln \mathcal{P}_i$$

was maximum under the constraints

$$\sum_i^{states} \mathcal{P}_i = 1 \quad ; \quad \sum_i^{states} E_i \mathcal{P}_i = U$$

where $\mathcal{P}_i$ is the probability of a state and E_i is the energy of a state.

For the grand canonical ensemble, states can have different numbers of particles; call N_i the number of particles for state i. We now add an additional constraint that the average value of N_i equals N, that is

$$\sum_i^{states} N_i \mathcal{P}_i = N$$

Note: we use E_i for the energy of a state and U for the average energy of the system, averaged over all states. With number of particles we use N_i and N for state and system values; hope this is not confusing but the notation is standard and probably necessary since we are running out of letters.

Using Lagrange multipliers, as before, we obtain that S is maximum when

$$\mathcal{P}_i = \frac{e^{-\beta E_i + \beta\mu N_i}}{\sum\limits_i e^{-\beta E_i + \beta\mu N_i}}$$

Since the constraint on N is similar to the constraint on E it should not be surprising that their terms in the above expression are similar.

The denominator in the above expression for $\mathcal{P}_i$ is the grand canonical partition function, $\mathcal{Q}$. It is useful to write it as

$$\begin{aligned} \mathcal{Q} &= \sum_i^{states} e^{-\beta E_i + \beta\mu N_i} \\ &= \sum_{N=0}^{\infty} e^{\beta\mu N} \sum_j e^{-\beta E_j} \qquad \text{(sum } j \text{ is over states with } N_i = N) \\ &= \sum_{N=0}^{\infty} e^{\beta\mu N} Q_N \end{aligned}$$

where Q_N is the canonical partition function with N particles.

Define $z = e^{\beta\mu}$ as the *fugacity* (sometimes called the absolute activity). Our partition function may then be written as

$$\mathcal{Q} = \sum_{N=0}^{\infty} z^N Q_N$$

We can compute other quantities in terms of $\mathcal{Q}$ using the following results

$$\begin{aligned} U &= \sum_i E_i \mathcal{P}_i = -\left(\frac{\partial}{\partial\beta}\right)_z \ln\mathcal{Q} \\ N &= \sum_i N_i \mathcal{P}_i = z\left(\frac{\partial}{\partial z}\right)_\beta \ln\mathcal{Q} \\ S &= -k\sum_i \mathcal{P}_i \ln\mathcal{P}_i = k\ln\mathcal{Q} + k\beta U - k\beta\mu N \end{aligned}$$

and

$$\frac{PV}{kT} = \ln\mathcal{Q}$$

Example: Given the canonical partition function for a highly relativistic ideal gas of N indistinguishable particles to be,

$$Q_N = \frac{1}{N!}\left[8\pi V(hc\beta)^{-3}\right]^N$$

Find the grand canonical partition function, $\mathcal{Q}$, and use it to obtain the equation of state, $P(T,V,N)$.

Solution: Right off the bat I should mention that since we have Q_N we could find all the thermodynamic quantities directly using it; this exercise is simply to demonstrate the manipulation of $\mathcal{Q}$.

The grand canonical partition function may be written in terms of Q_N as,

$$\mathcal{Q} = \sum_{N=0}^{\infty} z^N Q_N = \sum_{N=0}^{\infty} \frac{1}{N!}\left[8\pi V(hc\beta)^{-3} z\right]^N$$

Using

$$e^x = \sum_{N=0}^{\infty} \frac{x^N}{N!}$$

gives

$$\mathcal{Q} = \exp(8\pi V(hc\beta)^{-3} z)$$

so

$$\frac{PV}{kT} = 8\pi V(hc\beta)^{-3} z$$

To replace z with N in this expression we evaluate,

$$N = \sum_i N_i \mathcal{P}_i = z\left(\frac{\partial}{\partial z}\right)_\beta = 8\pi V(hc\beta)^{-3} z$$

which yields the familiar result, $P = NkT/V$.

Before leaving this chapter, you might ask:

Q: For a physical system (e.g. diatomic gas, crystal, semiconductor) don't all three ensembles (micro, plain, and grand) give the same answer for S, U, C_V, etc..?

A: Yes.

Q: Exactly the same answer?

A: Not quite, but, the differences are terms which vanish in the limit $N \to \infty, V \to \infty, N/V$ fixed.

Q: So for all practical purposes, the answers are the same?

A: Yes.

Q: So why bother with three different ensembles?

A: Because depending on the problem at hand one ensemble can be much easier to work with than another.

Q: How could the grand canonical ensemble ever be simpler since the above expression for $\mathcal{Q}$ involves the canonical ensemble partition function Q_N?

A: You are about to see an important example. The key will be to find a trick for getting $\mathcal{Q}$ without Q_N.

Chapter 4

Sample Exercises for First Midterm

Lecture 10

The problems below are typical of midterm problems and are good exercises for review.

Problem: (a) Show that for $S(T,V)$,

$$dS = \frac{C_V}{T}dT + \left(\frac{\partial P}{\partial T}\right)_V dV$$

(b) Show that for $U(T,V)$,

$$dU = C_V dT + \left[T\left(\frac{\partial P}{\partial T}\right)_V - P\right] dV$$

Hint: Use part (a).
(c) Consider a van der Waals gas with molar heat capacity C_V (i.e., heat capacity for one mole) and equation of state

$$P = \frac{nRT}{V - bn} - \frac{an^2}{V^2}$$

Find $(\partial U/\partial V)_T$ for this gas.
(d) Consider n moles of this gas undergoing free expansion from an initial state T_i, V_i to a final volume V_f. Find the final temperature T_f.

Solution: (a) For $S(T,V)$,

$$dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV$$

but

$$\left(\frac{\partial S}{\partial T}\right)_V = \frac{C_V}{T} \qquad \left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V$$

where we used the Maxwell relation (M3).

(b) Using $dU = TdS - PdV$ and the result from part (a) for dS gives the desired result.

(c) Comparing the mathematical expression

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV$$

with the result from part (b) gives

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

From the equation of state

$$\left(\frac{\partial P}{\partial T}\right)_V = \frac{nR}{V - bn}$$

so $(\partial U/\partial V)_T = an^2/V^2$. Notice that it is zero for an ideal gas (for which $a = 0$).

(d) A system does no work and exchanges no heat in a free expansion so $dU = 0$ and

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV = nC_V dT + \frac{an^2}{V^2}dV = 0$$

using the result from part (c). From this

$$dT = -\frac{an}{C_V V^2}dV$$

so integrating both sides gives

$$T_2 - T_1 = \frac{an}{C_V}\left(\frac{1}{V_2} - \frac{1}{V_1}\right)$$

Note that unlike an ideal gas, a van der Waals gas cools on expanding due to the potential energy between the molecules. In case you're concerned about a free expansion not being a reversible process, remember that we're working with exact differentials so we can always consider another path, which is reversible, which goes from the same initial and final states and has $\Delta U = 0$.

Problem: Consider the classical ideal gas (with N indistinguishable particles in a volume V) in the highly relativistic limit where the energy of a particle may be approximated as, $\epsilon = c(|p_x| + |p_y| + |p_z|)$, where p_i are the components of the momentum of the particle and c is the speed of light.
(a) Using the classical *micro-canonical* ensemble (fixed U, V, N), find the number of states Σ that have energy $E \leq U$. Note that the volume of a hyper-simplex in d dimensions (a pyramid whose edges are the lines connecting the origin and the points $x_1 = L$, $x_2 = L, \ldots, x_d = L$) is $L^d/d!$.
(b) From the result in part (a), find $U(V,T)$. Note: If you could not do part (a) then take $\Sigma = AV^B U^C$, where A, B, and C are constants.
(c) From the result in part (a), find the equation of state.
(d) Find $\gamma = C_P/C_V$, the ratio of the heat capacities.

Solution: (a) We want to evaluate the following expression

$$\Sigma = \frac{1}{N!h^{3N}}\int d\mathbf{r}_1 \ldots \int d\mathbf{r}_N \int d\mathbf{p}_1 \ldots \int d\mathbf{p}_N$$

where the position integrals are constrained such that $\mathbf{r}_i$ is within the volume V. The momentum integrals are constrained such that

$$0 \leq c(p_{x,1} + p_{y,1} + \ldots + p_{z,N}) \leq U$$

Each position integral is independent and gives a factor of V. The momentum integrals equal the volume of a hyper-simplex with sides $L = U/c$ and dimension $d = 3N$ multiplied by the number of quadrants in d dimensions, which is 2^d. Thus

$$\Sigma = \frac{1}{N!h^{3N}} V^N \frac{2^{3N}U^{3N}}{(3N)!c^{3N}} = \frac{V^N U^{3N}}{N!(3N)!(hc/2)^{3N}}$$

(b) The entropy is

$$S = k \ln \Sigma = kN \ln V + 3kN \ln U + f(N)$$

where $f = -k \ln(N!(3N)!(hc/2)^{3N})$. To obtain the energy in terms of temperature we may use

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V = \frac{3kN}{U}$$

so $U = 3kNT$.

(c) The equation of state may be found using

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_U = \frac{kN}{V}$$

which gives ideal gas law.

(d) The heat capacity at constant volume is

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = 3Nk$$

In general

$$C_P = C_V + \frac{TV\alpha^2}{K_T}$$

where α is the coefficient of thermal expansion and K_T is the isothermal compressibility. Since our equation of state is the ideal gas law, $\alpha = 1/T$ and $K_T = 1/P$, which gives $C_P = C_V + Nk = 4Nk$, thus $\gamma = C_P/C_V = 4/3$. Note that for the non-relativistic ideal gas $\gamma = 5/3$.

Problem: Consider the following simple model for rubber. A chain consisting of N independent, distinguishable links, each of length a, is attached to a weight of mass M. Each link points either up or down; the individual links have negligible mass and kinetic energy. The energy for any given configuration may be written as $E(L) = Mg[Na - L]$ where the length $L = a(n_1 + n_2 + \ldots + n_N)$, with $n_i = -1$ and $+1$ for links pointing up and down. Find the canonical partition function, Q_N, and use it to find the average length $\langle L \rangle = L(E = U)$ as a function of temperature. Hint: Unlike most materials, rubber expands when cooled; you can check that your answer confirms this observation.

Solution: The partition function is

$$\begin{aligned} Q_N = \sum^{states} e^{-\beta E} &= \sum^{states} e^{-\beta MgNa} \exp(\beta Mga(n_1 + n_2 + \ldots + n_N)) \\ &= e^{-\beta MgNa} \left(\sum_{n_1=-1}^{+1} e^{\beta Mgan_1} \right) \times \ldots \times \left(\sum_{n_N=-1}^{+1} e^{\beta Mgan_N} \right) \\ &= e^{-\beta MgNa} \left(\sum_{n=-1}^{+1} e^{\beta Mgan} \right)^N \\ &= e^{-\beta MgNa} \left(e^{-\beta Mga} + e^{\beta Mga} \right)^N \end{aligned}$$

Using

$$\begin{aligned}
U &= -\frac{\partial}{\partial\beta}\ln Q_N \\
&= \frac{\partial}{\partial\beta}\left[\beta MgNa - N\ln\left(e^{-\beta Mga} + e^{\beta Mga}\right)\right] \\
&= MgNa - MgNa\frac{e^{\beta Mga} - e^{-\beta Mga}}{e^{\beta Mga} + e^{-\beta Mga}}
\end{aligned}$$

Since $L(E) = Na - E/Mg$,

$$\begin{aligned}
\langle L\rangle &= Na - Na + Na\frac{e^{\beta Mga} - e^{-\beta Mga}}{e^{\beta Mga} + e^{-\beta Mga}} \\
&= Na\frac{e^{\beta Mga} - e^{-\beta Mga}}{e^{\beta Mga} + e^{-\beta Mga}} \\
&= Na\tanh(\beta Mga) = Na\tanh(Mga/kT)
\end{aligned}$$

Since $\tanh(x)$ increases as x increases the length increases as temperature decreases. A state of zero entropy occurs when all the links point in the downward direction, which gives the maximum length.

Problem: Consider a classical ideal gas of N indistinguishable particles (mass m) in a rectangular container (sides of length L_x, L_y, L_z) in a constant gravitational field with acceleration of g in the $-z$ direction. The container is very tall, specifically $L_z \gg L_T$ where $L_T = kT/mg$ is the thermal height (i.e., height at which the gravitational potential energy of a particle equals kT).
(a) Find the classical canonical partition function, Q_N. You may find this integral identity helpful

$$\int_0^\infty e^{-a^2x^2} = \frac{\sqrt{\pi}}{2a}$$

(b) Find the heat capacity per particle, $c_V = C_V/N$.
(c) Could the result in part (b) have been predicted from the equipartition theorem? You *must* justify your answer to receive credit.

Solution: (a) The gas is ideal so the canonical partition function may be formulated as

$$Q_N = \frac{1}{N!}(Q_1)^N$$

where the classical, single particle partition function is

$$Q_1 = \frac{1}{h^3}\int_0^{L_x} dx \int_0^{L_y} dy \int_0^{L_z} dz \int_{-\infty}^{\infty} dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z\, e^{-\beta E(\mathrm{r},\mathrm{p})}$$

where the energy is

$$E(\mathrm{r},\mathrm{p}) = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} + mgz$$

The x and y integrals give L_x and L_y, respectively; each momentum integral is similar, for example,

$$\frac{1}{h}\int_{-\infty}^{\infty} dp_x e^{-p_x^2/(2mkT)} = \frac{\sqrt{2\pi mkT}}{h} = \frac{1}{\lambda}$$

Finally, the z integral gives

$$\int_0^{L_z} dz\, e^{-\beta mgz} = \left[\frac{-1}{\beta mg} e^{-\beta mgz}\right]_0^{L_z} = \frac{-1}{\beta mg}\left(e^{-\beta mgL_z} - 1\right)$$

Since the box is tall $\beta mgL_z \gg 1$ so the exponential term is negligible and integral is approximately just $1/\beta mg = kT/mg = L_T$. Collecting these pieces gives

$$Q_N = \frac{1}{N!}\left(\frac{L_x L_y L_T}{\lambda^3}\right)^N$$

(b) The energy of the system is

$$U = -\frac{\partial}{\partial\beta}\ln Q_N = -N\frac{\partial}{\partial\beta}\ln L_T + 3N\frac{\partial}{\partial\beta}\ln\lambda = NkT + \frac{3}{2}NkT = \frac{5}{2}NkT$$

Finally, the heat capacity is

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = \frac{5}{2}Nk$$

and $c_V = \frac{5}{2}k$.

(c) The result in part (b) could not be predicted from the equipartition theorem, as presented in the notes and in class, because the gravitational potential energy is not quadratic in position so the theorem does not apply. Pathria derives a more general form (eqn. (3.7.4)) for which

$$\left\langle z\frac{\partial E}{\partial z}\right\rangle = \langle mgz\rangle = kT$$

which tells us that the average potential energy for a particle is kT. Since the average kinetic energy for a particle is $\frac{3}{2}kT$ we do arrive at the result obtained in part (b). Since either yes or no is a correct answer, credit is given only when your answer is justified.

Lecture
Midterm

Chapter 5

Quantum Ideal Gases

Classical Limit

Lecture 12

We already used the canonical ensemble to study the monatomic ideal gas using quantum mechanics (i.e., non-interacting particles in a box). However, we had to use the approximation,

$$\lambda \ll \sqrt[3]{V/N} \ll L/\sqrt[3]{N}$$

Recall that

$$\lambda = \frac{h}{\sqrt{2\pi mkT}}$$

is the de Broglie wavelength of a particle at temperature T so the above inequality indicates that λ is much smaller than the distance between particles.

This approximation puts us into the classical regime and gives us the classical result for the canonical partition function

$$Q_N = \frac{1}{N!}\left(\frac{V}{\lambda^3}\right)^N \qquad \text{(Classical ideal gas)}$$

The grand canonical partition function is

$$\mathcal{Q} = \sum_{N=0}^{\infty} z^N Q_N = \sum_{N=0}^{\infty} \frac{1}{N!}\left(\frac{zV}{\lambda^3}\right)^N = \exp\left(\frac{zV}{\lambda^3}\right)$$

We now return to the quantum ideal gas but without making the approximation that λ is small. Usually when the distance between particles is comparable to their deBroglie wavelength we may no longer neglect the interparticle forces (i.e., the gas is no longer ideal). However, there are some important exceptions:

- Conduction electrons in a solid (λ is large because m is small).
- Photon gas (i.e., blackbody radiation).
- Liquid Helium (interatomic forces are weak).
- Ultracold systems (λ is large because T is extremely small; $T \approx 10^{-7}$).

In these cases the system of particles can be approximated as a *quantum ideal gas*, that is, the interparticle forces can be neglected yet the gas cannot be treated classically.

Quantum mechanics seems to divide particles into two categories. Particles with integer spin can have more than one particle occupying the same quantum state; these types of particles are called *bosons*.

Examples of such particles (and their spin) are: Alpha (0), Pion (0), Photon (1), Deuteron (1), Gluon (1), W (1), Z^0 (1). Particles with 1/2 integer spin obey an exclusion principle (Pauli exclusion). Each quantum state can have at most one particle; these types of particles are called *fermions.* Examples of such particles (and their spin) are: Electron (1/2), Muon (1/2), Proton (1/2), Neutron (1/2), Quarks (1/2), ^3He (1/2), Ω (3/2), ^{55}Fe (3/2). Why is this the way nature works? Are there other possibilities (i.e. anyons)? Who knows? Until other types of particles are discovered, we'll stick with fermions and bosons.

Fermion Grand Partition Function (§6.2)

We first derive the grand partition function of fermions. Recall that this function may be written in terms of the canonical partition function Q_N as

$$\mathcal{Q} = \sum_{N=0}^{\infty} z^N Q_N$$

where $z = e^{\beta\mu}$ is the fugacity. While we use this definition of $\mathcal{Q}$ we won't explicitly compute Q_N.

Start with the definition of the canonical partition function,

$$\begin{aligned} Q_N &= \sum_{j}^{states} e^{-\beta E_j} \\ &= \sum_{n_1=0}^{1*} \sum_{n_z=0}^{1*} \ldots \sum_{n_M=0}^{1*} \exp\left(-\beta\left[n_1\epsilon_1 + n_2\epsilon_2 + \ldots + n_M\epsilon_M\right]\right) \end{aligned}$$

where n_i is the number of particles in energy state i. We have $i = 1, ..., M$ where M is the number of energy states (which might be infinite). The asterisks on the sums are to remind us of the constraint,

$$n_1 + n_2 + \ldots + n_M = N \qquad n_j = 0 \text{ or } 1$$

that is, there are N particles distributed among the M energy states with at most one particle per energy state.

By the way, don't confuse system states, which are all the possible configurations of the system, with energy states. The energy states are similar to energy levels however remember that energy states don't necessarily have unique energy values (i.e., $\epsilon_i = \epsilon_j$ is allowed, in fact it's common). If we picture the classical phase space for a single particle then each energy level is a parcel in that phase space. We can formulate our sum over system states in terms of just energy states of a single particle because we're still doing ideal gases where the energy of a particle only depends on its own state and not the state of the other particles (although with quantum ideal gases, the energy states that are available to a particle are affected by the number of particles in those energy states).

Our next step is to write Q_N as

$$\begin{aligned} Q_N &= \left(\sum_{n_1=0}^{1} e^{-\beta n_1\epsilon_1}\right)\left(\sum_{n_2=0}^{1*} \ldots \sum_{n_M=0}^{1*} \exp\left(-\beta\left[n_2\epsilon_2 + \ldots + n_M\epsilon_M\right]\right)\right) \\ &= \left(1 + e^{-\beta\epsilon_1}\right)\left(\sum_{n_2=0}^{1*} \ldots \sum_{n_M=0}^{1*} \exp\left(-\beta\left[n_2\epsilon_2 + \ldots + n_M\epsilon_M\right]\right)\right) \\ &= Q_N^{'} + Q_{N-1}^{'} e^{-\beta\epsilon_1} \end{aligned}$$

where the prime means “remove energy state 1 from the sums.”

The grand partition function is

$$\begin{aligned}\mathcal{Q} = \sum_{N=0}^{\infty} z^N Q_N &= \sum_{N=0}^{\infty} z^N \left(Q'_N + Q'_{N-1} e^{-\beta\epsilon_1}\right)\\ &= \sum_{N=0}^{\infty} z^N Q'_N + e^{-\beta\epsilon_1} z \sum_{N=1}^{\infty} z^{N-1} Q'_{N-1}\\ &= \left(1 + z e^{-\beta\epsilon_1}\right) \sum_{N=0}^{\infty} z^N Q'_N\end{aligned}$$

or

$$\mathcal{Q} = \left(1 + z e^{-\beta\epsilon_1}\right) \mathcal{Q}'$$

We can play the same game and pull out energy state 2 and 3 and so on. Ultimately we arrive at

$$\begin{aligned}\mathcal{Q} &= \left(1 + z e^{-\beta\epsilon_1}\right)\left(1 + z e^{-\beta\epsilon_2}\right)\ldots\left(1 + z e^{-\beta\epsilon_M}\right)\\ &= \prod_{i=1}^{M}\left(1 + z e^{-\beta\epsilon_i}\right)\end{aligned}$$

which is our grand partition function for fermions.

Boson Grand Partition Function (§6.2)

Our derivation here will follow the same lines as the formulation for fermions. We start with the canonical partition function,

$$\begin{aligned}Q_N &= \sum_{n_1=0}^{N*}\sum_{n_2=0}^{N*}\ldots\sum_{n_m=0}^{N*} \exp\left(-\beta\left[n_1\epsilon_1 + \ldots + n_M\epsilon_M\right]\right)\\ &= \left(\sum_{n_1=0}^{N} e^{-\beta n_1\epsilon_1}\right)\left(\sum_{n_2=0}^{N*}\ldots\sum_{n_M=0}^{N} \exp\left(-\beta\left[n_2\epsilon_2 + \ldots + n_M\epsilon_M\right]\right)\right)\end{aligned}$$

Note that for bosons the sum goes from 0 to N instead of from 0 to 1. Again the $*$ means that the sums are restricted by the condition $n_1 + n_2 + \ldots + n_M = N$.

As before, we explicitly expand the n_1 summation to get

$$Q_N = Q'_N + e^{-\beta\epsilon_1} Q'_{N-1} + e^{-2\beta\epsilon_1} Q'_{N-2} + \ldots$$

The grand partition function is

$$\begin{aligned}\mathcal{Q} &= \sum_{N=0}^{\infty} z^N Q_N\\ &= \sum_{N=0}^{\infty} z^N Q'_N + z e^{-\beta\epsilon_1}\sum_{N=1}^{\infty} z^{N-1} Q'_{N-1} + z^2 e^{-2\beta\epsilon_1}\sum_{N=2}^{\infty} z^{N-2} Q'_{N-2} + \ldots\\ &= \left(1 + z e^{-\beta\epsilon_1} + z^2 e^{-\beta\epsilon_1} + \ldots\right)\sum_{N=0}^{\infty} z^N Q'_N\\ &= \left(1 - z e^{-\beta\epsilon_1}\right)^{-1}\sum_{N=0}^{\infty} z^N Q'_N\end{aligned}$$

where we used the identity

$$\frac{1}{1-a} = \sum_{i=0}^{\infty} a^i = 1 + a + a^2 + \dots \qquad |a| < 1$$

Next we would apply the same procedure to strip out energy state 2 then 3 and so forth. Ultimately, we have

$$\begin{aligned} \mathcal{Q} &= \left(1 - ze^{-\beta\epsilon_1}\right)^{-1} \left(1 - ze^{-\beta\epsilon_2}\right)^{-1} \dots \left(1 - ze^{-\beta\epsilon_M}\right)^{-1} \\ &= \prod_{i=1}^{M} \left(1 - ze^{-\beta\epsilon_i}\right)^{-1} \end{aligned}$$

which is our expression for the grand partition function for bosons.

General Results for Fermi and Bose Gases (§6.2–6.3)

Before studying fermion and boson systems in detail, let's obtain some general results which apply to both types of particles.

We may write the grand partition function as

$$\mathcal{Q} = \prod_{i=1}^{M} \left(1 + \sigma ze^{-\beta\epsilon_i}\right)^{\sigma}$$

where

$$\sigma = \begin{cases} +1, & \text{for fermions} \\ -1, & \text{for bosons} \end{cases}$$

and M is the number of energy states (which can be infinite). From this

$$\ln \mathcal{Q} = \sigma \sum_{i=1}^{M} \ln\left(1 + \sigma ze^{-\beta\epsilon_i}\right)$$

The average number of particles in the system is

$$\begin{aligned} N &= z\left(\frac{\partial}{\partial z}\right)_{\beta} \ln \mathcal{Q} = z\sigma \sum_{i=1}^{M} \frac{\sigma e^{-\beta\epsilon_i}}{1 + \sigma ze^{-\beta\epsilon_i}} \\ &= \sum_{i=1}^{M} \left(z^{-1}e^{\beta\epsilon_i} + \sigma\right)^{-1} \qquad \left(\text{Note: } \sigma^2 = 1\right) \end{aligned}$$

Yet the average number of particles in an energy state, $\overline{n}_i$, is related to N as

$$N = \sum_{i=1}^{M} \overline{n}_i$$

so

$$\overline{n}_i = \frac{1}{z^{-1}e^{\beta\epsilon_i} + \sigma}$$

for fermions and bosons.

Some points: 1) Sometimes $\overline{n}_i$ is written as $\langle n_i \rangle$; 2) For fermions $0 \le \overline{n}_i \le 1$ while for bosons $0 \le \overline{n}_i \le N$; 3) In both cases, z is fixed by the above relation for N; 4) In the classical (high T) limit $z \ll 1$ so $\overline{n}_i \approx \exp(-\beta\epsilon_i)$, which gives the Maxwell-Boltzmann distribution.

The energy is

$$U = -\left(\frac{\partial}{\partial\beta}\right)_z \ln\mathcal{Q} = \sum_{i=1}^{M} \frac{\epsilon_i}{z^{-1}e^{\beta\epsilon_i}+\sigma} = \sum_{i=1}^{M} \overline{n}_i\epsilon_i$$

Many of our calculations for Fermi and Bose ideal gases will involve either the above expression for N or the expression for U.

Fermi and Bose Integrals (§7.1, 8.1, Appendices D,E)

Lecture 13

We now have the grand partition function for quantum ideal gases,

$$\mathcal{Q} = \prod_{i=1}^{M}\left(1+\sigma z e^{-\beta\epsilon_i}\right)^{\sigma}$$

where

$$\sigma = \begin{cases} +1, & \text{for fermions} \\ -1, & \text{for bosons} \end{cases}$$

We also have

$$\frac{PV}{kT} = \ln\mathcal{Q} = \sigma\sum_{i=1}^{M}\ln\left(1+\sigma z e^{-\beta\epsilon_i}\right)$$

and

$$N = z\left(\frac{\partial}{\partial z}\right)_\beta \ln\mathcal{Q} = \sum_{i=1}^{M} z e^{-\beta\epsilon_i}/\left(1+\sigma z e^{-\beta\epsilon_i}\right)$$

and

$$U = -\left(\frac{\partial}{\partial\beta}\right)_z \ln\mathcal{Q} = -\left(\frac{\partial}{\partial\beta}\right)_z \frac{PV}{kT}$$

We could write other thermodynamic variables in terms of $\ln\mathcal{Q}$ but the above are the most useful.

