The Department of Physics & Astronomy

The University of British Columbia

STATISTICAL MECHANICS

Physics 303 and 455 lectures by

Birger Bergersen

Department of Physics and Astronomy
University of British Columbia

Telephone: (604) 822-2754
birger@physics.ubc.ca

Revised Aug 2001

Contents

Preface

:
These notes contain material from lectures in two courses PHYS 303 and 455 at UBC. Sections 1-7 largely presents material given fall 1996. That year I went on to present the theory of Bose Einstein condensation and superfluidity following closely the discussion of a book that I co-authored with Michael Plischke at Simon Fraser University[1]. Since this material is in print I do not reproduce it here. Instead I have added a discussion of phase transitions from PHYS 455 some years earlier. Some of that material is also incorporated in the book mentioned above, but the emphasis is quite different. I have also added to section 2 some subsections on the Weibull distribution and on scale free distributions ("Pareto tail"). This material was taught in PHYS 455 but not in 303.

1 The historical origin of statistical mechanics

The foundation of statistical physics was laid towards the end of the nineteenth century by James Clerk Maxwell, Ludwig Boltzmann, Josiah Willard Gibbs[1], and largely completed by Albert Einstein in 1905.

Maxwell's kinetic theory of gases can be said to represent the starting point. We will get a flavor of his arguments from an outline of his derivation of what is now called the Maxwell velocity distribution.

Boltzmann made the argument more general and introduced the concept of ensembles. Instead of considering a single system he considered a large number of equivalent systems which had been prepared in the same way. He obtained probabilities for the possible states by calculating the relative frequency that a given state would occur in the ensemble, using the principle of a priori equal probabilities. Gibbs followed up by establishing the equivalence of statistical physics and thermodynamics. He did this by stressing an analogy with classical mechanics, which was the best understood branch of theoretical physics at the time. Finally Einstein rounded out the picture by his theory of fluctuations, diffusion and Brownian motion.

These developments happened before the advent of quantum mechanics. Einstein's theory of the photoelectric effect only appeared in 1905, and

[1]Readers who are interested in the history of science will enjoy Boltzmann's own description of his 'sabbatical' visit to California[2], and the article by Martin J Klein on the life and times of Gibbs[3].

a comprehensive theory of quantum mechanics only became available two decades later. However, statistical physics becomes simpler if one can appeal to some quantum concepts.

Our starting point is the idea that one can *count* the number of available states of a system. For us these will be discrete quantum states,although for a large system the states will be very closely spaced. In classical mechanics we describe a microscopic system by specifying the coordinates and momenta of the particles. The allowed value of these form a continuum. The procedure of counting requires, however discrete states. The modern way of getting around this difficulty is to consider classical mechanics as a limiting case of quantum mechanics, and we will take this approach, rather than following the historical route.

We begin by reviewing some concepts of probability theory.

2 Probability distributions

2.1 Probabilities and averages.

There are two main approaches to the problem of estimating how often possible outcomes of random events will occur. We may predict the frequency of allowed outcomes in repeated experiments from *a priori* (first principles) knowledge of the probabilities of contributing factors, making use of properties of permutations, combinations and binomial coefficients. Alternatively, we may be able to observe the possible outcomes as they occur, and we may wish to estimate *a posteriori* (after the fact) their probability from measured frequencies of occurrence.

In the latter case, one way to proceed is to plot the data in a bar chart (or histogram). As an example let us consider the distribution of *speeds* v of a system of N particles. We divide the range of v into intervals or *bins* of width Δ. The speed of a given particle lies in the $i-$th bin if its speed is between v_i and $v_i + \Delta$. We then count the number n_i in each bin and plot the result. If we only include a moderate number of particles, the resulting bar chart will typically have a ragged shape as in figure 1 a. On the other hand if we measure the speeds of a much larger number of particles it is likely to have a more regular shape as indicated in figure 1 b.

According to the *law of large numbers* the relative frequency after many measurements will almost always be close to a limiting value. This limiting value is the *probability* p_i, of any given particle being in the $i-$th bin

$$p_i = \lim_{N\to\infty} \frac{n_i}{N} \tag{1}$$

An important property of probabilities is the fact that if two possible events i and j are *exclusive* the probability that one or the other happens is

$$p(i.or.j) = p(i) + p(j) \tag{2}$$

Similarly if two possible events i and j are *independent* the probability that they both happen is

$$p(i.and.j) = p(i)p(j) \tag{3}$$

If the two events are not independent they are said to be *correlated*. The probability distribution for exclusive events satisfies the normalization condition

$$1 = \sum_i p_i \tag{4}$$

The histogram represents a *discrete* distribution. The particle speed can in general take on a continuous range of values. The *continuous probability distribution* obtains from the limit

$$p(v_i) = \lim_{\Delta\to 0} \frac{p_i}{\Delta} \tag{5}$$

The continuous probability distribution associated with the particle speed may then look like figure 4. The probability that the speed of a particle is between v_i and $v_i + dv$ is p$(v_i)dv$, and the continuous probability distribution satisfies the normalization condition

$$\int p(v)dv = 1 \tag{6}$$

Equation (6) is a statement of the fact that since all the particles must have some speed, the probabilities will add up to one. The *mean speed* is given by

$$\langle v \rangle = \int vp(v)dv \tag{7}$$

Let $f(v)$ be some property associated with the speed of a particle (e.g. the particle kinetic energy is given by $f(v) = \frac{1}{2}mv^2$, where m is the particle mass). The *average* value of $f(v)$ is then

$$\langle f(v) \rangle = \int f(v)p(v)dv \tag{8}$$

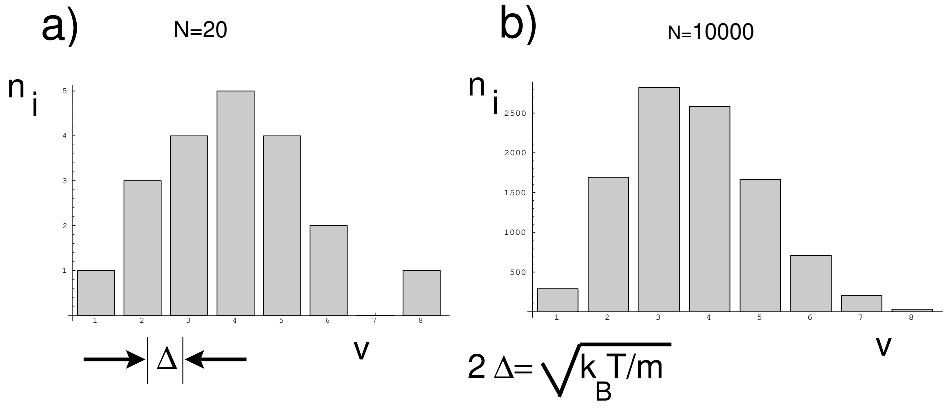


Figure 1: Speed distribution for (a) a large (b) a small number of particles.

In general, we will use angular brackets $\langle\ \rangle$ to describe averages. When applied to a probability distribution, the mean is the average outcome. Since we are dealing with probabilities, actual outcomes will typically deviate or *fluctuate* from the average. Let x be a *stochastic variable*, i.e. a quantity that can take on different values with some probability. A useful estimate of a typical deviation of the value of x from the mean is the *standard deviation* σ or root mean square fluctuation

$$\sigma = \sqrt{\langle (x - \langle x \rangle)^2 \rangle} = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \tag{9}$$

A closely related quantity is the *variance*: $var = \sigma^2$.

Sometimes it is useful to be able to *change variables.* Suppose a variable v has a probability distribution $p(v)$, i.e. the probability that v is between and $v+dv$ is $p(v)dv$. Let $g(v)$ be some function of v. What is the probability distribution $p'(g)$ for g?

The situation is simple if $g(v)$ has an inverse, i.e. only one value of v corresponds to each allowed value of g. Then, the probability that g is between g and $g + dv$ can be expressed as

$$p'(g)|dg| = p(v)|dv| \tag{10}$$

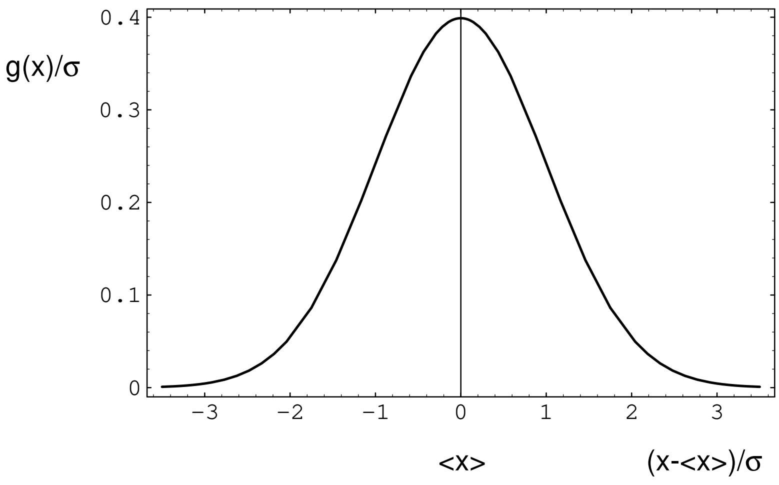


Figure 2: The normal distribution.

If more than one value of v corresponds to each value of g one must divide the range of v into intervals in which the derivative does not change sign. One can then add up contributions to the probability from each interval. *Question:* Why is it necessary to use the absolute value in 10?

2.2 The Gaussian distributions and the random walk.

A probability distribution which occurs in many context is the *Gaussian* or *normal distribution* which can be expressed in terms of the mean and variance as

$$g(x) = \frac{1}{\sqrt{2\pi(var)}} \exp\left[\frac{-(x - \langle x \rangle)^2}{2(var)}\right] \tag{11}$$

The normal distribution is plotted in figure 2. The importance of the this distribution is related to the *central limit theorem of statistics*:

> Suppose x_i is a random variable and X is the sum of N independent such variables all with the same mean $\langle x \rangle$ and variance $var = \langle x^2 \rangle - \langle x \rangle^2$ In the limit that N gets very large the probability distribution for X will be Gaussian with mean $N\langle x \rangle$ and variance $V = N\ var$.

Note that we are not assuming that the probability distribution for x is Gaussian, only that it has a finite mean and a variance. The above result

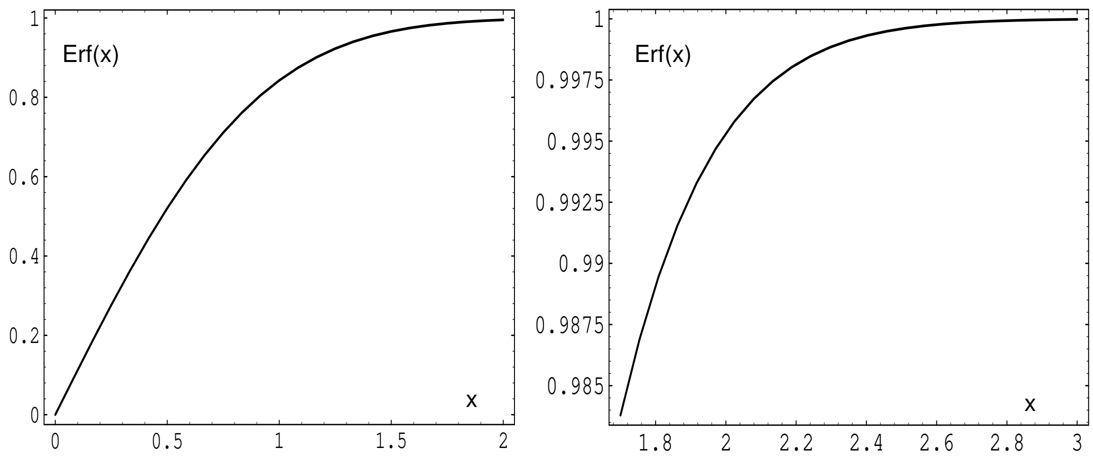


Figure 3: Error function

can be given a physical interpretation as a *random walk*. Consider a particle which is moving either to the left or right with speed v. From time to time the particle undergoes a collision, after which the particle with equal probability either changes direction or moves on. Let the particle collide at a constant rate. We first need to work out the survival probability $p(r)$ that a particle can travel a distance r without undergoing any collisions. The probability that a collision occurs between r and $r + dr$ is proportional to dr. Let the proportionality constant be $1/l$, where l is the *mean free path*. The probability $p(r + dr)$ that no collisions occur in an interval of length $r + dr$ is thus $p + dp = p(r + dr) = p(r)(1 - \frac{dr}{l})$ or

$$dp = -\frac{pdr}{l} \tag{12}$$

Since $p(0) = 1$ we find that $p(r)$ satisfies the *Poisson distribution*

$$p(r) = exp(-r/l) \tag{13}$$

The mean time between collisions is

$$\tau = \int_0^1 \frac{rdp}{v} = -\int\limits_0^\infty r\frac{dp}{vdr}dr = \frac{1}{lv}\int\limits_0^\infty r\exp(-r/l)dr = \frac{l}{v}$$

Since the particle travels with equal probability in either direction the mean distance traveled is zero, while the variance is

$$\langle r^2 \rangle - \langle r \rangle^2 = \int_0^1 r^2 dp = -\int\limits_0^\infty r^2 \frac{dp}{dr} dr = \frac{1}{l} \int\limits_0^\infty r^2 \exp(-r/l) dr = 2l^2$$

After some time t (assumed to be long compared to the mean time τ between collisions) the particle will have undergone

$$N \approx \frac{t}{\tau} = \frac{tv}{l}$$

collisions. From the central limit theorem the mean total distance traveled will be zero while the variance is $2l^2 N = 2lvt$. The probability distribution of distances traveled is

$$p(x,t) = \frac{1}{\sqrt{4\pi lvt}} \exp(\frac{-x^2}{4lvt}) \tag{14}$$

The probability distribution for the distance traveled after one or a few collisions will not be Gaussian, but the central limit theorem becomes approximately valid after a few collisions. Note that holds after many collisions When dealing with Gaussian distributions the following integrals are useful

$$\int\limits_{-\infty}^{\infty} \exp(-ax^2) dx = 2 \int\limits_0^\infty \exp(-ax^2) dx = \sqrt{\frac{\pi}{a}} \tag{15}$$

$$\int\limits_{-\infty}^{\infty} x^2 \exp(-ax^2) dx = 2 \int\limits_0^\infty x^2 \exp(-ax^2) dx = \frac{1}{2a} \sqrt{\frac{\pi}{a}} \tag{16}$$

$$\int\limits_{-\infty}^{\infty} x \exp(-ax^2) dx = 0 \tag{17}$$

$$\int\limits_0^\infty x \exp(-ax^2) dx = \frac{1}{2a} \tag{18}$$

The *error function* is defined as

$$\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}} \int\limits_{-x}^{x} \exp(-s^2) ds \tag{19}$$

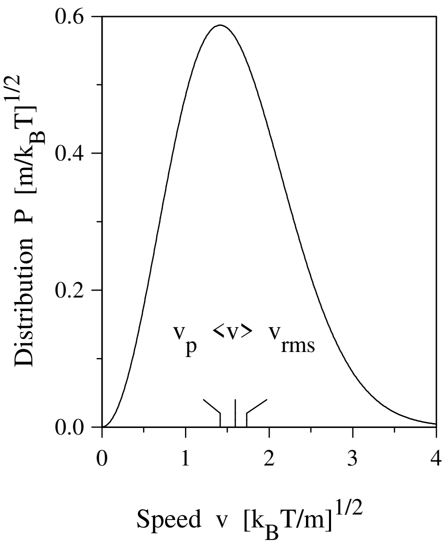


Figure 4: Maxwell speed distribution

and has the properties

$$\operatorname{erf}(0) = 0; \quad \operatorname{erf}(1) = 1; \quad \operatorname{erfc}(x) \equiv 1 - \operatorname{erf}(x) \tag{20}$$

where $\operatorname{erfc}(x)$ is called the complementary error function. Tables of the error function are readily available. Plots of the error function is given in figure 3.

2.3 Maxwell speed distribution

Consider a gas at a given temperature T. The molecules of the gas will move about with different speeds in random directions. The likelihood that a molecule has a given speed can be represented by a distribution function $p(v)$ according to which the probability that a given molecule has speed between v and $v + dv$ is $p(v)dv$. The Maxwell speed distribution is

$$p(v) = v^2 \sqrt{\frac{2}{\pi}} \left(\frac{m}{k_B T}\right)^{\frac{3}{2}} \exp\left(-\frac{mv^2}{2k_B T}\right), \tag{21}$$

where m is the mass of a molecule and $k_B = 1.38041 \times 10^{-23}$ $Joules/Kelvin$ is the Boltzmann constant. The distribution is normalized, i.e.

$$\int_0^\infty p(v)dv = 1. \tag{22}$$

The mean speed is

$$\langle v \rangle = \int_0^\infty vp(v)dv = \sqrt{\frac{8k_BT}{\pi m}} \tag{23}$$

The most probable speed v_p of the distribution is the position of the peak (see figure 4) and can be obtained by differentiating $p(v)$ and putting the result equal to zero. One finds

$$v_p = \sqrt{\frac{2k_BT}{m}}. \tag{24}$$

The mean square speed is

$$\langle v^2 \rangle = \int_0^\infty v^2 p(v)dv = \frac{3k_BT}{m} \tag{25}$$

The mean speed and the root mean square (rms) speed are often used in kinetic theory to describe a typical speed of a molecule.

To derive the result in Eq.(21) we must assume that the gas molecules are in thermal equilibrium and that they obey classical (Newtonian) mechanics. In particular, the only velocity dependence of the energy of a molecule is the kinetic energy $mv^2/2$.

Maxwell's original derivation of the distribution was published in 1860 and was based on the notions of isotropy of space and the assumption that the probability distributions for velocity components in orthogonal directions should be uncorrelated (i.e. independent of each other). Let $f(v_x)dv_x$ be the probability that the $x-$component of the velocity is between v_x and $v_x + dv_x$ and let $f(v_y)$ and $f(v_z)$ be the corresponding distributions for the $y-$ and $z-$components. The probability that a particle has velocity components between v_x and $v_x + dv_x$ and also between v_y and $v_y + dv_y$ and v_z and $v_z + dv_z$ will then be

$$f(v_x)f(v_y)f(v_z)dv_xdv_ydv_z \tag{26}$$

The orientation of the coordinate axes is arbitrary, and therefore the product $f(v_x)f(v_y)f(v_z)$ must depend only on the speed v which can be obtained from $v^2 = v_x^2 + v_y^2 + v_z^2$. We therefore can define a function ϕ by

$$f(v_x)f(v_y)f(v_z) = \phi(v_x^2 + v_y^2 + v_z^2) \tag{27}$$

Since f and ϕ are probability distributions they must be positive, and can be written as the exponential of a function. Since directions do not matter $f(v_x) = f(-v_x)$. We must thus be able to express f in terms of the square of the velocity component. Therefore

$$f(v_x) = e^{h(v_x^2)}; \quad \phi(v_x^2 + v_y^2 + v_z^2) = e^{\psi(v_x^2+v_y^2+v_z^2)} \tag{28}$$

or

$$h(v_x^2) + h(v_y^2) + h(v_z^2) = \psi(v_x^2 + v_y^2 + v_z^2) \tag{29}$$

implying that ψ must be a linear function. We also require that the probability becomes small for large speeds and find that the distribution must be on the form

$$f(v_x) = Ce^{-v_x^2/a}; \quad \phi(v^2) = C^3e^{-v^2/a} \tag{30}$$

From the normalization condition

$$1 = \int_{-\infty}^{\infty} dv_x f(v_x) = C\sqrt{\pi a} \tag{31}$$

$C = 1/\sqrt{\pi a}$. We can also calculate the mean kinetic energy by evaluating

$$\frac{m\langle v^2\rangle}{2} = \frac{3m}{2}\int_{-\infty}^{\infty} dv_x v_x^2 f(v_x) = \frac{3ma}{4} \tag{32}$$

The constant a is therefore closely related to the mean kinetic energy of the gas. If we require that $\frac{m\langle v^2\rangle}{2} = \frac{3k_BT}{2}$ we find $a = \frac{2k_BT}{m}$ To calculate the *speed* distribution from the distribution for the velocity components we go to polar coordinates $v_z = v\cos\theta$, $v_y = v\sin\theta\cos\phi$, $v_x = v\sin\theta\sin\phi$

$$dv_x dv_y dv_z = v^2 dv \sin\theta d\theta d\phi \tag{33}$$

The speed distribution can be obtained by integrating over the angles

$$p(v) = v^2\int_0^{\pi}\sin\theta d\theta\int_0^{2\pi} d\phi C^3e^{-v^2/a} = 4\pi v^2C^3e^{-v^2/a} \tag{34}$$

Substituting the values for C and a we finally obtain (21)

The result turns out to be valid even if the gas is not ideal i.e. velocity independent forces between the particles will not change the result. On the other hand, the speed distribution will need to be modified if the temperature is so high that speeds become relativistic. We will also find that there will be significant corrections to the Maxwell speed distribution at low temperatures when quantum effects are important, and one must correct for the Fermi-Dirac or Bose-Einstein characters of the particles in a gas. Finally the Maxwell speed distribution only holds if the gas is in thermal equilibrium The ionized gas in the outer atmosphere of the sun, or the ionosphere outside the earth is typically produced in a non equilibrium setting. Major deviations from the Maxwell velocity distribution are known to occur in such situations. It is far from obvious from the above derivation why these restrictions apply. Later on in this course we will find a more "modern" way of deriving the distribution which makes the limitations clear. This method involves the concept of the Boltzmann factor and the canonical ensemble.

2.4 Outcome of coin tosses

In this section we will attempt to make the central limit theorem plausible by demonstrating that it works in a particular case. We will also find an example system which does not satisfy the conditions set by the theorem.

A coin is tossed N times. We associate the value $x = +1$ for "heads" and $x = 0$ for "tails".The average value $\langle x \rangle$ of x is then $\frac{1}{2}$ if the coin is unbiased. The average value of x^2 is $\langle x^2 \rangle = \frac{1}{2}$. The variance associated with a single toss is thus $var = \langle x^2 \rangle - \langle x \rangle^2 = \frac{1}{4}$. After N tosses the number of times heads has come up will be X. By the central limit theorem the mean value of X will be $\langle X \rangle = N/2$, the variance will be $V = N/4$, and the probability distribution in the limit of large N will be

$$p(X) = \sqrt{\frac{2}{\pi N}} \exp\left[\frac{-(2X - N)^2}{2N}\right] \tag{35}$$

Since we did not derive the central limit theorem it is useful to derive (35) independently. An exact formula for the probability distribution of the outcome of N coin tosses is obtained from the *binomial distribution.* There are two possible outcomes of each toss - heads or tails. After N tosses there are thus 2^N possible outcomes. Of these

$$\frac{N!}{X!(N-X)!} \tag{36}$$

correspond to the outcome X. The probability distribution for X is thus in general

$$p(X) = \frac{N!}{2^N X!(N-X)!} \tag{37}$$

To see that (35) is a good approximation to (37) we define

$$y = X - \frac{N}{2} \tag{38}$$

as the deviation from the mean. We expect that $y << N$ "almost always" when N is large. We use the *Stirling approximation*[2]

$$\ln(z!) = \frac{1}{2}\ln(2\pi) + (z + \frac{1}{2})\ln z - z + \textit{terms small compared to unity} \tag{39}$$

as $z \to \infty$. After substitution of (38) and (39)into (37), taking the logarithm of both sides we find after some algebra

$$\ln p \approx \frac{1}{2}\ln\frac{2}{N\pi} - \frac{N+1}{2}\ln(1 - \frac{4y^2}{N^2}) - y\ln\frac{1+\frac{2y}{N}}{1-\frac{2y}{N}}$$

Assuming that $y/N << 1$ we can expand the logarithm $\ln(1+x) \approx 1+x$ and we obtain again after some algebra

$$p \approx \sqrt{\frac{2}{\pi N}}\exp(\frac{-2y^2}{N}) \tag{40}$$

Finally, substitution of X from (38) into (40) gives us the desired result (35).

Not all distribution have a finite mean or a variance. A famous counter-example, discussed already in the early days of probability theory (≈ 1730), is Daniel Bernoulli's *St. Petersburg paradox.* Imagine that you are allowed to play the following game against "the bank". The banker tosses a "fair" coin. If the outcome is "tails" you are given a ducat (or whatever they used for money in St. Petersburg in the 18'th century). If the outcome is "heads" the coin is tossed again until the outcome is "tails". If tails come up in the second trial you get 2 ducats. If tails only comes up after n trials you get 2^{n-1} ducats.

[2]It is frequently adequate to use an approximation for $\ln z!$ where the error is small compared to z. The approximation $\ln z! \approx z\ln z - z$ is then sufficient

The probability that tails comes up for the first time after n trials is the probability 2^{-n+1} that the first $n-1$ trials yielded "heads" times the probability $\frac{1}{2}$ that the next trial gives heads or 2^{-n}. The mean profit is then

$$\langle \$ \rangle = \frac{1}{2} + \frac{1}{4}2 + \ldots \frac{1}{2^n}2^{n-1} + \ldots = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \ldots = \infty \tag{41}$$

Thus, although the probability distribution is well defined with a finite value for each possible outcome, the mean is infinite! How much would you be willing to pay to be allowed to take part in this game?

2.5 School of hard knocks: lognormal distributions

The normal distribution results when the outcome of an event is the result of a large number of additive random factors. Sometimes situations arise in which an outcome depends on the successful completion of a number of subtasks. The probability of success is then the product of the probabilities for each subtask. This may give rise to a lognormal distribution - one in which the logarithm of the outcome follows a normal distribution. This distribution was first introduced by Francis Galton in 1879 but it is only relatively recently that is been realized that this distribution is common. According to the lognormal distribution the probability that a random variable has a probability between x and $x + dx$ is

$$p(x)dx = \frac{\exp[-(\ln \frac{x}{a})^2/2b]}{\sqrt{2\pi b}} \frac{dx}{x} \tag{42}$$

where a and b are the parameters of the distribution (the mean and the variance of the logarithm of the random variable). We obtain (42) by a change of variable, assuming that the logarithm of x has a Gaussian distribution with mean $\ln a$ and variance b.

As an example let us consider a model due to the Russian mathematician and physicist A.N. Kolmogorov for the size distribution of particles in crushed rock. Let x be a measure of the size of a rock (e.g. the volume). We start out with the distribution $f_0(x)$ of rocks of different sizes. We then let a crusher come down n times after which the probability distribution becomes $f_n(x)$. He constructed the following simple model to describe the situation:

Let us follow the history of a rock of size n through its parentage. The size of a fragment x_n will in general be some fraction of the parent rock

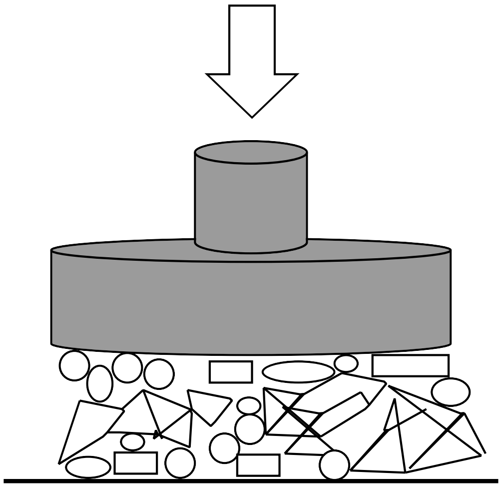


Figure 5: Crushing rocks

x_{n-1}.

$$x_n = R_n x_{n-1} = \prod_{i=1}^{n} R_i x_0 \tag{43}$$

giving

$$\sum_{n=1}^{N} \ln R_n = \ln \frac{x_N}{x_0} \tag{44}$$

We now assume that $\ln R_n$ is a random variable with a mean and a variance. As N becomes large the sum over $\ln R$ will approach a Gaussian from the central limit theorem. We see that the logarithm of the particle size will follow a normal distribution and the size distribution will be lognormal. Another examples of a distribution which is approximately lognormal is the distribution of family incomes for most members of society (all but the very rich).Many environmental distribution (e.g. concentrations of pollutants in a series of samples) also often follow an approximately lognormal distribution.

