Griffith Sciences (GSC)

6304NSC HONOURS STATISTICAL AND CONDENSED MATTER PHYSICS

STAT MECH NOTES 2021

Howard M. Wiseman
Based on the Notes of John F. Dobson
with Contributions from Tim Gould
School of Environment and Science
Griffith University

6304NSC (STATISTICAL MECHANICS SECTION)

Approximate Timetable.

Week 7 morning	Chap. 1 (Intro), Chap. 2 (Microcanonical) …
Week 7 afternoon	… continued. Chap. 3 (Canonical)
Week 8 morning	Tutorial
Week 8 afternoon	Chap. 4 (Grand Canonical), Chap. 5 (Shannon)
Week 9 morning	Chap. 6 (Non-interacting particles) …
Week 9 afternoon	… continued
Week 10 one slot	Tutorial
Week 10 the other	Chap 7 (Thermodynamics) NON-EXAMINABLE
Week 11 morning	Chap 8 (Weakly interacting particles)
Week 11 afternoon	Chap 9 (Strongly interacting systems)
Week 12 one slot	Tutorial
Week 12 the other	Chap 10 (Non-equilibrium Dynamics) NON-EXAMINABLE
Week 13	revision
Week 14	exam

CHAPTER 1: INTRODUCTION AND REVISION OF BASIC CONCEPTS

Statistical Mechanics is the study of systems that are so complex that we normally wouldn't want to know the full details of their motion, and usually couldn't know it even if we wanted to. This includes both systems that are large in themselves, but also small systems that are in "thermal contact" with a much larger system, conventionally called the "heat bath". By "thermal contact", also known as "weak contact" or "weak coupling" etc., we mean that the energy eigenstates of the small system are not much altered by the presence of the bath, but that the bath may induce transitions among the system's states. This idea can be incorporated into a formal definition of weak/thermal coupling/contact between any two systems. This is a condition where the states of the combined system (1+2) are in one-to-one correspondence with ordered pairs of states (*i,j*) of the separate systems, where *i and j can be chosen independently* and where the energy is approximately additive:

$$E_{ij} \approx E_i^{(1)} + E_j^{(2)} \qquad \text{(Weak Coupling)} \qquad\qquad (1.1)$$

Why energy eigenstates? When you were first introduced to statistical mechanics in second-year, you were (hopefully) familiar with energy eigenstates in quantum mechanics, and these counting arguments seemed natural. Since then you have (hopefully) been introduced to the idea that it is possible to create superpositions of energy eigenstates, and that these are *non-stationary* states (at least if the superposition involves energy eigenstates of different energies). Why, then, do we consider only energy eigenstates? It's not so easy to give a rigorous answer. It is certainly related to the largeness of systems, or to the weak coupling assumption for small systems, but the exact justification is still debated. We will say more about this in later chapters.

Inherently quantum? You may recall that to apply statistical mechanics to classical systems, with their continuously variable parameters, we adopted the technique of using quantum mechanics to obtain a way of counting states in the classical continuum. (Counting states is really the heart of statistical mechanics theories.) Moreover, Finally, some approaches to the foundations of statistical mechanics rely on the properties of quantum states and make no sense when applied to classical states (i.e. particles with definite positions and momentums). Thus it can be argued that statistical mechanics is inherently a quantum mechanical subject. However, this is by no means a universally accepted proposition. Regardless, we certainly need quantum mechanics to get the observed behaviour in the low-temperature limit for systems such as black-body radiation, gases of identical particles etc.

External parameters. The discrete quantum states or levels of a system, and their energies, are determined by certain **external parameters.**

> *Example of external parameter*: magnetic field. A spin-½ particle in a magnetic field has energy levels
>
> $$E = \pm\alpha B$$
>
> where α is the magnetic moment of the spin. The magnetic field also affects *orbital* motion of charged particles, giving rise to discrete Landau levels
>
> $$E = (n+\tfrac{1}{2})\hbar\omega_c \quad (n = 0,1,2,\ldots)$$

where $\omega_c = qB/M$ is the cyclotron frequency of charge q of mass M in transverse orbital motion with a longitudinal magnetic field B.

Constraints. There are also in general **constraints** on the system, usually associated with conserved quantities:

Example 1 of constraint: energy is constant for an isolated system;
Example 2 of constraint: total (system + bath) energy is constant for small system held by heat bath at temperature T;
Example 3 of constraint: total angular momentum is constant if no external torques are present. This is important in stat mech of rotating stars.

