PH350 Classical Physics   Problem Set 4   7.9.2009

Microstates, macrostates, probabilities, & all that

A good way to understand many of the basic concepts of statistical physics, such as microstates, macrostates, accessible microstates, probabilities, etc., is to work out a *toy* model in some detail. This is the purpose of this Problem Set.

The simplest model we can construct is a system consisting of N independent elementary constituents, each of which can have only a small number of states – in fact, just two states, to keep things really simple. The obvious example of such a "two-level" or "two-state" object is a coin, which can be either in the "heads" state (denoted by h), or in the "tails" state (denoted by t). Thus we have a much simpler situation than in a gas, in which each of the molecules has a very *large* number of possible states in general. But our example is not totally unphysical — there are numerous physical systems in which the constituents have just two (or some such small number of) possible states. Examples are binary alloys, paramagnets, etc. We will discuss some of these in brief here, and in greater detail later on.

Our system is thus a set of N coins. Each coin is tossed once, and the coins are laid out in a row from coin 1 to coin N.

- A **microstate** of our system is a *detailed* specification of the state (h or t) of *each* of the N coins: e.g., $hhhh\ldots hhh$, or $hthhthtth\ldots ht$, and so on. In other words, each "word" of N letters, made up of just the two letters h and t, is a possible microstate of our system.

- A **macrostate** of the system is specified as follows. Let H and T denote the *total* number of heads and tails, respectively, in the collection. Then the pair of values (H, T) specifies a macrostate of the system. Clearly, in general, a knowledge of H and T is not sufficient to tell us whether any *particular* coin is in the h state or the t state. It is therefore obvious that a macrostate of the system provides much less detailed information about the system than a microstate does. Equivalently, much less information is necessary to specify a macrostate.

  A macrostate of our system could also be specified by the pair (N, M) where N is the total number of coins, and M is the *difference* between the total

number of heads and the total number of tails, because $N = H + T$ and $M = H - T$. Further, if N is fixed once and for all, then M alone needs to be specified to label a macrostate of the system.

A few remarks on the analogy between our toy model and a paramagnet will be helpful. In its simplest form, a paramagnetic specimen consists of a large number N of independent elementary (or atomic) magnetic dipole moments, each of magnitude μ. Suppose each such moment can align itself either parallel or anti-parallel to an applied magnetic field. Then $N = N_\uparrow + N_\downarrow$ where $N_\uparrow$ is the total number of moments parallel to the applied magnetic field, and $N_\downarrow$ is the number antiparallel to it. The net magnetic moment of the system is $\mu(N_\uparrow - N_\downarrow) = M$. (This is why we chose the symbol M for $H - T$ in our toy model!) However, two important points must be noted.

First, each atomic moment may actually point at any angle θ ($0 \leq \theta \leq \pi$) to the field, contributing $\mu \cos\theta$ (rather than just $+\mu$ or $-\mu$) to the magnetization in the direction of the field. This is true in some cases. But in other cases in which the moments must be treated *quantum mechanically*, it turns out that each atomic moment can only have a discrete set of possible orientations – in particular, just orientations parallel or anti-parallel to an applied magnetic field, as assumed above.

The second point is that, unlike a "fair" or *unbiased* coin which can be in a state h or state t with equal probabilities ($= \frac{1}{2}$), an atomic moment may actually have *different* probabilities for being oriented parallel or antiparallel to an applied field. In fact, this is just what happens in general. Later on, we shall deduce the actual values of these probabilities. It will turn out that *at sufficiently high temperatures* the two probabilities (of parallel or antiparallel orientations with respect to the applied field) tend to become equal to each other. It is this situation, therefore, that our toy model simulates. It is easy to generalise it to the case of "unfair" or biased coins, to reproduce what happens in a paramagnet at finite temperatures. For the moment, let's stick to fair coins. Our system simply consists of N *fair* coins.

1. (a) What is the total number of microstates of the system?

   (b) Are all microstates equally probable?

   (c) What is the *à priori* probability of occurrence of each microstate? ("*A priori*" means *before* the event happens, i.e., before the coins are tossed.)

   (d) What is the number of *accessible microstates* in each of the following macrostates? (i) $M = N + 1$; (ii) $M = N$; (iii) $M = N - 1$; (iv) $M = 1$; (v) $M = 0$; (vi) $M = -N$; (vii) $M = -N - 1$.

(e) Hence determine the physical range of M, for a given value of N.

(f) Find an expression for the number of accessible microstates in a given macrostate M.

(g) Hence find the **most probable macrostate** (for a given N). Observe that all the *macrostates* of the system are **not** equally probable. Compare this with the situation in (b) above.

