Chapter 11

Energy

In previous chapters of these notes we introduced the Principle of Maximum Entropy as a technique for estimating probability distributions consistent with constraints.

In Chapter 9 we discussed the simple case which can be done analytically, in which there are three probabilities, one constraint in the form of an average value, and the fact that the probabilities add up to one. There are, then, two equations and three unknowns, and it is straightforward to express the entropy in terms of one of the unknowns, eliminate the others, and find the maximum. This approach also works if there are four probabilities and two average-value constraints, in which case there is again one fewer equation than unknown.

In Chapter 9 we also discussed a more general case in which there are many probabilities but only one average constraint, so that the entropy cannot be expressed in terms of a single probability. The method of Lagrange multipliers was used, and provided the summations can be done, a general method of solution was presented.

In the application of the Principle of Maximum Entropy to physical systems, the number of possible states is usually very large, so that neither analytic nor numerical solutions are practical. Even in this case, however, the Principle of Maximum Entropy is useful because it leads to relationships among different quantities. Here we look at general features of such systems including the dependence of some quantities on others. A relatively simple model of a magnetic dipole system will be used.

Because we are now interested in the application to physical systems, we will express the entropy in Joules per Kelvin rather than in bits, and use the natural logarithm rather than the logarithm to the base 2. We will assume the multi-state model of the physical system implied by quantum mechanics, as discussed in Chapter 10 of these notes. This chapter focuses on the flow of energy in physical systems.

11.1 Magnetic Dipole Model

This model was introduced in section 9.1.2, and a version with two dipoles rather than one is shown below in Figure 11.1. An array of magnetic dipoles (think of them as tiny magnets) is subjected to an externally applied magnetic field H, and therefore the energy of the system depends on the orientations of the dipoles and on the applied field. Our system contains exactly two such dipoles, but it will from time to time be able to interchange information and energy with either of two environments, which are much larger collections of dipoles. Each dipole, both in the system and in its two environments, can be either "up" or "down," so there are four possible states for the system, "up-up," "up-down," "down-up," and "down-down." The energy of a dipole depends on its orientation and is proportional to the strength of the applied field, and the energy
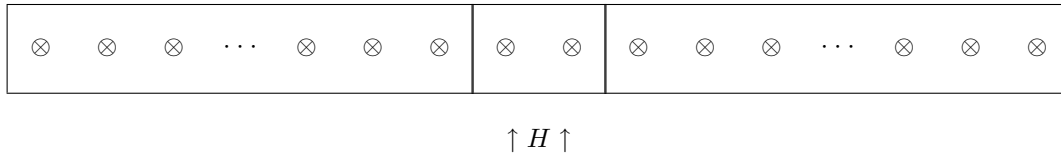
$$\uparrow H \uparrow$$

Figure 11.1: Dipole moment example.
(Each dipole can be either up or down.)

of each state is the sum of the energies of the two dipoles.

11.2   Principle of Maximum Entropy for Physical Systems

According to the multi-state model motivated by quantum mechanics (see Chapter 11 of these notes) there are a finite (or countable infinite) number of quantum states of the system. We will use i as an index over these states. The states have energy E_i, and might have other physical attributes as well. After these states are enumerated and described, then the Principle of Maximum Entropy can be used, as a separate step, to estimate which states are likely to be occupied.

We denote the occupancy of state i by the event A_i. The state i has probability $p(A_i)$ of being occupied. For simplicity we will write this probability $p(A_i)$ as p_i. We use the Principle of Maximum Entropy to estimate the probability distribution p_i consistent with the average energy E being a known (for example, measured) quantity. Thus

$$E = \sum_i p_i E_i \tag{11.1}$$

$$1 = \sum_i p_i \tag{11.2}$$

The entropy is

$$S = k_B \sum_i p_i \ln\left(\frac{1}{p_i}\right) \tag{11.3}$$

where $k_B = 1.38 \times 10^{-23}$ Joules per Kelvin and is known as Boltzmann's constant.

In the technique of Lagrange Multipliers, we do not use these equations to reduce the number of variables, but instead increase the number of unknowns. We define the "Lagrange multipliers" α and β and then the "Lagrangian" function L

$$L = S - k_B(\alpha - 1)\left[\sum_i p_i - 1\right] - k_B\beta\left[\sum_i E_i p_i - E\right] \tag{11.4}$$

The Lagrange multiplier α is dimensionless and β is expressed in inverse Joules.