2.6 Weibull distribution

There are a number of situations, e.g. in problems involving *fracture* or *failure* where what matters is not the *average* property of the system, but rather the behavior of the *weakest link*. A very practical approach to such problems was taken by the Swedish engineer Waloddi Weibull [4].

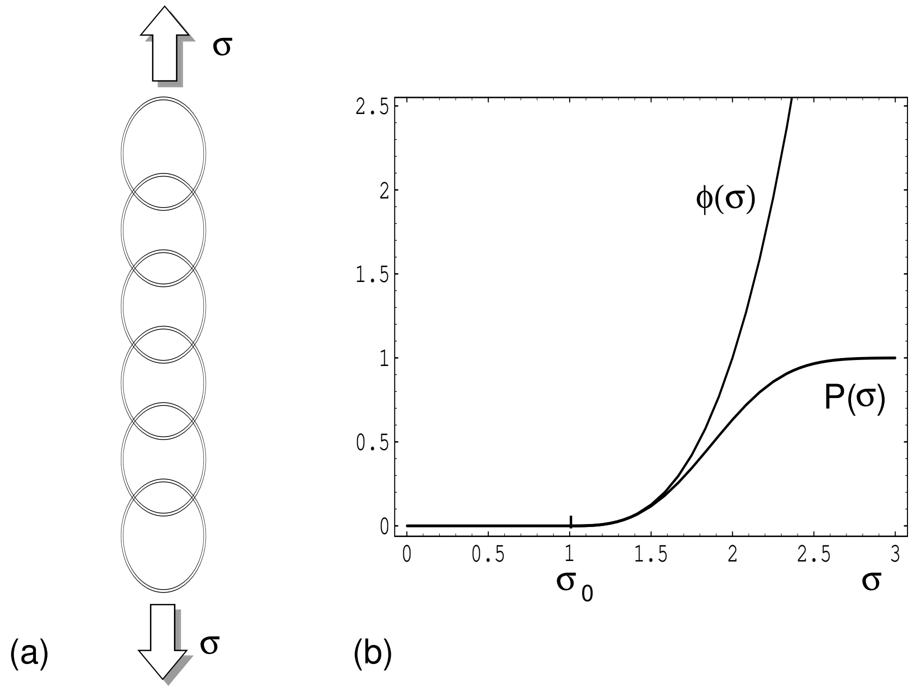


Figure 6: (a) A chain with $n = 5$ links subject to tensile stress σ. (b) Plot of $P(\sigma)$ and $\phi(\sigma)$ if $\sigma_0 = 1$ and $\phi = \sigma^3$ in appropriate units.

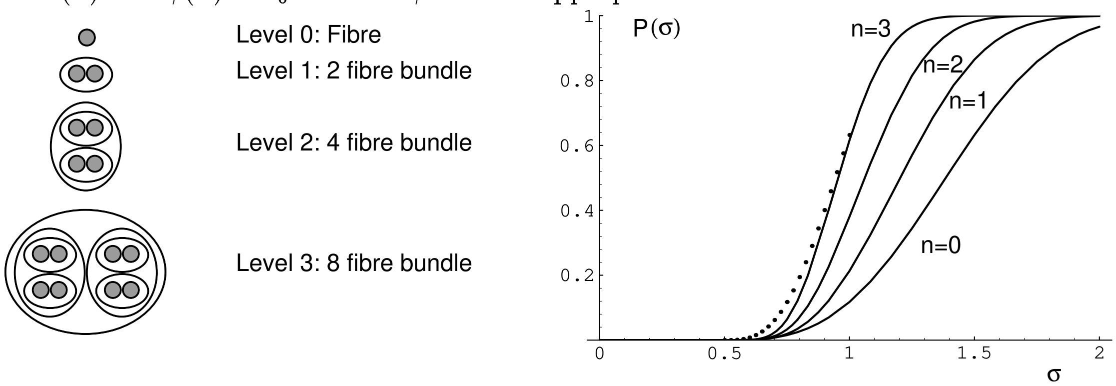


Figure 7: (a) A rope is made by bundling pairs of fibers and then bundling pairs of bundles. (b) Plot of $P(\sigma)$ and $\phi(\sigma)$ if $\sigma_0 = 0.5$ and $\phi = \sigma^3$ in appropriate units for the $n = 0, 1, 2, 3$. The dotted curve represents the approximation (49)

Consider a chain made up of n links (see figure 6(a)). Each link i has a yield threshold stress σ_i which may depend on microscopic flaws and will vary from link to link. We assume the probability that a link breaks with a stress σ is given by

$$P(\sigma) = 1 - e^{-\phi(\sigma)}$$

The probability that the link will *not* break is then $e^{-\phi(\sigma)}$, where $\phi(\sigma)$ is some monotonically increasing function of σ. If any one of the links fails the whole chain fails. Therefore, the probability that the whole chain survives is $e^{-n\phi(\sigma)}$ and the probability that the *chain* will break under stress σ is

$$P_{chain} = 1 - e^{-n\phi(\sigma)} \tag{45}$$

Weibull suggested the form

$$\phi(\sigma) \propto (\sigma - \sigma_0)^\rho; \quad for\ \sigma > \sigma_0 \tag{46}$$

where σ_0 is some minimum failure stress (which may be zero). The exponent ρ must be expected to be a property of the link material (Weibull found $\rho \approx 3$ for Bofors steel, and $\rho \approx 1.46$ for Indian cotton.). The proportionality constant in (46) will from (45) not be a material property, but will depend on the length of the chain and the size of the links.

Consider next a *rope* which is made by bundling fibers together (Figure 7(a), see also [5]) Let the *stress* σ be the average tensile force per fiber. $P(\sigma)$ will, as before, represent the cumulative probability of breakage under stress. Each fiber is paired with another fiber to form a bundle, with each bundle paired again etc. On the n'th level there are $m = 2^n$ fibers.

If a fiber breaks the stress is transferred to the other members of the pair. A bundle fails if (i) either both fibers fail under the original stress, or (ii) only one fails initially, but the redistributed stress is too much for the other member of the pair. The probability (i) is $P_n(\sigma)^2$, while the probability for (ii) is $2P_n(\sigma)[P_n(2\sigma) - P_n(\sigma)]$. The cumulative failure probability on level $n+1$

is then

$$P_{n+1} = P_n(\sigma)[2P_n(2\sigma) - P_n(\sigma)] \tag{47}$$

As before we assume a Weibull distribution for the failure of a single fiber

$$P_0(\sigma) = 1 - \exp[-a(\sigma - \sigma_0)^\rho]; \quad \sigma > \sigma_0 \tag{48}$$

The typical behavior of P_n for n=0,1,2,3 is illustrated in figure 7(b). We can find an upper bound for the failure probability by putting $P(2\sigma) = 1$ in

(47). We then solve for P_n and find for $\sigma > \sigma_0$

$$P_n(\sigma) = 1 - \exp[-am(\sigma - \sigma_0)^\rho] \tag{49}$$

i.e. we find that the Weibull form is retained for the rope, with an unchanged exponent ρ, but the constant a is proportional to the number of fibers. Of course, in practice it may not be realistic to assume that *all* the load is redistributed locally to the other member of the pair when failure occurs. Nevertheless, the above discussion suffices to show that material properties such as yield stress cannot normally be treated as an intensive thermodynamic variable analogous to bulk the modulus or the elastic constants.

2.7 Pareto tail

The income distribution in most societies is approximately lognormal for most people who do have an income, up to approximately the 97th percentile. However, the lognormal distributions does not to hold for the 2-3% who are extremely rich.

The Italian sociologist and economist Vilfredo Pareto tried, during the last part of the 19th and the early part of the 20th century, to make economics and sociology into an exact science by pursuing analogies with physics and mechanics. He was particularly interested into the dynamics of business cycles and the rise and fall of empires and elites, and his work remains somewhat controversial. He is possibly best known for collecting statistics on the income of individuals at various times and places. This convinced him in 1897 that

> In all places and at all times the distribution of income in a stable economy, when the origin of measurement is at a sufficiently high income level, will be given approximately by the empirical formula y= ax^{-v}, where y is the number of people having an income x or greater and v is approximately 1.5.

Income tax data in several countries are compatible with Pareto's observation also in more recent times. Indeed, numerous such power law distributions have been studied and popularized by Mandelbrot [6], and the reader is encouraged to peruse his fascinating book. Power law distributions are ubiquitous in critical phenomena associated with phase transitions, as we shall see later. Recently much interest have focused on 'self-organized critically' which occur in driven non equilibrium systems and give rise to Pareto like distributions for the size of avalanches, epidemics, energy dissipation in earthquakes, size of forest fires etc. Returning to the income distribution, there are probably several mechanisms responsible, and the distribution of very high income may be more complex than a single exponent. These mechanisms have in common that some form of *amplification* is involved. Montroll and Shlesinger [7] [8] builds a model to explain the data from the observation that

> The leverage people in the investment business have their style of amplification. During certain periods of prosperity easy money become available for investment, sometimes in stock, sometimes in real estate or perhaps in silver or Rembrandts. A common feature of such times is that the daring may exploit the easy money to acquire some speculative commodity through margin payment, say, 10% with a promise to pay the remainder. If the commodity doubles in price a 10% margin is amplified into a ninefold profit.

Following Mandelbrot [6] we will pursue a different mechanism originally due to Lydall [9]. Suppose the employees within an enterprise are arranged in the form of a pyramid, or hierarchy, with a director at the top and numerous levels of supervisory personnel below. Let y_i be the number of people at the i'th level and let $i+1$ be the level above. We assume that the ratio of personnel at the two levels are $n = y_i/y_{i+1}$. Suppose each operator on the i'th level earn their income x_i from a commission of a fraction λ of the income of the people in the level below in return a fraction λ of the income is paid to the immediate boss above. The income at the i+1 level is thus $x_{i+1} = n(1-\lambda)\lambda x_i$. If $p[x]$ is the probability distribution for income x we find

$$np[n(1-\lambda)\lambda x] = p[x] \tag{50}$$

Taking the logarithm on both sides we find that (50) leads to a power law distribution of the Pareto form with

$$v + 1 = \frac{\ln n}{\ln[n\lambda(1-\lambda)]}$$

In order for amplification to take place we must impose the restriction $n\lambda(1-\lambda)\rangle 1$. Note that v is the exponent of the cumulative distribution i.e. $n \sim x^{-v-1}$. It is worth noting that the exponent v is *not* universal, but depends on the parameters n and λ which must be expected to vary from society to society. This suggests that the mechanism suggested above is not quite satisfactory. An introduction to Pareto's own thoughts on the subject can be found in [10] and [11].

2.8 Problem set 1

1:

- a: Construct a histogram for the probabilities of the outcome of 5 tosses of a fair coin (5 heads, four heads one tail etc.). Are there any significant difference between results obtained using the binomial distribution and the Gaussian approximation?

- b: Estimate the probability of 1050 heads in 2000 tosses of a fair coin using the Gaussian approximation.

2:

- a: Estimate the parameters a and b of (42) to the Kolmogorov model of crushed rock. Assume that you are starting from a single rock of size 1, use the lognormal approximation, assume that R_n has a uniform distribution between 0 and 1, and that you have $N = 12$.

- b: Plot the lognormal distribution found under **a**.

- c: (Optional). You may wish to verify the theoretical result under **a:**and **b:** by making a computer simulation.

3 The information theoretic approach

3.1 Micro- and macro-states

Our description of the world takes place on many scales, for example:

?
galaxy cluster
galaxy
solar system
sun, planets and moons
continents and oceans
countries, mountain ranges
provinces and towns
things, people, animals and trees
matter
atoms and ions
elementary particles
quarks
?

At each level we avoid describing in detail what goes on the level below, replacing a detailed description by a statistical treatment: 'Things average out'. Similarly - what goes on in the next level up is often not considered: 'One doesn't see the forest for the trees'.

Conventional statistical mechanics is mainly concerned with the interface between 'matter' and 'atom and ions' levels.

Have: 'particles' with forces between them.
Want: macroscopic behavior, equations of state.

The terms used to describe systems at various levels in the hierarchy are quite different. A possible state of the system at the lower level is called a *micro-state*. At the higher level much detail is lost and many micro-state appears indistinguishable. The states at the higher level are called *macro-states*. Each macro-state can be realized by many micro-states.

> *Assumption:* It is possible to describe a large system by a small number of macro *state variables*. Some of these are independent variables, or *control parameters* others are dependent and given by an *equation of state*. The dependent variables will in general fluctuate about the mean value given by the equation of state, but for a large enough system these fluctuations are small.

In our coin toss example one dependent state variable would be X, the number of heads after N tosses (N being a control, or independent variable).

Another possible control variable would be the probability P that head will come up in a single toss ($P = 1/2$ in our example. The equation of state would then be $X = NP$. The microstate would be a description of the outcome of each toss. The number of microstates associated with a given macrostate will then be

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

One of the goals of statistical mechanics is to derive equations of state from a description of the microscopic dynamics. A partial list of state variables for a thermodynamic system is given in the table below. The most common variables are listed in the columns to the left. Special systems such as dielectrics, magnetic systems or liquid interfaces require further state variables as indicated in the columns to the right.

P	pressure	$\mathcal{P}$	polarization
V	volume	$\mathcal{E}$	electric field
μ	chemical potential	$\mathcal{M}$	magnetization
N	# of particles	$\mathcal{B}$	magnetic field
T	temperature	σ	surface tension
S	entropy	$\mathcal{A}$	surface area
U, E	internal energy		

The following basic assumption of thermodynamics is sometimes referred to as that of 'existence of matter'

> State variables can be separated into *extensive* variables which are proportional to the size of the system, and *intensive* variables which are independent of the size of the system.

We will come back to this assumption later and also note some important exceptions. Of the variables listed in the table above P, μ, T, $\mathcal{E}$, $\mathcal{B}$, σ, are intensive while V, N, S, U, $\mathcal{P}$, $\mathcal{M}$, $\mathcal{A}$ are extensive. In our coin toss example X and N are extensive P is intensive.

In the case of the Kolmogorov model of crushed rock it is the logarithm of the size, rather than the size itself, that is used. A thermodynamic description in terms of extensive and intensive variables therefore doesn't seem useful. Similar considerations apply to the Weibull distribution. Since the rupture stress of an object depends on its size in a non-trivial it is not a proper thermodynamic variable. Income and wealth as discussed earlier does not fit either into the thermodynamic picture of extensive and intensive variables. There is thus something very special about "matter".

3.2 A measure of uncertainty

Consider a system where there are a number of possible states that we label $i = 1, 2, 3....r$, with r the total number of possible states. Assume that we have only partial knowledge as to the state in which the system actually is. We represent this state of affairs by saying that there is a probability P_i that the system is in state i. In what follows it turns out to be extremely useful to find a measure $H(\{P_i\})$ of the degree of uncertainty associated with the probability distribution. In order to proceed, let us appeal to the *a posteriori* definition of probability. Let us imagine that we prepare the system N times and measure its state each time. Since this is theory, we can assume that N is very large. By the law of large numbers the state 1 will come up approximately NP_1 times, state 2 will come up approximately NP_2 times etc. A possible measure of the uncertainty would then be the number of ways these outcomes can be reordered

$$\Omega_N = \frac{N!}{\prod_{i=1}^{r}(NP_j!)} \tag{51}$$

Indeed, the more possible outcomes there are of our experiment, the more uncertain is the result! This is not the only possible measure of uncertainty, any function $H(\Omega)$ of the number of sequences Ω that leads to the same result will serve, provided that the function satisfies the condition that if $\Omega_1 > \Omega_2$ then $H(\Omega_1) > H(\Omega_2)$ (i.e. $H(\Omega)$ must be a monotonic function of Ω). Consider next a system made up of two independent subsystems. We imagine that we measure the state of subsystem 1 N_1 times and subsystem 2 N_2 times. The number of possible sequences is now $\Omega = \Omega_1\Omega_2$. We next limit the choice of possible functions H by requiring our uncertainty to be the sum of the un certainties associated with the two subsystem, i.e. we require that H is an *extensive variable*:

$$H(\Omega_1\Omega_2) = H(\Omega_1) + H(\Omega_2) \tag{52}$$

The choice of the function H is now essentially unique. To see this let us differentiate (52) two different ways

$$\Omega_2 \frac{d\,H(\Omega_1\Omega_2)}{d\,(\Omega_1\Omega_2)} = \frac{d\,H(\Omega_1)}{d\,\Omega_1} \tag{53}$$

$$\Omega_1 \frac{d\,H(\Omega_1\Omega_2)}{d\,(\Omega_1\Omega_2)} = \frac{d\,H(\Omega_2)}{d\,\Omega_2} \tag{54}$$

We multiply (53) by Ω_2 and (54) by Ω_1 and subtract the two equations. The result is

$$\Omega_1 \frac{d\,H(\Omega_1)}{d\,\Omega_1} = \Omega_1 \frac{d\,H(\Omega_2)}{d\,\Omega_2} \tag{55}$$

Since the left hand side of (55) is independent of Ω_2 and the right hand side is independent of Ω_1 they must both be equal to a constant. We call this constant c_0. We now drop the subscripts and find

$$\Omega \frac{d\,H(\Omega)}{d\,\Omega} = c_0$$

We integrate this equation to obtain

$$H(\Omega) = c_0 \ln \Omega + c_1$$

where c_1 is another constant. If there is no uncertainty, i.e. if we know the outcome then $\Omega = 1$ and we require that $H(1) = 0$. This gives $c_1 = 0$ and we are left with

$$H(\Omega) = c_0 \ln \Omega$$

The choice of the constant c_0 is arbitrary at this stage, the constant essentially determines the "unit of uncertainty". We will for the time being choose the constant to be $c_0 = 1$. Later on we will identify H with the entropy and choose c_0 to be equal to the Boltzmann constant k_B.

Going back to (51) we find that the uncertainty associated with N repeated experiments is

$$H_N = \ln \Omega_N$$

Since H is extensive we now define the "uncertainty" associated with the probability distribution itself to be $1/N$ times the uncertainty associated with N repetitions of the experiment. This gives for the "uncertainty" of the probability distribution

$$H = \lim_{N \to \infty} \frac{1}{N} \ln \frac{N!}{\prod_{i=1}^{r} (NP_i)!}$$

Since we only need to evaluate the logarithm of the factorials to accuracy N we can use the simple version of Stirling's formula

$$\ln N! \approx N \ln N - N$$

and we find after some algebra

$$H = -\sum_{i=1}^{r} P_i \ln P_i \tag{56}$$

Equation (56) was derived by Shannon[12] in shortly after the second world war and is generally considered to be one of the fundamental results of information theory

3.3 The principle of least bias

We have shown how to calculate the uncertainty associated with a given probability distribution. Now, we turn this question around and address how to find the most likely probability distribution given the information we have about the system. We do this by applying *the principle of least bias* which states that the most reasonable distribution is the one which maximizes the uncertainty about the outcome subject to what we know. If we use a distribution with less uncertainty we are biasing the probability by providing information which is not in evidence. As our first example consider the case of a single coin toss. Let P_h be the probability that the outcome is "heads" and let $P_t = 1 - P_h$ be the probability for "tails". The "uncertainty" about the outcome is

$$H = -P_h \ln P_h - (1 - P_h) \ln(1 - P_h)$$

by differentiating the above expression we find that the probability which produces the least bias satisfies

$$0 = \frac{d\ H}{d\ P_h} = -\ln \frac{P_h}{1 - P_h}$$

or $P_h = 1/2$, ie. the probabilities of the two outcomes are equal. Any other value of P_h would mean that the coin is biased.

Next consider the case of r possible outcome. Let P_i be the probability of the i'th outcome. The uncertainty is now

$$H = -\sum_{i=1}^{r} P_i \ln P_i \tag{57}$$

The different probabilities P_i can not take on completely arbitrary values since they are subject to the restriction

$$\sum_{i=1}^{r} P_i = 1 \tag{58}$$

We can calculate the maximum of (57) subject to the normalization restriction (58) by using the method of *Lagrange multipliers* according to which one first finds the maximum of

$$-\sum_{i=1}^{r} P_i \ln P_i + \lambda P_i \tag{59}$$

for arbitrary values of the Lagrange multiplier λ and the adjusts λ so that the restraint (58) is satisfied. Putting the derivative of (59) with respect to any of the P_is equal to zero yields

$$-\ln P_i - 1 + \lambda = 0$$

with solution

$$P_i = e^{\lambda - 1}$$

The normalization condition (58) then gives

$$e^{\lambda-1} = \frac{1}{r}$$

and we finally get

$$P_i = \frac{1}{r}$$

Put in words: If we do not have any reason to believe that any one outcome is more likely than any other, we must, according to the principle of least bias, assume that they are equally probable.

Consider next the case of a loaded die. When we throw a die any one of the numbers 1,2..6 my come up. Let P_i bet the probability that the number i comes up and let

$$a = \sum_{i=1}^{6} i P_i \tag{60}$$

be the average value of the number. If the die is unbiased $a = 3.5$ but for a loaded die we could have an arbitrary $1 < a < 6$. Suppose that by some means or other we know the average vale a. How should we modify the unbiased probabilities $P_i = 1/6$ to incorporate this new information? The simplest way to do this is to introduce one more Lagrange multiplier i.e. we maximize

$$\sum_{i=1}^{6} (-P_i \ln P_i + \lambda P_i + i\beta) \tag{61}$$

and to choose λ and β so that the normalization condition and (60) are both satisfied. If we differentiate (61) with respect to P_i and put the result to zero we now find

$$0 = -\ln P_i + \lambda - 1 + i\beta$$

with solution

$$P_i = e^{\lambda-1}e^{\beta i}$$

Eliminating λ from the normalization condition gives

$$P_i = \frac{e^{\beta i}}{\sum_{i=1}^{6} e^{\beta i}} \tag{62}$$

We can substitute (62) into (60) and require that $e^{\beta i}$ be real and nonzero. If we introduce the new variable $y = e^{\beta}$ this will lead to a fifth order polynomial equation for y which will have to be solved numerically. If $1 < a < 6$ there will be only one real root form which the probabilities easily can be calculated.

It should be noted that the solution (62) need not be accurate, given the particular manufacturing flaw that produced a bias. It is, however, according to the principle of least bias the best result we can come up with from the information available.

3.4 Statistical definition of thermodynamic variables

We are now in the position that we can make a connection between thermodynamics and a statistical description of physics. We do this by developing a theory of thermodynamics from scratch and establish the contact with conventional thermodynamics by showing that the new theory gives familiar results.

Our starting point is the idea that one can count the number of available states of a system. In principle, these are discrete quantum states. For a large system the states will be very closely spaced. We will later show how one can use the *correspondence principle of quantum mechanics* to count the number of states for systems obeying classical Newtonian mechanics. The number of possible states with energy between E and $E + \delta E$ is

$$\Omega(E) = g(E)\delta E, \tag{63}$$

where $g(E)$ is the density of states. Next, consider a closed system with fixed volume V, number of particles N, and energy E. In order to avoid problems associated with the discreteness of the quantum states we take the energy to be specified within a tolerance δE. This tolerance should be chosen so that for a large system the precise value of δE does not matter. We do not know in which of the Ω allowed states the system finds itself. In fact, our fundamental assumption is that at equilibrium our ignorance in this matter is complete, and that all the $\Omega(E, V, N)$ possible states are equally likely, i.e. all memory of how the system was initially prepared is lost, except for the values of the energy, volume, and number of particles.

We define the entropy as

$$S = k_B \ln \Omega(E, N, V). \tag{64}$$

Consider next an infinitesimally small change from an equilibrium state E, V, N to another, slightly different, equilibrium state $E + dE, V + dV, N + dN$. The change in the entropy is then

$$dS = \frac{\partial S}{\partial E} dE + \frac{\partial S}{\partial V} dV + \frac{\partial S}{\partial N} dN. \tag{65}$$

The change in energy in this process is given by

$$dE = đQ + đU, \tag{66}$$

We distinguish between two forms of energy *heat* and *work*. Heat is a form of energy associated with random or *thermal* motion of atoms and molecules. Consider a gas of low density. The molecules will move in straight trajectories until they collide with other molecules or the walls of the gas container. After a few collisions it becomes practically impossible to relate the velocity and position of the molecules to the corresponding quantities at an earlier time. The difficulty is not just the enormous amount of data required to describe a large number of particles. A more fundamental problem is the fact that after a few collisions the positions and the velocities of the particles become extremely sensitive to the initial conditions. A very similar situation occurs when throwing an unbiased die or tossing a coin. In principle, it should be possible to predict the outcome of the toss using Newton's laws and the initial velocity and position. In practice, the calculation will not be able to predict the behavior of real coins, because initial conditions that give rise to radically different outcomes are so close together that the problem of specifying the initial conditions and parameters of the problem with sufficient accuracy becomes severe. This type of motion has been described as *chaotic*. Each particle is just as likely to move in any direction as in any other, and the speed of the particles is frequently changing.

We distinguish between the random motion of a molecule and bulk (ordered) movement. An example of the latter is the flight of a solid object such as a pebble thrown in the air. We refer to changes in energy associated with bulk motion or transport of matter as *work*. In (66) $đQ$ is the heat *supplied to* the system and $đU$ the work *done on* the system. The internal energy E is a state variable and its differential is exact, i.e. dE depends only on the initial and final state and is independent of the process leading to the change. On the other hand "heat" and "work" are not state variables and the partition into heat and work depends on the process. Hence, the difference in notation: dE, but $đQ$ and $đU$. Please be warned that different texts use different sign conventions for heat and work, this can be tricky, particularly for magnetic and dielectric work where the sign is not obvious.

We have not yet defined the variables P, T and μ. We want to do this in such a way as to allow us to write

$$dE = TdS - PdV + \mu dN \tag{67}$$

or

$$dS = \frac{1}{T}dE - \frac{\mu}{T}dN + \frac{P}{T}dV.$$

We now define the temperature, pressure, and chemical potential as

$$T = \left(\frac{\partial S}{\partial E}\right)^{-1}_{N,V} \qquad \mu = -T\left(\frac{\partial S}{\partial N}\right)_{E,V} \qquad P = T\left(\frac{\partial S}{\partial V}\right)_{N,E} \tag{68}$$

It is important to note that our basic assumption is that all allowed states are equally likely. The second law of thermodynamics now becomes the statement that a closed system will tend to approach a macroscopic state which can be achieved the most possible ways according to the principle of least bias. The conventional mathematical formulation of the second law (67) on the other hand only becomes an essentially trivial matter of definition. We must next show that these definitions lead to familiar looking results- otherwise they would not be useful.

3.5 The zeroth law of thermodynamics

To establish the equivalence of our definitions and the conventional thermodynamic ones we shall make contact with the zeroth law of thermodynamics.

This law has an analogy with mechanics, where in equilibrium the forces are balanced. In particular if two subsystems (which we label 1 and 2) are in contact and in equilibrium:

$$\begin{aligned} T_1 &= T_2 \rightarrow \text{ thermal equilibrium} \\ P_1 &= P_2 \rightarrow \text{ mechanical equilibrium} \\ \mu_1 &= \mu_2 \rightarrow \text{ chemical equilibrium} \end{aligned} \tag{69}$$

The zeroth law has a fairly straightforward statistical interpretation and this will allow us to begin to establish the equivalence between the statistical definitions and the conventional thermodynamic ones.

