

Handout 7. Entropy

January 26, 2011

Contents

1	Reaching equilibrium after removal of constraint	2
2	Entropy and irreversibility	3
3	Boltzmann’s entropy expression	6
4	Shannon’s entropy and information theory	6
5	Entropy of ideal gas	10

In this lecture, we will first discuss the relation between entropy and irreversibility. Then we will derive the entropy formula for ideal gas,

$$S(N, V, E) = Nk_B \left[\ln \left(\frac{V}{N} \left(\frac{4\pi m E}{3N h^2} \right)^{3/2} \right) + \frac{5}{2} \right] \quad (1)$$

from the microcanonical (NVE) ensemble. To do so, we will

1. Establish Boltzmann’s entropy expression

$$S = k_B \ln \Omega(N, V, E) \quad (2)$$

where Ω is the number of microscopic states consistent with macroscopic state (N, V, E) .

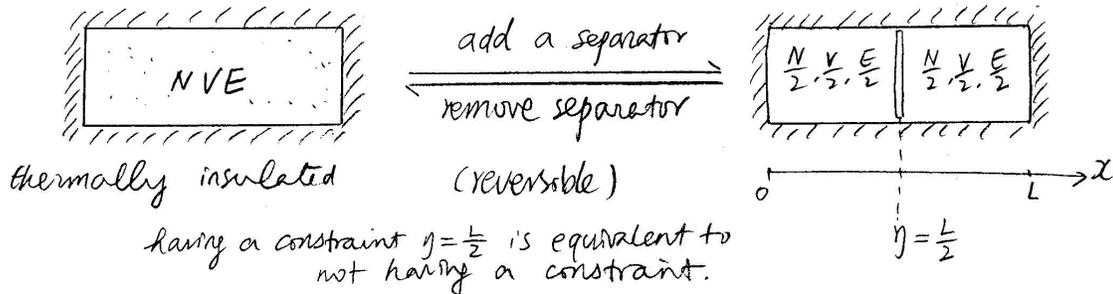
This is a special case of entropy defined in the information theory $S = \sum_{i=1}^n p_i \ln p_i$ when $p_i = \frac{1}{\Omega}$ for all i .

2. Count the number of microscopic state $\Omega(N, V, E)$, carefully.

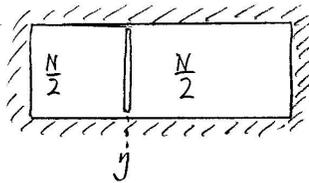
Reading Assignment, Reif §3.1-3.10.

1 Reaching equilibrium after removal of constraint

First let us consider a specific example of irreversible process caused by the removal of a constraint on a thermodynamic system (see Lecture Note 6.4. Irreversible processes).



now imagine a general situation (with constraint) with arbitrary η .



Imagine the separator is thermally conducting

Let the system settle down to an equilibrium state **under the constraint** η . In this state, the two sides should have the same temperature T . Given the ideal gas equation of state $PV = Nk_B T$, the two sides will not have the same pressure, unless $\eta = L/2$. This means that, in general, force must be applied on the separator to maintain the constraint η .

Let $S(N, V, E; \eta)$ be the entropy of the system in this state (with constraint η).

Now imagine removing the constraint η , by allowing the separator to slide in response to the pressure difference between the two sides.

Initially the separator may oscillate due to inertia effects. Imagine there is friction between the gas tank wall and the separator. Then the oscillation will eventually die down. Mechanical motion is converted to heat in the process, while the total energy remains constant (assuming the system is isolated from the rest of the world).

Entropy S will increase in this process.

When the system eventually settle down to the new equilibrium state (without constraint η), the new entropy is

$$S(N, V, E) = \max_{\eta} S(N, V, E; \eta) \quad (3)$$

If the system is in contact with a thermostat at temperature T , then N, V, T remain constants

during the irreversible process. One can show that the Helmholtz free energy decreases in the process (heat generated by friction flows to the thermostat). The Helmholtz free energy in the new equilibrium state is

$$A(N, V, T) = \min_{\eta} A(N, V, T; \eta) \tag{4}$$

2 Entropy and irreversibility

2.1 Birth of entropy

Entropy is first defined by German physicist Clasius, “On various forms of the laws of thermodynamics that are convenient for applications”, (1865).

