Jimmy Yang

PHYS 203 Notes

Notes for PHYS 203 adapted from Rio's Science One thermodynamics notes. Some sections might be poorly written and disorganized so please take it with a grain of salt. If you have any questions or comments related to the notes, free to email me at marethyu@student.ubc.ca.
Last modified: February 8, 2023

Contents

1   Probability and Statistics

Most of the material covered in this section is also covered in my MATH 318 notes.

1.1   Permutations

Permutations are combinations where order matters. How permutations are calculated depends on whether repetitions are allowed or not. The number of permutations made from choosing k from n items when repetitions are allowed is n^k.

When repetitions are not allowed, the number of permutations is just $n!$. Here is one way to understand where this formula came from: For the first item, there are n choices to place and since no repetitions are allowed, there are $n-1$ choices left for the second item and so on.

EXAMPLE 1.

Calculate the probability of obtaining exactly $k = 3$ heads and tails for a total of 6 coin tosses.

There are the total of $2^6 = 64$ possible outcomes. The number of times when the combination $HHHTTT$ occurs is $\frac{6!}{3!3!} = 20$. Therefore the probability is $\frac{20}{64} = \frac{5}{16}$.

For k values larger than 3, the probability of getting the equal number of heads and tails will be low. It will be close to zero as $k \to \infty$.

EXAMPLE 2.

Find the number of words created from rearranging letters $ALLELE$.

There are two Es and three Ls. Notice that there is no difference when you reorder these letters. Each word created from two Es and three Ls corresponds to 2!3! permutations.

The number of permutations is 6! which includes duplicates. To get the final answer, we divide it by 2!3! which gives $\frac{6!}{2!3!} = 60$ total words.

Speaking of probability, there are no arrangements of $ALLELE$ that is more likely than others. This makes $ALLLEE$ equally likely.

To calculate the number of permutations made from choosing k out of n items with no repetition:

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

It is usually denoted by $_n\mathrm{P}_k$.

1.2   Combinations

EXAMPLE 3.

How many ways to pick two out of $\{a, b, c, d\}$?

Initially, there are four choices for the first letter of a ordered pair. After picking one, there are three choices left. There are $4(3) = 12$ ordered pairs:

> $ab, ba, ac, ca, ad, da, bc, cb, bd, db, cd, dc.$ However, we don't care about the order (ie. *ab* and *ba* is considered the same) so we divide by 2! which gives: $\frac{12}{2} = 6$ combinations.

To pick k out of n items, the total number of combinations is the number of ways to create ordered combinations (permutations) of size k divided by the number of permutations of k items in combinations. The calculation process is outlined below:

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \frac{\frac{n!}{(n-k)!}}{k!} = \frac{n!}{k!(n-k)!}$$

We use $\binom{n}{k}$ to denote $\frac{n!}{k!(n-k)!}$. It is read as "n choose k." This is a binomial coefficient. Another way to think about combinations is that it is the same as finding how many ways to rearrange n cards where k cards are coloured red and $n-k$ cards are coloured green. Also note that:

$$\binom{n}{n-k} = \frac{n!}{(n-k)!(n-(n-k))!} = \frac{n!}{(n-k)!k!} = \binom{n}{k}$$

$\binom{n}{k}$ makes a Gaussian curve as shown below.

Check number two

Below are plots of $\binom{N}{k}$ for k on the horizontal axis for $N = 10, 40$ and 160.
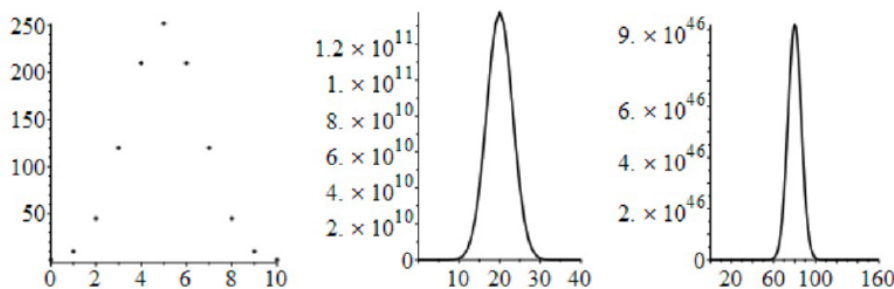
By using k/N as the variable on the horizontal axis and $\frac{\binom{N}{k}}{\binom{N}{k}_{MAX}} = \frac{\binom{N}{k}}{\binom{N}{N/2}}$ on the vertical axis, we scale these three plots and put them together:
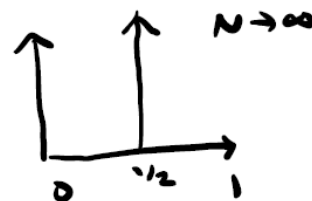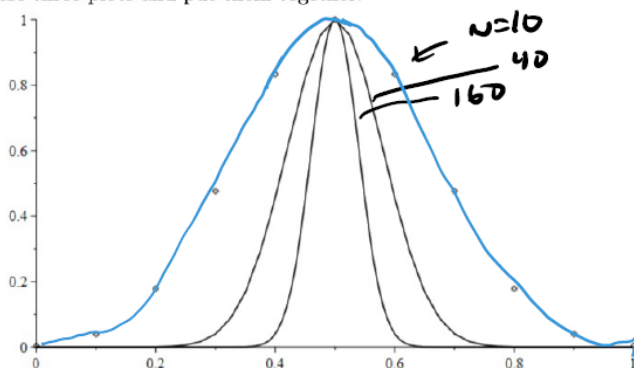
What do you learn about the answer to question 4 from examining these plots?

Figure 1: Gaussians for different N

$\binom{n}{k}$ is maximized when $k = n/2$.

There is an interesting observation that can be used to calculate binomial coefficients. Suppose you toss a coin n times and k of n tosses are heads. The number of possible ways to get k

heads out of n tosses is given by $\binom{n}{k}$. To find this binomial coefficient, consider two cases: the first toss is H and the first toss is not H. When the first toss is H, the remaining tosses $n-1$ must have $k-1$ heads and the number of possible ways it can happen is $\binom{n-1}{k-1}$. When the first toss is not H, the remaining tosses $n-1$ must have k heads and the number of possible ways it can happen is $\binom{n-1}{k}$. The number of possible ways to get k heads out of n tosses is the number of possible ways to get H as the first toss plus the number of possible ways not to get H as the first toss:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

This is called "Pascal's rule." It can be used to construct Pascal's triangle.

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
\end{array}
$$

$\binom{n}{k}$ gives the value at nth row and kth column (both zero-indexed) of the triangle. For example, in the third row and second column, $\binom{2}{1} = 2$. Notice that the total in each row is always 2^n.

1.3   Frequentist Definition of Probability

Consider an experiment with distinct outcomes $O_1, O_2, \ldots, O_i$. Repeat this experiment N times. Let N_i be the number of times outcome O_i occurs (out of N). Probability $P(O_i)$ of O_i is given by:

$$P(O_i) = \lim_{N \to \infty} \frac{N_i}{N}$$

If we know that the probability of some event E is $P(E)$, our best prediction for how many times we will see event E occur out of n attempts is $nP(E)$.

1.4   Example: Coin Toss Experiment

Consider an experiment where a fair coin is tossed k times. We repeat this process for N times. We define score is the number of heads appear after k flips which takes the values from 0 to k. We count how many times each score appears then divide by N to normalize so that the total is one.

Take $k = 4$ for example, there are $2^k = 16$ possible outcomes for the sequence of four coins which is listed below:

$(H,H,H,H), (T,H,H,H), (H,T,H,H), (H,H,T,H), (H,H,H,T), (T,T,H,H),$
$(T,H,T,H), (T,H,H,T), (H,T,T,H), (H,T,H,T), (H,H,T,T), (T,T,T,H)$
$(T,T,H,T), (T,H,T,T), (H,T,T,T), (T,T,T,T)$

The table below shows the number of sequences containing a given number of heads:

no. of heads	0	1	2	3	4
no. of sequences	1	4	6	4	1

To get the probability of getting a particular score after running this experiment for an infinite number of times, divide the bottom row with 16:

no. of heads	0	1	2	3	4
no. of sequences	1/16	1/4	3/8	1/4	1/16

If we define the score to be the number of heads divided by the number of fair coin tosses N. We notice that as N gets bigger, the fraction $f = \frac{\text{no. heads}}{N}$ will fluctuate between 0.25 and 0.75 then converges to $\frac{1}{2}$. See figure 1.
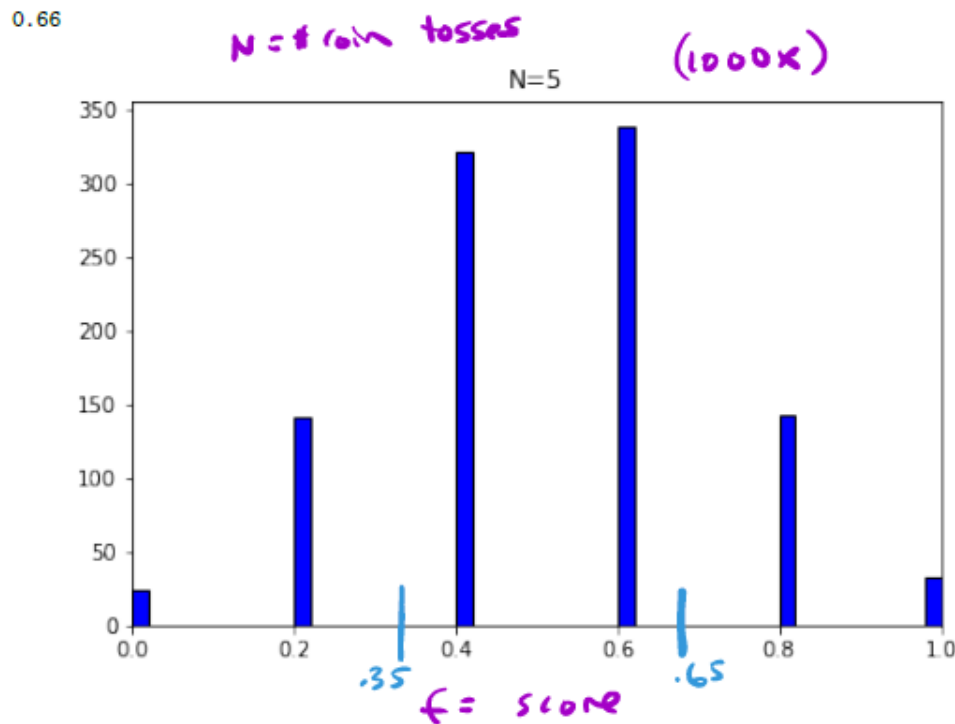

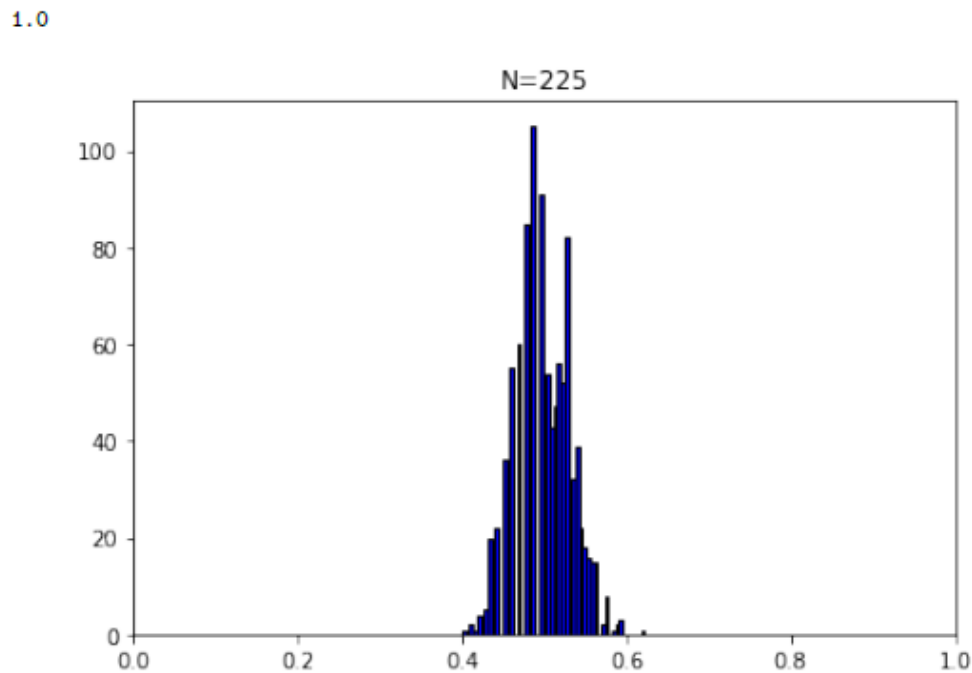Figure 2: Score histogram for 1000 coin tosses carried out $N = 5$ times

Figure 3: Score histogram for 1000 coin tosses carried out $N = 225$ times

If $N = \infty$, the histogram will look like the delta function centered at 0.5.

Regarding example one, this is a bit different compared to this example. Behaviour as something approaches infinity and behaviour at infinity are two different things. In example one, as $k \to \infty$, the probability of getting *exactly* k heads and tails will approach zero. Any finite k is far from infinity. Now back to the example we are discussing right now, when $N = \infty$ the fraction is exactly 0.5 and, of course, there is the exact match in heads and tails, both infinite.

1.5 Probability for Binomial Random Variable*

This section is optional.

Suppose you want to calculate the probability of getting a certain combination of heads and tails from an unfair coin. It depends on whether the order matters or not.

> EXAMPLE 4.
>
> Consider an unfair coin, let the probability of getting H on a single toss be 0.6 and the probability of getting T is 0.4. Find (1) the probability of obtaining the sequence $THHTTTHT$ (exactly, in this order) (2) the probability that a sequence of eight coin tosses would have 5 tails and 3 heads in any order.
>
> (1) Coin toss events are independent of each other. The probability will be calculated as follows $0.4 \cdot 0.6 \cdot 0.6 \cdot 0.4 \cdot 0.4 \cdot 0.4 \cdot 0.6 \cdot 0.4 = 0.6^3 0.4^5 = 0.00221$. By symmetry, this is the same as the probability of getting the sequence $TTTTTHHH$ (or any other arrangements involving 3 heads and 5 tails) in exact order.
>
> (2) The probability of getting 5 tails and 3 heads in any order can be calculated using the addition law for disjoint events:

$$P(3 \text{ heads}) = P(\{THHTTTHT\}) + P(\{THTHTTTH\})$$
$$+ P(\{TTHTHHTT\}) + \cdots$$
$$+ P(\{\text{other arrangements of } 3H \text{ and } 5T\})$$

In this case, we need to find the number of ways to permute 5 tails and 3 heads which is given by $\binom{8}{3} = 56$. Using our answer to (1), the probability is therefore:

$$P(3 \text{ heads}) = 56 \cdot 0.00221 = 0.12386$$

This example presents the basic intuition for how to calculate probabilities for binomial random variables.

