

Statistical Mechanics Revision Notes

Contents

1	Introduction	1
2	Fundamentals	1
2.1	The Microstate and the Macrostate	1
2.2	The link with Entropy	2
3	Ensembles	2
3.1	Microcanonical Ensemble	2
3.2	Maximum Probability - Stirling's approximation and Lagrange multipliers	2
3.3	The Canonical Ensemble	3
3.4	Links to Thermodynamics	4
3.5	Grand Canonical Ensemble	6
4	Particle Statistics	7

1 Introduction

This revision note covers the basic ideas of statistical mechanics, starting from the quantum mechanical basis.

2 Fundamentals

2.1 The Microstate and the Macrostate

The state of a system is specified by the *microstate*: that is, the quantum mechanical state of every particle in it. Note that it is not necessary for a one-particle picture to be valid — in, for example, a strongly interacting electron system the specification of the microstate may be through a many-electron quantum mechanical state. In classical mechanics, the state of an N -particle system is expressed as a microstate by giving the $3N$ coordinates and $3N$ momenta of all the particles. The $3N \times 3N$ - dimensional space of positions q_n and momenta p_n constitute the *phase space* of the problem.

In reality, of course, we do not *observe* the microstate. What we actually measure are properties such as pressure, temperature, volume, heat capacity and so on of a large number of particles — the *macrostate*. Although the microstate evolves rapidly with time (as, for example, the particles in a fluid collide with one

another), the macrostate is essentially constant¹, and there will be a large number of microstates which correspond to a given macrostate. This is where statistics enters the picture. The fundamental assumption of statistical mechanics is that all microstates with the same energy occur with equal probability² — it follows that the state which we are most likely to observe is that which has the largest number Ω_{\max} of possible microscopic realizations³.

2.2 The link with Entropy

The entropy of the system also has the property of being maximum for a system in equilibrium. We also know that entropy is an extensive quantity (that is, for two independent systems with entropies S_1 and S_2 the total entropy is $S = S_1 + S_2$). However, if the numbers of microstates in the two systems are Ω_1 and Ω_2 the total number of microstates in the two systems taken together is $\Omega = \Omega_1\Omega_2$ ⁴. The only way these two statements can be reconciled is if $S \propto \ln\Omega$. The constant of proportionality, as we shall see later, is Boltzmann's constant.

3 Ensembles

3.1 Microcanonical Ensemble

In order to determine the most probable macrostate, we can make a large number of 'copies' of our system, each with the same macroscopic energy, volume and number of particles (E, V, N). Suppose we make \mathcal{N} such copies, and in that collection of copies ('ensemble') n_i are in microstate i . Then, of course,

$$\mathcal{N} = \sum_i n_i. \quad (1)$$

Now count the number of ways of distributing this set of microstates $\{n_1, n_2, \dots\}$ over the \mathcal{N} copies: there are $\mathcal{N}!$ different ways of numbering the copies, but of these $n_i!$ exchanges of particle within one one-particle state do not make any difference to the macroscopic quantities. Thus the number of ways of generating

¹Of course, there will in fact be *fluctuations* in the macroscopic quantities, but we may safely neglect these except for small systems or systems close to changes of phase.

²Note that this is an *assumption* — but no case has yet been experimentally demonstrated to contradict it. Note also that we assume that the dynamics of the microstates is such that all regions of phase space are sampled (strictly, that the phase space trajectory passes arbitrarily close to all points on the constant energy surface) — the *ergodic hypothesis*.

³Note the difference between the statistical and thermodynamical views: in equilibrium thermodynamics, deviations from the equilibrium state do not occur, whereas in statistical mechanics they are merely very improbable.