The Principle of Maximum Entropy is carried out by finding α, β, and all p_i such that L is stationary. These values of p_i make S the largest it can be, subject to the constraints.

Since α only appears once in the expression for L, the quantity that multiplies it must be zero for the values that maximize L (otherwise a small change in α could increase L). Similarly, β only appears once in the expression for L, so the quantity that multiplies it must also vanish. Thus in the general case the p_i that we are seeking must satisfy

$$0 = \sum_i p_i - 1 \tag{11.5}$$

$$0 = \sum_i E_i p_i - E \tag{11.6}$$

The result of maximizing L with respect to each of the p_i is a relation between p_i and the Lagrange multipliers

$$\ln\left(\frac{1}{p_i}\right) = \alpha + \beta E_i \tag{11.7}$$

so

$$p_i = e^{-\alpha} e^{-\beta E_i} \tag{11.8}$$

The sum of the probabilities must be 1 and therefore

$$\alpha = \ln\left(\sum_i e^{-\beta E_i}\right) \tag{11.9}$$

As expressed in terms of the Principle of Maximum Entropy, the objective is to find the various quantities given the expected energy E. However, except in the simplest circumstances it is usually easier to do calculations the other way around. That is, it is easier to use β as an independent variable, calculate α in terms of it, and then find the p_i and then the entropy S and energy E.

Because the Lagrange multiplier plays a central role, it is helpful to understand intuitively how different values it may assume affect things. First, if $\beta = 0$, all probabilities are equal. This can only happen when there are a finite number of states.

If $\beta > 0$, then states with lower energy have a higher probability of being occupied. If $\beta < 0$, then states with higher energy have a higher probability of being occupied. Because of the exponential dependence on energy, unless $\mid \beta \mid$ is small, the only states with much probability of being occupied are either those with very large or very small energy.

11.2.1   General Properties

In earlier chapters of these notes, we assumed that it was possible to solve for the Lagrange multipliers, and then the probabilities could be found. If, as a practical matter, it is not possible to solve for the Lagrange multipliers, there are still general relations that can be derived.

First, we can multiply the equation above for $ln(1/p_i)$ by p_i and sum over i to obtain

$$S = k_B(\alpha + \beta E) \tag{11.10}$$

Second, we can look at a small change dE in E and inquire how the other variables change. Such relationships provide intuition which helps when the formulas are interpreted. The critical equations above are listed here for convenience

$$1 = \sum_i p_i \tag{11.11}$$

$$E = \sum_i p_i E_i \tag{11.12}$$

$$S = k_B \sum_i p_i \ln\left(\frac{1}{p_i}\right) \tag{11.13}$$

$$p_i = e^{-\alpha} e^{-\beta E_i} \tag{11.14}$$

$$\alpha = \ln\left(\sum_i e^{-\beta E_i}\right)$$

$$= \frac{S}{k_B} - \beta E \tag{11.15}$$

Now suppose E changes by a small amount dE. We will calculate from the equations above the changes in the other quantities, keeping only first-order variations (i.e., neglecting terms like $(dE)^2$ which, for small enough dE, are insignificantly small)

$$0 = \sum_i dp_i \tag{11.16}$$

$$dE = \sum_i E_i \, dp_i \tag{11.17}$$

$$dS = k_B \sum_i \ln\left(\frac{1}{p_i}\right) dp_i + k_B \sum_i p_i \, d\left[ln\left(\frac{1}{p_i}\right)\right]$$

$$= k_B \sum_i \ln\left(\frac{1}{p_i}\right) dp_i - k_B \sum_i \left(\frac{p_i}{p_i}\right) dp_i$$

$$= k_B \sum_i (\alpha + \beta E_i) \, dp_i$$

$$= k_B \beta \, dE \tag{11.18}$$

$$d\alpha = \left(\frac{1}{k_B}\right) dS - \beta \, dE - E \, d\beta$$

$$= - E \, d\beta \tag{11.19}$$

$$dp_i = p_i(-d\alpha - E_i \, d\beta)$$

$$= - p_i(E_i - E) \, d\beta \tag{11.20}$$

from which it is not difficult to show

$$dE = -\left(\sum_i p_i(E_i - E)^2\right) d\beta \tag{11.21}$$

$$dS = - k_B \beta \left(\sum_i p_i(E_i - E)^2\right) d\beta \tag{11.22}$$

These equations may be used in several ways. Note that all first-order variations are expressed as a function of $d\beta$ so it is natural to think of that as the independent variable. But this is not necessary; these equations remain valid no matter which change causes the other changes.

