

Lecture 3: Thermodynamics I

1 Introduction

Historically the way in which statistical physics developed was by first observing the macrostate of systems of many particles and deriving functional relations between large scale variables of significant interest to the behavior of the system. This field of empirical study is referred to generically as thermodynamics. Later as the molecular view of matter gained credibility amongst physicists an important topic was the justification of these macrostate rules using the known properties of the microscopic molecules. In order to understand this subject then it is useful to first gain an understanding of the basis and form of thermodynamics. As we noted in previous lectures the bulk variables for open systems such as the total energy are not constant but instead are subject to usually very small variations which are termed fluctuations. The assumption behind the study of thermodynamics is that such fluctuations can be ignored because the systems of interest consist of enough molecules to make them negligible. Macrostates can be specified in general by the mean of the bulk variables which for an equilibrium system are conserved quantities. In the limit that the number of molecules becomes infinite the fluctuations vanish and the macrostate specification is the same as specifying the bulk variable. In a later lecture we look at the fluctuations but for now we ignore them.

2 The principle of increasing entropy

Consider a closed system consisting of a series of (open) macroscopic subsystems with equal numbers of molecules. Assume that each subsystem is in equilibrium internally but not with each other¹. As argued in the previous lecture each subsystem will have (to a good approximation) a Gibbs distribution which in the simplest case is specified by the subsystem mean energy. Assume that these mean energies are not all the same. From the Gibbs distribution we can deduce the entropy of each subsystem using the formula derived in the first lecture. Now if we further assume², as we did in the last lecture, that the joint probability distribution for the entire closed system is the product of all the individual Gibbs distributions then it follows that the total entropy of the system is the sum of the subsystem entropies i.e. entropy is an additive macroscopic variable like energy.

As time proceeds it is empirically observed that the total system entropy monotonically increases³ and in equilibrium a maximum is achieved in which

¹The system must be considered on a sufficiently long time scale for this to be the case. Such a time scale is generally rather short relative to the timescale with which the system as a whole relaxes to equilibrium.

²This assumption may be verified empirically to be approximately correct.

³The fluctuations discussed previously may lead to a very small violation of this law. Thermodynamics is formally the limit that these fluctuations vanish.

the mean energies are all equal. This is commonly known as the second law of thermodynamics and is a central organizing principle in thermodynamics. It is important to note that this principle is appropriate only for a closed i.e. isolated system. In an open system entropy may be added from the system to the environment resulting in a local decrease but an overall total increase. The second law is an empirical fact of macroscopic systems and no violation has ever been observed⁴. Actually “proving” it is so from first principles is non-trivial. Indeed the second law implies an arrow in time which is absent from the classical dynamical equations since these are invariant under a reversal of time. This has been a topic of endless fascination since statistical mechanics was first proposed by Boltzmann and it is fair to say still remains not fully understood. When we deal with non-equilibrium systems later in the course we shall revisit this issue.

3 Temperature

To date we have met two macrovariables of interest namely the energy and entropy. As we noted in the previous section a system in equilibrium is characterized by having a maximum entropy if we take the second law of thermodynamics as given. Consider now a closed system consisting of two open systems that are in equilibrium with each other. The entropy and energy of each subsystem being denoted by E_1, E_2 and S_1, S_2 respectively with $E = E_1 + E_2$ and $S = S_1 + S_2$ being the total energy and entropy of the combined system and E is a constant. Consider now the simple case that entropy is solely a function of energy (we extend this later):

$$S_i = S(E_i)$$

Now for S to be a maximum i.e. the total system to be in equilibrium we require

$$\frac{dS}{dE_1} = 0$$

since there is only one independent energy variable. However we have

$$\begin{aligned} \frac{dS}{dE_1} &= \frac{dS_1}{dE_1} + \frac{dS_2}{dE_2} \frac{dE_2}{dE_1} \\ &= \frac{dS_1}{dE_1} - \frac{dS_2}{dE_2} \end{aligned}$$

because total energy is conserved. In other words

$$\frac{dS_1}{dE_1} = \frac{dS_2}{dE_2}$$

This argument generalizes to an arbitrary number of systems pairwise in equilibrium with each other and so we get the condition that

$$\frac{dS_i}{dE_i} = \text{Constant} \equiv \frac{1}{T} \tag{1}$$

⁴Apart from the fluctuation effect.

