

6 Entropy

Introduction

These notes are largely devoted to things you do not need to know in detail about entropy to complete the course, but which I hope will help to make your understanding deeper even if you only understand the outlines of the arguments. I hope that, by going into rather more depth than the general physics books go, I can leave you with something you will still remember a year from now.

6.1 Disorder and lack of information

The **entropy** of a system is a measure of its **disorder**. Another way of looking at this, introduced by E. T. Jaynes, is that entropy is a measure of the **lack of information** about a system.

The **information** contained in a message of a certain length is quantified as the logarithm of the number of possible messages of that length that could have been sent instead. If the logarithm is taken to base 2, the unit of information is a **bit**, and if it is taken to base 256 it is a **byte**. An empty floppy disk has no information recorded on it (other than formatting information), whereas a full floppy disk may have 1,400,000 bytes of information, or 11,200,000 bits. Much of this information is actually redundant, and can be coded more compactly when space is a serious issue. The conversion of a text file to its compressed .gzip form or of a digital photograph to a .jpg file are examples of what can be done with a more efficient coding than a blind coding of single characters to single bytes, or a pixel by pixel coding of a picture. One source of redundancy is that different characters in the set of characters defined by 8 binary digits occur with very different frequencies in normal text. Claude Shannon, in his mathematical theory of information, showed that the information capacity of a message with N characters, the logarithm of the number of different messages that can be sent, is given by

$$I = -N \sum_n P_n \log_2 P_n \text{ bits} , \quad (1)$$

where the P_n are the frequencies of the different characters. If the different 8 bit characters came with equal probabilities $1/256$, this formula would just give the information capacity of an N byte device as $I = 8N$ bits.

If we have a physical system such as a gas contained in a volume V , we may only know the enclosing volume, the number N of atoms of a particular sort, say ^{40}Ar atoms, and their average kinetic energy $3k_B T/2$. Where each atom is, and what is its momentum, is unknown to us, and the entropy is defined as k_B times the natural logarithm of the number $\Omega(N, V, T)$ of arrangements of the atoms compatible with what we know:

$$S = k_B \ln[\Omega(N, V, T)] . \quad (2)$$

Since this measures what we do not know rather than what we do know it is the negative of information; the fact that it is conventionally measured in joules per kelvin rather than bits or bytes is irrelevant. If we knew where all the atoms were and what their velocities were we would have a ridiculous amount of information, perhaps 10^{26} bytes, and the system would have no entropy. Rather quickly, however, our knowledge would become out of date as the atoms hit one another and bounced off at varied angles, our information would become irrelevant to the present situation, and the entropy would revert to the equilibrium value it would have had if we had not made this enormously complicated measurement.

6.2 Thermodynamics and statistical mechanics of entropy

The Carnot analysis of reversible heat engines showed that the heat added to a system in a reversible transfer of heat can be written as the integral of $T dS$, where S is the entropy, a conserved quantity in reversible processes only. This is combined with the first law of thermodynamics to give the equation

$$dU = -p dV + T dS , \quad (3)$$

which says in this new form that the change in internal energy is the sum of the work done on the system and the heat added. By writing it in this form we assume that at each stage of the process there is thermal and mechanical equilibrium, and that the system has a well-defined pressure and temperature at all times. When there are temperature gradients, pressure changes, or local flow velocities, so that more variables are needed to describe the system, this formula is no longer valid. The assumption of an equation of state involving only pressure and temperature hides an assumption of reversibility.

We can read from eq. (3) that, at constant volume, we can write the heat capacity in the two different forms

$$C_v = \left(\frac{\partial U}{\partial T}\right)_V = T \left(\frac{\partial S}{\partial T}\right)_V , \quad (4)$$

where I am using the convention common in thermodynamics that the subscript on the parentheses shows what is kept fixed when a partial derivative is taken. In the Notes on Kinetic Theory and Statistical Mechanics I wrote the average energy in the form

$$U = \sum_n E_n P_n , \quad (5)$$

where E_n is the energy of a state n of the system, and P_n is its probability, given by the Boltzmann distribution as

$$P_n = \frac{\exp(-E_n/k_B T)}{\sum_m \exp(-E_m/k_B T)} . \quad (6)$$

Equation (4) shows that the entropy can be written as

$$S(T, V) = \int^T \frac{1}{T} \left(\frac{\partial U}{\partial T} \right)_V dT . \quad (7)$$