Microstates. The system states referred to above are also called **microstates**. A microstate is a pure quantum state $|i\rangle$ of the system, specified by all the relevant quantum numbers need to be given – e.g. n, l, m, m_s. for an electron in a hydrogen atom.

Example of a microstate: Suppose our system is a collection of 5 non-interacting spin-1/2 particles, (distinguishable as they are tied to different sites on a crystal lattice). Each particle has only two states, "up" ($\uparrow$, $s_i = \frac{1}{2}$) and "down" ($\downarrow$, $s_i = -\frac{1}{2}$). [Note that here that the i subscript is *not* a label of microstates, but rather a label for individual spins. Don't assume that a symbol always has the same meaning in every equation.] In an external B-field the total energy is

$$E = 2\mu B \sum_{i=1}^{5} s_i = \mu B \left(N_\uparrow - N_\downarrow \right)$$

That is, the energy depends only on the number of spin ups and spin downs, but to specify the microstate we have to say which spins are up and down. E.g. a particular microstate ($|a\rangle$ say) has just the first spin up, $\uparrow\downarrow\downarrow\downarrow$, with energy $E_a = -3\mu B$.

Macrostates. A **macrostate** in general is a collection of 'similar' microstates, typically determined by specifying some of the external parameters (e.g. the B-field in above example) and some of the conserved quantities (e.g. total energy) $E = 2\mu B \sum_{i=1}^{5} s_i$.

Example of a macrostate: In the above system, assuming $B \neq 0$, specifying "total energy $= -3\mu B$" determines a macrostate. The 5 possible *micro*states conforming to these requirements (and therefore belonging to the *macro*state) are tabled below:

$$\uparrow\downarrow\downarrow\downarrow, \quad \downarrow\uparrow\downarrow\downarrow, \quad \downarrow\downarrow\uparrow\downarrow, \quad \downarrow\downarrow\downarrow\uparrow\downarrow, \quad \downarrow\downarrow\downarrow\uparrow$$

In large systems there are very many microstates to each macrostate. Such microstates are called the "accessible microstates" (e.g. the 5 states above): they satisfy the constraints. Obviously the concept of macrostate defined is a bit vague, because 'similar' could be in the eye of the beholder. But it generally relates to those things that are easy to measure. Speaking of measurement …

Fluctuations. If one were to *measure* some macroscopic property of a large system (or of a small system in a heat bath), this quantity will actually fluctuate in time; e.g. it is remotely

possible that the pressure on the wall of an ideal gas could be zero at some instant when no molecules happen to be hitting it: an instant later, a bunch of molecules might come in, giving a high pressure. What we are normally interested in is the average pressure over a time τ which is short on a macroscopic scale, but long on a microscopic scale so that jerky motions are averaged or "coarse-grained" away.

Ensembles. Rather than attempting to describe the exact motion of a big system and then averaging over a short time interval, a great simplification in stat mech is the notion of a **statistical ensemble**. This was introduced be the American physical scientist J. Willard Gibbs, who in fact coined the term "statistical mechanics" and "statistical ensemble" in his 1902 textbook. A Gibbsian ensemble is a fictitious collection of systems (copies), each having identical construction to the real system but with many different many microstates represented in the ensemble. In Gibbs' formulation, one considers the *relative frequency f_i* of a particular microstate $|i\rangle$ to occur in the fictitious ensemble. Since the ensemble is fictitious, we might as well assume it is infinitely large, and then we can simply say that f_i is the *probability* of finding state *i* within the fictitious ensemble describing the system at a particular time. f_i is also known as the **ensemble distribution function**.

Ensemble averages. If the operator $\hat{\theta}$ represents some property of the system, then the **ensemble average value of θ** is

$$\langle \hat{\theta} \rangle = \sum_i f_i \langle i | \hat{\theta} | i \rangle \qquad \text{(sum over all microstates } i \text{ of real system)} \qquad (1.4)$$

Note

$$1 = \sum_i f_i \qquad\qquad\qquad \text{(normalization)} \qquad (1.5)$$

Usefulness of ensemble averages. One reason that ensemble averages are useful is that for large systems (i.e. not, in this case, small systems coupled to a large bath), typical fluctuations are relatively small. Another reason is the "coarse-graining" referred to above. More formally, the **Ergodic hypothesis** states that time averages in the real system are identical to instantaneous averages over a suitable Gibbs ensemble.