While the above is complete, the sums are clumsy to work with. For this reason we will convert from sums over energy states (taking $M\to\infty$) to integrals over phase space,

$$\begin{aligned} \sum_{i=1}^{M} &\to \frac{g_d}{h^3}\int\vec{r}\int d\vec{p} \\ &= \frac{g_d}{h^3}(V)\left(\int_0^{2\pi} d\phi\int_0^{\pi} d\theta\sin\theta\int_0^{\infty} dp\,p^2\right) \\ &= \frac{4\pi g_d}{h^3}V\int_0^{\infty} dp\,p^2 \end{aligned}$$

where g_d is the degeneracy factor (number of states that have the same energy). Since our only integration is over momentum, we rewrite the energy of a state using $\epsilon = p^2/2m$ (which is legit as long as our particles do not move at relativistic speeds).

We are not reverting back to classical mechanics even though the notation is similar. Physical systems have many particles so there are a large number of states; the only approximation that we are making here to assume that systems have so many states that we can replace the sums with integrals. Our previous classical limit was a high temperature approximation.

Our expressions are now

$$\frac{P}{kT} = g_d \frac{4\pi}{h^3}\sigma \int_0^\infty dp\, p^2 \ln\left(1 + \sigma z e^{-\beta p^2/2m}\right)$$

and

$$\frac{N}{V} = g_d \frac{4\pi}{h^3} \int_0^\infty dp\, p^2 \frac{z e^{-\beta p^2/2m}}{1 + \sigma z e^{-\beta p^2/2m}}$$

Our previous classical approximation, taking $z \ll 1$ and using $\ln(1+x) \approx x$, we would write these expressions as

$$\begin{aligned}\frac{P}{kT} &\approx g_d \frac{4\pi}{h^3}\sigma \int_0^\infty dp\, p^2 \sigma z e^{-\beta p^2/2m} \\ &= \frac{g_d z}{\lambda^3} \qquad \text{(Classical approximation)}\end{aligned}$$

which agrees with our earlier result, $\ln \mathcal{Q} = PV/kT = zV/\lambda^3$, for nondegenerate ($g_d = 1$) energy states. Similarly, the classical approximation gives

$$\begin{aligned}\frac{N}{V} &\approx g_d \frac{4\pi}{h^3}\sigma \int_0^\infty dp\, p^2 \sigma z e^{-\beta p^2/2m} \\ &= \frac{g_d z}{\lambda^3} \qquad \text{(Classical approximation)}\end{aligned}$$

which using the previous equation gives $PV = NkT$.

We introduce the Fermi integrals

$$\begin{aligned}f_{5/2}(z) &= \frac{4}{\sqrt{\pi}} \int_\circ^\infty dx\, x^2 \ln\left(1 + z e^{-x^2}\right) \\ &= \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell+1} z^\ell}{\ell^{5/2}} \\ f_{3/2}(z) &= z \frac{\partial}{\partial z} f_{5/2}(z) \\ &= \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell+1} z^\ell}{\ell^{3/2}}\end{aligned}$$

For fermions, we thus have

$$\begin{aligned}\frac{P}{kT} &= \frac{g_d}{\lambda^3} f_{5/2}(z) \\ \frac{N}{V} &= \frac{g_d}{\lambda^3} f_{3/2}(z)\end{aligned}$$

which are equations (8.1.3) and (8.1.4) in Pathria (see pg. 231).

Similarly, we introduce the Bose integrals,

$$\begin{aligned}g_{5/2}(z) &= -\frac{4}{\sqrt{\pi}} \int_0^\infty dx\, x^2 \ln\left(1 - z e^{-x^2}\right) \\ &= \sum_{\ell=1}^{\infty} \frac{z^\ell}{\ell^{5/2}} \\ g_{3/2}(z) &= z \frac{\partial}{\partial z} g_{5/2}(z) \\ &= \sum_{\ell=1}^{\infty} \frac{z^\ell}{\ell^{3/2}}\end{aligned}$$

so for bosons,

$$\begin{aligned}\frac{P}{kT} &= \frac{g_d}{\lambda^3} g_{5/2}(z) \\ \frac{N}{V} &= \frac{g_d}{\lambda^3} g_{3/2}(z)\end{aligned}$$

Careful: don't confuse g_d with the Bose integrals; sorry but this is the standard notation.

If you compare the above with equations (7.1.5) and (7.1.6) in Pathria (pg. 180) you will find a discrepancy. We should actually write

$$\begin{aligned}\frac{P}{kT} &= \frac{1}{\lambda^3} g_{5/2}(z) - \frac{1}{V}\log(1-z) \\ \frac{N}{V} &= \frac{1}{\lambda^3} g_{3/2}(z) + \frac{1}{V}\frac{z}{1-z}\end{aligned}$$

The last term in each expression is a correction required for very low temperature problems. Since the origin of these corrections is intellectually interesting we return to this question later and consider it in some depth.

We may also write the internal energy in terms of the above integrals. Using

$$U = -\left(\frac{\partial}{\partial\beta}\right)_z \ln\mathcal{Q} = -\left(\frac{\partial}{\partial\beta}\right)_z \frac{PV}{kT}$$

Since $d\lambda/d\beta = \lambda/2\beta$, from our previous results,

$$U = \frac{3}{2} g_d V \frac{kT}{\lambda^3} \times \begin{cases} f_{5/2}(z), & \text{(Fermi)} \\ g_{5/2}(z), & \text{(Bose)} \end{cases}$$

We also find

$$U = \frac{3}{2}PV \qquad \text{(both Bose and Fermi)}$$

a result which is also true for the classical monatomic ideal gas.

Fermi Energy

We now consider a Fermi gas in some detail. Using the grand canonical ensemble, we already determined that

$$\frac{N}{V} = \frac{g_d}{\lambda^3} f_{3/2}(z)$$

where $\lambda = h/\sqrt{2\pi mkT}$ is the thermal wavelength and $z = e^{\beta\mu}$ is the fugacity. You should view this expression as an equation to obtain z, or equivalently, to give the chemical potential μ, given T, V, and N.

The Fermi function $f_{3/2}(z)$ may be written as

$$f_{3/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^\infty dx \frac{x^2}{z^{-1}e^{x^2}+1} = \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell+1} z^\ell}{\ell^{3/2}}$$

Figure 5.1 shows a sketch of this function. We will be interested in the approximations

$$f_{3/2}(z) \approx \begin{cases} z, & z \ll 1 \\ \frac{4}{3\sqrt{\pi}}(\log z)^{3/2}, & z \gg 1 \end{cases}$$

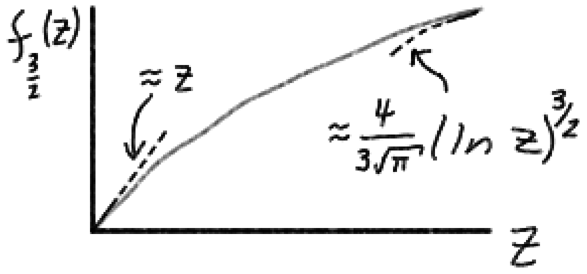

Figure 5.1: Graph of $f_{3/2}(z)$ indicating low and high z regions.

From $N\lambda^3/g_dV = f_{3/2}(z)$ we see that in the classical limit $(\lambda \to 0)$ then $f_{3/2}(z) \ll 1$ so $z \ll 1$. In the opposite case, $\left(\lambda^3 \gg V/N\right)$ we have $f_{3/2}(z) \gg 1$ so $z \gg 1$. Let's consider these separately.

Classical limit When the thermal de Broglie wavelength λ is small, then $f_{3/2}(z) \approx z$ so

$$z = \frac{N}{g_d}\frac{\lambda^3}{V}$$

The average occupation number for state i becomes

$$\langle n_i \rangle = \frac{1}{z^{-1}e^{\beta\epsilon_i}+1} \approx \frac{1}{z^{-1}e^{\beta\epsilon_i}} = \frac{N}{g_d}\frac{\lambda^3}{V}e^{-\beta\epsilon_i}$$

which is the classical Maxwell-Boltzmann distribution. In the classical limit $f_{5/2}(z) \approx z$ so we also recover the classical ideal gas equation of state $PV = NkT$.

Strong Quantum Limit Now we consider the opposite limit where the thermal de Broglie wavelength λ is much larger than the particle separation $\sqrt[3]{V/N}$. In this limit $z \gg 1$ and

$$\frac{N}{g_d}\frac{\lambda^3}{V} \approx \frac{4}{3\sqrt{\pi}}(\log z)^{3/2}$$

or

$$z \approx e^{\beta\mu_0}$$

where

$$\mu_0 \equiv \frac{\hbar^2}{2m}\left(\frac{6\pi^2 N}{g_d V}\right)^{2/3}$$

When $\lambda \to \infty$ (or $T \to 0$) then $z \to e^{\beta\mu_0}$ so μ_0 is the chemical potential at $T = 0$. A common notation is $\epsilon_F \equiv \mu_0$ where ϵ_F is called the *Fermi energy.* The energy level at the Fermi energy is called the Fermi level.

As $T \to 0$, the average occupation number is

$$\langle n_i \rangle = \frac{g_d}{z^{-1}e^{\beta\epsilon_i}+1} = \frac{g_d}{e^{-\beta\epsilon_F}e^{\beta\epsilon_i}+1} = \frac{g_d}{e^{\beta(\epsilon_i-\epsilon_F)}+1}$$

Yet as $T \to 0$, $\beta = 1/kT \to \infty$ so

$$\lim_{T\to 0} e^{\beta(\epsilon_i-\epsilon_F)} = \begin{cases} \infty, & \epsilon_i > \epsilon_F \\ 0, & \epsilon_i < \epsilon_F \end{cases}$$

and thus

$$\lim_{T\to 0} \langle n_i \rangle = \begin{cases} 0, & \epsilon_i > \epsilon_F \\ g_d, & \epsilon_i < \epsilon_F \end{cases}$$

In other words, as $T \to 0$, all states below the Fermi level are filled and all states above the Fermi level are empty. The particles filling the energy levels below ϵ_F are called the "Fermi sea" (a better term would be the "Fermi Seventh Street Garage During the First Week of Classes").

From the Fermi energy we can define a *Fermi temperature* as $\epsilon_F = kT_F$ or $T_F = \epsilon_F/k$. If $T \gg T_F$, then the system is in the classical regime. However, for most fermi gases of interest, $T \ll T_F$ so they are in the strong quantum limit.

It is easy to understand why the chemical potential at $T = 0$ equals the Fermi energy. At $T = 0$, every fermion is in its lowest possible energy level and stacked, due to Fermi exclusion, up to the Fermi level. The entropy is zero at $T = 0$; there is only one possible state for the system. Now add a single particle; this particle has to be place on the stack at the next available energy level. This means that the particle will have infinitesimally more energy than the Fermi energy. But the entropy remains zero so the energy of the system increases by ϵ_F thus the chemical potential

$$\mu = \left(\frac{\partial U}{\partial N}\right)_{S,V} = \epsilon_F \qquad (T = 0)$$

It turns out that $\mu \approx 0$ for $T \approx T_F$ and the chemical potential is negative when $T \gg T_F$, which was already pointed out for the classical ideal gas.

The chemical potential at $T = 0$ is $\mu_0 = \epsilon_F$; to obtain μ for $T > 0$, we need to keep the next order in the expansion for $f_{3/2}(z)$. Specifically,

$$\begin{aligned} f_{3/2}(z) &= \frac{4}{3\sqrt{\pi}} (\log z)^{3/2} \left[1 + \frac{\pi^2}{8} (\log z)^{-2} + \ldots\right] \\ &\approx \frac{4}{3\sqrt{\pi}} (\beta\mu)^{3/2} \left[1 + \frac{\pi^2}{8} (\beta\mu)^{-2}\right] \qquad (z \gg 1) \end{aligned}$$

Using this new approximation we get,

$$\frac{N\lambda^3}{g_d V} = \frac{4}{3\sqrt{\pi}} (\beta\mu)^{3/2} \left[1 + \frac{\pi^2}{8} (\beta\mu)^{-2}\right]$$

or

$$(\beta\mu)^{3/2} = \frac{3\sqrt{\pi}}{4} \frac{N\lambda^3}{g_d V} \left[1 + \frac{\pi^2}{8} (\beta\mu)^{-2}\right]^{-1}$$

or

$$\begin{aligned} \mu &= \frac{1}{\beta} \left(\frac{3\sqrt{\pi}}{4} \frac{N\lambda^3}{g_d V}\right)^{2/3} \left[1 + \frac{\pi^2}{8} (\beta\mu)^{-2}\right]^{-2/3} \\ &= \epsilon_F \left[1 + \frac{\pi^2}{8} (\beta\mu)^{-2}\right]^{-2/3} \end{aligned}$$

Only one hitch, we have μ in the correction term on the r.h.s.. However, since that term is the first order correction, we set $\mu = \mu_0 = \epsilon_F$ on the r.h.s. and get

$$\begin{aligned} \mu &\approx \epsilon_F \left[1 + \frac{\pi^2}{8} (\beta\epsilon_F)^{-2}\right]^{-2/3} \\ &\approx \epsilon_F \left[1 - \frac{\pi^2}{12} (\beta\epsilon_F)^{-2}\right] \\ &= \epsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F}\right)^2\right] \\ &= \epsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{T}{T_F}\right)^2\right] \end{aligned}$$

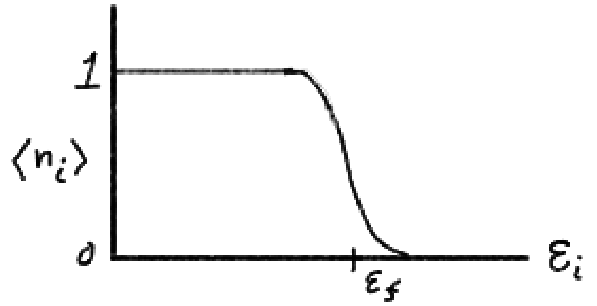

Figure 5.2: Average occupation number for a fermi gas at $T > 0$, $T \ll T_F$.

which gives μ for $T > 0$. Since it is common that $T \ll T_F$, we rarely need a second order correction.

At low (but non-zero) temperatures the average occupation number is,

$$\langle n_i \rangle = \frac{1}{e^{\beta(\epsilon_i - \mu)} + 1}$$

which is sketched in Fig. 5.2.

From the above, we can obtain the internal energy,

$$U = \frac{3}{5} N \epsilon_F \left[1 + \frac{5}{12} \pi^2 \left(\frac{T}{T_F} \right)^2 \right]$$

Notice that at $T = 0$ the average energy per particle is $\frac{3}{5}\epsilon_F$, which is the average ground state energy.

The heat capacity is

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = \frac{\pi^2}{2} N k \frac{T}{T_F}$$

the specific heat per mole (set $N = N_A$ and use $R = N_A k$) is

$$c_V = \frac{\pi^2}{2} R \frac{T}{T_F}$$

This result resolves an old mystery that tortured Boltzmann. Recall that the law of Dulong and Petit which said that due to molecular vibrations the specific heat per mole of a solid is $3R$. Yet it is strange that this result applies for both insulators and metals. After all metals have conduction electrons which should contribute 3 extra degrees of freedom (due to their translational kinetic energy). For metals we would expect the molar heat capacity to be,

$$c_V = 3R + \frac{3}{2} R$$

where the first term is the contribution from the lattice and the second is the contribution from the conduction electrons. But this result is not correct because conduction electrons do not behave as a classical ideal gas but rather as a Fermi ideal gas. The correct expression for the molar heat capacity of a metal is,

$$c_V = 3R + \frac{\pi^2}{2} R \frac{T}{T_F}$$

For copper $T_F \approx 50,000^\circ$C so it is not surprising that the conduction electrons have a negligible effect on the heat capacity.

An alternative way to view this phenomenon is to consider the other definition of heat capacity

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

As temperature increases, only a few particles rise above their ground states. Most particles remain in the ground state so long as $T \ll T_F$. For this reason, the number of accessible states Γ (and consequently the entropy S) increases very slowly. Since $\left(\frac{\partial S}{\partial T}\right)_V$ is small, so is C_V.

Finally, the pressure in a Fermi gas is

$$P = \frac{2}{3}\frac{U}{V} = \frac{2}{5}\frac{N\epsilon_F}{V}\left[1 + \frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2\right]$$

Notice that even when $T = 0$, the pressure is nonzero. Again this is due to the ground state energy of the particles. In the next section we consider an example in which this pressure supports a star and prevents its gravitational collapse.

White Dwarf Stars (§8.5)

Lecture 14

We now consider an interesting example of a Fermi gas. White dwarf stars have long been recognized as unusual in that they are exceptionally faint given their color. It is well established that this is due to their old age; in a white dwarf most of the hydrogen has been fused into heavier elements (outer shell is helium while the core is mostly carbon, oxygen and nughat).

From various types of measurements, we know that a typical white dwarf has a density of $\rho \approx 10^7 \mathrm{g/cm}^3$ (10 million times the density of water) and temperature $T \approx 10^7$K (about the temperature of the sun). Interestingly, because the density is high, the Fermi temperature of the electrons is $T_F \approx 10^{11}$ K. Since $T \ll T_F$ almost all the electrons are in their ground state. The star's temperature is relatively "cold" and we can treat the electrons as a Fermi gas in its ground state.

A star is an interesting container of gas. It has no outer wall but rather it relies on gravitational attraction to contain its gas. We may view the gravitational force as exerting a pressure, just as the walls of an ordinary container.

The question we now want to answer is: knowing the theory for a Fermi gas, what is the radius of a white dwarf given its mass? That is, at what radius does the pressure of the Fermi gas balance the gravitational force?

First, let's derive the results we need for a Fermi gas of electrons, (Note: we could work out the pressure contribution for the helium nuclei but would find it to be negligible compared with the electrons).

Unfortunately, our derivation for a Fermi gas assumed nonrelativistic dynamics (i.e., we used $\epsilon = \frac{1}{2}mv^2 = p^2/2m$). However, we must now use the relativistic expression

$$\epsilon = \sqrt{(pc)^2 + (mc^2)^2}$$

Fortunately, we can assume that all the electrons are in their ground state (since $T \ll T_F$).

The total internal energy is then

$$U = g_d \sum_i^{states} \epsilon_i n_i$$

with the number of electrons in state i being

$$n_i = \begin{cases} 1, & \text{if } \epsilon_i < \epsilon_F \\ 0, & \text{if } \epsilon_i > \epsilon_F \end{cases}$$

Electrons are spin $\frac{1}{2}$ so each energy state (level) can contain two electrons, a spin up and a spin down, so $g_d = 2$.

We define a Fermi momentum p_F as the momentum of a particle with Fermi energy. Since every state up to the Fermi level is filled,

$$\begin{aligned} N &= 2\sum_{\vec{r}}\sum_{\vec{p}\le\vec{p}_F} 1 \\ &= 2\left(\frac{1}{h^3}\int d\vec{r}\int d\vec{p}\right) \\ &= \frac{2V}{h^3}(4\pi)\int_0^{p_F} dp\, p^2 = \frac{8\pi V}{h^3}\frac{p_F^3}{3} \end{aligned}$$

so

$$p_F = \hbar\left(\frac{3\pi^2 N}{V}\right)^{1/3}$$

The total energy in the gas is found by summing up the energy of all the particles,

$$U = 2\sum_{\vec{r}}\sum_{p<p_F}\epsilon\left(\vec{r},\vec{p}\right)$$

As usual, we replace the sum over states with an integral over states by integrating over position and momentum,

$$\begin{aligned} U &= 2\frac{1}{h^3}\int\limits_V d\vec{r}\int\limits_{p<p_F} d\vec{p}\sqrt{\left(pc\right)^2 + \left(mc^2\right)^2} \\ &= 2\frac{1}{h^3}\left(V\right)\left(4\pi\right)\int\limits_0^{p_F} dp\, p^2\sqrt{\left(pc\right)^2 + \left(mc^2\right)^2} \end{aligned}$$

or

$$U = \frac{m^4c^5}{\pi^2\hbar^3}V f\left(x_F\right)$$

where $x_F \equiv p_F/mc$ and

$$f(x) \equiv \int_0^x dy\, y^2\sqrt{1+y^2} = \frac{1}{8}\sqrt{x^2+1}(2x^3+x) - \frac{1}{8}\sinh^{-1}(x)$$

Note that $f(x)$ is no a particularly interesting looking function; it basically looks like x^a where a is a small integer (around 2 to 5 depending on whether $x \to 0$ or $x \to \infty$).

The reason we wanted to get U is that we can get the pressure using

$$P = -\left(\frac{\partial U}{\partial V}\right)_S$$

Note: Entropy is fixed because all electrons are in the ground state. Putting in our expression for U,

$$\begin{aligned} P &= -\left(\frac{m^4c^5}{\pi^2\hbar^3}\right)\left[f\left(x_F\right) + \frac{\partial f}{\partial x_F}\left(\frac{\partial x_F}{\partial V}\right)_S\right] \\ &= \left(\frac{m^4c^5}{\pi^2\hbar^3}\right)\left[\frac{1}{3}x_F^3\sqrt{1+x_F^2} - f\left(x_F\right)\right] \end{aligned}$$

This gives us pressure in terms of x_F but we would prefer to know in terms of M, the star's mass and R, the star's radius. Yet it is simple to write x_F in terms of M and R since

$$x_F = \frac{p_F}{mc} = \frac{\hbar}{mc}\left(\frac{3\pi^2 N}{V}\right)^{1/3} = \frac{\overline{M}^{1/3}}{\overline{R}}$$

where

$$\overline{R} = \frac{R}{\hbar/mc}$$

is a dimensionless radius and

$$\overline{M} = \frac{9\pi}{8}\frac{M}{m_p}$$

is a dimensionless mass. We've used the fact that the star's mass, M, is related to the number of electrons, N, as

$$M \approx 2m_p N$$

where m_p is the mass of a proton. That is for each electron there is approximately one proton and one neutron (which have about the same mass). All miscellaneous constants are stuffed into $\overline{M}$ and $\overline{R}$ to keep things tidy.

Our expression for the pressure remains complicated. However, in the two approximations

$$P \approx \frac{4}{5}K x_F^5 \qquad (x_F \ll 1)$$

and

$$P \approx K\left(x_F^4 - x_F^2\right) \qquad (x_F \gg 1)$$

where

$$K = \frac{mc^2}{12\pi}\left(\frac{mc}{\hbar}\right)^3$$

we use these expressions in a moment.

The gravitational force of the star must balance the pressure of the Fermi gas in order to contain the gas. To do this right, the calculation requires solving some partial differential equations. Yet we can use a simple approximation for the pressure exerted by gravity.

$$\begin{aligned}
\text{(Force)} &= -\frac{\partial}{\partial R}\,\text{(gravitational potential energy)} \\
P \times \ \text{(surface area)} &= -\frac{\partial}{\partial R}\left(-\frac{GM^2}{R}\right) \\
P\left(4\pi R^2\right) &= \frac{GM^2}{R^2}
\end{aligned}$$

so

$$P = \frac{1}{4\pi}\frac{GM^2}{R^4}$$

Comparing this with our previous expression for pressure allows us to solve for radius in terms of mass. We find,

1) In the low density limit ($x_F \ll 1$),

$$\overline{R} \propto \overline{M}^{-1/3}$$

Notice that as mass increases, radius decreases.

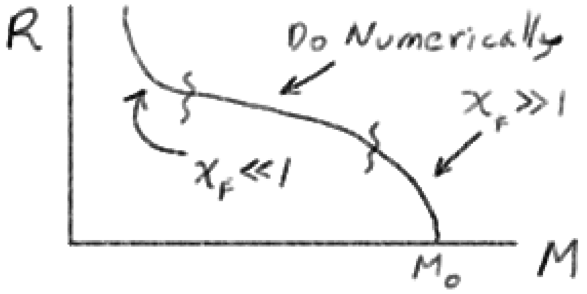

Figure 5.3: Radius as a function of mass for a white dwarf star.

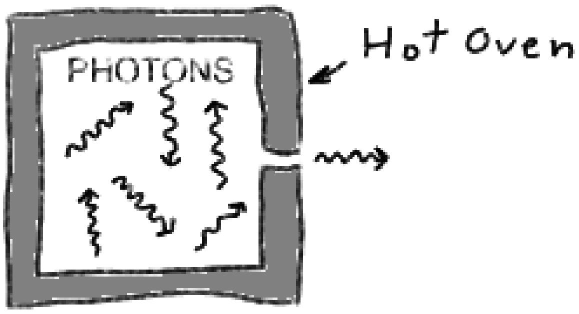

Figure 5.4: Blackbody kiln.

2) In the high density limit ($x_F \gg 1$)

$$\overline{R} \propto \overline{M}^{-2/3} \sqrt{1 - (M/M_0)}$$

where the constant M_0 turns out to be about 1.44 times the mass of the sun.

The graph of radius versus mass is sketched in Fig. 5.3.

Notice that the theory predicts that if the mass $M > M_0$, the Fermi gas pressure cannot hold up the star; this is called the Chandrasekhar limit. [*] If a star's mass exceeds M_0, it does not necessarily collapse into a black hole—internuclear forces can hold up the star in which case it becomes a neutron star.

Photon Gas (§7.2)

Lecture 15

A photon gas is a unique boson gas because the particles have zero mass. The number of particles is not conserved—photons can spontaneously appear and disappear. The picture to have in mind is a kiln, as shown in Fig. 5.4.

The system is maintained at fixed temperature, T, since the walls are at fixed temperature. The photon system is like a canonical ensemble except N is not fixed. Nor is the average number of particles, $\langle N \rangle$, fixed (as in the grand canonical ensemble). Removing this constraint on particle number is equivalent to setting $\mu = 0$, i.e., the chemical potential for a photon gas is zero.

---

[*] S. Chandrasekhar, one of the great physicists of our times, worked out this theory at the age of 25.

The canonical partition function is

$$Q=\sum_{n_1=0}^{\infty}\sum_{n_i=0}^{\infty}\dots\sum_{n_M=0}^{\infty}\exp\left(-\beta E\left(n_1,n_2,\dots,n_M\right)\right)$$

where n_i is the number of photons in energy level i and E is the energy for the state $(n_1,\dots,n_M)$. Notice that the sums are *unrestricted* since particle number is unconstrained.

Call ϵ_i the energy of a photon in state i and we may write

$$\begin{aligned}
Q &= \sum_{n_1=0}^{\infty}\dots\sum_{n_M=0}^{\infty} e^{-\beta[n_1\epsilon_1+n_2\epsilon_2+\dots+n_M\epsilon_M]}\\
&= \left(\sum_{n_1=0}^{\infty} e^{-\beta n_1\epsilon_1}\right)\left(\sum_{n_2=0}^{\infty} e^{-\beta n_2\epsilon_2}\right)\dots\left(\sum_{n_M=0}^{\infty} e^{-\beta n_M\epsilon_M}\right)\\
&= \left(\frac{1}{1-e^{-\beta\epsilon_1}}\right)\left(\frac{1}{1-e^{-\beta\epsilon_2}}\right)\dots\left(\frac{1}{1-e^{-\beta\epsilon_M}}\right)\\
&= \prod_{i=1}^{M}\frac{1}{1-e^{-\beta\epsilon_i}}
\end{aligned}$$

so

$$\ln Q=-\sum_{i=1}^{M}\ln\left(1-e^{-\beta\epsilon_i}\right)$$

Notice how easy those unrestricted sums were. For photons there is no physical upper bound on the number of states so we may take $M=\infty$.

As usual, to get much farther we need to replace the sum over states into integrals over phase space,

$$\begin{aligned}
\sum^{states} &\to \sum^{spins}\frac{1}{h^3}\int_V d\vec{r}\int d\vec{p}\\
&= (2)\,\frac{1}{h^3}\,(V)\int d\vec{p}\\
&= \frac{8\pi V}{h^3}\int_0^{\infty} dp\,p^2
\end{aligned}$$

since there are two spin states (corresponding to left and right polarized).