Suppose you are doing an experiment where the probability of success is p and failure is $1 - p$. The probability of getting exactly k successes out of n trials can be calculated by the below formula:

$$P(k \text{ successes out of } n \text{ trials}) = \binom{n}{k} p^k (1 - p)^{n-k}$$

EXAMPLE 5.

Using the same unfair coin from the previous example, define a score to be (number of Heads) / (number of tosses). Find the probability that the score is between 0.4 and 0.8 when $N = 9$ coins are tossed. What is the answer if N is very large?

To get the score between 0.4 and 0.8, we need to get 4 to 7 heads. So the probability is

$$P(\text{score between 0.4 and 0.8}) = P(4 \text{ or } 5 \text{ or } 6 \text{ or } 7 \text{ heads})$$
$$= P(4 \text{ heads}) + P(5 \text{ heads})$$
$$+ P(6 \text{ heads}) + P(7 \text{ heads})$$
$$= \sum_{k=4}^{7} \binom{9}{k} 0.6^k 0.4^{9-k}$$
$$= 0.8301$$

If N is very large, the probability will reach 1.

1.6   Standard Deviation

In the coin experiment, we noticed that as N tends to infinity, the score converges to $\frac{1}{2}$. This means the larger the N is more accurate the measurement is. This is reflected in the uncertainty of the measurement. This uncertainty is called "standard deviation," average distance from mean.

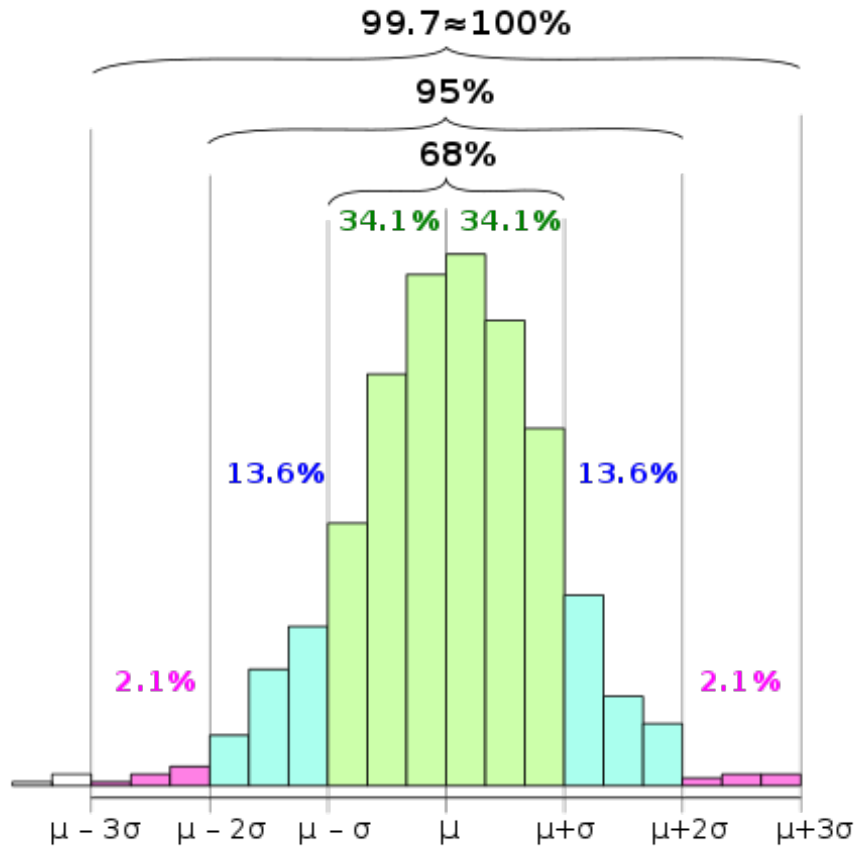One can estimate uncertainty using the 68–95–99.7 rule.

Figure 4: Distribution of mass in Gaussian

One standard derivation σ left and right from mean value μ will give about 68% of the overall mass. Two standard derivations left and right from the mean will give about 95% of the overall mass.

Standard deviation squared for a single measurement $N = 1$ is given by $\sigma_1^2 = \langle (x - \langle x \rangle)^2 \rangle$ where x are all possible outcomes of a single measurement and $\langle x \rangle$ is theoretical mean which is given by:

$$\langle x \rangle = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} x_i$$

EXAMPLE 6.

Find σ_1^2 for a dice roll experiment.

There are six possible outcomes for this experiment which are $x = 1, 2, 3, 4, 5, 6$. The mean of these outcomes is $\frac{1+2+3+4+5+6}{6} = 3.5$. The standard deviation squared is given by

$$\sigma_1^2 = \frac{1}{6} \sum_{x=1}^{6} (x - 3.5)^2 = 2.91667$$

The standard deviation squared after N measurements is given by $\sigma_N^2 = \langle (x_N - \langle x \rangle)^2 \rangle$ where x_N is the mean of x after N measurements:

$$x_N = \frac{1}{N} \sum_{i=1}^{N} x_i$$

EXAMPLE 7.

Consider a coin tossing experiment, find standard deviation after two measurements.

The possible outcomes after two measurements are HH, TH, HT, and TT.

$$\sigma_2^2 = \left\langle \left(\frac{1}{2} \sum_{i=1}^{2} x_i - \langle x \rangle \right)^2 \right\rangle = \left\langle \left(\frac{1}{2} (x_1 + x_2) - \langle x \rangle \right)^2 \right\rangle$$

$$= \frac{1}{4} \left[\left(\frac{1}{2}(0+0) - 0.5 \right)^2 + \left(\frac{1}{2}(0+1) - 0.5 \right)^2 \right.$$

$$\left. + \left(\frac{1}{2}(1+0) - 0.5 \right)^2 + \left(\frac{1}{2}(1+1) - 0.5 \right)^2 \right] = \frac{1}{8}$$

We have $\sigma_2 = \frac{1}{\sqrt{8}}$. The standard deviation for a single measurement is $\sigma_1 = \sqrt{\frac{1}{2}\left((0-0.5)^2 + (1-0.5)^2\right)} = \frac{1}{2}$. So $\sigma_2 = \frac{\sigma_1}{\sqrt{2}}$.

EXAMPLE 8.

Show that $\sigma_N^2 = \frac{\sigma_1^2}{N}$.

$$\sigma_N^2 = \langle (x_N - \langle x \rangle)^2 \rangle$$

$$= \left\langle \left(\left(\frac{1}{N} \sum_{i=1}^{N} x_i \right) - \langle x \rangle \right)^2 \right\rangle$$

$$= \left\langle \frac{1}{N^2} \left(\sum_{i=1}^{N} x_i \right)^2 - \frac{2\langle x \rangle}{N} \left(\sum_{i=1}^{N} x_i \right) + \langle x \rangle^2 \right\rangle$$

We will use the following fact that $\left(\sum_{i=1}^{N} x_i \right)^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i x_j$ and $\langle x_i \rangle = \langle x \rangle$ to simplify the equation further:

$$= \left\langle \frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j\right) - \frac{2\langle x\rangle}{N}\left(\sum_{i=1}^{N} x_i\right) + \langle x\rangle^2 \right\rangle$$

$$= \frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{j=1}^{N} \langle x_i x_j\rangle\right) - \frac{2\langle x\rangle}{N}\left(\sum_{i=1}^{N} \langle x_i\rangle\right) + \langle\langle x\rangle^2\rangle$$

$$= \frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{j=1}^{N} \langle x_i x_j\rangle\right) - \frac{2\langle x\rangle}{N}N\langle x\rangle + \langle x\rangle^2$$

$$= \frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{j=1}^{N} \langle x_i x_j\rangle\right) - \langle x\rangle^2$$

$$= \frac{1}{N^2}\left(\sum_{i=1}^{N} \langle x_i^2\rangle + \sum_{i\neq j} \langle x_i\rangle\langle x_j\rangle\right) - \langle x\rangle^2$$

$$= \frac{1}{N^2}\left(\sum_{i=1}^{N} \langle x^2\rangle + \sum_{i\neq j} \langle x\rangle^2\right) - \langle x\rangle^2$$

$$= \frac{1}{N^2}\left(N\langle x^2\rangle + (N^2 - N)\langle x\rangle^2\right) - \langle x\rangle^2$$

$$= \frac{1}{N}(\langle x^2\rangle - \langle x\rangle^2)$$

We have $\sigma_1^2 = \langle(x - \langle x\rangle)^2\rangle = \langle x^2 - 2x\langle x\rangle + \langle x\rangle^2\rangle = \langle x^2\rangle - 2\langle x\rangle\langle x\rangle + \langle x\rangle^2 = \langle x^2\rangle - \langle x\rangle^2$, so the final result is $\sigma_N^2 = \frac{\sigma_1^2}{N}$.

Since it doesn't matter how you group the experiments or number of trials, the std error of the mean for 10 coin tosses 100 times is the same as that for 100 coin tosses 10 times.

EXAMPLE 9.

Show that the standard deviation of 100 experiments of 10 coin flips is the same as that of 10 experiments of 100 coin flips.

$$\sigma(100 \text{ experiments of 10 coin flips}) = \frac{\sigma_{10}}{\sqrt{100}} = \frac{\sigma_1}{\sqrt{100}\sqrt{10}}$$

$$\sigma(10 \text{ experiments of 100 coin flips}) = \frac{\sigma_{100}}{\sqrt{10}} = \frac{\sigma_1}{\sqrt{10}\sqrt{100}}$$

In the coin toss experiment, we can approximate the standard deviation of scores using $\sigma_N \propto \frac{1}{\sqrt{N}}$. The score now is given by $f = \frac{1}{2} \pm \sigma_N = \frac{1}{2} \pm \frac{1}{\sqrt{N}}$. As $N \to \infty$, the score is $f = \frac{1}{2}$, as expected.

EXAMPLE 10.

A box, divided into two identical parts by a partition with a small hole in it, has 1 mole of ideal gas in it. After a while, the hole is plugged up, and the number of atoms on each side of the partition is counted. How large (roughly) would you expect the difference between the two counts to be within? Recall Avogadro's number $N_A = 6\times10^{23}$.

An ideal gas is a kind of gas where non-interacting particles each acts independently of

---

the others. This is like the coin-flipping experiment we discussed earlier. The amount of atoms on one part of the box is given by:

$$N_A \left(\frac{1}{2} \pm \frac{1}{\sqrt{N_A}} \right) = \frac{N_A}{2} \pm \sqrt{N_A}$$

We see that the uncertainty given by $\sigma = \sqrt{N_A}$ which corresponds to the order of the magnitude of about 10^{12} atoms.

2 Fundamentals

2.1 Microstates and Macrostates

A specific microscopic configuration of a system is called a microstate. The more loose term is macrostate which specifies the general or macroscopic properties of a system. In the coin-flipping experiment, a microstate contains the state of each coin and a macrostate is the number of heads or tails.
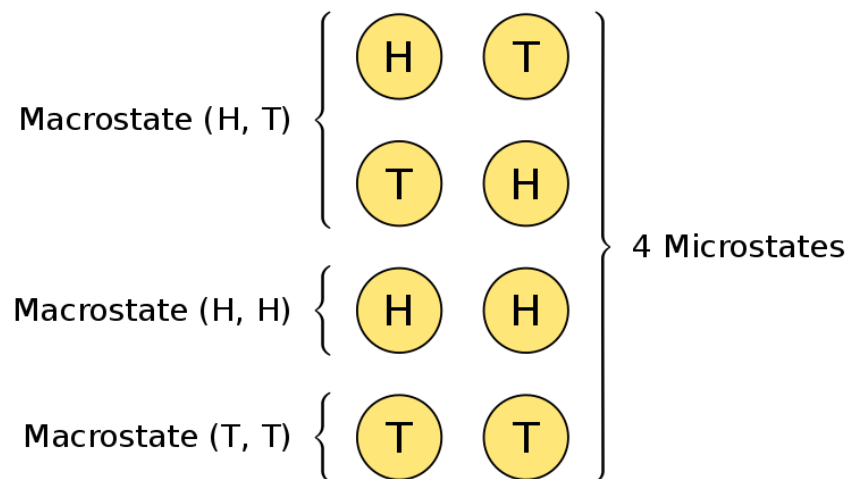

Figure 5: Microstate and macrostate of the experiment involving two coin flips

Each macrostate has some number of microstates and it is called multiplicity, denoted by Ω. For the image above, $\Omega(0 \text{ heads}) = 1$, $\Omega(1 \text{ heads}) = 2$, $\Omega(2 \text{ heads}) = 1$. The total number of microstates is $\Omega(\text{all}) = 1 + 2 + 1 = 4$. The probability of getting n heads is $\frac{\Omega(n)}{\Omega(\text{all})}$.

Macroscopic state: description of a system based on macroscopic observables only (not microscopic, atom-by-atom ones). E.g. total energy vs the distribution of energy. Other examples include temperature, volume and pressure.

EXAMPLE 11.

You have 13 cards, 6 red (R) and 7 black (B). You shuffle them and place them face up in two rows. The top row has 5 cards, the bottom row 8. If you do this many times, what number of red cards will the top row have most often?

The total number of arrangements is $\Omega(\text{all}) = \frac{13!}{6!7!} = 1716$.

To determine the most likely number of red cards, we count the number of arrangements when the top row has $0, 1, 2, 3, 4, 5$ red cards. For example, to count the number of arrangements for the top row with 2 red cards, $\binom{5}{2} = \frac{5!}{2!3!} = 10$. To get the total number of arrangements we multiply the results, for example, the total number of arrangements with 3 red cards at the top is $10(56) = 560$. After that, we calculate probabilities and compare.

Let n_T and n_B denote the number of red cards at the top and bottom rows respectively. Let Ω_T denote the number of arrangements for the top row created with n_T red cards, same for Ω_B. $\Omega_{TB} = \Omega_T \Omega_B$ is the total number of arrangements.

n_T	0	1	2	3	4	5
Ω_T	1	5	10	10	5	1
n_B	6	5	4	3	2	1
Ω_B	28	56	70	56	28	8
Ω_{TB}	28	280	700	560	140	8
$\frac{\Omega_{TB}}{\Omega(\text{all})}$	0.016	0.163	0.408	0.326	0.082	0.005

From the table, we see that the top row will most likely have 2 red cards.

Also notice that $28 + 280 + 700 + 560 + 140 + 8 = 1716$.