Consider two systems that are free to exchange energy but are isolated from the rest of the universe by an ideal insulating surface. The particle numbers N_1, N_2 and volumes V_1, V_2 are fixed for each subsystem. The total energy will be constant under our assumptions and we assume further that the two subsystems are sufficiently weakly interacting that

$$E = E_1 + E_2, \tag{70}$$

where E_1 and E_2 are the energies of the subsystems. Assume that the densities of state $g(E), g_1(E), g_2(E)$ are coarse grained so that $\Omega = g(E)\delta E$, $\Omega_1 = g_1(E1)\delta E$, $\Omega_2 = g_2(E_2)\delta E$. We then have

$$g(E) = \int dE_1 g_2(E - E_1) g_1(E_1), \tag{71}$$

If the subsystems are sufficiently large, the product $g_2(E - E_1)g_1(E_1)$ will be a sharply peaked function of E_1. The reason for this is that g_1 and g_2 are rapidly increasing functions of E_1 and $E - E_1$, respectively. From the definition of the entropy we note that it is a monotonically increasing function of g and that the product $g_1 g_2$ will be at a maximum when the total entropy

$$S(E, E_1) = S_1(E_1) + S_2(E - E_1) \tag{72}$$

is at a maximum. The most likely value $\langle E_1 \rangle$ of E_1 is the one for which

$$\frac{\partial S_1}{\partial E_1} + \frac{\partial S_2}{\partial E_2}\frac{\partial E_2}{\partial E_1} = 0. \tag{73}$$

Since $\partial E_2/\partial E_1 = -1$, we find using (68), that

$$\frac{1}{T_1} - \frac{1}{T_2} = 0, \tag{74}$$

or $T_1 = T_2$. The most probable partition of energy between the two systems is the one for which the two temperatures are the same.

Consider next two subsystems that are separated by a movable wall. The two systems are free to exchange energy, but the number of particles is fixed in each subsystem and the total volume $V = V_1 + V_2$ is constant. We write $E = E_1 + E_2$. The density of allowed states for the total system is then

$$g(E,V) = \sum_{volume\ settings} \int dE dE_1 g_1(E_1, V_1) g_2(E - E_1, V - V_1)$$

The integrand is sharply peaked for large systems and takes on its maximum value when $S_1(E_1, V_1) + S_2(E - E_1, V - V_1) = max$. Differentiation using

$$\left(\frac{\partial S}{\partial E}\right)_{VN} = \frac{1}{T},; \quad \left(\frac{\partial S}{\partial V}\right)_{EN} = \frac{P}{T}$$

implies that it is overwhelmingly probable that the system will be near a state for which

$$\frac{1}{T_1} = \frac{1}{T_2}; \quad \frac{P_1}{T_1} = \frac{P_2}{T_2}$$

or $T_1 = T_2, P_1 = P_2$. Conventionally, one would say that the pressure in the two compartments must be equal at equilibrium because the forces have to be in balance. The argument now being made is quite different, there are no forces, instead the movable wall is guided to its equilibrium by the invisible hand of the law of large numbers.

Similarly, consider two systems 1 and 2 which are free to exchange particles and energy. It is easy to show that the most probable configuration is the one for which $T_1 = T_2, \mu_1 = \mu_2$.

4 Boltzmann statistics

4.1 Boltzmann factor

Consider now a system in contact with a heat bath, or reservoir. System 1 is the one we are interested in, and we want to find the probability P(E_1) that it has energy E_1. We assume that system 2 is much larger than 1, so that $E_1 << E = E_1 + E_2$. Another way of putting this is to say that the heat capacity C_2 of system 2 is very large. We assume that all compatible microstates are equally likely. We have

$$P(E_1)dE_1 = \frac{g_1(E_1)g_2(E - E_1)dE_1}{\int dE_1 g_1(E_1) g_2(E - E_1)}$$

From the definition of entropy

$$g_2 = \frac{1}{\delta E} \exp\left[\frac{S_2(E - E_1)}{k_B}\right]$$

We expand in a Taylor series

$$S_2(E - E_1) = S_2(E) - E_1 \frac{\partial S_2}{\partial E} + \frac{E_1^2}{2} \frac{\partial^2 S_2}{\partial E^2} + \cdots \tag{75}$$

With T the temperature of the heat bath and C its heat capacity, the partial derivatives are given by

$$\frac{\partial S_2}{\partial E} = \frac{1}{T} \tag{76}$$

$$\frac{\partial^2 S_2}{\partial E^2} = \frac{\partial \frac{1}{T}}{\partial E} = -\frac{1}{T^2}\frac{\partial T}{\partial E} = -\frac{1}{T^2 C} \tag{77}$$

Since $E_1 << TC$ we neglect the last term in (75) giving

$$g_2 = const. \exp\left[\frac{-E_1}{k_B T}\right] = const. e^{-\beta E_1}$$

where we define $\beta = 1/(k_B T)$. We conclude

$$P(E_1) = const. g_1(E_1) e^{-\beta E_1} \tag{78}$$

The factor $e^{-\beta E_1}$ is the *Boltzmann factor.* When a system is in contact with a heat bath at a certain temperature, all possible microstates of the system are no longer equally likely. Instead, the Boltzmann factor acts as a weight factor biasing the distribution towards states with lower energy.

4.2 Partition function and the canonical distribution

The constant in (78) can be determined by *normalizing* the probability distribution i.e. requiring that

$$\int p(E)dE = 1$$

Let us define the *canonical partition function* ($\alpha = microstate$)

$$Z_c = \sum_\alpha e^{-\beta E(\alpha)} = \int dEg(E)e^{-\beta E}$$

We find that the probability $p(\alpha)$ that a state is in a given microstate

$$p(\alpha) = \frac{1}{Z_c} e^{-\beta E} \tag{79}$$

Eq. (79) is the *canonical distribution.* If $x(\alpha)$ is the value of some physical property in microstate α, and $E(\alpha)$ the energy of this state then the *canonical ensemble average* is given by

$$\langle x \rangle = \frac{1}{Z_c} \sum_\alpha x(\alpha) e^{-\beta E(\alpha)} \tag{80}$$

Equation (80) is a very useful formula, and we will give many examples of its use.

4.3 Isothermal atmosphere

Consider a gas in a gravitational field. We chose the potential energy to be zero at ground level. If the height Z is not so large that the z-dependence of the acceleration of gravity g needs to be taken into account the potential energy of a molecule at height z will be mgz, with g the m the mass of a molecule. The kinetic energy of the molecules depend only on the temperature and not on the height. At thermodynamic equilibrium the temperature of the atmosphere is constant, the density at height z must be proportional to the probability of finding a molecule there, i.e. to $\exp(-\beta mgz)$. With the density at ground level given as ρ_0 the density of height z will be

$$\rho = \rho_0 \exp(-\beta mgz) \tag{81}$$

The isothermal assumption is *not* a particularly realistic one, because the atmosphere is generally *not* in thermal equilibrium, being exposed to radiation from the sun and in turn radiating excess heat into space. Sometimes the *adiabatic atmosphere* is a better approximation.

4.4 Helmholtz free energy

For an isolated system $S = S(E, V, N)$, with E, V, N *independent variables.* For a system in contact with a heat bath at a given temperature, T becomes an independent variable, or control parameter. The energy E and entropy S will then fluctuate about their mean values $\langle E\rangle$ and $\langle S\rangle$. E and S become *dependent variables* given by equations of state. The change of variables is handled most efficiently by introducing he *Helmholtz free energy.* In thermodynamics it is defined as

$$F = E - TS$$

Imagine a reversible process which takes the system from one equilibrium state to another

$$dE = TdS - PdV + \mu dN = dE(S, V, N)$$

$$dF = dE - TdS - SdT = -SdT - PdV + \mu dN = dF(T, V, N)$$

We see that the Helmholtz free energy is should be considered to be dependent on the control variables T, V, N. We have

$$S = -\frac{\partial F}{\partial T}$$

$$P = -\frac{\partial F}{\partial T}$$

$$\mu = \frac{\partial F}{\partial N}$$

In statistical mechanics we define the Helmholtz free energy as

$$A = -k_B T \ln Z_c$$

We wish to show that for a large system

$$A = \langle E\rangle - T\langle S\rangle = \langle F\rangle \tag{82}$$

Proof:

The canonical partition function is

$$Z_c = \int dE g(E) e^{-\beta E}$$

$$= \int \frac{dE}{\delta E} \exp\left\{\beta[E - TS(E,V,N)]\right\} \tag{83}$$

We evaluate this integral using the *saddle point method.* Almost all the contribution to the integral will come from values of E near $E = \langle E \rangle$ the value for which $E - T(S,E,V,N) =$ minimum. We let $S(\langle E \rangle, V, N) = \langle S \rangle$ and

$$E - TS \simeq \langle E \rangle - T\langle S \rangle - \frac{1}{2}T(E - \langle E \rangle)^2 \frac{\partial^2 S}{\partial E^2} + \cdots$$

Substituting (77) into (83) we obtain

$$Z_c \simeq \exp[-\beta(\langle E \rangle - T\langle S \rangle)] \int \frac{dE}{\delta E} \exp\left\{\frac{-(E - \langle E \rangle)^2}{2Ck_BT^2}\right\} \tag{84}$$

Using (15) we find

$$Z_c \simeq \frac{\sqrt{2\pi k_B T^2 C}}{\delta E} \exp[-\beta(\langle E \rangle - T\langle S \rangle)]$$

and from (84)

$$A = -k_BT \ln Z_c = \langle E \rangle - T\langle S \rangle - k_BT \ln\left[\frac{\sqrt{2\pi k_B T^2 C}}{\delta E}\right] \tag{85}$$

The last term in (85) will be small compared to the first two terms for a large system, and it is possible to choose the tolerance δE so that it is identically zero. We have therefore shown that (82) is correct. We have also shown that for a large system with a constant number of particles, the volume and temperature will almost always be in a macrostate for which the Helmholtz free energy is a minimum.

4.5 Quantum harmonic oscillator

The energy levels of a quantum harmonic oscillator are $(n + \frac{1}{2})\hbar\omega$, where $\omega = 2\pi f$ is the angular frequency of oscillation. We wish to calculate the thermal properties of the oscillator when it is in contact with a heat bath at fixed temperature. The partition function is

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

The free energy is on the average

$$\langle F\rangle = \frac{-1}{\beta}\ln Z_c = \frac{\hbar\omega}{2} + \frac{1}{\beta}\ln(1-e^{-\beta\hbar\omega}) \tag{86}$$

We may also estimate the mean energy. For this purpose we derive a useful thermodynamic identity:

$$\langle E\rangle = \frac{\int dEEg(E)e^{-\beta E}}{\int dEg(E)e^{-\beta E}} = -\frac{\partial}{\partial\beta}\ln Z_c = \frac{\partial\beta\langle F\rangle}{\partial\beta} \tag{87}$$

If we apply (87) to the harmonic oscillator we get

$$\langle E\rangle = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\beta\hbar\omega}-1} \tag{88}$$

The first term in (88) is the *zero point energy.* It is important to realize that eq. (88) gives a mean value not the actual value of the energy of the oscillator.

4.6 Energy fluctuations in the canonical ensemble

In a finite size system there will be limitations on the accuracy by which we can determine thermodynamic quantities As an example let us consider the energy

$$\langle E\rangle = \frac{\sum_\alpha E_\alpha e^{-\beta E_\alpha}}{\sum_\alpha e^{-\beta E_\alpha}} = -\frac{1}{Z}\frac{\partial Z}{\partial\beta} \tag{89}$$

In order to calculate the mean square fluctuation $\langle(\delta E)^2\rangle = \langle E^2\rangle - \langle E\rangle^2$ we need

$$\langle E^2\rangle = \frac{1}{Z}\frac{\partial^2 Z}{\partial\beta^2}$$

As we shall see, the mean square fluctuation of the energy is closely related to the specific heat, which is given by

$$C_V = \frac{\partial\langle E\rangle}{\partial T} = \frac{\partial\langle E\rangle}{\partial\beta}\frac{\partial\beta}{\partial T} = -\frac{\partial\langle E\rangle}{\partial\beta}\frac{1}{k_B T^2}$$

By differentiating (89) we get

$$-\frac{\partial\langle E\rangle}{\partial\beta} = \frac{1}{Z}\frac{\partial^2 Z}{\partial\beta^2} - \frac{1}{Z^2}(\frac{\partial Z}{\partial\beta})^2 = \langle E^2\rangle - \langle E\rangle^2$$

$$\langle(\delta E)^2\rangle = k_B T^2 C_V \tag{90}$$

As we will discuss later in more detail this formula is very useful in computer simulations. In a simulation at constant temperature T the energy of the system will fluctuate during a run. We can then obtain the specific heat from an average of $\langle E^2\rangle - \langle E\rangle^2$. We can also approximate C_V by running at two different temperatures:

$$C_V \simeq \frac{\langle E(T+\delta) - E(T-\delta)\rangle}{2\delta} \tag{91}$$

If the results don't agree, this is a good indication that the simulation has not run long enough. Generally (90) is a more convenient way of calculating the heat capacity than (91) since it can be evaluated at a fixed temperature.

4.7 Problem set 2

1: A system is made up of two different kinds of atoms, A and B, and they can associate chemically to form an AB molecule. The energy of association of an AB molecule is $-|\epsilon|$. Assume that there are N_A and N_B atoms of each kind (either free or as part of an AB molecule). What is the equilibrium number of AB molecules at temperature T? Assume that the mixture forms an ideal gas with a Helmholtz free energy which is the sum of the contributions from each component. Contributions to the free energy of the AB molecule from internal degrees of freedom are assumed to be incorporated into ϵ. The mass of an A atom is m_A that of a B atom m_B.

2: In a certain system the internal energy E is related to the entropy S, particle number N, and volume V through

$$E = \text{const}N\left(\frac{N}{V}\right)^D \exp\left[\frac{DS}{Nk_B}\right]$$

(a) Show that the system satisfies the ideal gas law independent of the value of the constant D.

(b) Find the coefficient γ in the adiabatic equation of state $PV^\gamma = const.$ and the molar specific heats C_P and C_V of the system. (Adiabatic here means that the entropy is constant)

3:

Imagine that a hole has been drilled to the center of a small planet of radius R. If the acceleration of gravity at the surface of the planet is g it will, at distance $r < R$ from the center, be $\frac{gr}{R}$. If the atmospheric pressure at the surface is P, find the pressure $P(r)$, for $r < R$, assuming an isothermal ideal gas at temperature T. (Hint: If the gravitational force $mg(r)$ depends on the distance to the center of the planet r, the potential energy will not be just mgr).

5 Ideal gases and generalizations

5.1 Counting quantum states

In statistical mechanics one is very often dealing with a situation in which there are a very large number of possible microstates which a priori are equally likely. Each macrostate can be realized in very many different ways and the probability of the different macrostates will be proportional to the number of allowed realizations. This means that we must learn how to count states. In the examples so far the counting of microstates was not too difficult, it was fairly obvious what constituted a discrete microstate. In classical mechanics one is dealing with continuous states - can we still define the number of states'? As we shall see this can be done, but we will need to appeal to quantum mechanical ideas. Consider first a free particle in a cubic box (see fig. 8a) The volume of the box $V = L^3$., where L is the length of a side of the cube. First consider the simplest case of periodic boundary conditions. The wave function of the particle satisfies

$$\psi(x = 0) = \psi(x = L).$$

The quantum eigenstates of the system are then given by

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot r} = \frac{1}{\sqrt{V}} \exp(i[k_x x + k_y y + k_z z])$$

where

$$k_x = \frac{2\pi n_x}{L}, k_y = \frac{2\pi n_y}{L}, k_z = \frac{2\pi n_z}{L} \tag{92}$$

and n_x, n_y, n_z can take on the values $0, \pm1, \pm2...$ i.e. are integers. The number of allowed k-values inside a "volume" d^3k in k-space is from (92), noting the relation $p = \hbar k$ between momentum and wave-vector

$$\frac{V d^3k}{(2\pi)^3}$$

When summing over allowed k-states we can replace a sum by an integral in the large V, or continuum limit

$$\sum_{\mathbf{k}} \to \frac{V}{(2\pi)^3} \int d^3k \tag{93}$$

The same result holds if ψ is the amplitude of a wave instead of the wave-function of a particle and we want to sum over the allowed wave vectors. Later on when we get to counting states described by classical (Newtonian) mechanics we generalize the above result to the concept of *phase space*. The x-component of the momentum is $p_x = \hbar k_x = hn_x/L$ and we have similar expressions for the other components. Consider now a volume in *phase space* $W = L^3 \Delta p_x \Delta p_y \Delta p_z$ where the components are in the intervals

$$\begin{aligned} p_1 &< p_x < p_1 + \Delta p_x \\ p_2 &< p_x < p_2 + \Delta p_y \\ p_3 &< p_z < p_3 + \Delta p_z \end{aligned} \tag{94}$$

We find that the number of states in W is given by

$$\Omega = \frac{L^3 \Delta p_x \Delta p_y \Delta p_z}{h^3} = \frac{W}{h^3} \tag{95}$$

The rule (97) must be applied with some discretion. Suppose we change the boundary conditions and consider a box with hard walls. The boundary condition is $\psi = 0$ at surface. The eigenstates are now

$$\psi = \sqrt{\frac{8}{V}} \sin(k_x x) \sin(k_y y) sin(k_z z)$$

with

$$k_x = \frac{\pi n_x}{L},\ k_y = \frac{\pi n_y}{L},\ k_z = \frac{\pi n_z}{L} \tag{96}$$

Note the absence of the factor of 2 in (96). The number of states in the interval (94) is then

$$\frac{2^3 \Delta p_x \Delta p_y \Delta p_z L^3}{h^3} = \frac{8W}{h^3}$$

The reason for this apparent paradox is that the 'standing wave' states associated with the hard wall boundary condition must be considered to be mixtures of states propagating back and forth. The states k_x and -k_x are not distinct. So while

$$\frac{\#\ of\ states}{Volume\ of\ phase\ space} = 8\ times\ larger$$

only 1/8 th of the momentum states are distinct. The periodic boundary condition result corresponds to a classical picture where particles can move back and forth and the components of the momentum can be either positive or negative.

We generalize our results for free particles in a box through the *Bohr-Sommerfeld quantization rule.* Suppose we have a particle with coordinate q and momentum p. By integrating the equations of motion $\dot{q} = \frac{p}{m}$, $\dot{p} = force$, one can in principle construct the 'orbits' $p(q)$. The $p - q$ plane is called the *phase plane* (figure 8(a)). The rule is now that the orbit for which

$$\int_{orbit} p(q)dq = nh$$

correspond to quantum states if n is an integer and h is Planck's constant. Consider next an 'area' W in the phase plane containing many orbits.

$$\int dp \int dq = W$$

The number of quantum states in W is then $\Omega = W/h$. We can generalize to a system with many degrees of freedom $q_1, q_2..q_{\mathcal{N}} p_1, p_2..p_{\mathcal{N}}$. To a 'volume' in phase space

$$W = \int dq_1 \int dq_2 ... \int dq_{\mathcal{N}} \int dp_1 \int dp_2 ... \int dp_{\mathcal{N}}$$

corresponds a number of states

$$\Omega = \frac{W}{h^{\mathcal{N}}} \tag{97}$$

5.2 Black body radiation

Electromagnetic radiation is made up of transverse waves. For each wave vector $\mathbf{k}$ there are two modes, or *polarizations.* The quanta of electromagnetic radiation are called photons. We have

$$\omega = ck \tag{98}$$

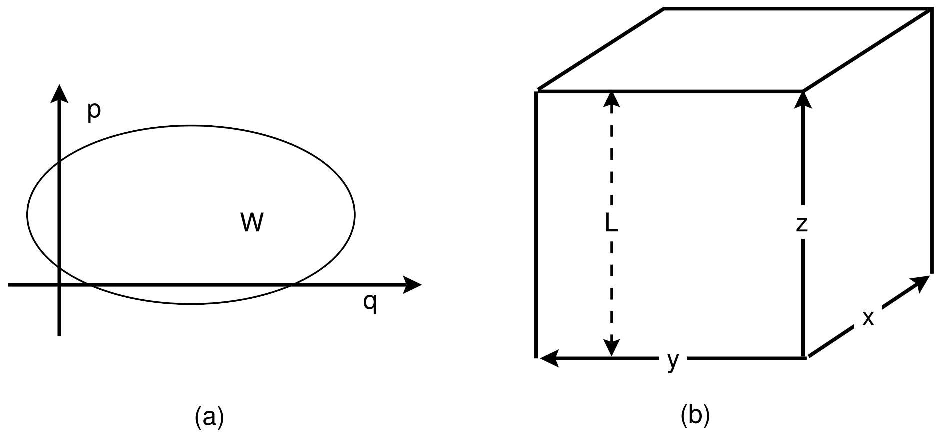


Figure 8: (a) Orbit in phase plane. (b) Particles in box.

where c is the velocity of light, and $\hbar\omega$ is the energy of the photon. The photon momentum is

$$p = \hbar k = \frac{hk}{2\pi}$$

The number of modes in an element d^3k of 'wave vector space' is thus

$$\frac{2V d^3k}{(2\pi)^3}$$

where V is the volume of the cavity containing the black body radiation. There may be 0,1,2,3... photons with any allowed wave vector. The partition function the modes with a given wave vector is

$$Z = \sum_{n=0}^{\infty} e^{-n\beta\hbar\omega} = \frac{1}{1 - e^{-\beta\hbar\omega}}$$

The free energy associated with the photons of any one wave vector is from our discussion of the harmonic oscillator

$$f(\omega) = k_B T \ln(1 - e^{\beta\hbar\omega})$$

From (98)

$$d^3k = \frac{4\pi\omega^2 d\omega}{c^3}$$

The free energy of the black body radiation is thus from (86)

$$F = \frac{V}{\beta \pi^2 c^3} \int_0^\infty \omega^2 d\omega \ln(1 - e^{-\beta\hbar\omega}) \tag{99}$$

We substitute $x = \beta\hbar\omega$ into (99) and get

$$F = \frac{-VI}{\beta^4 \pi^2 c^3 \hbar^3}$$

where one can show that

$$I = -\int_0^\infty x^2 dx \ln(1 - e^{-x}) = \frac{\pi^4}{45}$$

using the method of contour integration. We get

$$F = -\frac{\pi^2 V}{45\hbar^3 c^3 \beta^4} \tag{100}$$

From (87)

$$E = \frac{\pi^2 V}{15\hbar^3 c^3 \beta^4}$$

which the reader might recognize as the *Stefan-Boltzmann law* of radiation. The *radiation pressure* of black body radiation is

$$P = -\frac{\partial F}{\partial V} = \frac{E}{3V}$$

This can be considered as the equation of state for black body radiation. Since we have an explicit form for the free energy in terms of its independent variable we compute anything we want to know about the thermodynamics of black body radiation by using appropriate thermodynamic identities.

5.3 The classical ideal gas

We next show how we can calculate the partition function for an ideal gas and demonstrate that this partition function yields the expected expressions for the energy and pressure. In addition we find expressions for the entropy and chemical potential which cannot be obtained directly from the ideal gas laws. Consider a system made up of a single particle in contact with a heat bath. From our state counting rules the number of states with momentum components between p_x and $p_x + dp_x$, p_x and $p_x + dp_x$ is $V dp_x dp_y dp_z / h^3$ and

$$Z_1 = \frac{1}{h^3} V \int_{-\infty}^{\infty} dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z \exp\left(\frac{-p_x^2 + p_y^2 + p_z^2}{2mk_B T} \right) = \frac{V}{h^3} \left(\frac{2\pi m}{\beta} \right)^{3/2} \tag{101}$$

It is convenient to define the *thermal wavelength* λ as

$$\lambda = \sqrt{\frac{h^2 \beta}{2\pi m}} \tag{102}$$

With this definition we find that the single particle partition function is given by the simple expression

$$Z_1 = \frac{V}{\lambda^3} \tag{103}$$

From Eq.(87) we find the familiar result for the mean energy per particle of a monatomic gas with no internal degrees of freedom

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} = \frac{3}{2\beta} = \frac{3}{2} k_B T \tag{104}$$

Using

$$\langle P \rangle = \frac{\partial k_B T \ln Z}{\partial V}$$

we find mean contribution per particle to the pressure to be $k_B T / V$. If we assume that the total pressure is the sum of the contributions from each particle we find

$$P = N k_B T \tag{105}$$

again a familiar result. Getting the entropy or the chemical potential is a bit more tricky. The problem is that the free energy of N particles is not just N times the free energy of a single particle. If the particles in the gas are indistinguishable, quantum mechanics tells us that we must not treat as distinct states that only differ by which particles are in which state. Instead we find the number of states where there are particles in the system with momenta between $p_{x,i}$ and $p_{x,i}+dp_{x,i}$, $p_{y,i}$ and $p_{y,i}+dp_{y,i}$, $p_{z,i}$ and $p_{z,i}+dp_{z,i}$, for $i = 1, \dots N$.

$$\frac{V^N}{h^{3N} N!} \Pi_i^N dp_{x,i} dp_{y,i} dp_{z,i}. \tag{106}$$

If the factor $N!$ had not been included we would have counted the state where particle 1 has momentum p_1 and particle 2 has momentum p_2 as distinct from a state in which it is particle 2 that has momentum p_1 and particle 1 has momentum p_2. The need for the inclusion of the $N!$ term goes under the name Gibbs paradox. Before the introduction of quantum mechanics this term was hard to justify. We will come back to Gibbs paradox later on when we discuss entropy of mixing.

It is useful to notice the contrast between the way we treat the black body radiation problem and the ideal gas. In the case of the radiation we analyzed the *states* and we summed up the contributions to the partition function from situations in which there were 0,1,2.. photons in each state. Nothing is said about which photon is in which state, and the states are distinct. Hence there is no need for any $N!$ factor when we sum over the states. In the case of the ideal classical gas we number the particles and sum over all allowed states for each particle. If the particles are identical this leads to over counting of the number of different microstates, and we correct for this by dividing by $N!$.

We find for the partition function of an ideal gas of N particles

$$Z_N = \frac{Z_1^N}{N!} = \frac{V}{N! \lambda^{3N}} \tag{107}$$

Using Stirling's formula $\ln N! \approx N \ln N - N$ we find the free energy

$$\langle A \rangle = -N k_B T (\ln \frac{V}{N\lambda^3} + 1) \tag{108}$$

Differentiation of the expression for A, using

$$\langle S \rangle = \frac{\partial k_B T \ln Z}{\partial T}$$

we obtain the so-called Sackur–Tetrode formula

$$\langle S \rangle = N k_B \ln \frac{V}{N\lambda^3} + \frac{5}{2} N k_B \tag{109}$$

and

$$\langle\mu\rangle = k_B T \ln(\frac{N\lambda^3}{V}) \tag{110}$$

In the case of molecules containing more than one atom there will be corrections to the entropy and chemical potential from vibrational and rotational degrees of freedom. We will come back to this later.

5.4 Gibbs-Duhem relation, and the Gibbs energy.

Quantum systems are often more conveniently handled if we allow the number of particles to be variable, rather than considering a fixed number of particles. We will also frequently encounter systems in contact with a heat bath, that can exchange particles at a given chemical potential. Before carrying out the change in variables which does this, it is useful to derive another thermodynamic identity. Consider a system described by the micro canonical ensemble. The independent variables S, V, N and the energy $E(S, V, N)$ are assumed to be extensive while the dependent variables

$$T = \frac{\partial E}{\partial S},\ P = -\frac{\partial E}{\partial V},\ \mu = \frac{\partial E}{\partial N}$$

are intensive. Consider two systems the second γ times larger than the first but identical in all other respects.

$$\gamma E = E(\gamma S, \gamma V, \gamma N)$$

Since the temperature, pressure and chemical potential are the same in the two systems differentiation with respect to γ gives

$$E = ST - PV + \mu N \tag{111}$$

The Gibbs free energy is defined as

$$G = E - ST + PV$$

and we get the *Gibbs-Duhem relation*

$$\frac{G}{N} = \mu \tag{112}$$

This relationship will prove very useful in what follows.

5.5 The grand canonical ensemble

We now wish to consider systems in which the number of particles is a dependent variable. Just as we did earlier, we consider two systems in thermal contact with each other. We will also assume that the two systems are separated by a permeable wall, i.e. they are free to exchange particles. We write

$$E_1 + E_2 = E_T$$

$$N_1 + N_2 = N_T$$

The density of states for the two systems together is

$$g(E_T, N_T) = \sum_{N_1=0}^{N_T} \int dE_1 g_1(E_1, N_1) g_2(E_T - E_1, N_T - N_1)$$

As before we expect the integrand to be sharply peaked for values of N_1 and E_1 such that the sum of the entropies $S_1(E_1, N_1) + S_2(E_T - E_1, N_T - N_1) =$ *maximum* Using

$$\frac{\partial S}{\partial E} = \frac{1}{T} \quad \frac{\partial S}{\partial N} = \frac{\mu}{T} \tag{113}$$

we find that the maximum occurs when $\mu_1 = \mu_2, T_1 = T_2$ in agreement with the zeroth law of thermodynamics.