One usually prefers characterizing the state of a photon by its frequency rather than its momentum so using $p=\epsilon/c=\hbar\omega/c$,

$$\frac{8\pi V}{h^3}\int_0^{\infty} dp\,p^2\to\frac{V}{\pi^2c^3}\int_0^{\infty} d\omega\,\omega^2$$

and

$$\begin{aligned}
\ln Q &= -\frac{V}{\pi^2c^3}\int_0^{\infty} d\omega\,\omega^2\ln\left(1-e^{-\beta\hbar\omega}\right)\\
&= \frac{\pi^2 V}{45c^3}\left(\frac{kT}{\hbar}\right)^3
\end{aligned}$$

The pressure may be found using $PV=kT\ln Q$, and

$$P=\frac{\pi^2}{45c^3\hbar^3}\left(kT\right)^4=\frac{4}{3}\sigma T^4$$

where σ is Stefan's constant.

The energy is

$$U = -\left(\frac{\partial}{\partial\beta}\ln Q\right)_V = 4\sigma V T^4$$

so $PV = \frac{1}{3}U$.

For our final result, go back a moment to our integral expression for $\ln Q$,

$$\ln Q = -\frac{V}{\pi^2 c^3}\int_0^\infty d\omega\,\omega^2 \ln\left(1 - e^{-\beta\hbar\omega}\right)$$

Using $U = -\left(\frac{\partial}{\partial\beta}\right)_V \ln Q$

$$U = \frac{V}{\pi^2 c^3}\int_0^\infty d\omega\,\omega^2 \frac{\hbar\omega e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}$$

or

$$\frac{U}{V} = \int_0^\infty d\omega\, u(\omega, T)$$

where

$$u(\omega, T) = \frac{\hbar}{\pi^2 c^3}\frac{\omega^3}{e^{\beta\hbar\omega} - 1}$$

is the energy per unit volume per unit frequency, i.e., the Planck radiation law (see Fig. 7.7 in Pathria).

By the correspondence principle, in the limit, $\hbar \to 0, e^{\beta\hbar\omega} \simeq 1 + \beta\hbar\omega$, we recover the classical Rayleigh-Jeans law

$$u\left(\omega, T\right) = \frac{kT}{\pi^2 c^3}\omega^2$$

The ultraviolet catastrophe ($u \to \infty$ as $\omega \to \infty$) in the classical result led to the birth of quantum mechanics.

Phonons in Solids (§7.3)

The sound modes in a solid can be analyzed in a fashion very similar to blackbody radiation. This similarity is suggested by the name "phonon" which is a quantized vibrational excitation.

Just as with photons, our partition function is

$$\begin{aligned}\ln Q &= -\sum_{i=1}^{M} \ln\left(1 - e^{-\beta\epsilon_i}\right) \\ &= -\sum_{i=1}^{M} \ln\left(1 - e^{-\beta\hbar\omega_i}\right)\end{aligned}$$

Yet there are some differences. First, the number of states (i.e., modes), M, is not infinite but rather is limited to $3N$, where N is the number of atoms in the solid. This limitation may be understood by the fact that the smallest wavelength of the sound modes is bounded by the finite separation between atoms.

As before, we want to replace the sum over states with an integral and this is as with photons,

$$\sum_{i=1}^{M} \to \frac{3}{2}\frac{V}{\pi^2 c^3}\int_0^{\omega_{\max}} d\omega\,\omega^2$$

with the following differences:

(1) The polarization vector for phonons has 3 directions (2 traverse + 1 longitudinal) instead of 2 (just transverse) as with photons. This produces the factor 3/2.

(2) The speed c is the sound speed instead of the speed of light

(3) The finite number of modes introduces a maximum frequency $\omega_{\max}$. This $\omega_{\max}$ is found using the condition,

$$\frac{3}{2}\frac{V}{\pi^2 c^3}\int_0^{\omega_{\max}} d\omega\, \omega^2 = 3N$$

or

$$\omega_{\max} = c\left(\frac{6\pi^2 N}{V}\right)^{1/3}$$

for the maximum frequency.

I should mention that we have made an approximation (first introduced by Debye) that the distribution of modes is the same as for photons. In reality, we should write

$$\sum_{i=1}^{M} \to \int_0^{\omega_{\max}} f(\omega)\, d\omega$$

where $f(\omega)$ is the density of states.

The partition function, using the Debye approximation, is

$$\ln Q = -\frac{3V}{2\pi^2 c^3}\int_0^{\omega_{\max}} \ln\left(1 - e^{-\beta\hbar\omega}\right)\omega^2\, d\omega$$

The internal energy is

$$\begin{aligned} U = -\left(\frac{\partial}{\partial\beta}\ln Q\right)_V &= \frac{3V}{2\pi^2 c^3}\int_0^{\omega_{\max}} d\omega \quad \omega^2 \frac{\hbar\omega}{e^{\beta\hbar\omega}-1} \\ &= 9N\frac{(kT)^4}{(\hbar\omega_{\max})^3}\int_0^{\beta\hbar\omega_{\max}} \frac{t^3}{e^t - 1}\, dt \end{aligned}$$

This last integral is tabulated (see Debye function). If we define a Debye temperature $T_0 = \hbar\omega_m/k$ then we have the approximations,

$$U = 3NkT \times \begin{cases} 1 - \frac{3}{8}\frac{T_0}{T} + \ldots, & T \gg T_0 \\ \frac{\pi^2}{5}\left(\frac{T}{T_0}\right)^3 + \ldots, & T \ll T_0 \end{cases}$$

Typically, $T_0 \approx 200K$ so at room temperature $U \approx 3NkT$ and $C_V = \left(\frac{\partial U}{\partial T}\right)_V \approx 3Nk = 3Rn$, where n is the number of moles. This last expression is the familiar law of Dulong and Petit. For low temperatures, $U \propto T^4$ so $C_V \propto T^3$, as seen experimentally.

Bose-Einstein Condensation (§7.1-2)

Lecture 16

We already worked out the grand partition function for bosons,

$$\mathcal{Q} = \prod_i^{states}\left(1 - ze^{-\beta\epsilon_i}\right)^{-1}$$

where ϵ_i is the energy of state i. We also have

$$\frac{PV}{kT} = \ln \mathcal{Q} = -\sum_i^{states} \ln\left(1 - ze^{-\beta\epsilon_i}\right)$$

$$N = z\frac{\partial}{\partial z}\ln \mathcal{Q} = \sum_i^{states} \frac{ze^{-\beta\epsilon_i}}{1 - ze^{-\beta\epsilon_i}}$$

Normally, we replace the sums over states with integrals over phase space

$$\sum_i^{states} \to \frac{4\pi}{h^3}V\int_0^\infty dp\, p^2$$

so, for example

$$\begin{aligned} N &= \frac{4\pi}{h^3}V\int_0^\infty dp\, p^2 \frac{ze^{-\beta p^2/2m}}{1 - ze^{-\beta p^2/2m}} \\ &= \frac{V}{\lambda^3} g_{3/2}(z) \end{aligned}$$

where the function $g_{3/2}(z)$ is the Bose integral we met before.

I already mentioned that, in some cases, this replacement of the sum over states was not a good approximation and had to be modified. Let's consider why.

An aside: For bosons, $0 \le z < 1$. You can prove this by considering the average occupation number

$$\overline{n}_i = \frac{1}{z^{-1}e^{\beta\epsilon_i} - 1} = \frac{z}{e^{\beta\epsilon_i} - z}$$

For the ground state, $i = 1, \epsilon_1 = 0$, we have

$$\overline{n}_1 = \frac{z}{1 - z}$$

Since $0 \le n_1 < \infty$, the fugacity must be between 0 and 1. (Note: for fermions $0 \le z < \infty$.)

Consider the bose integral $g_{3/2}(z)$ in its summation form,

$$g_{3/2}(z) = \sum_{l=1}^\infty \frac{z^l}{l^{3/2}}$$

It is easy to show that

$$\max\left\{g_{3/2}(z)\right\} = g_{3/2}(1) = \zeta(3/2) \simeq 2.612$$

See Fig. 12.3 in Huang for a graph of $g_{3/2}(z)$.

Now for the paradox. Since the thermal wavelength is

$$\lambda = \frac{h}{\sqrt{2\pi mkT}}$$

then $\lambda \to \infty$ as $T \to 0$. Yet if

$$N = \frac{V}{\lambda^3} g_{3/2}(z)$$

then $N \to 0$ as $T \to 0$, that is, the number of particles in the system goes to zero as the temperature goes to zero. Weird. As one author (Robertson) puts it

> "... not even the magic of quantum mechanics can make particles disappear with cooling or reappear when warmed."

The key to resolving this paradox is to realize that at low temperatures, a significant fraction of the particles will occupy the ground state. We explicitly separate the first term from the sum over states, and replace the remaining sum with an integral, as before,

$$\begin{aligned} N &= \sum_{i=1}^{\infty} \frac{1}{z^{-1}e^{\beta\epsilon_i}-1} \\ &= \frac{1}{z^{-1}-1} + \sum_{i=2}^{\infty} \frac{1}{z^{-1}e^{\beta\epsilon_i}-1} \\ &= \frac{z}{1-z} + \frac{V}{\lambda^3} g_{3/2}(z) \\ &= N_{\text{gr}}(z) + N_{\text{ex}}(z) \end{aligned}$$

where $N_{\text{gr}}, N_{\text{ex}}$ are the number of particles in the ground state and non-ground states respectively. As we saw, the $N_{\text{ex}} \to 0$ as $T \to 0$ yet the first term, $N_{\text{gr}} = z/(1-z)$ can be arbitrarily large as $z \to 1$.

As we saw, the maximum value for $g_{3/2}(z)$ is $\zeta(3/2)$ ($\simeq 2.612$) so the maximum number of particles outside the ground state is

$$\max\{N_{\text{ex}}(z)\} = \frac{V}{\lambda^3}\zeta(3/2) = N_{\text{ex}}(z=1)$$

Recall that $\lambda \propto T^{-3/2}$; we define a critical temperature, T_c, by the condition,

$$N_{\text{ex}}(1) = N$$

or

$$\frac{V}{\lambda(T_c)^3}\zeta(3/2) = N$$

In other words, we ask: at what minimum temperature could we still fit all the particles into non-ground states? Solving the above for this temperature

$$kT_c = \left(\frac{h^2}{2\pi m}\right)\left(\frac{N}{V\zeta(3/2)}\right)^{2/3}$$

By the way, the value of kT_c is approximately the same as the Fermi energy so quantum effects appear in the same energy range for fermions and bosons.

The graph of N_{gr} and N_{ex} is shown in Fig. 5.5. Specifically, the functions are

$$N_{\text{gr}} = \begin{cases} N\left(1-(T/T_c)^{3/2}\right), & T < T_c \\ 0, & T > T_c \end{cases}$$

Notice that below $T = T_c$, particles start collecting in the ground state; this phenomenon is called Bose-Einstein condensation.

To obtain the other thermodynamic properties, we use the standard expressions but treat the ranges $T < T_c$ and $T > T_c$ separately:

- For $T < T_c$: The value of z in this case is *very* close to 1. For example, if the ground state contains a mere 10^6 particles then from

$$N_{\text{gr}} = \frac{z}{1-z}$$

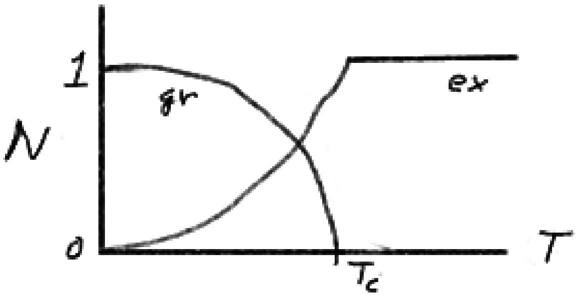

Figure 5.5: The number of particles in the ground state and in excited states as a function of temperature.

we have that $z \approx 0.999999$. Thus for $T < T_c$, we may set $z = 1$ everywhere *except* the above expression for N_{gr}.

- For $T > T_c$: We can take $N_{\text{gr}} = 0, N_{\text{ex}} = N$ so

$$N = \frac{V}{\lambda^3} g_{3/2}(z)$$

We must solve this expression to find z; probably need to do this numerically.

Some of the other thermodynamic results we obtain from the partition function include

$$U = \begin{cases} \frac{3}{2}kT\frac{V}{\lambda^3} g_{5/2}(z), & (T > T_c) \\ \frac{3}{2}kT\frac{V}{\lambda^3} g_{5/2}(1), & (T < T_c) \end{cases}$$

(see Fig. 7.10b in Pathria and Beale) and

$$C_V = \begin{cases} \frac{15R}{4}\frac{g_{5/2}(z)}{g_{3/2}(z)} - \frac{9R}{4}\frac{g_{3/2}(z)}{g_{1/2}(z)} & (T > T_c) \\ \frac{15R}{4}\frac{g_{5/2}(1)}{g_{3/2}(1)}\left(\frac{T}{T_c}\right)^{3/2} & (T < T_c) \end{cases}$$

Graphing the specific heat, we get a curve with a noticeable cusp; see Fig. 5.6 or Pathria and Beale's Fig. 7.11. Quantum effects in a Fermi gas of electrons are significant, even at room temperature, because the mass of an electron is small, giving it a large deBroglie wavelength. We don't have a similar situation for bosons. For a He^4 atom (spin 0), the critical temperature at which Bose-Einstein condensation would occur is $T_c = 2.18K$. Experimentally, at $T = 3.14K$, He^4 does have a phase transition into a superfluid phase. The specific heat near the transition is reminiscent of the ideal Bose gas result. Unfortunately, a liquid is *not* well approximated as an ideal gas; the interaction between atoms is not negligible. Thus our B-E condensation is only a first step toward understanding superfluid helium.

Bose-Einstein condensation is observed in the laboratory [†] and significant experimental work has been done in the field of quantum ideal gases in the past decade. Finally, white dwarfs are filled with helium nuclei and are so dense that the critical temperature is quite high $\left(T_c \approx 2.5 \times 10^6\text{K}\right)$. Maybe we have B-E condensation in their cores? There is a white dwarf nearby (8 light years away, near Sirius), we should go check this out.

[†] M.H. Anderson, J.R. Ensher, M.R. Matthews, C.E. Wieman, and E.A. Cornell, Science **269** 198 (1995).

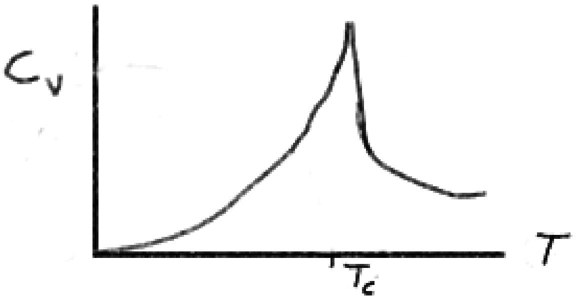

Figure 5.6: Heat capacity of an ideal boson gas as a function of temperature; notice cusp at $T = T_c$.

Chapter 6

Non-ideal Systems – Ising Model

Ideal vs. Non-ideal Systems §12.3

Lecture 17

Statistical mechanics has given us a general recipe for obtaining thermodynamic properties given the microscopic dynamics of a system. For example, using the canonical ensemble, one computes the partition function

$$Q_N = \sum_{i}^{\text{states}} e^{-\beta E_i}$$

and obtain all thermodynamics from this.

Thus far we have concentrated on ideal systems, that is, systems where the energy of each state was a sum over the energy of each particle

$$E_i = \epsilon_1(j') + \epsilon_2(j'') + \ldots + \epsilon_N(j' \cdots ')$$

and the energy of each particle is independent of the state of the other particles. For distinguishable particles, the partition function simplifies

$$Q_N = \underbrace{\left(\sum_{j}^{\text{states}} e^{-\beta\epsilon_1(j)} \right)}_{\text{Particle 1}} \times \ldots \times \underbrace{\left(\sum_{j}^{\text{states}} e^{-\beta\epsilon_N(j)} \right)}_{\text{Particle } N} = (Q_1)^N$$

where Q_1 is the partition function for a single particle. Sometimes counting states is a little more difficult (e.g., Fermi and Bose ideal gases) but still doable within one lecture.

In this chapter we consider the much more difficult problem of *non-ideal systems* for which the energy does not decompose as a simple sum of single particle energies because of interactions between the particles.

To give you a picture to have in mind, consider the following two examples:

Classical Non-ideal Gas In this system the energy of a single particle has both a kinetic and potential energy contribution,

$$\epsilon_i = \frac{1}{2m} p_i^2 + \sum_{k \neq i}^{N} \mathcal{U}(|\mathbf{q}_i - \mathbf{q}_k|)$$

where $\mathcal{U}$ is the potential energy due to the force between a pair of particles, which is here taken to only be a function of their relative separation. The total energy for a given state is

$$E(q_1, \ldots, q_N, p_1, \ldots, p_N) = \sum_{i=1}^{N} e_i(q_1, \ldots, q_N, p_i)$$

This system is not an ideal gas, that is,

$$E(q_1, \ldots, q_N, p_1, \ldots, p_N) \neq \sum_{i=1}^{N} e_i(q_i, p_i)$$

because the potential energy for particle j is a function of the positions of *all* the other particles.

Quantum Ising Spin System Consider a lattice of particles with spin $\frac{1}{2}$ for which the state of particle i is $s_i = +1$ (“up”) or -1 (“down”).[*] Two typical configurations could be pictured as

$$\begin{array}{cccc} + & + & - & + \\ - & + & - & - \\ + & + & - & + \\ + & - & + & - \end{array} \qquad\qquad \begin{array}{cccc} + & + & - & + \\ + & + & + & + \\ + & - & + & + \\ + & + & + & + \end{array}$$

In the left configuration the spins appear randomly oriented while on the right they are aligned to mostly point up. The state of the system is given by the set of quantum numbers $(s_1, \ldots, s_N)$.

The spins interact in this system and we write the total energy for a given state as

$$E(s_1, \ldots, s_N) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathcal{U}_{ij}(s_i, s_j) + \sum_{i=1}^{N} \mathcal{U}_i^H(s_i)$$

where $\mathcal{U}_{ij}$ is the potential energy due to the interaction between spins i and j; $\mathcal{U}_i^H$ is the potential energy of spin i in the presence of an external magnetic field H.

The Ising model simplifies this system even further to write the energy as

$$E(s_1, \ldots, s_N) = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i}^{\gamma} \epsilon s_i s_j - \sum_{i=1}^{N} H s_i$$

where the sum “$j \in \mathcal{N}_i$” is restricted to the spins that are “neighbors” of spin i; each spin has γ neighbors.[†] The interaction energy is ϵ so if $\epsilon > 0$ then the lower energy state is when spins are aligned (s_i and s_j both $+1$ or both -1). The potential energy due to the external field is also lower when H and s_i have the same sign (i.e., when the spin points in the same direction as the field).

The Ising model is a simplified representation of a ferromagnet. You can check that the energy cannot be written as the sum of independent terms, that is,

$$E(s_1, \ldots, s_N) \neq \sum_{i=1}^{N} e(s_i)$$

because of the nonlinear interaction between the spins. The Ising model is non-ideal system due to the coupling of a spin and its neighbors. Despite its simplicity it is a very rich theoretical model since it is one of the simplest non-ideal systems with a direct relation to a physical material.

---

[*] NOTATION: Pathria uses σ_i instead of s_i

[†] NOTATION: Pathria uses J instead of ϵ and μB instead of H

Approaches for Non-ideal Systems

Now that we've seen the general form and two examples of non-ideal systems, let's consider some ways of tackling these problems.

Brute Force Approach One way to proceed would be to numerically evaluate the partition function Q_N using its definition as a sum over states. This turns out to be impractical; let's see why.

Consider the Ising model for a small three-dimensional system with $10\times10\times10$ lattice sites ($N = 1000$ spins). The partition function is

$$\begin{aligned} Q_N &= \sum_i^{\text{states}} e^{-\beta E_i} \\ &= \sum_{s_1=-1}^{+1} \sum_{s_2=-1}^{+1} \cdots \sum_{s_N=-1}^{+1} \exp(-\beta E(s_1, s_2, \ldots, s_N)) \end{aligned}$$

Each sum has only two terms but there are 1000 sums so the total number of terms to evaluate is $2^{1000} \approx 10^{300}$.

Modern supercomputers can evaluate about 10^{15} operations per second (1 Petaflop). Even if each term required only one operation, this small system would require 10^{285} s $\approx 10^{275}$ years to evaluate Q_N for a *single value of* β!

Phenomenological Approach This approach can be generously termed "intelligent guessing" or cynically termed "fudging to fit the answer." The paradigm for this approach is the formulation of the van der Waals model for dense gases and liquids. Starting from the ideal gas, van der Waals proposed two corrections to account for the short-ranged repulsion force and the long-ranged attraction force. For many phenomena (e.g., behaviour near the gas-liquid critical point) the model is in good agreement with experimental data, indicating that it contains some elements of truth.

The phenomenological approach produces thermodynamic models that are only partially derived from statistical mechanics. This is less satisfying to the pure theorist but the models can be very handy for general applications. And sometimes the model can be derived from fundamental principles, as with the Kac-Uhlenbeck-Hemmer analysis of the van der Waals model, thought this usually takes decades (90 years for the van der Waals model).

Perturbation Approach If a system is close to ideal then the interactions can be added as a correction and solved by a perturbation expansion. For a non-dilute gas we can work out the virial coefficients, B_i, to formulate an equation of state of the form

$$\frac{PV}{NkT} = B_1(T) + B_2(T)\frac{N}{V} + B_3(T)\frac{N^2}{V^2} + \ldots$$

At low densities only the first few terms contribute. The calculation of these virial coefficients is rather tedious and composes the bulk of chapter 9 in Pathria. *Second quantizaton*, a specialized perturbation methodology for quantum gases, is described in Chapter 10 of Pathria. We will not cover either of these chapters.

Mean Field Approach In this approach we approximate the system to be a sum of noninteracting particles by replacing the interactions with a fake external field. The idea is that, on average, a particle feels a "mean field" of neighbors. But don't we need to know the distribution for the system to find this mean field? Yes, the calculation is circular—assume we know the answer, compute the answer, then use the answer to get the answer. This is best understood by applying the method, which is what is done in the next section.

Mean Field Solution for the Ising Model (§12.5)

Lecture 18

Recall that in the Ising model the state of the system is given by the set of quantum numbers $(s_1, \ldots, s_N)$ and the total energy

$$E(s_1, \ldots, s_N) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i}^{\gamma} \epsilon s_i s_j - \sum_{i=1}^{N} H s_i$$

where $\epsilon > 0$ is the interaction energy of the spins, γ is the number of neighboring sites that interact with each spin, and H is the component of the external magnetic field parallel to the spins.

Before formulating the mean field solution for the Ising model, let's discuss some general features we expect to find in the model. To start, take the external field $H = 0$ so we only have the interaction among the spins. The energy is lower when the spins are aligned so if we consider the two configurations:

$$\begin{array}{cccc} + & + & - & + \\ - & + & - & - \\ + & + & - & + \\ + & - & + & - \end{array} \qquad\qquad \begin{array}{cccc} + & + & - & + \\ + & + & + & + \\ + & - & + & + \\ + & + & + & + \end{array}$$

the configuration on the right has a lower energy than the one on the left. In the canonical ensemble the probability of a state goes as $\exp(-E/kT)$ so the state on the left is more probable.

Does this mean that we expect to find the spins aligned? Not necessarily. Again, considering the two configurations above, the one on the left has 9 up spins and 7 down spins while the configuration on the right has 14 up spins and 2 down spins. The number of combinations of $M = 9$ up spins chosen from $N = 16$ sites is ${}_N C_M = N!/M!(N-M)! = 11440$; the number of combinations for $M = 14$ up spins is 120. Though the configuration on the right is more probable there are many *more* configurations like the one on the left.

At low temperatures the "energy advantage" will highly favor configurations with aligned spins. At high temperatures the "entropy advantage" will favor configurations with random spins because there are so many of them. Experimentally, ferromagnets are observed to undergo a phase transition at the *Curie temperature*, T_c. Below the Curie temperature the ferromagnet can sustain a non-zero spontaneous magnetization but above T_c the material transitions from ferromagnetic to paramagnetic behavior. For iron $T_c = 1043$ K, making it a ferromagnet at room temperature. Finally, when $H \neq 0$ the spins are more likely to align with the field than against it but the general behavior discussed above is qualitatively unchanged.

Now for the mean field approximation in the Ising model. In general, the energy of a single spin, at site i, may be written as

$$e_i = -\epsilon s_i \sum_{j\in\mathcal{N}_i}^{\gamma} s_j - H s_i$$

Figure 6.1 illustrates an individual spin in a two-dimensional lattice with the neighbors being the four spins immediately adjacent to site i (so $\gamma = 4$). The energy for this single spin may be written as $e_i(s_i, s_{\text{South}}, s_{\text{West}}, s_{\text{East}}, s_{\text{North}})$, with the compass directions pointing to the four neighbors.

The mean field approximation replaces the specific values of the neighboring spins with an average spin, $\langle s \rangle$,[‡] so the energy of spin i is

$$\begin{aligned} e_i &= (-\epsilon s_i)(\gamma\langle s \rangle) - H s_i \\ &= -H^* s_i \end{aligned}$$

[‡] NOTATION: Pathria uses L instead of $\langle s \rangle$

$$\begin{array}{ccc} & \text{Original System} & \\ \cdot & s_{\text{North}} & \cdot \\ s_{\text{West}} & s_i & s_{\text{East}} \\ \cdot & s_{\text{South}} & \cdot \\ \multicolumn{3}{c}{e_i(s_i, s_{\text{South}}, s_{\text{West}}, s_{\text{East}}, s_{\text{North}})} \end{array} \qquad \begin{array}{ccc} & \text{Mean Field Approx.} & \\ \cdot & \langle s \rangle & \cdot \\ \langle s \rangle & s_i & \langle s \rangle \\ \cdot & \langle s \rangle & \cdot \\ \multicolumn{3}{c}{e_i(s_i, \langle s \rangle)} \end{array}$$

Figure 6.1: Schematic of the Ising system focusing on spin i in the original system and in the mean field approximation

where

$$H^* = H + \gamma\epsilon\langle s \rangle$$

is the "effective mean field" felt by spin i. The energy of site i is thus $e_i = \mp H^*$ for the two states $s_i = \pm 1$. Of course we don't know the average spin $\langle s \rangle$ yet but the strategy will be to formulate an expression for $\langle s \rangle$ in terms of $\langle s \rangle$.

For a single lattice site the partition function is

$$\begin{aligned} Q_1 &= \sum_{s_i=-1}^{+1} e^{-\beta e_i(s_i)} \\ &= e^{-\beta H^*} + e^{+\beta H^*} = 2\cosh(\beta H^*) \end{aligned}$$

The probability that site i is in state s_i is

$$\mathcal{P}(s_i) = \frac{1}{Q_1} e^{-\beta e_i(s_i)} = \frac{1}{Q_1} e^{\beta H^* s_i}$$

The average value is

$$\begin{aligned} \langle s_i \rangle &= \sum_{s_i=-1}^{+1} s_i \mathcal{P}(s_i) \\ &= (+1)\mathcal{P}(+1) + (-1)\mathcal{P}(-1) \\ &= \frac{1}{Q_1} e^{\beta H^*} - \frac{1}{Q_1} e^{-\beta H^*} \\ &= \frac{2}{Q_1} \sinh(\beta H^*) \\ &= \frac{\sinh(\beta H^*)}{\cosh(\beta H^*)} = \tanh(\beta H^*) \end{aligned}$$

But since there is nothing special about site i so we write $\langle s_i \rangle = \langle s \rangle$. Recall that $H^* = H + \gamma\epsilon\langle s \rangle$ so

$$\langle s \rangle = \tanh(\beta H + \beta\gamma\epsilon\langle s \rangle) \qquad (*)$$

which gives us a transcendental equation for $\langle s \rangle$.

