Answers to Physics 176 One-Minute Questionnaires

Lecture date: February 3, 2011

What motivated entropy?

I don't know the history well (few scientists take the time to learn historical details, including myself) and my comments below should be taken with a grain of salt. Take a look at the Wikipedia article "Entropy" and also look through Chapter 4 of Schroeder, which summarizes the traditional thermodynamics approach to entropy via heat engines and refrigerators. There are also many books, popular and technical, that discuss the history, and some of these are listed at the back of Schroeder (e.g., the book "The Second Law" by P. Atkins).

My understanding is that much of the motivation for thinking about entropy was driven by the practical needs of the industrial revolution, by people who wanted to replace animals and people with machines such as water wheels and steam engines. Builders of steam engines and scientists (which were sometimes the same person) started trying to figure out empirically and conceptually how well could one do in converting energy from fuel, say burning wood, to getting useful work out of a machine. Over time, scientists realized (and it is here that I don't know the history) that an important quantity for understanding steam engines was the ratio Q/T, heat Q exchanged between two systems divided by some temperature T, either the temperature of the system receiving the heat or the system giving up the heat, and this ratio was called "entropy". It was discovered empirically that entropy always increased, which prevented perfectly efficient machines from existing, and also discovered theoretically that entropy set hard limits on the efficiency of engines, as you saw a little bit in the recent homework problem (Problem 4.1 in Schroeder).

Toward the end of the 19th century, several decades after the entropy was identified as Q/T, scientists like Boltzmann developed a new understanding of entropy by using kinetic theory to interpret heat in terms of microscopic details. This led to Boltzmann's great discovery (also one of the highlights of this course), that the thermodynamic entropy could be understood as counting possible microstates of specified macrostates, $S = k\ln(\Omega)$. (Boltzmann was so proud of this insight that he asked that this formula be engraved on his tomb in Vienna after his death.) But Boltzmann's insight was limited by the inability of classical mechanics to identify correctly what was a microstate as shown by the wrong predictions of the equipartition theorem. It was not until the invention of quantum mechanics in the 1920s that a definitive understanding of "microstate" was achieved and then entropy became a clear valuable concept that could be used quantitatively and unambiguously in all areas of science.

In the 1950s and later, entropy became widely appreciated and important in a new context after Claude Shannon at Bell Labs showed in 1948 that understanding how much information can be transmitted from one device to another could be analyzed in terms of entropy. Shannon's theory had a huge role in electrical engineering, computer design, Internet design, experimental science, and currently in neurobiology, and in turn has motivated many physicists to reinterpret physical phenomena as manipulating information rather than just changing one physical state to another state.

How can something allow atoms to pass through but not heat?

Actually, I am not sure that this can done physically, at least for an isolated system. It could be accomplished if one were allowed to maintain a temperature difference across the partition, for example arrange for the partition to be thick and have a temperature T_1 on one side and temperature T_2 on the other side, and then molecules could diffuse through without exchanging heat between two subsystems with temperatures T_1 and T_2. But this scenario is not allowed for a closed system.

The relaxation time for thermal equilibrium is usually not the same for diffusive equilibrium and the times can differ by an order of magnitude or more (and the ratio of relaxation times can also be easily varied by choosing the material properties of the partition). So two subsystems separated by a partition might reach thermal equilibrium before reaching diffusive equilibrium in which case no heat is being transported.

Why does the max value of $S \Rightarrow dS/dU_1 = 0$? Why with respect to U_1?

In an isolated macroscopic system with a fixed amount of energy U and with two macroscopic subsystems with energies U_1 and U_2 with $U_1 + U_2 = U$, the entropy S of the entire system will depend on the value of U_1. (We will soon study this quantitatively for some simple models which should help make this point clear.) For the entropy S of the system to reach a maximum value, it must also be a maximum with respect to changes in any internal variables such as U_1, N_1, and V_1. So the derivative dS/dx must be zero for x representing some quantity that can be exchanged between subsystems.

For the example $U = (1/2)m_1v_1^2 + (1/2)m_2v_2^2$, would the number of microstates be infinite based on the number of possible velocities? (Unless we consider the energy to be quantized?)

According to classical mechanics, the state of a point particle is given by two numbers, the position x of the particle (say moving in one dimension for simplicity) and its momentum p_x (or speed v_x). For two point particles with states $(x_1, p_{1,x})$ and $(x_2, p_{2,x})$, with a fixed total energy specified by $U = p_{1,x}^2/(2m) + p_{2,x}^2/(2m)$, there is indeed an infinite number of microstates for a specified volume V and energy U and so the multiplicity seems to be a useless concept since one can't compute with infinities.

But quantum mechanics changes this situation fundamentally and leads to a finite although large set of microstates for specified macrovariables V and U. The flavor of the argument is based on the position-momentum uncertainty principle in the form $\Delta x \times \Delta p_x \geq \hbar$. As we will discuss soon when we get to Section 2.5 in Schroeder, the uncertainty principle breaks up the phase space (or state space) of the two particles into finite volumes of size $\Delta x \Delta p_x \approx \hbar$ which leads to a finite number of microstates for given V and U.