Now, assume that system 2 is much larger than system 1 so that the former can be consider a reservoir of energy and particles. System 1 is a system whose properties we wish to study as a function of temperature and chemical potential. We have

$$g_2(E_T - E_1, N_T - N_1) = \frac{\exp[S_2(E_T - E_1, N_T - N_1)/k_B]}{\delta E}$$

$$= g_2(E_T, N_T) \exp\left[\frac{-1}{k_B}(E_1 \frac{\partial S_2}{\partial E} + N_1 \frac{\partial S_2}{\partial N})\right]$$

Using (113) we find for the probability that system 1 has energy E_1 and N_1 particles

$$p(E_1, N_1) = \frac{g_1(E_1, N_1)}{Z_G} e^{-\beta(E_1 - \mu N_1)}$$

where Z_G is the *grand canonical partition function*

$$Z_G(\mu, T, V) = \sum_{N=0}^{\infty} \exp[\beta \mu N] Z_c(N, T, V) \tag{114}$$

In the *grand canonical ensemble* a physical property x(α) has the average value

$$\langle x \rangle = \frac{1}{Z_G} \sum_\alpha x(\alpha) e^{-\beta(E(\alpha)-\mu N(\alpha))} \tag{115}$$

T, μ, V are now the control parameters and S, N and P are dependent variables. We can proceed in analogy with what we did for the canonical ensemble and write for a large system

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} \exp[\beta\mu N] Z_c \\ &\simeq \sum_{N=0}^{\infty} \exp[-\beta(F - \mu N)] \end{aligned}$$

I leave as an exercise to the reader to convert the sum to an integral and to use the saddle point method to get

$$-k_B T \ln Z_G \equiv L = F - \mu N$$

where L is called the *grand potential* or the Landau free energy[3]. If we use

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

We find for an infinitesimal reversible process

$$dL = -SdT - PdV - Nd\mu = dL(T, V, \mu)$$

i.e. the natural independent variable for the Landau potential are the temperature, volume and chemical potential and

$$\begin{aligned} S &= -\frac{\partial L}{\partial T} \\ P &= -\frac{\partial L}{\partial V} \\ N &= -\frac{\partial L}{\partial \mu} \end{aligned} \tag{116}$$

We interpret (115) as implying that for a system with fixed temperature, volume and chemical potential the equilibrium state - the state most likely to occur - is the one for which the Landau potential is as small as possible.

From the Gibbs-Duhem relation (112) $E = TS - PV + \mu N$ we find that

$$L = -PV \tag{117}$$

[3] After Lev Davidovich Landau 1908-68

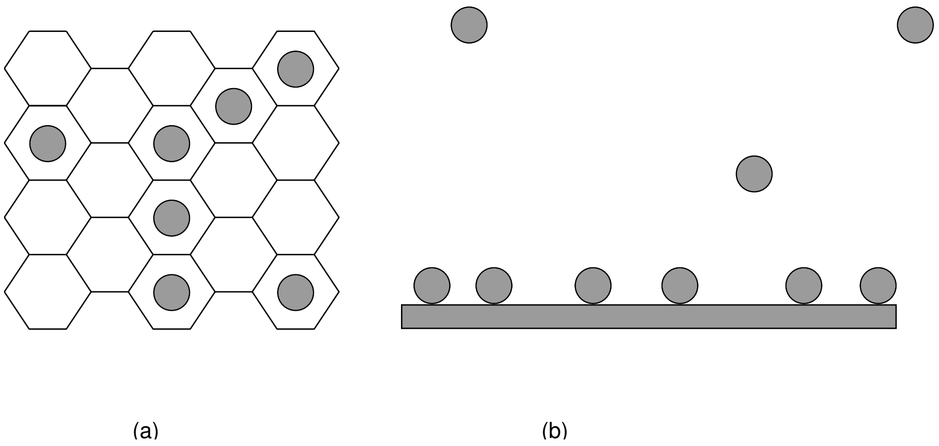


Figure 9: Physisorption at a surface. (a) Honeycomb graphite surface

5.6 Langmuir adsorption isotherm

We wish to construct a model for *physisorption*, the adsorption of gas atoms at adsorption sites on the surface of a solid (figure 9). We define $|\epsilon| = -\epsilon$ as the adsorption energy per site and let μ be the chemical potential the surrounding gas. We assume that we are dealing with a dilute gas. The ideal gas formula (110) for μ should therefore be a good approximation.

A material which adsorbs many gases is graphite. The carbon atoms in graphite form a planar honey-comb lattice (figure 9(a)). The centers of the honey-combs often form good adsorption sites. Graphite can be exfoliated, and then has an enormous surface to volume ratio. The attractive interaction responsible for the adsorption is the van der Waals force between the carbon layer and the gas atoms. This interaction is typically stronger than the gas atom gas atom interaction. We make the following approximations:

> (1) The atoms occupy discrete adsorption sites that are either occupied or empty. The adsorbed phase is taken to be a *lattice gas*.
>
> (2) We neglect the interaction between adsorbed gas atoms. At low temperatures the adsorbed atoms may condense to form puddles and our model will then break down.

In the gas phase the chemical potential is given by (110). It is conventional to define the *fugacity* as

$$z = e^{\beta\mu}$$

For the ideal gas

$$z = \frac{N}{V}\lambda^3 = \frac{P\lambda^3}{k_B T}$$

Consider now a system consisting of a single adsorption site. This system has only two microstates: empty or occupied. The grand canonical partition function for this system is

$$Z_G = 1 + e^{-\beta(\epsilon - \mu)}$$

the probability that the site is occupied is thus

$$f = \frac{e^{-\beta(\epsilon-\mu)}}{e^{-\beta(\epsilon-\mu)} + 1} = \frac{z}{e^{\beta\epsilon} + z} = \frac{P}{P_0 + P} \tag{118}$$

where

$$P_0 = \frac{k_B T}{\lambda^3} e^{\beta\epsilon}$$

Equation (118) is called the *Langmuir isotherm*. If $|\epsilon| >> k_B T$ $f \to 1$. Note that if $f \simeq 1$ one must consider the possibility that more than one layer is formed.

5.7 Fermi Dirac distribution

Subatomic particles such as electrons, protons, neutrons and muons are *Fermions*, they satisfy the Pauli exclusion principle: two Fermions of the same species cannot occupy the same state. Let us consider a system to be a single particle state characterized by a wave vector $\mathbf{k}$ and spin σ which is either $\uparrow$ or $\downarrow$. This state is either occupied or not.

$$Z_{G1} = 1 + e^{-\beta(\epsilon_{k\sigma} - \mu)}$$

where $\epsilon_{k\sigma} is$ the energy of the state. The probability that a state is occupied is then

$$\langle n_{k\sigma} \rangle = \frac{1}{\exp[\beta(\epsilon_{k\sigma} - \mu)] + 1} \tag{119}$$

the probability that the state is empty is

$$1 - \langle n_{k\sigma} \rangle = \frac{1}{1 + \exp[-\beta(\epsilon_{k\sigma} - \mu)]}$$

The qualitative behavior of the *Fermi-Dirac* distribution (119) is drawn in figure (87). As $T \to 0, \beta \to \infty$ the distribution approaches a step function

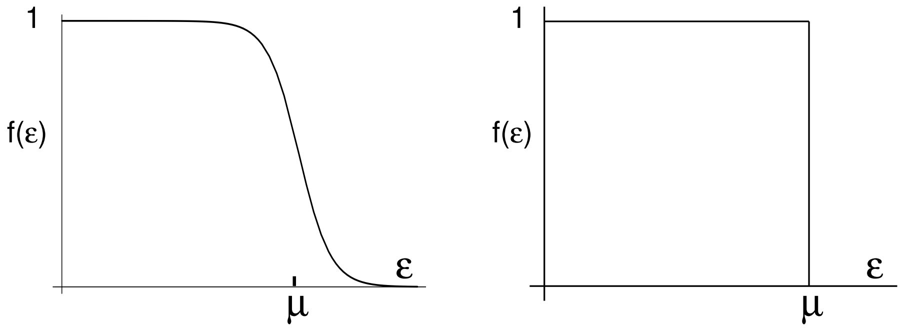


Figure 10: (a) Fermi Dirac distribution at nonzero temperature. (b) Fermi Dirac distribution at T=0.

$n_{k\sigma} = 1$, for $\epsilon_{k\sigma} < \mu, n_{k\sigma} = 0$ for $\epsilon_{k\sigma} > \mu$ (figure 10(b)). The *Fermi energy* (121) is related to the chemical potential by

$$\epsilon_F = \mu(T = 0)$$

If the Fermions are non-interacting the different states are independent and we get for the grand canonical partition function of an ideal gas of Fermions

$$Z_G = \prod_{\mathbf{k}\sigma}(1 + e^{-\beta(\epsilon_{k\sigma} - \mu)}) \tag{120}$$

5.8 Free electron gas at zero temperature

At zero temperature the system will be in its lowest energy state, the 'ground state'. Electrons are *Fermions* and the Pauli exclusion principle for Fermions states that no two particles can occupy the same state. The single particle states can be characterized by a wave vector $\mathbf{k}$ and spin σ. The number of allowed k-values inside a 'volume' d^3k in k-space is from (95), noting the relation $p = \hbar k$ between momentum and wave-vector

$$\frac{V d^3 k}{(2\pi)^3}$$

When summing over allowed k-states we can replace a sum by an integral in the large V, or continuum limit as shown in(93).We will neglect the interaction between the electrons. The energy of a state with wave vector k

is $E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}$, i.e. it increases monotonically with increasing k. The lowest energy state then has all momentum states occupied up to a maximum wave vector k_F, the *Fermi wave vector.* The number of states with wave vector less than k_F is then (including a factor 2 for spin)

$$N = V \frac{2 \times 4\pi {k_F}^3}{3(2\pi)^3} = \frac{V {k_F}^3}{3\pi^2}$$

We refer to the k-space volume with $k < k_F$ as the *Fermi sphere* and its surface as the *Fermi surface.* The momentum of a particle with wave vector k_F is called the *Fermi momentum* $p_F = \hbar k_F$. The energy of a particle with wave vector k_F is the *Fermi energy*

$$\epsilon_F = \frac{{p_F}^2}{2m} \tag{121}$$

The free electron gas is often used as a starting point to describe the electronic properties of metals. Let us make some order of magnitude estimates from this application.

Length scale: The *Wigner-Seitz radius* r_s is defined as the ratio

$$\frac{radius\ of\ sphere\ containing\ one\ electron}{Bohr\ radius}$$

so that

$$\frac{V}{N} = \frac{4\pi {a_0}^3 {r_s}^3}{3}$$

Here $a_0 = 0.529 \times 10^{-10} m$ is the Bohr radius. If Z is the valence of a metal then the number of conduction electrons is $N = ZN_{atoms}$. From tables of the density, atomic weight and Avogadro's number one can then work out the value of r_s. For the metallic elements one typically has $2 < r_s < 6$. Aluminum is a high density metal with $r_s = 2.07$, cesium is a low density metal with $r_s = 5.62$.

Wave vectors: We have

$$\frac{{k_F}^3}{3\pi^2} = \frac{N}{V} = \frac{3}{4\pi {a_0}^3 {r_s}^3} \tag{122}$$

giving

$$a_0 k_F = (\frac{9\pi}{4})^{1/3} \frac{1}{r_s} = \frac{1.92}{r_s} \tag{123}$$

The Fermi wave vector k_F is of the order of the inverse Bohr radius.

Velocity: The *Fermi velocity* is given by

$$v_F = \frac{\hbar k_F}{m} = \frac{4.2}{r_s} \times 10^6 m s^{-1}$$

This velocity is of the order 1% of the velocity of light, or 1000 times a typical sound velocity. We conclude that electrons in metals are supersonic, but non-relativistic.

Energy: It is instructive to compare the Fermi energy with the Rydberg, the ground state energy of the hydrogen atom. We have

$$1Ry = \frac{e^2}{4\pi\epsilon_0 2a_0} = 13.6eV$$

From (123)

$$\epsilon_F = \frac{\hbar^2 {k_F}^2}{2m} = \frac{\hbar^2}{2m{a_0}^2}\left(\frac{9\pi}{4}\right)^{2/3}\frac{1}{{r_s}^2} \tag{124}$$

Substituting numbers gives

$$\epsilon_F = \frac{50.1eV}{{r_s}^2}$$

For aluminum, $\epsilon_F = 11.7eV$, for cesium $\epsilon_F = 1.59eV$. The model which we have described above is commonly called the Sommerfeld model. We have neglected the electron-electron interaction. Real electrons are charged particles. The energy associated with the Coulomb repulsion between two electrons a distance r apart is

$$\frac{e^2}{4\pi\epsilon_0 r}$$

A typical nearest neighbor distance between electrons is $\sim a_0 r_s$ and we see that the Coulomb energy and the kinetic energy is comparable. Nevertheless, the Sommerfeld model explains many of the qualitative properties of the physics of metals. Since the kinetic energy is $\sim 1/{r_s}^2$ and the model will work best when r_s is small- and for many purposes aluminum is the most free electron like metal.

Temperature: Thermal energies available to electrons are of the order $k_B T$. It is instructive to compare this quantity to the Fermi energy. For this purpose one can define the Fermi temperature as

$$T_F = \epsilon_F / k_B$$

The Fermi temperature is $13.6\times10^4 K$ for Al and 1.84×10^4 for cesium. These are high temperatures compared to room temperature.

5.9 Gravitational collapse of a star

We next will next make some order of magnitude estimates for a situation which is very different from what obtains here on earth. Our argument will show that when the properties of matter are dominated by long range forces such as gravitational interactions we do not have a neat separation into extensive and intensive thermodynamic variables. Since gravitational interactions are attractive any body held together by gravitational forces alone would find it energetically favorable to become as compact as possible. If the object has a net electric charge the Coulomb interaction would overcome the gravitational forces so we will consider a star as made up of an equal number N of electrons and protons that are contained inside a radius R, volume $V = 4\pi R^3/3$. Somewhat arbitrarily we also assume that there is an equal number of neutrons and protons.

On earth the gravitational pressure is not large enough to overcome the repulsive forces between atoms and molecules at short distances. Inside the sun matter will not exist in the form of atoms and molecules, but since it is still burning there is radiation pressure which keeps it from collapsing. Let us consider a burnt out star such as a white dwarf.

We assume that the temperature of the star is low enough compared to the electron Fermi temperature that the electrons can be approximated by a $T = 0$ electron gas. The kinetic energy of the protons and neutrons can be neglected in comparison (see (124) noting that the mass of protons and neutrons is much larger that the mass of electrons).

Let us first assume that the electron gas is *non relativistic* the kinetic energy of the electron gas is then with m_e the electron mass

$$E_{kin} = 2\sum_{k<k_F} \frac{\hbar^2 k^2}{2m_e} = \frac{2V}{(2\pi)^3}\frac{\hbar^2}{2m_e}4\pi\int_0^{k_F} k^4 dk = \frac{2R^3\hbar^2 {k_F}^5}{15\pi m_e}$$

From (122) we find $Rk_F = (\frac{9\pi N}{4})^{1/3}$ and we conclude that the expression for the kinetic energy can be rewritten as

$$E_{kin} = \frac{3\hbar^2}{10m_e}\left(\frac{9\pi}{4}\right)^{2/3}\frac{N^{5/3}}{R^2}$$

The gravitational potential energy is dominated by the neutrons and protons. Let m_n be the nucleon mass. The mass density will then be approximately $\rho = 2Nm_n/V$ if there is an equal number of protons and neutrons. With G the gravitational constant we find for the potential energy

$$E_{pot} = -\int_0^R dR 4\pi R^2 \frac{G 4\pi R^3}{3R}\rho^2 = -\frac{12}{5}{m_n}^2 G\frac{N^2}{R} \tag{125}$$

Suppose N is fixed. The radius for which the kinetic plus potential energy is a minimum can then be found by differentiation. We find

$$R_0 = \frac{1}{4}\left(\frac{9\pi}{4}\right)^{2/3} \frac{\hbar^2}{N^{1/3} m_e {m_n}^2 G}$$

The gravitational constant is $6.67 \times 10^{-11} Nm^2/kg^2$, the mass of a nucleon $1.67 \times 10^{-27} kg$, the mass of an electron is 9.11×10^{-31}. The solar mass is $1.99 \times 10^{30} kg$. The radius of the sun is $6.96 \times 10^8 m$. The radius for which the energy of a white dwarf with the same mass as the sun will be minimum will then be $\sim 7.2 \times 10^6 m$ or $\sim 1/100$ that of the sun.

If the density is too large the Fermi velocity becomes comparable with the velocity of light and we should use the relativistic formula

$$\epsilon(p) = \sqrt{m^2c^4 + p^2c^2}$$

for the relationship between energy and momentum. Consider next an ultra relativistic gas ($\epsilon \simeq cp$) the electron kinetic energy will now be

$$E_{kin} = 2\sum_{k<k_F} \hbar kc = \frac{2V}{(2\pi)^3} 4\pi\hbar c \int_0^{k_F} k^3 dk = \frac{N^{4/3}}{R}\frac{\hbar c}{3\pi}\left(\frac{9\pi}{4}\right)^{4/3} \tag{126}$$

while the potential energy is unchanged. Comparing (126) and (125) we find that for fixed N the potential and gravitational energies are both inversely proportional to R. Since $N^2 >> N^{4/3}$ for large N we find that if the mass of the star is large enough the potential energy will dominate. This suggests that if a star is sufficiently heavy it will undergo a gravitational collapse. The critical value of N is

$$N_{crit} = \left(\frac{5\hbar c}{36\pi {m_n}^2 G}\right)^{3/2} \left(\frac{9\pi}{4}\right)^2 \tag{127}$$

Substituting numbers we find that this corresponds to approximately 1.71 solar masses.

5.10 Low temperature electronic specific heat

At nonzero temperatures not all the electrons will occupy states inside the Fermi sphere. Some electrons will be excited from states inside the Fermi sphere to an electronic state outside the sphere leaving a *hole* behind, see figure 11(b). Let $f(\epsilon)$ be the probability that a state with energy ϵ is occupied in a system with chemical potential μ

$$f(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)}+1} \tag{128}$$

The mean number of particles is then given by

$$\langle N \rangle = 2\frac{V}{(2\pi)^3}\int d^3k f(\epsilon_k) \tag{129}$$

The condition (129) determines the chemical potential μ. For $T = 0$ we have $\mu = \epsilon_F$, while for $T \neq 0$ μ will be a slowly varying function of T. When discussing thermodynamic properties of the Fermi gas it is convenient to change the integration variable from wave vector to energy. We have

$$\epsilon = \frac{\hbar^2k^2}{2m}, d^3k = 4\pi k^2 dk, \quad d\epsilon = \frac{\hbar^2 k}{m}dk$$

$$\langle N \rangle = \int_{-\infty}^{\infty} d\epsilon\ \mathcal{D}(\epsilon)f(\epsilon)$$

where the *density of states* is given by (see figure 11)

$$\mathcal{D}(\epsilon) = \frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{\epsilon}$$

The density of states allows us to evaluate thermal averages simply. The internal energy of the ideal Fermi gas is e.g.

$$U = \int_{-\infty}^{\infty} \mathcal{D}(\epsilon)\epsilon f(\epsilon)d\epsilon \tag{130}$$

Let us use this result to find a formula for the specific heat

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

The only quantity in (130) which depends on temperature is $f(\epsilon)$. The trouble is that the other independent variable is μ not N. When differentiating

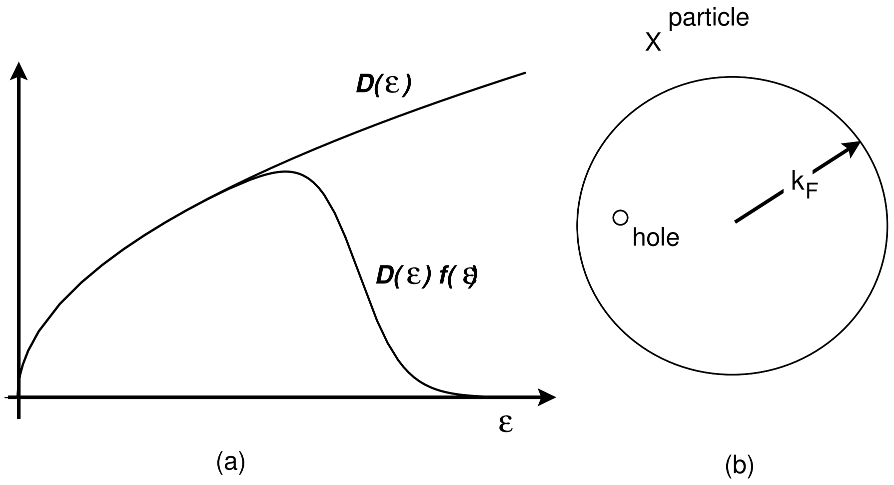


Figure 11: (a)Density of states of free electrons. (b) Particles and holes.

f with respect to T we must therefore consider μ to be an implicit function of T, N, V. We can obtain the leading term in a low temperature expansion for the specific heat by noting that for $T = 0, \mu = \epsilon_F$. The temperature dependence of U at constant N comes about because of *thermal excitations* in which particles are excited into previously unoccupied states outside the Fermi surface leaving *holes* behind (figure 11 (b)).

We use ϵ_F as our *reference energy*: Then $\epsilon - \epsilon_F$ is the excitation energy of a particle outside Fermi surface. Similarly the excitation energy of a hole, i.e. a particle missing from a single particle state of energy ϵ is $\epsilon_F - \epsilon$. The change in internal energy due to thermal excitations is then

$$\Delta U = \int_{\epsilon_F}^{\infty} d\epsilon(\epsilon - \epsilon_F) f(\epsilon)\mathcal{D}(\epsilon) + \int_{-\infty}^{\epsilon_F} d\epsilon(\epsilon_F - \epsilon)(1 - f(\epsilon))\mathcal{D}(\epsilon)$$

We differentiate inside the integrand to get

$$C_V = \int_{-\infty}^{\infty} d\epsilon(\epsilon - \epsilon_F)\frac{\partial f}{\partial T}\mathcal{D}(\epsilon) \tag{131}$$

We now approximate $\mu \simeq \epsilon_F$ and neglect the temperature dependence of μ

$$\frac{\partial f}{\partial T} = \frac{\epsilon - \epsilon_F}{k_B T^2} \frac{e^{\beta(\epsilon - \epsilon_F)}}{(e^{\beta(\epsilon - \epsilon_F)} + 1)^2}$$

Because of the factor $\partial f/\partial T$ the integrand in (131) will be sharply peaked near $\epsilon = \epsilon_F$. If we assume that near $\epsilon = \epsilon_F, \mathcal{D}(\epsilon) \simeq \mathcal{D}(\epsilon_F)$ and put $x = \beta(\epsilon - \epsilon_F)$ we get

$$C_V \simeq k_B{}^2 T \mathcal{D}(\epsilon_F) \int_{-\infty}^{\infty} dx \frac{x^2 e^x}{(e^x+1)^2} \tag{132}$$

The integral in (132) can be evaluated to yield $\frac{\pi^2}{3}$ and we get the formula for the electronic specific heat

$$C_V = \frac{T\pi^2 k_B{}^2}{3} \mathcal{D}(\epsilon_F) \tag{133}$$

Note that we have not made use of the formula $\epsilon = \frac{\hbar^2 k^2}{2m}$ for the electron energy. In a metal electrons occupy *energy bands* where $\epsilon_{\mathbf{k}}$ may be a complicated function of $\mathbf{k}$. Equation (133) still remains valid in this situation.

If we substitute the free electron value for the energy we get

$$C_V = \frac{\pi^2 N k_B}{2} \frac{T}{T_F} \tag{134}$$

An important aspect of (133) and (134) is the linear temperature dependence. We will later show that for low temperatures the contribution to the specific heat from lattice vibrations will be proportional to T^3. Therefore, the specific heat of metals at low temperatures (T$\leq$ $10K$) will be dominated by the electronic contribution. This contribution is still small. The factor T/T_F in (134) indicates that only a small fraction of the conduction electrons will be thermally excited even at room temperature. At ordinary temperatures the lattice specific heat dominates.

5.11 Problem set 3:

1 A cylinder of cross sectional area A and height L is sealed at both ends. It contains an airtight piston (mass m) and N ideal gas molecules on each side of the piston (see figure 12). The temperature is T. Find the equilibrium height of the piston by minimizing the free energy of the gas + piston. Does this result agree with what you would expect by considering mechanical equilibrium?

2. *Pressure of the ideal Fermi gas.*

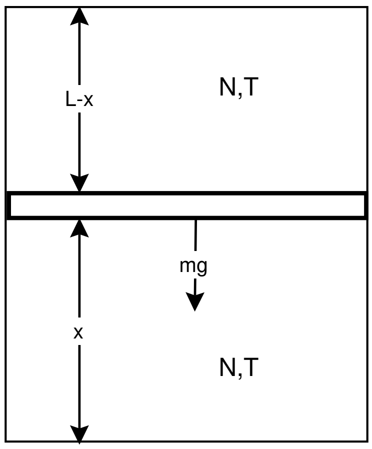


Figure 12:

Consider an ideal Fermi gas with N particles in a volume V. The energy of a one-particle state with wave vector $\mathbf{k}$ is $\frac{\hbar^2 k^2}{2m}$. There are two spin states for each orbital $\mathbf{k}$.

(a). Express the ground state energy in terms of N and V.

(b). What is the pressure of the Fermi gas at $T = 0$?

(c). What is the bulk modulus $-V\frac{\partial P}{\partial V}$?

3. *Cosmic background radiation*

Show that $TV^{1/3}$ is constant for adiabatic (isentropic) expansion of a volume V containing black body radiation in thermal equilibrium. Assume that the temperature of the cosmic black body radiation was decoupled from that of matter when both were at $3000K$ and that it is at $3K$ now. By what factor have distances expanded since then?

6 Kinetic theory

6.1 Drude model

In this section we will discuss *transport properties* associated with the conduction electrons, mainly the conductivity of a metal subject to a DC electric field or an alternating electromagnetic field. The latter extension of the theory will also will allow us to discuss the optical properties of a metal. We will employ a simple classical model called the *Drude model.* For simplicity we will assume that the metal is isotropic and neglect all band structure effects.

Let us first consider the *DC conductivity* of a metal containing n conduction electrons per unit volume. Each electron will have a velocity $\mathbf{v}$ which will be some fraction of the Fermi velocity, but in the absence of an electric field the electrons are equally likely to move in any direction, so the average velocity $\langle \mathbf{v} \rangle = 0$. However, if we apply an electric field $\vec{\mathcal{E}}$ the electrons will be accelerated in the field, and there will be an average nonzero *drift velocity* $\mathbf{v}_{drift}$. We let the charge of the electrons be $e = -|e|$ and the mass be m_e. From time to time the electrons will undergo collisions with impurities and vibrating atoms in the lattice and we assume that after each collision the electron will be equally likely to travel in any direction, so that immediately after each collision the average velocity is zero. Let $\mathbf{v(0)}$ be the velocity immediately after a collision. Some time t after that collision, but before the next collision the electron velocity will be

$$\mathbf{v}(t) = \mathbf{v}(0) + \frac{e\vec{\mathcal{E}}t}{m_e} \tag{135}$$

The average time between collisions, τ, is commonly called the *relaxation time*. We see from (135) that the drift velocity in our simple model will be

$$\mathbf{v}_{drift} = \langle \mathbf{v}(\tau) \rangle = \frac{e\vec{\mathcal{E}}\tau}{m_e} \tag{136}$$

The electric current density (current per unit area) is then

$$\mathbf{j} = ne\mathbf{v}_{drift} = \frac{ne^2\tau}{m_e}\vec{\mathcal{E}} \tag{137}$$

According to Ohm's law the electric field needed to produce a given current is given by

$$\vec{\mathcal{E}} = \rho \mathbf{j}$$

where ρ is the *resistivity*. Conversely we put

$$\mathbf{j} = \sigma_0 \vec{\mathcal{E}}$$

with σ_0 the *DC conductivity*. We thus find in the Drude model

$$\sigma_0 = \frac{ne^2\tau}{m_e} \tag{138}$$

It is instructive to construct a differential equation for the average velocity $\langle \mathbf{v}(t) \rangle$. If the average time between collisions is τ, the probability of a collision taking place in the time interval dt will be dt/τ. The mean speed will then increase by

$$d\langle \mathbf{v}(t) \rangle = \frac{e\vec{\mathcal{E}}dt}{m_e} \tag{139}$$

with probability $(1 - dt/\tau)$ or drop to zero with probability dt/τ. Neglecting terms of order dt^2 we find

$$\frac{d\langle \mathbf{v}(t) \rangle}{dt} = \frac{e\vec{\mathcal{E}}}{m_e} - \frac{\langle \mathbf{v}(t) \rangle}{\tau}$$

In the steady state $\langle \mathbf{v}(t) \rangle \to \mathbf{v}_{drift}$ and

$$\mathbf{v}_{drift} = \frac{e\vec{\mathcal{E}}\tau}{m_e}$$

and we get the same result as before.