⁴Each microstate in system 1 can be set against all Ω_2 states of system 2.

a distribution $\{n_1, n_2, \dots\}$ is

$$W(\{n_i\}) = \frac{\mathcal{N}!}{\prod_i n_i!}. \quad (2)$$

Now we ask what is the probability of finding this distribution. Suppose the probability that one system is in microstate i is w_i : then the probability that n_i are in that microstate will be $w_i^{n_i}$. Thus the probability of finding the distribution $\{n_1, n_2, \dots\}$ will be

$$p(\{n_i\}) = \mathcal{N}! \prod_i w_i^{n_i} / n_i!. \quad (3)$$

3.2 Maximum Probability - Stirling's approximation and Lagrange multipliers

In order to find the most probable distribution, we maximise the probability given by Equation 3 with respect to the numbers n_i . In fact, rather than maximise p it is more convenient to maximise $\ln(p)$. Providing both \mathcal{N} and n_i are large we may use Stirling's approximation

$$\ln(n!) \approx n \ln n - n \quad (4)$$

to find

$$\ln p \approx \mathcal{N} \ln \mathcal{N} - \mathcal{N} + \sum_i [n_i \ln w_i - (n_i \ln n_i - n_i)] \quad (5)$$

and taking derivatives with respect to the n_i (remembering that \mathcal{N} is constant) gives

$$\sum_i (\ln n_i - \ln w_i) dn_i = 0 \quad (6)$$

in which, because of Equation 1, not all the dn_i are independent. We restore the independence by the method of Lagrange multipliers, subtracting the differential form

$$dN = \sum_i dn_i = 0 \quad (7)$$

of Equation 1 multiplied by an unknown constant λ from Equation 6 giving

$$\sum_i (\ln n_i - \ln w_i - \lambda) dn_i = 0 \quad (8)$$

in which we may regard the n_i as independent, so that the coefficient of each dn_i must be zero. Hence

$$n_i = w_i e^\lambda. \quad (9)$$

What this shows is that the probability of finding the system in the ensemble in a particular microstate i , $p_i = n_i/\mathcal{N}$, is a constant (because λ is a constant, and by assumption all microstates have equal likelihood w_i). Remember, though, that

the only microstates which are being considered here are those which have exactly the correct total energy E , volume V and number of particles N ⁵.

3.3 The Canonical Ensemble

The microcanonical ensemble has not really told us much, apart from serving as a justification of the ensemble concept. Things get moving when we look at the question of a system in which we specify the temperature T instead of the energy E of the system. In order to do this we add to our original system A a heat bath B : the heat bath is taken to be of infinite extent, and we assume that its microstates have no effect on those of the system A . We still use the ensemble concept to handle the system, but now we do not know the total energy — we must find it by averaging

$$E = \langle E_i \rangle = \sum_i p_i E_i \quad (10)$$

or, as $p_i = n_i/\mathcal{N}$,

$$\mathcal{N}E = \sum_i n_i E_i. \quad (11)$$

In equilibrium, the energy will take on some mean value, with $dE = 0$, and our set of equations for the most probable distribution now becomes

$$\sum_i dn_i = 0 \quad (12)$$

$$\sum_i E_i dn_i = 0 \quad (13)$$

$$\sum_i (\ln n_i - \ln w_i) dn_i = 0 \quad (14)$$

which we solve again by using Lagrange multipliers to obtain

$$n_i = w_i e^\lambda e^{-\beta E_i}. \quad (15)$$

We can determine the factor e^λ , because

$$\sum_i n_i = \mathcal{N} \quad (16)$$

and therefore

$$p_i = \frac{e^{-\beta E_i}}{\sum_i e^{-\beta E_i}} \quad (17)$$

which is conveniently abbreviated by introducing the canonical partition function

$$Z = \sum_i e^{-\beta E_i}. \quad (18)$$

It is essential to remember that the sum in Z is over microstates and not over energy levels.

⁵In a classical treatment, rather than the exact energy E we consider a narrow range dE of energies centred on the energy E .

3.4 Links to Thermodynamics

The partition function Z is the central quantity in statistical mechanics. We have to remember, when we use it, that the energies E_i of the microstates involved in Z will, in general, be functions of the volume of the system and the number of particles in the system.