As an example of the insight gained from these equations, note that the formula relating dE and $d\beta$ implies that if E goes up then β goes down, and vice versa.

Third, other general properties are based on the dependence of some of the constraints on external parameters. In our magnetic-dipole example, the energy of each state depends on the externally applied magnetic field H. Each E_i could be written in the form $E_i(H)$ to emphasize this dependence. Thus the constraint could be written to show this dependence explicitly:

$$E = \sum_i p_i E_i(H) \tag{11.23}$$

Therefore each of the probabilities p_i depends on both E and H, as do α, β, and S. (In the case of our magnetic-dipole model, the energy $E_i(H)$ happens to be proportional to H with a constant of proportionality that depends on i but not on H. Other models, for other physical systems, might have dependencies on H or other parameters, that are different.)

Consider what happens if both E and H vary slightly, by amounts dE and dH, from the values used to calculate p_i, α, β, and S. There will be small changes dp_i, $d\alpha$, $d\beta$, and dS in those quantities which can be expressed in terms of the small changes dE and dH. The changes due to dE have been calculated above. The changes due to dH enter through the change in the energies associated with each state, $dE_i(H)$ (formulas like the next few could be derived for changes caused by any external parameter, not just the magnetic field).

$$0 = \sum_i dp_i \tag{11.24}$$

$$dE = \sum_i E_i(H)\, dp_i + \sum_i p_i\, dE_i(H) \tag{11.25}$$

$$dS = k_B \beta\, dE - k_B \beta \sum_i p_i\, dE_i(H) \tag{11.26}$$

$$d\alpha = -E\, d\beta - \beta \sum_i p_i\, dE_i(H) \tag{11.27}$$

$$dp_i = -p_i(E_i(H) - E)\, d\beta - p_i \beta\, dE_i(H) + p_i \beta \sum_j p_j\, dE_j(H) \tag{11.28}$$

$$dE = -\left[\sum_i p_i(E_i(H) - E)^2\right] d\beta + \sum_i p_i(1 - \beta(E_i(H) - E))\, dE_i(H) \tag{11.29}$$

$$dS = -k_B \beta \left[\sum_i p_i(E_i(H) - E)^2\right] d\beta - k_B \beta^2 \sum_i p_i(E_i(H) - E)\, dE_i(H) \tag{11.30}$$

For the particular magnetic dipole model considered here, the terms involving $dE_i(H)$ can be simplified by noting that each state's energy $E_i(H)$ is proportional to the parameter H and therefore

$$dE_i(H) = \left(\frac{E_i(H)}{H}\right) dH \tag{11.31}$$

$$\sum_i p_i dE_i(H) = \left(\frac{E}{H}\right) dH \tag{11.32}$$

so these formulas simplify to

$$0 = \sum_i dp_i \tag{11.33}$$

$$dE = \sum_i E_i(H)\, dp_i + \left(\frac{E}{H}\right) dH \tag{11.34}$$

$$dS = k_B \beta\, dE - \left(\frac{k_B \beta E}{H}\right) dH \tag{11.35}$$

$$d\alpha = -E\, d\beta - \left(\frac{\beta E}{H}\right) dH \tag{11.36}$$

$$dp_i = -p_i(E_i(H) - E)(d\beta + \left(\frac{\beta}{H}\right) dH) \tag{11.37}$$

$$dE = -\left[\sum_i p_i(E_i(H) - E)^2\right](d\beta + \left(\frac{\beta}{H}\right) dH) + \left(\frac{E}{H}\right) dH \tag{11.38}$$

$$dS = -k_B \beta \left[\sum_i p_i(E_i(H) - E)^2\right](d\beta + \left(\frac{\beta}{H}\right) dH) \tag{11.39}$$

There formulas can be used to relate the trends in the variables. For example, the last formula shows that a one percent change in β produces the same change in entropy as a one percent change in H.