If we were to compute the average number of red cards in the top row, we use the following relation: $\sum_{n=0}^{5} nP_n$ where P_n is the probability of getting n red cards in the top row. This gives:

$$0(0.016) + 1(0.163) + 2(0.408) + 3(0.326) + 4(0.082) + 5(0.005) = 2.31$$

This playing card example has applications in statistical mechanics for modelling two-state systems which we will see later.

2.2 Thermal Equilibrium and Zeroth Law of Thermodynamics

The basic fact of thermodynamics is when objects of different temperature contact each other, heat flows from the warmer object to the colder object. The colder object warms up until thermal equilibrium is reached. This transfer of energy is an irreversible process.

Recall that heat is a spontaneous flow of energy on a microscopic scale. "warmer" means to be able to give more heat and "colder" means to be able to give off less heat. Temperature is the measure of the ability to give off heat.

Zeroth law of thermodynamics states that if two objects A and B are immediately in thermal equilibrium when brought together, and A and C are immediately in thermal equilibrium when brought together, then B and C are also in thermal equilibrium.

Thermal equilibrium is a state that any two objects will reach if they are in thermal contact (able to exchange heat) long enough. They might have different energy but have the same temperature. Being in equilibrium means things are not changing macroscopically. A thermometer is a system that allows us to visualize equilibrium.

A heat bath or reservoir is a system large enough s.t. the temperature is essentially unchanged after supplying or absorbing energy with much smaller systems.

EXAMPLE 12.

You have a standard deck of 52 cards which consists of 26 red cards and 26 black cards. The deck is shuffled well, what is the probability that the first five cards contain 3 red cards and 2 black cards? What is the probability if the deck is large enough?

We count the number of ways to draw 3 red cards from the set of 26 red cards: $\binom{26}{3}$ and the number of ways to draw 2 black cards from the set of 26 black cards: $\binom{26}{2}$. Then multiply the result and divide it by the number of ways to draw 5 cards from the deck: $\binom{52}{5}$. This gives $P(3R, 2B) = \frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}}$.

Alternative approach: Count the number of combinations of a 5 card hand that contains 3 red and 2 black: $\binom{5}{3}$. Then count the number of ways to rearrange the remaining cards in the deck: $\binom{52-5}{26-3}$. Finally count the number of ways to arrange the deck: $\binom{52}{26}$. This gives $P(3R, 2B) = \frac{\binom{5}{3}\binom{52-5}{26-3}}{\binom{52}{26}}$.

To approximate the probability when the deck is large, we rewrite things a bit:

$$
\begin{aligned}
P(3R, 2B) &= \frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} \\
&= \frac{\frac{26!}{3!23!}\frac{26!}{2!24!}}{\frac{52!}{5!47!}} \\
&= \frac{5!}{3!2!}\frac{\frac{47!}{23!24!}}{\frac{52!}{26!26!}} \\
&= \frac{5!}{3!2!}\frac{(52-5)!26!26!}{52!(26-3)!(26-2)!} \\
&= \frac{5!}{3!2!}\frac{26(26-1)(26-2)26(26-1)}{52(52-1)(52-2)(52-3)(52-4)} \\
&\approx \frac{5!}{3!2!}\frac{26^5}{56^5} \\
&\approx \frac{5!}{3!2!}\frac{1}{2^5}
\end{aligned}
$$

We conclude that the suitable approximation is $\frac{\binom{5}{3}}{2^5}$. This gets better when the deck is larger.

This example illustrates the concept of heat bath. The larger the deck is, taking out a few cards will have almost no effect on the probability.

As a demonstration of the thermodynamic process, consider two coupled lattices of oscillators where each particle has either 0 or 1 quanta of energy.
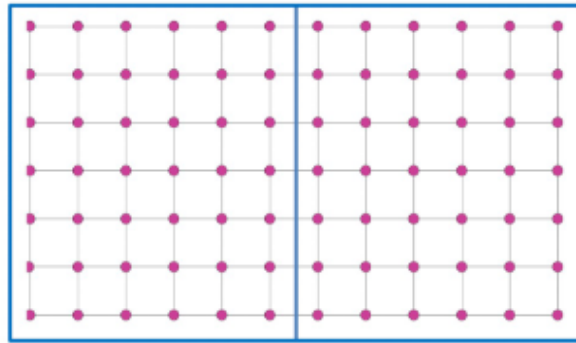
Figure 6: Two-state system example

The two systems can exchange energy.

EXAMPLE 13.

We have two systems A and B. A has 5 particles and B has 8 particles. There are 6ϵ units of energy. How many units of energy is likely in A? Average units of energy in A?
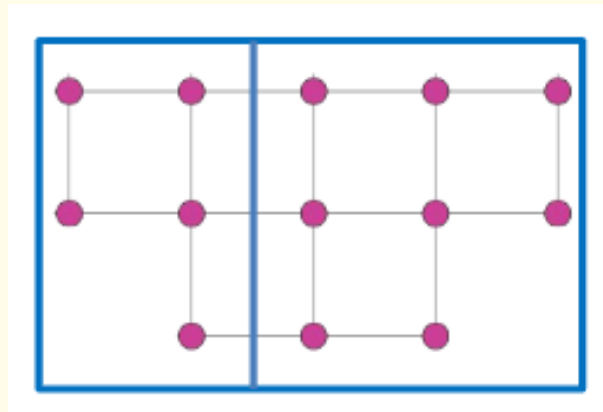
Figure 7: Two-state system

This is the same as the card rearrangement example we did earlier, so the answer is 2ϵ. The average energy units in A is also 2.31ϵ. This is the amount of energy reached when the system reaches thermal equilibrium.

Figure 8: If we start with 5 units of average energy in A, as time increases it will decrease to 2.31ϵ

This example illustrates that a system can flow its energy spontaneously (with some fluctuations) until it stops at the most likely macrostate (the one with the highest

---

multiplicity). This is the second law of thermodynamics.
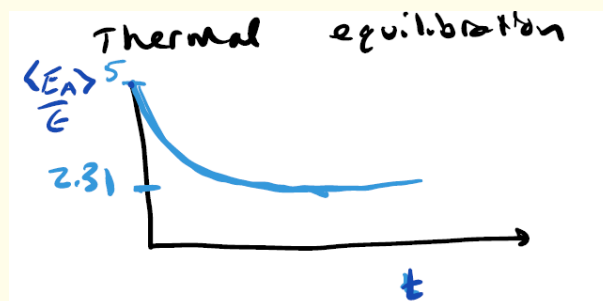
So why we can treat thermodynamical systems as card games? It is based on the fundamental assumption in statistical mechanics: "All microstates are equally likely" $P(s) = \frac{1}{\Omega_{total}}$.

We will now explore the probability distribution for the two-state system discussed previously.

A=5 atoms
B=8 atoms
n=6 quanta of energy
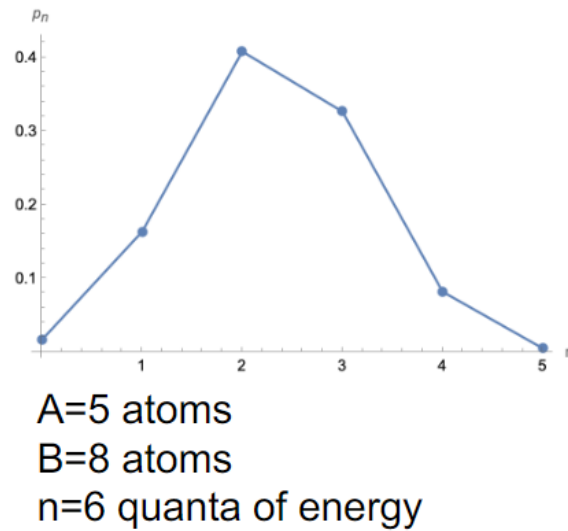
Figure 9: Probability distribution for the two-state system

If we increase the number of atoms and energy quantas, we get

A=500 atoms
B=800 atoms
n=600 quanta of energy

Figure 10: Increasing atoms and energy

The curve becomes more smooth and sharp. As the number of atoms and energy increases to infinity, the system is said to be approaching the thermodynamic limit.

Thermodynamic limit is when the number of components of the system is large enough so that measurable fluctuations away from the most likely macroscopic state practically never occur.

Fluctuations happen when there is a sudden deviation of energy from its average value. It happens often for small systems.

EXAMPLE 14.

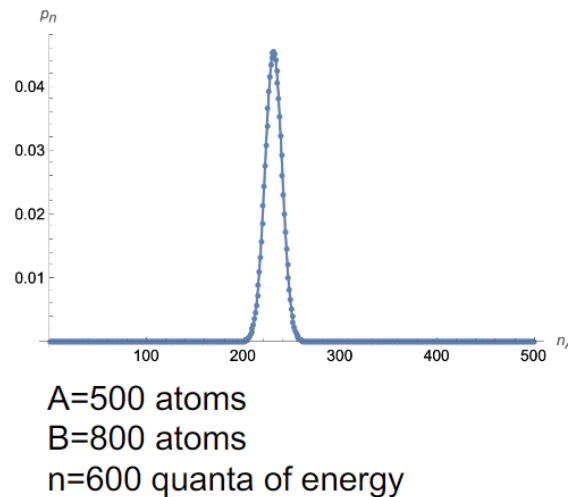What is the likelihood of fluctuation to pack all the possible quanta of energy to the small container for: (a) 6ϵ in $(5, 8)$ atoms (b) 60ϵ in $(50, 80)$ atoms (c) 600ϵ in $(500, 800)$ atoms.

We need to find the probability for a fluctuation to pack almost all energy units into the small container.

(a) From the previous example, we found that $P = \frac{\Omega_{TB}(5,1)}{\Omega(\text{all})} = 0.005$.

(b) $P = \frac{\Omega_{TB}(50,10)}{\Omega(\text{all})} = \frac{\binom{50}{50}\binom{80}{10}}{\binom{130}{60}} = 3e - 26$

(c) $P = \frac{\Omega_{TB}(500,100)}{\Omega(\text{all})} = \frac{\binom{500}{500}\binom{800}{100}}{\binom{1300}{600}} = 3e - 259$

EXAMPLE 15.

- Ω counts the number of microstates with associated energy, thus is a function of energy.

- Thermal equilibrium is reached when $\Omega_A \Omega_B$ is maximized.

- As you increase Ω_A, Ω_B will decrease and vice versa (so arrows point in opposite directions).

- The slope of $\ln \Omega$ curve will give inverse temperature β.

- $\ln \Omega$ curves have a global maximum, so as energy increases, the curve will look like a downward parabola.

To make sense of the above example, we will now derive the formula for temperature. Consider two systems A and B in direct thermal contact with each other, the total energy shared between

systems is E_T. The multiplicity (number of microstates) of the AB system when A has energy E_A and B has energy $E_B = E_T - E_A$ is

$$\Omega_{AB}(E_A, E_B) = \Omega_A(E_A)\Omega_B(E_B)$$

Suppose system A receives Δ units of energy from system B which loses Δ units of energy in exchange, define the difference between systems to be

$$\text{diff} = \ln \Omega_{AB}(E_A + \Delta, E_B - \Delta) - \ln \Omega_{AB}(E_A, E_B)$$

We apply the logarithm identity $\ln(ab) = \ln(a) + \ln(b)$ to get

$$\text{diff} = (\ln \Omega_A(E_A + \Delta) + \ln \Omega_B(E_B - \Delta)) - (\ln \Omega_A(E_A) + \ln \Omega_B(E_B))$$

$$\text{diff} = \ln \Omega_A(E_A + \Delta) - \ln \Omega_A(E_A) + \ln \Omega_B(E_B - \Delta) - \ln \Omega_B(E_B)$$

Let $f(E) = \ln \Omega(E)$ and assume $\Delta \ll E$, derivative $\frac{\partial f}{\partial E}$ at E_0 can be approximated as follows:

$$\left. \frac{\partial f}{\partial E} \right|_{E=E_0} \approx \frac{f(E_0 + \Delta) - f(E_0)}{\Delta}$$

With this approximation in mind, diff becomes:

$$\text{diff} = \left. \frac{\partial \ln \Omega_A}{\partial E} \right|_{E=E_A} \Delta + - \left. \frac{\partial \ln \Omega_B}{\partial E} \right|_{E=E_B} \Delta$$

$$\text{diff} = \left(\left. \frac{\partial \ln \Omega_A}{\partial E} \right|_{E=E_A} - \left. \frac{\partial \ln \Omega_B}{\partial E} \right|_{E=E_B} \right) \Delta$$

We define a factor

$$\beta = \frac{\partial \ln \Omega}{\partial E}$$

which is porportional to inverse temperature $\frac{1}{T}$. The equation now is

$$\boxed{\text{diff} = (\beta_A - \beta_B)\Delta}$$

The physical interpretation of the system depends on the value of diff.

- If diff > 0, system A gains energy which indicates A is cooler. $T_A < T_B$ and $\beta_A > \beta_B$.

- If diff < 0, system A loses energy which indicates A is warmer. $T_A > T_B$ and $\beta_A < \beta_B$.

- If diff $= 0$, the system AB is in thermal equilibrium. $T_A = T_B$ and $\beta_A = \beta_B$.

EXAMPLE 16.

Each table below shows the number of configurations (denoted with Ω) that a system is able to reach at different energies (E). At each energy, all configurations available to a given system are equally likely. There are two systems, A and B. These two systems are in thermal contact with each other and share a total of 100 J of energy.

E_A	Ω_A
0 J	1
20 J	$10^{6.92\text{e}21}$
40 J	$10^{1.32\text{e}22}$
60 J	$10^{1.86\text{e}22}$
80 J	$10^{2.26\text{e}22}$
100 J	$10^{2.52\text{e}22}$

E_B	Ω_B
0 J	1
20 J	$10^{6.30\text{e}21}$
40 J	$10^{1.10\text{e}22}$
60 J	$10^{1.54\text{e}22}$
80 J	$10^{1.95\text{e}22}$
100 J	$10^{2.32\text{e}22}$

(a) How much energy is in system A at thermal equilibrium? Give your answer in J.

Thermal equilibrium happens when $\Omega_A \Omega_B$ is maximized. We will make a similar table from the cards example but since numbers are very large, we will use logarithms to turn multiplications into additions (natural base is usually used but base 10 is for convenience).

E_A	0 J	20 J	40 J	60 J	80 J	100 J
$\log_{10} \Omega_A$	0	6.92e21	1.32e22	1.86e22	2.26e22	2.52e22
E_B	100 J	80 J	60 J	40 J	20 J	0 J
$\log_{10} \Omega_B$	2.32e22	1.95e22	1.54e22	1.10e22	6.30e21	0
$\log_{10} \Omega_{AB}$	2.32e22	2.642e22	2.86e22	2.96e22	2.89e22	2.52e22

It seems like $\Omega_A \Omega_B$ is maximized when $E_A = 60\,\text{J}$ and $E_B = 40\,\text{J}$.
(b) Estimate the value of β at thermal equilibrium (in units of J^{-1}).