Types of ensemble in common use:
> **Microcanonical ensemble** (used for isolated system, typically "large" ones i.e. having many possible microstates);
> **Canonical ensemble** (used for systems of any size which exchange energy, but not particles, with a much larger "heat bath");
> **Grand Canonical ensemble** (used for systems of any size which exchange both energy and particles, with a much larger "heat bath");

In the next chapters we will proceed to specify the equilibrium distribution function for each of these ensembles, and hence obtain a useful set of recipes for carrying out equilibrium statistical mechanics, which we will apply in later chapters. We will also look at some approaches to the foundations of stat mech, its relation to thermodynamics, and a little bit of dynamics (approach to equilibrium).

CHAPTER 2: MICROCANONICAL ENSEMBLE

The microcanonical ensemble is meant to describe thermal equilibrium of a large *isolated* system. Since energy is conserved during the motions of such a system, a macrostate can be defined by specifying the energy E (as well as *external parameters* such as volume, B-field, etc.). Thus all members i of the ensemble have energies E_i in a small range near E:

$$E < E_i < E + \Delta E \tag{2.1}$$

and the number of members in this energy range is sometimes denoted $D(E)\Delta E$. We then make the following postulate, which seem reasonable but which is ultimately justified by agreement with experiment:

> Postulate of equal *a priori* probability. For an isolated equilibrium system, all "accessible microstates", with energies satisfying (2.1), are equally likely. $\hspace{1cm}$ (2.4)

Why do we introduce any range at all, if energy is conserved? One rationale is that we often break a system down into subsystems which interact weakly. Then the total system energy is approximately the sum of subsystem energies. However, it is the true fully interacting energy which is conserved, so our sum of subsystem energies is only approximately conserved – hence a range of "energies". Actually, for typical large noninteracting quantum systems each allowed energy is highly degenerate, so we can make our range ΔE arbitrarily small but still have very many microstates because of the degeneracy. Transitions between these degenerate levels occur because the interactions aren't exactly zero.

In this quantum mechanical case, then, we simply need to look at the **degeneracy** $g_N(E)$ which is also known as the **"multiplicity function"**: it is the number of quantum microstates of an N-particle system with energy E. Thus the ensemble distribution function is

$$f_i = \begin{cases} \dfrac{1}{g_N(E)}, & i \ such \ that \ E_i = E \\[2mm] 0 & for \ all \ other \ i \end{cases} \tag{2.5}$$

$g_N(E)$ is an extremely rapidly increasing function of E, that can be calculated in principle for any particular system by counting the number of many-particle quantum states that have a given energy E. It is in general very difficult in practice to obtain an analytic expression for $g_N(E)$. It can be done, however, for simple cases such as a *weakly interacting* collection of many harmonic oscillators, or of spin-1/2 systems, or of free particles in a box. These examples were treated in the 2nd year Statistical Mechanics course.

In the 2nd year Notes we also applied the above postulate to a pair of systems in thermal contact via a rigid immovable wall. Next we postulated that thermal equilibrium of the two systems corresponds to the most likely sharing of energy between them (i.e. the sharing with the largest value for the degeneracy).

Discussion: Why might this be a reasonable postulate for equilibrium? Does nature "want to" maximize degeneracy, or is there some other explanation?

From this one can derive (see 2nd year Notes) the following prescription for obtaining thermodynamic properties of a **large system**:

Microcanonical Prescription for Calculating Internal Energy $U(T)$ and Entropy $S(T)$

(i) Find the multiplicity function $g_N(U)$ by counting many-particle quantum states.

(ii) Calculate the entropy S from $S(U) = k_B \ln[g_N(U)]$. ← *Boltzmann's epitaph.*

(iii) Find T in terms of U from $1/T = \partial S(U)/\partial U$ (note d{ln(n!)}/dn ≈ ln(n))

(iv) Invert the relationship to get U as a function of T.