It is possible to go from eqs. (5) and (6) to the solution

$$S = -k_B \sum_n P_n \ln P_n \quad (8)$$

of eq. (7), but it is much easier to differentiate the two equations (5) and (8), and verify that eq. (4) is satisfied. In these manipulations, the constant volume condition implies that the energy levels E_n remain constant. Equation (8) for the entropy is of exactly the same form as (1) for the information capacity of a system, except that the information capacity is defined in a dimensionless form, and the logarithm that gives the entropy is multiplied by a constant with dimensions of J/K.

6.3 Ideal gas again

For the ideal monatomic gas we can get the entropy from eq. (3) and the two relations

$$pV = Nk_B T , \quad U = \frac{3}{2} Nk_B T . \quad (9)$$

At constant temperature we know the internal energy of an ideal gas is constant, so the first law of thermodynamics, with the heat expressed as TdS , gives

$$pdV = TdS , \quad \text{or} \quad \int Nk_B dV/V = \int dS .$$

and this gives the change of entropy along an isotherm as

$$\Delta S = Nk_B \Delta(\ln V) .$$

At constant volume we have

$$dU = TdS = (3/2)Nk_B dT , \text{ or } (3/2) \int Nk_B dT/T = \int dS .$$

and this gives the change of entropy at constant volume as

$$\Delta S = (3/2)Nk_B \Delta(\ln T) .$$

These combine to give the result

$$S = Nk_B \left\{ \ln \left[\frac{V}{N} (k_B T)^{3/2} \right] + \text{constant} \right\} . \quad (10)$$

We have actually gone a little further than the equations allow us to go, by writing V/N in the argument of the logarithm rather than just V , since this makes the entropy an **extensive quantity** that doubles when V and N are doubled simultaneously.

L. Boltzmann and J. Willard Gibbs used eq. (2) to work out the entropy from kinetic theory. There are two parts of this calculation. First we consider the number of ways the N atoms can be embedded in the real space of volume V . Then we consider the number of ways they can be embedded in the three-dimensional momentum space. This looks like a silly question, because you cannot count the number of ways that a continuous variable can be assigned, but it becomes a meaningful question if you say that you are not concerned with exactly where each atom is and how fast it is traveling, but are only concerned with the approximate values of its position and momentum. After this theory was developed, quantum theory came along to specify the limits to how closely we can hope to specify position and momentum.

First, we consider the volume V divided up into \mathcal{N} small volumes (cells) of size v , where

$$\mathcal{N}v = V, \quad \mathcal{N} \gg N . \quad (11)$$

For each atom there are \mathcal{N} different places to put it, so the total number of ways of assigning the atoms to cells v is \mathcal{N}^N , which would give a contribution to the entropy, from eq. (2), of

$$S_1 = k_B \ln \mathcal{N}^N = Nk_B \ln(V/v) . \quad (12)$$

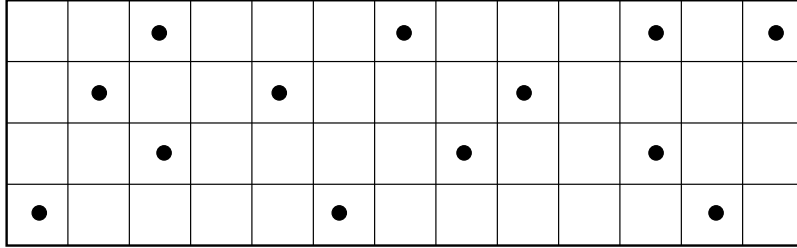


Figure 1: 52 cells of volume $v = V/52$ containing 13 atoms (or one typical bridge hand).

Gibbs did not like this answer, because it gives the result that if V and N are both doubled, the entropy does not just double, but there is also an additional term $Nk_B \ln 2$. He got rid of this by arguing that you cannot tell atoms apart, so you do not care which atom goes into a particular cell; the most you could hope to know (the zero of entropy) is which cells have atoms in them. If \mathcal{N} is sufficiently large that most atoms are in cells with no other atoms in them ($\mathcal{N} \gg N$), the number of ways of rearranging the N atoms in a given set of cells is $N!$, so the total number of ways of choosing which N cells are occupied is approximately $\mathcal{N}^N/N!$.

