Lecture 5: Statistical Equilibrium Ensembles

1   Introduction

It is clearly important to know what the probability density is for a system in equilibrium. This will be a function of the microstate i.e. phase space variables. Given such a density we can then calculate the system entropy using the Gibbs/Shannon entropy functional on such densities and from this determine most thermodynamical variables of interest in the manner described in the last two lectures. This is the basic approach of equilibrium statistical mechanics which derives a system thermodynamics from the properties of the constituent fine grained elements. For materials this means molecules but can sometimes be extended to more general situations such as fluids.

The particular form of this target probability density will depend obviously on the energetics of the molecules i.e. the precise form of $E(\boldsymbol{p}, \boldsymbol{q})$. It will also depend on the way in which equilibrium is set up for the large scale system. Historically three different configurations have been considered:

1. Completely closed to the environment. Here conserved quantities for the dynamical system will be constant in time. Such a configuration is termed <u>microcanonical</u>.

2. Open to exchanges of energy only with the environment. In this case if the system reaches equilibrium then the total energy will show very small fluctuations from a mean value as discussed in previous lectures. This is due to the energy interchange with the environment. Such a configuration is termed <u>canonical</u>.

3. Open to exchange of energy and molecules with the environment. This is an extension of the previous case. Here <u>both</u> the total energy and number of particles will be subject to small fluctuation about means of both quantities. This configuration is termed <u>grand canonical</u>.

Now as time proceeds in these three equilibrium configurations the microstates will not be fixed in general which is the reason for seeking an equilibrium probability density. We hypothesise that this exists and is time invariant. It is sometimes referred to as an invariant measure for the system i.e. a measure which remains the same under time translations. A (large) sample of microstates drawn from such a density is commonly called an ensemble and was introduced as an analytical tool originally by Gibbs. Behind it however lies the central concept of an equilibrium density or invariant measure. The densities for the three different configurations above differ and the last two are in practise the easiest to compute. In the limit that the number of molecules becomes infinite they are supposed to converge to each other (this is called the thermodynamical limit).

2 Microcanonical ensemble

The closed property for this case ensures that the system lies always in a particular submanifold of phase space. This is specified by the values of invariants for the system with energy being the most commonly considered. We have therefore

$$E(\boldsymbol{p}, \boldsymbol{q}) = E_0$$

and this defines a generally closed and compact submanifold C of phase space of dimension $2n - 1$ where n is the number of particles. Since the system is so constrained then all probability densities (equilibrium or otherwise) must be defined as non-zero only on this submanifold. The $2n-1$ dimensional volume for C is of central significance to this situation and we denote it by $V(C)$ and have

$$\int_C 1 \equiv V(C)$$

Now suppose we have two different densities p and q defined on C then their relative entropy is

$$D(p||q) = \int_c p \log \frac{p}{q} \geq 0.$$

Now select q to be uniform on C. This makes sense if C is compact of finite $2n - 1$ dimensional volume. Now it follows from the relative entropy non-negativity that

$$\int_c p \log p \geq \log q$$

i.e.

$$S(p) \leq \log \frac{1}{q} = S(q)$$

where S here is the Shannon entropy. This immediately implies that q is the density of maximum entropy for the system and therefore by the principle outlined in the third lecture the equilibrium density for the microcanonical ensemble. It is further easily seen that

$$
\begin{aligned}
q &= \frac{1}{V(C)} \\
S(q) &= \log\left[V(C)\right]
\end{aligned}
$$

This formulation of entropy was first proposed by Boltzmann in a slightly different form to take into account the units of phase space volume and to set the temperature unit to degrees rather than joules.

Boltzmann hypothesised that the maximum entropy density uniform on C was actually achieved in equilibrium. One can actually examine this issue in tractable simple dynamical systems and it turns out that the situation is somewhat more complicated than this: One needs to be rather careful from a mathematically technical viewpoint in stating this result. Indeed there can exist

members of the microcanonical ensemble that are never visited in any equilibrium configuration. These have however Lebesgue measure zero. The hypothesis that all microcanonical ensemble members are "equally likely" was first stated by Boltzmann and was called the ergodic hypothesis. It was shown to be strictly untrue around 1913 but was replaced by the quasi-ergodic hypthosesis which very roughly means any coarse graining of the submanifold C with equal Lebesgue measure will have equal probability measure. This was proven by Birkhoff and Von Neumann in the 1930s. More details in the review [1].

3   Canonical ensemble

The calculation of the entropy in the case of the microcanonical ensemble can be difficult since it involves the volume of a constraint submanifold. A somewhat easier approach mathematically comes with the more practically realistic case of the canonical ensemble where energy exchange is permitted with the environment. As a consequence when equilibrium is achieved total energy fluctuates slightly since molecules inside and outside the system of interest are always exchanging energy. Some reflection indicates however that as the number of molecules gets larger such fluctuations become smaller. We analyze this further in later lectures. In the "thermodynamical limit" as the number of molecules becomes infinite these fluctuations approach zero and the canonical and microcanonical ensembles converge.