(v) From this we can find experimentally verifiable quantities such as $C_V = \partial U/\partial T|_V$

Note that (iii) can be written in differential form as $dU = TdS$. In the case of the rigid wall assumed above, no work is done on the system, so $dU = dQ$. Thus $dQ = TdS$, which is the "equality" version of the **second law**, applicable for slow processes close to equilibrium.

$$(2.6)$$

<u>General form of Second Law.</u> By the same reasoning used above to "derive" thermal equilibrium, one can further postulate that an isolated system always moves towards more towards the maximum-degeneracy state which is energetically possible. Thus for an isolated system the multiplicity g, and hence the entropy $S = k_B \ln g$, can never decrease, i.e. $dS \geq 0$ (By contrast, the above "equality" form of the second law would imply $dS = 0$). This is the famous (and controversial) **arrow of time** in statistical mechanics. By considering also cases where work can be done one obtains the most general form of the second law, namely $TdS \geq dQ$.

<u>Disadvantages of the microcanonical ensemble:</u>

(i) The calculation of $g_N(E)$ and $\dfrac{\partial \ln(g_N(E))}{\partial E}$ is very difficult for most systems of interest;

(ii) The assumptions around the microcanonical ensemble makes no sense for a small isolated system (e.g. a single spin).

When such a small system is coupled to a bath, a statistical mechanical description only makes sense if it coupled to a large bath, so we can treat them together as an isolated system.
This can be done in a very general way by deriving a probability distribution for a system in contact with an arbitrary large heat bath. One then does not have to count states of the heat bath afresh every time one treats a new system. This probability distribution is known as the "canonical" or "Boltzman" distribution.

<u>"Derivation" of canonical distribution from microcanonical ensemble.</u> The following "derivation" is found in most books[1]). Consider again a system in contact with a (much larger) heat bath. If the total energy of the small system (1) plus bath (2) is fixed at ε, then the

---

[1] e.g. Reif, *Statistical Physics (SP)* 4.3; Reif, *Foundations of Statistical and Thermal Physics (FSTP)*, 6.2

probability of finding the small system in any one particular, non-degenerate, state i of energy E_i is

$$P_i = const\ g^{(2)}(\varepsilon\text{-}E_i)$$

How does P_i vary with E_i? The standard approach is to expand the "slowly varying" logarithm of P_i.

$$\ln P_i = const' + \ln g^{(2)}\left(\varepsilon - E_i\right)$$

$$= const' + \ln g^{(2)}\left(\varepsilon\right) + \frac{\partial \ln g^{(2)}}{\partial \varepsilon} \cdot \left(-E_i\right) \qquad \text{(1st order Taylor expansion)}$$

$$= const'' - \beta^{(2)}E_i$$

$$P_i = const''' \ e^{-\beta E_i} = const\ e^{-E_i/kT} \qquad T = \text{bath temp}$$

N.B.: *const''* (and *const'''*) is temperature-dependent but E_i - independent.

This "derivation" is a bit of a swindle: the only reason we got an exponential dependence on E_i is that we chose to expand the logarithm of P_i in a linear approximation.

Discussion: what if we expanded P_i itself to linear order in E_i? What about $1/P_i$?

Alternative approach: (see 2nd year Statistical Mechanics notes): by considering examples of large systems with many constituent parts (e.g. many modes of oscillation) we can look at $g_N(E)$ and argue from its basic combinatorial character that it increases approximately exponentially with energy over a substantial range of energy. More specifically, $g_N(E)$ changes rapidly with E, but $\ln g_N(E)$ changes slowly, so it makes sense to Taylor expand the latter but not the former.

CHAPTER 3: SUMMARY OF THE CANONICAL ENSEMBLE
+ SOME FURTHER DEVELOPMENTS

(See Chap 4 Notes from 2nd year SM for derivation of these results.)

Rigid thermally conducting wall

S₁

S₂ (heat bath)

Rigid thermally insulating
outer walls
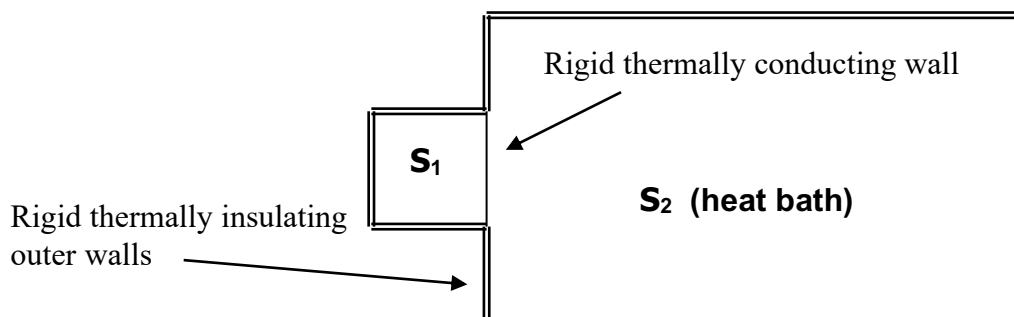
3a: SUMMARY OF THE CANONICAL ENSEMBLE

A system **S₁** of *any size* will come to equilibrium in contact with a *much larger* heat bath **S₂** characterised by its inverse temperature $\beta = (k_B T)^{-1}$.
(Compare with the *micro*canonical ensemble which applies to a *large* system in *isolation*, and which could be applied to the heat bath here.)