This is like the difference between the number of different ways in which 13 cards can be dealt from a deck of 52 cards, for which the answer is

$$52 \times 51 \times 50 \times \dots \times 40 = \frac{52!}{39!}, \quad (13)$$

and the number of different bridge hands of 13 cards that can be dealt from a deck of 52 cards, if you do not care what order they are dealt in, for which the answer is

$$\frac{52 \times 51 \times 50 \times \dots \times 40}{13!} = \frac{52!}{39!13!}. \quad (14)$$

To use this result to calculate the entropy we need to know the approximation for $N!$ found by Stirling in the eighteenth century, which is

$$N! \approx (N/e)^N \sqrt{2\pi N}; \quad (15)$$

see my notes on Statistical Mechanics. The corrected result for the contribution to the entropy is

$$S_1 = k_B \ln[\mathcal{N}^N/N!] \approx k_B \ln\left[\frac{Ve}{vN}\right]^N = Nk_B \left[\ln \frac{V}{N} + \ln \frac{e}{v}\right], \quad (16)$$

where I have dropped the relatively small term contributed by the logarithm of $\sqrt{2\pi N}$. This now has the same volume dependence as eq. (10).

There is a similar argument for the number of ways to assign the momenta of the particles. This is complicated by the fact that we have to assign them in such a way that the average kinetic energy works out to be $3k_B T/2$, and to do this I would have to do real statistical mechanics, and would derive the Maxwell-Boltzmann distribution properly. I do not do that, but instead put in the temperature sloppily, by saying that the kinetic energy of the particles can have any value up to a maximum value ϵ_{\max} , which I take to be of the order of $k_B T$, say $\alpha k_B T$. The three components of momentum of an atom define a position in **momentum space**. The volume of momentum space which corresponds to energy less than ϵ_{\max} is a sphere whose radius p_{\max} is given by $p_{\max}^2/2M = \epsilon_{\max}$, and this has a volume

$$V' = \frac{4\pi}{3} p_{\max}^3 = \frac{4\pi}{3} (2M\epsilon_{\max})^{3/2}. \quad (17)$$

This is divided into cells each of volume v' , and the number of ways of putting N atoms into these V'/v' cells is $(V'/v')^N$, and the contribution to the entropy is

$$S_2 = k_B \ln(V'/v')^N = Nk_B [\ln(2M\epsilon_{\max})^{3/2} + \ln \frac{4\pi}{3v'}]. \quad (18)$$

The dimensions of the cell volume v' , or of the sphere volume V' , are momentum cubed.

Equations (16) and (18) can be put together to yield

$$S = S_1 + S_2 = Nk_B \left\{ \ln \left[\frac{V}{N} (2M\alpha k_B T)^{3/2} \right] + \ln \frac{4\pi e}{3vv'} \right\}. \quad (19)$$

This is of the form we want, just like equation (10), and a proper derivation would have given the same answer except for a slightly different constant factor in the argument of the logarithm. The only unsatisfactory feature is the term $-Nk_B T \ln(vv')$, which seems quite arbitrary, because there was no reason to pick a particular size of the cells. Here again, quantum theory came to the rescue afterwards, because quantum theory says that the right value of vv' is h^3 , where h is Planck's constant.

We have discussed monatomic gases in detail. More complex gases have a higher entropy per molecule because there is an extra disorder associated with the internal motion, primarily rotational, of the molecule, and this leads to a higher power of T in the argument of the logarithm; $5/2$ instead of $3/2$ for the familiar diatomic gases like oxygen, nitrogen and hydrogen.