We make the assumption that the equilibrium density is invariant under time translations which seems reasonable for a statistical equilibrium. It follows now that the expectation value of total energy with respect to the equilibrium density is also invariant[1]. From this we can derive the form of the equilibrium density using the maximum entropy principle as we did for the microcanonical case. We do this using Lagrange multipliers: Suppose the required probability density is $\varrho(\boldsymbol{p}, \boldsymbol{q})$. Assume that the following relations hold

$$\int \varrho(\boldsymbol{p}, \boldsymbol{q}) E(\boldsymbol{p}, \boldsymbol{q}) = E_0 \tag{1}$$

where E_0 is the fixed mean energy of the ensemble and the integral is over all phase space not just the constraint submanifold as in the last section.

The problem then becomes to find $\varrho(\boldsymbol{p}, \boldsymbol{q})$ subject to the requirement that the associated differential entropy S is maximized and (1) is met as well as the normalization condition:

$$\int \varrho(\boldsymbol{p}, \boldsymbol{q}) = 1$$

Let the Lagrange multiplier for the energy constraint be λ_1 and that for the normalization condition be λ_0 and maximize the augmented functional with respect to ϱ

$$J \equiv -\int \varrho \log \varrho + \lambda_1 \left(\int \varrho E - E_0 \right) + \lambda_0 \left(\int \varrho - 1 \right)$$

---

[1]Put another way, the canonical ensemble mean energy is fixed.

Taking the functional derivative with respect to ϱ and setting to zero we get

$$\frac{\delta J}{\delta \varrho} = -\log \varrho - 1 + \lambda_1 E + \lambda_0 = 0$$

which implies that the maximum entropy distribution must have the form

$$\varrho(\boldsymbol{p}, \boldsymbol{q}) = Z^{-1}(\lambda_1) \exp\left(\lambda_1 E(\boldsymbol{p}, \boldsymbol{q})\right) \qquad (2)$$

$$Z(\lambda_1) \equiv \int \exp\left(\lambda_1 E(\boldsymbol{p}, \boldsymbol{q})\right) d\boldsymbol{q} d\boldsymbol{p} = \exp(1 - \lambda_0) \qquad (3)$$

where the partition function Z is need to ensure that ϱ is normalized as a density. Notice that this maximum entropy density has precisely the form deduced in Lecture 2 using the steady Liouville equation and the additivity of system invariants across subsystems. The density in equation (2) is of central importance to equilibrium statistical mechanics and is called the Gibbs density. Empirically it can be shown to apply in many different systems of practical interest. It is particularly convenient since it requires only knowledge of the system total energy as a function of phase space co-ordinates.

Using Lagrange multipliers we have established that the Gibbs density is a neccessary condition for a maximum entropy density but is it sufficient? The information theoretic argument given at the end of Lecture 2 actually establishes this since it shows that the entropy of any density with mean energy E_0 is bounded above by that of the Gibbs density with equality iff they are equal.

Notice that there is a different Gibbs density for each choice of λ_1. We can interpret this parameter by holding it fixed and calculating the entropy:

$$\begin{aligned} S - S_{ref} &= -\int \varrho \log \varrho \\ &= \langle \log Z - \lambda_1 E \rangle_\varrho \\ &= \log Z - \lambda_1 E_0 \end{aligned} \qquad (4)$$

Now the E_0 is the thermodynamical energy discussed in Lecture 3 so for the equilibrium system under consideration

$$\frac{dS}{dE_0} = \frac{1}{T} = -\lambda_1 \qquad (5)$$

thus the Lagrange multiplier is simply minus the reciprocal of the system temperature. Note also that the partition function Z is an important part of the entropy and is a non-obvious function of the inverse temperature via equation (2) and the total energy function of the phase space co-ordinates.

4 Grand canonical ensemble

This is the most realistic situation in which the open system exchanges both energy and molecules with the environment but has a steady equilibrium density.

The mathematical formulation is a little more technically complicated since the phase spaces of systems with different numbers of molecules have differing dimensions but this is not important to the final result obtained: Define an outcome space

$$\boldsymbol{X} = \mathbb{N} \otimes R^\infty$$

and a probability density P which satisfies the restriction

$$P(N, \boldsymbol{x}) = 0 \qquad when \ \boldsymbol{x} \notin R^{2N} \subset R^\infty \tag{6}$$

where N is the number of molecules and R^{2N} consists of real sequences with zeros beyond the $2N$ position.