We can get qualitative behavior of the solution of this equation by graphing the left and right sides of $(*)$ and seeing where the curves cross (i.e., where the two sides are equal). For $H = 0$ (no external magnetic field) we may write the left and right sides as

$$\mathcal{L}(\langle s \rangle) = \langle s \rangle \qquad \mathcal{R}(\langle s \rangle) = \tanh(\langle s \rangle T_c/T)$$

where $T_c = \gamma\epsilon/k$. These functions are sketched in Pathria Fig. 12.6; notice that at high temperatures ($T > T_c$) the only solution is $\langle s \rangle = 0$.

At low temperatures ($T \leq T_c$) there are three solutions, $\langle s \rangle = 0$, $+\bar{s}(T)$, and $-\bar{s}(T)$. Being a transcendental equation there is no simple analytic form for $\bar{s}(T)$ but using $\tanh(x) = x - x^3/3 + \ldots$ you can show that

$$\bar{s}(T) \approx \sqrt{3\left(1 - \frac{T_c}{T}\right)}$$

for $T \approx T_c$. For $T \ll T_c$, we find

$$\bar{s}(T) \approx 1 - 2e^{-2T_c/T}$$

The function $\bar{s}$ is sketched in Pathria Fig. 12.7, where it is shown to be in qualitative agreement with experimental data. Unfortunately the mean field approximation is too crude to give quantitative agreement.

Ising Model & Physical Systems

A variety of physical systems may be represented by the Ising model. Some examples are:

Ferromagnets It was already mentioned that the Ising model was originally conceived as a simple representation of a ferromagnet. What may not be obvious is that the interaction energy between the spins is *not* due to magnetic forces. Having neighboring spins aligned is energetically favored because when spins are aligned Pauli exclusion keeps electrons separated thus lowering the Coulombic energy.

Antiferromagnets In some materials (e.g., oxides of iron, copper and nickel) the anti-aligned state is favored. The Ising model may be used with simply taking the interaction coefficient $\epsilon < 0$. At $T = 0$ the spins take a checkerboard pattern of alternating up and down spins.

Binary Alloys Some alloys, such as zinc/copper, will form a regular alternating structure below the critical temperature, similar to the checkerboard pattern of an antiferromagnet.

Bragg–Williams Approximation (§12.5)

Lecture 19

Let's consider an alternative way of approximately solving the Ising model. This new approach doesn't give a better solution (in fact it gives the *same* solution) but it does help us to understand interacting systems.

Picture our lattice of spins as follows

$$\begin{array}{ccccc}
- & \overset{+-}{\Longleftrightarrow} & + & \overset{+-}{\Longleftrightarrow} & - \\
\Updownarrow_{--} & & \Updownarrow_{++} & & \Updownarrow_{+-} \\
- & \overset{+-}{\Longleftrightarrow} & + & \overset{++}{\Longleftrightarrow} & + \\
\Updownarrow^{+-} & & \Updownarrow^{+-} & & \Updownarrow^{+-} \\
+ & \overset{+-}{\Longleftrightarrow} & - & \overset{--}{\Longleftrightarrow} & -
\end{array}$$

where a "+" is an up spin state ($s = +1$) and "−" is a down spin state ($s = -1$). Here we take each site to have $\gamma = 4$ neighbors and label each pair of sites as "++" or "−−" if the pair are both up or both down; pair that are unaligned are labeled as "+−". Define

$$\begin{aligned}
N_+ &= \text{Number of up spins} \\
N_- &= \text{Number of down spins}
\end{aligned}$$

$$\begin{aligned} N_{++} &= \text{Number of up-up pairs} \\ N_{--} &= \text{Number of up-down pairs} \\ N_{+-} &= \text{Number of up-down pairs} \end{aligned}$$

Since the total number of spins is N and the total number of pairs is $\frac{1}{2}\gamma N$,

$$\begin{aligned} N_{+} + N_{-} &= N \\ N_{++} + N_{--} + N_{+-} &= \frac{1}{2}\gamma N \end{aligned}$$

Notice that the average value of the spin state may be written as

$$\langle s \rangle = \frac{1}{N}\sum_{i=1}^{N} s_i = \frac{N_+ - N_-}{N}$$

so

$$\begin{aligned} N_{+} &= \frac{N}{2}(1 + \langle s \rangle) \\ N_{-} &= \frac{N}{2}(1 - \langle s \rangle) \end{aligned}$$

For $\langle s \rangle = 0$, $N_+ = N_- = N/2$.

The total energy in the system may be written as

$$\begin{aligned} U &= -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \epsilon_{i,j} - H\sum_{i=1}^{N} s_i \\ &= -\epsilon(N_{++} + N_{--} - N_{+-}) - H(N_+ - N_-) \end{aligned}$$

since the interaction coefficient $\epsilon_{i,j} = \epsilon$ if spins i and j are neighbors and is zero otherwise.

So far we've made no approximations (and not gone very far). We now introduce the *Bragg-William approximation* by assuming that

$$\mathcal{P}_{++} = (\mathcal{P}_+)^2$$

where $\mathcal{P}_{++}$ is the probability that a pair is "++" and $\mathcal{P}_+$ is the probability that a spin is up. This is similar to the situation in which we flip a coin twice; the probability of getting two heads is the square of the probability of getting a single head. Note that this is true even if it's an unfair coin (i.e., change of heads different from 50%).

Since

$$\begin{aligned} \mathcal{P}_{+} &= \frac{N_+}{N} \\ \mathcal{P}_{++} &= \frac{N_{++}}{\frac{1}{2}\gamma N} \end{aligned}$$

the Bragg-William approximation may be written as

$$N_{++} = \frac{\gamma}{2N} N_+^2$$

By a similar line of argument we also have

$$\begin{aligned} N_{--} &= \frac{\gamma}{2N} N_-^2 \\ N_{+-} &= \frac{\gamma}{N} N_+ N_- \end{aligned}$$

Putting all this into our energy expression gives

$$\begin{aligned} U &= -\epsilon(N_{++} + N_{--} - N_{+-}) - H(N_+ - N_-) \\ &= -\frac{\epsilon\gamma}{2N}(N_+^2 + N_-^2 - 2N_+N_-) - H(N_+ - N_-) \\ &= -\frac{\epsilon\gamma}{2N}(N_+ - N_-)^2 - H(N_+ - N_-) \end{aligned}$$

which may be written as

$$U(N_+) = -\frac{\epsilon\gamma}{2N}(2N_+ - N)^2 - H(2N_+ - N)$$

An alternative expression for the energy, using $N_+ = \frac{N}{2}(1 + \langle s\rangle)$, is

$$U(\langle s\rangle) = -\frac{\epsilon\gamma}{2}N\langle s\rangle^2 - HN\langle s\rangle \qquad (*)$$

Unfortunately, we still don't know $\langle s\rangle$ so we're not finished.

At this point we can use the microcanonical ensemble since we've formulated the energy as a function of just N_+. This means that the energy is fixed if N_+ is fixed. But it is easy to count the total number of states at fixed N_+ since this is just the number of ways that we can pick N_+ up spins from a set of N, that is,

$$\Gamma(N_+) = \frac{N!}{N_+!(N - N_+)!} = \frac{N!}{N_+!N_-!}$$

In the microcanonical ensemble the entropy is

$$S = k\ln\Gamma = k\ln N! - k\ln N_+! - k\ln N_-!$$

Using $\ln M! \approx M\ln M$ gives

$$S = kN\ln N - kN_+\ln N_+ - kN_-\ln N_-$$

which may be written as

$$S = -\frac{kN}{2}\left[(1 + \langle s\rangle)\ln(1 + \langle s\rangle) + (1 - \langle s\rangle)\ln(1 - \langle s\rangle)\right] \qquad (**)$$

You can check that $S \to 0$ as $\langle s\rangle \to \pm 1$.

We know the sort of behavior to expect: at high temperature the average spin is zero but below the critical temperature the spins are favored to align. Since temperature is the key, let's get an expression for $\langle s\rangle$ in terms of T by evaluating

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial U}\right)_{H,N} \\ &= \frac{\left(\frac{\partial S}{\partial\langle s\rangle}\right)_{H,N}}{\left(\frac{\partial\langle s\rangle}{\partial U}\right)_{H,N}} \end{aligned}$$

Using (**) to evaluate the numerator and (*) to evaluate the denominator gives

$$\frac{1}{T} = \frac{-\frac{kN}{2}\ln\left(\frac{1+\langle s\rangle}{1+\langle s\rangle}\right)}{-N\epsilon\gamma\langle s\rangle + NH} = \frac{1}{2\epsilon\gamma\langle s\rangle + 2H}\ln\left(\frac{1 + \langle s\rangle}{1 - \langle s\rangle}\right)$$

We can do a bunch of algebra (but no further approximations) to arrive at the expression

$$\langle s \rangle = \tanh(\epsilon\gamma\langle s \rangle/kT + H/kT)$$

which is *exactly* the same transcendental equation we obtained from mean field theory. Naturally the results obtained from mean field theory (e.g., critical temperature $T_c = \gamma\epsilon/k$) are the same for the Bragg-William approximation.

Finally, the reason that the mean field approximation and the Bragg-William approximation are equivalent is that both represent the "long range order" in the Ising model but neglect the "short range order." In the real physical system, if a spin is up then it is more probable that a neighbor spin will also be up; both approximations neglect this short-range correlation. If many spins are up then $\langle s \rangle$ will be positive so in the mean field approximation the value of s_i will tend to be positive since spin i feels this global field. In this sense the long-range correlation is included. For the Bragg-William approximation, the number of "++" pairs (N_{++}) depends only on the total number of up spins (N_+) and not on how they are distributed, which is equivalent to assuming the up spins are randomly distributed.

Any improvement of mean field theory will have to include short-range order. The simplest correction is called the Bethe approximation (section 12.6 in Pathria). Advanced methods are presented in Chapters 13 and 14. The full story of the Ising model (and its variants) could easily occupy us for several months. Instead of continuing its study we'll move on to the general theory of critical phenomena. But before leaving the Ising model, let's consider one last important case which happens to be exactly solvable.

One Dimensional Ising Chain (§13.2)

We consider a system of N Ising spins arranged as a one-dimensional chain with nearest-neighbor interactions, as pictured below

$$\begin{array}{cccccc} + & - & \dots & + & \dots & - \\ 1 & 2 & \dots & i & \dots & N \end{array}$$

In the absence of an external field (i.e., $H = 0$), the total energy is

$$E(s_1, \dots, s_N) = -\epsilon \sum_{i=1}^{N-1} s_i s_{i+1}$$

so the canonical partition function is

$$\begin{aligned} Q_N &= \sum_{s_1=-1}^{+1} \dots \sum_{s_N=-1}^{+1} e^{\beta E} \\ &= \sum_{s_1=-1}^{+1} \dots \sum_{s_N=-1}^{+1} \exp\left(-\epsilon\beta \sum_{i=1}^{N-1} s_i s_{i+1}\right) \\ &= \sum_{s_1=-1}^{+1} \dots \sum_{s_N=-1}^{+1} \prod_{i=1}^{N-1} \exp(\beta\epsilon s_i s_{i+1}) \end{aligned}$$

Using the identities

$$\begin{aligned} e^x &= \cosh(x) + \sinh(x) \\ \cosh(-x) &= \cosh(x) \\ \sinh(-x) &= -\sinh(x) \end{aligned}$$

and the fact that $s_i s_{i+1} = \pm 1$ gives

$$\exp(\beta\epsilon s_i s_{i+1}) = \cosh(\beta\epsilon) + s_i s_{i+1}\sinh(\beta\epsilon) = (1 + s_i s_{i+1} y)\cosh(\beta\epsilon)$$

where $y = \tanh(\beta\epsilon)$.

The partition function is now

$$\begin{aligned} Q_N &= \sum_{s_1=-1}^{+1} \cdots \sum_{s_N=-1}^{+1} \prod_{i=1}^{N-1} (1 + s_i s_{i+1} y)\cosh(\beta\epsilon) \\ &= \cosh^{N-1}(\beta\epsilon) \sum_{s_1=-1}^{+1} \cdots \sum_{s_N=-1}^{+1} \prod_{i=1}^{N-1} (1 + s_i s_{i+1} y) \end{aligned}$$

Let's explicitly write out the product,

$$(1 + s_1 s_2 y)(1 + s_2 s_3 y)\ldots(1 + s_{N-1} s_N y)$$

Doing the first sum, $\sum_{s_1=-1}^{+1}$, makes the product

$$(1 - s_2 y)[(1 + s_2 s_3 y)\ldots(1 + s_{N-1} s_N y)] + (1 + s_2 y)[(1 + s_2 s_3 y)\ldots(1 + s_{N-1} s_N y)]$$

or

$$2[(1 + s_2 s_3 y)\ldots(1 + s_{N-1} s_N y)]$$

In fact, each time we do a sum we eliminate a term in the sum and pick up a factor of two.

The partition function for the one-dimensional Ising chain is just

$$Q_N = 2^N \cosh^{N-1}(\beta\epsilon)$$

Compare this with the partition function for non-interacting spins with energy levels of 0 and ϵ,

$$Q_N = (2\cosh(\beta\epsilon))^N$$

which is the same as the 1D Ising partition function when N is large.

Interestingly, the 1D Ising chain has *no critical point.* Ising worked this out in writing his Ph.D. thesis, published it in a short paper (his first) in 1925, became a high school teacher and never wrote another paper! Though at the time he thought his model was a failure because of the negative result he obtained for the one-dimensional case, to this day the Ising model is the topic of over 1000 papers a year.

Chapter 7

Critical Phenomena

Universality in Critical Phenomena

Lecture 20

We found that the average spin $\langle s\rangle$ (or total magnetic moment $M \propto N\langle s\rangle$) had the form of a "pitchfork" bifurcation. That is above the critical temperature $\langle s\rangle = 0$ while below T_c we have non-zero magnetization, as shown in Fig. 12.10 in Pathria and Beale. What is particularly interesting about this result is that it appears to agree with all data for ferromagnetic materials (see Fig. 12.7 in Pathria and Beale) at least when $T \approx T_c$. This is surprising since different materials have different crystal structure, material properties, etc. Furthermore, the Ising model is *very* simplistic.

These results suggest a universal behavior for systems near the critical point. For a specific material the details of its structure and the nature of the interactions among its constituent atoms will affect the value of T_c but the thermodynamic properties, when normalized by the values at the critical point, will be universal.

Another example of this universality near the critical point is the liquid-vapor transition. Recall the $P - V$ diagram for the van der Waals model and the region of coexistence formulated using the Maxwell construction (minimizing the Gibbs free energy). When the pressure, volume and temperature are normalized by their values at the critical point then the experimental data for all materials fall on the same curve, at least close to T_c (see Fig. 7.1) and have approximate quantitative agreement with van der Waals. Given how difficult statistical mechanics can be for non-ideal systems, it seems worthwhile to pursue these exceptional results.

Critical Exponents – A First Look (§12.2)

To quantify how well mean field theory agrees with experiment (and with more advanced theories) we define a set of power law exponents for the behavior of various thermodynamic quantities near the critical point. To start, we define the critical exponent β as

$$\langle s\rangle \propto \left(\frac{T_c - T}{T_c}\right)^{\beta} \qquad\qquad (T < T_c, T \to T_c)$$

It's not difficult to show that mean field theory gives

$$\langle s\rangle \propto \sqrt{\frac{T_c - T}{T_c}}$$

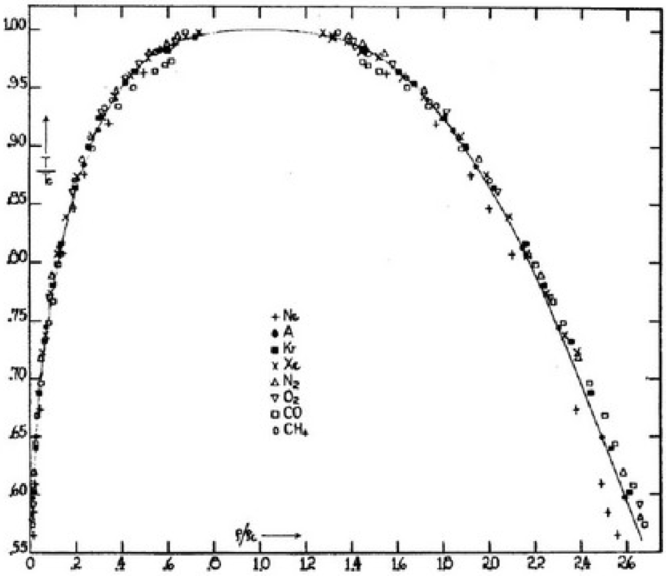

Figure 7.1: Normalized vapor/liquid coexistence boundary for various substances. From E. A. Guggenheim, *J. Chem. Phys.* **13**, 253 (1945).

so $\beta = 1/2$ in mean field theory. Experiments on ferromagnets find that for iron $\beta = 0.37 \pm 0.01$ and for nickel $\beta = 0.358 \pm 0.003$. Obviously mean field theory is not completely wrong but it is also not the final word.

Next, let's do a similar analysis for the van der Waals equation near the liquid-gas critical point. Recall that the equation

$$\left(P + \frac{an^2}{V^2}\right)(V - bn) = nRT$$

may be written as

$$\left(\frac{P}{P_c} + 3\frac{V_c^2}{V^2}\right)\left(3\frac{V}{V_c} - 1\right) = 8\frac{T}{T_c}$$

After some rearranging we find that this may be written as

$$(2-\eta)F = 8\tau(\eta+1) + 3\eta^3 \qquad (*)$$

where

$$\tau = \frac{T - T_c}{T_c}$$

is the dimensionless temperature. We define the *order parameter*

$$\eta = \frac{\rho - \rho_c}{\rho_c} = \frac{n - n_c}{n_c} = \frac{V_c - V}{V}$$

where $\rho = mN/V$ is the mass density and $n = N/V$ is the number density. The order parameter is the variable that bifurcates at the critical point; below T_c it has a high value on the liquid branch and a low value on the gas branch of the coexistence diagram. For the Ising model the corresponding order parameter is $\eta = \langle s \rangle$.

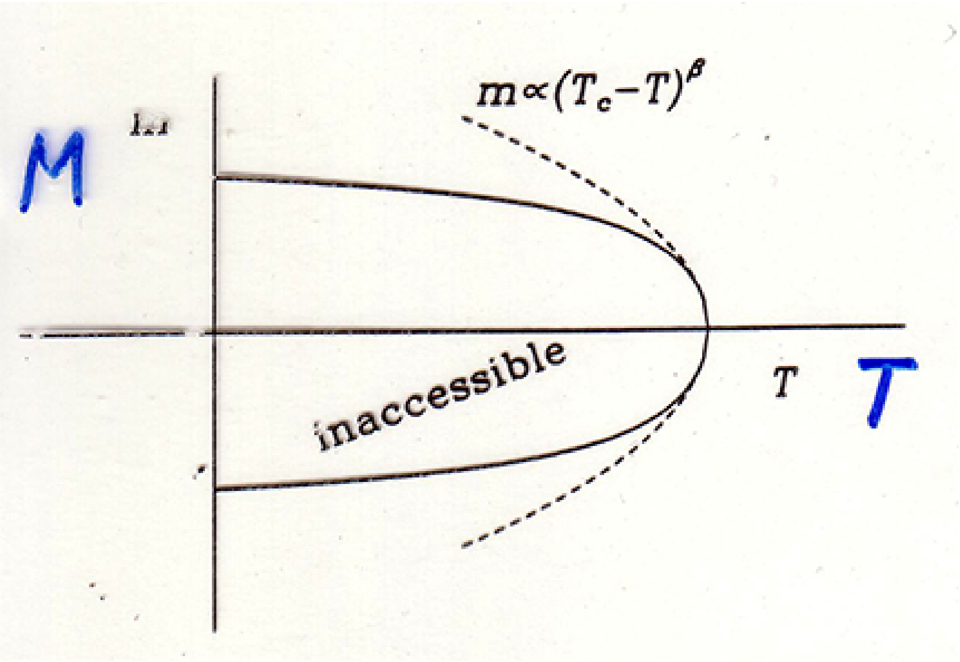

Figure 7.2: In the absence of an applied field ($F = 0$) the order parameter goes as $\eta \propto (T_c - T)^\beta$ as it bifurcates below the critical temperature.

The variable F is the dimensionless applied external field, specifically the pressure for which the order parameter vanishes,

$$F = \frac{P_0(T) - P_c}{P_c}, \qquad \text{where} \qquad P_0(T) = P(T, V = V_c)$$

For the Ising model we simply have $F = H$. In summary, the analogous variables are:

$$\langle s \rangle \Leftrightarrow \frac{\rho - \rho_c}{\rho_c} = \eta \qquad \text{(Order Parameter)}$$

$$H \Leftrightarrow \frac{P - P_c}{P_c} = F \qquad \text{(External Forcing)}$$

with $\tau = (T - T_c)/T_c$. See Figure 7.2 for a graphical illustration of the definition of the exponent β and Figure 7.3 for a 3D graph of the equation of state $\eta(\tau, F)$ in the vicinity of a critical point.

To find the critical exponent β for van der Waals we need to approach the critical point along the "phase boundary" which is a line that gives $\eta = 0$ as we pass the critical point. Using (*) and setting $\eta = 0$ we find $F = 4\tau$; holding this value of F from (*) we get

$$(2 - \eta)(4\tau) = 8\tau(\eta + 1) + 3\eta^3$$

or

$$\eta^3 = -4\tau\eta$$

so

$$\eta = 0 \quad \text{or} \quad \pm 2\sqrt{-\tau}$$

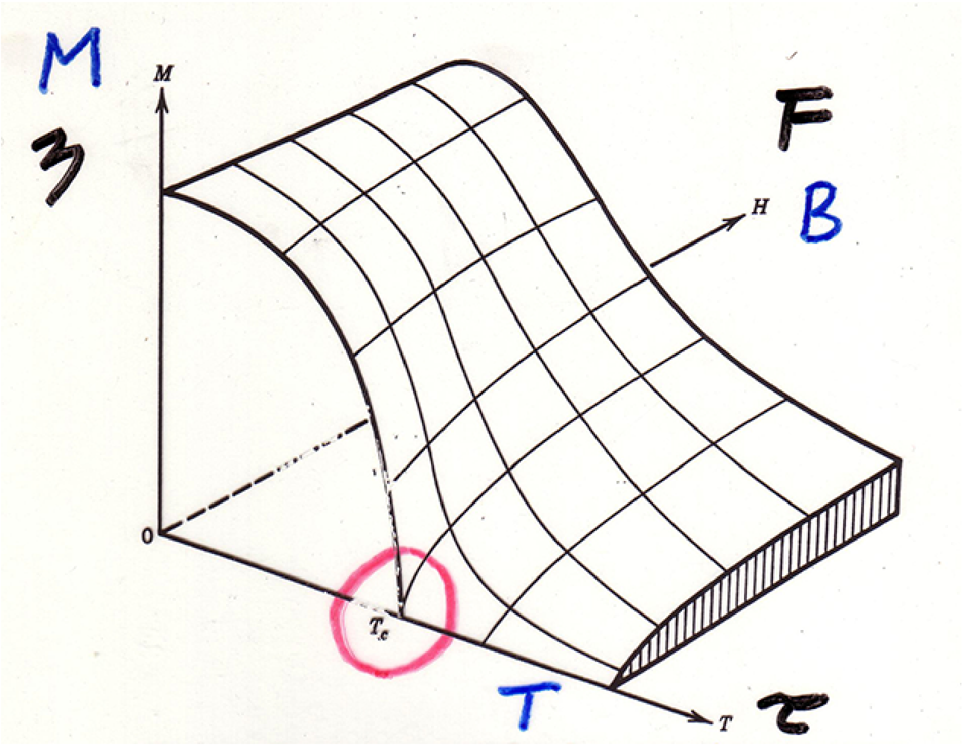

Figure 7.3: Graph of the equation of state $\eta(\tau, F)$ in the vicinity of a critical point (only the positive η surface is shown).

Above the critical temperature ($\tau > 0$) the only solution is $\eta = 0$. But below T_c we have a pitchfork bifurcation and the critical exponent is $\beta = 1/2$, just as we found with the Ising model. In fact, it turns out that the van der Waals equation of state may be shown to arise from a mean field approximation of the intermolecular attraction.

Landau Theory of Critical Phenomena

Given the results above and Maxwell's formulation of the coexistence boundary based on minimizing the Gibbs free energy, Landau suggested the following idea: The equilibrium state of a system is found by minimizing the appropriate thermodynamic potential; for systems at constant temperature and pressure that potential is the Gibbs free energy. Landau assumed that we may write the appropriate thermodynamic potential as $G(\eta, T)$ and that above the critical temperature this potential has a single minimum value, $G_0(T) = G(0, T)$ for $T > T_c$. Below the critical temperature there should be two minima and Landau assumed that sufficiently close to the critical point those minima are symmetrically positioned about $\eta = 0$. Finally, the function $G(\eta, T_c)$ must have an inflection point to smoothly transition from one minima to two minima as we pass the critical point. In brief, above T_c the thermodynamic potential $G(\eta, T)$ is single-well and below T_c it is double-well with minima on each side of $\eta = 0$.

Since the order parameter is small near T_c, we can Taylor expand our thermodynamic potential as,

$$G(\eta, T) = G_0(T) + \frac{1}{2}G_2(T)\eta^2 + \frac{1}{4}G_4(T)\eta^4 + \ldots$$

where the odd terms are absent due to the assumption of symmetry. The simplest way to get the desired transition of single-well to double-well is for $G_2(T)$ to change sign from positive to negative as the temperature drops below T_c. In the vicinity of the critical point we'll assume that $G(\eta, T)$ has the very simple form,

$$G(\eta, T) = \mathcal{C}_0 + \frac{1}{2}\mathcal{C}_2\tau\eta^2 + \frac{1}{4}\mathcal{C}_4\eta^4 + \ldots$$

where $\mathcal{C}_0, \mathcal{C}_2$ and $\mathcal{C}_4$ are positive constants. Landau's justification was that if the form of $G(\eta, T)$ isn't simple then the behavior won't be universal.

To find the value of the order parameter at equilibrium we find the minima of $G(\eta, T)$; keeping only the terms written above,

$$\left(\frac{\partial G}{\partial \eta}\right)_T = \mathcal{C}_2\tau\eta + \mathcal{C}_4\eta^3 = 0$$

so

$$\eta = 0 \quad \text{or} \quad \pm\sqrt{\frac{\mathcal{C}_2}{\mathcal{C}_4}}\sqrt{-\tau}$$

which, once again, produces a pitch-fork bifurcation of the order parameter as we pass through the critical point with a critical exponent of $\beta = 1/2$.

Landau's theory is the thermodynamic analog of mean field theory and it nicely ties together all the results we've seen thus far. Unfortunately we only have approximate agreement with experimental measurements so we don't have a complete theory at this point.