We can estimate β using $\frac{\ln \Omega_2 - \ln \Omega_1}{E_2 - E_1}$. We choose two points in A before and after equilibrium point which are $E_1 = 40\,\text{J}$ and $E_2 = 80\,\text{J}$. This gives

$$\beta = \frac{\ln \Omega_2 - \ln \Omega_1}{E_2 - E_1} = 5.411\text{e}20$$

2.3 Boltzmann Statistics

2.3.1 Boltzmann Factor

There exists a simple formula for finding the probability to find a system in a particular microstate. Last time, we have seen that the multiplicity of a system typically increases the more energy it gets (at least up to some maximum). We focused on systems with fixed energy (microcanonical ensemble) where the fundamental assumption of statistical mechanics holds but now we will study systems with fixed β. Imagine an atom in a solid with fixed temperature β that is in a particular quantum state (ie. having a certain amount of energy). The atom is the system of interest. We assume that each quantum state corresponds to a single microstate in this atom.
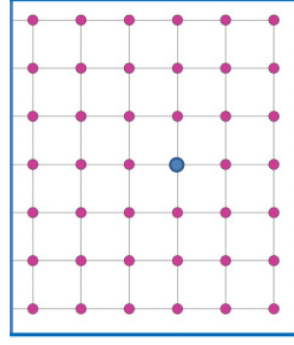
Figure 11: Atom in Solid

This atom is surrounded by the bath which is the rest of the solid. It is in thermal contact with the bath. The whole system (atom+bath) is isolated and this means the total energy of the whole system is constant no matter how much energy is passed back and forth between the atom and the bath. The energy that is available to the bath depends on how much energy the atom has in its quantum state. Since the atom can have many quantum states, the atom's energy is discrete while the bath's energy is continuous.

If the atom is isolated from the rest of the universe, it will have fixed energy. But if it interacts with lots of other atoms via bath by exchanging energy, it can change its quantum state (microstate) and some of its microstates will be more likely than others. This is because its energy is not constant since it is in thermal contact with the bath.

Suppose the total energy of the atom+bath system is E_0. Let E denote the energy of the bath and Δ denote the energy of the atom which will change depending on its quantum state. By conservation of energy,

$$E + \Delta = E_0$$

The total multiplicity for the atom+bath system is (sys is atom)

$$\Omega_{tot}(E, \Delta) = \Omega_{bath}(E_0 - \Delta)\Omega_{sys}(\Delta)$$

We have $\Omega_{sys}(\Delta) = 1$ (the number of ways to arrange a particular microstate itself is one).

$$\Omega_{tot}(E, \Delta) = \Omega_{bath}(E_0 - \Delta)$$

While it looks like the multiplicity of the bath is independent of the atom's state but not really. Depending on how much energy Δ the bath loses to the atom, Ω_{bath} will vary. To express the relationship between Ω_{tot} and Ω_{bath}, we apply the first order Taylor expansion on $\ln \Omega_{tot}$ about E_0,

$$
\begin{aligned}
\ln \Omega_{tot}(E, \Delta) &= \ln \Omega_{bath}(E) \\
&= \ln \Omega_{bath}(E_0) + \left. \frac{\partial \ln \Omega_{bath}}{\partial E} \right|_{E=E_0} (E - E_0) \\
&= \ln \Omega_{bath}(E_0) - \left. \frac{\partial \ln \Omega_{bath}}{\partial E} \right|_{E=E_0} \Delta \\
&= \ln \Omega_{bath}(E_0) - \beta \Delta
\end{aligned}
$$

Raising e to the power of both sides gives

$$\boxed{\Omega_{tot} = \Omega_{bath}(E_0)e^{-\beta\Delta}}$$

$e^{-\beta\Delta}$ is known as the Boltzmann factor.

EXAMPLE 17.

What is the relative likelihood of the atom having a quantum of energy ϵ, vs. having no energy at all? (ie. $p(\epsilon)/p(0)$)

$$\frac{p_\epsilon}{p_0} = \frac{\Omega_{tot}(E_0, \epsilon)}{\Omega_{tot}(E_0, 0)} = \frac{\Omega_{bath}(E_0)e^{-\beta\epsilon}}{\Omega_{bath}(E_0)e^{-\beta(0)}} = e^{-\beta\epsilon}$$

We got $e^{-\beta\epsilon} < 1$ which indicates that the atom will most likely have zero energy.

EXAMPLE 18.

The states of a system are labelled below, and their energies are shown (if you don't know how to read the graph, energies are macrostates, each has some number microstates. Each microstate is labelled with numbers).
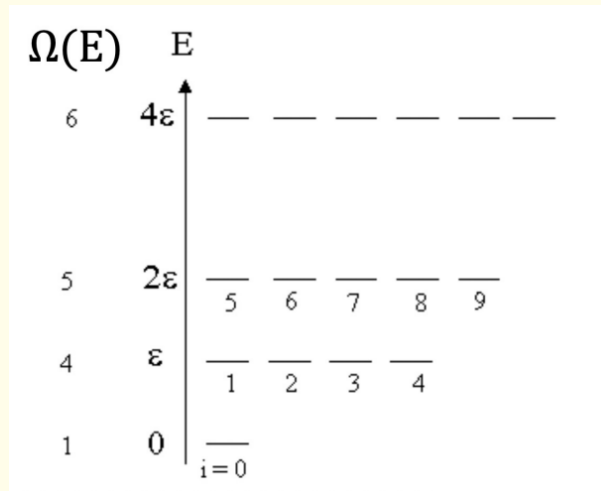
Find the relative likelihoods for the following pairs of microstates: (a) 2 and 4 (b) 2 and 7.

For microstates 2 and 4, the relative likelihood is

$$\frac{p_4}{p_2} = \frac{\Omega_{bath}(E_0)e^{-\beta E_4}}{\Omega_{bath}(E_0)e^{-\beta E_2}} = e^{-\beta(E_4 - E_2)} = e^{-\beta(\epsilon - \epsilon)} = 1$$

They are equally likely. Similarly for microstates 2 and 7,

$$\frac{p_7}{p_2} = e^{-\beta(E_7 - E_2)} = e^{-\beta(2\epsilon - \epsilon)} = e^{-\beta\epsilon} < 1$$

This means microstate 2 is more likely than microstate 7: $P_2 > P_7$.

EXAMPLE 19.

Continuing from the last example, decide whether it is true that $P(E = 0) = P(E = 2\epsilon)$.

We have

$$\frac{P(E = 0)}{P(E = 2\epsilon)} = \frac{e^{-\beta(0)}}{5e^{-\beta(2\epsilon)}} = \frac{1}{5}e^{2\beta\epsilon}$$

It looks like whether $\frac{1}{5}e^{2\beta\epsilon} > 1$ or $\frac{1}{5}e^{2\beta\epsilon} < 1$ depends on the value of β. The former is true for low temperatures ($\beta \to \infty$) and the system is more likely to be in the state of zero energy. For high temperatures ($\beta \to 0$), the ratio becomes $\frac{1}{5} < 1$ and the system is more likely to be in the state of 2ϵ energy. So the answer is undecidable.

2.3.2   Partition Function

The total probability of finding the atom in some state or other must add up to one:

$$\sum_{j}^{\Omega_{total}} p_j = 1$$

We can use this fact to find the expression for p_i, the probability of finding the atom in a particular microstate state i. We start with the relative probability expression we derived from the previous section:

$$\frac{p_j}{p_i} = e^{-\beta(E_j - E_i)}$$

$$\frac{1}{p_i} \underbrace{\left(\sum_j p_j \right)}_{=1} = e^{\beta E_i} \left(\sum_j e^{-\beta E_j} \right)$$

Rearranging the terms, we get

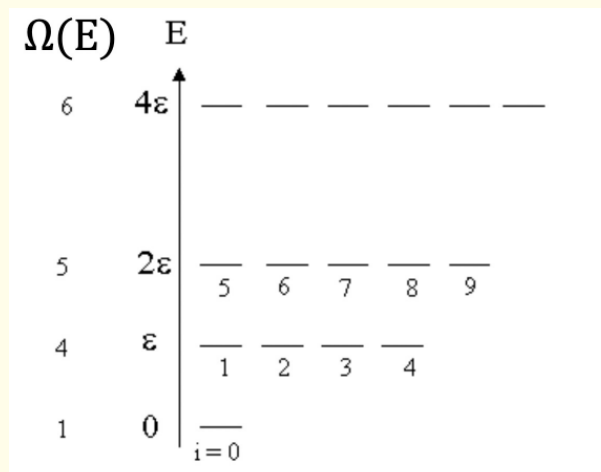$$\boxed{p_i = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}} = \frac{e^{-\beta E_i}}{Z(\beta)}}$$

The $Z(\beta)$ is called the partition function and it represents the sum of all state's Boltzmann factors. It depends on temperature only. From this formula, we see that the microstates with the same energy will still have the same probability.

EXAMPLE 20.

What is the partition function for this system?

We have

$$Z = \sum_j e^{-\beta E_j} = 1 + 4e^{-\beta \epsilon} + 5e^{-2\beta \epsilon} + 6e^{-4\beta \epsilon} = \sum_E \Omega(E)e^{-\beta E}$$

From the previous example, we can deduce that the probability of the system having an energy E (particular macrostate) is

$$P(E) = \frac{\Omega(E)e^{-\beta E}}{\sum_E \Omega(E)e^{-\beta E}} = \frac{\Omega(E)e^{-\beta E}}{Z(\beta)}$$

EXAMPLE 21.

A cup of water is in a large room at temperature. Is it more likely that the cup is in a particular state containing zero quanta of energy than it is in a particular state with the average number of quanta of energy?

The particular state here means a particular microstate.

The answer is yes because if we look at this formula we derived earlier $p_i = \frac{e^{-\beta E_i}}{Z(\beta)}$, p_i is largest if E_i is zero.
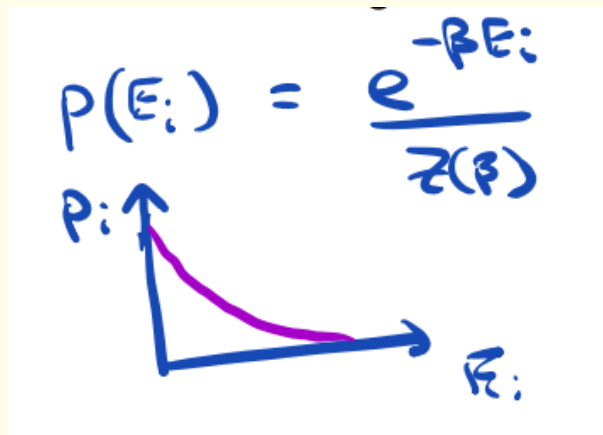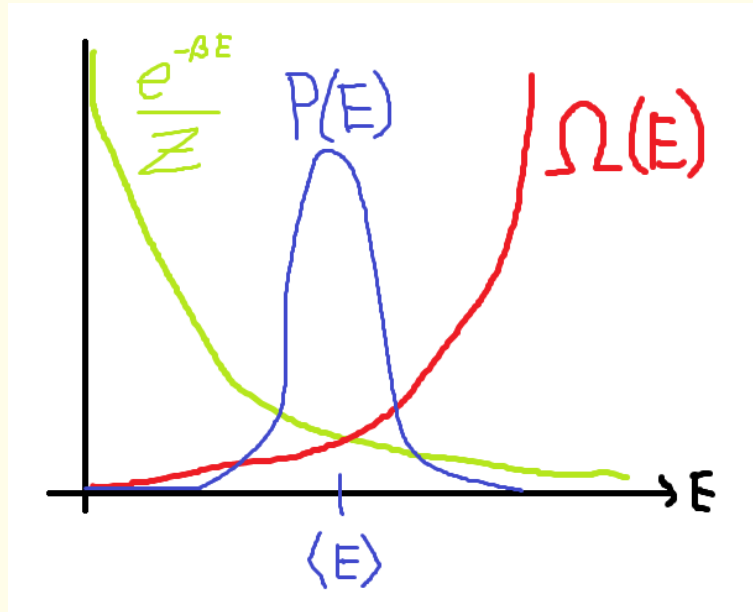

Figure 12: Graph of p_i

As a side note, even though the cup is more likely to be in the state of zero energy, its energy in thermal equilibrium is nonzero. Take a look at the following graph

Figure 13: Graph of p_i, Ω and $P(E)$

The green line represents $p_i = \frac{e^{-\beta E_i}}{Z(\beta)}$, the red line represents number of microstates with energy E: $\Omega(E)$ and the blue line represents the product of the green and red lines:

$$P(E) = \frac{\Omega(E)e^{-\beta E}}{Z(\beta)}$$

It is averaged at $\langle E \rangle$.

EXAMPLE 22.

A particle has 3 possible states: $0\,\text{eV}$, $2\,\text{eV}$ and $3\,\text{eV}$. If this particle is interacting with a large object with $\beta = 0.5\,\text{eV}^{-1}$, what is its average energy?

We have

$$Z = \sum_E \Omega(E)e^{-\beta E} = e^{-0.5(0)} + e^{-0.5(2)} + e^{-0.5(3)} = 1.59$$

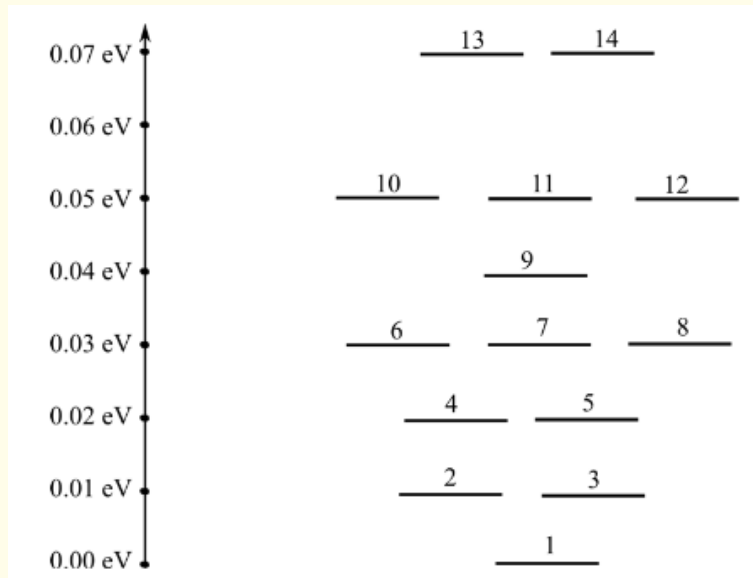The probabilities are $P(0) = \frac{1}{1.59}$, $P(2) = \frac{e^{-1}}{1.59}$, $P(3) = \frac{e^{-1.5}}{1.59}$.