The energy levels E_i of the system are assumed known. The system can exchange energy with the bath and thus fluctuate between states of different energies.

The probability of state i is equal to the **Canonical or Boltzmann distribution:**

$$f_i \text{ or } p_i = Z^{-1}\exp(-\beta E_i) , \qquad (3.1)$$

where $\qquad Z = \Sigma_i \exp(-\beta E_i)$ is the **partition function**. $\qquad (3.2)$

If θ_i is the value (or, in the quantum case, the quantum expectation value) of quantity θ in state i, then the thermal average of property θ is:

$$\langle\theta\rangle = \Sigma_i \theta_i p_i : \text{ for example we have } U = \langle E\rangle = \Sigma_i E_i p_i \qquad (3.3)$$

This way of computing thermal averages as sums is particularly convenient if the system has a finite number of states i (e.g. as for one spin in a paramagnet). For other cases see (3.5-7).

Reversible heat and work: $\quad dQ^{\text{rev}} = \Sigma_i E_i \, dp_i, \quad dW^{\text{rev}} = - \Sigma_i p_i \, dE_i, \quad dU = dQ - dW \qquad (3.4)$
(Recall that dW is the infinitesimal work done **by** the system)

When the number of states is infinite it may be convenient first to calculate the infinite sum for Z, then use the following " ln Z " formulas

internal energy $\qquad\qquad\qquad\qquad\qquad$ $U = - \partial \ln Z/\partial \beta$ $\qquad\qquad$ (3.5)
pressure (derived from $dW = PdV$) $\qquad$ $P = k_B T \partial \ln Z/\partial V$ $\qquad$ (3.6)
Helmholtz free energy (from $dQ = TdS$) $\quad$ $F = U - TS = - k_B T \ln Z$ $\qquad$ (3.7)

Note: the expression for F (also written as A) can then be used to get entropy S using (3.5). It also can be used to write an alternate expression for the probability: $p_i = \exp[-\beta(F - E_i)]$.

3b. Work and generalised forces in the Canonical ensemble (Reif, *FSTP* 3.8,3.9)

The external parameters affecting the set of allowed quantum energy levels $\{E_i\}$ of a system are often denoted $\{x_j\}$. For example, consider a gas particle with nonzero spin and nonzero dipole moment in a box of volume V, with an applied electric field $\mathbf{E}$ and a magnetic field $\mathbf{B}$. Ignoring the vector nature of the fields (except for continuing to use bold font so as to avoid confusion between $\mathbf{E}$ and the energy E), we could label these external parameters as
$$x_1 = V, \ x_2 = \mathbf{E} \ , \ x_3 = \mathbf{B}.$$
Then every energy level of a gas atom could be written as

$$E_i(V,\mathbf{E},\mathbf{B}) = \boldsymbol{p}_i^2/(2m) - d_i\,|\mathbf{E}| - \mu_i\,|\mathbf{B}| \equiv E_i(x_1,x_2,x_3) = E_i(\{x_j\}).$$

Here the atom has linear momentum $\boldsymbol{p}_i$ (whose quantized values depend upon the volume V), electric dipole moment projection d_i along the $\mathbf{E}$-field, and spin magnetic moment projection μ_i along the $\mathbf{B}$-field.