11.3   System and Environment

Other general properties of these physical systems can be found if the system and its environment are considered separately. This partition into system and environment was already inherent in the magnetic-dipole model presented earlier (in fact that model has two environments, a property that will be essential when we consider energy conversion machines based on that model).

11.3.1   Partition Model

Let us model the system and its environment (for the moment consider only one such environment) as systems that each have their own set of possible states, and which can be isolated or can be in contact. That is, the system, considered apart from its environment, has states which, at least in principle, can be described. Each has an energy associated with it, and perhaps other physical properties as well. This description is separate from the determination of which state is actually occupied – that determination is made using the Principle of Maximum Entropy.

We also assume that the environment has its own set of states, each with its own energy and possibly other physical properties. Again this description of the states is independent of which states are actually occupied.

Our model for the interaction between these two (or what is equivalent, our model for the way the total combination is partitioned into the system and the environment) is that the combination has states that consist of one state from the environment and one from the system. Thus, for example, if the system has four states (as our simple two-dipole model does) and the environment has 100 states, then the combination would have 400 states. Each state of the combination corresponds to exactly one state of the system and exactly one state of the environment.

We need a notation to keep things straight. We will use the index i for the system and the index j for the environment. Then we can denote the states of the total combination using both i and j, in the form i,j just like the notation for joint probability (which is exactly what it is). A sum over the states of the total combination is then a sum over both i and j.

We will assume that it is possible for the system and the environment to be isolated from one another (the dipole drawing shows a vertical bar which is supposed to represent a barrier to interaction) and then, at other times, for the two to be interacting. Whether they are isolated or interacting does not affect the states or the physical properties associated with the states, although it may affect the probability of occupancy of the states.

11.3.2 Interaction Model

The reason for our partition model is that we want to control interaction between the system and its environment. Different physical systems would have different modes of interaction, and different mechanisms for isolating different parts. Here is described a simple model for interaction of magnetic dipoles that are aligned in a row. It is offered as an example.

Suppose that the apparatus that holds the magnetic dipoles allows adjacent dipoles to influence each other. This influence might be to cause one dipole to change from up to down or vice versa. Naturally, if one dipole influences its neighbor, then its neighbor at the same time influences it. It is reasonable to suppose that if one dipole changes its status from, say, up to down, then the neighbor that is interacting with it should change its status in the opposite direction. In that way the total number of dipoles oriented in each direction stays fixed.

If both start out aligned the same way, this exchange would have no effect. However, if they were differently aligned (one up and one down), this exchange would have the effect of moving the alignments around without moving the dipoles themselves. It is as though the orientations migrated around.

If the two dipoles exchanging dipole moment are both in the system, or both in the environment, then there is no effect on the energy of the system or the energy of the environment. Only if they are located on the boundary, one on each side, does this mechanism lead to energy being transferred to or from the system. If the system and the environment have different amounts of average energy per dipole then the random interchange of dipole moment would result in a net flow of energy across the boundary between them. For example if most of the dipoles in the system start out in the down state, with high energy, then random interchanges will, on average, result in fewer dipoles in that state, so the energy of the system has gone down, and the energy of the environment has gone up an equal amount.

Sometimes this kind of a process is referred to as "mixing" because the effect is like different kinds of particles being mixed together. However, here the particles do not move, but rather the property of dipole orientation is what moves around.

Energy that is transferred to or from the system as a result of mixing is referred to as "heat." A formula for heat in terms of changes of probability distribution is given below.

Let us also assume that we can, by placing or removing appropriate barriers, either inhibit or allow this process. For example, the process might be inhibited by simply moving the system away from its environment physically. Energy conversion devices often use sequences where mixing is encouraged or prevented at different times.

11.3.3 Extensive and Intensive Quantities

This partition model leads to an important property that physical quantities can have. Some physical quantities will be found to be "extensive" and others "intensive." These properties are a natural consequence of our model.