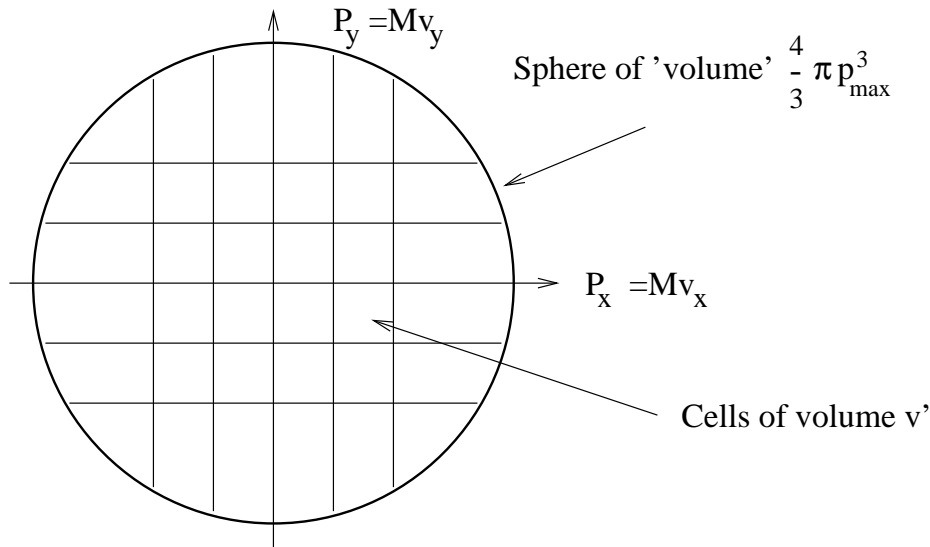


Figure 2: Sphere of momentum space divided into cells

The entropy of a mixture of gases can be worked out, at reasonably low pressures, just as if each constituent occupied the entire volume independently of the other constituent.

6.4 Entropy of liquids and solids

The latent heat of vaporization of a liquid is T times the entropy change when the liquid turns to vapor. Most of this difference comes from the change in volume. There may be an even greater difference than one would estimate from the change in volume, because the molecules in a liquid are pressed quite tightly together, are more ordered in their arrangement in space and have less freedom to rotate, than the molecules of an ideal gas. This is why the latent heat of vaporization is typically rather large.

A solid usually has a somewhat higher density than the liquid, but the main difference is that the atoms or molecules are arranged in the regular array that is characteristic of a crystal. This extra order means that the solid has a lower entropy than the liquid at the same temperature, but there is not a lot of difference between the two phases, so the latent heat of melting is considerably smaller, typically, than the latent heat of vaporization.

When the temperature is lowered the solid gets more and more ordered, until, in most cases, the disorder would completely disappear if zero Kelvin

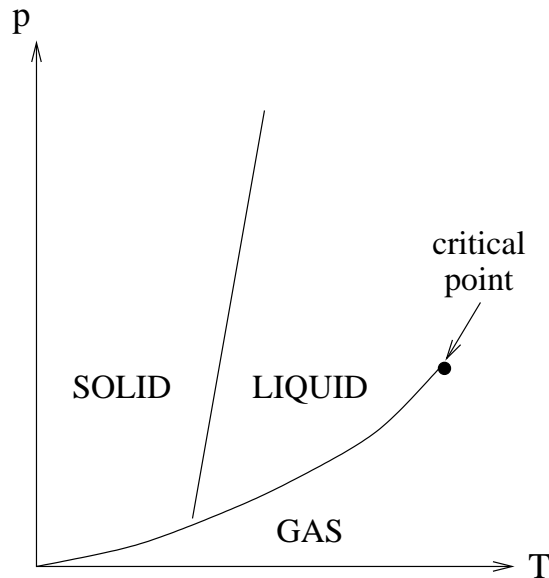


Figure 3: Phase diagram for solid, liquid and vapor

could be reached. This is Nernst's Third Law of Thermodynamics, that the entropy goes to zero in the limit of zero Kelvin. Liquids freeze (except for helium, which is special), solids become completely ordered, alloys either phase separate or form regularly ordered compounds, and so on.

6.5 Thermodynamics of phase transitions

One of the earliest applications of thermodynamic reasoning was to determine the slope of the coexistence line in the p - T plane for two phases such as water and ice, or water and steam. This was worked out by Clausius, who developed thermodynamics in ignorance of Carnot's work, and Clapeyron, who knew Carnot's work and passed it on to William Thomson (Lord Kelvin).