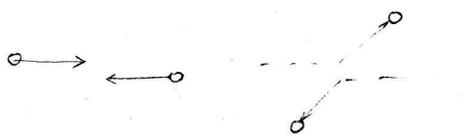
Entropy is the Greek word for “transformation” — Hans C. von Baeyer, “Maxwell’s Demon”, (1998), p.61.

Entropy — stays constant in reversible processes.

Entropy — always increases in irreversible processes.

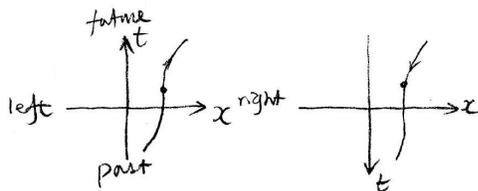
2.2 Entropy increase defines arrow of time

This should be very puzzling for everybody, because all microscopic theories of nature (e.g. classical mechanics, electromagnetism, relativity, quantum mechanics) are time reversible.



In classical mechanics, the trajectories of individual particles are completely reversible. One cannot tell whether the movies is playing forward or backward.

Einstein was very puzzled by the arrow of time.



In the theory of relativity, time is just one of the axes of the 4-dimension “space-time”.

Past-v.s.-future is not so different from left-v.s.-right.

Einstein remarked “... this separation between part, present and future is an illusion, albeit a stubborn one.” — “Maxwell’s Demon” p.129.

Indeed the illusion of “the arrow of time” is a very stubborn one. It is one of the most fundamental aspect of human experience.

Obviously, we remembers the past, but are usually not so good at predicting the future.

Why cannot we “remember” the future?

Clausius stated that: Entropy always increases as we move into the future.

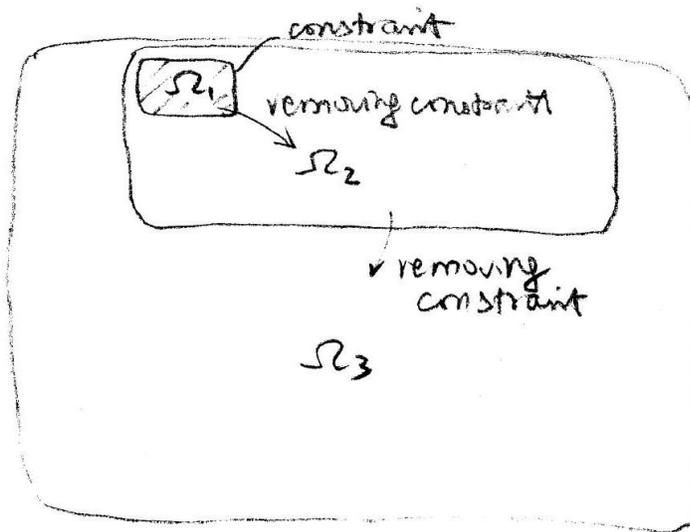
What is this entropy anyway? How can we explain it in terms of the microscopic particles (atoms)?

Why does it always increase with time?

2.3 Boltzmann’s entropy expression

$$S = k_B \ln \Omega \quad (5)$$

where Ω is the number of microscopic states consistent with the macroscopic state, e.g. (N, V, E) .



Phase space

- As we remove some internal constraint on the system, it “diffuses” out of its original volume in the phase space into a much larger volume in phase space: $\Omega_1 \rightarrow \Omega_2$.
- Microscopic dynamics is reversible. It does not prevent the system from spontaneously moving back into region Ω_1 — but to see that happen you will have to wait a LONG time.
- The waiting time easily exceeds the age of the universe, which has been only 14 billion years anyway. The age of the earth is about 4.5 billion years.

- If you want to reverse the change in your (very limited) life time, you will have to spend work.
- Hence irreversibility is connection to the finite span of our existence (yes, mortality, alas). — After all, God (being immortal) may not see irreversibility.
- Entropy increases because the system's initial condition has lower entropy than the equilibrium state. No irreversibility can be observed if the system is already in thermal equilibrium.
- Almost all irreversible processes (e.g. life) on earth are fuelled (ultimately) by sun light. The evolution of stars obeys thermodynamics, i.e. entropy always increases in a burning star. This means entropy is increasing everywhere in the universe.
- Future — will it be the “heat death of the universe”? Maybe. But not so eminent. I suggest we don't worry about it.
- Past — If the entropy has been always increasing, then the universe must have a beginning. The initial state of the universe must have very low entropy.
- “Big bang” — 14 billion years ago, all energy/mass of the universe is concentrated at one point (smaller than one atom). This initial state has very low entropy. Boltzmann (1844-1906) already realized that!