(h) What is the probability of obtaining the most probable macrostate?

(i) What is the probability $P_N(M)$ of obtaining a general macrostate labelled by M?

(j) Analyse the variation of $P_N(M)$ with M, for a given N, in the case when N is very large. In particular, compare the probabilities of the *most* probable and the *least* probable macrostates.

Factorials of large numbers: Stirling's Formula

It is instructive to put in some actual numbers in the model above. Let's take N to be 100, a decently large number (but nowhere near 10^{23}, of course). We find that factorials occur in the expressions we have to calculate, via ${}^nC_r$, etc. Now $N!$ is an extremely rapidly increasing function of N – so rapidly increasing that standard pocket calculators won't even give you 70!, because this number exceeds 10^{99}. Fortunately, there's a very interesting and ubiquitous[1] formula called Stirling's approximation, that gives an excellent estimate of $N!$ for large values of N. This formula is

$$N! = N^N e^{-N} (2\pi N)^{\frac{1}{2}} \left\{ 1 + \frac{1}{12N} + \text{a term of order } \frac{1}{N^2} + \cdots \right\} \quad .$$

For sufficiently large N, it suffices to retain just the term 1 in the curly brackets. To see how good this approximation is, try it out for $N = 1, 2, 5, 10, 20, \ldots, 60$ – you can see how rapidly the approximate formula approaches the actual value. Clearly, even for $N = 10^3$, you can't distinguish between the LHS and RHS for all practical purposes – let alone for the case $N = 10^{23}$. A good bit of statistical physics is based on the fantastic accuracy of this formula – which is an example of what is called a "law of large numbers". It turns out that in physical applications we need the *logarithm* of the factorial, rather than the factorial itself. Stirling's formula is then

$$\ln N! \simeq (N + \frac{1}{2}) \ln N - N + \frac{1}{2} \ln(2\pi) \; + \; (\text{terms that tend to zero as } N \to \infty) \quad .$$

---

[1] It would be sad if you have only encountered this word in some GRE word list, or if you're looking it up solely for reasons connected with the GRE.

The point to remember is that $\ln N!$ is of the order of $(N \ln N - N)$ for very large N. Stirling's formula is one of the few things that are worth memorizing[2].

2. Use Stirling's formula to go over questions (a) to (j) for the case $N = 100$.

3. When N is very large, we can think of M as essentially varying continuously. (Steps of unity are neglible compared to N.) Then we could define $\Omega(M)$, the number of states (microstates) in a small range δM about the value M. What does this look like? Does it increase monotonically as M increases, the way $\Omega(E)$ does for a collection of particles? Explain the reason for the difference, if any. (Going back to the paramagnet analogy, observe that if H is the applied magnetic field, the total energy of the system *cannot be greater* than $N\mu H$ in magnitude.)

4. Density fluctuations in a gas: Consider an ideal gas of N particles in equilibrium in a container of volume V. The average number density is N/V, a constant. Call this ratio ρ. Consider a sub-volume v of the container.

    (a) Find the probability $p(n)$ that the sub-volume contains exactly n particles at a given instant of time. Verify that $p(n)$ is properly normalized, i.e., that

$$\sum_{n=0}^{N} p(n) = 1 \, .$$

This is called the **binomial distribution**, for what should be an obvious reason.

    (b) Use Stirling's formula to work out what happens to $p(n)$ in the limit in which $N \to \infty$ and $V \to \infty$, keeping the number density $\rho = N/V$ constant and finite. This is called the **thermodynamic limit** in statistical mechanics. You must show that $p(n)$ is the **Poisson distribution** in this limit, i.e.,

$$p(n) = (\rho v)^n e^{-\rho v}/n! \quad (n = 0, 1, 2, \ldots)$$

in this case. Find the (i) mean value of n, (ii) the mean squared value of n, (iii) the standard deviation of n, and (iv) the relative fluctuation in n, defined as the ratio (standard deviation)/(mean).

---

[2]GRE word lists are not, no matter what your "seniors" tell you. Mere memorization is no substitute for understanding.

(c) Suppose now that the system is a mixture of two noninteracting gases, A and B. Suppose the probability that v contains n_A molecules of gas A is given to be

$$p_A(n_A) = (\lambda_A)^{n_A} e^{-\lambda_A}/n_A! \qquad (\lambda_A = \text{positive constant.})$$

Similarly, suppose the probability that v contains n_B molecules of gas B is given by

$$p_B(n_B) = (\lambda_B)^{n_B} e^{-\lambda_B}/n_B! \qquad (\lambda_B = \text{positive constant.})$$

Assuming the gases to be independent of each other, find the probability that v contains a total of n molecules at any instant of time.