The fact that we have an expression for the probability of each microstate means that we can go on to find the average value of any thermodynamic quantity. For example, the pressure (using P to distinguish pressure from probability p) in microstate i is

$$P_i = - \left(\frac{\partial E_i(N, V)}{\partial V} \right)_N \quad (19)$$

and so we may write the average pressure as

$$P = \sum_i P_i p_i = - \frac{\sum_i \left(\frac{\partial E_i}{\partial V} \right) e^{-\beta E_i}}{\sum_i e^{-\beta E_i}}. \quad (20)$$

Now if we differentiate the average energy

$$E = \sum_i E_i p_i = \frac{\sum_i E_i e^{-\beta E_i}}{Z} \quad (21)$$

with respect to V , keeping N and β fixed, we find

$$\left(\frac{\partial E}{\partial V} \right)_{N, \beta} = -P + \beta E P - \beta \frac{\sum_i \left(\frac{\partial E_i}{\partial V} \right) E_i e^{-\beta E_i}}{Z}. \quad (22)$$

On the other hand, if we differentiate Equation 20 with respect to β we find

$$\left(\frac{\partial P}{\partial \beta} \right)_{N, V} = -E P + \frac{\sum_i \left(\frac{\partial E_i}{\partial V} \right) E_i e^{-\beta E_i}}{Z}. \quad (23)$$

Putting these last two equations together we obtain

$$\left(\frac{\partial E}{\partial V} \right)_{N, \beta} + \beta \left(\frac{\partial p}{\partial \beta} \right)_{N, V} = -P, \quad (24)$$

where it is crucial to note that whereas the E_i are functions only of N and V the average energy E is a function of N , V and β . But thermodynamics tells us⁶ that

$$\left(\frac{\partial E}{\partial V} \right)_{N, T} - T \left(\frac{\partial P}{\partial T} \right)_{N, V} = -P \quad (25)$$

⁶The energy we are calling E is the internal energy U , and $dU = TdS - PdV$ and the Maxwell relation $(\partial P/\partial T)_V = (\partial S/\partial V)_T$ give us this result.

which we may rewrite as

$$\left(\frac{\partial E}{\partial V}\right)_{N,T} + \frac{1}{T} \left(\frac{\partial P}{\partial 1/T}\right)_{N,V} = -P \quad (26)$$

from which we may deduce that $\beta = 1/kT$, where k is some as yet unknown constant. Indeed, at present we have no reason to say that k does not vary from system to system. We can demonstrate the universality of k by considering a slightly more complex situation of two systems A and A' in thermal contact with each other and with the thermal bath. Simultaneously optimising the probability with respect to the n_i of the two systems tells us that the two partition functions are

$$Z_A = \sum_i e^{-\beta E_i} \quad (27)$$

and

$$Z_{A'} = \sum_{i'} e^{-\beta E_{i'}}. \quad (28)$$

Thus systems in thermal contact have the same β : but thermodynamics tells us that they have the same T , and therefore k must be independent of the system. From then on it is an experimental matter to select a system (such as an ideal gas) and to determine the value of this universal constant.

With the connection between thermodynamics and statistical mechanics established, we may go on to write

$$E = kT^2 \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} \quad (29)$$

$$P = kT \left(\frac{\partial \ln Z}{\partial V}\right)_{N,T} \quad (30)$$

$$S = kT \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} + k \ln Z. \quad (31)$$

The last relation is more neatly expressed via the free energy

$$F = -kT \ln Z \quad (32)$$

$$S = - \left(\frac{\partial F}{\partial T}\right)_{N,V}. \quad (33)$$

3.5 Grand Canonical Ensemble

The final elaboration is to allow the number of particles in the system to vary. Situations in which this is significant include, of course, pn junctions in semiconductors, where the number density of carriers varies as a function of position. In this situation there is a reservoir of particles which establishes the chemical

potential, much as the heat reservoir in the canonical ensemble established the temperature.