We can also consider the response to a AC electric field

$$\vec{\mathcal{E}} = Re[\vec{\mathcal{E}}(\omega)e^{-i\omega t}]$$

Assuming a steady state response

$$\langle \mathbf{v}(t) \rangle = \mathbf{v}(\omega)e^{-i\omega t} \tag{140}$$

we find

$$-i\omega\mathbf{v}(\omega) = \frac{e\vec{\mathcal{E}}(\omega)}{m_e} - \frac{\mathbf{v}(\omega)}{\tau}$$

Putting $\mathbf{j}(t) = Re[\mathbf{j}(\omega)e^{-i\omega t}]$ we write

$$\mathbf{j}(\omega) = \sigma(\omega)\vec{\mathcal{E}}(\omega)$$

and we find after some algebra

$$\sigma(\omega) = \frac{\sigma_0}{1 - i\omega\tau} \tag{141}$$

We see that (141) exhibits two distinct régimes: a low frequency régime where $\omega\tau << 1$ in which the current is approximately in phase with the field, and a high frequency régime $\omega\tau >> 1$ in which the current will be approximately 90^o out of phase .

Another way of looking at the system is in terms of a complex dielectric constant.

$$\epsilon(\omega) = \epsilon_r(\omega)\epsilon_0$$

where ϵ_r is the relative dielectric constant and ϵ_0 is the permittivity of vacuum. We have

$$\vec{\mathcal{D}} = \epsilon_0\vec{\mathcal{E}} + \vec{\mathcal{P}}$$

where $\vec{\mathcal{D}}$ is the *electric displacement* and $\vec{\mathcal{P}}$ the *electric polarization* per unit volume, $\vec{\mathcal{P}} = ne\mathbf{u}$ and $\mathbf{u}$ is the displacement of an electron, caused by the applied field

$$\frac{d\mathbf{u}}{dt} = \mathbf{v}$$

In analogy with (140) we put $\mathbf{u} = \mathbf{u}(\omega)e^{-i\omega t}$ and find $-i\omega\mathbf{u}(\omega) = \mathbf{v}(\omega)$ and after a little algebra we find

$$\epsilon_r = 1 + \frac{i\sigma(\omega)}{\epsilon_0\omega} \tag{142}$$

We define the *Plasma frequency* as

$$\Omega_{pl} = \sqrt{\frac{ne^2}{\epsilon_0 m}} \tag{143}$$

The formula (142) for the relative dielectric constant can be rewritten as

$$\epsilon_r = 1 - \frac{\Omega_{pl}}{\epsilon_0\omega(\omega + i/\tau)} \tag{144}$$

You will learn from other courses that an electromagnetic wave with frequency ω propagates in a medium with dielectric constant ϵ_r with wave vector

$\mathbf{q} = \mathbf{q}_0\sqrt{\epsilon_r}$ where $q_0 = \omega/c$ is the wave vector in vacuum. If the dielectric constant is complex the wave will be damped and it is customary to write

$$\sqrt{\epsilon_r} = n + ik$$

where n and k are *optical constants*. The normal incidence reflectivity from vacuum can be shown to be given in terms of the optical constants as

$$R = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2} \tag{145}$$

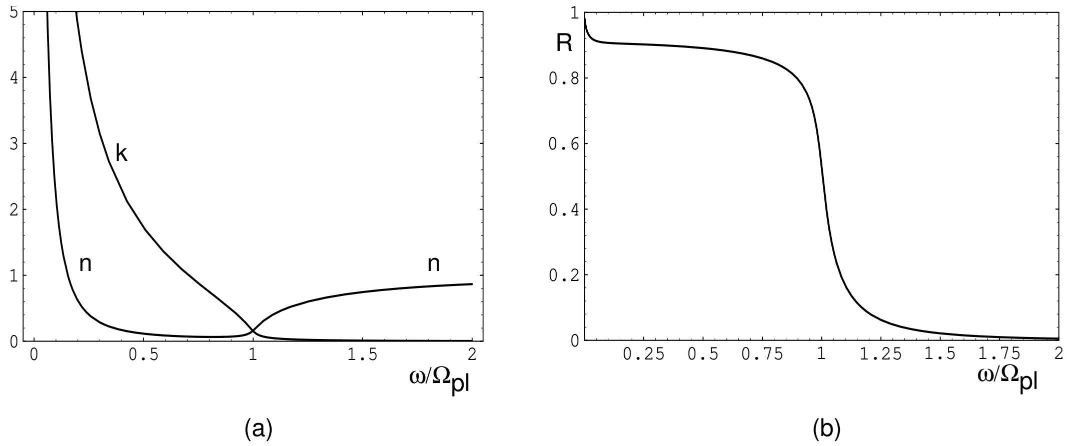


Figure 13: (a) Optical constants n and k and (b) normal incidence reflectivity in the Drude model. Frequencies are shown in units of the plasma frequency Ω_{pl} and the curves are drawn for the case $\Omega_{pl}\tau = 20$

In figure 13 we plot the optical constants and the reflectivity in the Drude model. For frequencies which are small compared to the plasma frequency the imaginary part of the index of refraction will be large and electromagnetic waves can only penetrate a short distance in a metal (skin effect). The intensity of the wave will decay according to

$$I \propto \mathcal{E}^2 \propto \exp\left(-\frac{2\omega k}{c} z\right)$$

where z is the penetration depth.

For frequencies higher than the plasma frequencies k will typically be very small and the metal will be transparent to electromagnetic radiation. Similarly, the reflectivity of a metal will be high for frequencies below the plasma frequency, and low for higher frequencies.

6.2 Problem set 4

1: *Fermi energy*

(a) Estimate the Fermi energy in eV of the conduction electrons in sodium metal. The atomic weight of sodium is $23g/mol$, the density of sodium is $0.950g/cm^3$ and Avogadro's number is $6.022 \times 10^{23}/mol$.

(b) ^3He atoms obey Fermi statistics. (Consider each ^3He atom to be single Fermi particle. At low temperatures ^3He forms a liquid with a volume $4.62 \times 10^{-29} m^3$ per helium atom. The mass of a ^3He atom is $5.01 \times 10^{-27} kg$ Calculate the Fermi temperature ($k_B = 1.38 \times 10^{-23} J/K$).

2: *Phase and group velocities.* Consider a metal for which $\omega_{pl}\tau >> 1$. The dielectric constant for frequencies larger than the plasma frequencies will then be approximately real, but less than one. The *phase velocity* for electromagnetic waves is ω/q, with the wave vector $q = 2\pi/\lambda$. The phase velocity for $\omega > \Omega_{pl}$ will then be faster than the speed of light. This will not violate special relativity since signals propagate with the *group velocity* given by $v_{group} = d\omega/dq$.

(a). Show that the group velocity of light is less that the speed of light for $\omega > \Omega_{pl}$ in the Drude model.

(b). One consequence of the fact that for $\omega > \Omega_{pl}$ the index of refraction $n < 1$ in the Drude model is that there will be *total reflection* of X-rays impinging with an angle of incidence larger than a critical angle given by $\sin(\theta_c) = n$. Calculate the critical angle of incidence for light with frequency $2\Omega_{pl}$.

3: Calculate the penetration depth at which the intensity of an electromagnetic wave in a sodium sample with relaxation time $\tau = 10^{-14} s$ will be reduced by a factor of 2 at the frequency of

(a) $\omega = 2\pi 10^7 Hz$.

(b) $\omega = 2\pi 10^{14} Hz$.

6.3 The Boltzmann Equation: Fields drift and collisions

In this section we discuss transport theory from the point of view of the Boltzmann equation. We assume, that the system of interest can be adequately described in terms of a single-particle distribution, $f_{\mathbf{p}}(\mathbf{r})$, which is the density of particles with momentum $\mathbf{p}$ at position $\mathbf{r}$. The distribution is normalized, so that

$$\int d^3k \int d^3r f_{\mathbf{k}}(\mathbf{r}) = N \tag{146}$$

where N is the number of particles.

In the case of a quantum system, a particle with momentum $\mathbf{p}$ is represented as having wave vector $\mathbf{k} = \mathbf{p}/\hbar$, energy $\epsilon_{\mathbf{k}}$. velocity $\mathbf{v_k}$[4]. Because of the uncertainty principle, we must assume that $f_{\mathbf{p}}(\mathbf{r})$ is *coarse grained* over a sufficiently large volume that $\mathbf{r}$ can be considered to be a *macroscopic* variable. In order to be able to do this we must assume that densities and fields are sufficiently slowly varying that they are well defined locally. We will take a *semiclassical* approach to quantum systems, that is, we assume that in the presence of macroscopic electric and magnetic fields the equation of motion for particles of charge e is

$$\dot{\mathbf{p}} = \hbar\dot{\mathbf{k}} = e\left(\vec{\mathcal{E}} + \mathbf{v_k} \times \mathbf{B}\right) \tag{147}$$

If the acceleration of the particles due to the fields were the only effect causing changes in the distribution function, we would have

$$f_{\mathbf{k}}(t + \delta t) = f_{\mathbf{k}-\delta\mathbf{k}}(t)$$

or

$$\left.\frac{\partial f_{\mathbf{k}}(\mathbf{r})}{\partial t}\right|_{field} = -e\left(\vec{\mathcal{E}} + \mathbf{v_k} \times \mathbf{B}\right) \cdot \frac{\partial f_{\mathbf{k}}}{\hbar\partial\mathbf{k}} \tag{148}$$

If we take the fields to be slowly varying, we can visualize the states as wave packets that are accelerated by the fields according to the classical equations of motion. This assumption is difficult to justify rigorously and we will not attempt to do so. In some cases, such as in inhomogeneous semiconductors, near surfaces, or in insulators subjected to intense fields, the electric fields are strong enough to cause *tunneling.* The semiclassical approach is then not appropriate.

Particles from time to time undergo *collisions* with obstacles in their path. Electrons in solids are scattered by impurities, vacancies, dislocations, and phonons. Because of screening, the interactions responsible for scattering are generally short-ranged. We will assume that the scattering events are discrete (quantum) events Let $W(\mathbf{k}, \mathbf{k}')$ be the transition rate (transition probability per unit time) from state $\mathbf{k}$ to state $\mathbf{k}'$. The distribution function then changes in time due to transitions *into* and *out* of state $\mathbf{k}$ and we may write

$$\left.\frac{\partial f_{\mathbf{k}}}{\partial t}\right|_{coll} = \sum_{\mathbf{k}'}[f_{\mathbf{k}'}(1 - f_{\mathbf{k}})W(\mathbf{k}', \mathbf{k}) - f_{\mathbf{k}}(1 - f_{\mathbf{k}'})W(\mathbf{k}, \mathbf{k}')]$$

[4]For plane waves the velocity of propagation is $\mathbf{p}/m$. In the case of electrons in a crystalline solid the relationship between momentum and velocity and momentum is more complicated and we should use the more general formula $\mathbf{v_k} = \partial\epsilon_{\mathbf{k}}/\partial\hbar\mathbf{k}$.

The calculation of the transition probability would take us deeply into quantum mechanics and solid state physics and we will not attempt this. In what follows we also limit ourselves to the simplest treatment of collisions and work within the *relaxation time* approximation.

We assume that the external fields produce only a small change in the distribution function and write

$$f_{\mathbf{k}} = f^0_{\mathbf{k}} + g_{\mathbf{k}} \tag{149}$$

where $f^0_{\mathbf{k}}$ is the equilibrium (zero-field) distribution given, as appropriate by the Boltzmann, Bose–Einstein, or, in our case, the Fermi–Dirac distributions. In the relaxation-time approximation one assumes that if the external fields were switched off, the non equilibrium part of the distribution function would decay exponentially with time:

$$g_{\mathbf{k}}(t) = g_{\mathbf{k}}(0)e^{-t/\tau}$$

where τ is the relaxation time. We thus obtain

$$\left.\frac{\partial f^0_{\mathbf{k}}}{\partial t}\right|_{coll} = 0 \qquad \frac{\partial g_{\mathbf{k}}}{\partial t} = -\frac{g_{\mathbf{k}}}{\tau} \tag{150}$$

If the distribution is inhomogeneous, it will change in time due to *drift.* If the particles are not subject to any forces,

$$f_{\mathbf{k}}(\mathbf{r}, t + \delta t) = f_{\mathbf{k}}(\mathbf{r} - \mathbf{v}_{\mathbf{k}}\delta t, t) \tag{151}$$

and hence

$$\left.\frac{\partial f_{\mathbf{k}}}{\partial t}\right|_{drift} = -\mathbf{v}_{\mathbf{k}}\cdot\frac{\partial f_{\mathbf{k}}}{\partial \mathbf{r}} \ . \tag{152}$$

Combining the various terms, we obtain the *Boltzmann equation* for the distribution function:

$$\frac{df_{\mathbf{k}}}{dt} = \left.\frac{\partial f_{\mathbf{k}}}{\partial t}\right|_{field} + \left.\frac{\partial f_{\mathbf{k}}}{\partial t}\right|_{coll} + \left.\frac{\partial f_{\mathbf{k}}}{\partial t}\right|_{drift} \tag{153}$$

In a steady-state situation we require $df_{\mathbf{k}}/dt = 0$.

6.4 DC conductivity of a metal revisited

Consider a metal subject to a weak electric field but no magnetic field. The field term is

$$-\frac{e\vec{\mathcal{E}}}{\hbar}\cdot\frac{\partial(f^0_{\mathbf{k}}+g_{\mathbf{k}})}{\partial\mathbf{k}} \tag{154}$$

In the case of a weak $\vec{\mathcal{E}}$, the non equilibrium part, $g_{\mathbf{k}}$, will be proportional to $\vec{\mathcal{E}}$ and thus be small compared to $f^0_{\mathbf{k}}$. Thus, for a weak field, the non-equilibrium part of the distribution can be neglected in (154). Since the equilibrium part of f does not contribute to the collision term the non equilibrium part must be taken into account in the collision term. We obtain the *linearized* Boltzmann equation

$$-\frac{e\vec{\mathcal{E}}}{\hbar}\cdot\frac{\partial f^0_{\mathbf{k}}}{\partial\mathbf{k}}-\frac{g_{\mathbf{k}}}{\tau}=0$$

or

$$g_{\mathbf{k}}=-\frac{e\vec{\mathcal{E}}\tau}{\hbar}\frac{\partial f^0_{\mathbf{k}}}{\partial\epsilon_{\mathbf{k}}}\cdot\frac{\partial\epsilon_{\mathbf{k}}}{\partial\mathbf{k}}=-\tau e\vec{\mathcal{E}}\cdot\mathbf{v}_{\mathbf{k}}\frac{\partial f^0_{\mathbf{k}}}{\partial\epsilon_{\mathbf{k}}}$$

(Remember our sign convention according to which the charge of an electron to is e.) The electrical current density is given by

$$\mathbf{j}=\frac{2}{V}\sum_{\mathbf{k}}e\mathbf{v}_{\mathbf{k}}g_{\mathbf{k}} \tag{155}$$

This in turn yields the conductivity

$$\sigma=-\frac{2e^2\tau}{V}\sum_{\mathbf{k}}\mathbf{v}_{\mathbf{k}}:\mathbf{v}_{\mathbf{k}}\frac{\partial f^0_{\mathbf{k}}}{\partial\epsilon_{\mathbf{k}}} \tag{156}$$

We here use the *dyadic* notation, i.e. $\mathbf{v}_{\mathbf{k}}:\mathbf{v}_{\mathbf{k}}$ is a tensor. This formula illustrates that the current for an anisotropic system will not necessarily flow in the direction of the field, the conductivity is a tensor. In component form

$$j_i=\sum_m\sigma_{im}E_m$$

and we have, for example,

$$\sigma_{xx}=-\frac{2e^2\tau}{V}\sum_{\mathbf{k}}(v_k)^2_x\frac{\partial f^0_{\mathbf{k}}}{\partial\epsilon_{\mathbf{k}}}$$

For a metal at moderate temperatures $\frac{\partial f^0_{\mathbf{k}}}{\partial \epsilon_{\mathbf{k}}}$ will be sharply peaked at the Fermi energy. Only electrons close to the Fermi surface thus contribute to the conductivity.

It is useful to interpret the result

$$f_{\mathbf{k}} = f^0_{\mathbf{k}} + g_{\mathbf{k}} = f^0_{\mathbf{k}} - \tau e(\mathcal{E} \cdot \mathbf{v}_{\mathbf{k}}) \frac{\partial f^0_{\mathbf{k}}}{\partial \epsilon_{\mathbf{k}}} \tag{157}$$

in terms of the drift velocity. To first order in $\mathcal{E}$, we can rewrite (157) in the form

$$f_{\mathbf{k}} = f^0_{\mathbf{k}}\{\epsilon_{\mathbf{k}} - \tau e(\vec{\mathcal{E}} \cdot \mathbf{v}_{\mathbf{k}})\} \ .$$

The right-hand side of this equation is simply the equilibrium distribution of the system with all energies shifted by an amount

$$\delta \epsilon_{\mathbf{k}} = \tau e(\vec{\mathcal{E}} \cdot v_k)$$

i.e., by precisely the amount expected classically for particles moving with constant velocity $\mathbf{v}$ for a time τ in a force field $e\vec{\mathcal{E}}$. The extra energy gained in this way can be interpreted in terms of a drift velocity $\delta \mathbf{v}_{\mathbf{k}}$ in the direction of the field so that

$$\delta \mathbf{v}_k \cdot \frac{\partial \epsilon_{\mathbf{k}}}{\partial \mathbf{v}_{\mathbf{k}}} = e\tau(\mathbf{v}_{\mathbf{k}} \cdot \vec{\mathcal{E}})$$

If

$$\frac{\partial \epsilon_{\mathbf{k}}}{\partial \mathbf{v}_{\mathbf{k}}} = \mathbf{p}_{\mathbf{k}} = m\mathbf{v}_{\mathbf{k}}$$

we obtain

$$\delta \mathbf{v}_{\mathbf{k}} = \frac{e\tau}{m} \vec{\mathcal{E}}$$

For n particles per unit volume, we have for the current,

$$\mathbf{j} = ne\delta\mathbf{v}$$

and we recover our previous result for the DC conductivity.

In the case of a metal the drift velocities are typically very small compared to the Fermi velocity v_F, mainly because the electric fields inside a metal tend to be small. In a semiconductor one sometimes deals with fields which are large enough that non ohmic effects are important. It is then not adequate to linearize the Boltzmann equation and one must consider the nonlinear problem. In such situations collisions often occur so frequently that one cannot describe them as independent events, and the whole Boltzmann approach becomes suspect.

6.5 Thermal conductivity and thermoelectric effects

To give an example of the use of the Boltzmann equation when the system is not spatially uniform, and the drift term comes into play, we discuss the case of a time-independent temperature gradient maintained across a metallic sample. In analogy with Ohm's law, there will be a heat current whenever there is a temperature gradient.

We also allow for an electric field $\vec{\mathcal{E}}$, with corresponding scalar potential $\phi(\mathbf{r})$. Let u, $đq$, n be, respectively, the internal energy, heat supplied and electron density per unit volume. From thermodynamics we have

$$du = đq + (\mu + e\phi)dn = đq + \mu' dn$$

where $\mu' = \mu + e\phi(\mathbf{r})$ is the *electrochemical potential* at point $\mathbf{r}$. Hence the heat current density is given by

$$\mathbf{j}_Q = \mathbf{j}_U - \mu' \mathbf{j}_N \ . \tag{158}$$

where $\mathbf{j}_U$ is the energy current density and $\mathbf{j}_N$ is the particle current density.

We assume that the heat current is due entirely to the motion of electrons and neglect the lattice thermal conductivity. In the presence of the electrostatic potential $\phi(\mathbf{r})$ the electronic energies will be locally shifted by an amount $e\phi(\mathbf{r})$ so that the *energy current* density is then given by

$$\mathbf{j}_U(\mathbf{r}) = \frac{2}{(2\pi)^3} \int d^3k \, [\epsilon_{\mathbf{k}} + e\phi(\mathbf{r})] \, \mathbf{v}_{\mathbf{k}} f_{\mathbf{k}}(\mathbf{r})$$

and the particle current density is

$$\mathbf{j}_N(\mathbf{r}) = \frac{2}{(2\pi)^3} \int d^3k \, \mathbf{v}_{\mathbf{k}} f_{\mathbf{k}}(\mathbf{r}) \tag{159}$$

which, of course, also implies an electrical current density $\mathbf{j}_C = e\mathbf{j}_N$. Thus the heat current (158) is given by

$$\mathbf{j}_Q(\mathbf{r}) = \frac{2}{(2\pi)^3} \int d^3k [\epsilon_{\mathbf{k}} - \mu] \mathbf{v}_{\mathbf{k}} f_{\mathbf{k}}(\mathbf{r}) \ . \tag{160}$$

As before we write

$$g_{\mathbf{k}}(\mathbf{r}) = f_{\mathbf{k}}(\mathbf{r}) - f^0_{\mathbf{k}}$$

and, in addition, assume that the thermal gradient is small enough that it is meaningful to talk about a local temperature and electro- chemical potential. With these assumptions, the Boltzmann equation becomes

$$-\mathbf{v}_{\mathbf{k}} \cdot \frac{\partial f_{\mathbf{k}}}{\partial \mathbf{r}} - \frac{e}{\hbar} \vec{\mathcal{E}} \cdot \frac{\partial f_{\mathbf{k}}}{\partial \mathbf{k}} + \left. \frac{\partial f_{\mathbf{k}}}{\partial t} \right|_{coll} = 0 \ .$$

We take $f^0_{\mathbf{k}}(\mathbf{r})$ to be the equilibrium distribution with the local temperature $T(\mathbf{r})$ and the local electro-chemical potential $\mu'(\mathbf{r})$ controlling the density at point $\mathbf{r}$. Noting that $\epsilon_{\mathbf{k}} + e\phi(\mathbf{r}) - \mu'(\mathbf{r}) = \epsilon_{\mathbf{k}} - \mu(\mathbf{r})$, we have

$$f^0_{\mathbf{k}} = f^0\{\epsilon_{\mathbf{k}}, \mu(\mathbf{r}), T(\mathbf{r})\} = \left[\exp\left\{\frac{\epsilon_{\mathbf{k}} - \mu(\mathbf{r})}{k_B T(\mathbf{r})}\right\} + 1\right]^{-1}$$

and hence

$$\frac{\partial f^0_{\mathbf{k}}}{\partial \mathbf{r}} = \frac{\partial f^0_{\mathbf{k}}}{\partial T}\nabla T + \frac{\partial f^0_{\mathbf{k}}}{\partial \mu}\nabla\mu$$

We next make the relaxation-time approximation (151) and in the spirit of the linearized Boltzmann equation, neglect terms such as

$$\frac{\partial g}{\partial \mathbf{r}} \quad \text{and} \quad \frac{e}{\hbar}\vec{\mathcal{E}}\cdot\frac{\partial g}{\partial \mathbf{k}}\ .$$

Using

$$\frac{\partial f^0_{\mathbf{k}}}{\partial T} = -\frac{\epsilon_{\mathbf{k}} - \mu}{k_B T^2}\frac{\partial f^0_{\mathbf{k}}}{\partial \epsilon_{\mathbf{k}}}; \qquad \frac{\partial f^0_{\mathbf{k}}}{\partial \mu} = -\frac{\partial f^0_{\mathbf{k}}}{\partial \epsilon_{\mathbf{k}}}$$

and collecting terms, we obtain

$$\frac{1}{\tau}g_{\mathbf{k}} = -\frac{\partial f^0_{\mathbf{k}}}{\partial \epsilon_{\mathbf{k}}}\mathbf{v}_{\mathbf{k}}\cdot\left[-\frac{\epsilon_{\mathbf{k}} - \mu}{k_B T^2}\nabla T + (e\vec{\mathcal{E}} - \nabla\mu)\right] \tag{161}$$

The potential difference measured by, say, a voltmeter will be given by a line integral of a field along the path of the current, However, it is not given by

$$\int \vec{\mathcal{E}}\cdot d\mathbf{l}$$

but rather by the quantity

$$\Psi = \int\left(\vec{\mathcal{E}} - \frac{1}{e}\nabla\mu\right)\cdot d\mathbf{l}$$

We therefore introduce the "electromotive field" or "observed" field

$$\mathbf{E} = \vec{\mathcal{E}} - \frac{1}{e}\nabla\mu = -\frac{1}{e}\nabla\mu' \tag{162}$$

Clearly, $\mathbf{E}$ is of more interest than the electric field $\vec{\mathcal{E}}$ itself. We now define the kinetic coefficients $\mathcal{L}_{CC}, \mathcal{L}_{CQ}, \mathcal{L}_{QC}$, and $\mathcal{L}_{QQ}$ through [5]

$$\begin{aligned} \mathbf{j}_C &= \mathcal{L}_{CC}\mathbf{E} - \mathcal{L}_{CQ}\nabla T \\ \mathbf{j}_Q &= \mathcal{L}_{QC}\mathbf{E} - \mathcal{L}_{QQ}\nabla T \end{aligned} \tag{163}$$

Using (159) and (160), we see that the kinetic coefficients can all be expressed in terms of the integral

$$I_\alpha = \int d\epsilon \left(-\frac{\partial f^0}{\partial \epsilon}\right)(\epsilon - \mu)^\alpha \sigma(\epsilon) \tag{164}$$

where

$$\sigma(\epsilon) = e^2\tau \int \frac{d^3k}{4\pi^3}\delta(\epsilon - \epsilon_{\mathbf{k}})\mathbf{v}_{\mathbf{k}} : \mathbf{v}_{\mathbf{k}} \tag{165}$$

is the generalized energy-dependent form of the conductivity tensor (156). We shall evaluate I_α for conditions appropriate to a metal and in this case

$$-\frac{\partial f^0}{\partial \epsilon} = \frac{\beta \exp\{\beta(\epsilon - \mu)\}}{[\exp\{\beta(\epsilon - \mu)\} + 1]^2} \tag{166}$$

can be taken to be nonzero only in a narrow energy range of order k_BT around ϵ_F. We introduce the new variable $z = \beta(\epsilon - \mu)$ and expand

$$\sigma(k_BTz + \mu) = \sigma(\mu) + k_BTz\frac{\partial \sigma}{\partial \mu} + \dots \ .$$

Substituting in (165), we then have the transport coefficients expressed in terms of

$$I_\alpha \approx (k_BT)^\alpha \int_{-\infty}^{\infty} dz \frac{z^\alpha e^z}{(1+e^z)^2}\left[\sigma(\mu) + k_BT\frac{\partial \sigma}{\partial \mu}\right] \ . \tag{167}$$

Defining

$$Q_j = \int_{-\infty}^{\infty} dz \frac{z^j}{(e^z+1)(e^{-z}+1)}$$

we have $Q_0 = 1$, $Q_1 = 0$, $Q_2 = \pi^2/3$, and $Q_3 = 0$. Taking $\mu \approx \epsilon_F$, we thus obtain

$$\mathcal{L}_{CC} = \sigma(\epsilon_F) = \sigma \tag{168}$$

[5]The kinetic is sometimes defined using $\nabla(1/T)$ rather than ∇T as a driving term. The resulting coefficients will differ by factors of T^2.

$$\mathcal{L}_{CQ} = T\mathcal{L}_{QC} = \frac{\pi^2}{3e} k_B^2 T^3 \left. \frac{\partial \sigma(\epsilon)}{\partial \epsilon} \right|_{\epsilon=\epsilon_F} \tag{169}$$

$$\mathcal{L}_{QQ} = \frac{\pi^2}{3e^2} k_B^2 T^3 \sigma \ . \tag{170}$$

We see that $_{CQ} = T\mathcal{L}_{QC}$. This result turns out to have a much more general validity than our derivation here suggests and is known as an Onsager reciprocity relation. We note also that for electrons ($e = -|e|$), L_{QC} and L_{CQ} are negative. If these coefficients are found experimentally to be positive, it is an indication that the charge carriers are not electrons but holes.