The average energy is therefore

$$\langle E \rangle = \sum_i E_i P(E_i) = 0\,\text{eV}\frac{1}{1.59} + 2\,\text{eV}\frac{e^{-1}}{1.59} + 3\,\text{eV}\frac{e^{-1.5}}{1.59} = 0.88\,\text{eV}$$

EXAMPLE 23.

A system described by the energy diagram below is in thermal equilibrium with a heat bath at temperature T.

(a) What is the most likely energy when T is very large? What if T is very small? (b) Compare probabilities of $0.01\,\text{eV}$ and $0.03\,\text{eV}$ at high and low temperatures. (c) Finally, You have a collection of one million (10^6) atoms with the spectrum shown in the diagram above. The heat bath these atoms are interacting with has a very, very high temperature. How large would you expect the fluctuations in the total energy of the one million atoms to be (make an estimate)?

(a) The most likely energy when T is large is either $0.03\,\text{eV}$ or $0.05\,\text{eV}$ since they both have the same number of microstates. But let's compare:

$$\frac{P(0.05\,\text{eV})}{P(0.03\,\text{eV})} = \frac{3e^{-\beta(0.05)}}{3e^{-\beta(0.03)}} = e^{-\beta(0.02)} < 1$$

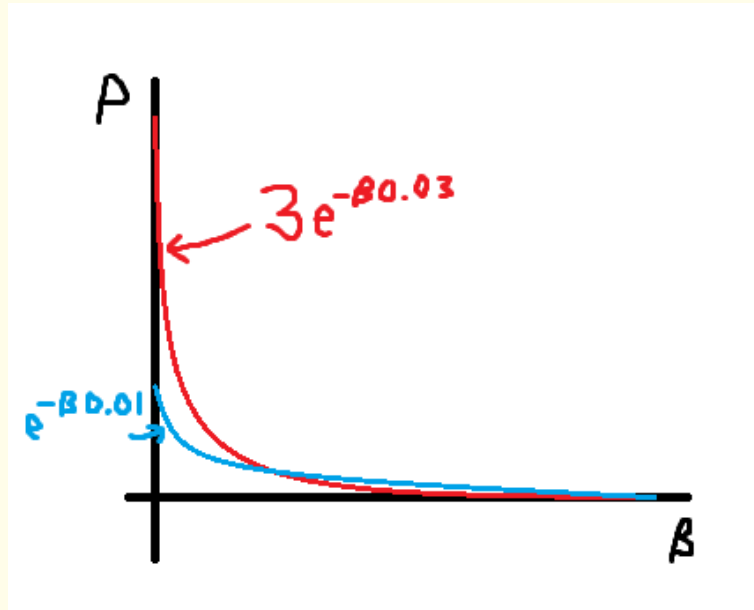This makes $0.03\,\text{eV}$ more likely energy when T is high.

When T is very small, β is very large which makes $e^{-\beta E} \approx 0$. In this case, $0.00\,\text{eV}$ is most likely since $P(0.05\,\text{eV}) \sim e^{-\beta 0.00} = 1$.

(b) We have

$$P(0.01\,\text{eV}) \sim e^{-\beta 0.01}$$
$$P(0.03\,\text{eV}) \sim 3e^{-\beta 0.03}$$

The graph of these probabilities looks like the following

We see that at low temperatures (high beta), $0.01\,\text{eV}$ is more likely than $0.03\,\text{eV}$.

(c) The expectation value (or average energy for a single atom) is

$$\langle E \rangle = \frac{1}{14}\left(0.00(1) + 0.01(2) + 0.02(2) + 0.03(3) + 0.04(1) + 0.05(3) + 0.07(2)\right) = 0.0343\,\text{eV}$$

For the system of N atoms, the total energy is $N\langle E \rangle$. The standard deviation is therefore

$$\sigma_N = \frac{N\langle E \rangle}{\sqrt{N}} = 34.3\,\text{eV}$$

2.3.3   Ideal Gas Kinetic Model

In this section, we will derive the alternative form of the ideal gas law step-by-step using the statistical tools we just learned.

An ideal gas is a theoretical gas model made of particles that are noninteracting (have no forces between them) but interact with each other through elastic collisions, and particles have no volume exclusion (i.e. no size) but have identical mass.
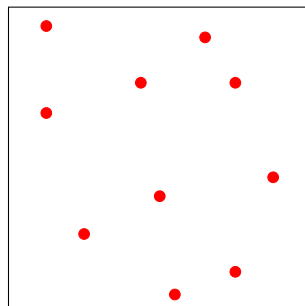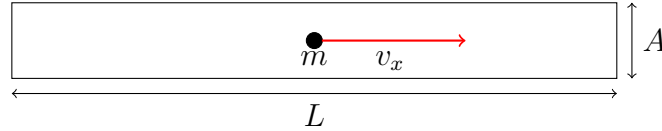

Figure 14: Cartoon of Ideal Gas

Its properties can be summarized by the following empirical formula

$$PV = nRT$$

where P is gas pressure, V is container's volume, $n = N/N_A$ is number of moles in gas, R is the ideal gas constant and T is gas temperature.

We start by expressing gas pressure in terms of particles' velocity. Consider the below one-dimensional setup

It shows a particle of mass m travelling with speed v_x inside the tube of length L to the right. It will hit the front wall of area A and bounce back to its starting point.

The total time for the whole journey is

$$\Delta t = \frac{2L}{v_x}$$

The pressure applied to the wall is given by

$$P = \frac{F}{A}$$

We need to calculate the average force exerted on the wall by the particle which is given by

$$F = -\frac{\Delta p}{\Delta t}$$

(The minus sign exists because of Newton's third law.) The particle's change of momentum after it bounces off the wall is $\Delta p = -mv_x - mv_x = -2mv_x$. The force is now

$$F = \frac{-(-2mv_x)}{\frac{2L}{v_x}} = \frac{mv_x^2}{L}$$

Substituting this result into our pressure formula results in

$$P = \frac{mv_x^2}{AL} = \frac{mv_x^2}{V}$$

where $V = AL$ denotes the volume of the tube. If there are more than one particle (say N particles), the pressure becomes

$$P = \sum_{i=1}^{N} \frac{mv_{x,i}^2}{V}$$

We can express the pressure in terms of average velocity squared v_x^2

$$P = \frac{Nm}{V}\left(\frac{1}{N}\sum_{i=1}^{N} v_{x,i}^2\right) = \frac{Nm}{V}\langle v_x^2 \rangle$$

$$PV = Nm\langle v_x^2 \rangle$$

Next, we will express $\langle v_x^2 \rangle$ in terms of temperature. The mean-square velocity depends only on temperature, not how many particles are present to some extent. The partition function for the particle is

$$Z = \sum_i e^{-\beta E_i} = \sum_i e^{-\beta \frac{1}{2} m v_i^2}$$

Remember that index i denotes possible microstates of a single particle, not each individual particle in gas. This function applies to a single particle moving in 3D space. Since each particle microstate can be expressed in terms of possible combinations of velocities (v_x, v_y, v_z), we rewrite the partition function as follows

$$Z = \frac{1}{\Delta v_x \Delta v_y \Delta v_z} \sum_{v_x} \sum_{v_y} \sum_{v_z} e^{-\beta \frac{1}{2} m (v_x^2 + v_y^2 + v_z^2)} \Delta v_x \Delta v_y \Delta v_z$$

We add $\Delta v_x \Delta v_y \Delta v_z$ to divide the whole momentum space into little 3D cubes, since velocities are continuous quantities. $\Delta v_x \Delta v_y \Delta v_z$ can be thought of as the volume of a small portion of the momentum space. As $\Delta v_x, \Delta v_y, \Delta v_z$ gets smaller, we get the following integral

$$
\begin{aligned}
Z &= \frac{1}{\Delta v_x \Delta v_y \Delta v_z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dv_x dv_y dv_z e^{-\beta \frac{1}{2} m (v_x^2 + v_y^2 + v_z^2)} \\
&= \frac{1}{\Delta v_x \Delta v_y \Delta v_z} \left[\int_{-\infty}^{\infty} dv_x e^{-\beta \frac{1}{2} m v_x^2} \right] \left[\int_{-\infty}^{\infty} dv_y e^{-\beta \frac{1}{2} m v_y^2} \right] \left[\int_{-\infty}^{\infty} dv_z e^{-\beta \frac{1}{2} m v_z^2} \right] \\
&= \frac{1}{\Delta v_x \Delta v_y \Delta v_z} Z_x Z_y Z_z \\
&= \frac{1}{\Delta v_x \Delta v_y \Delta v_z} (Z_x)^3
\end{aligned}
$$

where the last line follows from symmetry. The Z_x integral can be evaluated by rewriting it as a Gaussian integral

$$
\begin{aligned}
Z_x &= \int_{-\infty}^{\infty} dv_x e^{-\beta \frac{1}{2} m v_x^2} \\
&= \left(\frac{2}{\beta m} \right)^{\frac{1}{2}} \underbrace{\int_{-\infty}^{\infty} du \, e^{-u^2}}_{\sqrt{\pi}} = \left(\frac{2\pi}{\beta m} \right)^{\frac{1}{2}}
\end{aligned}
$$

Now, the average velocity squared $\langle v_x^2 \rangle$ can be written as

$$\langle v_x^2 \rangle = \sum_i v_{x,i}^2 p_i = \frac{1}{Z} \sum_i v_{x,i}^2 e^{-\beta E_i} = \frac{1}{Z} \sum_i v_{x,i}^2 e^{-\beta \frac{1}{2} m (v_x^2 + v_y^2 + v_z^2)}$$

Using the similar argument as before, the average velocity squared becomes

$$
\begin{aligned}
\langle v_x^2 \rangle &= \frac{\Delta v_x \Delta v_y \Delta v_z}{Z_x Z_y Z_z} \frac{1}{\Delta v_x \Delta v_y \Delta v_z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dv_x dv_y dv_z v_x^2 e^{-\beta \frac{1}{2} m (v_x^2 + v_y^2 + v_z^2)} \\
&= \frac{1}{Z_x Z_y Z_z} \left[\int_{-\infty}^{\infty} dv_x v_x^2 e^{-\beta \frac{1}{2} m v_x^2} \right] \underbrace{\left[\int_{-\infty}^{\infty} dv_y e^{-\beta \frac{1}{2} m v_y^2} \right]}_{Z_y} \underbrace{\left[\int_{-\infty}^{\infty} dv_z e^{-\beta \frac{1}{2} m v_z^2} \right]}_{Z_z} \\
&= \frac{1}{Z_x} \int_{-\infty}^{\infty} dv_x v_x^2 e^{-\beta \frac{1}{2} m v_x^2}
\end{aligned}
$$

27

We will use our result for Z_x are rewrite this integral as another Gaussian integral

$$\langle v_x^2 \rangle = \left(\frac{\beta m}{2\pi} \right)^{\frac{1}{2}} \int_{-\infty}^{\infty} dv_x v_x^2 e^{-\beta \frac{1}{2} m v_x^2}$$

$$= \left(\frac{\beta m}{2\pi} \right)^{\frac{1}{2}} \left(\frac{2}{\beta m} \right)^{\frac{3}{2}} \underbrace{\int_{-\infty}^{\infty} du \, u^2 e^{-u^2}}_{\frac{\sqrt{\pi}}{2}} = \frac{1}{\beta m}$$

Substituting our average velocity squared into our ideal gas law equation gives

$$PV = Nm \langle v_x^2 \rangle = \frac{Nm}{\beta m} = \frac{N}{\beta}$$

We are almost done! Next, we will try to relate it to the empirical gas law to derive the relationship between β and T

$$PV = nRT = \frac{N}{N_A} RT = \frac{N}{\beta}$$

$$\boxed{\beta = \frac{1}{k_B T}}$$

We defined the new constant $k_B = R/N_A$, known as Boltzmann constant. Substituting this result into our ideal gas law equation gives

$$\boxed{PV = N k_B T}$$

We derived the ideal gas law in an alternate form.

As a corollary, we can also calculate the root mean square speed and the total kinetic energy of an ideal gas with results from this section. We showed that the mean velocity squared is

$$\langle v_x^2 \rangle = \frac{1}{\beta m}$$

Using our result for β, we get

$$\langle v_x^2 \rangle = \frac{k_B T}{m}$$

Now for each gas particle, its *average* kinetic energy in the x-direction is

$$\langle E_x \rangle = \frac{1}{2} m \langle v_x^2 \rangle = \frac{1}{2} m \frac{k_B T}{m} = \frac{1}{2} k_B T$$

Since each gas particle moves in each of the three dimensions, the overall average kinetic energy is

$$\langle E \rangle = \langle E_x \rangle + \langle E_y \rangle + \langle E_z \rangle = 3 \langle E_x \rangle = \frac{3}{2} k_B T$$

by symmetry. Equating this result with kinetic energy will give the formula for calculating the root mean square speed:

$$\frac{1}{2} m \langle v^2 \rangle = \frac{3}{2} k_B T$$

$$v_{rms} = \sqrt{\langle v^2 \rangle} = \sqrt{\frac{3k_B T}{m}}$$

or

$$v_{rms} = \sqrt{\frac{3RT}{M}}$$

where $M = m/n$ is molar mass of a gas molecule.

Note that the root mean square speed is not always equal to the average speed $\langle v \rangle$. It will only give an *estimate* of the average speed of gas molecules.

Finally, for N gas particles, the total thermal energy of an ideal gas is:

$$\boxed{U_{tot} = \frac{3}{2} N k_B T}$$

This should be the average thermal energy but for large N, fluctuations from the average will be negligible.

We will see the more general result of this in the later section.

2.4 Ideal Gas Law

As discussed in the previous section, the properties of an ideal gas is summarized by the following formula

$$PV = nRT$$

or

$$PV = N k_B T$$

where P is gas pressure (in pascals), V is container's volume (in cubic meters), N is number of molecules, $n = N/N_A$ is number of moles in gas (Avogadro's number $N_A = 6.02\text{e}23$), $R = 8.314\,462\,\text{m}^3\,\text{Pa}\,\text{K}^{-1}\,\text{mol}^{-1}$ is the ideal gas constant, $k_B = R/N_A = 1.380\,649\text{e}{-}23\,\text{J}\,\text{K}^{-1}$ is Boltzmann constant, and T is gas temperature (in kelvins). Sometimes pressure is measured in atmospheres ($1\,\text{atm} = 101\,325\,\text{Pa}$) and bars ($1\,\text{bar} = 1\text{e}5\,\text{Pa}$) and volume is sometimes in liters ($1\,\text{L} = 0.001\,\text{m}^3$).

It is nothing but a combination of three simple gas laws. There is Boyle's Law, which tells us that for fixed temperature T and the number of gas molecules N, the pressure P is inversely proportional to the volume V:

$$P \propto \frac{1}{V}$$

And there is Charles' Law, which tells us that for fixed pressure P and the number of molecules N, the volume V is directly proportional to the temperature T:

$$V \propto T$$

Finally, Avogadro's Law, which tells us that for constant temperature T and constant pressure P, The volume of gas V is directly proportional to the number of molecules N:

$$V \propto N$$

Combining all these three laws will derive the ideal gas law shown at the beginning.