For all systems to be pairwise in equilibrium. Thus the derivative of entropy with respect to energy is a fundamental marker of equilibrium and the inverse of this constant is called the temperature T of the total system in equilibrium. Recall now the way we have defined entropy in earlier lectures as the Shannon entropy plus a constant. Suppose we take the probability density of the macrostate to be the Gibbs distribution i.e.

$$\varrho = A \exp(aE)$$

where

$$A^{-1}(a) = \int d\mathbf{p}d\mathbf{q} \exp(aE(\mathbf{p}, \mathbf{q})) \equiv Z(a) \quad (2)$$

where Z is commonly called the partition function. Let us fix the parameter a then we obtain

$$S = -a \langle E \rangle_{\varrho} + \log Z(a) + C \quad (3)$$

where the angle brackets denote expectation with respect to the given Gibbs distribution. Now it is easily seen that

$$u \equiv \langle E \rangle_{\varrho} = Z^{-1} \int d\mathbf{p}d\mathbf{q} E(\mathbf{p}, \mathbf{q}) \exp(aE(\mathbf{p}, \mathbf{q}))$$

and so $u = G(a)$. Now

$$\frac{dS}{du} = Z^{-1} \frac{dZ}{du} - \frac{da}{du} u - a$$

but using (2)

$$\frac{dZ}{du} = \int d\mathbf{p}d\mathbf{q} \frac{da}{du} E(\mathbf{p}, \mathbf{q}) \exp(aE(\mathbf{p}, \mathbf{q})) = \frac{da}{du} Z u$$

so we obtain due to the first two terms cancelling simply

$$\frac{dS}{du} = -a \equiv \beta$$

This mean energy u is simply the thermodynamical energy already discussed above and so we obtain

$$a = -\frac{1}{T}$$

which reveals the significance of the particular parameter a in the Gibbs distribution. It is simply minus the inverse absolute temperature.

Now consider two systems not in equilibrium and follow their evolution toward equilibrium i.e. toward states with equal temperatures. The second law insists that

$$\frac{dS}{dt} = \frac{dS_1}{dt} + \frac{dS_2}{dt} = \frac{dS_1}{dE_1} \frac{dE_1}{dt} + \frac{dS_2}{dE_2} \frac{dE_2}{dt} > 0$$

We now invoke conservation of energy between the two system so that

$$\frac{dE_1}{dt} = -\frac{dE_2}{dt}$$

and thus we obtain from above

$$\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \frac{dE_1}{dt} > 0$$

Let us assume that the first system has a higher temperature than the second. Then this inequality implies that

$$\frac{dE_1}{dt} < 0$$

i.e. that energy must be flowing from the hot system to the cold system as the system equilibrates.

Returning now to the entropy S we see that its definition as the Shannon entropy implies that it must be dimensionless. Thus we see from equation (1) that temperature the way we have defined it has units of energy. If the temperature of typical bodies are calculated we get very small values for the temperature so it is convenient to introduce effectively a unit of temperature (the degree Celsius) which is much smaller than a Joule. The ratio of the two is given by a constant called Boltzmann's constant

$$k = 1.38 \times 10^{-23} \text{ Joules/deg}$$

If the temperature is measured using the small unit of energy namely degrees then it is conventional often to rescale the definition of entropy by multiplying our original definition by k . This ensures that all equations derived above work with temperature measured in degrees rather than in Joules.