More Critical Exponents

We already discussed one of the critical exponents, β, defined as,

$$\eta \propto |\tau|^{\beta}$$

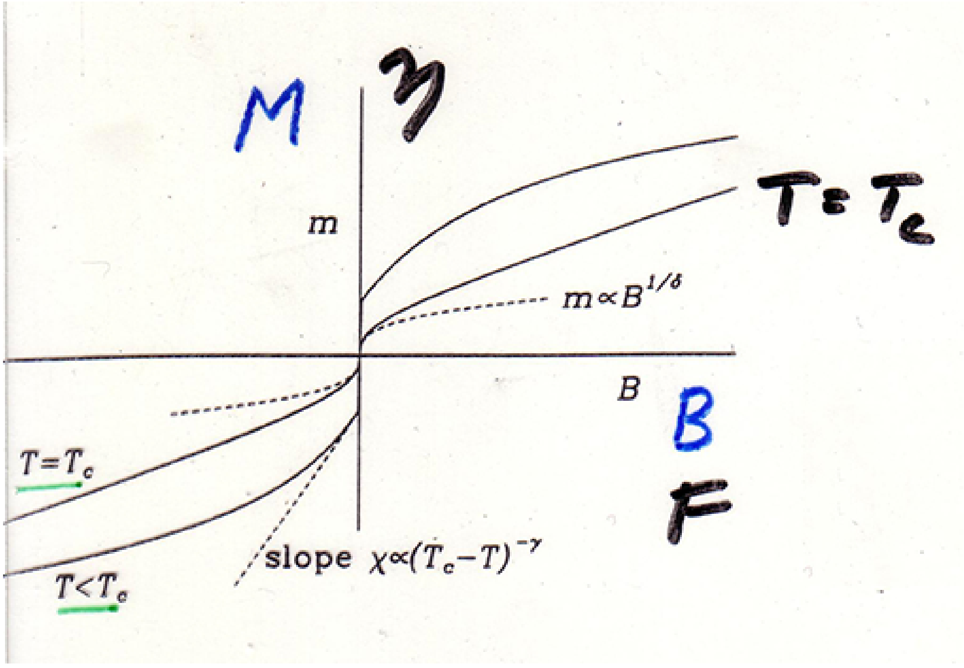

Figure 7.4: Graphical illustration for the critical exponents γ and δ on equation of state curves.

where η is the order parameter (e.g., $\eta = \langle s \rangle$ for the Ising model) and $\tau = (T - T_c)/T_c$; Landau theory and the mean field approximation give $\beta = 1/2$. While β is probably the most interesting critical exponent, there are three others that are studied.

The critical exponent α is defined as the temperature dependence of the heat capacity at constant F near T_c,

$$C_F \propto |\tau|^{-\alpha}$$

In reality there's α and α' depending on whether we are above or below the critical temperature but experimentally $\alpha \cong \alpha'$. The critical exponent γ gives the temperature dependance of the susceptibility (or compressibility) near T_c,

$$\left(\frac{\partial \eta}{\partial F}\right)_T \propto |\tau|^{-\gamma}$$

Again, there's a γ and γ' depending on whether we are above or below the critical temperature but experimentally $\gamma \cong \gamma'$. For both α and γ the external field, F, is fixed at $F = 0$. The fourth and final critical exponent, δ, is defined,

$$\eta \propto F^{1/\delta} \quad \text{at } T = T_c$$

which is the power law dependance of the order parameter on the external field near the critical point. See Figure 7.4 Graphical illustration for the critical exponents γ and δ on equation of state curves.

Now let's generalize Landau theory to include an external applied field and compute these new critical exponents. We now write the thermodynamic potential, say the Gibbs free energy, as

$$G(\eta, \tau, F) = \mathcal{C}_0 + \frac{1}{2}\mathcal{C}_2 \tau \eta^2 + \frac{1}{4}\mathcal{C}_4 \eta^4 - \mathcal{C}_* F \eta$$

where the $\mathcal{C}$'s are positive constants.* Notice that the external field energetically favors the case where

---

*Again we invoke Occam's razor to justify this simple form for G.

F and η have the same sign. The order parameter at equilibrium is obtained by finding the minimum of G so, as before, we evaluate

$$\left(\frac{\partial G}{\partial \eta}\right)_{\tau,F} = \mathcal{C}_2\tau\eta + \mathcal{C}_4\eta^3 - \mathcal{C}_*F = 0 \qquad (*)$$

We could solve this cubic but we don't need the explicit solution for the purpose of finding the critical exponents. The general shape of G is still either a single well or double well potential depending on whether we are above or below the critical temperature; for $F > 0$ the largest positive root of this cubic is the equilibrium state since it is the deeper well.

To evaluate the critical exponent γ we use $(*)$ to evaluate

$$\left(\frac{\partial F}{\partial \eta}\right)_{\tau} = \left(\frac{\partial}{\partial \eta}\right)_{\tau}\left[\frac{\mathcal{C}_2}{\mathcal{C}_*}\tau\eta + \frac{\mathcal{C}_4}{\mathcal{C}_*}\eta^3\right] = \frac{\mathcal{C}_2}{\mathcal{C}_*}\tau + 3\frac{\mathcal{C}_4}{\mathcal{C}_*}\eta^2$$

For $T > T_c$, $F \to 0$, we know $\eta \to 0$ (this is just the disordered state). From our earlier analysis we also know that for $T < T_c$ the order parameter $\eta \propto |\tau|^{1/2}$ near the critical point. In summary, near the critical point,

$$\left(\frac{\partial \eta}{\partial F}\right)_{\tau} = \begin{cases} \left(\frac{\mathcal{C}_2}{\mathcal{C}_*}\tau\right)^{-1} & \text{for} \quad T > T_c \\ \left(\frac{\mathcal{C}_2}{\mathcal{C}_*}\tau + 3\frac{\mathcal{C}_4}{\mathcal{C}_*}|\tau|\right)^{-1} & \text{for} \quad T < T_c \end{cases}$$

From the second law of thermodynamics we know that the isothermal susceptibility (or compressibility) must be positive and so we find that both above and below the critical point,

$$\left(\frac{\partial \eta}{\partial F}\right)_{\tau} \propto \frac{1}{|\tau|}$$

so $\gamma = 1$ in Landau theory.

Using (*) again but setting $T = T_c$ ($\tau = 0$) immediately gives

$$\eta = \left(\frac{\mathcal{C}_*}{\mathcal{C}_4}F\right)^{1/3}$$

so the critical exponent $\delta = 3$ in Landau theory.

Finally, to get the exponent α we need to find the heat capacity; we'll evaluate this using

$$C_F = T\left(\frac{\partial S}{\partial T}\right)_F = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_F = -\frac{1+\tau}{T_c}\left(\frac{\partial^2 G}{\partial \tau^2}\right)_F$$

since $dG = f(\eta)dF - SdT$ (e.g., for van der Waals $dG = VdP - SdT$). Using our expression for G we get,

$$\begin{aligned}\left(\frac{\partial^2 G}{\partial \tau^2}\right)_F &= 2\mathcal{C}_2\eta\left(\frac{\partial \eta}{\partial \tau}\right)_F + \mathcal{C}_2\tau\left(\frac{\partial \eta}{\partial \tau}\right)_F^2 + \mathcal{C}_2\tau\eta\left(\frac{\partial^2 \eta}{\partial \tau^2}\right)_F \\ &\quad + 3\mathcal{C}_4\eta^2\left(\frac{\partial \eta}{\partial \tau}\right)_F^2 + \mathcal{C}_4\eta^3\left(\frac{\partial^2 \eta}{\partial \tau^2}\right)_F - \mathcal{C}_*F\left(\frac{\partial^2 \eta}{\partial \tau^2}\right)_F\end{aligned}$$

Since the critical exponent is defined at zero field ($F = 0$) the last term drops out. Above the critical temperature η is constant (and equal to zero) so all the terms in the expression above are zero. Below T_c we know that

$$\eta \propto |\tau|^{1/2} \qquad \text{so} \quad \left(\frac{\partial \eta}{\partial \tau}\right)_F \propto |\tau|^{-1/2} \propto \frac{1}{\eta} \qquad \text{and} \left(\frac{\partial^2 \eta}{\partial \tau^2}\right)_F \propto |\tau|^{-3/2} \propto \frac{1}{\eta^3}$$

so we find that each term in $(\partial^2 G/\partial\tau^2)_F$ is constant so C_F is a constant. Given the definition of α being $C_F \propto |\tau|^{-\alpha}$ we have that $\alpha = 0$ for Landau theory.

In summary, the Landau theory (and mean field theory) predict that the critical exponents are

$$\alpha = 0 \qquad \beta = 1/2 \qquad \gamma = 1 \qquad \delta = 3$$

Pathria and Beale list experimental data for a variety of physical systems (see Table 12.1) and find values to be,

$$\alpha = 0.0 - 0.2 \qquad \beta = 0.30 - 0.36 \qquad \gamma = 1.0 - 1.4 \qquad \delta = 4.0 - 5.0$$

On the one hand the experimental values are comparable to the theoretical predictions and similar for vastly different systems (magnets, gas-liquid mixtures, binary alloys, etc.). On the other hand, there is some disagreement with theory and the values are similar among different systems but not entirely universal.

Scaling Hypothesis §12.10

By definition, a homogeneous function, $f(x)$, has the property that

$$f(ax) = g(a)f(x)$$

where a is a constant. In fact, since this means that

$$f(abx) = g(ab)f(x) \qquad \text{and} \qquad f(abx) = g(a)g(b)f(x)$$

then $g(a)$ must be a power law, that is,

$$f(ax) = a^r f(x)$$

We may write this result as

$$f(a^p x) = af(x)$$

where $p = 1/r$.

The scaling hypothesis assumes that the free energy (near T_c) is a homogeneous function of the form,

$$G(a^p\tau, a^q F) = aG(\tau, F)$$

To obtain a relation between the coefficients p,q and the critical exponents, we apply the derivative to both sides of this equation,

$$\left(\frac{\partial}{\partial F}\right)_\tau G(a^p\tau, a^q F) = \left(\frac{\partial}{\partial F}\right)_\tau aG(\tau, F)$$

Define the re-scaled values $\tau' = a^p\tau$ and $F' = a^q F$ and we may write this equation as,

$$\left[\left(\frac{\partial}{\partial F'}\right)_\tau G(\tau', F')\right]\left[\left(\frac{\partial F'}{\partial F}\right)_\tau\right] = a\left(\frac{\partial}{\partial F}\right)_\tau G(\tau, F)$$

Near the critical point the order parameter is proportional to the derivative of G with respect to the external field, that is,

$$\eta \propto \left(\frac{\partial G}{\partial F}\right)_\tau$$

Using this observation we have,

$$\eta(\tau', F')a^q = a\eta(\tau, F)$$

or

$$\eta(a^p\tau, a^qF) = a^{1-q}\eta(\tau, F)$$

To grasp the meaning of this result, let's consider a numerical example. It turns out that in Landau theory $p = 1/2$ and $q = 3/4$. To get round numbers we'll pick $a = 16$, so the result above gives,

$$\eta((16)^{(}1/2)\tau, (16)^{(}3/4)F) = (16)^{1/4}\eta(\tau, F)$$

or

$$\eta(4\tau, 8F) = 2\eta(\tau, F)$$

That is, if we raise the temperature τ by a factor of 4 and also raise the applied field F by a factor of 8 then according to the scaling hypothesis the order parameter will double.

Now let's directly link the scaling coefficients to the critical exponents. Suppose we take $a = (-\tau)^{-1/p}$ and set $F = 0$ so

$$\eta(((-\tau)^{-1/p})^p\tau, 0) = ((-\tau)^{-1/p})^{1-q}\eta(\tau, 0)$$

or

$$\eta(-1, 0) = ((-\tau)^{-(1-q)/p})\eta(\tau, 0)$$

The left hand side is a constant so

$$\eta(\tau, 0) \propto |\tau|^{(1-q)/p} \qquad \text{thus} \qquad \beta = (1-q)/p$$

Working along similar lines we can also derive that

$$\alpha = 2 - 1/p \qquad \gamma = (2q-1)/p \qquad \delta = q/(1-q)$$

These give two relations among the critical exponents so, for example, given β and γ we could find α and δ. A simple form for the two relations is,

$$\begin{aligned} \alpha + 2\beta + \gamma &= 2 \\ \alpha + \beta(1+\delta) &= 2 \end{aligned}$$

Recalling that the critical exponents in Landau theory are $\alpha = 0$, $\beta = 1/2$, $\gamma = 1$, and $\delta = 3$ we see that they agree with this result from the scaling hypothesis. In fact these scaling hypothesis relations appear to be exact. For example, for the Ising model in 2-dimensions can be solved exactly via a rather long calculation (by Onsager) and one finds that,

$$\alpha = 0 \qquad \beta = 1/8 \qquad \gamma = 7/4 \qquad \delta = 15$$

which satisfy the two scaling hypothesis relations and give coefficients of $p = 1/2$ and $q = 15/16$. Finally, the universality of the scaling hypothesis result has its origin in the long-ranged, self-similar correlations that arise near a critical point, which is our last topic in the theory of critical phenomena.

Critical Fluctuations §12.11, 12.12

In describing thermodynamics systems we have taken them to be homogeneous. For the air in this room we would assign an average number density, n, so the average number of particles in any given (imaginary) volume is $\langle N \rangle = nV$. Yet we know that the actual number of molecules in such a volume would fluctuate, either if we looked at two different locations at the same time or looked in the same location at two different times. [†] Typically the standard deviation of the fluctuations in N is a fraction of the average value with

$$\frac{\text{Standard Deviation}}{\text{Average}} = \frac{\sqrt{\langle (N - \langle N \rangle)^2 \rangle}}{\langle N \rangle} = \frac{\mathcal{C}}{\sqrt{N}}$$

where the factor $\mathcal{C} = \sqrt{nkTK_T}$ depends on the isothermal compressibility K_T; for an ideal gas $\mathcal{C} = 1$.

In air at STP the number density is $n = 2.7 \times 10^{25}\text{m}^{-3}$ so taking the volume as $V = \lambda_{\text{Blue}}^3$ where $\lambda_{\text{Blue}} = 4.5 \times 10^{-7}$ m is the wavelength of blue light we get about 2.5 million air molecules. The fractional variation, $1/\sqrt{\langle N \rangle}$, is a little less than a tenth of one percent; this is not enough to notice directly but this variation in density (and the corresponding variation in the index of refraction) does produce noticeable scattering of blue light over large distances, which is why the sky looks blue. Nevertheless it's a small effect so distant mountains have a bluish tint but not my face as viewed from across the room.

At the critical point it's a whole different ballgame. For example, for the liquid-vapor system the isothermal compressibility goes as,

$$K_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \propto \frac{1}{|T - T_c|^\gamma}$$

where this critical exponent is $\gamma \approx 1.2$ for water. At the critical point the compressibility diverges so it takes no energy to expand or contract an element of fluid; the density undergoes wild fluctuations which are macroscopically visible in a laboratory volume. Because these fluctuations turn the vapor mucky the effect is known as *critical opalescence*; see `www.youtube.com/watch?v=0gfxO1OeoJ0` for example. Analogously, the isothermal susceptibility of a ferromagnet near the critical point goes as

$$\chi_T \propto \left(\frac{\partial \langle s \rangle}{\partial H}\right) \propto \frac{1}{|T - T_c|^\gamma}$$

where $\gamma \approx 1.3$ for iron. From simulations of the Ising model we see that near the critical point we find "islands" of all sizes, from single outcroppings to huge continents of aligned spins (with corresponding "lakes" of the opposite alignment); see Figure 7.5. These structures at all sizes are analogous to the density variations that create critical opalescence.

Correlations

To investigate these fluctuations near the critical point it makes sense to define a local order parameter, $\eta(\vec{r})$, such as the local number density at position $\vec{r}$ for liquid-vapor condensation. We define the correlation function as,

$$g(\vec{r}, \vec{r}') = \langle \eta(\vec{r})\eta(\vec{r}') \rangle - \langle \eta(\vec{r}) \rangle \langle \eta(\vec{r}') \rangle$$

[†] Assuming that the two locations or the two times were macroscopically distinct.

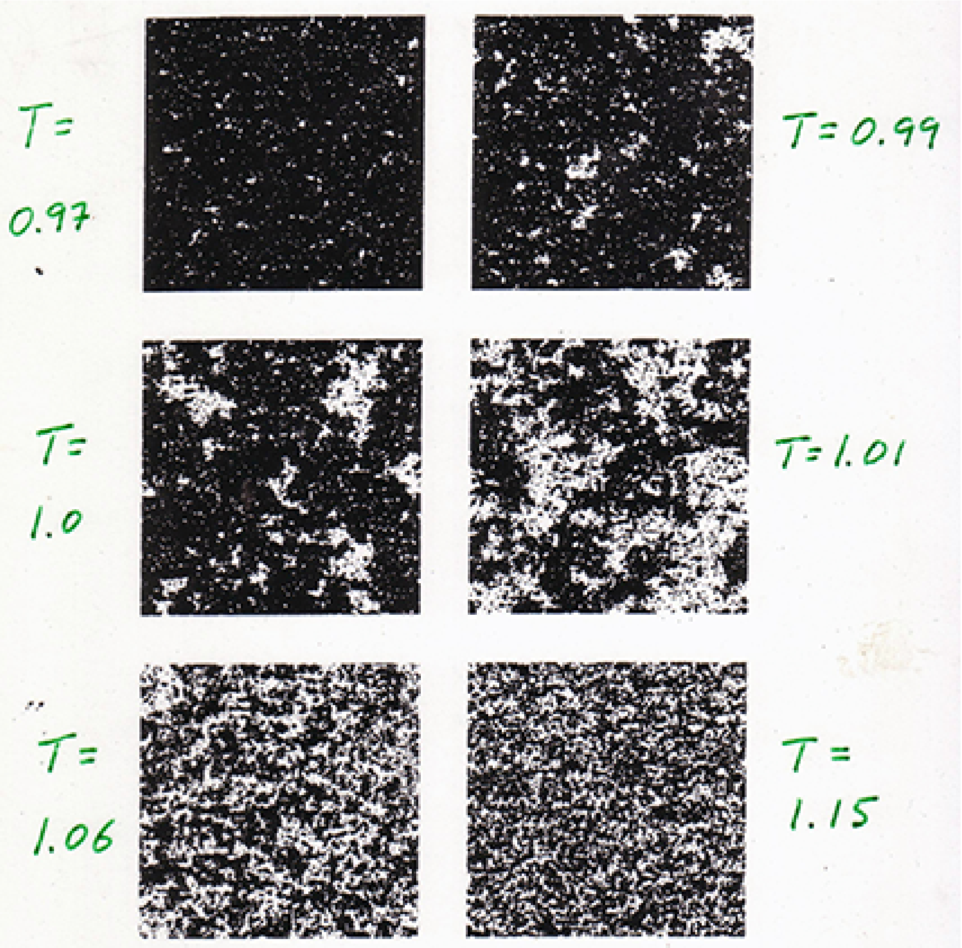

Figure 7.5: Simulation results for the Ising model at various temperatures above and below the critical point (with $T_c = 1$). Systems have 512 by 512 sites with periodic boundary conditions; black/white dots indicate up/down spins.

If $\eta(\vec{r})$ and $\eta(\vec{r}')$ are independent then $g = 0$. We can interpret this average as a time average, that is,

$$\langle f(\eta(\vec{r}))\rangle = \lim_{t\to\infty} \frac{1}{t}\int_0^t f(\eta(\vec{r},t))dt$$

Since the order parameter is a mechanical variable (such as number of particles or number of up spins) there's no difficulty in defining the instantaneous value, $\eta(\vec{r},t)$.[‡]

We still consider our equilibrium system to be statistically homogeneous (though not instantaneously homogeneous) so we have $\langle\eta(\vec{r})\rangle = \langle\eta(\vec{r}')\rangle$ and

$$g(\vec{r},\vec{r}') = g(|\vec{r}-\vec{r}'|) = g(\Delta r)$$

This correlation function is easy to measure in computer simulations and can be probed indirectly by neutron scattering or light scattering in the laboratory. In both simulations and experiments one finds that both above and below the critical point,

$$g(\Delta r) \propto \frac{1}{\Delta r} e^{-\Delta r/\xi}$$

for 3D systems. The correlation length, ξ, is not constant, in fact it increases and diverges as we approach the critical point. Specifically,

$$\xi \propto \frac{1}{|T-T_c|^\nu} \qquad\qquad (T \approx T_c)$$

The critical exponent, ν, is measured to be $\nu \approx 0.5 - 0.8$ (see Table 12.1 in Pathria and Beale). At $T = T_c$ the correlation function $g(r) \propto 1/r$ in 3D systems, so it is very long-ranged.

Ginzburg-Landau Theory

Let's extend Landau theory to deal with spatial fluctuations of the order parameter. We now define the local free energy density as,

$$G(\vec{r}) = \mathcal{C}_0 + \frac{1}{2}\mathcal{C}_2\tau\eta(\vec{r})^2 + \frac{1}{4}\mathcal{C}_4\eta(\vec{r})^4 - \mathcal{C}_*F(\vec{r})\eta(\vec{r}) + \frac{1}{2}\mathcal{C}_\nabla(\nabla\eta)^2$$

This should be familiar by now except for the last term, which depends on the spatial gradient of the order parameter. This term accounts for the fact that although it is energetically favorable for the system to be uniform (i.e., with zero gradients) and this term is constructed as the simplest possible form that gives this behavior.

At equilibrium the total free energy,

$$G_\Sigma = \int d\vec{r} G(\vec{r})$$

is minimum; to find that minimum we want to evaluate

$$\frac{\partial}{\partial\eta}G_\Sigma = \int d\vec{r}\ [\mathcal{C}_2\tau\eta(\vec{r}) + \mathcal{C}_4\eta(\vec{r})^3 - \mathcal{C}_*F(\vec{r}) + \mathcal{C}_\nabla\nabla^2\eta(\vec{r})]$$

[‡] Defining instantaneous values for non-mechanical variables, such as temperature, is a more delicate issue.

where the last term comes from

$$\begin{aligned}\frac{\partial}{\partial\eta}\int d\vec{r}\,(\nabla\eta)^2 &= \int d\vec{r}\,2(\nabla\eta)\frac{\partial}{\partial\eta}(\nabla\eta)\\ &= -2\int d\vec{r}\,(\nabla^2\eta)\frac{\partial}{\partial\eta}\eta \qquad \text{(Integration by parts)}\\ &= -2\int d\vec{r}\,\nabla^2\eta\end{aligned}$$

and the last step is obtained using integration by parts. For the integral to vanish requires that

$$\mathcal{C}_2'\tau\eta(\vec{r}) + \mathcal{C}_4'\eta(\vec{r})^3 - \mathcal{C}_\nabla'\nabla^2\eta(\vec{r}) = F(\vec{r})$$

where $\mathcal{C}_2' = \mathcal{C}_2/\mathcal{C}_*$, etc.

Since we are interested in fluctuations we use perturbation theory and write,

$$\begin{aligned}\eta(\vec{r}) &= \eta_0 + \eta_1(\vec{r})\\ F(\vec{r}) &= F_0 + F_1(\vec{r})\end{aligned}$$

where $\eta_0 = \langle\eta(\vec{r})\rangle$ and $F_0 = \langle F(\vec{r})\rangle$ are constants. Keeping only zero-th order terms we have

$$\mathcal{C}_2'\tau\eta_0 + \mathcal{C}_4'\eta_0^3 = F_0$$

which is our previous result for Landau theory (i.e., η_0 is zero for $T > T_c$ and has a pitchfork bifurcation for $T \leq T_c$). At first order in perturbation theory,

$$\mathcal{C}_2'\tau\eta_1(\vec{r}) + \mathcal{C}_4'\eta_0^2\eta_1(\vec{r}) - \mathcal{C}_\nabla'\nabla^2\eta(\vec{r}) = F_1(\vec{r})$$

or

$$(\nabla^2 - \xi^{-2})\eta_1(\vec{r}) = -\frac{1}{\mathcal{C}_\nabla'}F_1(\vec{r})$$

which is the inhomogeneous Helmholtz equation with

$$\xi = \sqrt{\frac{\mathcal{C}_\nabla'}{\mathcal{C}_2'\tau + \mathcal{C}_4'\eta_0^2}}$$

Notice that for $T > T_c$ we have $\eta_0 = 0$ while for $T < T_c$ we have $\eta \propto \sqrt{|\tau|}$ so both above and below the critical point we find

$$\xi \propto \frac{1}{|\tau|^{1/2}}$$

In a moment we'll see that this ξ is indeed the correlation length so Ginzburg-Landau theory predicts the critical exponent $\nu = 1/2$.

To finish the derivation, let's pin the value of the order parameter at location $\vec{r}'$ by setting $f_1(\vec{r}) = \delta(\vec{r} - \vec{r}')$. The solution to this Green's function problem for the Helmholtz equation is (in 3D),

$$\hat{\eta}_1(\vec{r}) \propto \frac{1}{|\vec{r} - \vec{r}'|}e^{-|\vec{r}-\vec{r}'|/\xi} \qquad \text{(Fixed } \eta(\vec{r}')\text{)}$$

This means that in the absence of an applied field the correlation function is

$$\begin{aligned}g(\Delta r) &= \langle\eta(\vec{r})\eta(\vec{r}')\rangle - \langle\eta(\vec{r})\rangle\langle\eta(\vec{r}')\rangle\\ &= \langle\eta_1(\vec{r})\eta_1(\vec{r}')\rangle\\ &\propto \hat{\eta}_1(\vec{r}) \propto \frac{1}{|\vec{r} - \vec{r}'|}e^{-|\vec{r}-\vec{r}'|/\xi}\end{aligned}$$

where the last step is justified by the fact that the $\hat{\eta}_1(\vec{r})$ gives the value of the order parameter for fixed $\eta(\vec{r}')$.

Chapter 8

Classical Non-ideal Gases

We now continue our study of systems with interacting particles (i.e., non-ideal) but in a new direction. Up to now we've considered mainly the universal behavior that we find near the critical point. We now turn to the more mundane problem of the specific behavior of specific systems, such as dense gases. The advantage of studying the universal behavior was that since it was universal we could pick the simplest systems, such as the Ising model, to study the behavior near the critical point. The disadvantage is that we're restricted to that small area of phase space.

In this chapter we'll consider some basic approximations for treating dense gases, essentially finding corrections to the ideal gas equation of state. We'll also see how some of the ideas may be extended to simple liquids. For the temperatures and pressures of interest the molecular interaction are well approximated by classical potentials so we'll restrict our attention to the classical canonical ensemble. Pathria and Beale consider more advanced techniques, including some quantum mechanics calculations, in their Chapter 10.

8.1 Canonical ensemble for non-ideal gases

Recall that in classical ensemble theory the state of a system of N particles is specified by the $\vec{r}_1, \ldots, \vec{r}_N$ positions and $\vec{p}_1, \ldots, \vec{p}_N$ momenta, which marks a point in the $6N$ dimensional phase space. The canonical partition function may be written as,

$$\begin{aligned} Q_N &= \frac{1}{N!h^{3N}} \int d\vec{r}_1 \ldots \int d\vec{r}_N \int d\vec{p}_1 \ldots \int d\vec{p}_N \exp\{-\beta E(\vec{r}_1, \ldots, \vec{r}_N, \vec{p}_1, \ldots, \vec{p}_N)\} \\ &= \frac{1}{N!h^{3N}} \int d\vec{r}_1 \ldots \int d\vec{r}_N \exp\{-\beta \mathcal{V}(\vec{r}_1, \ldots, \vec{r}_N)\} \int d\vec{p}_1 \ldots \int d\vec{p}_N \exp\{-\beta \mathcal{K}(\vec{p}_1, \ldots, \vec{p}_N)\} \end{aligned}$$

assuming that the total energy may be split as $E = \mathcal{K} + \mathcal{V}$ where the kinetic energy, $\mathcal{K}$, only depends on the momenta and the potential energy, $\mathcal{V}$, only depends on the positions of the particles. The $1/N!$ factor accounts for the indistinguishability of the particles (recall Gibbs' paradox) and the factor $1/h^{3N}$ is primarily for convenience, keeping the units consistent.

