Apr. 8

Entropy

In the previous lecture, we introduced the concept of *entropy.* The goal of this lecture is to explain this concept further, from several points of view. Entropy can be looked at as a measure of *irreversibility*, as a measure of *heat flow*, and as a measure of *information.* I will discuss these viewpoints in turn.

Begin with irreversibility. My viewpoint on the approach to thermal equilibrium is that systems move from special to typical points on the surface of fixed energy. Typical points can be represented by uniform averages over the energy surface. Special points are represented by averages over subspaces or restricted regions of the energy surface in which these states are typical. In each case, we can define an entropy. The area of the energy surface or volume of phase space associated with the average is $\Omega(E)$, with

$$S(E) = \log \Omega(E)$$

The evolution from a special to a typical state then corresponds to an increase in $\Omega(E)$, or an increase in entropy.

This idea combined with the idea of conservation of energy gives rise to the *laws of thermodynamics*:

1. The total *energy* is constant.
2. The total *entropy* tends to a maximum.

Entropy is also associated with *heat.* Heat is a concept that we have not yet discussed in this course. How should we define it? In general, when a macroscopic system changes its energy, that change has three components. First, there is a change associated with work that is done on the system by large-scale mechanical motion. For example, if the volume of a gas is decreased and the gas has a nonzero pressure, work is done on the gas. Second, there is a change in the energy of a system if the number of particles in the system is increased or decreased. But, finally, there is energy tied up in the disorganized motion of the atoms or molecules that make up a macroscopic material. A change in this atomic-scale motion is a change in the energy. This is the sort of change that we would like to describe as a flow of heat.

In the previous lecture, I wrote the equation

$$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{N,V}$$

Equivalently

$$dE\Big|_{N,V} = T\,dS\Big|_{N,V}$$

According to this equation, S precisely accounts for that part of the energy of a body that is not associated with work done on the system or the exchange of particles. It represents energy lost to atomic or molecular motion. Often, then, we write

$$dQ = T\,dS$$

where dQ is a flow of heat.

Energy in heat is not necessarily energy that is inaccessible to practical use. There are engines that run with heat as their energy source. However, it is not possible to extract heat energy with 100% efficiency. Normally, a heat engine will transfer heat from a hotter body to a cooler body. Some fraction of the energy transfered can be diverted to a more useful form.

In the nineteenth century, Carnot analyzed this situation in great generality. His simplest example, an ideal gas providing a contact between a body at high temperature T_1 and a cooler body at temperature T_2, is worth close attention. Carnot argued that the best possible situation is that in which we transfer energy between the bodies *reversibly*, that is, with *zero total change of entropy*. A reversible process will give the highest possible efficiency. If some entropy is created, this will consume some of the available energy in extra heating.

Before we consider Carnot's process, we must analyze the change in volume of a gas with zero change in entropy. This is a *reversible* or *adiabatic* expansion or contraction. I will assume that N is fixed in this process. Then the change in energy is entirely accounted for by the work done on the system,

$$dE = -p\,dV$$

If the system remains in thermal equilibrium during the process,

$$E = \frac{3}{2} N T \qquad p = \frac{NT}{V}$$

so the temperature must change by

$$\frac{3}{2} N\,dT = -\frac{NT}{V}\,dV$$

that is

$$\frac{3}{2}\frac{dT}{T} = -\frac{dV}{V}$$

We see, finally, that the temperature change is given by

$$\left(V_1/V_2\right) = \left(T_1/T_2\right)^{-3/2}$$

Adiabatic compression increases the temperature; adiabatic expansion lowers the temperature. The relation between pressure and volume is then different from that in *isothermal* expansion, with T fixed. Since $p = NT/V$, isothermal expansion gives

$$\frac{P_1}{P_2} = \left(\frac{V_1}{V_2}\right)^{-1}$$

We can obtain the corresponding formula for adiabatic expansion by substituting $T_i = p_i V_i / N$ in the expression above. Then