Whether the system is isolated from the environment or is interacting with it, and whatever the probability distributions $p_{s,i}$ of the system, $p_{e,j}$ of the environment, and $p_{t,i,j}$ of the combination, the energies of the system state and of the environment state add up to form the energy of the corresponding total state:

$$E_{t,i,j} = E_{s,i} + E_{e,j} \tag{11.40}$$

The probability of occupancy of total state k is the product of the two probabilities of the two associated states i and j:

$$p_{t,i,j} = p_{s,i}p_{e,j} \qquad (11.41)$$

With this background it is easy to show that the expected value of the total energy is the sum of the expected values of the system and environment energies:

$$
\begin{aligned}
E_t &= \sum_{i,j} E_{t,i,j} p_{t,i,j} \\
&= \sum_{i,j} \left[E_{s,i} + E_{e,j} \right] p_{s,i}p_{e,j} \\
&= \sum_i \sum_j \left[E_{s,i} + E_{e,j} \right] p_{s,i}p_{e,j} \\
&= \sum_i p_{s,i} \sum_j E_{e,j}p_{e,j} + \sum_j p_{e,j} \sum_i E_{s,i}p_{s,i} \\
&= \sum_j E_{e,j}p_{e,j} + \sum_i E_{s,i}p_{s,i} \\
&= E_e + E_s \qquad (11.42)
\end{aligned}
$$

This result holds whether the system and environment are isolated or interacting. It states that the energy of the system and the energy of the environment add up to make the total energy. It is a consequence of the fact that the energy associated with each total state is the sum of the energies associated with the corresponding system and environment states.

A quantity with the property that its total value is the sum of the values for the two (or more) parts is known as an **extensive** quantity. Energy has that property, as was just demonstrated. Entropy is also extensive. That is,

$$
\begin{aligned}
S_t &= \sum_{i,j} p_{t,i,j} \ln\left(\frac{1}{p_{t,i,j}} \right) \\
&= \sum_{i,j} p_{s,i}p_{e,j} \left[\ln\left(\frac{1}{p_{s,i}} \right) + \ln\left(\frac{1}{p_{e,j}} \right) \right] \\
&= \sum_i \sum_j p_{s,i}p_{e,j} \left[\ln\left(\frac{1}{p_{s,i}} \right) + \ln\left(\frac{1}{p_{e,j}} \right) \right] \\
&= \sum_i p_{s,i} \sum_j p_{e,j} \ln\left(\frac{1}{p_{e,j}} \right) + \sum_j p_{e,j} \sum_i p_{s,i} \ln\left(\frac{1}{p_{s,i}} \right) \\
&= \sum_j p_{e,j} \ln\left(\frac{1}{p_{e,j}} \right) + \sum_i p_{s,i} \ln\left(\frac{1}{p_{s,i}} \right) \\
&= S_e + S_s \qquad (11.43)
\end{aligned}
$$

Again this result holds whether or not the system and environment are isolated or interacting.

Not all quantities of interest are extensive. In particular, the Lagrange multipliers α and β are not extensive. Consider β. This is an example of a quantity for which the values associated with the system, the environment, and the total configuration may or may not be related. If the system and environment are isolated, so that a separate application of the Principle of Maximum Entropy is made to each, then there is no reason why β_s and β_e would be related. On the other hand, if the system and environment are interacting

so that they are exchanging energy, the distribution of energy between the system and the environment may not be known and therefore the Principle of Maximum Entropy can be applied only to the total combination, not to the system and environment separately. Then, the same value of β would apply throughout, even though the energy and entropy are extensive quantities.

Quantities like β that are the same throughout a system when analyzed as a whole are called **intensive**. Lagrange multipliers which multiply extensive quantities are intensive.

11.3.4   Equilibrium

The partition model leads to interesting results when the system and its environment are allowed to come into contact after having been isolated. In thermodynamics this process is known as the total configuration coming into equilibrium.

Let us suppose that the system and the environment have been in isolation and therefore are characterized by different, unrelated, values of energy, entropy, and other quantities. Then suppose they are allowed to interact, using a model of interaction in which the total energy is unchanged. Energy may flow from the system to the environment or vice versa because of mixing, and this flow of energy is called heat. As a result, the probabilities of occupancy will change, although the descriptions of the states and their properties, including their energies, do not change.

We have developed general formulas that relate small changes in probabilities, and in E, S, α, and β which can be used now. If the energy of the system is assumed to change somewhat (because of mixing), that fact could be incorporated into a new application of the Principle of Maximum Entropy to the system that would result in modified probabilities, E, S, α, and β. In particular, we saw earlier that the signs of dE and $d\beta$ are opposite, so that if E goes up, β then goes down, and vice versa.