When we studied the classical ideal gas we assumed that $U = 0$ which is a good approximation when the particles are far apart (i.e., dilute gas) since the inter-molecular potentials are very short ranged (typically going like $1/r^6$). For the classical ideal gas we found,

$$Q_N^{\text{ideal}} = \frac{V^N}{N!h^{3N}} \int d\vec{p}_1 \ldots \int d\vec{p}_N \exp\{-\beta \mathcal{K}(\vec{p}_1, \ldots, \vec{p}_N)\}$$

$$\begin{aligned} &= \frac{1}{N!h^{3N}} \left(\int d\vec{p}\ \exp\{-\beta|\vec{p}|^2/2m\} \right)^N \\ &= \frac{V^N}{N!\lambda^{3N}} \end{aligned}$$

where $\lambda = h/\sqrt{2\pi mkT}$ is the thermal wavelength. Using this result, we may write the partition function for the non-ideal system as,

$$Q_N = Q_N^{\text{ideal}} \frac{1}{V^N} \int d\vec{r}_1 \ldots \int d\vec{r}_N \exp\{-\beta\mathcal{V}(\vec{r}_1, \ldots, \vec{r}_N)\} = \frac{Q_N^{\text{ideal}} I_N}{V^N}$$

Note that for an ideal gas we have $I_N = V^N$.

To find the equation of state for this non-ideal system we could use,

$$\begin{aligned} P(T,V,N) &= -\left(\frac{\partial A}{\partial V}\right)_{T,N} = \frac{1}{\beta}\left(\frac{\partial}{\partial V}\right)_{T,N} \ln Q_N \\ &= \frac{1}{\beta}\left(\frac{\partial}{\partial V}\right)_{T,N} (\ln Q_N^{\text{ideal}} + \ln I_N - N\ln V) \\ &= \frac{NkT}{V} - kT\left[\frac{N}{V} - \left(\frac{\partial}{\partial V}\right)_{T,N} \ln I_N\right] \\ &= kT\left(\frac{\partial}{\partial V}\right)_{T,N} \ln I_N \end{aligned}$$

Unfortunately the potential energy integral, I_N, has a much more complex form, in general, than the kinetic energy so we'll have to make some approximations or limit ourselves to some simplified intermolecular potentials.

8.2 Hard Core Model

Suppose that the particles interact with a pure exclusion potential with no attraction and that the volume occupied by each particle is v_0. The simple physical picture to have in mind is that the particles are hard spheres (e.g., tiny billiard balls). The potential energy for such a potential may be written as,

$$\mathcal{V}(\vec{r}_1, \ldots, \vec{r}_N) = \begin{cases} 0 & \text{no particles overlap} \\ \infty & \text{otherwise} \end{cases}$$

The probability of a point in phase space $\mathcal{P}(\vec{r}_1, \ldots, \vec{r}_N)$ is zero if any particles have positions that overlap; all other points in phase space are equally probable. This might seem like it would be difficult to work with but it turns out to be simple if we assume that the gas is not too dense.

Recall that we want to evaluate

$$I_N = \int d\vec{r}_1 \ldots \int d\vec{r}_N \exp\{-\beta\mathcal{V}(\vec{r}_1, \ldots, \vec{r}_N)\}$$

We can do the integral for the position of the first particle easily since it may be anywhere within the volume so $\int d\vec{r}_1 = V$. The second particle may be anywhere within the volume *except* at positions where it would overlap with the first particle. Since the particles have volume v_0 we have $\int d\vec{r}_2 = V - v_0$. For the next particle we will assume that the system is not so dense that we have to worry about geometry corrections that depend on the positions of the first two particles; in brief, we assume that the available volume of phase space for the third particle is $\int d\vec{r}_3 = V - 2v_0$.

Continuing in this way we find,

$$I_N = (V)(V - v_0)(V - 2v_0)\ldots(V - (N-1)v_0) = \prod_{i=0}^{N-1}(V - iv_0)$$

so

$$\ln I_N = \sum_{i=0}^{N-1} \ln(V - iv_0)$$

To find the equation of state we evaluate,

$$\begin{aligned}
P(T,V,N) &= kT\left(\frac{\partial}{\partial V}\right)_{T,N} \ln I_N \\
&= kT\sum_{i=0}^{N-1}\frac{1}{V - iv_0} = \frac{kT}{V}\sum_{i=0}^{N-1}(1 + iv_0/V) \\
&= \frac{kT(N-1)}{V} + kT\frac{v_0}{V^2}\frac{(N-1)(N-2)}{2} \\
&\approx \frac{NkT}{V}\left(1 + \frac{bN}{V}\right) \approx \frac{NkT}{V - bN}
\end{aligned}$$

where $b = v_0/2$. Notice that we get the van der Waals equation of state with only repulsion (i.e., van der Waals attraction coefficient $a = 0$).

8.3 Intermolecular Potentials

The hard core model is obviously a highly simplified picture of intermolecular interactions. A more accurate representation is to assume that the potential energy may be expressed in terms of pair-potentials that only depend on the relative distance between particles, that is,

$$\mathcal{V}(\vec{r}_1,\ldots,\vec{r}_N) = \sum_{i=1}^{N}\sum_{j=i+1}^{N} v(|\vec{r}_i - \vec{r}_j|) == \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} v(|\vec{r}_i - \vec{r}_j|)$$

This is a good approximation for simple molecules; far more complicated intermolecular models exist but they're mostly used in computer simulations.

Theory and measurement indicate that the potential $v(r)$ has two parts: a weak attraction term due to dipole-dipole interactions and a strong but shorter-ranged repulsion due to Fermi exclusion of the electron clouds. From E&M we know that the dipole-dipole potential goes as r^{-6} while the repulsion term is not so short-ranged that it's precise form is not too important so for convenience it's usually taken to be r^{-12}. We'll write this model potential, known as the *Lennard-Jones* potential, as,

$$v_{LJ}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

where the parameter σ indicates the range of the potential and ϵ indicates the strength. See Figure 8.1 for a graph of the potential; note that the minimum of the potential well is at $r_{min} = 2^{1/6}\sigma \approx 1.12\sigma$ and $v_{LJ}(r_{min}) = -\epsilon$.

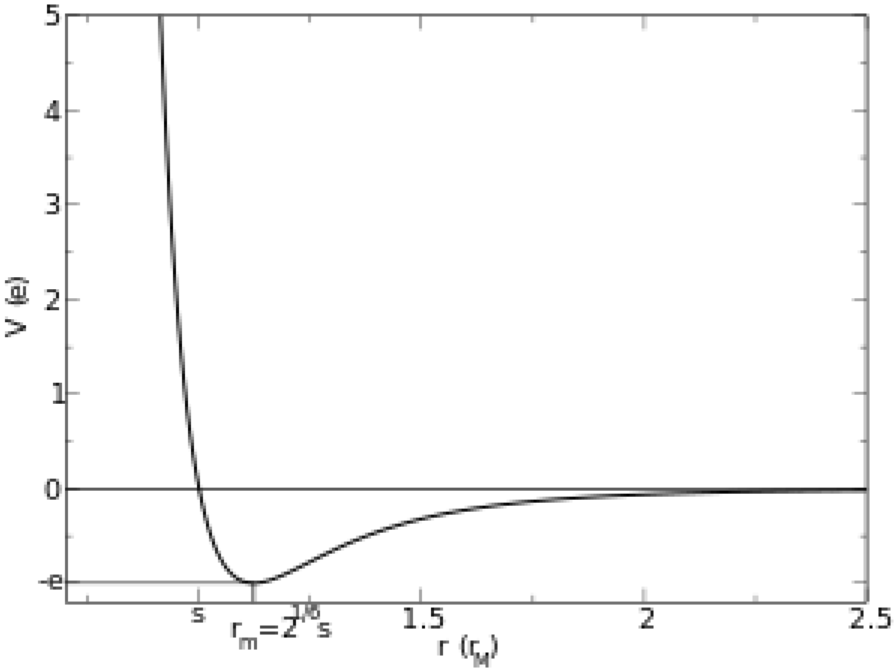

Figure 8.1: Lennard-Jones potential.

A potential that is half-way between the hard core model and the Lennard-Jones model is hard core with dipole attraction,

$$v_G(r) = \begin{cases} \infty & r < r_{min} \\ \epsilon(r_{min}/r)^6 & r \geq r_{min} \end{cases}$$

Since I can't find the real name for it we'll call this the Garcia model.

8.4 Homogeneous Approximation

A way to account for the contribution of the inter-molecular forces is to assume that the total potential energy may be written as

$$U_{pot} = \frac{1}{2} N \langle v \rangle_T$$

where we define $\langle v \rangle_T$ as the average potential energy of a given (tagged) particle due to the presence of all the other particles. We can write this average as

$$\langle v \rangle_T = \int d\vec{r}\; v(r) \rho_T(\vec{r})$$

where $\rho_T(\vec{r})$ is the number density of particles located at a position $\vec{r}$ relative to the tagged particle.[*]

In the homogeneous approximation we take $\rho_T(\vec{r}) = \rho = N/V$, that is, a constant equal to the total number density. For intermolecular potentials with a strong repulsion, such as a hard core repulsion, we improve this approximation by changing it to,

$$\rho_T(\vec{r}) = \begin{cases} 0 & |\vec{r}| < r_{min} \\ N/V & |\vec{r}| \geq r_{min} \end{cases}$$

---

[*] For convenience we can simply take the origin to be at the center of the tagged particle.

where r_{min} is the minimum separation between particles due to the strong repulsion. The homogeneous approximation is reasonable for dilute systems and later we'll discuss corrections for more realistic forms of $\rho_T(\vec{r})$.

As an example of the homogeneous approximation let's evaluate $\langle \upsilon \rangle_T$ for the Garcia model (hard core with dipole attraction). The integrals are easy to do,

$$\begin{aligned}\langle \upsilon \rangle_T &= \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \int_{r_{min}}^{\infty} r^2\, dr \upsilon_G(r) \left(\frac{N}{V}\right) \\ &= 4\pi\epsilon r_{min}^6 \frac{N}{V} \int_{r_{min}}^{\infty} r^{-4} \\ &= -\frac{4}{3}\pi\epsilon r_{min}^3 \frac{N}{V}\end{aligned}$$

From this, the total potential energy is

$$U_{pot} = \frac{1}{2} N \langle \upsilon \rangle_T = -\frac{aN^2}{V}$$

where $a = \frac{2}{3}\pi\epsilon r_{min}^3$. Notice that the factor a increases if ϵ increases (strength of the potential increases) or r_{min} increases (range of the potential increases).

To relate this result to the pressure we use

$$P = -\left(\frac{\partial A}{\partial V}\right)_{T,N} = -\left(\frac{\partial U}{\partial V}\right)_{T,N} + T\left(\frac{\partial S}{\partial V}\right)_{T,N} \qquad (*)$$

The total energy is

$$U = U_{kin} + U_{pot} = \frac{3}{2} NkT - \frac{aN^2}{V}$$

so the contribution to the pressure due to the dipole attraction is

$$P_{\text{attraction}} = -\left(\frac{\partial U}{\partial V}\right)_{T,N} = -\frac{aN^2}{V^2}$$

For the second term in (*), that is the entropy contribution, we can use our previous result for the hard core model equation of state,

$$P_{\text{repulsion}} = \frac{NkT}{V - bN}$$

and putting these two pieces together gives the full van der Waals EoS.

8.5 Pair Correlation Function

Returning to the general case, for dense, isotropic fluids it's more convenient to express the number density about a tagged particle as

$$\rho_T(\vec{r}) = \rho g(r)$$

where $g(r)$ is the radial (or pair) distribution function. In the homogeneous approximation for the hard core potential we assumed that

$$g(r) = \begin{cases} 0 & r < r_{min} \\ 1 & r > r_{min} \end{cases}$$

But even for true hard sphere molecules this is not correct except at low densities because the molecules tend to form an approximate lattice structure at the higher densities (see Figure 8.2). In general,

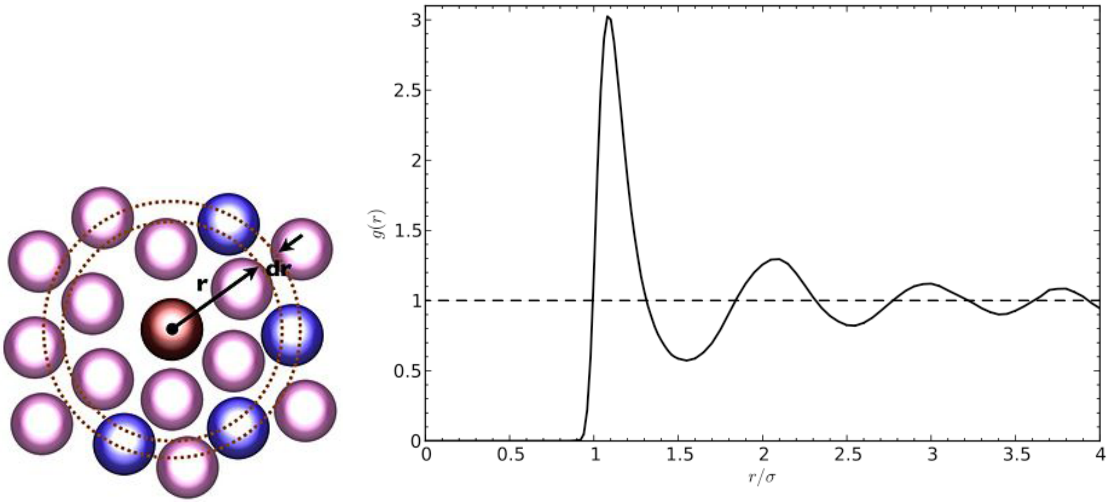

Figure 8.2: (Left) Illustration of hard sphere particles at higher densities; notice that there is ordering of the positions even in the absence of an attraction force. (Right) Pair correlation function in a dense Lennard-Jones fluid.

$g(r) \to 1$ as $r \to \infty$ because as we move far away from the tagged particle any correlations are lost (i.e., $\rho_T(\vec{r}) \to \rho$ as $r \to \infty$).

One of the big advantages of formulating our analysis in terms of the pair correlation function is that it is easily measured in computer simulations. We can also formulate theories to estimate it for a given potential. Rather than do the more general case we'll only consider the simple approximation,

$$g(r) \approx e^{-\beta v(r)}$$

which is a good approximation for moderate densities. Notice that for the hard core potential this is the same as the homogeneous approximation but for more complicated potentials it includes more structure. Using it we have

$$U = \frac{3}{2}NkT + \frac{1}{2}\frac{N^2}{V}\int d\vec{r} v(r) e^{-\beta v(r)}$$

From which we can formulate the equation of state.

Because this approach is only valid for moderate densities we'll express our equation of state in the form of a virial expansion, that is, an expansion in $\rho = N/V$,

$$P = \rho kT(1 + B_2(T)\rho + B_3(T)\rho^2 + \ldots)$$

and find an expression for the first expansion coefficient, $B_2(T)$, in terms of the potential $v(r)$.

We'll be using the Helmholtz free energy again and since we want to find the difference from the ideal gas law we'll write,

$$\Delta U = U - U_{\text{ideal}} = U - \frac{3}{2}NkT \qquad \text{and} \qquad \Delta A = A - A_{\text{ideal}}$$

To find ΔA we use

$$\begin{aligned} U &= A + TS = A + \frac{1}{k\beta}\left(-\frac{\partial A}{\partial T}\right)_{V,N} \\ &= A + \beta\left(\frac{\partial A}{\partial \beta}\right)_{V,N} = \left(\frac{\partial}{\partial \beta}\right)_{V,N}(\beta A) \end{aligned}$$

so

$$\Delta U = \left(\frac{\partial}{\partial \beta}\right)_{V,N} (\beta \Delta A)$$

This is useful once you notice that,

$$\begin{aligned}\Delta U &= \frac{1}{2}\frac{N^2}{V}\int d\vec{r} v(r) e^{-\beta v(r)} \\ &= -\left(\frac{\partial}{\partial \beta}\right)_{V,N} \frac{1}{2}\frac{N^2}{V}\int d\vec{r} e^{-\beta v(r)}\end{aligned}$$

so by comparing the two expression above for ΔU we have,

$$\beta \Delta A = -\frac{1}{2}\frac{N^2}{V}\int d\vec{r}\left[e^{-\beta v(r)} + C\right]$$

where C is a constant of integration; we know that $C = -1$ since $\Delta A = 0$ if $v(r) = 0$ (ideal gas limit).

We can now get the deviation from the ideal gas pressure, Δp, from

$$\Delta P = -\left(\frac{\partial}{\partial V}\right)_{T,N} \Delta A = -\frac{1}{2}\frac{N^2}{V^2}\int d\vec{r}\left[e^{-\beta v(r)} - 1\right]$$

so

$$P = P_{\text{ideal}} + \Delta P = \rho k T\left(1 - \frac{1}{2}\rho\int d\vec{r}\left[e^{-\beta v(r)} - 1\right]\right)$$

Comparing with the virial expansion we see that

$$B_2(T) = -\frac{1}{2}\int d\vec{r}\left[e^{-\beta v(r)} - 1\right] = -\frac{1}{2}\int d\vec{r} f(r)$$

where $f(r) = e^{-\beta v(r)} - 1$ is known as the Mayer function.

Chapter 9

Fluctuations and Stochastic Processes

Thermodynamic Fluctuations (14.1)

Lecture xx

A system at thermodynamic equilibrium is extremely homogeneous. Most fluctuations we observe in daily phenomena (e.g., dust particles dancing in the air, density ripples of air above hot asphalt) are *non-equilibrium* phenomena. Though equilibrium fluctuations are typically extremely small in magnitude they are fundamentally important (and sometimes have important effects).

Recall that in the micro-canonical ensemble a system with fixed energy U and volume V is at equilibrium when its entropy, $S(U, V)$, is maximum. Let's take such a system and split it into two parts ($U_\Sigma = U_1 + U_2$, $V_\Sigma = V_1 + V_2$, $N_\Sigma = N_1 + N_2$) The first being much smaller than the second ($U_1 \ll U_2$, $V_1 \ll V_2$, $N_1 \ll N_2$); for simplicity take N_1, N_2 fixed.

Call T^* and P^* the temperature and pressure at equilibrium. The average temperature of each part is $\overline{T}_1 = \overline{T}_2 = T^*$, similarly for pressure.

We ask the following question: "Suppose that we measure T_1 at random, what is the probability of a fluctuation $\Delta T_1 = T_1 - T^*$?" This is the formal version of the question, "What's the chance that water will spontaneously boil in the corner of your glass of water?"

Call S^*_Σ the entropy at equilibrium, that is,

$$S^*_\Sigma = S_1(\overline{U}_1, \overline{V}_1) + S_2(\overline{U}_2, \overline{V}_2) = \overline{S}_1 + \overline{S}_2$$

For different values the system has a lower entropy, that is,

$$S_\Sigma = S_1(U_1, V_1) + S_2(U_2, V_2) \leq S_0$$

The probability of this fluctuation, $\mathcal{P}(S)$, may be found by using

$$S = k \ln \Gamma$$

where Γ is the number of accessible states for the system. Specifically,

$$\mathcal{P}(S) \propto \frac{\Gamma}{\Gamma_0} = \frac{e^{S_\Sigma/k}}{e^{S^*_\Sigma/k}} = e^{(S_\Sigma - S^*_\Sigma)/k} = e^{\Delta S_\Sigma/k}$$

We'll be OK without having to find the normalization so don't worry that the above is "goes as" rather than "equals."

Now we want to analyze ΔS_Σ in terms of the states of parts 1 and 2 of the system; specifically,

$$\Delta S_\Sigma = S_\Sigma - \overline{S}_\Sigma = (S_1 - \overline{S}_1) + (S_2 - \overline{S}_2) = \Delta S_1 + \Delta S_2$$

Since the fluctuations are small we can use the differential version of the first and second laws

$$\Delta U_2 = T_2 \Delta S_2 - P_2 \Delta V_2$$

or

$$\Delta S_2 = \frac{1}{T_2}(\Delta U_2 + P_2 \Delta V_2)$$

There's a similar expression for ΔS_1 but we won't need it.

Can now write

$$\Delta S_\Sigma = \Delta S_1 + \frac{1}{T_2}(\Delta U_2 + P_2 \Delta V_2)$$

and make the following two observations: 1) Since $U = U_1 + U_2$ is fixed, $\Delta U_2 = -\Delta U_1$ and similarly $\Delta V_2 = -\Delta V_1$. 2) Since system 2 is very large $T_2 \approx T^*$ and $P_2 \approx P^*$ (i.e., fluctuations only have a significant effect in the smaller part). From these observations,

$$\Delta S_\Sigma = \Delta S_1 - \frac{1}{T^*}\Delta U_1 - \frac{P^*}{T^*}\Delta V_1 \qquad (*)$$

We now have ΔS in terms of fluctuations in system 1; the next step is to replace ΔU_1 with an expression involving only temperature, pressure, etc.

We now write $U_1(S_1, V_1)$ and Taylor expand it about the equilibrium state,

$$\begin{aligned} U_1(S_1, V_1) &= U_1(\overline{S}_1, \overline{V}_1) + \overline{\left(\frac{\partial U_1}{\partial S_1}\right)_V} \Delta S_1 + \overline{\left(\frac{\partial U_1}{\partial V_1}\right)_S} \Delta V_1 \\ &+ \frac{1}{2}\left[\overline{\left(\frac{\partial^2 U_1}{\partial S_1^2}\right)_V}(\Delta S_1)^2 + \overline{\left(\frac{\partial^2 U_1}{\partial S_1 \partial V_1}\right)}(\Delta S_1)(\Delta V_1) + \overline{\left(\frac{\partial^2 U_1}{\partial V_1^2}\right)_S}(\Delta V_1)^2\right] + \ldots \end{aligned}$$

where the overbar indicates that the quantity is evaluated at equilibrium. For example,

$$\overline{\left(\frac{\partial U_1}{\partial S_1}\right)_V} = T^* \qquad \overline{\left(\frac{\partial U_1}{\partial V_1}\right)_S} = -P^*$$

To deal with the quadratic term in the square brace, use the Taylor expansion

$$T_1 = \left(\frac{\partial U_1}{\partial S_1}\right)_V = \overline{\left(\frac{\partial U_1}{\partial S_1}\right)_V} + \overline{\left(\frac{\partial^2 U_1}{\partial S_1^2}\right)_V}\Delta S_1 + \overline{\left(\frac{\partial^2 U_1}{\partial V_1 \partial S_1}\right)}\Delta V_1 + \ldots$$

and note that the first term on the r.h.s. equals T^* so

$$\Delta T_1 = T_1 - T^* = \overline{\left(\frac{\partial^2 U_1}{\partial S_1^2}\right)_V}\Delta S_1 + \overline{\left(\frac{\partial^2 U_1}{\partial V_1 \partial S_1}\right)_S}\Delta V_1$$

and so multiplying both sides by ΔS_1 gives,

$$(\Delta T_1)(\Delta S_1) = \overline{\left(\frac{\partial^2 U_1}{\partial S_1^2}\right)_V}(\Delta S_1)^2 + \overline{\left(\frac{\partial^2 U_1}{\partial V_1 \partial S_1}\right)}(\Delta S_1)(\Delta V_1)$$

By a similar line of reasoning, we obtain

$$-(\Delta P_1)(\Delta V_1) = \overline{\left(\frac{\partial^2 U_1}{\partial V_1^2}\right)_S}(\Delta V_1)^2 + \overline{\left(\frac{\partial^2 U_1}{\partial V_1 \partial S_1}\right)}(\Delta S_1)(\Delta V_1)$$

Collecting the above,

$$\begin{aligned} U_1(S_1, V_1) &= U_1(\overline{S}_1, \overline{V}_1) + T^* \Delta S_1 - P^* \Delta V_1 \\ &\quad + \frac{1}{2} \left[(\Delta T_1)(\Delta S_1) - (\Delta P_1)(\Delta V_1) \right] \end{aligned}$$

and putting this into (*) gives, *

$$\begin{aligned} \Delta S &= \Delta S_1 - \frac{P^*}{T^*} \Delta V_1 - \frac{1}{T^*} \Bigg\{ T^* \Delta S_1 - P^* \Delta V_1 \\ &\quad + \frac{1}{2} \left[(\Delta T_1)(\Delta S_1) - (\Delta P_1)(\Delta V_1) \right] \Bigg\} \\ &= -\frac{1}{2T^*} \left[(\Delta T_1)(\Delta S_1) - (\Delta P_1)(\Delta V_1) \right] \end{aligned}$$

We have one more step but before we finish let's simplify the notation and write this expression as just,

$$\Delta S_\Sigma = -\frac{1}{2T} \left[(\Delta T)(\Delta S) - (\Delta P)(\Delta V) \right]$$

so implicitly all the quantities are for system 1 and evaluated at equilibrium, that is, we'll simply write T instead of T^* for intensive variables and V instead of $\overline{V}_1$ for extensive variables.

We are almost finished but our expression for ΔS_Σ needs one last manipulation since the four variables $\Delta T, \Delta S, \Delta P$ and ΔV are not independent. Let's take ΔT and ΔV as the independent variables; to replace ΔS and ΔP we use,

$$\begin{aligned} \Delta S &= \left(\frac{\partial S}{\partial T} \right)_V \Delta T + \left(\frac{\partial S}{\partial V} \right)_T \Delta V = \frac{C_v}{T} \Delta T + \left(\frac{\partial P}{\partial T} \right)_V \Delta V \\ \Delta P &= \left(\frac{\partial P}{\partial T} \right)_V \Delta T + \left(\frac{\partial P}{\partial V} \right)_T \Delta V = \left(\frac{\partial P}{\partial T} \right)_V \Delta V - \frac{1}{V K_T} \Delta V \end{aligned}$$

Finally, the probability of a fluctuation in temperature and volume is,

$$\begin{aligned} \mathcal{P}(\Delta T, \Delta V) = \mathcal{P}(\Delta S_\Sigma) &\propto \exp \left\{ -\frac{C_v}{2kT^2} (\Delta T)^2 - \frac{1}{2kT K_T V} (\Delta V)^2 \right\} \\ &= \exp(-(\Delta T)^2 / 2\sigma_T^2) \; \exp(-(\Delta V)^2 / 2\sigma_V^2) \end{aligned}$$

We see that $\mathcal{P}(\Delta T, \Delta V)$ factors into a pair of Gaussian distributions† of the form,

$$\mathcal{P}(x) = A\, e^{-x^2/2\sigma^2} \qquad \text{(Gaussian distribution)}$$

where A is the normalization constant and σ is the standard deviation In this form for the Gaussian the mean, $\overline{x} = 0$, and the variance $\overline{(x - \overline{x})^2} = \overline{x^2} = \sigma^2$ so,

$$\overline{\Delta T^2} = \frac{kT^2}{C_v} \qquad \overline{\Delta V^2} = kT K_T V$$

Finally, since the distribution factors into a pair of distributions for ΔT and ΔV that means that these variables are independent (i.e., uncorrelated) so $\overline{\Delta T \Delta V}$. By the way, this result for the variances in

---

[*] You see why we kept N fixed for simplicity; as an exercise repeat this analysis to find ΔS_Σ with U, V, and N as variables.

[†] Also known as the normal distribution.

temperature and volume is unchanged if we allow N to vary as well. Pathria and Beale discuss the variances for other variables, such as $\overline{\Delta S^2}$, $\overline{\Delta P^2}$, etc.

Let's finish by putting in some values into these expressions to get a sense of the magnitude of the fluctuations. For a Gaussian distribution the probability that a random value is 3σ above the mean is about 0.2% and the probability that it is 9σ above the mean is about 10^{-19}. The heat capacity of one milliliter[‡] of water at room temperature is about 4 Joules per Kelvin so for temperature fluctuations,

$$\sigma_T = \sqrt{\overline{\Delta T^2}} = \sqrt{\frac{kT^2}{C_v}} \approx 5 \times 10^{-10} \text{ Kelvin}$$

Even if you've been watching that glass of water since the start of the Universe (about 10^{18} seconds) you're not going to see a corner start to spontaneously boil (or even get noticeable warm). This result isn't surprising since, in general, the ratio of the fluctuation to the mean values go as $1/\sqrt{N}$ so for macroscopic volumes where N is on the order of Avogadro's number these deviations are negligible.