Also the ideal gas law is an equation of state; it depends only on the current state of the system, but not how the system got there.

There are numerous implications from the ideal gas law. For a given amount of gas at a given temperature, doubling the pressure squeezes the gas into exactly half as much space. Or, at a given volume, doubling the temperature causes the pressure to double.

EXAMPLE 24.

Consider a box of ideal gas. What happens to the gas if we (a) increase the box's volume (b) poke a hole in the box? (c) If we connect the box with another box of ideal gas with the same volume but with warmer gas, which box has more molecules?
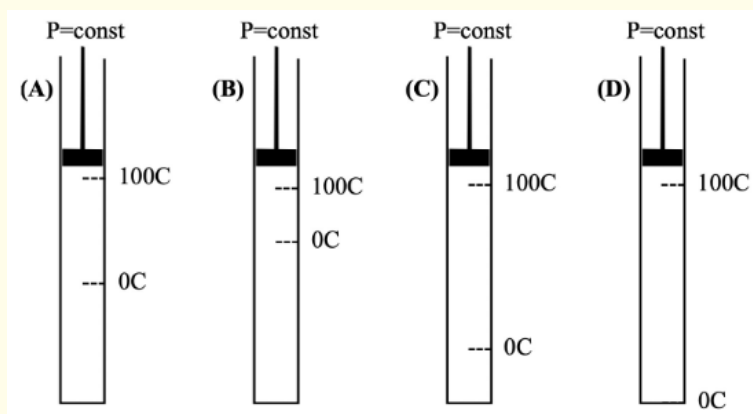
(a) Temperature T and number of molecules N are fixed (to change temperature, one must do something with molecules' average kinetic energy) but since we increased volume V, according to the ideal gas law $PV = Nk_BT$, we should expect pressure P to decrease.

(b) If we poke a hole in the box, it will leak which will decrease the number of particles N. The temperature T and volume V are the same, so only pressure P decreases.

(c) Since the boxes are connected, the pressure P of both boxes stays the same, as well as volume V. They both share the same PV value, in other words. Since the other box is warmer, it has a higher temperature T while our box has a lower temperature. According to the gas law $PV = Nk_BT$, we should expect our box to have more molecules N than the other box.

EXAMPLE 25.

A cylinder filled with gas and closed with a piston that can slide up and down while maintaining constant pressure inside can be used as a thermometer, by converting the volume into a temperature reading. To calibrate it, it was submerged in an ice bath and in boiling water; the position of the piston was noted. Assuming ideal gas behaviour between $0\,°C$ and $100\,°C$, which of the pictures are reasonable?

We are given that the pressure P and the number of molecules N are fixed. We will

use the ideal gas law $PV = Nk_BT$ to figure out the ratio of volumes from different temperatures:

$$\frac{V_{100\,°C}}{V_{0\,°C}}$$

First we convert temperatures to its SI unit kelvins by adding 273.15, this gives $0 + 273.15 = 273.15\,\text{K}$ and $100 + 273.15 = 373.15\,\text{K}$. The ratio becomes:

$$\frac{V_{100\,°C}}{V_{0\,°C}} = \frac{\frac{N\cancel{k_B}}{\cancel{P}}(373.15\,\text{K})}{\frac{N\cancel{k_B}}{\cancel{P}}(273.15\,\text{K})} = 1.3661$$

It is almost $\frac{4}{3}$, so the reasonable picture is (B).

EXAMPLE 26.

Nitrogen gas (N_2) is compressed to $10\,\text{atm}$ and stored in a $10\,\text{L}$ cylinder at room temperature ($21\,°C$). How much does the nitrogen weigh? The molar mass of a *single* nitrogen atom is $14\,\text{g}\,\text{mol}^{-1}$.

We are given that

$$P = 10\,\text{atm} = 1\text{e}6\,\text{Pa}$$

$$T = 273.15 + 21 = 294.15\,\text{K}$$

$$V = 10\,\text{L} = 10\text{e}{-}3\,\text{m}^3$$

$$M_{N_2} = 28\,\text{g}\,\text{mol}^{-1}$$

We use the ideal gas law $PV = nRT$ to calculate the amount of Nitrogen n:

$$n = \frac{PV}{RT} = 4.13\,\text{mol}$$

The mass of the gas is therefore

$$m = nM_{N_2} = 116\,\text{g}$$

2.5 Degrees of Freedom and Equipartition Theorem

2.5.1 Introduction

Ideal gases contain kinetic energy only but not all gases in the real world are ideal because they have other forms of energy. The total thermal energy of a real gas contains the following terms:

$$U_{tot} = U_k + U_{rot} + U_{vib} + U_g + U_{int} + U_{chem} + U_{rest}$$

where U_k is kinetic energy, U_{rot} is rotational kinetic energy, U_g is gravitational potential[1], U_{int} is interaction energy, U_{chem} is chemical energy, U_{rest} is rest energy. There could be more forms of energy we forgot to add.

---

[1]We don't typically include U_g because the effects of gravitational potential are very small compared to kinetic energy, so they can be neglected. This approximation works for gases on ground level but for gases in the atmosphere, gravity must be taken in account.

In the context of thermal physics, degrees of freedom refers to the number of ways a particle can store energy. For monoatomic particles in an ideal gas, there are three translational degrees of freedom, each specified by v_x, v_y and v_z. The other degrees of freedom include rotational motion, vibration, etc. For these types of energy, the formulas take the following forms:

$$\frac{1}{2}mv_x^2, \quad \frac{1}{2}mv_y^2, \quad \frac{1}{2}mv_z^2, \quad \frac{1}{2}I\omega^2, \quad \frac{1}{2}kx^2, \quad \ldots$$

We see that they all share the same format $\frac{1}{2}cx^2$ where c is some constant and x^2 is some (quadratic) degree of freedom. We will calculate the average energy at temperature T of any degree of freedom.

We start by calculating the partition function for a degree of freedom:

$$Z = \frac{1}{\Delta x} \int_{-\infty}^{\infty} dx e^{-\beta \frac{c}{2} x^2} = \frac{1}{\Delta x} \left(\frac{2}{\beta c} \right)^{\frac{1}{2}} \int_{-\infty}^{\infty} du\, e^{-u^2} = \frac{1}{\Delta x} \left(\frac{2\pi}{\beta c} \right)^{\frac{1}{2}}$$

Next, the average energy of a degree of freedom is:

$$\langle E \rangle = \frac{1}{2}c\langle x^2 \rangle = \frac{c}{2}\frac{1}{\Delta x} \int_{-\infty}^{\infty} dx\, x^2 \frac{e^{-\beta \frac{c}{2} x^2}}{Z} = \left(\frac{\beta c}{2\pi} \right)^{\frac{1}{2}} \left(\frac{2}{\beta c} \right)^{\frac{1}{2}} \frac{1}{\beta} \int_{-\infty}^{\infty} du\, u^2 e^{-u^2} = \frac{1}{2\beta} = \frac{1}{2}k_B T$$

Notice the energy is independent of c. This important result is known as the equipartition theorem. It states that at temperature T, the system's thermal energy is equally divided in all degrees of freedom in such way that the average energy of each degree of freedom is

$$\frac{1}{2}cx^2 = \frac{1}{2}k_B T$$

For a particle with f degrees of freedom, its average energy is

$$\langle E \rangle = \frac{f}{2}k_B T$$

For a system with N particles with f degrees of freedom, its total *thermal* energy is

$$U_{tot} = N\langle E \rangle = \frac{f}{2}N k_B T$$

As a side note, this equation allows us to express temperature in terms of average energy:

$$T = \frac{2}{f k_B}\langle E \rangle$$

We see that temperature increases as particles' average energy increases (by moving faster for instance).

2.5.2 Degrees of Freedom of Various Molecules

Besides monoatomic molecules (He, Ne, etc.), there are also diatomic (N_2, O_2, etc.) and triatomic molecules (H_2O, CO_2, etc.) which have more than three degrees of freedom. The below table summarizes the number of degrees of freedom in 2D for each molecule type at room temperature.

Type	# of Degrees of Freedom	Types of Degrees of Freedom
Monoatomic	2	2 translational
Diatomic	3	2 translational, 1 rotational
Triatomic	3	2 translational, 1 rotational

Here is the table for 3D

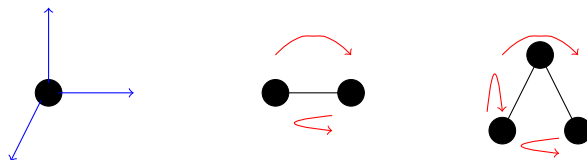| Type	# of Degrees of Freedom	Types of Degrees of Freedom
Monoatomic	3	3 translational
Diatomic	5	3 translational, 2 rotational
Triatomic	6	3 translational, 3 rotational

Figure 15: Visualization of the degrees of freedom of monoatomic, diatomic, and triatomic molecules in 3D. Monoatomic molecules only have three translational degrees of freedom (specified by velocities v_x, v_y, v_z), diatomic molecules have three translational plus two rotational degrees (it cannot rotate along the rotation axis parallel to its length due to quantum mechanical reasons), and the triatomic molecules have three translational and rotational degrees of freedom each.

Intramolecular bonds in diatomic and triatomic molecules can also vibrate like springs. The reason we didn't include vibrational degrees of freedom in the table above is that vibrational modes are, in some sense, frozen out at room temperature but once the temperature is high enough, vibrational modes do need to be taken into account. Vibrational modes are excited only if the room's temperature is high enough to overcome the energy gap between the ground and first excited state for that degree of freedom (ie. $\Delta E < \frac{1}{2}k_B T$). When vibration is taken into account, it will count as two degrees of freedom - one for spring kinetic energy and one for spring potential energy. There are also other modes of vibration besides stretching: twisting or flexing and they each count as two degrees of freedom.

The frozening of degrees of freedom can be explained using quantum mechanics. Degrees of freedom that are not in the classical regime can be described as "frozen out" (frozen at their quantum zero state) and don't contribute to the total energy. At low temperatures gas molecules only have translational degrees of freedom, then as they gain more energy they can also have rotational degrees of freedom, and then at really high temperatures they can also have vibrational degrees of freedom. But at those temperatures they have so much energy the molecules just break apart into plasma, so those vibrational degrees of freedom don't really exist classically for gases.
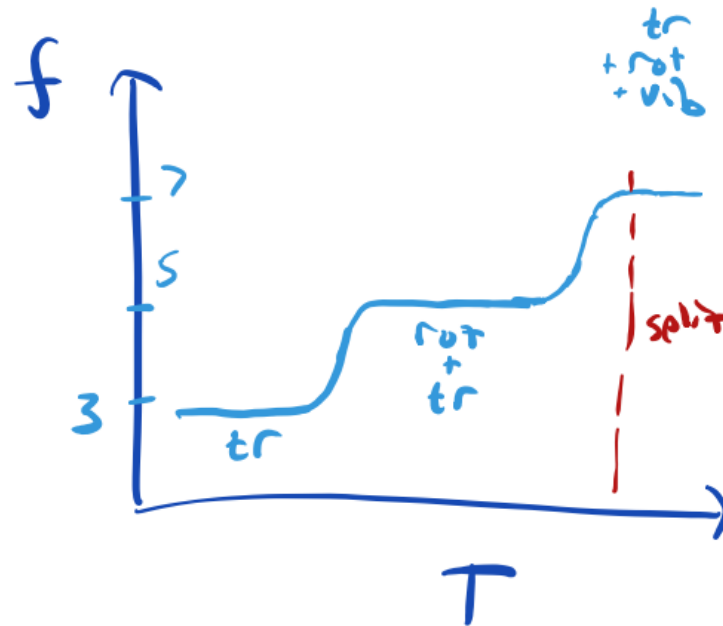
Figure 16: Graph showing the relationship between DoF and temperature. The red line is where molecules ionize to become plasma.

The frozening of degrees of freedom also explains why we don't typically take in account of rest-mass energies, or chemical energy stored in the molecular bonds. They are frozen at their quantum mechanical values unless the temperature is high enough for the gas to ionize.

EXAMPLE 27.

How many degrees of freedom does a diatomic molecule in gas exposed to high temperature have in two dimensions? Find the total energy for each diatomic molecule.

There are 2 translational degrees of freedom v_x, v_y for center of mass to move independently in x and y directions, 1 rotational degree of freedom to rotate on xy plane, and 2 vibrational degrees of freedom: spring kinetic energy term and spring potential energy term. So there are $f = 5$ degrees of freedom in total. The total energy for a diatomic molecule is:

$$E = \underbrace{\frac{1}{2}M(v_x^2 + v_y^2)}_{\text{translational}} + \underbrace{\frac{1}{2}I\omega^2}_{\text{rotational}} + \underbrace{\frac{1}{2}\mu\dot{x}^2 + \frac{1}{2}kx^2}_{\text{vibrational}}$$

where $M = m_1 + m_2$ is total mass, I is moment of inertia, ω is angular velocity, $\mu = \frac{m_1 m_2}{m_1 + m_2}$ is reduced mass, x is distance from spring equilibrium position, and k is spring constant.

Note in three dimensions, there is an extra translational and rotational degrees of freedom, so the total will be $f = 7$.

EXAMPLE 28.

How many degrees of freedom does a H_2O molecule have in high temperatures?

This is an example of triatomic molecule. There are 3 translational degrees of freedom for

its CoM, 3 rotational degrees of freedom, and 6 vibrational degrees of freedom. Stretching of each bond contributes 2 vibrational degrees of freedom. But the molecules can bend as well which contributes 2 extra vibrational degrees of freedom. So the total is $f = 6$.
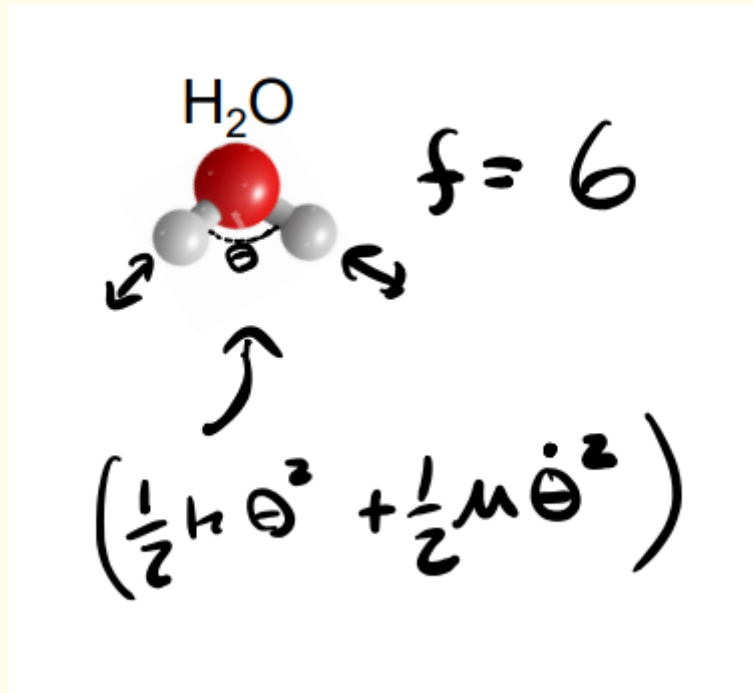

Figure 17: Vibrational DoFs in H_2O

EXAMPLE 29.