$$\frac{P_1}{P_2} = \left(\frac{V_1}{V_2}\right)^{-5/3}$$

This formula for adiabatic expansion or contraction is canonically written

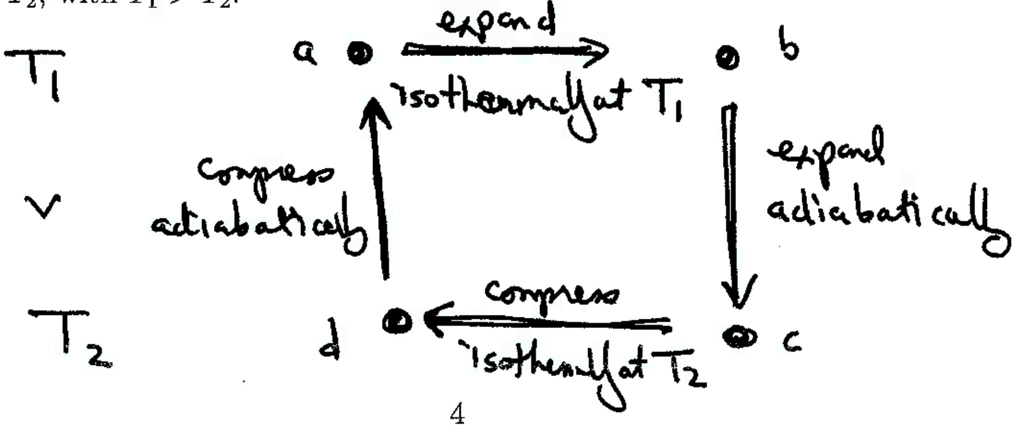
$$PV^{\gamma} = (\text{const})$$

We see that, for a structureless atomic gas,

$$\gamma = \frac{5}{3}$$

We will find different values for γ when we consider the statistical mechanics of molecules with internal degrees of freedom.

Now we can link adiabatic and isothermal reversible processes to make a *Carnot cycle* that transfers energy from a reservoir at temperature T_1 to a reservoir at temperature T_2, with $T_1 > T_2$:



For two bodies at the same temperature,

$$\left.\frac{dS}{dE}\right|_1 = \left.\frac{dS}{dE}\right|_2$$

so transfer of energy from one to the other does not change the total entropy. In an adiabatic process, the entropy does not change by construction. So all steps of this cycle are reversible.

We can extract useful work from running the cycle in the direction indicated. In isothermal expansion, the gas does work on its surroundings

$$\int_a^b dW = \int_a^b p\, dV = \int_a^b NT \frac{dV}{V}$$

Then, in the isothermal expansion step,

$$W_{ab} = N T_1 \log \frac{V_b}{V_a}$$

Similarly, in the isothermal contraction step work must be done on the gas,

$$W_{cd} = - N T_2 \log \frac{V_c}{V_d}$$

In the adiabatic stages of the cycle, work is done on the gas, but this work is precisely what is needed to change the internal energy of the gas from

$$\frac{3}{2} N T_1 \quad \text{to} \quad \frac{3}{2} N T_2$$

and back again. The net work around the cycle is zero. Then the total work extracted from the gas in this cycle is

$$NT_1 \log \frac{V_b}{V_a} - NT_2 \frac{V_c}{V_d}$$

The ratio of volumes is related to the ratio of temperatures by

$$\frac{V_b}{V_c} = \left(\frac{T_1}{T_2}\right)^{-3/2} = \frac{V_a}{V_d}$$

It follows also that

$$\frac{V_b}{V_a} = \frac{V_c}{V_d}$$

Then, finally, we find

$$\frac{(\text{net work extracted})}{(\text{energy input at } T_1)} = \frac{T_1 - T_2}{T_1}$$

This represents the maximum possible efficiency of a heat engine built on this principle. Note that the efficiency goes to zero for $(T_1 - T_2) \to 0$.