On the other hand, at microscopic scales they are significant. Consider a yoctoliter[§] of water at room temperature; for this volume the standard deviation of temperature fluctuations is,

$$\sigma_T \approx (5 \times 10^{-10} \text{ Kelvin}) \times \sqrt{\frac{10^{-6}}{10^{-27}}} \approx 15 \text{ Kelvin}$$

This microscopic volume has roughly a hundred water molecules so the fluctuations in temperature are noticeable.

Brownian Motion

Lecture xx

We now analyze a specific phenomenon connect to spontaneous thermal fluctuations. Consider a small particle, such a pollen grain, that is neutrally buoyant in water. The grain is much larger than a molecule of water so it is visible under a microscope. But it is also small enough that the random impacts of the water molecules cause noticeable effects. We'll refer to such a particle as a Brownian particle.

Consider a particle with a radius of 1 micron (which is about twice the wavelength of visible light) and a density of 1 g/cm^3 (about the density of wood and of water); its mass is about 4×10^{-15} kg. From the equipartition theorem we expect the mean kinetic energy of the Brownian particle is given by,

$$\langle \tfrac{1}{2} m v^2 \rangle = \tfrac{3}{2} kT$$

so the r.m.s. speed of the particle is,

$$\sqrt{\langle v^2 \rangle} = \sqrt{3kT/m}$$

For $T = 300$ K we find that the particle's speed is almost 2 millimeters per second. If it traveled ballistically then in one second it would fly a distance of about 2000 radii; under the microscope it would look like a blur. But due to the constant collisions with the water molecules the direction and magnitude of the Brownian particle's velocity changes randomly.[¶]

[‡] One milliliter equals one cubic centimeter equals 10^{-6} m^3.

[§] One yoctoliter equals one cubic nanometer equals 10^{-27} m^3.

[¶] The ratio of the r.m.s. momentum of the Brownian particle compared with a water molecule equals the square root of the ratio of their masses.

Random Walk Model

Consider the following cartoon for the motion of the Brownian particle. For simplicity we'll restrict the motion to one-dimensional motion along the x-axis with the particle starting at $x = 0$ at $t = 0$; the generalization to 3D motion will be straight-forward. To further simplify the formulation we'll assume that at integer time steps the particle moves a unit distance with equal probability of going to the left or to the right. This is called the random walk model.[‖] Notice that if the particle moved ballistically (e.g., all steps were to the right) then the particle would have unit speed.

At time $t = n$ the particle has taken n steps with n_+ steps to the right and $n_- = n - n_+$ steps to the left. The probability of taking n_+ steps to the right is

$$\mathcal{P}_n(n_+) = \left(\frac{1}{2}\right)^n \frac{n!}{n_+!(n-n_+)!}$$

The problem is similar to finding the probability of getting exactly n_+ "heads" when flipping a fair coin n times. Note that we could do the more general case where the probability of taking a step to the right is p_+ and to the left is $p_- = (1 - p_+)$ for which we'd have,

$$\mathcal{P}_n(n_+) = p_+^{n_+}(1-p_+)^{n-n_+} \frac{n!}{n_+!(n-n_+)!}$$

But for simplicity we'll stay with the symmetric case where $p_+ = 1/2$.

The position of the particle is $m = n_+ - n_- = 2n_+ - n$ so the probability of the particle being at position m is

$$\mathcal{P}_n(m) = \left(\frac{1}{2}\right)^n \frac{((n+m)/2)!}{((n-m)/2)!}$$

We can now compute a few average quantities, starting with the average position,

$$\langle m \rangle = \sum_{m=-n}^{n} m\mathcal{P}_n(m)$$

For this calculation it will actually be easier to work with n_+ instead of m and convert using

$$\langle m \rangle = 2\langle n_+ \rangle - n$$

Remember that since taking an average is a linear operation (either summing or integrating over possible values) we have, in general,

$$\langle af(x) + bg(x) \rangle = a\langle f(x) \rangle + b\langle g(x) \rangle$$

where the average is over the probability $\mathcal{P}(x)$.

To find $\langle n_+ \rangle$ we use the binomial expansion,

$$(p+q)^n = \sum_{n_+=0}^{n} p^{n_+} q^{n-n_+} \frac{n!}{n_+!(n-n_+)!} = q^n + npq^{n-1} + \ldots + np^{n-1}q + p^n$$

We now use a common trick in manipulating these types of series and differentiate both sides,

$$\frac{d}{dp}(p+q)^n = \frac{d}{dp} \sum_{n_+=0}^{n} p^{n_+} q^{n-n_+} \frac{n!}{n_+!(n-n_+)!}$$

[‖] The historical, less politically correct name is the "drunkard's walk" model.

so

$$n(p+q)^{n-1} = \sum_{n_+=0}^{n} n_+ p^{n_+-1} q^{n-n_+} \frac{n!}{n_+!(n-n_+)!}$$

Now if we set $p = q = 1/2$ in this expression,

$$n = \sum_{n_+=0}^{n} n_+ \left(\frac{1}{2}\right)^n - 1 \frac{n!}{n_+!(n-n_+)!} = 2\sum_{n_+=0}^{n} n_+ \mathcal{P}_n(n_+)$$

so

$$\langle n_+ \rangle = \sum_{n_+=0}^{n} n_+ \mathcal{P}_n(n_+) = \frac{n}{2}$$

which is not surprising since, on average, we take $n/2$ steps to the right and $n/2$ steps to the left. From this result we directly find that $\langle m \rangle = 0$, which is also no surprise.

The more interesting result comes from repeating the differentiation trick again, which after a little algebra gives $\langle n_+^2 \rangle = n(n+1)/4$ so

$$\langle m^2 \rangle = \langle (2n_+ - n)^2 \rangle = 4\langle n_+^2 \rangle - 4n\langle n_+ \rangle + n^2 = n$$

The r.m.s. displacement of the particle is $\sqrt{\langle m^2 \rangle} = \sqrt{n}$; contrast this with the ballistic displacement (distance traveled if all the steps were in the same direction), which is n.

Simple Diffusion

To make a closer connection between the random walk model and the diffusive, random motion of a Brownian particle we'll formulate the probability distribution $\mathcal{P}_n(m)$ in the limit where n is large. Naturally, the first step is to use Sterling's approximation for the factorials,

$$\ln n! \approx n \ln n - n + \frac{1}{2}\ln 2\pi n$$

Note that we're using one more term than usual. Recalling our earlier result,

$$\begin{aligned}
\ln \mathcal{P}_n(m) &= -n\ln 2 + \ln n! - \ln(\tfrac{1}{2}(n+m)!) - \ln(\tfrac{1}{2}(n-m)!) \\
&\approx n\ln n - \tfrac{1}{2}(n+m)\ln(n+m) - \tfrac{1}{2}(n-m)\ln(n-m) + \tfrac{1}{2}\ln 2\pi n - \tfrac{1}{2}\ln(\pi^2(n^2-m^2)) \\
&= n\ln n - \tfrac{1}{2}n\ln(n^2-m^2) - \tfrac{1}{2}m\ln\frac{n+m}{n-m} + \tfrac{1}{2}\ln\frac{2n}{\pi(n^2-m^2)}
\end{aligned}$$

Next we'll take the limit of $n \gg m$, which is reasonable since we know that $\sqrt{\langle m^2 \rangle} = \sqrt{n}$ so the probability $\mathcal{P}_n(m)$ drops off rapidly with increasing $|m|$. This allows us to write,

$$\begin{aligned}
\ln \mathcal{P}_n(m) &= n\ln n - \tfrac{1}{2}n\ln\left[n^2(1-m^2/n^2)\right] - \tfrac{1}{2}m\ln\left(1+m/n)^2\right] + \tfrac{1}{2}\ln(2/\pi n) \\
&= -\tfrac{1}{2}n\ln\left[1-\frac{m^2}{n^2}\right] - m\ln(1+m/n) + \tfrac{1}{2}\ln(2/\pi n)
\end{aligned}$$

Finally, using $\ln(1+x) \approx x$ for $x \ll 1$ gives,

$$\ln \mathcal{P}_n(m) \approx \frac{1}{2}\frac{m^2}{n} - \frac{m^2}{n} + \frac{1}{2}\ln(2/\pi n)$$

or

$$\mathcal{P}_n(m) \propto \sqrt{\frac{1}{n}}\; e^{-m^2/2n}$$

Not surprisingly, after all those approximations we have to multiply by the appropriate constant to normalize the distribution.** Doing so yields,

$$\mathcal{P}_n(m) = \sqrt{\frac{1}{2\pi n}}\; e^{-m^2/2n}$$

which is a Gaussian distribution with mean $\langle m\rangle = 0$ and variance $\langle (m - \langle m\rangle)^2\rangle = \langle m^2\rangle = \sqrt{n}$.

Passing from discrete to continuous space ($m \to x$) and from discrete steps to continuous time ($n \to t$) we obtain,

$$\mathcal{P}(x,t) = \sqrt{\frac{1}{4\pi D t}}\; e^{-x^2/4Dt} \qquad (*)$$

where D is the diffusion coefficient. This coefficient is implicit in the assumption in the random walk model that the steps are unit length and occur at unit times. The mean and variance of this distribution are,

$$\begin{aligned}\langle x\rangle &= \int_{-\infty}^{\infty} dx\, x\mathcal{P}(x,t) = 0\\ \langle (x-\langle x\rangle)^2\rangle &= \int_{-\infty}^{\infty} dx\, (x-\langle x\rangle)^2\mathcal{P}(x,t) = 2Dt\end{aligned}$$

so the r.m.s. displacement is $\sqrt{\langle (x-\langle x\rangle)^2\rangle} = \sqrt{\langle x^2\rangle} = \sqrt{2Dt}$.

You may already be familiar with the fact that this Gaussian function (*) is the solution of the diffusion equation,

$$\frac{\partial}{\partial t}\mathcal{P}(x,t) = D\frac{\partial^2}{\partial x^2}\mathcal{P}(x,t)$$

for the initial condition,

$$\mathcal{P}(x,t) = \delta(x)$$

and no boundary conditions (i.e., x extends from $-\infty$ to ∞). From inspection of (*) the amplitude of $\mathcal{P}(x,t)$ at $x = 0$ diverges for $t = 0$ while the standard deviation goes to zero as $t \to 0$, which verifies that the Gaussian satisfies the Dirac delta function initial condition.

All of the analysis may easily be extended to three-dimensional systems by assuming that the motion in each of the x, y, and z directions is independent. From this,

$$\mathcal{P}(\vec{r},t) = \mathcal{P}(x,t)\mathcal{P}(y,t)\mathcal{P}(z,t) = \left(\sqrt{\frac{1}{4\pi D t}}\right)^3 e^{-(x^2+y^2+z^2)/4Dt} = (4\pi D t)^{-3/2}\, e^{-r^2/4Dt}$$

with $\langle \vec{r}\rangle = 0$, $\langle r^2\rangle = \langle x^2+y^2+z^2\rangle = 6Dt$, and

$$\frac{\partial}{\partial t}\mathcal{P}(\vec{r},t) = D\nabla^2\mathcal{P}(\vec{r},t)$$

as the corresponding three-dimensional diffusion equation.

On last comment: Notice that in this section we, for the first time in this course, introduced the element of time. Equilibrium statistical mechanics is time-less, not in the conventional meaning of the term but literally; at equilibrium it is meaningless to speak about the past or the future. With the study of Brownian motion we are entering the realm of non-equilibrium statistical mechanics.

[**] We normalize by requiring that the integral of $\mathcal{P}_n(m)$ from $m = -\infty$ to ∞ equals one.

Langevin Theory of Brownian Motion

In the analysis above we considered Brownian motion from a microscopic point-of-view, starting from a random walk model and taking the macroscopic limit. We now analyze Brownian motion from the other direction, starting from a macroscopic, mechanical perspective and introducing microscopic effects due to fluctuations. From Newton's second law of motion we know that the motion of the Brownian particle (mass M) suspended in the fluid is given by,

$$M\frac{d\vec{v}}{dt} = \vec{F}_{\rm drag} + \vec{F}_{\rm random}$$

For large objects only the drag force is significant but for Brownian particles we must also consider the random force due to fluctuations. As before, we assume the Brownian particle is neutrally buoyant, that is, the buoyant force exactly balances gravity; the main simplification is that we won't have to consider the falling or rising motion due to these forces (although that's certainly possible to include in this framework).

For a spherical Brownian particle of radius a the drag force is given by Stokes' law,

$$\vec{F}_{\rm drag} = -6\pi\eta a\vec{v} = -\vec{v}/B$$

where η is the viscosity of the fluid suspending the Brownian particle. The mobility, B, is defined as the terminal velocity for a particle that is pulled by an external force of unit magnitude. Note that the smaller the particle or the lower the viscosity, the higher the mobility.

The random force is, on average, zero but we cannot neglect it since it is the source of the Brownian motion. The random force has a non-zero variance, but instead of tackling that question directly we'll take an easier, if indirect route. By the way, Pathria and Beale make the distinction between a time-average (indicated by overbar, $\overline{f}$) and ensemble-average (indicated by angle braces, $\langle f\rangle$). I won't bother with this subtlety and will use the latter for both; use the context to decide which is more appropriate.

To start, because $\langle\vec{F}_{\rm random}\rangle = 0$,

$$M\frac{d}{dt}\langle\vec{v}\rangle = \langle\vec{F}_{\rm drag}\rangle = -\frac{1}{B}\langle\vec{v}\rangle$$

This ODE is trivial to solve,

$$\langle\vec{v}(t)\rangle = \vec{v}(0)e^{-t/MB} = \vec{v}(0)e^{-t/\tau}$$

where $\tau = MB$ is the relaxation time for the drag force. Returning to our original equation, we may now write it as,

$$\frac{d}{dt}\vec{v} = -\frac{1}{\tau}\vec{v} + \frac{1}{M}\vec{F}_{\rm random}(t)$$

Of course it's rather meaningless for us to try to solve this for $\vec{v}(t)$ due to the random force; for an ensemble of Brownian particles each one will have it's own trajectory and velocity. But it is meaningful to compute variances for position and velocity, which is what we will now do.

First, apply $\vec{r}\cdot$ to both sides of the equation above,

$$\vec{r}\cdot\frac{d}{dt}\vec{v} = -\frac{1}{\tau}\vec{r}\cdot\vec{v} + \frac{1}{M}\vec{r}\cdot\vec{F}_{\rm random}(t)$$

or

$$\frac{d}{dt}(\vec{r}\cdot\vec{v}) - \frac{d\vec{r}}{dt}\cdot\vec{v} = -\frac{1}{\tau}\vec{r}\cdot\vec{v} + \frac{1}{M}\vec{r}\cdot\vec{F}_{\rm random}(t)$$

But since

$$\vec{r}\cdot\vec{v} = \vec{r}\cdot\frac{d\vec{r}}{dt} = \frac{1}{2}\frac{d}{dt}(\vec{r}\cdot\vec{r}) = \frac{1}{2}\frac{d}{dt}r^2$$

we have

$$\frac{1}{2}\frac{d^2}{dt^2}r^2 - v^2 = -\frac{1}{2\tau}\frac{d}{dt}r^2 + \frac{1}{M}\vec{r}\cdot\vec{F}_{\mathrm{random}}(t)$$

and finally taking the average of both sides,

$$\frac{1}{2}\frac{d^2}{dt^2}\langle r^2\rangle - \langle v^2\rangle = -\frac{1}{2\tau}\frac{d}{dt}\langle r^2\rangle + \frac{1}{M}\langle\vec{r}\cdot\vec{F}_{\mathrm{random}}(t)\rangle$$

Recall that our random walk description of Brownian motion gave $\langle r^2\rangle \propto t$ and this is the result we're working towards. But we're not simply duplicating the previous results because ultimately we also hope the find how this variance depends on physical quantities, such as the fluid viscosity and the particle's mass and radius.

The next two steps of the analysis are: 1) Assume that the random force is uncorrelated with the particle's position so $\langle\vec{r}\cdot\vec{F}_{\mathrm{random}}(t)\rangle = 0$. 2) Use equipartition theorem,

$$\langle \tfrac{1}{2}Mv^2\rangle = \frac{3}{2}kT$$

so $\langle v^2\rangle = 3kT/M$. With these two steps,

$$\frac{1}{2}\frac{d^2}{dt^2}\langle r^2\rangle + \frac{1}{2\tau}\frac{d}{dt}\langle r^2\rangle = \frac{6kT}{M}$$

You can check that the solution to this ODE may be written as,

$$\langle r^2\rangle = \frac{6kT\tau^2}{M}\left(\frac{t}{\tau} - c_1(1-e^{-t/\tau})\right) + c_2$$

where the constants c_1 and c_2 are set by the initial conditions (for example, if initially $\langle r^2\rangle = d\langle r^2\rangle/dt = 0$ then $c_1 = 1$ and $c_2 = 0$). For times longer than the initial transient, that is, for $t \gg \tau$ the result is,

$$\langle r^2\rangle = \left(\frac{6kT\tau}{M}\right)t = \left(\frac{kT}{\pi\eta a}\right)t$$

Interestingly, this was one of the first experimental methods for measuring Boltzmann's constant, k, and subsequently for computing Avogadro's number.

This derivation gives the correct results in the long time limit but it's not entirely rigorous because we assumed that $\langle v^2\rangle = 3kT/M$, which is correct as $t\to\infty$ but depending on the initial conditions for the Brownian particle this may not be accurate (e.g., if the particle is initially at rest). Let's back up to the original ODE,

$$\frac{d}{dt}\vec{v}(t) = -\frac{1}{\tau}\vec{v} + \frac{1}{M}\vec{F}_{\mathrm{random}}(t)$$

The general solution of this inhomogeneous ODE is,

$$\vec{v}(t) = \vec{v}(0)e^{-t/\tau} + \frac{1}{M}\,e^{-t/\tau}\int_0^t e^{t'/\tau}\vec{F}_{\mathrm{random}}(t')\,dt'$$

Squaring both sides and applying the average we find,

$$\langle v^2(t)\rangle = v^2(0)e^{-2t/\tau} + \frac{1}{M^2}e^{-2t/\tau}\int_0^t dt'\int_0^t dt''e^{(t'+t'')/\tau}\langle\vec{F}_{\mathrm{random}}(t')\cdot\vec{F}_{\mathrm{random}}(t'')\rangle$$

Notice that the cross-term is zero since $\langle\vec{F}_{\mathrm{random}}\rangle = 0$.

After lots of manipulations Pathria and Beale arrive at the result,

$$\langle v^2(t)\rangle = v^2(t)e^{-2t/\tau} + C\tau(1 - e^{-2t/\tau})$$

where

$$C = \frac{1}{M^2}\int_{-\infty}^{\infty} \langle \vec{F}_{\text{random}}(t') \cdot \vec{F}_{\text{random}}(0)\rangle \; dt'$$

is the integral of the time-correlation of the random force. Fortunately, we don't have to evaluate this integral because we can, instead, use the equipartition function to fix the value of C. Specifically, since $\langle v^2(t)\rangle \to 3kT/M$ as $t \to \infty$ then it must be the case that $C = 6kT/M\tau$.

I'll just quote the final result for the variance of the position,

$$\begin{aligned}\langle r^2(t)\rangle &= \frac{6kT\tau}{M}t + v^2(0)\tau^2(1 - e^{-t/\tau})^2 \\ &\quad -\frac{3kT}{M}\tau^2(1 - e^{-t/\tau})(3 - e^{-t/\tau})\end{aligned}$$

For long times ($t \gg \tau$) only the first term remains, which is the same as our earlier result. See Figures 15.3 and 15.4 in Pathria and Beale for graphs of $\langle r^2(t)\rangle$ and $\langle r^2(t)\rangle$ as functions of time for various initial conditions.

One last thing about this Langevin theory analysis: Since we know that the time-correlation function of the random force is given by,

$$\frac{6kTM}{\tau} = \int_{-\infty}^{\infty} \langle \vec{F}_{\text{random}}(t') \cdot \vec{F}_{\text{random}}(t'')\rangle \; dt'$$

we can use this to model the random force term. For example, we can use the white noise model, for which the random force varies so rapidly that it immediately de-correlates, that is,

$$\langle \vec{F}_{\text{random}}(t)\vec{F}_{\text{random}}(t')\rangle = 0 \quad \text{if} \quad t \neq t'$$

Then we can write the random force as,

$$\vec{F}_{\text{random}}(t) = \sqrt{\frac{6kTM}{\tau}}\mathcal{W}(t)$$

where the white noise function has the properties that

$$\langle \mathcal{W}(t)\rangle = 0 \qquad\qquad \langle \mathcal{W}(t)\mathcal{W}(t')\rangle = \delta(t - t')$$

This formulation is useful for both theoretical analysis and numerical stochastic modeling.

Chapter 10

Computer Simulations

Metropolis Monte Carlo Method (16.2)

Lecture xx

A natural (and early) application of computers is the calculation of statistical properties for interacting systems. For example, consider a quantity of interest, X; we'll assume that for any given state in the ensemble we can evaluate the value of X_i for state i. For example if you tell me the up or down value of all the spins in an Ising lattice then it's easy enough to calculate the net magnetization. In general we evaluate the mean value as,

$$\langle X \rangle = \sum_{i=1}^{M} X_i \mathcal{P}_i$$

In the microcanonical ensemble we evaluate the mean value as,

$$\langle X \rangle = \frac{1}{M} \sum_{i=1}^{M} X_i$$

since $\mathcal{P}_i = 1/M$ (i.e., all states are equally probable). In the canonical ensemble we evaluate the mean value as,

$$\langle X \rangle = \sum_{i=1}^{M} X_i \mathcal{P}_i = \frac{\sum_{i=1}^{M} X_i e^{\beta E_i}}{\sum_{i=1}^{M} e^{-\beta E_i}} = \frac{1}{Q_N} \sum_{i=1}^{M} X_i e^{-\beta E_i}$$

where $\beta = 1/kT$. As mentioned before, the direct evaluation of these summations is not practical due to the huge numbers of states, even in small systems. For example, for a three-dimensional Ising lattice with 10 by 10 by 10 sites the number of states is $M = 2^{1000} \approx 10^{300}$ configurations. The next generation of supercomputers will be able to perform 10^{18} operations per second (an exaFLOP) so we're still 282 orders of magnitude short of the necessary computational horsepower for the brute-force calculation.

The next option is not to evaluate the sums for all M states but instead to select a sub-sample of M' states with $M' \ll M$; in the micro-canonical ensemble this would be just,

$$\langle X \rangle = \sum_{i=1}^{M} X_i \mathcal{P}_i = \frac{1}{M} \sum_{i=1}^{M} X_i \approx \frac{1}{M'} \sum_{i'=1}^{M'} X_{i'}$$

This is analogous to doing a sample poll of a population instead of a full census. Of course the set of states M' that are selected must be an unbiased sample otherwise our estimate of $\langle X \rangle$ will be skewed.

For micro-canonical ensemble this approach is not practical because it's too difficult to find states that satisfy the constraint on total energy. The canonical ensemble is much more liberal in allowing us to pick states but then each state is weighed with probability $\mathcal{P}_i \propto e^{-\beta E_i}$. Unfortunately, if we pick states at random we find that $\mathcal{P} \approx 0$ for nearly all of the states. Nearly any random sample will have a large error because, given that M is enormous, it's unlikely that a random subset of M' states will contain any of the important states for which $\mathcal{P}_i$ is significantly different from zero. This is analogous to doing a random poll of people in the US asking them what it was like to walk on the Moon.

The alternative is to not choose states entirely at random but rather to pick mostly the important ones; this is called importance sampling. Instead of choosing M' states at random we will choose M^* states and Π_i the probability of choosing state i for this sample; in the canonical ensemble we would evaluate,

$$\langle X \rangle = \sum_{i=1}^{M} X_i \mathcal{P}_i = \sum_{i=1}^{M} \frac{X_i}{\Pi_i} \Pi_i \mathcal{P}_i \approx \frac{1}{M^*} \sum_{i^*=1}^{M^*} \frac{X_{i^*}}{\Pi_{i^*}} \mathcal{P}_{i^*}$$

where the states i^* are selected with probability Π_{i^*}. Notice that if states are chosen at random then $\Pi_{i^*} = 1/M^*$, which is what we had before.

Ideally, we'd like to choose the most important states so we'd like to use $\Pi_i \approx \mathcal{P}_i$. That means that we need to find an efficient way of choosing states with the specified probability Π_i. The key to finding these important states in an efficient manner is not to draw them independently but to pick the states close to each other in phase space; typically if a state has high probability then nearby states are also high probability states. To continue our analogy, if you want to poll persons who've walked on the moon you start with Neil Armstrong's address book.

To locate states near each other in phase space we'll use the transition probability, $W_{i\to j}$, which is the probability that starting from state i we pick state j as our next sample state. Moving between states in this fashion, if we call $\hat{\Pi}_i(t)$ the probability that we're at state i at time t then this probability obeys a Master equation,

$$\frac{d}{dt}\hat{\Pi}_i(t) = \sum_{j}^{M} \left[-W_{i\to j}\hat{\Pi}_i + W_{j\to i}\hat{\Pi}_j \right]$$

The first term inside the sum is the rate at which we leave state i, which is equal to the probability that we're at i times the rate at which we jump to any other state j. The second term is the rate at which we jump into i from any other state, which equals $W_{j\to i}$ times the probability that we're in state j (i.e., $\hat{\Pi}_j(t)$). Of course for the total rate of change both the "loss" and "gain" terms are summed over all states j, whether they are the destination state (in the first term) or the source state (in the second term).[*]

Assuming that this process has a steady state ($d\hat{\Pi}_i/dt = 0$) we have that as $t \to \infty$, $\hat{\Pi}_i \to \Pi_i$ with

$$\sum_{j}^{M} \left[-W_{i\to j}\Pi_i + W_{j\to i}\Pi_j \right] = 0$$

To get this we'll impose the condition of *detailed balance* on our transition probability, that is,

$$\Pi_i W_{i\to j} = \Pi_j W_{j\to i}$$

This ensures that if we take a large number of steps the probability of landing on state i is Π_i.

---

[*] This type of formulation also appears in several statistical mechanical theories, such as the Boltzmann equation for dilute gases.

Since we'd like to have $\Pi_i = \mathcal{P}_i$ in the canonical ensemble we have,

$$\frac{W_{i\to j}}{W_{j\to i}} = \frac{\Pi_j}{\Pi_i} = \frac{\mathcal{P}_j}{\mathcal{P}_i} = \frac{(Q_N)^{-1}e^{-\beta E_j}}{(Q_N)^{-1}e^{-\beta E_j}} = e^{-\beta(E_j - E_i)} \qquad (*)$$

Note an important feature: Because we only need the ratio of $\mathcal{P}_j/\mathcal{P}_i$ we don't need $\mathcal{P}_i$ explicitly; this is key because computing Q_N for the normalization of $\mathcal{P}_i$ is just as difficult as computing $\langle X\rangle$.