The procedure is similar to that for the canonical ensemble, except that now we need to consider configuration in which the number of particles N in each copy of the system in the ensemble is allowed to vary. We denote this by an extra subscript N , and include the different N values in the summations, the full set of equations being

$$\sum_{i,N} dn_{i,N} = 0 \quad (34)$$

$$\sum_{i,N} N dn_{i,N} = 0 \quad (35)$$

$$\sum_{i,N} E_i dn_{i,N} = 0 \quad (36)$$

$$\sum_{i,N} (\ln n_{i,N} - \ln w_{i,N}) dn_{i,N} = 0 \quad (37)$$

and using Lagrange multipliers as usual gives

$$p_{i,N} = \frac{e^{-\beta E_i + \alpha N}}{\mathcal{Z}} \quad (38)$$

where now

$$\mathcal{Z} = \sum_{i,N} e^{-\beta E_i + \alpha N}. \quad (39)$$

Note that the summation over i, N means that the sum is taken over all the microstates i for which the particle number is N , and over all the values of N .

We can make the links to thermodynamics again, using

$$TdS = dU + pdV - \mu dN \quad (40)$$

to show that

$$\alpha = \frac{\mu}{kT} \quad (41)$$

whereas

$$\beta = \frac{1}{kT} \quad (42)$$

as before.

Note that this means that \mathcal{Z} is just a weighted sum of canonical partition functions:

$$\mathcal{Z}(T, V, \mu) = \sum_{N=1}^{\infty} \left(e^{\frac{\mu}{kT}} \right)^N Z(T, V, N). \quad (43)$$

Compare this with the relationship of the canonical partition function to the microcanonical partition function (barely worth calling a partition function - it is just the degeneracy g) as the sum weighted by the Boltzmann factor

$$Z(T, V, N) = \sum_E e^{-\frac{E}{kT}} g(V, N, E). \quad (44)$$

4 Particle Statistics

So far we have said nothing about the statistics which our particles obey, that is, whether they are distinguishable or indistinguishable, classical or quantum, and if quantum whether they are Fermions or Bosons. Let us suppose that we are dealing with single particle states, labelled by k , with n_k particles in the state k with energy ϵ_k . We may define a *statistical weight* $g(\{n_k\})$ which is a function of the whole set of occupation numbers, which will depend on the type of particle. For bosons, it does not matter how many particles are in each state, and

$$g^{\text{BE}}(\{n_k\}) = 1. \quad (45)$$

For fermions, however, we must discount any arrangement in which any state contains more than one particle

$$g^{\text{FD}}(\{n_k\}) = \begin{cases} 1 & \text{if all } n_k = 0 \text{ or } 1 \\ 0 & \text{otherwise.} \end{cases} \quad (46)$$

For classical particles, the situation is a little more complicated. In principle, the particles are distinguishable, so there are $N!$ ways of changing the labels on the particles. However, we can shuffle round the labels on the particles in a particular one-particle state without producing a new microstate, so we have a statistical weight

$$g^{\text{Cl}}(\{n_k\}) = \frac{N!}{n_1!n_2!\dots}. \quad (47)$$

With this definition of statistical weight, the grand partition function may be written as

$$\mathcal{Z} = \sum_{i,N} e^{-\beta(E_i - \mu N)} = \sum_{\{n_k\}} g(\{n_k\}) \exp \left[-\beta \sum_k n_k (\epsilon_k - \mu) \right]. \quad (48)$$

For the Bose-Einstein system

$$\mathcal{Z}^{\text{BE}} = \sum_{n_1, n_2, \dots = 0}^{\infty} [\exp(-\beta(\epsilon_1 - \mu))]^{n_1} [\exp(-\beta(\epsilon_2 - \mu))]^{n_2} \dots \quad (49)$$

$$= \prod_{k=1}^{\infty} \sum_{n_k=0}^{\infty} [\exp(-\beta(\epsilon_k - \mu))]^{n_k}. \quad (50)$$