Why are all of the derivatives of entropy like $\partial S/\partial U$, $\partial S/\partial V$, and $\partial S/\partial N$ parametrized in terms of T? I thought only $\partial S/\partial U$ had this definition. Is temperature a sort of proxy for entropy?

For the expression $T(\partial S/\partial V)_{U,N}$ which turns out to be the pressure P, one has to introduce the factor of T to make the definition reduce correctly to familiar cases like the ideal gas law $PV = NkT$. We will see this soon when we discuss Section 2.5 in Schroeder.

For the chemical potential $\mu = -T(\partial S/\partial N)_{U,V}$, I don't know the historical reason why the temperature factor was added to the entropy derivative but am guessing that there also had to be some case for which μ had to reduce to something already familiar and this wouldn't happen unless there was a factor of T included.

Note that from a formal point of view, the factors T do not matter whether they are there or not since, for a system in thermodynamic equilibrium, $T_1 = T_2$ for two subsystems and in the condition $\mu_1 = \mu_2$ or $P_1 = P_2$, the common temperatures would divide out from both sides, leaving the entropy derivatives equal as needed.

I'm not sure what the motivation for some of the initial definitions (like $1/T = dS/dU$) came from, they just seemed to work.

It is tricky to track down how and why various definitions were introduced and I don't know the history. People in the 19th century already appreciated that, whatever temperature was, equality of temperature between two systems implied that the systems were equilibrated with each other. We showed in lecture that dS/dU had this same property of being the same for two subsystems in equilibrium . So it would be natural to guess that dS/dU must be related to a temperature. As I mentioned briefly in lecture, the straightforward definition $T = dS/dU$ would predict incorrectly that energy flows from low temperature to high temperature systems, contrary to how we interpret temperatures via thermometers. One easy way to fix this is to define T via $1/T = dS/dU$ and that turns out to work in all cases.

Can you explain the significance of $dS_1/dU_1 = dS_2/dU_2$ again?

Not sure what to say beyond what I said in lecture and in my lecture notes. If that was not clear, I would be glad to meet with you in person and try to help you through the parts of the argument that were not clear.

Can you more fully explain the probability/counting methods discussed in class that led to the equation for $\Omega(N_H, N)$?

This is discussed on pages 50-51 of Schroeder. If this or my discussion was not clear, I would suggest we meet and try to work out a few simple examples to convince you how this works.

The argument is independent of Schroeder and is a classical result of elementary combinatorics, often taught in high school Algebra II. One problem in combinatorics is to determine how many ways you can arrange N distinct symbols, say the letters a, b, and c. The answer turns out to be $N!$ since there are N ways to choose the first symbol, times $N - 1$ independent ways to choose one of the remaining symbols, and so on. For three symbols, $3! = 3 \times 2 \times 1 = 6$ and the corresponding arrangement of symbols would be abc, acb, bac, bca, cab, and cba.

A second classic problem in combinatorics is to count how many ways you can place K identical objects in N slots, where $0 \leq K \leq N$. For example, if I have two identical coins labeled by the letter C and I have three slots SSS labeled by S, then there are three ways to place the two coins in two of the slot: CCS, CSC, and SCC. The argument on page 51 of Schroeder shows that the number of ways to place K objects in N slots is given by the binomial coefficient $B(N,K)$ given by

$$B(N,K) = \binom{N}{K} = \frac{N!}{K!(N-K)!}. \tag{1}$$

Fortunately, this is about all you have to know about combinatorics for the entire semester.

It is useful to observe that binomial coefficients arise naturally when raising expressions to integer powers:

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \ldots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n. \tag{2}$$

(You should try to see if you can understand directly *why* the binomial coefficients show up in this way: when you multiply $(a+b)$ by itself n times and expand everything out to get a sum of 2^n terms, how many ways will you obtain k factors of a and $n-k$ factors of b?) This connection allows one to use calculus and algebra to solve many difficult combinatoric problems, although we won't take advantage of this during the semester.

I know that it does, but how does the binomial distribution relate to gases and deriving the Boltzmann distribution?

The binomial distribution is not useful for gases or for deriving the Boltzmann distribution. It is useful only for counting the number of microstates for a system that consists of microscopic objects that each have only two possible states. A familiar example would be a set of coins each of which can be heads or tails, a quantum example would be a set of spin-1/2 particles in a magnetic field (a paramagnet). For reasons I will discuss in Tuesday's lecture, spin-1/2 particles can only be parallel or antiparallel to the external magnetic field and so act like pennies where "heads" means parallel, "tails" means antiparallel.

So the binomial distribution will let us count the number of microstates of a spin-1/2 paramagnet whose total magnetization (total number of up spins or total number of heads) is specified at a macroscopic level.

One of the conditions of thermodynamic equilibrium is chemical potential μ is constant but not necessarily the number of particles N. How does the entropy derivation account for variable N but is still in equilibrium?

Let me postpone explaining this until we get to Chapter 3, where we will calculate μ explicitly for several physical systems and then we can see how the condition $\mu_1 = \mu_2$ for two subsystems in diffusive equilibrium allows particle numbers to change.