To obtain the thermal conductivity we require that there be no electric current, or from (163),

$$\mathcal{E} = -L_{CC}^{-1} L_{CQ} \nabla \left(\frac{1}{T} \right) \ . \tag{171}$$

Substituting into the second equation of (163), we have for the thermal conductivity

$$\kappa = \frac{L_{QQ} - L_{QC} L_{CC}^{-1} L_{CQ}}{T^2} \ . \tag{172}$$

We now argue that in a metal the second term in (172) is small compared to the first. We first note that the second law of thermodynamics implies that κ is positive or $L_{CC}L_{QQ} > L_{CQ}L_{QC}$. To obtain an order-of-magnitude estimate, we make the approximation (on dimensional grounds) $\partial\sigma/\partial\epsilon \sim \sigma/\epsilon$ (for free electrons $\partial\sigma/\partial\epsilon = 3\sigma/2\epsilon$) and thus have

$$\frac{L_{CQ}L_{QC}}{L_{CC}L_{QQ}} \approx \frac{\pi^2}{3} \left(\frac{k_B T}{\epsilon_F} \right)^2 \sim 10^{-4}$$

for a typical metal at room temperature. Neglecting the second term in (172), we therefore find

$$\kappa = \frac{L_{QQ}}{T^2} = \frac{\pi^2}{3e^2} k_B^2 T \sigma \ . \tag{173}$$

This result is known as the Wiedemann–Franz law, and is in general agreement with experiment for a number of metals.

The coupled transport equations (163) suggest a number of thermoelectric effects. Consider a circuit consisting of two wires made of different material with different transport coefficients. The two wires are joined at one end and the junction is kept at temperature T_1 at the other end the temperature is T_2 if the circuit is open at the other end there will be a voltage difference between the two ends at temperature T_2 which can be picked up by a voltmeter (Seebeck effect). Similarly, if one forms a closed circuit of the two wires and forces a current through by applying a voltage, one junction will heat and the other cool depending on the direction of the current (Peltier effect)[13]

6.6 Problem set 5

1:
(a) Show that the energy-dependent conductivity tensor

$$\sigma(\epsilon) = e^2\tau \int \frac{d^3k}{4\pi^3}\delta(\epsilon - \epsilon_{\mathbf{k}})\mathbf{v_k} : \mathbf{v_k}$$

is diagonal (i.e $\sigma_{xy} = \sigma_{xz} = \sigma_{yz} = 0$) for the free Fermi gas $\epsilon_k = \frac{\hbar k^2}{2m}$.
(b) Determine the diagonal components $\sigma_{xx} = \sigma_{yy} = \sigma_{zz}$ if $\epsilon_k = \frac{\hbar k^2}{2m}$.
(c) Use the result under (b) to find the transport coefficients of the Fermi gas

$$L_{CC} = \sigma(\epsilon_F) = \sigma$$

$$L_{CQ} = TL_{QC} = \frac{\pi^2}{3e}k_B^2T^2\left.\frac{\partial\sigma(\epsilon)}{\partial\epsilon}\right|_{\epsilon=\epsilon_F}$$

$$L_{QQ} = \frac{\pi^2}{3e^2}k_B^2T^3\sigma\ .$$

2: Consider an *anisotropic* material for which the relationship between energy is

$$\epsilon_k = \frac{\hbar^2k_x^2}{2m_x} + \frac{\hbar^2k_y^2}{2m_y} + \frac{\hbar^2k_z^2}{2m_z}$$

where m_x, m_y and m_z are constants.
(a) The Fermi surface will now be an ellipsoid with axes $(k_F)_x, (k_F)_y$ and $(k_F)_c$. Find $(k_F)_x, (k_F)_y$ and $(k_F)_c$. [hint: it is a good idea to rescale the three components of $\mathbf{k}$ so that the equal energy surface becomes a sphere]
(b) Using the expression

$$\mathbf{v_k} = \frac{\partial\epsilon_k}{\partial\hbar\mathbf{k}}$$

show that the conductivity tensor will still be diagonal in the x, y, z coordinate system, but the diagonal components will be different in the three directions if m_x, m_y and m_z are different.

c The result under (b) implies that if the electric field is at an angle to the three *principal directions* {x,y,z}, the current will not be in the direction of the field. Plot the angle ψ of the current with respect to the $z-$axis against the angle ϕ of the field with respect to the $z-$axis if $\sigma_{xx} = \sigma_{yy} = 5\sigma_{zz}$.

7 Phonons

7.1 The harmonic chain

We leave the electrons for a while and concentrate instead on the thermal properties of a solid due to lattice vibrations. We will not have time to discuss the solid state aspects in any detail and we will make use of a very simple model due to Debye which captures many of the essential aspects. To motivate this model we first consider a one dimensional caricature of a solid, the harmonic chain. Consider a long chain consisting of N masses m connected by springs with spring constant K (figure 14). Let a be the equilibrium spacing between the masses, so that $x = na$ is the equilibrium position of the n'th mass. The actual position of the n'th mass is

$$r_n = na + u(n)$$

and the potential energy of the spring connecting the n'th and the $(n+1)$'st mass is $\frac{K}{2}[u(n+1) - u(n)]^2$. We assume *periodic boundary conditions,* i.e. we let the N'th mass be connected with the first mass. Formally, we do this by requiring that $u(N+1) = u(1)$ and in general

$$u(N+n) = u(n) \tag{174}$$

The force on the n'th mass is $K[u(n+1) + u(n-1) - 2u(n)]$ and Newton's second law gives rise to a coupled set of differential equations

$$m\ddot{u}(n) = K[u(n+1) + u(n-1) - 2u(n)] \tag{175}$$

We can find solutions to (175) on the form

$$u(n) = Ae^{i(nka-\omega t)} \tag{176}$$

where k is the wave vector . Substitution of (176) into (175) yields

$$-m\omega^2 Ae^{i(nka-\omega t)} = AK\left\{e^{i[(n+1)ka-\omega t]} + e^{i[(n-1)ka-\omega t]} - 2e^{i[nka-\omega t]}\right\}$$

or

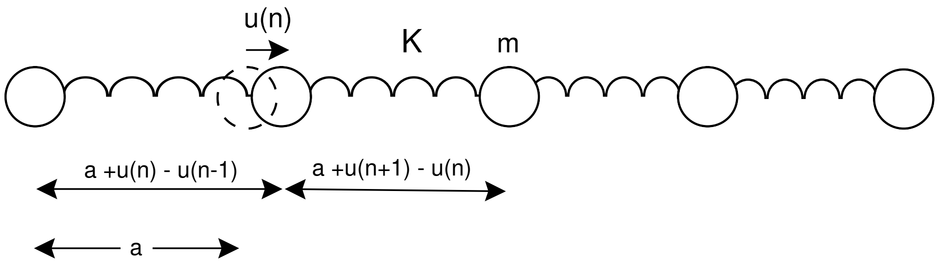


Figure 14: The harmonic chain

$$\omega^2 = \frac{2K}{m}[1 - \cos(ka)]$$

We use the trigonometric identity

$$2\sin^2(\frac{ka}{2}) = [1 - \cos(ka)] \tag{177}$$

By convention the frequency ω is positive and we find

$$\omega = 2\sqrt{\frac{K}{m}}\left|\sin\frac{ka}{2}\right| \tag{178}$$

The periodic boundary condition (174) determines which values of k are allowed. Substitution of (176) into (174) gives

$$e^{ikNa} = 1$$

or $k = \frac{2\pi l}{Na}$ with l an *integer.* From (176) we see that $l' = l + N$ and l give rise to identical solutions, and we can without loss of generality restrict l to the values

$$l = -\frac{N}{2}, \frac{N}{2} - 1, \cdots 1, 2, .. \frac{N}{2} - 1 \tag{179}$$

From (179) we note that there are N distinct values of l. This result is compatible with the fact that the system has N degrees of freedom. The restriction on the wave vector is

$$-\frac{\pi}{a} \leq k < \frac{\pi}{a} \tag{180}$$

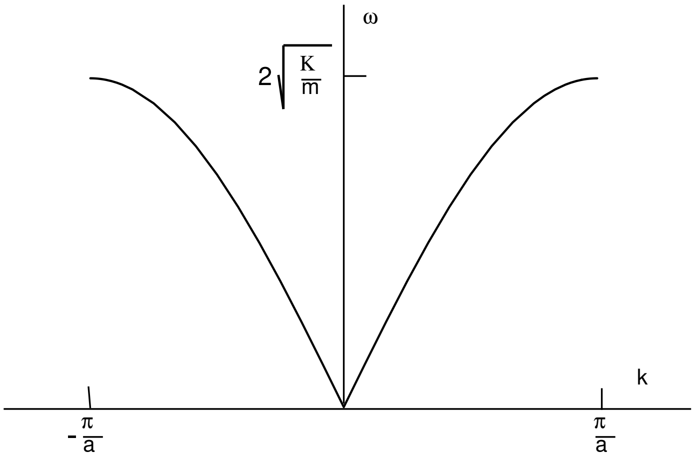


Figure 15: The harmonic chain

The relationship (178) between frequency and wave vector is plotted in figure (15) The physical displacement is, of course, real. Since the differential equation (175) is linear and real, both the real and imaginary part of a complex solution are also solutions. With these considerations the general solution to (175) is then with A, B, C_l and S_l arbitrary constants of integration

$$u(n) = \sum_{l=-\frac{N}{2}, l\neq 0}^{\frac{N}{2}-1} \left[C_l \cos(\frac{2\pi l n}{N} - \omega_l t) + S_l \sin(\frac{2\pi l n}{N} - \omega_l t) \right] + A + Vt \quad (181)$$

The case $l = 0$ in (181 requires special attention. If $l = 0$ the wave vector $k = 0$. This means that all displacements u are the same. The differential equation (175) now becomes

$$m\ddot{u}(n) = 0$$

with solution $u = A + Vt$ where A and V are constants. Usually the situation where the chain moves uniformly (as a whole) is not of much interest in lattice vibrational problems.

It is instructive to solve the harmonic chain problem with various boundary conditions. First, consider the situation where we "pluck" the chain by displacing the atom at the origin by a fixed amount δ.

The initial conditions are then

$$u(0, t = 0) = \delta; \;\; u(n, t = 0) = 0 \;\; for \;\; n \neq 0 \tag{182}$$

$$\dot{u}(n, t = 0) = 0 \; for \; all \; n \tag{183}$$

It is convenient to use the complex form for the normal mode expansion

$$u(n) = \sum_{l=-N/2, l\neq 0}^{N/2-1} A_l e^{i(2\pi nl/N - \omega t)} + A + Vt$$

giving

$$u(n, t = 0) = \sum_{l=-N/2, l\neq 0}^{N/2-1} A_l e^{i(2\pi nl/N)} + A \tag{184}$$

$$0 = \dot{u}(n, t = 0) = \sum_{l=-N/2, l\neq 0}^{N/2-1} -i\omega A_l e^{i(2\pi nl/N)} + V \tag{185}$$

The fundamental formula for inverting discrete Fourier series is

$$\sum_{n=0}^{N-1} e^{2\pi i(l-l')n/N} = \begin{cases} \frac{1-e^{2\pi i(l-l')}}{1-e^{2\pi i(l-l')/N}} = 0; & for \;\; l \neq l' \\ N; & for \;\; l = l' \end{cases} \tag{186}$$

Equating (182) and (184) and multiplying by $e^{i\pi l'n/N}$ and summing over n gives

$$A = A_l = \frac{\delta}{N}$$

i.e. all modes are excited with equal amplitude. The actual motion is the real part of the solution giving

$$u(n) = \frac{\delta}{N}\left(\sum_{l=-N/2, l\neq 0}^{N/2-1} \cos(2\pi nl/N - \omega t) + 1\right) \tag{187}$$

Next consider the case where the atom at the origin is given a "kick" so that it has velocity v (in units of lattice spacings per unit time) starting at the origin with all the other atoms at rest. The initial condition is now

$$\dot{u}(0, t = 0) = v; \;\; u(n, t = 0) = 0 \;\; forn \neq 0 \tag{188}$$

$$u(n, t = 0) = 0 \; for \; all \; n \tag{189}$$

We find equating (185) and (188) using the Fourier transform formula (186)

$$V = \frac{v}{N}$$

and

$$A_l = \frac{iv}{\omega N}$$

The actual motion of the masses is obtained by taking the real part or

$$u(n) = \frac{v}{N}\left(-\sum_{l=-N/2, l\neq 0}^{N/2-1} \frac{1}{\omega}\sin(2\pi nl/N - \omega t) + t\right) \tag{190}$$

Returning to the general case: we refer to the individual terms in the [ ] of 181 as *normal modes* and to (180) as a normal mode expansion.

From (178) we note that there is a maximum frequency

$$\omega_{max} = 2\sqrt{\frac{K}{m}}$$

If an atom is subject to a periodic force with frequency $\omega < \omega_{max}$ this will set up traveling waves with wave vectors satisfying (178). The *phase velocity* of this wave is

$$v_p = \frac{\omega}{k} = \frac{2}{k}\sqrt{\frac{K}{m}}\left|\sin(\frac{ka}{2})\right|$$

while the group velocity is

$$v_g = \frac{\partial\omega}{\partial k} = a\sqrt{\frac{K}{m}}\cos(ka)$$

Note that the group velocity $\to 0$ as $\omega \to \omega_{max}$. A local disturbance with frequency $> \omega_{max}$ cannot propagate through the lattice, but stays trapped. In the long wave length limit $k \to 0$, $\lambda = 2\pi/k \to \infty$ the phase and group velocities approach a common limit

$$s = a\sqrt{\frac{K}{m}}$$

We refer to long wave length lattice vibrations as *sound waves* with s the speed of sound.

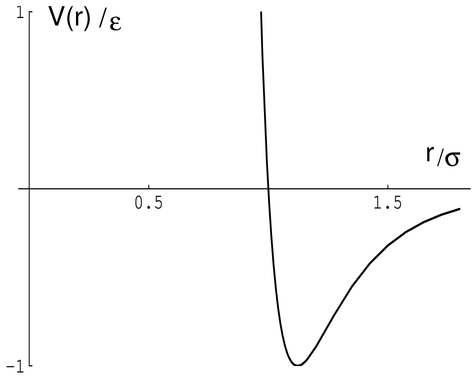


Figure 16: Lennard-Jones potential.

7.2 Lattice vibrations in a solid

Let us now try to form a more realistic picture of the lattice vibrations in a solid. To a reasonable approximation, the interaction energy between two atoms, one at $\mathbf{r}_l$ one at $\mathbf{r}_m$, can be described by a *pair potential* $v(\mathbf{r}_l - \mathbf{r}_m)$. This potential can be thought of as made up of two parts. At short range when the atomic cores overlap there is a *strong repulsion*, whose origin lies in the *Pauli exclusion principle*. Since two electrons cannot occupy the same state, the core states will have to be distorted when two atoms get too close together, and there is a large energy cost associated with such distortions.

At larger distances there is no significant overlap between the electron states of neighboring atoms. The interaction would then tend to be attractive (otherwise the solid would not stick together). The simplest cases are the crystals formed from the inert gases neon (Ne), argon (Ar), krypton (Kr) and xenon (Xe). The dominant force is then due to the *van der Waals attraction*. In this case the atoms have spherical symmetry and the potential is *central* $v = v(|\mathbf{r}_l - \mathbf{r}_m|)$ The corresponding potential falls off as the sixth power of the distance– and since force is minus the gradient of the potential, the corresponding force will be $\sim 1/R^7$. This r dependence can be understood from the following argument:

If an atom is exposed to an electric field $\mathcal{E}$ the force on the negatively charged electrons and the positively charged nucleus will be in opposite directions. This will induce a *dipole moment*

$$\mathbf{p}_l = \alpha\vec{\mathcal{E}} \tag{191}$$

where α is the *polarizability*. If the electric field is absent the average dipole moment will be zero, but there will still be a fluctuating dipole moment due to quantum effects (zero point motion). The electric field due to a dipole falls off as the inverse third power of the distance, and it will induce a dipole moment in neighboring atoms which is proportional to the field according to (191). The *interaction energy* will then be of the form

$$\langle -\mathbf{p}_l \cdot \vec{\mathcal{E}}_m \rangle = \alpha \langle \mathcal{E}_m^2 \rangle$$

where $\mathbf{p}_l$ is the induced dipole moment on atom l due to the fluctuating field $\vec{\mathcal{E}}_m$ from atom m. The angular brackets $\langle\rangle$ indicate *average value.* The average value of the fluctuating electric field is $\langle \mathcal{E} \rangle = 0$, while the mean square value will be non-zero. Since the fluctuating field falls off as the inverse third power of the distance, the mean square field will be proportional to the inverse sixth power of the distance to the fluctuating atom. The van der Waals interaction energy is thus $\sim 1/r^6$.

An approximate way of combining the repulsive and attractive interactions between the atoms in a molecular crystal is the *Lennard-Jones potential*

$$v(r) = \frac{A}{r^{12}} - \frac{B}{r^6} = 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6] \tag{192}$$

where A and B are constants which depend on which atom or molecule is involved. It is conventional to parameterize the potential in terms of an energy parameter ϵ and length parameter σ, in terms of which A= $4\epsilon\sigma^{12}$, B= $4\epsilon\sigma^6$. We plot the Lennard -Jones potential in figure 16.

The potential changes sign at $r = \sigma$. By differentiating (192) we find that the potential has minimum at $r = 2^{1/6}\sigma$. The value of the potential at the minimum is $-\epsilon$.

In other solids the mechanism for the attractive part of the interactions will be different. In *ionic solids* the crystal is made up of positively and negatively charged ions (e.g. common salt is Na^+ Cl^-). The crystal structure will be such that the nearest neighbors have opposite charge and thus attract each other). In a metal all the ions are positively charged but each ion is surrounded by a negatively charged *screening cloud.* Since this cloud will be closer to the neighboring ions than the ions themselves there will be a net attraction. We have already mentioned that conduction electrons, because of their light mass move much faster than the ions. On the time scale of the ion movement the electrons will have time to adjust to that motion and the screening cloud can be thought of as effectively attached to its ion.

Consider now a chain of atoms interacting with a potential similar to that of figure 16. When at rest the atoms in the chain will be spaced at a distance a which corresponds to the minimum of the potential. For small amplitude oscillations ($r = a + u$, where u is the displacement) the potential will be approximately given by

$$v(r) \approx v(a) + \frac{u^2}{2}\frac{d^2v(r)}{dr^2}\bigg|_{r=a+u}$$

The "spring constant" for the atomic chain is then just the second derivative of the interatomic potential evaluated at the equilibrium distance between the atoms.

We now wish to establish a more general formalism based on the above ideas. Let $\mathbf{R}_l$ be the equilibrium position of the l'th ion, $\mathbf{u}_l$ the displacement of this ion from the rest position. The actual position of the ion is then

$$\mathbf{r}_l = \mathbf{R}_l + \mathbf{u}_l$$

We write $u_{l\alpha}$ ($\alpha = x, y, z$) for the components of the displacement. The components of the force on the l'th ion is

$$f_{l\alpha} = -\frac{\partial V(\mathbf{r}_1, \mathbf{r}_2, ..\mathbf{r}_l, ..\mathbf{r}_N)}{\partial r_{l\alpha}}$$

or in the more compact vector notation

$$\mathbf{f}_l = -\frac{\partial V}{\partial \mathbf{r}_l} \tag{193}$$

For small amplitude vibrations we can make a Taylor series expansion of the potential energy function

$$V(\mathbf{r}_1, \mathbf{r}_2, ..\mathbf{r}_N) = V(\mathbf{R}_1, \mathbf{R}_2, ..\mathbf{R}_N) + \sum_{l\alpha} u_{l\alpha} \frac{\partial V(\mathbf{r}_1, \mathbf{r}_2, ..\mathbf{r}_N)}{\partial r_{l\alpha}}\bigg|_{r_{l\alpha}=R_{l\alpha}\ for\ all\ l} +$$

$$+\frac{1}{2}\sum_{lm\alpha\beta} u_{l\alpha}u_{m\beta} \frac{\partial^2 V(\mathbf{r}_1, \mathbf{r}_2, ..\mathbf{r}_N)}{\partial r_{l\alpha}\partial r_{m\beta}}\bigg|_{r_{l\alpha}=R_{l\alpha}\ for\ all\ l} + \cdots \tag{194}$$

When all the ions are at their equilibrium positions, the net force (193) on them is zero. The second term on the right hand side of (194) must therefore vanish.

The *harmonic approximation* consists of stopping at the last term on the right hand side of (194). This allows us to think of the lattice as made up a set of masses connected by springs. Each spring is represented by a spring constant

$$\Phi_{lm\alpha\beta} = \left.\frac{\partial^2 V(\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_N)}{\partial r_{l\alpha} \partial r_{m\beta}}\right|_{r_{l\alpha}=R_{l\alpha}} \quad \text{for all } l \tag{195}$$

The equation of motion for the α'component of the displacement of the l'th atom is then according to Newton's second law

$$m\ddot{u}_{l\alpha} = f_{l\alpha} = -\sum_{m\beta} \Phi_{lm\alpha\beta} u_{m\beta} \tag{196}$$

If the crystal potential and structure is known, e.g. if the atoms interact pairwise via a Lennard–Jones potential (192), the force constants (195) can be computed by straightforward if tedious algebra. The force constant (195) will be a function of the equilibrium distances $\mathbf{s} = \mathbf{R}_l - \mathbf{R}_m$.

$$\Phi_{lm\alpha\beta} = \Phi_{\alpha\beta}(\mathbf{s})$$

For simplicity we assume that the crystal only contains one kind of atom (or ion), and that all the equilibrium sites in the lattice are equivalent (have the same symmetry). We can now make a few general statements about properties of $\Phi_{\alpha\beta}(\mathbf{s})$ that do not depend on the details of the potential.

First note that me may exchange the order of differentiation in (195). Hence

$$\Phi_{\alpha\beta}(\mathbf{s}) = \mathbf{\Phi}_{\beta\alpha}(-\mathbf{s})$$

Since there is an equivalent neighbor at **-s** for every neighbor at **s** we also have

$$\Phi_{\alpha\beta}(\mathbf{s}) = \mathbf{\Phi}_{\alpha\beta}(-\mathbf{s})$$

There can be no change in energy associated with a uniform translation of the crystal as a whole i.e. if $\mathbf{u}_l$ =constant for all l. This means that

$$\sum_{\mathbf{s}} \Phi_{\alpha\beta}(\mathbf{s}) = \mathbf{0}$$

We can now solve (196) by looking for solutions on the form

$$\ddot{\mathbf{u}}_l(t) = \vec{\epsilon} e^{i\mathbf{k}\cdot\mathbf{R}_l - \omega t} \tag{197}$$

The amplitude of the oscillations is now the *polarization vector* $\vec{\epsilon}$, with components ϵ_α $(\alpha = x, y, z)$. Substitution of (197) into (196) gives

$$-m\omega^2 \epsilon_\alpha = -\sum_{\mathbf{s},\beta} \Phi_{\alpha\beta}(\mathbf{s}) e^{-i\mathbf{k}\cdot\mathbf{s}} \epsilon_\beta$$

The Fourier transform of the force constant matrix Φ is the *dynamical matrix*

$$D_{\alpha\beta}(\mathbf{k}) = \sum_{\mathbf{s}} \mathbf{\Phi}_{\alpha\beta}(\mathbf{s})\mathbf{e}^{-\mathbf{ik\cdot s}}$$

which because of the symmetries discussed above can be rewritten

$$D_{\alpha\beta}(\mathbf{k}) = \frac{\mathbf{1}}{\mathbf{2}}\sum_{\mathbf{s}} \mathbf{\Phi}_{\alpha\beta}(\mathbf{s})[\mathbf{e}^{\mathbf{ik\cdot s}} + \mathbf{e}^{-\mathbf{ik\cdot s}} - \mathbf{2}]$$

or

$$D_{\alpha\beta}(\mathbf{k}) = -\mathbf{2}\sum_{\mathbf{s}} \mathbf{\Phi}_{\alpha\beta}(\mathbf{s})\sin^{\mathbf{2}}(\frac{\mathbf{k\cdot s}}{\mathbf{2}}) \tag{198}$$

This form is very similar to what we had for the harmonic chain. In that case there is only one component of the matrix and only two neighbors $\mathbf{s}$ and the "home site" $\mathbf{s} = 0$ which contribute to the sum. The normal mode frequencies can thus be found by solving the cubic determinantal equation

$$\begin{vmatrix} D_{xx}(\mathbf{k}) - m\omega^2 & D_{xy}(\mathbf{k}) & D_{xz}(\mathbf{k}) \\ D_{yx}(\mathbf{k}) & D_{yy}(\mathbf{k}) - m\omega^2 & D_{yz}(\mathbf{k}) \\ D_{zx}(\mathbf{k}) & D_{zy}(\mathbf{k}) & D_{zz}(\mathbf{k}) - m\omega^2 \end{vmatrix} = 0 \tag{199}$$

and the polarization vector ϵ is an eigenvector of the dynamical matrix. The dynamical matrix will be *real*. From the symmetries discussed above we find that it also will be *symmetric*. It is a theorem of linear algebra that a real symmetric matrix has real eigenvalues and that the eigenvectors $\vec{\epsilon}$ will be orthogonal to each other. The eigenvalues must also be positive, if not we would have solutions the dynamical equations with imaginary frequencies. This would allow solutions which grow exponentially in time, and would imply that the crystal is *unstable*. Mathematically a negative eigenvalue would mean that there would be some direction in which the restoring force associated with a displacement would be negative, i.e. the crystal potential at the equilibrium site has a maximum or a saddle point. The cubic equation which results when one multiplies out the determinant in (199) will thus have three positive roots for ω^2. If we take the square root of ω^2 we thus get real frequencies.

Let us consider the case where the wave vector k is small and let $\mathbf{k} = k\hat{\mathbf{k}}$ where $\hat{\mathbf{k}}$ is a unit vector in the direction of $\mathbf{k}$. For small x we have that $\sin x \approx x$ and

$$\omega_{k\alpha}^2 = -\frac{k^2}{2}\sum_{\mathbf{s}} \Phi_{\alpha\beta}(\mathbf{s})(\hat{\mathbf{k}}\cdot)^2 \tag{200}$$

We see that, if the lattice potential is sufficiently short range that the sum in (200) is convergent, the frequency will be proportional to k. The proportionality constant is the velocity of sound, which in principle will depend on both the direction of k and of the polarization $\vec{\epsilon}$.

One can show that for small values of $\mathbf{k}$ one of the eigenvectors $\vec{\epsilon}$ will be parallel to $\mathbf{k}$, i.e. the vibrations will be *longitudinal*. The two other modes will have $\vec{\epsilon}$ perpendicular to $\mathbf{k}$ and the vibrations will be *transverse*. In the general case the amplitude vector will not be exactly parallel or perpendicular to $\mathbf{k}$, but the three vectors $\vec{\epsilon}$ corresponding to a given $\mathbf{k}$ will be orthogonal to each other.