In general, the infinitesimal work done **by** the system was identified as that part of the **loss** of internal energy coming from changes in the allowed energy levels $\{E_i\}$:
$$dW = -\sum_i f_i dE_i, \quad dQ = \sum_i E_i df_i \tag{3.8}$$
Since the $\{E_i\}$ are affected by changes in the $\{x_j\}$ (but not, for example, by changes in temperature) we can write the "total differential"
$$dE_i = \sum_j \frac{\partial E_i}{\partial x_j} dx_j \tag{3.9}$$
and then the work done can be written

$$dW = -\sum_i f_i dE_i = -\sum_i f_i \left(\sum_j \frac{\partial E_i}{\partial x_j} dx_j\right) = \sum_j X_j dx_j \tag{3.10}$$

where the jth *generalised force* is defined as
$$\boxed{X_j = -\sum_i f_i \frac{\partial E_i}{\partial x_j}} \tag{3.11}$$

First example of generalised force: the pressure P, or generalised force corresponding to external parameter x_1 = volume V.

$$P = X_1 = -\sum_i f_i \frac{\partial E_i}{\partial x_1} = -\sum_i f_i \frac{\partial E_i}{\partial V} \tag{3.12}$$

Second example of generalised force: the polarisation (macroscopic electric dipole moment) D, or generalised force corresponding to external parameter x_2 = $\mathbf{E}$ (electric field).

$$D = X_2 = -\sum_i f_i \frac{\partial E_i}{\partial x_2} = -\sum_i f_i \frac{\partial E_i}{\partial \mathbf{E}}.$$

Third example of generalised force: the magnetisation (macroscopic magnetic moment) M, or generalised force corresponding to external parameter x_2 = $\mathbf{B}$ (magnetic field).

$$M = X_3 = -\sum_i f_i \frac{\partial E_i}{\partial x_3} = -\sum_i f_i \frac{\partial E_i}{\partial \mathbf{B}}$$

The work done by the gas (per particle) in the canonical ensemble is then

$$dW = \sum_j X_j dx_j = PdV + Dd\mathbf{E} + Md\mathbf{B}$$

The generalized forces can also be written as derivatives on the partition function:

$$X_j = -\sum_i f_i \frac{\partial E_i}{\partial x_j} = -Z^{-1}\sum_i \exp(-\beta E_i)\frac{\partial E_i}{\partial x_j} = Z^{-1}\beta^{-1}\sum_i \frac{\partial}{\partial x_j}\exp(-\beta E_i) = \beta^{-1}Z^{-1}\frac{\partial Z}{\partial x_j} = k_\mathbf{B} T \frac{\partial(\ln Z)}{\partial x_j}$$

i.e.
$$\boxed{X_j = k_B T \frac{\partial(\ln Z)}{\partial x_j}}$$
(3.13)

Example: pressure due to 1 particle of an ideal monatomic gas (see 2nd year notes Chap. 4).

$$Z = V / \lambda_{\text{th}}^3$$

where the thermal de Broglie wavelength λ_{th} is independent of volume. Thus

$$P = k_B T \frac{\partial \ln Z}{\partial V} = k_B T \frac{\partial}{\partial V}[\ln V - 3\ln\lambda_{\text{th}}] = k_B T \frac{1}{V} + 0.$$

so the pressure for N particles satisfies $PV = N k_B T$ as expected.

3c A generalised expression for entropy

To recap, deriving (3.5) is easy. The preceding section gave a general derivation which includes (3.6). But we have not yet derived (3.7), which is the simplest way to obtain an expression for entropy, via

$$S = U/T - F = U/T + k_B \ln Z$$
(3.14)

One way can be found by integrating $\int dQ/T$ as in the rather complicated method in the 2nd year Notes Chap. 4, Sec. 4e, using the expression (3.4) above for dQ and integrating β from ∞ to the actual value. A simpler derivation is given below which also has the advantage that it yields an expression for the equilibrium entropy in a form is simply a function of the f_i's. Thus, although it is derived for equilibrium systems, it could be applied to *nonequilibrium* systems, where the probability (i.e. the proportion of time) that it will occupy the microstate i doesn't decay exponentially with E_i as it would according to the equilibrium Boltzmann distribution.

Since $E_i = -k_B T \ln(\exp(-E_i/k_B T))$ by definition, we have $\boxed{E_i = -k_B T \ln(f_i Z)}$. We can thus write the reversible heat change as

$$T\, dS = dQ = \Sigma_i df_i\, E_i \qquad\qquad \text{from (3.8)}$$

$$= \Sigma_i df_i\, (-k_B T)(\ln f_i + \ln Z) = -k_B T\, (\Sigma_i \ln f_i\, df_i + \ln Z\, \Sigma_i df_i)$$

$$= -k_B T\, \Sigma_i \ln f_i\, df_i \qquad\qquad \text{since } \Sigma_i\, df_i = 0 \text{ by normalization.}$$

Then $dS = -k_B\, \Sigma_i \ln f_i\, df_i = -k_B\, \Sigma_i\, [d(f_i \ln f_i) - \Sigma_i(f_i/f_i)\, df_i] = -k_B\, \Sigma_i\, d(f_i \ln f_i) = d(-k_B\, \Sigma_i(f_i \ln f_i))$

since, again, $\Sigma_i\, df_i = 0$.