One way to derive the Clausius-Clapeyron equation is to consider a Carnot cycle that works entirely with a mixture of the two phases – let us say water and steam. In the coexistence region the isotherms in the p - V diagram are completely flat, so that the volume changes at constant temperature without any change of pressure. The heat absorbed during the isothermal expansion at the higher temperature is the latent heat of vaporization times the mass of fluid evaporated, while the work done in the cycle is the area enclosed

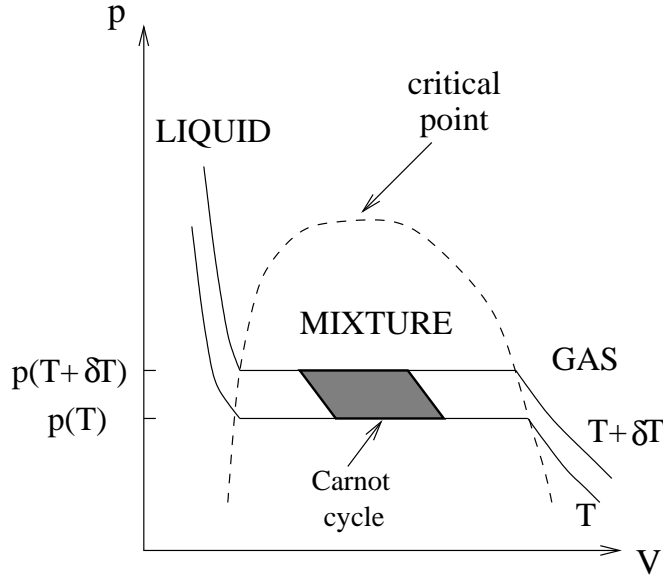


Figure 4: Isotherms and coexistence curve in pV diagram near liquid and vapor critical point

between the two flat portions of the isotherm. If this is done at an arbitrary temperature difference, as it would have to be for an idealized steam engine or refrigerator compressor, there would be an awkward problem of handling the adiabatic portions of the cycle, but it turns out to be very easy if we take the cycle to be between two very close temperatures T and $T + \delta T$. In this case the pressure on the low temperature isotherm is $p(T)$ and on the higher temperature isotherm is $p(T + \delta T) \approx p(T) + p'(T)\delta T$, where $p(T)$ gives the vapor pressure of the water (the coexistence curve in the p - T plane). If a volume V_l of liquid is converted into a volume V_v of vapor, the work done in the cycle is

$$W \approx (V_v - V_l)[p(T + \delta T) - p(T)] \approx (V_v - V_l)p'(T)\delta T, \quad (20)$$

where I have neglected the work done in the adiabatic expansion and compression, and the difference between the volume change at the higher temperature and at the lower temperature, because both of these give contributions of order $(\delta T)^2$. This process is shown in fig. 4.

The heat absorbed at the higher temperature, the latent heat of the liquid

which is evaporated, can be written as

$$Q_2 = T(S_v - S_l) , \quad (21)$$

where S_l and S_v are the entropies of the same mass of liquid and vapor. We know that the efficiency of the engine must be $\delta T/T$, so we have now got

$$\frac{\delta T}{T} = \frac{W}{Q_2} = \frac{V_v - V_l}{S_v - S_l} \frac{p'(T)\delta T}{T} . \quad (22)$$

If the left and right side of this equation are equated, the equation

$$\frac{dp}{dT} = \frac{S_v - S_l}{V_v - V_l} = \frac{L_{lv}}{T(V_v - V_l)} \quad (23)$$

is obtained, where L_{lv} is the latent heat of vaporization. This is the Clausius-Clapeyron equation.

Exactly the same argument can be used for the melting transition, and gives for the melting pressure as a function of temperature

$$\frac{dp}{dT} = \frac{S_l - S_s}{V_s - V_l} = \frac{L_{sl}}{T(V_l - V_s)} . \quad (24)$$

In most cases the melting pressure is an increasing function of temperature, because the latent heat is positive, and the solid is more dense than the liquid, but for ice the solid is less dense than the liquid, so the melting temperature is lowered by increasing pressure and the slope given by eq. (24) is negative. There is actually at least one very odd case where the solid is more dense than the liquid, but the slope is negative because the latent heat is negative. This is the isotope of helium ^3He below about 0.5 K.