5. **Model of a random binary alloy:** Here's a very simple model of a random binary alloy, made of atoms of two different elements (like copper and zinc alloying to give brass). To give the matter a touch that might strike a familiar chord, let's call the elements Apexium (symbol, A) and Brilliantium (symbol, B). The alloy consists of N atoms of A and N atoms of B, and let's call it Aprillass[3]. For simplicity, the structure of the alloy is taken to be two interpenetrating square lattices (solid and dotted lines, as shown in the figure), referred to as sub-lattices 'a' and 'b' respectively. Each sub-lattice has exactly N sites (vertices). In the *perfectly ordered* form of Aprillass, all N atoms of Apexium sit on 'a'-type sites , and all N atoms of Brilliantium sit on 'b' type sites. (In practice this can happen only at $T = 0$ K.) At higher temperatures, the alloy gets disordered – some A atoms sit on 'b' type sites, while a corresponding number of B atoms sit on 'a' type sites, thus making Aprillass a *random* alloy. Let the number of "wrongly seated" A atoms be denoted by w; clearly, $0 \leq w \leq N$. All Apexium atoms are indistinguishable from each other; similarly for the Brilliantium atoms. It is evident that we can regard w as labelling a *macrostate* of the system, while each distinct configuration of the $2N$ atoms is one of its *microstates*. We'll get back to more on the statistical physics of binary alloys later on.

(a) What is the total number of microstates of our sample of Aprillass?

(b) What is the number of accessible microstates in a given macrostate, i.e., for a given value of w (where $0 \leq w \leq N$), how many distinct configurations can the system have?

[3]There is no unanimity of opinion on the pronunciation of this word, but some experts stress the last three letters.

(Perfectly ordered)   ○ : **A atom**   (Disordered)

☐ : **B atom**
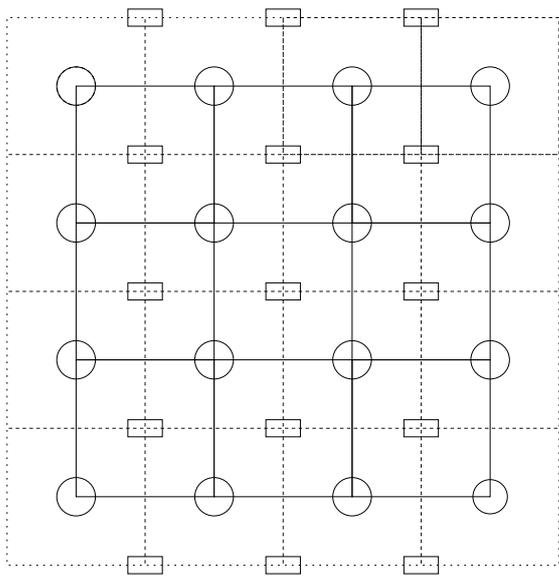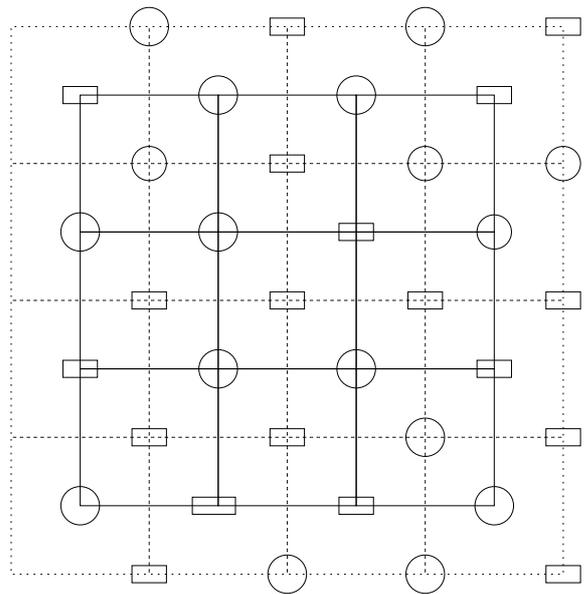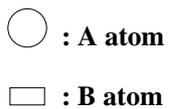
(c) Instead of using w to label a macrostate, it will turn out to be more convenient to use a variable p with the following properties: p can vary continuously between -1 and $+1$; when $p = 1$, all N atoms of element A are on sub-lattice 'a'; when $p = -1$, all A atoms are on sub-lattice 'b'. Find p in terms of w and N. What does the macrostate $p = 0$ correspond to?