A further complication concerns the phase space of identical particles: If one permutes identical particles then the microstate obtained remains the same. To avoid multiple counting of the same microstate, one divides phase space volumes by the number of possible permutations i.e. by $N!$. The concept of indistinguishability of particles is fundamentally a quantum one since in classical mechanics one could in principle label each molecule and keep track of them. This is not possible in quantum mechanics which of course is the more fundamental (and correct) physical theory[2]. We therefore weight volume elements in phase space as follows

$$dV = \frac{d^N \boldsymbol{p} d^N \boldsymbol{q}}{\hbar^{3N} N!}$$

where Planck's constant $\hbar$ has also appeared for dimensional consistency as discussed in Lecture 1. Now the expected (mean) number of molecules $\overline{N}$ is then

$$\overline{N} = \sum_{N=0}^{\infty} \int_{R^{2N}} N P(N, \boldsymbol{x}) dV$$

while the expected energy $\overline{E}$ is

$$\overline{E} = \sum_{N=0}^{\infty} \int_{R^{2N}} E(N, \boldsymbol{x}) P(N, \boldsymbol{x}) dV$$

and the normalization condition for the density P is

$$1 = \sum_{N=0}^{\infty} \int_{R^{2N}} P(N, \boldsymbol{x}) dV$$

For a time invariant P these expectations must always be fixed and so macrostates are specified using both the mean number of molecules and the

---

[2]Note that this issue is not important for the other two ensembles discussed earlier since there this additional factor is a constant. Note also that quantum mechanics is making an unavoidable intrusion into our basically classical presentation. In this course we make little mention of quantum statistical mechanics which can in certain situations be very important. Our presentation relies on the correspondence principle of quantum theory which says that under many other circumstances classical theory is asymptotically appropriate.

mean energy. We can now solve the maximum entropy constrained optimization problem with three constraints[3] rather than two and obtain in a similar way to the canonical ensemble

$$-\log P - 1 + \lambda_1 E(N, \boldsymbol{x}) + \lambda_2 N + \lambda_0 = 0$$

Thus we obtain

$$P(N, \boldsymbol{x}) = Z_G^{-1} \exp\left(\lambda_1 E(N, \boldsymbol{x}) + \lambda_2 N\right) \qquad (7)$$

$$Z_G(\lambda_1, \lambda_2) \equiv \sum_{N=0}^{\infty} \int_{R^{2N}} \exp\left(\lambda_1 E(N, \boldsymbol{x}) + \lambda_2 N\right) dV \qquad (8)$$

In a similar fashion to the canonical ensemble P may be shown to achieve the maximum entropy. Also in a manner similar to the canonical ensemble we can calculate the entropy and after differentiation by the mean number of molecules obtain

$$\lambda_2 = \frac{\mu}{T} \qquad (9)$$

hence this second multiplier is related to the chemical potential of the system.

5 Partition function and thermodynamics

As we have seen for the canonical ensemble $Z(\lambda_1)$ is an important component of the Gibbs density. It can be related in an interesting way to a thermodynamical potential we met in Lecture 4. Combining (4) and (5) we get up to a constant reference entropy[4] that

$$\log Z = (TS - E_0)/T = -F/T \qquad (10)$$

where F is the Helmholtz free energy thermodynamical potential introduced in the previous lecture and because the mean energy E_0 is the thermodynamical energy. Now in general Z will depend on both T since this is minus the inverse Lagrange multiplier, as well as the volume V of the system since the integral used to obtain Z depends on integrals over the locations of molecules. Thus if we are able to calculate Z for a particular system then we also have the free energy F as a function of T and V. We are then able to compute pressure and entropy by simple partial differentiation and we have the full set of thermodynamical variables for the system. A similar procedure extends this to the grand canonical ensemble where the chemical potential makes an appearance as a further thermodynamical variable. For the canonical ensemble the calculation of Z depends crucially on the form of the total energy which for many systems can be decomposed into two pieces:

$$E(\boldsymbol{p}, \boldsymbol{q}) = K(\boldsymbol{p}) + V(\boldsymbol{q})$$

---

[3]Strictly speaking we should impose further constraints to ensure (6) is met but these can easily be seen to amount to imposing these conditions a posteriori on the maximum entropy P.

[4]This can be eliminated by redefining Z by dividing by a constant

i.e. into kinetic and potential energy. The kinetic energy is a simple quadratic form for the momenta and so the integral over this part of phase space is rather easy since it is the integral of a Gaussian function. Most of the work then usually occurs in integrating the potential energy over the molecular position variables. For an ideal gas this potential term vanishes. We consider this simple case in the next lecture.

References

[1] J.L. Lebowitz and O. Penrose. Modern ergodic theory. *Physics Today*, 26(2):23–29, 1973.