Soon, the transfer of energy between the system and the environment may result in our not knowing the energy of each separately, but only the total energy (which does not change as a result of the mixing). In that case, it would be appropriate to use the Principle of Maximum Entropy on the total combination of system and environment together. When that is done, there will be a new single value of β and a new total entropy. What can be said about these values?

First, the new entropy is the sum of the new entropy of the system and the new entropy of the environment, because entropy is an extensive quantity. Furthermore, the old total entropy (at the time the interaction started) is the sum of the old system entropy and the old environment entropy, for the same reason. However, what is interesting is the new total entropy compared with the old total entropy.

The new entropy, because it is evaluated with the probability distribution that comes from the Principle of Maximum Entropy, is the largest value consistent with the total energy. Any other probability distribution consistent with the same total energy would lead to a smaller (or possibly equal) entropy. One such probability distribution is the distribution prior to the mixing, the one that led to the old entropy value. Therefore the total entropy has increased (or at best stayed constant) as a result of the interaction between the system and the environment. It may be that the entropy of the system alone has gone down, but if so then the entropy of the environment must have gone up at least as much.

The energies of the system and the environment have changed, and as a result the values of β_s and β_e have changed, in opposite directions. Their new values are the same (each is equal to β_t), and therefore this new value lies between the two old values.

11.3.5   Energy Flow, Work and Heat

Let us return to the magnetic dipole model as shown in Figure 11.1.

In this section we will consider interactions with only one of the two environments. If both are used then the machine can be used for practical energy conversion as a heat engine or a refrigerator. These uses are considered in Chapter 12 of these notes.

Consider first the case that the system is isolated from its environment, as shown above (the vertical bars represent barriers to interaction). The system is in some state, and we do not necessarily know which one, although the probability distribution p_i can be obtained from the Principle of Maximum Entropy. A

change in state generally requires a nonzero amount of energy, because the different states have different energies. We can always imagine a small enough change dH in H so that the magnetic field cannot supply or absorb the necessary energy to change state. Then we can imagine a succession of such changes in H, none of which can change state, but when taken together constitute a large enough change in H to be noticeable. We conclude that changing H for an isolated system does not change the state. Thus the probability distribution p_i is unchanged. Of course changing H by an amount dH does change the energy through the resulting change in $E_i(H)$:

$$dE = \sum_i p_i \, dE_i(H) \tag{11.44}$$

This change is reversible: if the field is changed back, the energy could be recovered in electrical or magnetic or mechanical form (there is nowhere else for it to go in this model). Energy flow of this sort, that can be recovered in electrical, magnetic, or mechanical form (or some other forms) is referred to as **work**. If $dE > 0$ then we say that work is positive, in that it was done by the external source on the system; if $dE < 0$ then we say that work is negative, in that it was done on the external source by the system. Naturally, in energy-conversion devices it is important to know whether the work is positive or negative. In many cases simply running the machine backwards changes the sign of the work; this is not always true of the other form of energy transfer, discussed below.

Changes to a system caused by a change in one or more of its parameters, when it cannot interact with its environment, are known as **adiabatic** changes. Since the probability distribution is not changed by them, they produce no change in entropy of the system. This is a general principle: adiabatic changes do not change the probability distribution and therefore conserve entropy.

First-order changes to the quantities of interest were given above in the general case where E and the various E_i are changed. If the change is adiabatic, then dE is caused only by the changes dE_i and the general equations simplify to

$$dp_i = 0 \tag{11.45}$$

$$dE = \sum_i p_i \, dE_i(H) \tag{11.46}$$

$$dS = 0 \tag{11.47}$$

$$d\alpha = -E \, d\beta - \beta \sum_i p_i \, dE_i(H) \tag{11.48}$$

$$0 = \left[\sum_i p_i (E_i(H) - E)^2 \right] d\beta + \beta \sum_i p_i (E_i(H) - E) \, dE_i(H) \tag{11.49}$$

If, as in our magnetic-dipole model, the energies of the states are proportional to H then these adiabatic formulas simplify further to

$$dp_i = 0 \tag{11.50}$$

$$dE = \left(\frac{E}{H} \right) dH \tag{11.51}$$

$$dS = 0 \tag{11.52}$$

$$d\alpha = 0 \tag{11.53}$$

$$d\beta = - \left(\frac{\beta}{H} \right) dH \tag{11.54}$$

Next, consider the system no longer isolated, but instead interacting with its environment. The interaction model permits heat to flow between the system and the environment, and by convention we will say the heat

is positive if energy flows into the system from the environment, and negative if the energy flows the other way. Energy can be transferred by heat and work at the same time. Work is represented by changes in the energy of the individual states dE_i, and heat by changes in the probabilities p_i. Thus the formula for dE above becomes

$$dE = \sum_i E_i(H)\, dp_i + \sum_i p_i\, dE_i(H) \tag{11.55}$$

where the first term is heat and the second term is work.