4 Further conditions for equilibrium

The two equilibrating systems discussed in the previous section exchanged only energy. In general other quantities may be exchanged as well and if we assume that such quantities are conserved quantities for closed systems then the Gibbs distribution will involve them as well as the energy in other words the entropy will be a function of not just the energy but also these other invariant quantities. One example is the number of particles N_1 and N_2 in the two subsystems. If the systems are able to exchange particles as well then the entropy becomes a function of the total number of particles as well. Using similar arguments as we did in the previous section we derive the further condition for equilibrium of our two subsystems using the maximum entropy principle and obtain

$$\frac{\partial S_1}{\partial N_1} = \frac{\partial S_2}{\partial N_2} = \text{Constant} \equiv -\frac{\mu}{T} \quad (4)$$

where the quantity μ has been introduced by this equilibration condition and is referred to as the chemical potential for the system. Another quantity which might be exchanged between subsystems is their volume V_i if we allow

the wall between the two subsystems to vary in location. Again the entropy will be a function of this important variable and using similar arguments we obtain

$$\frac{\partial S_1}{\partial V_1} = \frac{\partial S_2}{\partial V_2} = \text{Constant} \equiv \frac{p}{T} \quad (5)$$

where the equilibrium “label” p is referred to as the pressure. We discuss this identification of pressure in more detail later. The variables T , μ and p are of central importance to thermodynamics and their derivation from equations (1), (4) and (5) are called equations of state. Notice the crucial role of entropy to these equations. Evidently the calculation of entropy in terms of the invariant quantities is of fundamental significance and reflecting on the definition of this quantity given earlier, is clearly intimately related to the microscopic character of the system.

5 Internal energy, equilibrium and non-negative temperatures

Consider now for simplicity the case where entropy is only dependent on energy. Clearly in any system this energy cannot include the bulk kinetic energy of the system since one would not expect entropy to depend on the velocity of the object being observed due to a principle of relativity. We therefore introduce the concept of internal energy which is the total energy of a system minus it’s overall kinetic energy

$$I = E - KE = E - \frac{P^2}{2M}$$

where the system has momentum P and mass M and we have $S = S(I)$. Clearly we must now define temperature as $\frac{dS_a}{dI_a}$ for a given subsystem a . Consider now a series of open systems which together make up a large closed system. The total entropy is then

$$S = \sum_a S_a(I_a)$$

$$I_a = E_a - \frac{P_a^2}{2M_a}$$

Now the fact that the total system is closed means that total momentum and angular momentum must be conserved (as well as energy). Thus we have

$$\sum_a \mathbf{P}_a = \text{constant}$$

$$\sum_a \mathbf{r}_a \times \mathbf{P}_a = \text{constant}$$

where \mathbf{r}_a is the position vector for the particular subsystem. For entropy to be maximized we require S maximum subject to the previous two constraints.

A necessary condition for this can be derived using a constrained optimization method i.e. Lagrange multipliers. Thus our optimization function is

$$L = \sum_a S_a + \mathbf{b} \cdot \mathbf{P}_a + \mathbf{c} \cdot (\mathbf{r}_a \times \mathbf{P}_a) \quad (6)$$

where \mathbf{b} and \mathbf{c} are 3D vector Lagrange multipliers. Now using the chain rule we have

$$\frac{\partial}{\partial \mathbf{P}_a} S_a \left(E_a - \frac{P_a^2}{2M_a} \right) = \frac{-\mathbf{P}_a}{M_a T} = -\frac{\mathbf{v}_a}{T}$$

where T is the equilibrium temperature for all systems and \mathbf{v}_a is the subsystem velocity vector. Returning to the optimization function L in equation (6) and differentiating and setting to zero gives

$$-\frac{\mathbf{v}_a}{T} + \mathbf{a} + \mathbf{b} \times \mathbf{r}_a = 0$$

or

$$\mathbf{v}_a = \mathbf{u} + \mathbf{\Omega} \times \mathbf{r}_a$$

where $\mathbf{u} \equiv T\mathbf{a}$ and $\mathbf{\Omega} \equiv T\mathbf{b}$ are the same for all subsystems. This equation has the interpretation that in equilibrium the total system must move as a solid body with velocity \mathbf{u} and angular velocity $\mathbf{\Omega}$. When not in equilibrium this need not be the case of course.

Now an interesting question arises regarding temperature since it is the derivative of entropy with respect to (internal) energy. Does it always have the same sign? In all classical systems of practical interest the temperature is non-negative however it is possible to construct certain quantum systems in which a maximum energy occurs for fewer microstates than for a lower energy. These results are to date however somewhat controversial.