The interpretation of entropy as heat fives us a more general way to derive this formula for the efficiency of a reversible process. We can consider a very general idealized heat engine as being built out of a cycle of isothermal and adiabatic steps. No net work will be done in the adiabatic steps, and no net energy is transfered from the engine in these steps. In the isothermal steps, there is a transfer of heat to the engine at T_1, given by

$$\Delta Q_1 = T_1 \Delta S_1$$

This balances the extracted work. Similarly, at T_2, the heat

$$\Delta Q_2 = +T_2 \, \Delta S_2$$

is extracted from the gas by supplying an equal amount of work. Since the total change in the entropy of the gas around the cycle is zero, we must have

$$\Delta S_1 + \Delta S_2 = 0$$

Then the useful work extracted is

$$\frac{\Delta Q_1 - \Delta Q_2}{\Delta Q_1}$$

In this argument, the detailed properties of the ideal gas have disappeared. This estimate is then an upper bound on the efficiency of any heat engine.

Entropy has another aspect related to *information*. To explain this, I will first recast our formula for entropy into a form that would apply to a discrete, rather than a continuous, form of the microcanonical average. The formula we obtain will anticipate the formula for entropy that we will see later in quantum statistical mechanics.

In the classical microcanonical ensemble, we average over a continous array of systems spanning the surface of fixed energy. However, we could approximate this average as an average over cells in phase space, choosing a representative point in each phase space cell. For example, to approximated the situation in quantum statistical mechanics, we might take the volume of each cell to be $(2\pi\hbar)^{D-1}$, where D is the total number of degrees of freedom. In that case, each cell has an equal probability equal to

$$\frac{(2\pi\hbar)^{D-1}}{\Omega(E)}$$

and the formula for the entropy becomes

$$S = \sum_{\text{cells}} \frac{(2\pi\hbar)^{D-1}}{\Omega(E)} \log \frac{\Omega(E)}{(2\pi\hbar)^{D-1}} = -\sum_{\text{cells}} \frac{(2\pi\hbar)^{D-1}}{\Omega(E)} \log \frac{(2\pi\hbar)^{D-1}}{\Omega(E)}$$

This expression naturally generalizes to the expression, applicable to cells with a range of probabilities p_i,

$$S = -\sum_i p_i \log p_i$$

This expression is the canonical general formula for entropy. We will see many examples of its use as the course proceeds.

This expression for S has the following nice interpretation: Consider a process in which there are N possible outcomes. The true outcome is not known, but the various outcomes will appear with probabilities p_i, with

$$\sum_i p_i = 1$$

Then S is the *information that is gained* when the actual outcome is revealed. When S is interpreted as information, it is conventional to measure it in *bits*, by writing

$$S_S = -\sum_i p_i \log_2 p_i$$

This expression is called the *Shannon entropy.* Since

$$\log_2 P = \frac{\log P}{\log 2}$$

so $S_S = S/\log 2$. This use of entropy was introduced by Shannon to quantify how much information could be conveyed over a noisy communication channel.

I will now present and derive the basic properties of Shannon entropy. (In the rest of this lecture, I will write $\log x$ for $\log_2 x$.)

1. One unknown bit contains one unit of information.

For this case, there are two outcomes, 0 and 1. They appear with equal probabilities,

$$p(0) = \frac{1}{2} \qquad p(1) = \frac{1}{2}$$

When we measure the bit, we recover information

$$S_S = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$$

2. If an outcome has zero probability, it has no effect on information.

This is the case $p_i = 0$ The statement follows from the fact that

$$p \log p \to 0 \quad \text{as} \quad p \to 0$$

3. If one outcome is certain, the measurement generates no information.

This is the case in which $p_1 = 1$ and all other p_i are equal to zero. The statement follows from the fact that

$$p \log p \to 0 \quad \text{as} \quad p \to 1$$

4. If we compare a process in which two outcomes have probabilities p_1 and p_2 to one where these two outcomes are p_1' and p_2', S_S is higher for the process in which these probabilities are closer to being equal.