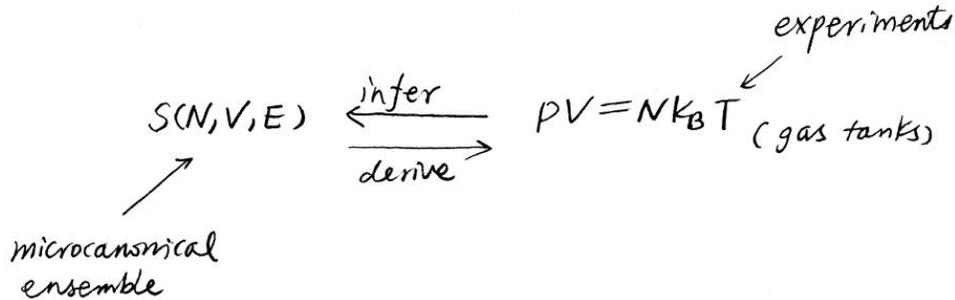


3 Boltzmann's entropy expression

S, V, N, E are all state functions in thermodynamics.

Furthermore, S is a function of (N, V, E) . $S(N, V, E)$ is a material specific equation of state from which all thermodynamic relations of this material can be derived.

In thermodynamics, the equation of state is obtained from experiments:



In statistical mechanics, we can derive $S(N, V, E)$ from fundamental dynamics of the microscopic particles. Other equation of state (e.g. $pV = Nk_B T$) can be derived from this fundamental relation.

Statistical Mechanics	Thermodynamics
Microcanonical (N, V, E) ensemble. There are many microscopic states $\{q_i, p_i\}$ consistent with (N, V, E) . Suppose the number of such microscopic states is Ω .	Macroscopic variables: S, N, V, E

Boltzmann proposes:

$$S = k_B \ln \Omega \tag{6}$$

where $k_B = 1.38 \times 10^{-23} \text{ J}\cdot\text{K}^{-1}$ is Boltzmann's constant.

Boltzmann's tombstone in Vienna reads: $S = K \log W$.

4 Shannon's entropy and information theory

Later on, people realize that Boltzmann's entropy formula is a special case of the entropy expression in Shannon's information theory.

$$S = -K \sum_{i=1}^n p_i \log(p_i) \tag{7}$$

This expression is called Shannon Entropy or Information Entropy.

Unfortunately, in the information theory, the symbol for entropy is H and the constant k_B is absent. We have changed their notation to avoid confusion.

4.1 How to understand Shannon's information entropy

Entropy measures the degree of our lack of information about a system. Suppose you throw a coin, which may land either with head up or tail up, each with probability $\frac{1}{2}$. Then we have some uncertainty about the outcome of each "experiment". The uncertainty can be quantified by a positive number S .

Now suppose you modified the coin (somehow) that you know for sure that each time you throw it, it will always land with head up (i.e. probability = 1). Then there is no uncertainty about the possible outcome of each "experiment". The information entropy should be $S = 0$.

In general, consider an experiment with n possible outcomes, each with probability p_i , $i = 1, \dots, n$ with normalization condition $\sum_{i=1}^n p_i = 1$.

We are looking for a general formula $S(p_1, p_2, \dots, p_n)$ that can characterize the uncertainty in all these experiments.

Intuitively, we expect

- (1) $S(p_1 = 1/n, p_2 = 1/n, \dots, p_n = 1/n)$ should be the maximum among all values $S(p_1, \dots, p_n)$ with a fixed n .
- (2) $S(p_1 = 0, p_2 = 1, \dots, p_n = 0) = 0$ should be the minimum (no uncertainty).

But to develop a general formula for arbitrary $\{p_i\}$ seems impossible!

That's why Shannon is so smart. He did it!