Each sum in this expression is a geometric series, with the value

$$\sum_{n_k=0}^{\infty} [\exp(-\beta(\epsilon_k - \mu))]^{n_k} = \frac{1}{1 - \exp(-\beta(\epsilon_k - \mu))} \quad (51)$$

and so

$$\mathcal{Z}^{\text{BE}} = \prod_{k=1}^{\infty} \frac{1}{1 - \exp(-\beta(\epsilon_k - \mu))}. \quad (52)$$

The expression for fermions follows from taking the upper limit on the sum to be 1 rather than infinity

$$\mathcal{Z}^{\text{FD}} = \sum_{n_1, n_2, \dots = 0}^1 [\exp(-\beta(\epsilon_1 - \mu))]^{n_1} [\exp(-\beta(\epsilon_2 - \mu))]^{n_2} \dots \quad (53)$$

$$= \prod_{k=1}^{\infty} \sum_{n_k=0}^1 [\exp(-\beta(\epsilon_k - \mu))]^{n_k} \quad (54)$$

$$= \prod_{k=1}^{\infty} 1 + \exp(-\beta(\epsilon_k - \mu)). \quad (55)$$

For classical particles we have

$$\mathcal{Z}^{\text{Cl}} = N! \sum_{n_1, n_2, \dots = 0}^1 \frac{1}{n_1! n_2! \dots} [\exp(-\beta(\epsilon_1 - \mu))]^{n_1} [\exp(-\beta(\epsilon_2 - \mu))]^{n_2} \dots \quad (56)$$

$$= N! \prod_{k=1}^{\infty} \sum_{n_k=0}^1 \frac{1}{n_k!} [\exp(-\beta(\epsilon_k - \mu))]^{n_k} \quad (57)$$

$$= N! \prod_{k=1}^{\infty} \exp[\exp(-\beta(\epsilon_k - \mu))]. \quad (58)$$

Actually, there is one extra modification to make to the classical grand partition function. We have assumed that the classical particles are distinguishable, and so they are if, for example, they are atoms in a crystal identifiable by their positions in the crystal. For atoms in a gas, though, the assumption of distinguishability leads to the Gibbs paradox⁷. To correct for this, the prefactor of $N!$ should be dropped.

The grand canonical potential may be defined by

$$\Phi(T, V, \mu) = -kT \ln \mathcal{Z}(T, V, \mu), \quad (59)$$

and from this entropy, pressure, particle number, etc. may be obtained by

$$S(T, V, \mu) = - \left(\frac{\partial \Phi}{\partial T} \right)_{V, \mu} \quad (60)$$

$$P(T, V, \mu) = - \left(\frac{\partial \Phi}{\partial V} \right)_{T, \mu} \quad (61)$$

$$N(T, V, \mu) = - \left(\frac{\partial \Phi}{\partial \mu} \right)_{V, T}. \quad (62)$$

⁷The simple model here is of a container in which identical quantities of a gas under the same conditions are placed either side of a partition which is then removed. If the particles can be identified, there will be an associated entropy of mixing. Experimentally, there is no entropy change. The way to correct this is to correct the statistical weight by a factor $1/N!$

It is clear that as the grand partition function is a product of one-particle terms, the logarithm will convert the grand canonical potential into a sum of one-particle terms. It is then straightforward to extract the individual level occupancies, because

$$N = \sum_k \langle n_k \rangle \quad (63)$$

giving, by differentiating with respect to μ

$$\langle n_k \rangle^{\text{BE}} = \frac{1}{\exp[\beta(\epsilon_k - \mu)] - 1} \quad (64)$$

$$\langle n_k \rangle^{\text{FD}} = \frac{1}{\exp[\beta(\epsilon_k - \mu)] + 1} \quad (65)$$

$$\langle n_k \rangle^{\text{CI}} = \frac{1}{\exp[\beta(\epsilon_k - \mu)]}. \quad (66)$$

Again, remember that k here labels a state: if one wishes to label by energy then one must insert an appropriate degeneracy factor.

A.H. Harker,
September 30, 2005