For definiteness, consider a range of situations parametrized by α in which $p_1(\alpha) > p_2(\alpha)$ and

$$p_1(\alpha) + p_2(\alpha) = \text{constant}$$

which implies

$$\frac{dp_1}{d\alpha} + \frac{dp_2}{d\alpha} = 0$$

Then

$$\frac{d}{d\alpha}(-p_1 \log p_1 - p_2 \log p_2)$$

$$= \frac{dp_1}{d\alpha}(-\log p_1 - 1) + \frac{dp_2}{d\alpha}(-\log p_2 - 1)$$

$$= \frac{dp_2}{d\alpha}\left(\log \frac{p_1}{p_2}\right)$$

This is positive if

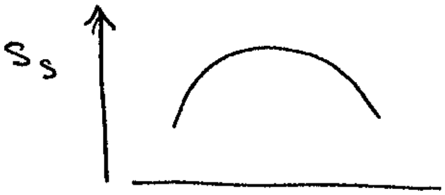
$$\frac{dp_2}{d\alpha} > 0$$

which implies the result claimed.

This result implies, in particular, that S_S is maximized when all of the probabilities are equal. In that case

$$p_i = \frac{1}{M} \qquad i = 1 - M \qquad S_S = \log_2 M$$

More generally, $S_S(\{p_i\})$ can be shown to be convex downward



along any line in the space of the p_i.

5. The Shannon entropy of the composite of two independent processes is the sum of their individual entropies.

Let process 1 have probabilities p_{1i}, let process 2 have probabilities p_{2i}, and let all of these outcomes be independent. Then the probability of the outcome (i and j) or (ij) is

$$p_{ij} = p_{1i} \cdot p_{2j}$$

The total Shannon entropy is

$$S_S^{(12)} = - \sum_{ij} p_{1i} p_{2j} \log(p_{1i} p_{2j})$$

$$= - \sum_{ij} p_{1i} p_{2j} \left(\log p_{1i} + \log p_{2j} \right)$$

Now use

$$\sum_i p_{1i} = \sum_j p_{2j} = 1$$

Then

$$S_S^{(12)} = -\sum_i p_{1i} \log p_{1i} - \sum_j p_{2j} \log p_{2j}$$

$$= S_S(1) + S_S(2)$$

In statistical mechanics, this property gives the important outcome that the entropy is *extensive*, that is

$$S \sim N$$

for a system with many degrees of freedom.

Even more nicely, if the outcomes i and j are correlated, S_S behaves correctly when one of i or j is measured. Let p_{ij} be the probability of the outcome (ij). The probability of the outcome j in the second process is

$$P_j = \sum_i p_{ij}$$

If j is measured, the *conditional probability* of i given j is

$$c(i|j) = \frac{p_{ij}}{P_j}$$

Note that

$$\sum_i c(i|j) = 1$$

Now there are three entropies in play,

$$S_S(12) = -\sum_{ij} p_{ij} \log p_{ij}$$ information in the choice of (ij)

$$S_S(2) = -\sum_{j} P_j \log P_j$$ information in the choice of j

$$S_S(1j) = -\sum_{i} C(i|j) \log C(i|j)$$ information in the choice of i given j

The average of $S_S(1j)$ over the outcomes j is

$$\begin{aligned} \langle S_S(1j) \rangle &= \sum_j P_j \left[-\sum_i C(i|j) \log C(i|j) \right] \\ &= -\sum_{ij} P_j \frac{p_{ij}}{P_j} \log \frac{p_{ij}}{P_j}, \\ &= -\sum_{ij} p_{ij} \log p_{ij} + \sum_{ij} p_{ij} \log P_j \end{aligned}$$

The final result is $$= -\sum_{ij} p_{ij} \log p_{ij} + \sum_j P_j \log P_j$$

$$\langle S_S(1j) \rangle = S_S(12) - S_S(2)$$

That is, the information gained in measuring i *after* j is less then the information gained in measuring the pair (ij) precisely by the information contained in the choice of j.

The idea of entropy as missing information has many applications in computer science, pattern recognition, and neuroscience.