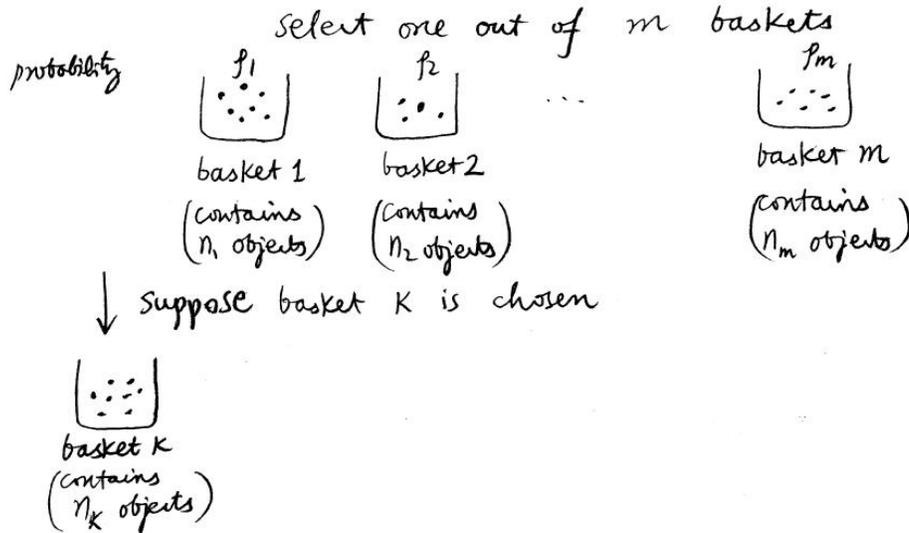
4.2 Derivation of Shannon entropy

Shannon showed that if we assume the entropy function should satisfy a set of reasonable properties then there is only one possible expression for it!

These conditions are:

- (1) $S(p_1, p_2, \dots, p_n)$ is a continuous function.
- (2) $f(n) \equiv S(1/n, 1/n, \dots, 1/n)$ is a monotonically increasing function of n .
- (3) Composition law for compound experiments:

$$S(AB) = S(A) + \sum_{k=1}^m p_k S(B|A) \quad (8)$$



To understand the composition law, consider the following compound experiment AB .

Experiment A : Select one out of m baskets. (Suppose basket k is selected.)

Experiment B : Select one out of n_k objects from basket k .

In the end, all we did was to select one object from $N = \sum_{k=1}^m n_k$ objects.

Let the total uncertainty of the compound experiment be $S(AB)$.

The composition law states that we should be able to quantify the uncertainty at each step and add them together.

The uncertainty in 1st step is simply $S(A) = S(p_1, \dots, p_m)$.

The uncertainty in 2nd step depends on the outcome of the 1st step. When basket k is selected, let the uncertainty of 2nd step be $S(B|A_k)$.

The expected value of the uncertainty in 2nd step is $\sum_{k=1}^m p_k S(B|A_k)$.

Hence we expect $S(AB) = S(A) + \sum_{k=1}^m p_k S(B|A_k)$

To show Shannon's entropy is the only expression that satisfy these three conditions, we design a special compound experiment.

Consider an experiment in which we randomly pick 1 object out of N objects. The probability of picking any object is $1/N$. The uncertainty of this experiment is

$$S(1/N, 1/N, \dots, 1/N) \equiv f(N) \tag{9}$$

Now imagine that we artificially break down the experiment into two steps. Imagine that we divide the N objects into m groups.

Each group k contains n_k objects and k runs from 1 to m , i.e. $\sum_{k=1}^m n_k = N$. In the first step, we randomly pick one of the m groups, the probability of picking group k is

$$p_k = \frac{n_k}{N} \quad (10)$$

In the second step, we pick one object from the selected group. Suppose group k is selected in the first step, then the probability of selecting one object in the second step is $1/n_k$.

The uncertainty in the first step is $S(p_1, \dots, p_m)$.

The expected value for the uncertainty in the second step is $\sum_{k=1}^m p_k f(n_k)$. Hence

$$f(N) = S(p_1, \dots, p_m) + \sum_{k=1}^m p_k f(n_k) \quad (11)$$

Now consider a special case of $n_1 = n_2 = \dots = n_m = n$, $p_k = 1/m$ for all k .