How many degrees of freedom does an atom in a solid in three dimensions have?

An atom in a solid can't move or rotate because they are trapped in lattice with other atoms however it can vibrate in three perpendicular directions, each contributing two vibrational degrees of freedom. Unlike liquids, vibrational modes of solids are not frozen out at room temperature so they do contribute to degrees of freedom. So there is a total of $f = 6$ degrees of freedom.
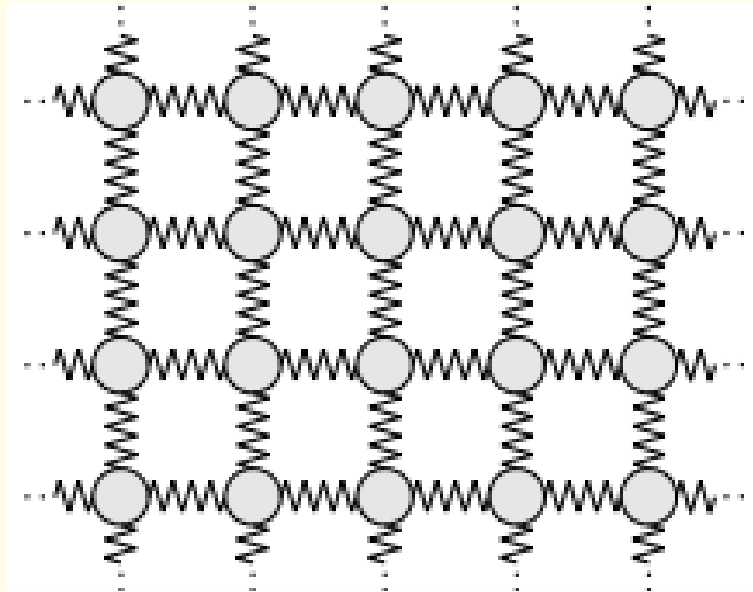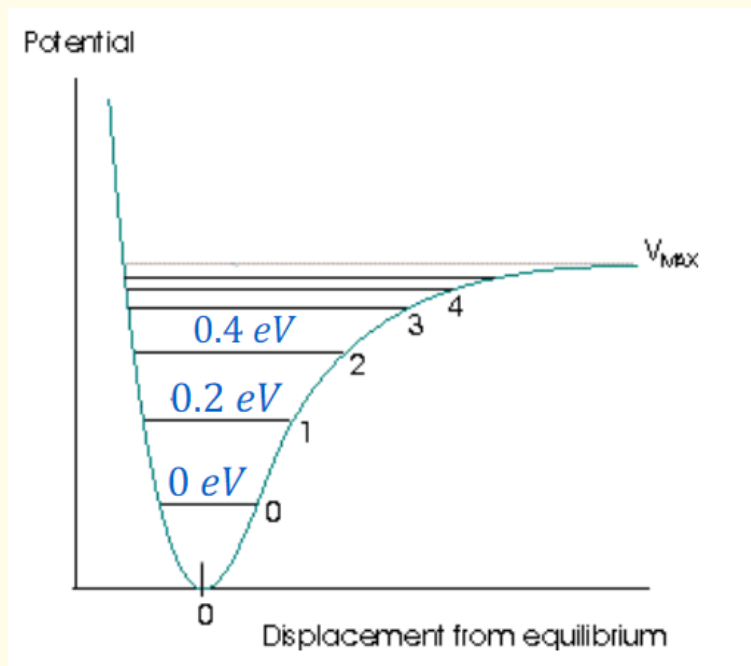
Figure 18: Ball-spring Model of Solid

EXAMPLE 30.

The vibrational energy levels of O_2 are measured spectroscopically as shown. At what temperatures would O_2 undertake vibrational excitations above the ground state?

We can approximate using $\Delta E \approx k_B T$ where ΔE is the energy difference between the ground and first excited state. So the minimum temperature is

$$T \approx \frac{\Delta E}{k_B} = \frac{0.2\,\text{eV}(1.6\text{e}{-}19\,\text{J eV}^{-1})}{1.38\text{e}{-}23\,\text{J K}^{-1}} \approx 2300\,\text{K}$$

Note on root mean square speed: this formula $v_{rms} = \sqrt{\frac{3k_B T}{m}}$ also applies to polyatomic molecules not just monoatomic gas molecules. This is because only translational DoFs contribute to gas kinetic energy.

2.6   Work, Heat and First Law of Thermodynamics

Work and heat are fundamentally two different things. Work is defined as a mechanical transfer of energy into or out of a system. Examples of doing work include squeezing a bottle of a gas, or pushing a piston, or shaking a hot water kettle. Heat refers to any non-mechanical transfer of energy which does no work. For example, when you boil water in a kettle, the water will heat up and gain energy but since there is no mechanical process involved, so no work is done. Remember that heat is a spontaneous flow of energy due to temperature. Other examples include how your laptop feels hot after you perform some CPU-intensive task.

Let ΔU be the change in energy of a system, Q_{in} be the heat that flows *in* the system, and W_{on} the work done *on* the system. The first law of thermodynamics simply states that

$$\Delta U = Q_{in} + W_{on}$$

or the total change in energy of a system is equal to the sum of its mechanical and non-mechanical transfers of energy[2]. Be careful with signs in this equation! Q_{in} is negative if heat leaves the system and W_{on} is negative if the system is being worked on.

It is worth noting here that while W_{on} and Q_{in} are both **functions of path**; how much work we do or a system does, or how much heat flows in/out of a system depends on how a process happens. However, the internal energy is a **function of state**, like the ideal gas law. The moral is that while how a system might gain or lose energy is a function of the path, the internal energy itself is a function of the state of the system only.
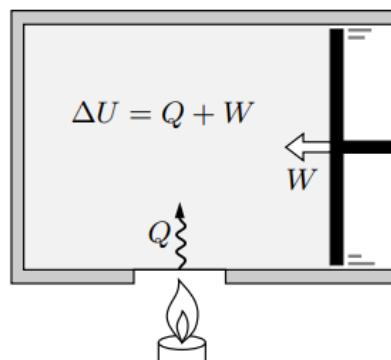

Figure 19: Illustration of the First Law

2.7   Compression Work on Gas

Work is generally defined as

$$W = \int_{x_1}^{x_2} F dx$$

However, in thermodynamics we don't have position or force as variables to work with. Instead, we have pressure and volume. Force due to pressure is $F = PA$ and we substitute it into the equation we get

---

[2]Some authors prefer to define W as the amount of work done *by* the system and the equation would instead be $\Delta U = Q - W$. But in this notes, we will go with W defined as the work done on the system.

$$W = \int_{x_1}^{x_2} PA\,dx$$

The product of area and distance is volume. So, since dx is a tiny displacement, which ultimately is a length, $A\,dx$ is a tiny change in volume! Note that we are assuming that volume changes are quasistatic (slow) so that pressure is equally distributed throughout the gas. Therefore:

$$W = \int_{V_1}^{V_2} P\,dV$$

However, this work was based on the definition of the amount of work done **by** a system on another system. We want the amount of work done **on** the system. Due to the conservation of energy, we can just throw a negative sign in front to get

$$W_{on} = -\int_{V_1}^{V_2} P\,dV$$

This is the proper definition of the work done on a gas. Note that since pressure P can change with volume so P is a function of volume. But if pressure is constant, the work becomes

$$W_{on} = -P\int_{V_1}^{V_2} dV = -P\Delta V$$

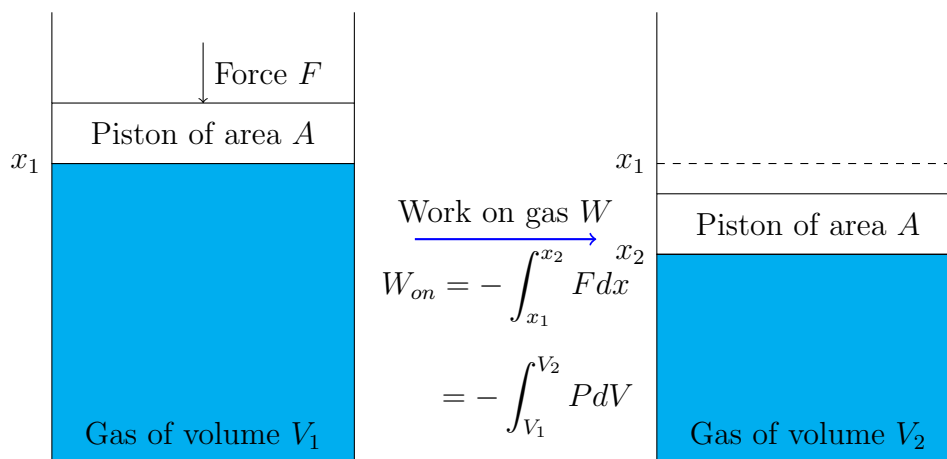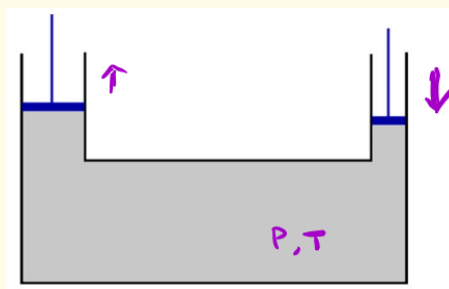It also applies when ΔV is infinitesimally small.

Figure 20: Illustration of Gas Being Compressed

EXAMPLE 31.

In the setup shown below, the gas inside is kept at constant pressure and temperature. The right piston slowly lowers and left piston slowly raises, while the pressure and temperature inside is kept constant. What is the net work done on the gas?

Since the volume does not change, the net work done on the gas is $W_{on} = -P\Delta V = 0$.

EXAMPLE 32.

Assuming the same setup from the previous example: The right piston slowly lowers and left piston slowly raises, while the pressure and temperature inside is kept constant. Ignoring friction (so that the system doesn't lose energy from heat), what do you say about external forces acting on either piston? What about work?

External forces[a] acting on either piston are pointing down all the time, both with the same magnitude: $F_L = F_R$. The total work applied to this system is still zero:

$$W = F_L \Delta x_L + F_R \Delta x_R = 0$$

In the case when $A_L = A_R$, we have $\Delta x_L = -\Delta x_R$ and $v_L = -v_R$[b] at some time.

---

[a]If it is vacuum in the surroundings, it will point up and the pistons will explode outwards but it is not the case here

[b]Note that force balance does not necessarily imply $v = 0$. Zero acceleration means constant velocity.

2.8 PV Diagrams

Pressure-Volume (PV) diagrams are the nice way to illustrate various thermodynamic processes. Shown below is a PV graph of a thermodynamic process that takes the gas from starting temperature T_1 (at pressure P_1 and volume V_1) to temperature T_2 (at pressure P_2 and volume V_2). Of course, there's nothing stopping us from assigning numbers to P, T, V, as well, if the situation calls for it. The process pictured below is an expansion of the gas, but we could also very well draw the arrow the other way and depict a compression of the gas.
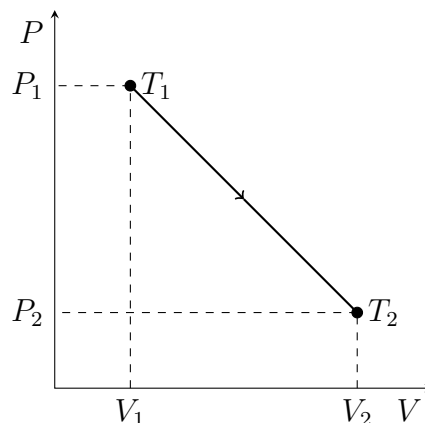
Figure 21: PV Diagram for Gas Expansion

In the previous section, we defined the work done on the gas as:

$$W_{on} = -\int_{V_1}^{V_2} P dV$$

And looking at the PV diagram, it now becomes clear that this is nothing more than the (negative) area under the PV curve! As an example, let's figure out the work done on the gas in the process shown above. All we have to do is calculate the area underneath the curve. For convenience, let us split it into two parts of the triangle and the rectangle. For the triangle, we

have area $\frac{1}{2}\left(P_1 - P_2\right)\left(V_2 - V_1\right)$, and for the rectangle, we have area $P_2\left(V_2 - V_1\right)$. Therefore, the total area under the graph is:

$$\text{Area } = \frac{1}{2}\left(P_1 - P_2\right)\left(V_2 - V_1\right) + P_2\left(V_2 - V_1\right) = \frac{1}{2}\left[\left(P_1 + P_2\right)\left(V_2 - V_1\right)\right]$$

This area is in fact the work done **by** the gas, and the work done on the gas is just its negative:

$$W_{on} = \frac{1}{2}\left[\left(P_1 + P_2\right)\left(V_1 - V_2\right)\right]$$

When using this area argument, do be careful of which direction the curve is going; for example, if the process was a compression of the gas rather than an expansion, we would have to introduce a negative sign as things would be going in the opposite direction.

EXAMPLE 33.

A thermodynamic system, close to equilibrium, follows a closed path on the PV diagram in the direction shown. Consider one full cycle, starting and ending at the blue dot. Is the total net heat added to the system over the course of this cycle Q_{in} less than, greater than, or equal to zero?
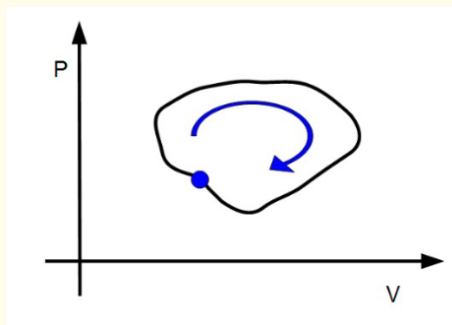
Since the system start and end at the same point, $U_i = U_f$ which means the change of energy $\Delta U = 0$. According to the first law of thermodynamics,

$$Q_{in} = -W_{on}$$
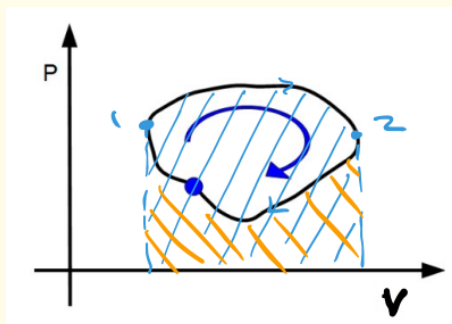
W_{on} is the sum of the work on the system to get from point 1 to point 2 and the work on the system to get from point 2 to point 1:

$$W_{on} = W_{1\to 2} + W_{2\to 1}$$
$$W_{on} = -A_{1\to 2} + A_{2\to 1}$$

Q_{in} is

$$Q_{in} = -W_{on} = A_{1\to2} - A_{2\to1}$$

Therefore Q_{in} is greater than zero.