7.3 Phonons

We treat the normal modes of the lattice as a set of independent harmonic oscillators (see section 4.5) each with frequency $\omega_{\mathbf{k}\alpha}$.

$$\langle E \rangle = \frac{\hbar\omega_{\mathbf{k}\alpha}}{2} + \frac{\hbar\omega_{\mathbf{k}\alpha}}{e^{\beta\hbar\omega_{\mathbf{k}\alpha}} - 1}$$

Each mode will contribute a term

$$\frac{\partial\langle E \rangle}{\partial T} = \frac{\partial\langle E \rangle}{\partial\beta}\frac{\partial\beta}{\partial T} = \frac{1}{k_B T^2}\frac{(\hbar\omega_{\mathbf{k}\alpha})^2 e^{\beta\hbar\omega_{\mathbf{k}\alpha}}}{(e^{\beta\hbar\omega_{\mathbf{k}\alpha}} - 1)^2}$$

to the specific heat. At high temperatures $\beta \to 0$, and

$$e^{\beta\hbar\omega} \to 1 + \beta\hbar\omega = 1 + \frac{\hbar\omega}{k_B T}$$

and the contribution to the specific from each mode approaches k_B.

In what follows we will need to calculate thermal averages which involves summing over contributions from all the individual modes. For this purpose we need to find a way to convert sums over allowed $\mathbf{k}$-values into a volume integral over d^3k, just as we did in the case of black-body radiation and for electrons. For simplicity we let the crystal contain $N = L \times L \times L$ cubic cells, of side a, each of which contains one atom. With periodic boundary conditions the allowed values of the $\mathbf{k}$-vector are

$$\mathbf{k} = \frac{n_1}{L}\hat{\mathbf{x}} + \frac{n_2}{L}\hat{\mathbf{y}} + \frac{n_3}{L}\hat{\mathbf{z}} \tag{201}$$

Consider the modes for which the integers n_1, n_2 and n_3 in (7-22) are between

n_1 and $n_1 + dn_1$
n_2 and $n_2 + dn_2$
n_3 and $n_3 + dn_3$
The total number of such modes is

$$dn_1 dn_2 dn_3 = d^3 n \tag{202}$$

From (201) and (202) we see that these modes occupy a 'volume' in k-space

$$d^3 k = \frac{(2\pi)^3}{V} d^3 n$$

Only the modes for which $\frac{L}{2} \leq n_1 < \frac{L}{2}$, $\frac{L}{2} \leq n_2 < \frac{L}{2}$, $\frac{L}{2} \leq n_3 < \frac{L}{2}$, can be considered to be distinct and there will be N such $\mathbf{k}$-values. If the volume V is macroscopically large the k-values for successive normal modes are very close together and we can replace a sum of allowed k-states by an integral according to

$$\sum_{\mathbf{k}} = \frac{V}{(2\pi)^3} \int d^3 k \tag{203}$$

Since there are three modes for each $\mathbf{k}$-value, there are 3N normal modes.

7.4 Debye model

In general, the lattice heat capacity will be the sum of the contributions from the individual modes. We find

$$C_V = \frac{1}{k_B T^2} \sum_{k,\alpha} \frac{[\hbar\omega(\mathbf{k},\alpha)]^2 e^{\beta\hbar\omega_{\mathbf{k},\alpha}}}{(e^{\beta\hbar\omega_{\mathbf{k},\alpha}} - 1)^2} \tag{204}$$

Here, $\omega_{\mathbf{k},\alpha}$ is the frequency of one of the three modes with wave vector $\mathbf{k}$.

We now describe a simplified model which allows us to evaluate (204), to a good approximation. In the *Debye model* one makes the following assumptions:

(1). The frequency of a normal mode with wave vector $\mathbf{k}$ is approximated as

$$\omega(\mathbf{k},\alpha) = sk \tag{205}$$

where s is an average sound velocity.

(2). The region of k-space that corresponds to distinct modes is approximated by a sphere. The radius q_D of the sphere is given by the requirement that the number of allowed wave vectors inside the sphere should be equal to the number of atoms N

$$\frac{4\pi}{3}{q_D}^3\frac{V}{(2\pi)^3} = N$$

or

$$q_D = (6\pi^2\frac{N}{V})^{1/3} \tag{206}$$

The maximum phonon frequency in the Debye model is called the *Debye frequency*

$$\omega_D = sq_D \tag{207}$$

We also define the *Debye temperature* as

$$\theta_D = \frac{\hbar\omega_D}{k_B} \tag{208}$$

We substitute (205) into (204)

$$C_V = \frac{3V}{(2\pi)^3}\frac{4\pi}{k_B T^2}\int_0^{q_D} k^2 dk \frac{\hbar^2 s^2 k^2 e^{\beta\hbar sk}}{(e^{\beta\hbar sk}-1)^2}$$

(The factor of 3 comes from the sum over α.). We introduce the dimension less variable $x = \beta\hbar sk$ and find

$$C_V = \frac{3V}{2\pi^2}\frac{{k_B}^2 T^3}{\hbar^3 s^3}\int_0^{\beta\hbar sq_D}\frac{x^4 e^x dx}{(e^x-1)^2}$$

which simplifies, using (206) (207) and (208), to

$$C_V = 9Nk_B(\frac{T}{\theta_D})^3\int_0^{\theta_D/T}\frac{x^4 e^x dx}{(e^x-1)^2} \tag{209}$$

Equation (209) can be evaluated analytically in some limits. At low temperatures $\theta_D/T \to \infty$. The integral in (209) can the be evaluated analytically to

$$\int_0^{\infty}\frac{x^4 e^x dx}{(e^x-1)^2} = \frac{4\pi^4}{15}$$

Table 1: Debye temperatures of some elements and compounds

Element	$\theta_D[K]$	Compound	$\theta_D[K]$
Li	335	$NaCl$	280
Na	156	KCl	230
K	91	CaF_2	470
Cu	343	LiF	680
Ag	226	Quartz	255
Au	162		
Al	428		
Ge	378		
Si	647		
Diamond	1860		

and we get

$$\lim_{T\to 0} C_V = \frac{12\pi^4}{5} Nk_B \left(\frac{T}{\theta_D}\right)^3 \tag{210}$$

i. e. the low temperature specific heat will be proportional to T^3. For high temperatures we note that for small values of x the integrand in (209) is $\simeq x^2$. The integral is then $\simeq \frac{1}{3}(\theta_D/T)^3$ and we recover the Dulong and Petit heat capacity.

$$C_V = 3Nk_B$$

Finally, we note that if we plot the specific heat per mole vs. the temperature in units of the Debye temperature, (209) predicts a universal curve (see figure 17). This prediction is remarkably accurate. To get the good agreement the Debye temperature should be considered as a fitting parameter, i.e. the value of θ_D was chosen which gave the best overall fit. Some fitted Debye temperatures are tabulated above.

7.5 Problem set 6

1: Consider a linear chain of atoms. Each atom interacts with its nearest neighbor on either side via a Lennard-Jones potential (192). Assume parameter values appropriate to krypton ($\sigma = 3.65$Å, $\epsilon = 0.0140$ eV).

(a). Find the equilibrium spacing between the atoms.

(b). Find the sound velocity.

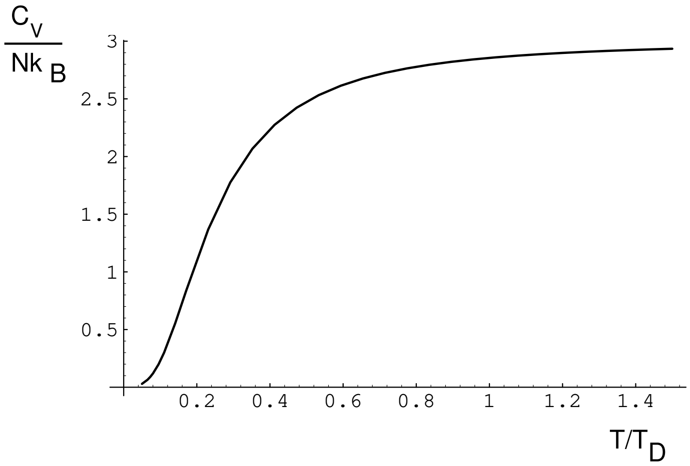


Figure 17: Heat capacity in the Debye model.

(c). What is the maximum frequency?

2: Plot by evaluating (190), for a sequence of times, the amplitude of a monatomic chain of 100 atoms with periodic boundary conditions. Give a physical interpretation of the evolving pattern, by making use of the fact that the sound velocity of the harmonic chain increases with increasing wavelength.

8 Phase transitions

8.1 Gibbs phase rule

Our next topic is phase transitions. We first consider the thermodynamics of the situation. A pure substance can typically exist as either a solid, a liquid or a gas. The Helmholtz free energy per particle $f(T, v) = F/N$ can be considered as a function of the temperature and volume per particle, $v = V/N$. The pressure is given by the equation of state

$$P = -\frac{\partial F}{\partial V} = -\frac{\partial f}{\partial v}$$

The equation of state can be thought of as defining a surface in P, v, T space (see figure 18). For coexistence of two phases, labeled by subscripts 1 and 2, (e.g. liquid and vapor) we must have $P_1 = P_2,\ T_1 = T_2,\ \mu_1 = \mu_2$. The restriction on the chemical potential represents an additional constraint on the equation of state surface and defines a line on that surface. In figure 8-2 we project such lines on the P, T plane.

We notice that there are special points, the critical point C and the triple point T. The triple point involves three phase coexistence. The additional constraint $\mu_1 = \mu_2 = \mu_3$ restricts us to an isolated point on the phase boundary. Similarly the condition that the partial volumes be the same at the critical end point of the gas liquid co-existence line restricts us to an isolated point.

Consider next an r-component mixture containing s coexisting phases. Let $c(i, j)$ be the mole fraction of constituent i in the j'th coexisting phase. We have $i = 1, 2...r$, and $j = 1, 2, ...s$. How many coexisting phases s can we have?

Let us consider the restriction on the variables. T, v are independent variables (control parameters). The mole-fractions are restricted by the condition

$$\sum_{i=1}^{r} c(i, j) = 1 \ \ for\ all\ j$$

There are therefore $s(r-1)$ independent concentration variables. Equating the s chemical potentials of the coexisting phases gives rise to $r(s-1)$ equations. For these equations to have solutions we must not have more equations than we have variables or:

$$s(r-1) + 2 \geq r(s-1)$$

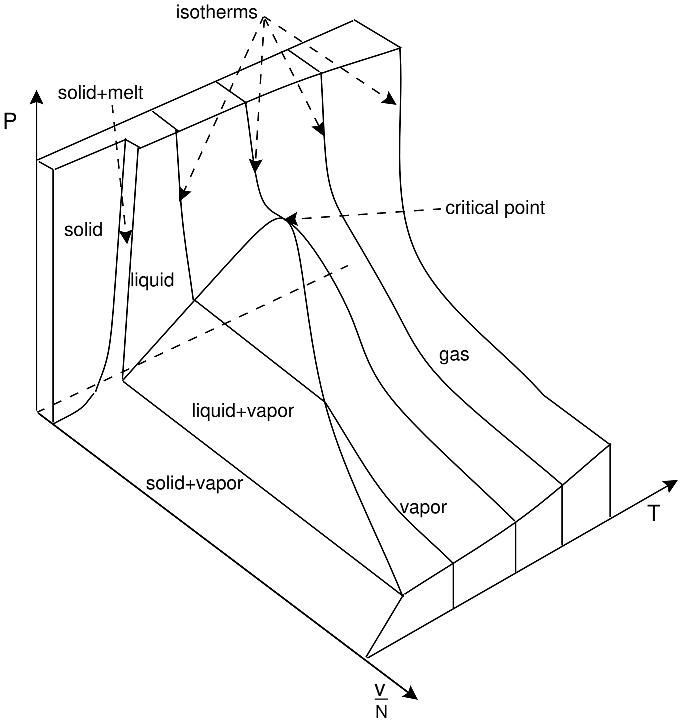


Figure 18: Equation of state surface

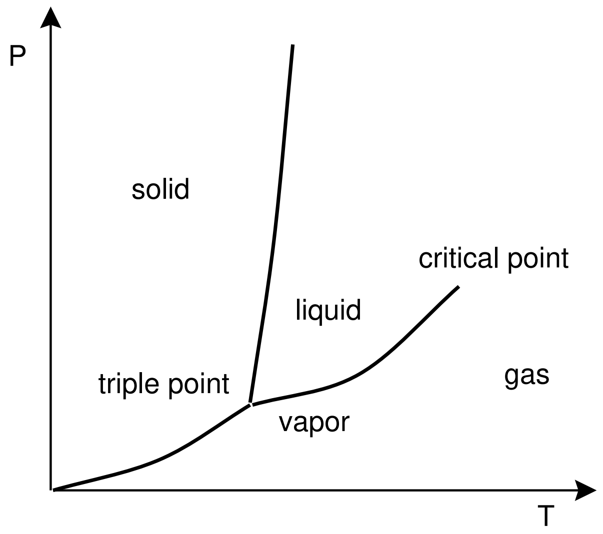


Figure 19: Projection of equation of state surface on P, T plane

8.2 'Magnetic' models

According to current theory there are two main types of types of magnetism:

> (1). The magnetic moments are localized on distinct sites (ions).
>
> (2). Magnetization arises because of difference in occupation of electronic bands for electrons of different spins (itinerant electrons). The magnetism in the two most important magnetic materials, iron and nickel are of this type.

In both cases there is an energy of interaction between the electronic spins (exchange energy) which favors parallel (ferromagnetic) or anti-parallel (anti-ferromagnetic) alignment of the spins.

We will here restrict our attention to case (1), since as we have already seen models of this types have many applications also outside the field of magnetism. There are three main types of discrete site models.

(a). *Ising models.*

The magnetic energy (or enthalpy if we stick with the good intentions established earlier) can be written

$$H = -\frac{1}{2}\sum_{i \neq j} J_{ij} s_i s_j - B\mu \sum_i s_i \tag{211}$$

Here J_{ij} is the *exchange energy* . If $J_{ij} > 0$ the coupling between the spins is ferromagnetic, if $J_{ij} < 0$ the coupling is anti-ferromagnetic while if J_{ij} is a random variable with sometimes positive sometimes negative sign (as e.g. in the neural networks) the coupling is of the *spin glass* type. The site variable s_i can only take on one of two values $s_i = \pm 1$. It is possible to generalize to more than two possibilities, but in the Ising model their number is always finite and they are discrete.

(b). *$x - y$ model.*

In this model the spin is a two component vector $\mathbf{s}_i = (s_{ix}, s_{iy})$. The interaction between any pair of spins is of the type

$$J_{ij} \mathbf{s}_i \cdot \mathbf{s}_j$$

The spins can be classical in which case the components are continuous variables, but with the length $(s)^2$ fixed. In that case the model is sometimes referred to as a 'clock model'. Alternatively we might assume quantum spins and let the components of $\mathbf{s}$ be Pauli spin matrices.

(c). *Heisenberg model*

This model is similar to the $x - y$ model but the spins are three dimensional vectors.

8.3 Mean field theory for the Ising model

We will here concentrate on the Ising model, both because it is the simplest of the above models, and also because this model has most applications outside the field of magnetism.

Let us first consider the case

$$J_{ij} = constant\ for\ all\ spins = \frac{K}{N}$$

where we assume that $K > 0$ (ferromagnetic case). For a large system the magnetization fluctuates about its mean value

$$\sigma = \langle s_i \rangle = \frac{1}{N}\sum_i s_i + small\ fluctuations$$

We refer to σ as the *order parameter*. Introducing the reduced magnetic field $h = B\mu$ we have for the Hamiltonian (211)

$$H = -\frac{K}{2N}\sum_{i\neq j} s_i s_j - h\sum_i s_i = -\frac{K}{2N}(\sum_i s_i)(\sum_j s_j) + \frac{K}{2N}\sum_i {s_i}^2 - h\sum_i s_i$$

$$= -\frac{KN\sigma^2}{2} - Nh\sigma + \frac{K}{2} + fluctuations \qquad (212)$$

the last two terms on the right hand side of (212) will be small compared to the first two in the limit $N \to \infty$. We next turn to the entropy. If the magnetization is $N\sigma$ there will be $\frac{N}{2}(1+\sigma)+$ spins and $\frac{N}{2}(1-\sigma)-$ spins and we find for the entropy

$$S = k_B \ln\left\{\frac{N!}{[\frac{N}{2}(1+\sigma)]![\frac{N}{2}(1-\sigma)]!}\right\} = -Nk_B\left\{\frac{1+\sigma}{2}\ln\left[\frac{1+\sigma}{2}\right] + \frac{1-\sigma}{2}\ln\left[\frac{1-\sigma}{2}\right]\right\} \qquad (213)$$

where we have used the Stirling formula. Collecting terms we find for the free energy

$$G = H - TS = N\left\{-\frac{K\sigma^2}{2} - h\sigma + \frac{1}{\beta}\left[\frac{1+\sigma}{2}\ln\left(\frac{1+\sigma}{2}\right) + \frac{1-\sigma}{2}\ln\left(\frac{1-\sigma}{2}\right)\right]\right\}$$

At equilibrium the order parameter σ will take on the value for which F is a minimum. Differentiation gives

$$\frac{\partial G}{\partial \sigma} = N\left\{-K\sigma - h + \frac{1}{2\beta}\ln\left(\frac{1+\sigma}{1-\sigma}\right)\right\} \qquad (214)$$

We put the right hand side of (214) to zero and get

$$\exp[2\beta(K\sigma + h)] = \frac{1+\sigma}{1-\sigma}$$

or

$$\sigma = \tanh[\beta(K\sigma + h)] \qquad (215)$$

In the special case $h = 0$ (215) reduces to

$$\sigma = \tanh(\beta K\sigma) \qquad (216)$$

We see that $\sigma = 0$ is always a solution to (216). If we differentiate (214) once more we find

$$\frac{1}{N}\frac{\partial^2 G}{\partial \sigma^2} = -K + \frac{1}{2\beta}\left(\frac{1}{1+\sigma} + \frac{1}{1-\sigma}\right)$$

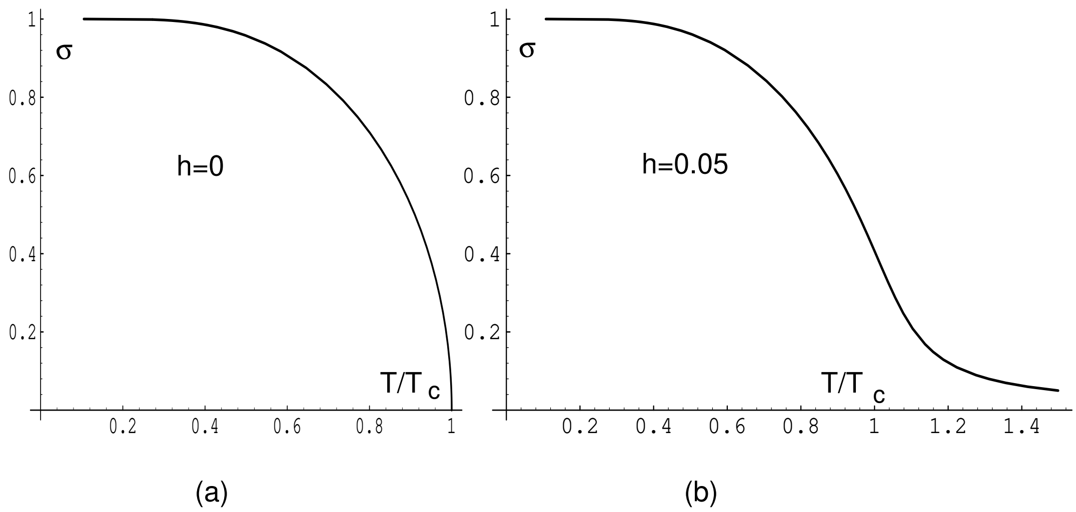


Figure 20:

For $\beta < 1/K$ or $T > T_c = K/k_B$ the second derivative is positive for $\sigma = 0$. The non magnetic state $\sigma = 0$ will then represent a minimum of the free energy. For $\beta > 1/K$ or $T < T_c$ the second derivative becomes positive and $\sigma = 0$ represents a maximum in the free energy. The non magnetic state then becomes unstable and there will be a *spontaneous magnetization*.

The equilibrium value of the magnetization for $T < T_c$ can easily be computed by choosing a value of $x = \beta K \sigma$. One can then find σ from (216) and T from

$$T = \frac{K\sigma}{k_B x}$$

The resulting magnetization is plotted in the figure 20(a). The same method can be used to calculate the magnetization in a non-zero field and the result is sketched in figure 20(b). In that case there will also be a solution in which the magnetization is opposite to the field. It is easy to verify that the free energy of this solution is larger than when σ and h have the same sign, and we can discard this solution. We next turn to a non magnetic application of the ferromagnetic Ising model.

8.4 Solubility gaps

Consider a mixture of two molecular species A and B and let x be the mole-fraction of A molecules and $1 - x$ the mole-fraction of B molecules. We wish to address the well known phenomenon that certain fluids, such as oil and vinegar, do not mix at low temperatures for all concentrations. For simplicity assume that the molecules are of the same size so that the volume stays constant as the mole-fraction is varied. We introduce

$u_{AA}(r)$ = interaction energy between two A molecules at distance r.
$u_{BB}(r)$ = interaction energy between two B molecules at distance r.
$u_{AB}(r)$ = interaction energy between an A and a B molecule separated by r.

Again for simplicity we assume that the interactions are long range so that we can work with the average energy

$$u_{AA} = \frac{e_{AA}}{N},\ u_{BB} = \frac{e_{BB}}{N},\ u_{AB} = \frac{e_{AB}}{N}$$

We have

$$\begin{array}{rll} \frac{N_A(N_A-1)}{2} & \simeq \frac{1}{2}x^2N^2 & = \#\ of\ AA\ pairs \\ \frac{N_B(N_B-1)}{2} & \simeq \frac{1}{2}(1-x)^2N^2 & = \#\ of\ BB\ pairs \\ N_AN_B & \simeq x(1-x)N^2 & = \#\ of\ AB\ pairs \end{array}$$

The energy is then

$$E = \frac{N}{2}[x^2e_{AA} + (1-x)^2e_{BB} + 2x(1-x)e_{AB}] \tag{217}$$

In analogy with (213) we find for the entropy

$$S = -Nk_B[x\ln x + (1-x)\ln(1-x)] \tag{218}$$

One method to search for the possible existence of solubility gap is to consider the grand potential

$$\Omega = E - TS - \mu_AN_A - \mu_BN_B \tag{219}$$

where μ_A, μ_B are the chemical potentials for the two species. At the equilibrium concentration we require that the grand potential is a minimum and get the equation of state

$$\frac{\partial\Omega}{\partial x} = 0$$

The solubility gap then arises, because certain x values do not give rise to a minimum in Ω. We find after straightforward algebra

$$\frac{\Omega}{N} = \frac{x^2}{2}(e_{AA} + e_{BB} - 2e_{AB}) + x(e_{AB} - e_{BB}) + \frac{1}{2}e_{BB}$$
$$+\frac{1}{\beta}(x \ln x + (1-x)\ln(1-x)) - (\mu_A - \mu_B)x - \mu_B \tag{220}$$

Define

$$\epsilon = 2e_{AB} - e_A - e_B$$
$$t = \frac{1}{\beta\epsilon}$$
$$\mu = \frac{1}{\epsilon}(\mu_A - \mu_B - e_{AB} + e_{BB})$$
$$\omega = \frac{\Omega}{N\epsilon} = -\frac{x^2}{2} - \mu x + t[x \ln x + (1-x)\ln(1-x)]$$

We assume that $\epsilon < 0$ (otherwise it wouldn't be energetically favorable for the mixture to separate). We have

$$\frac{\partial\omega}{\partial x} = -x - \mu + t \ln\frac{x}{1-x} \tag{221}$$
$$\frac{\partial^2\omega}{\partial x^2} = -1 + t(\frac{1}{x} + \frac{1}{1-x})$$

It is easy to show that

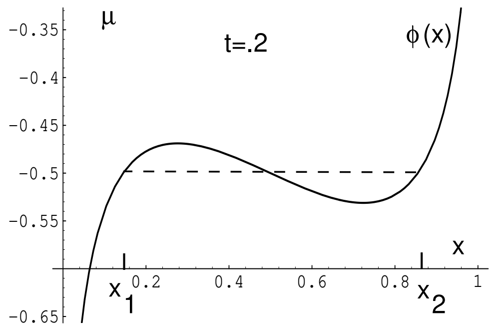
$$\frac{1}{x} + \frac{1}{1-x}$$

has its smallest value 4 in the interval $0, x, 1$ for $x = .5$. We see that if the reduced temperature $t > \frac{1}{4}$ then $\partial^2\omega/\partial x^2 > 0$, always. On the other hand if $t < \frac{1}{4}$ the second derivative will take on both positive and negative values. Putting the derivative in (221) to zero gives

$$\mu = -x + t \ln\frac{x}{1-x} \equiv \phi(x)$$

A plot of $\phi(x)$ vs. x is given in figure 21 for the two cases $t < \frac{1}{4}$, $t > \frac{1}{4}$.

We conclude that for $t > 1/4$ the mixture is stable for all concentrations. For $t < 1/4$ there will be two concentrations x_1, x_2 for which two mixtures with different concentrations x_1 and $x_2 can$ co-exist. To find the precise vales x_1 and x_2 we have to make use of the additional condition $\omega(x_1) = \omega(x_2)$.



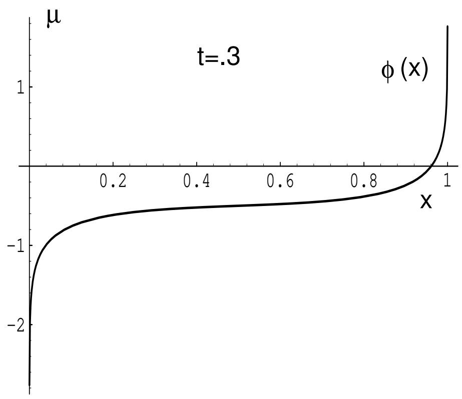


Figure 21:

In a mixture for which $x_1 < x < x_2$ the system will split into two coexisting phases of concentration x_1 and x_2 respectively. Finally if x< x_1 or x> x_2 the mixture is stable. Note the analogy between the present situation and the isotherms of a van der Waals fluid. The t vs. x *phase diagram* is plotted in figure 22.

8.5 Self-consistent field method

It is instructive to derive some of the results of the previous section using a different method. Consider the Hamiltonian associated with long range interactions in zero field

$$H = -\frac{K}{2N}\sum_{i\neq j} s_i s_j \tag{222}$$

We consider, as a subsystem, a given site spin s_0 . The terms in H which depend on s_0 come from the terms in (222) for which either $i = 0$ or j= 0. These terms are

$$H_0 = \frac{-K}{N} s_0 \sum_{i\neq 0} s_i = -\frac{K s_0 (N-1)\sigma}{N} + fluctuations \simeq -K s_0 \sigma$$

The partition function for the subsystem is

$$Z_0 = e^{\beta K\sigma} + e^{-\beta K\sigma} = 2\cosh(\beta K\sigma)$$

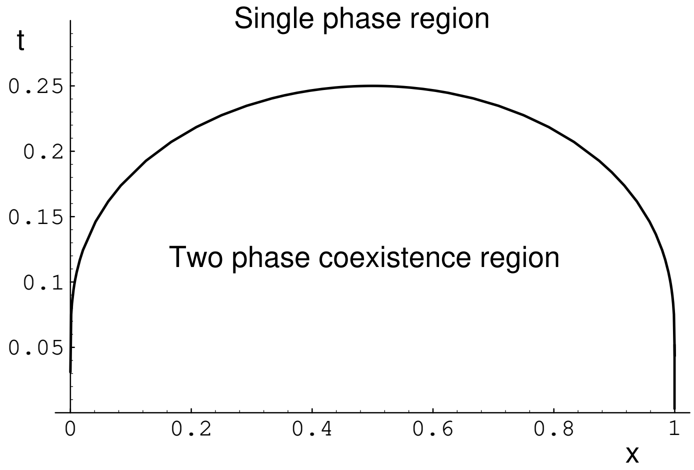


Figure 22:

and we have

$$\langle s_0 \rangle = \frac{e^{\beta K \sigma} - e^{-\beta K \sigma}}{e^{\beta K \sigma} + e^{-\beta K \sigma}} = \tanh(\beta K \sigma)$$

Since the selection of the 'home spin' s_0 was completely arbitrary self consistency requires that $\langle s_0 \rangle = \sigma$ giving

$$\sigma = \tanh(\beta K \sigma)$$

which is the same as (222). This is an exact result for the system with long range interactions. It is equivalent to the mean field theory discussed in the previous section. Both formulations can be generalized as an approximate method for systems with short range forces.