11.3.6   Reversible Energy Flow

We saw in a previous section that when a system is allowed to interact with its environment, total entropy generally increases. In this case it is not possible to restore the system and the environment to their prior states by further mixing, because such a restoration would require a lower total entropy. Thus mixing in general is irreversible.

The limiting case where the total entropy stays constant is one where, if the system has changed, it can be restored to its prior state. It is easy to derive the conditions under which such changes are, in this sense, reversible.

From the formulas given earlier, the change in system entropy is proportional to the part of the change in energy due to heat. Thus

$$dS_s = k_B \beta_s\, dE_s - k_B \beta_s \sum_i p_{s,i}\, dE_{s,i}(H) \tag{11.56}$$

$$= k_B \beta_s \left[dE_s - \sum_i p_{s,i}\, dE_{s,i}(H) \right] \tag{11.57}$$

$$= k_B \beta_s\, dq_s \tag{11.58}$$

where dq_s stands for the heat that comes into the system due to the interaction mechanism.

This formula applies to the system and a similar formula applies to the environment:

$$dS_e = k_B \beta_e\, dq_e \tag{11.59}$$

The two heats are the same except for sign

$$dq_s = -dq_e \tag{11.60}$$

and it therefore follows that the total entropy $S_s + S_e$ is unchanged if and only if the two values of β for the system and environment are the same:

$$\beta_s = \beta_e \tag{11.61}$$

Reversible changes (with no change in total entropy) can involve work and heat and therefore changes in energy and entropy for the system, but the system and the environment must have the same value of β. Also, we showed earlier that interactions between the system and the environment result in a new value of β intermediate between the two starting values of β_s and β_e, so reversible changes result in no change to β.

The first-order change formulas given earlier can be written to account for reversible interactions with the environment by simply setting $d\beta = 0$

$$0 = \sum_i dp_i \tag{11.62}$$

$$dE = \sum_i E_i(H)\, dp_i + \sum_i p_i\, dE_i(H) \tag{11.63}$$

$$dS = k_B \beta\, dE - k_B \beta \sum_i p_i\, dE_i(H) \tag{11.64}$$

$$d\alpha = -\beta \sum_i p_i\, dE_i(H) \tag{11.65}$$

$$dp_i = -p_i \beta\, dE_i(H) + p_i \beta \sum_j p_j\, dE_j(H) \tag{11.66}$$

$$dE = \sum_i p_i (1 - \beta(E_i(H) - E))\, dE_i(H) \tag{11.67}$$

$$dS = -k_B \beta^2 \sum_i p_i (E_i(H) - E)\, dE_i(H) \tag{11.68}$$

As before, these formulas can be further simplified in the case where the energies of the individual states is proportional to H

$$0 = \sum_i dp_i \tag{11.69}$$

$$dE = \sum_i E_i(H)\, dp_i + \left(\frac{E}{H}\right) dH \tag{11.70}$$

$$dS = k_B \beta\, dE - k_B \beta \left(\frac{E}{H}\right) dH \tag{11.71}$$

$$d\alpha = -\left(\frac{\beta E}{H}\right) dH \tag{11.72}$$

$$dp_i = -\left(\frac{p_i \beta}{H}\right)(E_i(H) - E)\, dH \tag{11.73}$$

$$dE = \left(\frac{E}{H}\right) dH - \left(\frac{\beta}{H}\right) \left[\sum_i p_i (E_i(H) - E)^2\right] dH \tag{11.74}$$

$$dS = -\left(\frac{k_B \beta^2}{H}\right) \left[\sum_i p_i (E_i(H) - E)^2\right] dH \tag{11.75}$$

These formulas will be used in the next chapter of these notes to derive constraints on the efficiency of energy conversion machines that involve heat.