Every group has n objects, $n \cdot m = N$.

$$f(N) = S(1/m, \dots, 1/m) + \sum_{k=1}^m \frac{1}{m} f(n) \quad (12)$$

$$f(mn) = f(m) + f(n), \quad \text{for arbitrary } m, n \quad (13)$$

Thus

$$f(m) = K \log m \quad (14)$$

where K is arbitrary positive constant.

Plug it back to the general case

$$K \log N = S(p_1, \dots, p_m) + \sum_{k=1}^m p_k K \log n_k \quad (15)$$

$$\begin{aligned} S(p_1, \dots, p_m) &= K \log N - \sum_{k=1}^m p_k K \log n_k \\ &= K \sum_{k=1}^m p_k \log N - K \sum_{k=1}^m p_k \log n_k \\ &= K \sum_{k=1}^m p_k \log \frac{N}{n_k} \\ &= -K \sum_{k=1}^m p_k \log p_k \end{aligned} \quad (16)$$

Of course, in the special case of $p_k = 1/\Omega$, $k = 1, 2, \dots, \Omega$, we get back to Boltzmann's entropy expression

$$\begin{aligned} S &= -K \sum_{k=1}^{\Omega} (1/\Omega) \log(1/\Omega) \\ &= K \log \Omega \end{aligned} \tag{17}$$

5 Entropy of ideal gas

Now, let's count how many (Ω) microscopic states are consistent with macroscopic state variables N, V, E .

In the microcanonical ensemble, all microstates with $E \leq H(\{q_i\}, \{p_i\}) \leq E + \Delta E$ have equal probability density. (Yes, we need to talk about probability density because q_i and p_i are continuous variables).

From Lecture Notes "Microcanonical ensemble"

$$\Omega = \frac{\tilde{\Omega}(E + \Delta E) - \tilde{\Omega}(E)}{N! h^{3N}} \tag{18}$$

$\tilde{\Omega}(E)$ is the phase space volume occupied by all micro states with $H(\{q_i\}, \{p_i\}) \leq E$. $N!$, h^{3N} are correction factors from quantum mechanics.

For an ideal gas, $H = \sum_{i=1}^{3N} \frac{p_i^2}{2m}$, $\tilde{\Omega}(E)$ is the volume of $3N$ -dimensional sphere ($R = \sqrt{2mE}$) times V^N .

$$\tilde{\Omega} = \frac{(2\pi m E)^{3N/2}}{\left(\frac{3N}{2}\right)!} \cdot V^N \tag{19}$$

$$\begin{aligned} \Omega &= \frac{1}{N! h^{3N}} \cdot \frac{\partial \tilde{\Omega}(E)}{\partial E} \cdot \Delta E \\ &= \frac{1}{N! h^{3N}} \cdot \frac{(2\pi m)^{3N/2} \cdot E^{3N/2-1}}{(3N/2 - 1)!} \cdot V^N \cdot \Delta E \end{aligned} \tag{20}$$

$$\begin{aligned} \log \Omega &= -\log N! - 3N \log h + (3N/2) \log(2\pi m) \\ &\quad + (3N/2 - 1) \log E - \log(3N/2 - 1)! + N \log V + \log \Delta E \end{aligned} \tag{21}$$

Apply Stirling's formula, and only keep terms of the order N ,

$$\begin{aligned}
 \log \Omega &\approx -N \log N + N - N \log h^3 + N \log(2\pi m E)^{3/2} \\
 &\quad - (3N/2) \log(3N/2) + 3N/2 + N \log V \\
 &= N[-\log N + 1 - \log h^3 + \log(2\pi m E)^{3/2} \\
 &\quad - \log(3N/2)^{3/2} + 3/2 + \log V] \\
 &= N \left[\log \frac{(2\pi m E)^{3/2} V}{N h^3 (3N/2)^{3/2}} + 5/2 \right] \\
 &= N \left[\log \left(\frac{V}{N} \left(\frac{4\pi m E}{3N h^2} \right)^{3/2} \right) + \frac{5}{2} \right] \tag{22}
 \end{aligned}$$

Notice that Planck's constant h remains in the final expression, but ΔE does not, because the latter does not contribute a term of the order N .

Given the fundamental equation of state $S(E, V, N)$, we can derive all other properties of the ideal gas (Handout 6. Thermodynamics §3.2).

The statistical distribution of momentum of one particle in ideal gas in the microcanonical ensemble has been derived in Handout 5. Microcanonical Ensemble §2.2.