Can a system come to equilibrium at a negative T?

The answer is yes, and in Chapter 3 we will discuss a physical example, which is a paramagnet consisting of many non-interacting atomic magnetic moments. Negative temperatures are not a common situation (and do not occur spontaneously in nature), it is interesting mainly for being logically possible and demonstrating that temperature can be defined in contexts that make no sense classically, e.g., the equipartition result $(1/2)m\langle v^2 \rangle = (3/2)kT$ suggests incorrectly that only non-negative temperatures are possible.

Why does mass contribute to entropy in an ideal gas?

We will discuss the specific reason soon when we come to Section 2.5 of Schroeder. Briefly, for an ideal gas consisting of N free non-interacting identical particles, the total energy U of the gas is related to the kinetic energy of the particles by this equation:

$$U = \frac{\mathbf{P}_1^2}{2m} + \frac{\mathbf{P}_2^2}{2m} + \ldots + \frac{\mathbf{P}_N^2}{2m}, \tag{3}$$

where $\mathbf{P}_i = (p_{x,i}, p_{y,i}, p_{z,i})$ is the three-dimensional momentum vector of the ith particle. This can be rewritten in the form

$$\left(\sqrt{2mU}\right)^2 = \mathbf{P}_1^2 + \ldots \mathbf{P}_N^2, \tag{4}$$

which defines mathematically the set of points that lie on the surface of a $3N$-dimensional hypersphere of radius $\sqrt{2mU}$. Thus increasing or decreasing mass m changes the surface area which in turn changes the number of possible momentum states (bigger surface area means more states are allowed). This in turn determines how many microstates are defined by the momentum vectors of the particles and this then affects the multiplicity Ω and so the entropy $S = k \ln \Omega$. It is rather neat that we will need to use the surface area of a $3N$-dimensional hypersphere to derive an expression for the entropy of an ideal gas consisting of N identical atoms.

I'm a little confused about the connection between the whole entropy topic and the probability/microstates/macrostates topic.

You are right to be confused because I have not yet explained the connection, this is the purpose of most of Chapter 2.

In the last lecture, I showed that *if* there was a quantity $S(U, V, N)$ that was additive over subsystems and that reached a maximum when an isolated system was in thermodynamic equilibrium, then we could start to understand why energy spontaneously flows from hot to cold regions, and that there was a quantity dS/dU that had to have the same value for all subsystems when equilibrium was attained, and this quantity was related to temperature.

And then I started saying something that won't be clear for several more lectures, that the way to discover the hypothetical S that makes all these ideas work was to consider a new concept called the multiplicity $\Omega(U, V, N)$ of an isolated macroscopic system, which counts the number of microstates consistent with specified global values of U, V, and N. It will then turn out that the quantity $k \ln \Omega$ behaves exactly like the entropy we postulated, e.g., it is additive over subsystems and increases monotonically with time until it reaches a maximum corresponding to equilibrium, and we will end up defining $S = k \ln \Omega$.

I remember from general chemistry that the entropy of the Earth is constantly increasing and from your lecture today, I am assuming it is because the Earth is a non-equilibrium system. What would be an example of something on Earth that is not in equilibrium and has an entropy of 0 or constant?

It is impossible for an isolated macroscopic nonequilibrium system to have a constant entropy, for reasons we will soon discuss in detail. However, if you consider a finite region of a nonequilibrium system, it is possible to force its entropy to be constant, but only by increasing the entropy of other systems so that the net entropy of the universe increases.

Thus there are refrigerators that can bring a substance down to extremely low temperatures (the world record is less than a nanokelvin (10^{-9} K) and maintain that low temperature so the entropy of that frozen substance is effectively zero and constant. But if you enlarge your point of view to the laboratory containing the refrigerator and cold substance, the entropy of the lab is steadily increasing over time as the refrigerator consumes energy and generates entropy.

It remains a difficult frontier scientific question why the universe is not already in a maximumm entropy state. The universe is clearly nonequilibrium and there are various processes heading toward equilibrium such as the nuclear reactions in the cores of stars, the equilibration of energy within galaxies, and so on. Since entropy is increasing now, it meant that entropy had to be much less in the past, but no one understands how to deduce the entropy of the universe at the time of the Big Bang nor understand why it was so small.

What is the connection between quantum mechanics and entropy?

It is given by Boltzmann's relation $S = k \ln \Omega$, where the multiplicity Ω counts the number of quantum microstates compatible with a specified macrostate labeled by values of U, V, and N. We will see numerous examples of this during the next few weeks so I don't give further details here, but the key point is that quantum mechanics provides a precise way to define the concept of a microstate and so count microstates.

What's up with the hovering frog?

I will show next lecture two short videos to illustrate the difference between paramagnetic materials and diamagnetic materials. The frog is diamagnetic and so can be suspended in space (without harm) by a sufficiently strong magnetic field gradient.

How does sepia coloring work?

I don't know how to answer this question, do you mean how does a software program like Photoshop apply sepia coloring to an image or historically where did sepia coloring arise? (And what is the connection of this question to the course?)