Thus we interpret entropy

$$\boxed{\mathbb{S} = -k_B \sum_i f_i \ln(f_i)}$$
(3.15)

$\mathbb{S}$ (engraved S) has been used for this formula since it will be used for arbitrary distributions f_i, not just the equilibrium (Boltzmann) distribution. $\mathbb{S}$ is sometimes called the "**Shannon**

entropy", as will be discussed in Chap. 5. The regular symbol S will be reserved for the entropy of an *equilibrium* system (i.e. one with a well-defined temperature).

Note that, at $T = 0$, the entropy $S = 0$, provided that the ground state is non-degenerate.

Exercise: show this using the expression for $\mathcal{S}$. Remember that $\ln(1)=0$ and $\lim_{x \to 0} x \ln(x) = 0$.

For nonzero temperature, the thermal entropy S can be evaluated more generally using the equilibrium Boltzmann distribution for f_i. Using the useful relation $-E_i / k_B T - \ln(Z) = \ln(f_i)$,

$$S = -k_B \sum_i f_i [-E_i/(k_B T) - \ln Z] = + T^{-1} \sum_i f_i E_i + k_B \ln Z = U/T + k_B \ln Z.$$

This relationship is equivalent to that given in Eq. (3.7),

$$A = F \equiv U - TS = -k_B T \ln Z$$

where the Helmholtz free energy is denoted A or F.

3d. Properties of the Shannon entropy

Here we summarise some properties of $\mathcal{S} = - k_B \Sigma f_i \ln f_i$.

(i) $\mathcal{S} = $ "$< -k_B \ln f >$". Formally, the expectation of $k_B \ln f$ (but note that $k_B \ln f$ isn't a usual property of the system because it depends on the statistical distribution, unlike, say, the position operator $\hat{\mathbf{x}}$). The quantity $(-\ln f)$ actually has a name – the "surprisal". (3.16)

(ii) $\mathcal{S}$ reduces to S (the thermodynamic entropy) in equilibrium. In fact we <u>constructed</u> $\mathcal{S}$ to ensure this for the <u>canonical</u> distribution $f_i = Z^{-1} \exp(-E_i/k_B T)$. (3.17)

(iii) For the equilibrium <u>microcanonical</u> ensemble at fixed internal energy U we have

$$f_i = \begin{cases} \dfrac{1}{g_N(E)}, & i \text{ such that } E_i = E \\ 0 & \text{for all other } i \end{cases}$$

Thus $\mathcal{S} = -k_B \dfrac{1}{g_N(U)} \sum_{i:E_i=U} \ln(\dfrac{1}{g_N(U)}) = -k_B \dfrac{1}{g_N(U)} (g_N(U))(-\ln(g_N(U)) = k_B \ln(g_N(U))$.

(3.18)

Thus $\mathcal{S}$ gives the thermodynamic entropy S also for the microcanonical ensemble.

(iv) $\mathcal{S}$ is a <u>measure of chaos</u>. That is, it is bigger the more "spread-out" the distribution. It is always positive, because $0 \leq f_i \leq 1$ so $\ln f_i \leq 0$. It is zero if the system is with certainty in one particular microstate (such as the groundstate). (3.19)

(v) $\mathcal{S}$ is a <u>measure of ignorance</u>, or, equivalently, <u>amount of information</u> (3.20) required to determine <u>which</u> microstate the system is actually in, for a given macrostate (distribution function f). As an example, consider the microcanonical case where $\mathcal{S} = k_B \ln(g)$. Here $g = \#$ of microstates the system could be in, consistent with the macrostate (determined

in this case solely by the total energy U). Why is it the <u>logarithm</u> of g which measures information? Consider amount of computer storage (information) needed to specify numbers of a certain size, g. This is determined by number of decimal (or, more fundamentally, binary) places required, and this goes as $\log g$. For example, to specify a number up to $g=10^8$, we need $\log_{10}(10^8) = 8$ decimal locations (one for each digit). This corresponds to $\log_2(10^8) \approx 27$ binary digits (bits). It is possible to base the entire equilibrium theory of statistical mechanics on this <u>information theoretic</u> approach. We will outline this approach in Chap. 5.