6.6 Free energy

The first and second laws of thermodynamics can be summarized by eq. (3) for changes in the internal energy, which can be written in terms of partial derivatives as

$$\left(\frac{\partial U}{\partial V}\right)_S = -p, \quad \left(\frac{\partial U}{\partial S}\right)_V = T , \quad (25)$$

where the subscript with each partial derivative is our usual reminder of what variable is being kept constant. Since the mixed second partial derivative $\partial^2 U / \partial V \partial S$ is independent of the order in which the derivatives are taken, this gives the Maxwell relation

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V . \quad (26)$$

There are three closely related quantities that are used in similar ways to the internal energy. These are the enthalpy

$$H = U + pV, \quad dH = V dp + T dS, \quad (27)$$

the Helmholtz free energy

$$F = U - TS, \quad dF = -p dV - S dT, \quad (28)$$

and the Gibbs free energy

$$G = U + pV - TS, \quad dG = V dp - S dT. \quad (29)$$

Each of these quantities yields another Maxwell relation, and each quantity is heavily used in various branches of science. Enthalpy is easier to use than internal energy under the constant pressure conditions that are more commonly encountered than constant volume conditions. The heat capacity at constant pressure is given by

$$C_p = \left(\frac{\partial H}{\partial T} \right)_p = T \left(\frac{\partial S}{\partial T} \right)_p,$$

and, if you want to find latent heats of fusion in a book of tables like *CRC Handbook of Chemistry and Physics*, you may find them listed as “enthalpies of fusion”. Latent heats are measured at constant pressure and temperature, and give

$$T(S_l - S_s) = H_l - H_s.$$

Enthalpy also gets used in discussions of engines and of heats of reaction. Helmholtz free energy is used a lot in statistical mechanics when one wants to work out the state of a system with a given number of molecules in a given volume and at a given temperature. The usual procedure is to work out the Helmholtz free energy as a function of volume and temperature, and then other quantities such as internal energy heat capacity, pressure, and bulk modulus can be found from the partial derivatives of F with respect to V and T .

The Gibbs free energy is a function of the two intensive variables, pressure and temperature, which are imposed by external conditions, and it is used for discussions of phase changes and of chemical equilibrium. The *chemical potential* of a molecule is essentially the Gibbs free energy per molecule.

The Maxwell relation that comes from eq. (29) by working out $\partial^2 G / \partial p \partial T$ in two different ways is

$$\left(\frac{\partial V}{\partial T} \right)_p = - \left(\frac{\partial S}{\partial p} \right)_T, \quad (30)$$

which relates the volume coefficient of thermal expansion to the heat given out under isothermal compression. This is obviously a useful relation, and not altogether expected.

If you look at eq. (29) and think about the fact that, when ice is melted isothermally or water is evaporated at constant temperature, neither the pressure nor the temperature change, you should realize that melting or vaporization at constant temperature must produce no change in G . This tells you that if two phases of a substance, like ice and water, are in equilibrium with one another they have the same Gibbs free energy per unit mass. Furthermore, on the vapor pressure curve, the coexistence curve, since $G_v = G_l$, eq. (29) gives

$$0 = dG_v - dG_l = (V_v - V_l)dp - (S_v - S_l)dT, \quad (31)$$

which gives the Clausius-Clapeyron equation (23) again.

As a last example of the kind of thing you can do with thermodynamic relations when you become expert I will give the relation between heat capacity at constant volume and at constant pressure, which is written in terms of entropy as

$$C_p - C_V = T\left(\frac{\partial S}{\partial T}\right)_p - T\left(\frac{\partial S}{\partial T}\right)_V. \quad (32)$$

From the chain rule for partial derivatives we can write

$$\left(\frac{\partial S}{\partial T}\right)_p = \left(\frac{\partial S}{\partial T}\right)_V + \left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p. \quad (33)$$

At constant temperature the partial derivatives are just like ordinary derivatives, so we can write

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial S}{\partial p}\right)_T \left(\frac{\partial p}{\partial V}\right)_T, \quad (34)$$

and this can be rewritten, using the Maxwell relation (30), as

$$\left(\frac{\partial S}{\partial V}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_p \left(\frac{\partial p}{\partial V}\right)_T. \quad (35)$$

Combination of eqs. (32), (33) and (35) gives

$$C_p - C_V = -T\left(\frac{\partial V}{\partial T}\right)_p^2 \left(\frac{\partial p}{\partial V}\right)_T. \quad (36)$$

This formula, which you can check for the ideal gas if you feel like it, gives the difference in heat capacity per unit volume as the product of the temperature in Kelvin, the bulk modulus, and the square of the volume coefficient of expansion. If you can follow this last paragraph you are really doing well.