A simple way to construct $W_{i\to j}$ such that it satisfies (*) is to use,

$$W_{i\to j} = \begin{cases} e^{-\beta(E_j - E_i)} & E_i < E_j \\ 1 & E_i \geq E_j \end{cases}$$

That is, if the energy of state j is higher than that of state i by and amount $\Delta E = E_j - E_i$ then make the transition to state j with probability proportional to $e^{-\beta\Delta E}$; if state j has lower energy ($\Delta E < 0$) then automatically take the step. Since this transition rate gives $\Pi_i = \mathcal{P}_i$ we finally have our estimate,

$$\langle X\rangle \approx \frac{1}{M^*}\sum_{i^*=1}^{M^*} \frac{X_{i^*}}{\Pi_{i^*}}\mathcal{P}_{i^*} = \frac{1}{M^*}\sum_{i=1}^{M^*} X_{i^*}$$

where the M^* states are generated by stepping from one to the next. This procedure was first proposed by Metropolis *et al.*[†] in 1953 but was independently discovered by a number of other computational physicists and chemists at about the same time. It's now known as the Metropolis Monte Carlo method or sometimes as just the Monte Carlo method.[‡]

Let's see how the Metropolis Monte Carlo method might be used in practice. For the Ising model we could take a lattice of N spin sites and ask what is the value of $\langle N_+\rangle$, the average number of "up" sites at a given temperature T. [§] Starting from an arbitrary initial configuration we take a step to the next configuration by choosing a spin site at random and then making a random choice as to whether or not to flip that spin. If flipping the selected spin site causes it to have the same or more aligned neighbors (i.e., if $\Delta E \leq 0$) then we automatically make that transition. On the other hand if the transition is energetically unfavorable then we generate a random number, $\mathcal{R}$, between zero and one; if $\mathcal{R} \leq \exp(-\beta\Delta E)$ then the spin is flipped otherwise the system remains unchanged. Whatever the outcome we now have our new state and a sample value for N_+. We continue repeating this process for the desired number of M^* samples until we get the desired accuracy for $\langle N_+\rangle$; the details on how to estimate the accuracy of this sample average are discussed by Pathria and Beale in Section 16.1.

To find the equation of state for dense gas the procedure is similar. We select random values for the positions of the particles in a volume V; as we saw before we only need to estimate the pair-correlation function to get the non-ideal contribution for the equation of state. We pick a random particle, change it's position by a random increment $\Delta\vec{r}$ and compute ΔE, the difference in total energy ΔE based on the intermolecular potential energy $\mathcal{V}(\vec{r}_1, \ldots, \vec{r}_N)$. Typically $|\Delta\vec{r}|$ is taken as a fraction of a molecular diameter to minimize the chance of particles overlapping (i.e., the strong repulsion term can make ΔE very large, which means almost certain rejection of the move). Once the move is accepted or rejected according to the Metropolis Monte Carlo procedure the quantities of interest are measured and the procedure repeats. The original Metropolis Monte Carlo calculations were of this type since the equations of state for materials at extreme conditions were of particular interest to the scientists at Los Alamos in the early 1950's.

[†] J. Chem. Phys. **21** 1087 (1953)

[‡] The latter term can be confusing given the wide variety of stochastic algorithms that are generally called Monte Carlo methods.

[§] Note that $\langle s\rangle = 2\langle N_+\rangle - N$.

One last note: It's important to keep in mind that the transitions between states in the Metropolis Monte Carlo (MMC) method have nothing to do with the actual physical dynamics of an interacting system. That is, if we took an initial condition in phase space, say the positions and velocities of particles, and computed the real dynamics (say, from the Lagrangian) this would trace out a trajectory in phase space. At equilibrium the states visited by that trajectory have the same equilibrium distribution as the states visited by the MMC method. But these two ways of traveling through phase space have completely different trajectories. Another way to say this is that with MMC the only quantity of physical interest is the set of sample values; the order in which we get those samples has no physical meaning.

Molecular Dynamics

The interactions between particles in a gas or liquid often may be accurately modeled using classical mechanics. Assume the particles interact by a pairwise force that depends only on the relative separation

$$\mathbf{F}_{ij} = \mathbf{F}(\mathbf{r}_i - \mathbf{r}_j) = -\mathbf{F}_{ji}$$

where $\mathbf{F}_{ij}$ is the force on particle i due to particle j; the positions of the particles are $\mathbf{r}_i$ and $\mathbf{r}_j$, respectively. The explicit form for $\mathbf{F}$ may either be approximated from experimental data or computed theoretically using quantum mechanics.

Once we fix the interparticle force, the dynamics is given by the equation of motion

$$\frac{d^2}{dt^2}\mathbf{r}_i = \frac{1}{m}\sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{F}_{ij}$$

where m is the mass of a particle. From the initial conditions, in principle, the future state can be computed by evaluating this system of ODEs. This numerical approach is called *molecular dynamics* and it has been very successful in computing microscopic properties of fluids; Pathria and Beale give a short introduction in Sections 16.3 and 16.4.

Though simple in concept the simulations can be computationally expensive unless sophisticated numerical techniques are used. For example, the calculation of $\mathbf{F}_{ij}$ for all pairs is a computation that requires $O(N^2)$ operations; when dealing with millions of particles this makes the calculation intractable. But since the intermolecular force is typically short-ranged it really only needs to be computed for particles closer than a specified radius of influence. The sophisticated approach is to maintain neighbor lists so that the computation of $\mathbf{F}_{ij}$ can be reduced to $O(N)$ or $O(N \ln N)$.

Kinetic theory

Let's turn our attention now to ideal gases. Although an ideal gas is simple at equilibrium, when the gas is out of equilibrium, say flowing due to a pressure gradient, it has all the rich, interesting behavior found in hydrodynamics. Let's consider how to model the microscopic dynamics of a gas so as to formulate efficient numerical methods for simulating it.

Consider the following model for a monatomic gas: a system of volume V contains N particles. These particles interact, but since the gas is dilute, the interactions are always two-body collisions. The criterion for a gas to be dilute is that the distance between the particles is large compared to d, the effective diameter of the particles. This effective diameter may be measured, for example, by scattering experiments. Our criterion for a gas to be considered dilute may be written as

$$d \ll \sqrt[3]{V/N}$$

An alternative view of this criterion is to say that a gas is dilute if the volume occupied by the particles is a small fraction of the total volume.

In Boltzmann's time there was no hope of evaluating the equations of motion numerically, and even today molecular dynamics is limited to relatively small systems. To understand the scale of the problem, consider that in a dilute gas at standard temperature and pressure, the number of particles in a cubic centimeter, Loschmidt's number, is 2.687×10^{19}. A molecular dynamics simulation of a dilute gas containing a million particles represents a volume of 0.037 cubic microns. Even on a supercomputer, an hour of computer time will evolve the system for only a few nanoseconds of physical time.

Maxwell-Boltzmann Distribution

Instead of being overwhelmed by the huge numbers, we can use them to our advantage. The basic idea of statistical mechanics is to abandon any attempt to predict the instantaneous state of a single particle. Instead, we obtain probabilities and compute average quantities, for example, the average speed of a particle. The large numbers of particles now work in our favor because even in a very small volume we are averaging over a very large sample.

For a dilute gas we usually take the gas to be ideal; that is, we assume that a particle's energy is all kinetic energy,

$$E(\mathbf{r}, \mathbf{v}) = \frac{1}{2} m |\mathbf{v}|^2$$

In a dilute gas this is a good approximation, since the interparticle forces are short-ranged. The probability that a particle in this system is at a position between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ with a velocity between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$ is

$$P(\mathbf{r}, \mathbf{v})\, d\mathbf{r} d\mathbf{v} = A \exp(-E(\mathbf{r}, \mathbf{v})/kT)\, d\mathbf{r} d\mathbf{v}$$

The constant A is a normalization that is fixed by the condition that the integral of the probability over all possible states must equal unity. The differential elements, $d\mathbf{r}d\mathbf{v}$, on each side of the equation serve to remind us that $P(\mathbf{r}, \mathbf{v})$ is a *probability density.*

Since a particle's energy is independent of $\mathbf{r}$, the probability density may be written as

$$\begin{aligned} P(\mathbf{r}, \mathbf{v})\, d\mathbf{r} d\mathbf{v} &= [P_{\mathbf{r}}(\mathbf{r}) d\mathbf{r}][P_{\mathbf{v}}(\mathbf{v}) d\mathbf{v}] \\ &= \left[\frac{1}{V} d\mathbf{r}\right] [P_{\mathbf{v}}(\mathbf{v}) d\mathbf{v}] \end{aligned}$$

The particle is equally likely to be anywhere inside the volume V. For example, suppose that we demark a subregion α inside our system. The probability that the particle is inside α is

$$\iiint_\alpha P_{\mathbf{r}}(\mathbf{r})\, d\mathbf{r} = \frac{1}{V} \iiint_\alpha d\mathbf{r} = \frac{V_\alpha}{V}$$

where V_α is the volume of subregion α.

We may further simplify our expression for the probability by making use of the isotropy of the distribution. In spherical coordinates, the probability that a particle has a velocity between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$ is

$$\begin{aligned} P_{\mathbf{v}}(\mathbf{v})\, d\mathbf{v} &= P_{\mathbf{v}}(v, \theta, \phi) v^2 \sin\theta\, dv\, d\theta\, d\phi \\ &= A(e^{-\frac{1}{2} m v^2 / kT} v^2\, dv)(\sin\theta\, d\theta)(d\phi) \end{aligned}$$

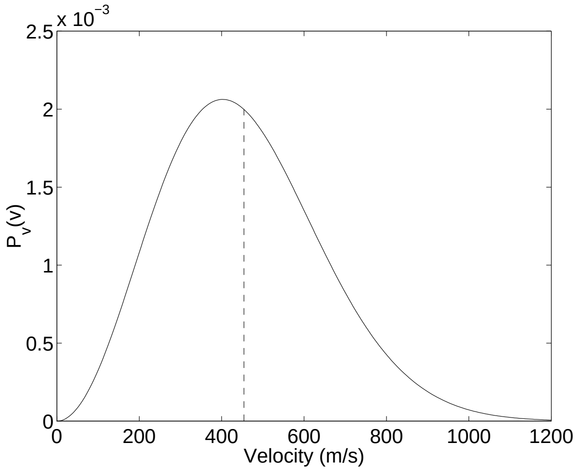

Figure 10.1: Maxwell-Boltzmann distribution of particle speed for nitrogen at $T = 273$ K. The dashed line marks the average speed, $\langle v\rangle$.

Since the distribution of velocities is isotropic, the angular parts can be integrated to give

$$P_v(v)\,dv = 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} v^2 e^{-\frac{1}{2}mv^2/kT} dv$$

where $P_v(v)dv$ is the probability that a particle's speed is between v and $v + dv$. Notice that we finally fixed the normalization constant A by imposing the condition that $\int_0^\infty P_v(v)dv = 1$. This velocity distribution is known as the *Maxwell-Boltzmann distribution* (see Figure 10.1).

Using the Maxwell-Boltzmann distribution it is not difficult to compute various average quantities. For example, the average particle speed

$$\langle v\rangle = \int_0^\infty vP_v(v)\,dv = \frac{2\sqrt{2}}{\sqrt{\pi}}\sqrt{\frac{kT}{m}}$$

and the root mean square (r.m.s.) particle speed

$$\sqrt{\langle v^2\rangle} = \sqrt{\int_0^\infty v^2 P_v(v)\,dv} = \sqrt{3}\sqrt{\frac{kT}{m}}$$

From this, the average kinetic energy of a particle is,

$$\langle K\rangle = \langle \frac{1}{2}mv^2\rangle = \frac{3}{2}kT$$

in agreement with the equipartition theorem.

Finally, the most probable speed, v_{mp}, is not an average, but rather is the speed at which $P_v(v)$ has a maximum. Solving

$$\left.\frac{d}{dv}P_v(v)\right|_{v=v_{\mathrm{mp}}} = 0$$

gives

$$v_{\mathrm{mp}} = \sqrt{2}\sqrt{\frac{kT}{m}}$$

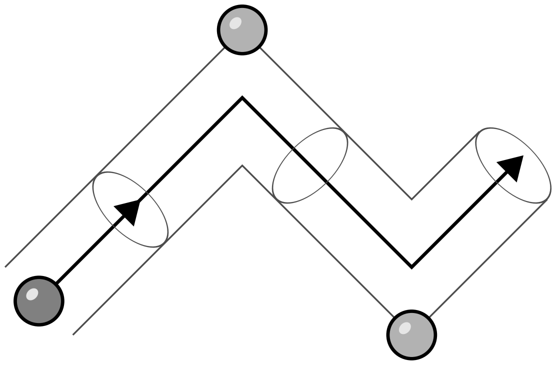

Figure 10.2: Test particle moving amid a field of stationary particles.

Notice that $v_{\mathrm{mp}} < \langle v \rangle < \sqrt{\langle v^2 \rangle}$, but they are all of comparable magnitude and approximately equal to the speed of sound

$$v_{\mathrm{s}} = \sqrt{\gamma}\sqrt{\frac{kT}{m}}$$

where $\gamma = c_{\mathrm{p}}/c_{\mathrm{v}}$ is the ratio of specific heats. In a monatomic gas, $\gamma = 5/3$.

Collision Frequency and Mean Free Path

In general, the particles in a dilute gas interact when their separation is of the order of their effective diameter. While the interaction between particles is continuous, it is also short-ranged, so it is useful to think of the particles as colliding. In the *hard-sphere model* we picture the dilute gas as a cloud of tiny billiard balls of diameter d. Particles collide elastically when their separation equals this diameter. We use this model throughout this chapter because it does a surprisingly good job of representing a real gas.

The average number of collisions per particle per unit time is called the *collision frequency*, f. A related quantity is the *mean free path*, λ, which is the average distance traveled by a particle between collisions. To compute these two quantities for a hard-sphere gas, consider the following picture. A test particle moves with speed v_{r} amid a sea of stationary particles (Figure 10.2). The particle travels a zigzag path much like the ball in a pinball machine.[¶] In a time increment t_0, the test particle travels a distance $l = v_{\mathrm{r}} t_0$.

Imagine a cylindrical tube of radius d centered on the path of the test particle. This tube has "elbows" at the locations of collisions, but we approximate its volume as $\pi d^2 l$. The number of stationary particles contained within this cylinder equals the number of collisions experienced by the test particle. On the other hand, since the gas is homogeneous, the number of particles in the tube is

$$\begin{aligned} M_{\mathrm{tube}} &= \text{(number of particles per unit volume)} \times \text{(volume of tube)} \\ &= (N/V)(\pi d^2 v_{\mathrm{r}} t_0) \end{aligned}$$

which is also the number of collisions in a time t_0.

The collision frequency is

$$f = \frac{\langle M_{\mathrm{tube}} \rangle}{t_0} = \frac{N}{V}\pi d^2 \langle v_{\mathrm{r}} \rangle$$

[¶] A primitive mechanical entertainment device in common use before the advent of video games.

There remains one problem: We don't really have just one moving particle; all particles move. This issue is resolved by identifying $\langle v_\mathrm{r}\rangle$ as the average relative speed between particles

$$\begin{aligned}\langle v_\mathrm{r}\rangle &= \langle|\mathbf{v}_1-\mathbf{v}_2|\rangle = \int\!\!\int |\mathbf{v}_1-\mathbf{v}_2| P_\mathbf{v}(\mathbf{v}_1) P_\mathbf{v}(\mathbf{v}_2)\, d\mathbf{v}_1\, d\mathbf{v}_2 \\ &= \frac{4}{\sqrt{\pi}}\sqrt{\frac{kT}{m}}\end{aligned}$$

where $P_\mathbf{v}(\mathbf{v})$ is given by the Maxwell-Boltzmann distribution. From the collision frequency, the mean free path is obtained as

$$\begin{aligned}\lambda &= \text{(average particle speed)} \times \text{(average time between collisions)} \\ &= \langle v\rangle \frac{1}{f} = \frac{V}{N\pi d^2}\frac{\langle v\rangle}{\langle v_\mathrm{r}\rangle}\end{aligned}$$

Using the expressions above for the average particle speed and the mean free path,

$$\lambda = \frac{V}{\sqrt{2}N\pi d^2}$$

Notice that the mean free path depends only on density and particle diameter; it is independent of temperature.

Using kinetic theory, we can design a numerical simulation of a dilute gas. Instead of solving the deterministic equations of motion, we will build a stochastic model. The probability arguments discussed in this section give us the framework for the numerical scheme.

Direct Simulation Monte Carlo (DSMC)

General Algorithm

We now turn to the problem of constructing a numerical simulation for a dilute gas. Again, we don't want to compute the trajectory of every particle. Instead, we'll use kinetic theory to build a stochastic model. The scheme is loosely based on the Boltzmann equation; it was popularized as a practical numerical algorithm by G. A. Bird. He named it direct simulation Monte Carlo (DSMC) and it has been called "the dominant predictive tool in rarefied gas dynamics for the past decade."

The DSMC algorithm is like molecular dynamics in that the state of the system is given by the positions and velocities of the particles, $\{\mathbf{r}_i, \mathbf{v}_i\}$, for $i = 1, \ldots, N$. A useful concept in these types of simulations is that of representative particles. If each particle were to represent only a single molecule, a simulation of ambient air would need about 27 million particles per cubic micron. Instead, each particle in the simulation is assumed to represent N_ef molecules in the physical system that are roughly at the same position with roughly the same velocity. The totally democratic dynamics of the real system is represented in the simulation by a parliamentary subset. Of course, the simulation will not be accurate if the number of particles, N, is too small. Surprisingly, using 20 or more particles per cubic mean free path is usually sufficient. The concept of representative particles allows us to rescale length and time to model larger systems. For example, the system volume is $V = (NN_\mathrm{ef})/n$, where n is the number density.

The evolution of the system is integrated in time steps, τ, which are typically on the order of the mean collision time for a particle. At each time step, the particles are first moved as if they did not interact with each other. Every particle's position is reset as $\mathbf{r}_i(t+\tau) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\tau$. In this section we study homogeneous problems, but formulate the DSMC algorithm for inhomogeneous systems, in anticipation of the next section. After the particles move, some are selected to collide. The rules for this random selection process are obtained from kinetic theory. After the velocities of all colliding particles have been reset, the process is repeated for the next time step.

Collisions

Intuitively, we would want to select only particles that were near each other as collision partners. In other words, particles on opposite sides of the system should not be allowed to interact. To implement this condition, the particles are sorted into spatial cells and only particles in the same cell are allowed to collide. We could invent more complicated schemes, but this one works well, as long as the dimension of a cell is no larger than a mean free path.

In each cell, a set of representative collisions is processed at each time step. All pairs of particles in a cell are considered to be candidate collision partners, regardless of their positions within the cell. In the hard-sphere model, the collision probability for the pair of particles, i and j, is proportional to their relative speed,

$$P_{\mathrm{coll}}[i,j] = \frac{|\mathbf{v}_i - \mathbf{v}_j|}{\sum_{m=1}^{N_\mathrm{c}} \sum_{n=1}^{m-1} |\mathbf{v}_m - \mathbf{v}_n|}$$

where N_c is the number of particles in the cell. Notice that the denominator serves to normalize this discrete probability distribution.

It would be computationally expensive to use the above expression directly, because of the double sum in the denominator. Instead, the following acceptance-rejection scheme is used to select collision pairs:

1. A pair of candidate particles, i and j, is chosen at random.

2. Their relative speed, $v_\mathrm{r} = |\mathbf{v}_i - \mathbf{v}_j|$, is computed.

3. The pair is accepted as collision partners if $v_\mathrm{r} > v_\mathrm{r}^{\max}\Re$, where $v_\mathrm{r}^{\max}$ is the maximum relative velocity in the cell and $\Re$ is a uniform deviate in [0, 1).

4. If the pair is accepted, the collision is processed and the velocities of the particles are reset.

5. After the collision is processed or if the pair is rejected, return to step 1.

This acceptance-rejection procedure exactly selects collision pairs according to the desired collision probability. The method is also exact if we overestimate the value of $v_\mathrm{r}^{\mathrm{new}}$, although it is less efficient in the sense that more candidates are rejected. On the whole, it is computationally cheaper to make an intelligent guess that overestimates $v_\mathrm{r}^{\mathrm{new}}$ rather than recompute it at each time step.

After the collision pair is chosen, their postcollision velocities, $\mathbf{v}_i^*$ and $\mathbf{v}_j^*$, need to be evaluated. Conservation of linear momentum tells us that the center of mass velocity remains unchanged by the collision,

$$\mathbf{v}_{\mathrm{cm}} = \frac{1}{2}(\mathbf{v}_i + \mathbf{v}_j) = \frac{1}{2}(\mathbf{v}_i^* + \mathbf{v}_j^*) = \mathbf{v}_{\mathrm{cm}}^*$$

From conservation of energy, the magnitude of the relative velocity is also unchanged by the collision,

$$v_\mathrm{r} = |\mathbf{v}_i - \mathbf{v}_j| = |\mathbf{v}_i^* - \mathbf{v}_j^*| = v_\mathrm{r}^*$$

These equations give us four constraints for the six unknowns in $\mathbf{v}_i^*$ and $\mathbf{v}_j^*$.

The two remaining unknowns are fixed by the angles, θ and ϕ, for the relative velocity

$$\mathbf{v}_\mathrm{r}^* = v_\mathrm{r}[(\sin\theta\cos\phi)\hat{\mathbf{x}} + (\sin\theta\sin\phi)\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}]$$

For the hard-sphere model, these angles are uniformly distributed over the unit sphere. The azimuthal angle is uniformly distributed between 0 and 2π, so it is selected as $\phi = 2\pi\Re_1$. The θ angle is distributed according to the probability density,

$$P_\theta(\theta)\,d\theta = \tfrac{1}{2}\sin\theta\,d\theta$$

Using the change of variable $q = \sin\theta$, we have $P_q(q)\,dq = (\frac{1}{2})\,dq$, so q is uniformly distributed in the interval $[-1, 1]$. We don't really need to find θ; instead we compute

$$\begin{aligned} q &= 2\Re_2 - 1 \\ \cos\theta &= q \\ \sin\theta &= \sqrt{1-q^2} \end{aligned}$$

to use for computing the components of the relative velocity. The post-collision velocities are set as

$$\mathbf{v}_i^* = \mathbf{v}_\mathrm{cm}^* + \frac{1}{2}\mathbf{v}_\mathrm{r}^*$$
$$\mathbf{v}_j^* = \mathbf{v}_\mathrm{cm}^* - \frac{1}{2}\mathbf{v}_\mathrm{r}^*$$

and we go on to select the next collision pair.

Finally we ask, "How many total collisions should take place in a cell during a time step?" From the collision frequency, the total number of collisions in a cell during a time τ is

$$M_\mathrm{coll} = \frac{1}{2}(N_\mathrm{c}-1)N_\mathrm{ef}f\tau = \frac{N_\mathrm{c}(N_\mathrm{c}-1)N_\mathrm{ef}\pi d^2\langle v_\mathrm{r}\rangle\tau}{2V_\mathrm{c}}$$

where V_c is the volume of the cell. Each collision between simulation particles represents N_ef collisions among molecules in the physical system. However, we don't really want to compute $\langle v_\mathrm{r}\rangle$, since that involves doing a sum over all $\frac{1}{2}N_\mathrm{c}(N_\mathrm{c}-1)$ pairs of particles in the cell.

Recall that collision candidates go through an acceptance-rejection procedure. The ratio of total accepted to total candidates is

$$\frac{M_\mathrm{coll}}{M_\mathrm{cand}} = \frac{\langle v_\mathrm{r}\rangle}{v_\mathrm{r}^\mathrm{max}}$$

since the probability of accepting a pair is proportional to their relative velocity. Using the expressions above,

$$M_\mathrm{cand} = \frac{N_\mathrm{c}(N_\mathrm{c}-1)N_\mathrm{ef}\pi d^2 v_\mathrm{r}^\mathrm{max}\tau}{2V_\mathrm{c}}$$

which tells us how many candidates we should select over a time step τ. Notice that if we set $v_\mathrm{r}^\mathrm{max}$ too high, we still process the same number of collisions on average, but the program is inefficient because many candidates are rejected.

Figure 10.3 shows the results from a DSMC simulation of the relaxation of a gas towards equilibrium from an initial state in which all the particles have the same speed (but random velocity direction). We see that after 10 steps (and 2720 collisions) for a system of 3000 particles the distribution has already significantly relaxed toward equilibrium despite the extremely improbable initial condition and despite the fact that each particle has only been in fewer than two collisions. Figure 10.4 shows the distribution after 50 steps (and 14,555 collisions). This latter histogram shows that the system has almost completely relaxed to equilibrium in about a nanosecond.

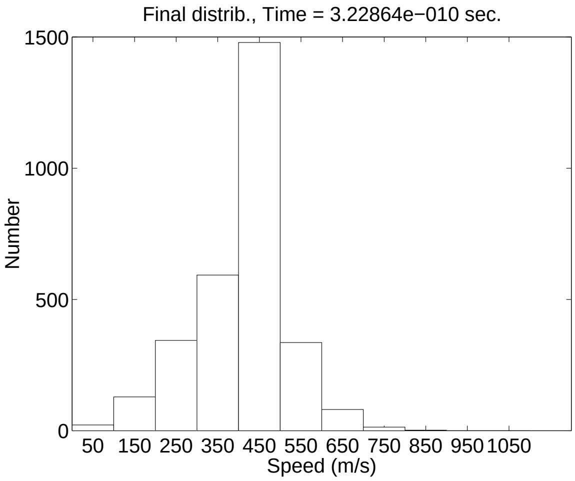

Figure 10.3: Speed distribution as obtained from `dsmceq` for $N = 3000$ particles. After 10 time steps, there have been 2720 collisions.

We can also use this simulation to compute

$$H(t) = \sum_{\text{bins}} \Delta v \frac{N_h(v)}{N} \ln \frac{N_h(v)}{N}$$

where $N_h(v)$ is the number of particles in a histogram bin of width Δv. This H-function was introduced by Boltzmann and it is proportional to the negative of the entropy in the system. Since the system is initially out of equilibrium, H decreases with time until the system equilibrates, as shown in Figure 10.5.

For a full description of the DSMC methods, including programming examples, see my textbook *Numerical Methods for Physics*. Makes for great summer reading at the beach!

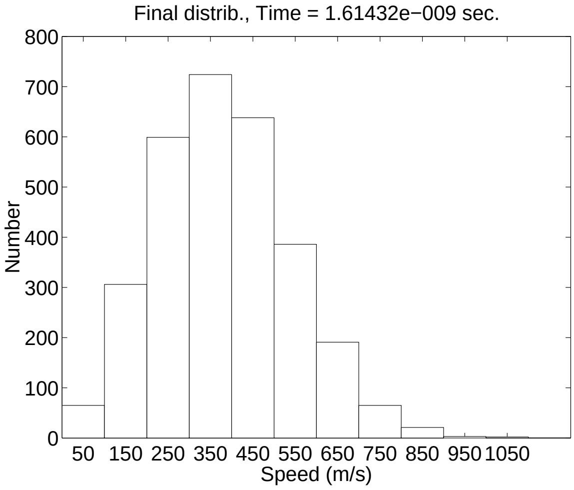

Figure 10.4: Speed distribution as obtained from `dsmceq` for $N = 3000$ particles. After 50 time steps, there have been 14,555 collisions.

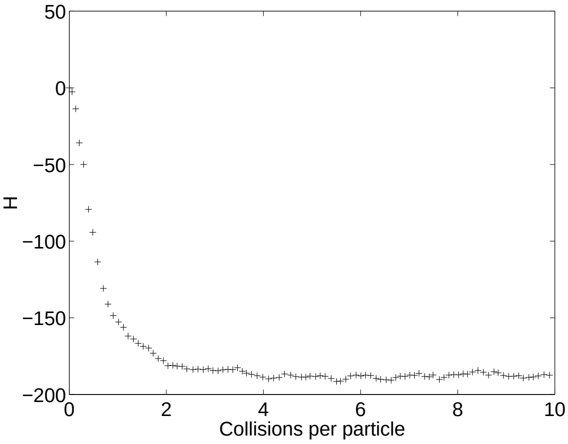

Figure 10.5: Boltzmann H-function measured in DSMC simulation.