8.6 Order disorder transition in a binary alloy

When the mean field theory is applied to systems with short range interaction it is called the *Bragg-Williams approximation*. As an example we consider the order disorder transition in a binary alloy. The most well known example of such a transition occurs in $\beta - brass$ an alloy consisting of approximately 50% Cu and 50% Zn.

The alloy has a body centered cubic structure (figure 23). We may think of the ordered phase as one in which the corners of the cube are mainly occupied by Zn atoms while the cube centers are mainly occupied by Cu

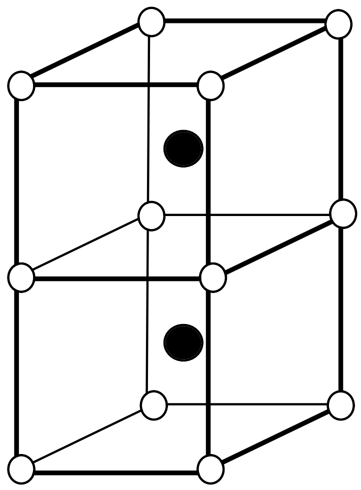


Figure 23: Ordered phase of β-brass

atoms. The corners and the cube centers both constitute two simple cubic sublattices.

In the high temperature disordered phase the atoms are randomly distributed between the cube corners and cube centers. We will use a notation in which A and B refer to the two atomic species and where the subscript 1 refers to a corner position and 2 a body center. For example, N_{A1} is the number of A atoms in corner positions. We assume that the interactions are short range so short range that there are only nearest neighbor interactions (body center to corner).

We define

$$\begin{aligned} N_{AA} &= \# \text{ of nearest neighbor } AA \text{ pairs} \\ N_{AB} &= \# \text{ of nearest neighbor } AB \text{ pairs} \\ N_{BB} &= \# \text{ of nearest neighbor } BB \text{ pairs} \end{aligned}$$

and write for the energy

$$E = N_{AA}e_{AA} + N_{AB}e_{AB} + N_{BB}e_{BB}$$

For simplicity we assume that exactly 50% of the atoms are of the A type and 50% are of the B type. This allows us to write

$$\frac{N}{2} = N_{A1} + N_{A2} = N_{B1} + N_{B2} = N_{A1} + N_{B1} = N_{A2} + N_{B2}$$

We introduce the *order parameter*

$$m = \frac{N_{A1} - N_{A2}}{\frac{1}{2}N}$$

where $-1 \leq m \leq 1$. We have

$$N_{A1} = N_{B2} = \frac{N}{2}(1+m)$$

$$N_{A2} = N_{B1} = \frac{N}{2}(1-m)$$

We define q as the number of nearest neighbors ($q = 8$ for the body centered cubic lattice). In the mean field approximation we put for the number of nearest neighbor pairs of different type

$$N_{AA} = qN_{A1}\frac{2N_{A2}}{N}$$

$$N_{BB} = qN_{B1}\frac{2N_{B2}}{N}$$

$$N_{AB} = \frac{2q}{N}(N_{A1}N_{B2} + N_{A2}N_{B1})$$

We define

$$\epsilon = q(e_{AA} + e_{BB} - 2e_{AB})$$

In the present problem we assume that $\epsilon > 0$. This is the opposite of the situation that we had in case of phase separation. We now assume that it is energetically more favorable for A atoms to associate with B atoms than with their own kind. The coupling is now *anti-ferromagnetic*. The energy is now given by

$$\frac{E}{N} = -\epsilon m^2 + q(e_{AA} + e_{BB} + 2e_{AB})$$

We write for the entropy

$$S = k_B \ln\left(\frac{N_{A1}!N_{B1}!N_{A2}!N_{B2}}{(\frac{1}{2}N!)^2}\right) \simeq Nk_B\left\{\frac{1+m}{2}\ln\frac{1+m}{2} + \frac{1-m}{2}\ln\frac{1-m}{2}\right\} \tag{223}$$

The last step in (223) follows after some algebra using Stirling's formula. The approximate free energy is then

$$f = \frac{F}{N} = const. - \epsilon m^2 + k_B T\left\{\frac{1+m}{2}\ln\frac{1+m}{2} + \frac{1-m}{2}\ln\frac{1-m}{2}\right\}$$

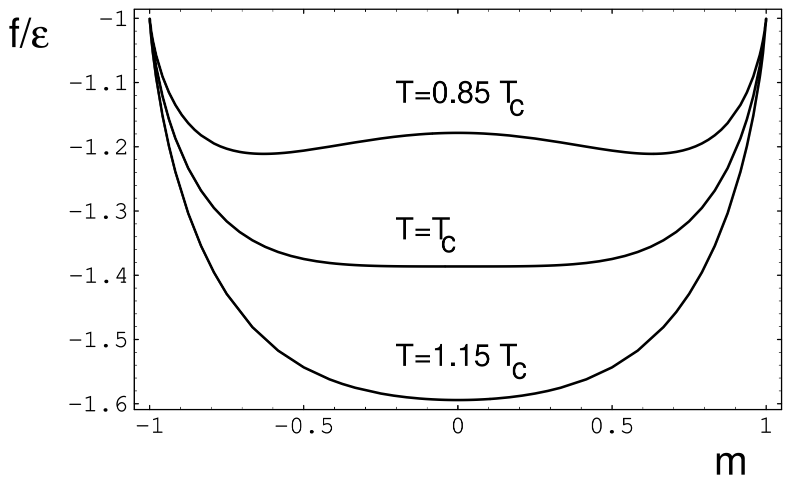


Figure 24: Free energy as a function of order parameter for different temperatures

The equilibrium of the order parameter m is the one for which the free energy is a minimum. We have

$$\frac{\partial f}{\partial m} = 0 = -2\epsilon m + \frac{k_B T}{2} \ln \frac{1+m}{1-m}$$

$$\frac{\partial^2 f}{\partial m^2} = -2\epsilon + \frac{k_B T}{2} \left(\frac{1}{1+m} + \frac{1}{1-m} \right)$$

We see that for $k_B T > \epsilon/2$ the free energy has a minimum for $m = 0$. This is the high temperature disordered state. For $k_B T < \epsilon/2$ there are two symmetric minima with $m \neq 0$ corresponding to the ordered alloy. The behavior of the free energy above and below $T = T_c = \frac{\epsilon}{2k_B}$ is shown in figure 24 respectively. The plot of the temperature dependence of the order parameter will look like figure 20(a) except σ is replaced by m.

8.7 Critical Phenomena

The physics near critical points in the phase diagram has interested physicists for many years as a research topic. In recent years particular attention has been paid to critical *non equilibrium phenomena*, although we will here only discuss briefly the equilibrium theory. Many properties of a system such as the order parameter, the specific heat, critical isotherms and the susceptibility will exhibit singular behavior near a critical point. Near such a point small changes in parameter values causes a qualitative change in the system behavior. Because of the lack of stability in the response the system will then exhibit large fluctuations. At the critical point fluctuations occur on all scales.

Let us revisit the mean field theory for the Ising model. We rewrite the self consistent equation for the order parameter σ in zero field

$$\sigma = \tanh\left(\frac{K\sigma}{k_B T}\right) = \tanh\left(\frac{T_c\sigma}{T}\right)$$

Below, but near, the critical temperature the order parameter σ will be small. We use the Taylor expansion for the hyperbolic tangent

$$\tanh(x) = x - \frac{x^3}{3} + \cdots$$

to find

$$\sigma = \left(\frac{T_c}{T}\right)\sigma - \frac{1}{3}\left(\frac{T_c}{T}\right)^3 \sigma^3 + \cdots \tag{224}$$

giving

$$\sigma(T) \simeq 3^{1/2}\left(\frac{T}{T_c}\right)^{3/2}\left(\frac{T_c}{T} - 1\right)^{1/2}$$

The order parameter will thus approach zero in a singular manner as $T \to T_c$ from below. The asymptotic form is

$$\sigma \sim \left(\frac{T_c}{T} - 1\right)^{1/2}$$

The exponent for the power law behavior of the order parameter is conventionally given the symbol β. In more realistic theories, and in real ferromagnets it will not be given by the simple fraction $1/2$ found here.

Next consider the magnetic susceptibility per spin which is defined as

$$\chi(h,T) = \left(\frac{\partial\sigma}{\partial h}\right)_T$$

where h is the magnetic field in reduced units. Recall that the self-consistent equation in the presence of a field is

$$\sigma = \tanh\left[\frac{K\sigma + h}{k_B T}\right]$$

by differentiating we find

$$\frac{\partial \sigma}{\partial h} = \cosh^{-2}\left[\frac{K\sigma + h}{k_B T}\right]\frac{1}{k_B T}\left(K\frac{\partial \sigma}{\partial h} + 1\right)$$

where Using $T_c = \frac{K}{k_B}$ and taking the limit h→ 0 we find

$$\chi(0,T) = \frac{1}{k_B T}\frac{1}{\cosh^2(\frac{T_c \sigma}{T})}\frac{1}{(1 - \frac{T_c}{T}\cosh^{-2}(\frac{T_c \sigma}{T}))}$$

For small x

$$\cosh(x) \simeq 1 + \frac{x^2}{2} \tag{225}$$

For $T > T_c$, $\sigma = 0$ and

$$\chi(0,T) = \frac{1}{k_B(T - T_c)}$$

For $T < T_c$, but close to T_c, we find after some algebra using (224) and (225)

$$\chi(0,T) = \frac{1}{2k_B(T_c - T)}$$

In both cases the susceptibility diverges as $T \to T_c$. It is conventional to write for χ near T_c

$$\chi(0,T) \sim |T - T_c|^{-\gamma}$$

In our mean field theory $\gamma = 1$ while different values are found in other models and in real magnets.

Finally let us consider the specific heat. The internal energy in our mean field theory is for $h = 0$

$$H = -\frac{KN\sigma^2}{2}$$

From (224) we find that as $T \to T_c$ from below

$$C_h = \left(\frac{\partial H}{\partial T}\right)_{h=0} \to \frac{3}{2}k_B$$

while $C_h = 0$ above T_c. Thus, the mean field theory produces a discontinuity at the transition. Other models exhibit a power law singularity in the specific heat

$$C_h \sim |T - T_c|^{-\alpha}$$

where α is the conventional symbol for the specific heat exponent.

It is natural to ask oneself on what do these exponents depend? Landau showed some 50 years ago that if the free energy is an analytic function of the order parameter near $\sigma = 0$, one finds that under very general conditions $\alpha = 0, \beta = \frac{1}{2}, \gamma = 1$. On the other hand, in the Ising model with short range interactions one can show that for a two dimensional lattice $\alpha = 0, \beta = .125, \gamma = 1.75$. There is no transition in one dimension. In three dimensions the exact results are not known, but simulation studies and advanced analytical methods are consistent with $\alpha \simeq 0.124, \beta \simeq 0.31, \gamma \simeq 1.25$. The value of the exponents does not depend on the lattice type. Some other critical phase transitions such as the gas liquid transition appears to have the same exponents as the Ising model. The Heisenberg model on the other hand has no transition in one or two dimensions while the exponents in three dimensions differs from the Ising exponents.

9 Some old exam questions

9.1 April 1990 PHYS 455 Exam

This paper contains 2 parts (I) Problems, (II) Questions. Solve 4 of the 5 problems and answer 4 out of the 5 questions. Allowed aids: 1 double sided 'cheat sheet'. Time 3 hours.

I Problems:

1: *Entropy of mixing*

(a). Consider two ideal gases both at the same temperature T and pressure P. One gas occupies the volume V_1 the other the volume V_2. The gases are mixed so they occupy the volume $V_1 + V_2$. Derive a formula for the entropy of mixing.

(b). 1 mole of argon and one mole of krypton are mixed at a fixed temperature. Originally the volume of each gas was V. After the mixing the volume is $2V$, Subsequently the gas is compressed isothermally to the volume V. What is the difference in entropy of the gases before and after mixing, before and after the isothermal compression.

(c). A 10 kg rock falls off a cliff 100m above the surface of a 100m deep lake. The rock settles at the bottom of the lake. What is the change of entropy of the rock? of the lake ? The density of the rock is 2000 kg/m^3 of the lake 1000 kg/m^3. Assume that initially the temperature of the rook and of the lake was 300K and that the potential energy of the rock is dissipated as heat in the lake. **4:** *Polymer model.*

N monomeric molecules are arranged along a straight line to form a chain molecule. Each monomeric unit can be in an α state or a γ state. In the former case the length of a monomer is a and the energy is E_α. The corresponding values in the latter state are g and E_γ.

(a). Derive the relation between the length L of the chain molecule and the tension $\mathcal{X}$ applied between the two ends of the molecule.

(b). Under what conditions will the chain contract when heated under constant tension.

5: *A dipolar gas.*

Show that the electric polarization P of an ideal gas consisting of N diatomic molecules having a constant electric dipole moment μ is given by

$$P = \frac{N\mu}{V}\left\{\coth(\frac{\mu E}{k_B T}) - \frac{k_B T}{\mu E}\right\}$$

II Questions:

Answer briefly 4 of the 5 questions.

3: Give an argument why ^4He is superfluid at low temperatures while an ideal Bose gas below the Bose Einstein transition temperature is not.

4: What is the Dulong Petit heat capacity?

Some formulas:

$$dE = TdS - PdV + \mu dN;\ F = E - TS;\ G = F + PV;\ d\mu = vdP - sdT$$

$$E = \frac{\partial(\beta F)}{\partial\beta};\ \beta = \frac{1}{k_B T};\ C_V = \left.\frac{\partial E}{\partial T}\right|_V;\ {\lambda_{th}}^{-3} = \left(\frac{mk_B T}{2\pi\hbar^2}\right)^{3/2}$$

$$F_{ideal\ gas} = -Nk_B T\left\{1 + \ln\left(\frac{V}{N{\lambda_{th}}^3}\right)\right\}$$

9.2 April 1991 PHYS 455 Exam

This paper contains 2 parts (I) problems (II) questions. Solve 4 out of the 5 problems and answer 4 out of the 5 questions. Time 3 hours. Closed book exam

I Problems:

1:

A liquid is in equilibrium with its vapor at temperature T and pressure P. An inert gas which is insoluble in the liquid, is introduced into the container. The partial pressure of the inert gas in the vapor phase is p_i. The

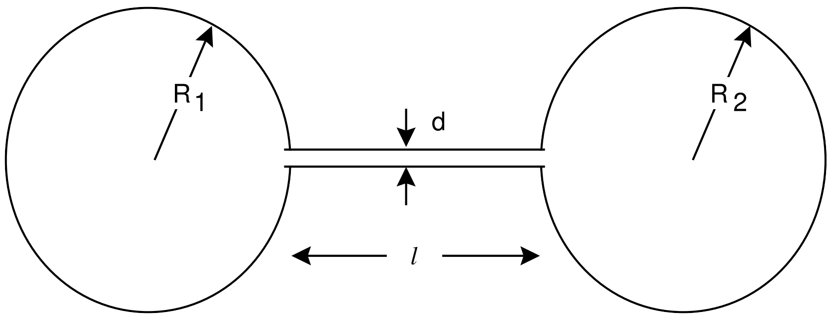


Figure 25:

volume change of the liquid can be neglected and $p_i << P$. Show that vapor pressure will increase by an amount δp which is approximately given by

$$\frac{\delta p}{P} = \frac{p_i v_L}{k_B T - P v_L} \simeq \frac{p_i v_L}{k_B T}$$

where v_L is the specific volume in the liquid phase (volume per particle).

2:

The surface energy of soap bubble is $2\sigma A$, where σ is the surface tension, A is the area and the factor of 2 comes from the fact that the bubble has an inside and an outside.

(a). Show that the excess pressure inside a soap bubble is $4\sigma/r$, where r is the radius of the bubble.

(b). Consider two equal bubbles attached to the ends of a straw as shown in figure 25 The system is perturbed from equilibrium by making one bubble slightly larger than the other. Common intuition would suggest that the bubbles would then undergo small oscillations about an equilibrium size. Instead the smaller bubble collapses. Explain why!

(c). A professor wishes to demonstrate the soap bubble experiment in his class. The professor is unable to create soap bubbles simultaneously at both ends of a straw and decides to use rubber balloons attached to a glass tube instead. When the balloons are blown up and attached to the ends of the glass tube they undergo small oscillations in size about an equilibrium position. Explain why!

[Hint: the energy associated with stretching the surface of the balloon is elastic. If the balloon when blown up is much larger than its original size the energy can be taken to be proportional to the square of the area.]

3:

(a). Show that the energy fluctuations in the canonical ensemble are given by

$$\langle E^2\rangle - \langle E\rangle^2 = \langle \delta(E)^2\rangle = k_B T^2 C_V$$

where C_V is the specific heat at constant volume.

(b). Evaluate the energy fluctuations $\langle E^2\rangle - \langle E\rangle^2$ for a quantum harmonic oscillator with frequency ω.

5:

(b). A system contains spins at two distinct sites. Each spin is either $\uparrow$ or $\downarrow$. When both spins point up or both spins point down the energy of the system is $-\epsilon$. When the two spins point in opposite directions the energy is 0. For what value of the temperature parameter $\beta = 1/k_B T$ will the probability that the two spins point in the opposite direction be 1/3.

II Questions:

Answer 4 out of 5 questions.

2: Sketch the behavior of the chemical potential of an ideal Bose gas as the temperature is varied. What is the qualitative difference between the isotherm (pressure vs. volume at constant T) of an ideal Bose gas and a classical ideal gas.

3: Explain the fountain effect in superfluid helium.

Some formulas:

$$d\mu = -sdT;\ s = S/N;\ v = V/N;\ G = \mu N = E - TS + PV;\ dE = TdS - PdV$$

$$F = E - TS;\ \ln N! = N\ln N - N;\ S = k_B \ln g\ \ C_V = \left.\frac{\partial E}{\partial T}\right|_{VN}$$

$$Z_c = \sum_\alpha e^{-\beta E_\alpha};\ Z_G = \sum_{\alpha N} e^{-\beta(E_\alpha - \mu N)}$$

9.3 April 1992 PHYS 455 Exam

Closed book exam. Time 3 hours. Answer 4 out of the 6 questions. If you answer more than 4 questions you will be given credit for the four best marks. All questions have equal value.

1:

A system is made up of two distinguishable particles. Each particle can occupy one of three possible states labeled 1,2 and 3. If the two particles are in a state with the same label the energy of the system is $-\epsilon$, if the labels are different the energy is 0. The temperature is T= $1/(\beta k_B)$.

(a). Find the mean energy of the system.

(b). Calculate the entropy as a function of the temperature.

(c). What is the mean entropy of the system *(i)* in the limit of high temperature, *(ii)* in the limit of low temperature. Can you give a simple argument why these result are expected?

2:

Consider a system of particles in the grand canonical ensemble.

(a). If the temperature is given in terms of $\beta = 1/(k_B T)$, the chemical potential is μ and the volume V is held fixed show that the mean number of particles is

$$\langle N \rangle = \frac{\partial \ln(Z_G)}{\beta \partial \mu}$$

where Z_G is the partition function.

(b). Show that the statistical variance in the number of particles is

$$\langle (\delta N)^2 \rangle = \langle N^2 \rangle - \langle N \rangle^2 = \frac{\partial \langle N \rangle}{\beta \partial \mu}$$

(c). A surface contains M potential adsorption sites, each site is occupied with probability

$$p = \frac{1}{\exp[\beta(\epsilon - \mu)] + 1}$$

where $\epsilon \langle 0$ is the energy of adsorption. Calculate the mean number of particles on the surface and the variance in the number of absorbed particles.

3: Two *3-dimensional* vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ of unit length (${s_1}^2 = {s_2}^2 = 1$) are free to rotate about the origin. The temperature is T and we can ignore the kinetic energy. The potential energy of the system is, with θ the angle between the vectors and J a constant.

$$U = -J\mathbf{s}_1 \cdot \mathbf{s}_2 = -J\cos(\theta)$$

(a). Calculate the partition function for the system.

(b). Calculate the mean value of $\langle \cos\theta \rangle$.

(c). Is $\langle \theta \rangle = \cos^{-1}\langle \cos\theta \rangle$?

5:

A sealed cylindrical vessel is in contact with a heat bath at constant temperature T. A friction-less airtight piston of weight mg divides the container into two volumes V_1 and $V_2 = V - V_1$. There are N_1 ideal gas atoms in the top partition and N_2 in the bottom partition.

(a). Find the equilibrium height of the piston.

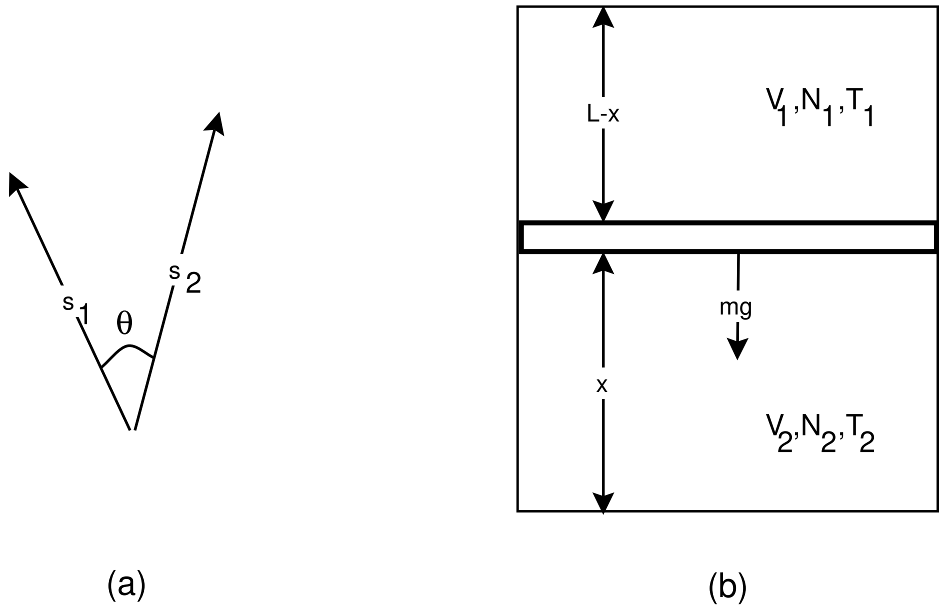


Figure 26:

(b). The ideal gas is replaced by a real single component gas. At a certain temperature the bottom partition is found to contain a puddle of liquid coexisting with its vapor. Which of the following statements may be true at equilibrium:

(i). The top partition contains a liquid in coexistence with its vapor.
(ii). The top partition contains only vapor.
(iii). The top partition contains only liquid.

(c). Assume that the cylinder is turned around so that the piston moves horizontally and that $N_1 = N_2$. Which of the three statements in part (b) may now be true.

6:

(a). Explain the difference between ^4He below the λ point and an ideal Bose gas, that makes the former a superfluid, while the latter is not.

(b). What is second sound and how can it be demonstrated?

Formulas:

$$dE = TdS - PdV + \mu dN;\ F = E - TS = -k_B T \ln Z_C$$

$$dF = -SdT + \mu dN - PdV;\ d\mu = vdP - sdT$$

$$E = \frac{\partial \beta F}{\partial \beta};\ \beta = \frac{1}{k_B T}$$

$$C_V = \left.\frac{\partial E}{\partial T}\right|_V;\ S = k_B \ln g;\ Z_C = \sum_\alpha \exp(-\beta E_\alpha)$$

$$Z_G = \sum_{\alpha N} \exp(-\beta[E_\alpha - \mu N]); \lambda = \sqrt{\frac{2\pi\hbar^2}{mk_B T}}$$

$$F_{ideal\ gas} = -Nk_B T \left\{ \ln(\frac{V}{N\lambda^3}) + 1 \right\}\ E = ST - PV + \mu N$$

9.4 April 1994 PHYS 455 exam

Time 2 1/2 hours. 'Closed book exam'. Answer 4 out of the 5 questions. All questions have equal value. If you answer more than 4 questions you will be given credit for the 4 best answers.

————————————————————-

2:

Consider a system of four spins each can be in either a $s = +1$ or a $s = -1$ state. If all spins are in the same state the energy is -6ϵ, where ϵ is a positive constant. If three spin are in one state and one in the other the energy is zero, if there are two spins in each state the energy is $+2\epsilon$.

a: What is the mean energy of the system at temperature T.

b: What is the mean entropy of the system at temperature T.

c: What value does the entropy approach as $T \to \infty$? as $T \to 0$? Give a physical interpretation of your result.

———————————————-

———————————————-

4:

A fiber has an average length $\langle L \rangle$ when pulled with force $\mathcal{X}$. The enthalpy of the fiber is

$$H = U(L,T) - L\mathcal{X}$$

where the internal energy $U(L,T)$ is a function of the length of the fiber and the temperature. Derive an expression for the Young modulus

$$Y = \left(\frac{\partial \mathcal{X}}{\partial \langle L \rangle} \right)_T$$

in terms of the temperature and the mean square fluctuation

$$\langle (\delta L)^2 \rangle = \langle L^2 \rangle - \langle L \rangle^2$$

———————————————

5:

N_A molecules of one species and N_B molecules of another species are initially kept separate in two containers each at pressure P and temperature T. The free energy of a mixture of the two gases can be approximated by the hard sphere expression by

$$F(N_A, N_B, V, T) = -(N_A + N_B)k_B T \{\ln \left[\frac{V - N_A b_A - N_B b_B}{\lambda^3} \right] + 1\}$$

where $\lambda = \sqrt{\frac{2\pi\hbar^2}{mk_BT}}$, and b_A and b_B and m are constants.

a: If the two gases are mixed at constant temperature and pressure, what is the volume of the mixture?

b: What is the entropy of mixing?

Some formulas

$$\langle (\delta E)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 = k_B T^2 C_V$$

$$Z = \sum_\alpha e^{-\beta H_\alpha}$$

$$S = -\frac{\partial F}{\partial T} = k_B \ln g$$

$$F = -k_B T \ln Z_C$$

References

[1] M. Plischke and B. Bergersen 1994, *Equilibrium Statistical Physics* Second edition. World Scientific.

[2] L. Boltzmann 1992, *A German professor's trip to El Dorado*, Physics today January issue p. 44.

[3] Klein M.J. 1990,*The physics of J, Willard Gibbs in his time*, Physics Today September issue p.40.

[4] Weibull W. 1951 J. Applied Mechanics Sept. issue 294.

[5] Newman W.I and Gabrielov A.M. 1991 International Journal of Fracture **50** 1.

[6] Mandelbrot B.B. 1983 *The fractal geometry of nature* W.H. Freeman and Co, QA447.M357.

[7] Montroll E.W. and Shlesinger M.F. 1982, Proc. Natl. Acad. Sci. **79** 3380.

[8] Montroll E.W. and Shlesinger M.F. 1983 Journal of Statistical Physics **39**, 203.

[9] Lydall H.F. 1959, Econometrica **27**, 110.

[10] Pareto V. 1968 *The rise and fall of the elites* Bedminster Press (Originally published as *Un applicazione di teorie sociologiche*, Revista Ital. di Sociol. 1901 402).

[11] Pareto V. 1984 *The transformation of democracy* Transaction Books, JC423.P25313, translation of *transformazione dell democrazia.* This edition contains an introductory chapter describing the life and times of Vilfredo Pareto.

[12] Shannon C E. [1948]. Bell Syst. Tech. J. **27**: 379, 623. (Reprinted in *The Mathematical Theory of Communication.* Urbana: University of Illinois Press (1949).)

[13] For further reading see Ashcroft, N. and Mermin, N. D. 1976, *Solid State Physics.* New York:Holt, Rinehart and Winston.