(vi) $\mathcal{S}$ <u>is additive</u> when two systems are simply put together without interaction and treated as one system. For the joint system has states (i,j) with probability $f_{i,j} = f_i^{(1)} f_j^{(2)}$.

<div style="border:1px solid #000; background:#dbe5f1; padding:4px;">

Exercise: show that this leads to $\mathcal{S} = \mathcal{S}^1 + \mathcal{S}^2$. (3.21)

</div>

(vii) The equilibrium canonical distribution f_i is the solution of the problem of minimizing the so-called "trial Helmholtz free energy"

$$F = U - T\mathcal{S} = \sum_i (f_i E_i + k_B T.f_i \ln f_i) \tag{3.22}$$

Here 'trial' means we allow for an arbitrary distribution, subject only to the normalization constraint $\Sigma_i f_i = 1$. This is easily proved using Lagrange multipliers (see Ch 5). Minimization of F can be viewed as a struggle between the forces of order (1st term is lowest in an ordered low-energy state) and chaos (2nd term is lowest in a chaotic state). Chaos wins at high temperature because of factor T. This minimum principle is very useful for interacting systems, as we will see later.

3e. Fluctuations in the canonical ensemble

A property of a system S, such as its energy E, is not fixed when S is in contact with a heat bath B. This occurs because of the interaction between S and the bath B – simplistically, we can think of B as continually "kicking" system S, causing properties of S to vary over time. An alternative way to conceptualize fluctuations is variation between copies in the Gibbs ensemble. Thus, a measurement of a quantity θ at some chosen time yields unpredictable results, even in classical physics. (Of course in quantum physics there is the additional unpredictability associated with the intrinsic quantum uncertainties in variables (other than energy) when the system is in a certain quantum microstate $|i\rangle$.)

A common way to measure how much a quantity θ varies is to consider the **mean square fluctuation of θ.** Recall that the thermal average of a quantity is $\langle\theta\rangle \equiv \Sigma_i \theta_i f_i$, where θ_i is the value (or, in the quantum case, the quantum expectation value) of quantity θ in state i. The *thermal* fluctuation, $\Delta\theta$, of θ is defined as its departure from this thermal average:

$$\Delta\theta = \theta - \langle\theta\rangle \tag{3.23}$$

Notice that the mean fluctuation is zero, because fluctuations above and below the mean are equally likely in equilibrium:

$$\langle\Delta\theta\rangle = \Sigma_i f_i (\theta_i - \langle\theta\rangle) = \langle\theta\rangle - (\Sigma_i f_i)\langle\theta\rangle = \langle\theta\rangle - 1\langle\theta\rangle = 0 . \tag{3.24}$$

However the **squared fluctuation** $(\Delta\theta)^2$ cannot be negative, so its thermal average is typically positive. The ensemble average of this quantity is the **mean square fluctuation**

$$\left\langle (\Delta\theta)^2 \right\rangle = \sum_i f_i (\Delta\theta_i)^2 = \sum_i f_i \left(\theta_i - \langle\theta\rangle \right)^2 \tag{3.25}$$

and this is a sensible measure of the variability of the quantity θ.

The square root if this quantity is the Root of the Mean of the Square of the fluctuation, (the **RMS** fluctuation) and this is convenient because it has the same dimensions as the quantity θ itself. The RMS is a familiar concept from the theory of a.c. electric circuits.

$$\theta_{RMS} = \sqrt{\left\langle (\Delta\theta)^2 \right\rangle} \tag{3.26}$$

Examples calculating this quantity in the Canonical and Grand ensembles are given on **Problem Sheets 1 and 2**. A decomposition that is sometimes useful is as follows:

$$\begin{aligned}
\left\langle (\Delta\theta)^2 \right\rangle &= \sum_i f_i \left(\theta_i - \langle\theta\rangle \right)^2 = \sum_i f_i \left(\theta_i^2 - 2\langle\theta\rangle\theta_i + \langle\theta\rangle^2 \right) \\
&= \left\langle \theta^2 \right\rangle - 2\langle\theta\rangle^2 + \langle\theta\rangle^2 \\
&= \left\langle \theta^2 \right\rangle - \langle\theta\rangle^2
\end{aligned} \tag{3.27}$$