EXAMPLE 34.

Imagine some helium in a cylinder with an initial volume of $1\,\text{L}$ and an initial pressure of $1\,\text{atm}$. Somehow the helium is made to expand to a final volume of $3\,\text{L}$ (like by increasing temperature during process), in such a way that its pressure rises in direct proportion to its volume. Calculate the amount of heat added to the helium.
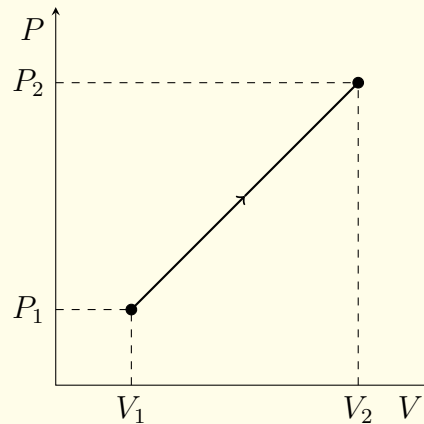
We have

$$V_1 = 0.001\,\text{m}^3$$
$$V_2 = 0.006\,\text{m}^3$$
$$P_1 = 101\,325\,\text{Pa}$$

The PV diagram for the process is shown below:

We are given that $P(V) = cV$ where c is some constant. The constant is calculated with the given initial condition:

$$c = \frac{P_1}{V_1} = 10\,132\,500\,\text{Pa}\,\text{m}^{-3}$$

P_2 is calculated as follows:

$$P_2 = cV_2 = 607\,950\,\text{Pa}$$

With all pieces in place, we calculate the work done on the gas:

$$W_{on} = -\int_{V_1}^{V_2} cV\,dV = -c\left(\frac{V_2^2}{2} - \frac{V_1^2}{2}\right) = -1773.1875\,\text{J}$$

The change in helium's internal energy is:

$$\Delta U = \frac{3}{2}\Delta(PV) = \frac{3}{2}(P_2V_2 - P_1V_1) = 5319.5625\,\text{J}$$

The heat added to the gas by the first law of thermodynamics is therefore:

$$Q = \Delta U - W_{on} = 7092.75\,\text{J}$$

2.9 Thermodynamic Processes

2.9.1 Isochoric Process

Isochoric processes are heating/cooling processes where the volume of the gas does not change ($\Delta V = 0$). For example, when you heat the gas sealed in a jar, the gas does not change its volume since the jar walls are rigid but its energy changes. V does not change but P and T changes.

On a PV diagram, isochoric processes look like a straight vertical line. As the gas gets heated up, the volume of the gas remains unchanged, so the process should be a graph of a line with constant V. Conversely, as the gas goes from a lower temperature T_1 to a higher temperature T_2, we would expect the pressure to increase.
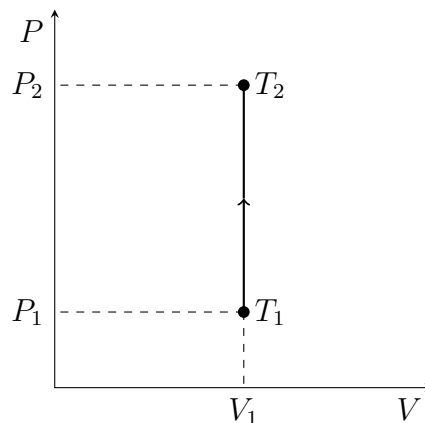

Figure 22: PV Diagram for Isochoric Process

Since a straight vertical line has no area, the work done on the gas during the process is

$$W = -\int_{V_1}^{V_2} P\,dV = 0$$

since there is no change in volume. According to the first law of thermodynamics,

$$\Delta U = Q + W = Q$$

The change in internal energy is just the heat that flows in/out of the system. We can express it in terms of temperature change:

$$\Delta U = \frac{f}{2} N k_B \Delta T = Q$$

We define the quantity C heat capacity:

$$C = \frac{Q}{\Delta T}$$

This is the heat needed to change the temperature of a system by one degree, or the ratio of the heat added to the system to the change in temperature.

For isochoric processes, we have heat capacity at the **fixed volume** for ideal gases:

$$C_v = \frac{f}{2} N k_B$$

We obtain heat as a function of the change in temperature and amount (and degrees of freedom!) of the gas:

$$Q = C_v \Delta T$$

2.9.2 Isobaric Process

Isobaric processes are compression/expansion processes in which the pressure of the gas does not change ($\Delta P = 0$). For example, consider a box of gas where the top face is covered by a massless piston that slides without friction. If the gas heats up slowly then the piston gets pushed up slowly[3] by the warming, expanding gas, while the pressure of the gas stays constant. P stays constant but V and T increases. Heating a balloon is another example of an isobaric process. It heats up while it warms up and expands. The pressure inside the balloon matches the surroundings.
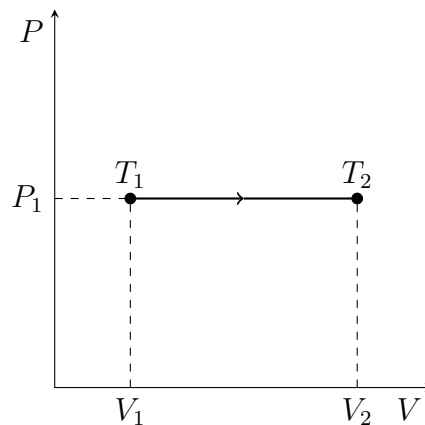
Figure 23: PV Diagram for Isobaric Process

The work done in this process is just

$$W = -P\Delta V$$

since the pressure is constant. Applying the ideal gas law, the change in temperature is

$$\Delta T = \frac{P}{Nk_B}\Delta V$$

and the change in internal energy is

$$\Delta U = \frac{f}{2}Nk_B\Delta T = \frac{f}{2}P\Delta V$$

The amount of heat added to the gas is calculated using the first law of thermodynamics:

$$Q = \Delta U - W = \frac{f+2}{2}P\Delta V$$

is related to temperature by

$$Q = \frac{f+2}{2}Nk_B\Delta T$$
$$Q = C_p\Delta T$$

We also found the formula for heat capacity at the **fixed pressure**:

---

[3]The piston should rise slowly (or quasistatically) so that the gas pressure matches the surroundings.

$$C_p = \frac{f+2}{2} N k_B$$

Note that $C_p > C_v$, or that the heat capacity at constant pressure is higher than the heat capacity at constant volume for gases. This tells us that when gas is allowed to expand as it is heated (retaining constant pressure), it requires more energy to increase its temperature than if it is kept at a fixed volume.

EXAMPLE 35.

Consider $1 \, \text{cm}^3$ of methane (CH_4) gas at $2 \, \text{atm}$ of pressure and at temperature equal to $25 \, °C$. A molecule of methane has tetrahedral shape. Find the amount of heat that has to be removed while compressing the gas to $0.4 \, \text{cm}^3$ to stop the pressure from rising.

This is an isobaric process. We have

$$\Delta V = -0.6 \, \text{cm}^3 = -6\text{e}{-7} \, \text{m}^3$$

$$P = 2 \, \text{atm} = 202 \, 650 \, \text{Pa}$$

$$f = 6$$

Applying the first law of thermodynamics,

$$Q = \Delta U - W = \frac{f}{2} N k_B \Delta T - W = \frac{f}{2} P \Delta V + P \Delta V = \frac{f+2}{2} P \Delta V = -0.49 \, \text{J}$$

2.9.3 Isothermal Process

Isothermal processes are compression/expansion processes in which the temperature of the gas remains constant ($\Delta T = 0$). These processes occur very slowly. As an example, consider a box of gas where the top face is a piston again. When the piston is pushed down very very slowly, the gas inside the piston remains at thermal equilibrium with its surroundings. Hence, as we do work on the gas, an equal amount of energy leaves the gas in the form of heat, and the gas inside the piston stays at the same temperature.
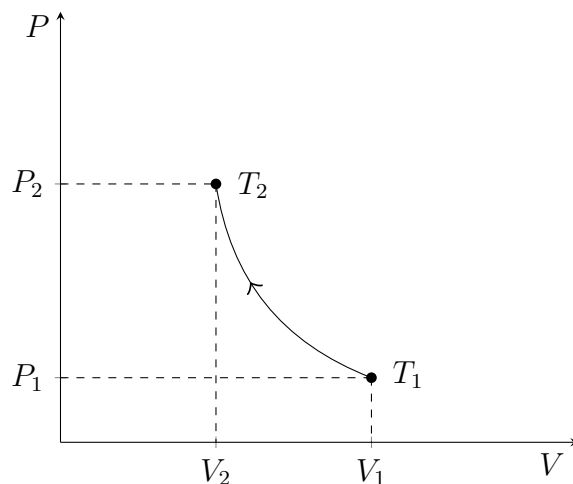

Figure 24: PV Diagram for Isothermal Compression

The relationship between pressure and volume according to the ideal gas law is

$$P = \frac{Nk_BT}{V}$$

The work done on the gas during this process is

$$W = -\int_{V_1}^{V_2} \frac{Nk_BT}{V}dV = -Nk_BT \int_{V_1}^{V_2} \frac{1}{V}dV = Nk_BT \ln\left(\frac{V_i}{V_f}\right)$$

Since there is no temperature change, the change in internal energy is zero[4]:

$$\Delta U = \frac{f}{2}Nk_B\Delta T = \frac{f}{2}(P\Delta V + V\Delta P) = 0$$

[5] From the first law of thermodynamics, it follows that

$$Q = -W = Nk_BT \ln\left(\frac{V_f}{V_i}\right)$$

EXAMPLE 36.

4 mol if ideal gas is compressed at $0\,°\text{C}$ so that the volume decreases by a factor of 3. How much work is done on the gas?

This is an isothermal process. The work done is

$$W = Nk_BT \ln\left(\frac{V_i}{V_f}\right) = Nk_BT \ln(3) = nRT \ln(3) = 9980\,\text{J}$$

2.9.4 Adiabatic Process

Adiabatic processes are compression/expansion processes in which the gas has no heat flow ($Q = 0$). These processes tend to be either fast, or occur under insulated conditions; for example, consider the same gas cylinder with the piston we discussed in the isothermal section. If instead of pushing in the piston very slowly, we push down forcefully, the compression will occur so quickly that there wouldn't be time for heat to flow out of the gas. This would be an adiabatic compression. As another example, if we had a cylinder of gas surrounded with thermal insulation, then even if we were to compress the piston or let the gas expand, there would be no heat flow and again we would have an adiabatic process.

Considering infinitesimal change in volume ΔV, the work done in this process is

$$W = -\int_{V_0}^{V_0+\Delta V} PdV = -P\Delta V = -\frac{Nk_BT}{V}\Delta V$$

The change in internal energy is just $\Delta U = W$ since $Q = 0$ which gives

$$\Delta U = \frac{f}{2}Nk_B\Delta T = -\frac{Nk_BT}{V}\Delta V$$

---

[4]For ideal gases, this is true but for real gases, this is not always true. Change in internal energy U contains many other factors like pressure.

[5]$\Delta(PV) = P\Delta V + V\Delta P$ arises from product rule. $P\Delta V + V\Delta P$ is equal to zero because $\frac{dV}{dP} = \frac{d}{dP}\left(\frac{Nk_BT}{P}\right) = -\frac{Nk_BT}{P^2}$ and $P\Delta V = P\left(-\frac{Nk_BT}{P^2}\Delta P\right) = -\frac{Nk_BT}{P}\Delta P = -V\Delta P$ which cancels out the $V\Delta P$ term

$$\frac{f}{2}\Delta T = -\frac{T}{V}\Delta V$$

For infinitesimal ΔV and ΔT, we get the following differential equation:

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

Integrating both sides from initial to final state gives

$$\frac{f}{2}\int_{T_1}^{T_2}\frac{dT}{T} = -\int_{V_1}^{V_2}\frac{dV}{V}$$

$$\frac{f}{2}\ln\left(\frac{T_2}{T_1}\right) = -\ln\left(\frac{V_2}{V_1}\right)$$

$$\ln\left(\left(\frac{T_2}{T_1}\right)^{\frac{f}{2}}\right) = \ln\left(\frac{V_1}{V_2}\right)$$

$$\left(\frac{T_2}{T_1}\right)^{\frac{f}{2}} = \frac{V_1}{V_2}$$

We conclude that

$$V_1 T_1^{\frac{f}{2}} = V_2 T_2^{\frac{f}{2}}$$

or

$$VT^{\frac{f}{2}} = \text{const}$$
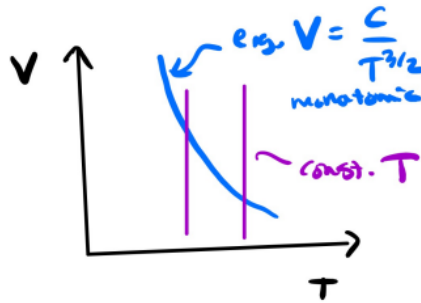
Figure 25: VT graph for adiabatic transform

There is the alternative form for this equation. We apply the ideal gas law $PV = Nk_BT$ to eliminate temperatures T_1 and T_2 which gives

$$V_1\left(\frac{P_1 V_1}{Nk_B}\right)^{\frac{f}{2}} = V_2\left(\frac{P_2 V_2}{Nk_B}\right)^{\frac{f}{2}}$$

$$P_1^{\frac{f}{2}}V_1^{1+\frac{f}{2}} = P_2^{\frac{f}{2}}V_2^{1+\frac{f}{2}}$$

Raising both sides to the power of $\frac{2}{f}$ gives

$$P_1 V_1^{\frac{2}{f}+1} = P_2 V_2^{\frac{2}{f}+1}$$

Denote constant $\gamma = \frac{2}{f} + 1 = \frac{f+2}{f}$ to be adiabatic index[6], we conclude that

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

$$V^{\gamma-1}T = \text{const}$$

So, an adiabatic curve on a PV diagram obeys the following equation

$$P = \frac{\text{const}}{V^\gamma}$$

Since $\gamma > 1$, an adiabatic curve looks quite a bit like the isothermal curve, but is steeper/lies above it.
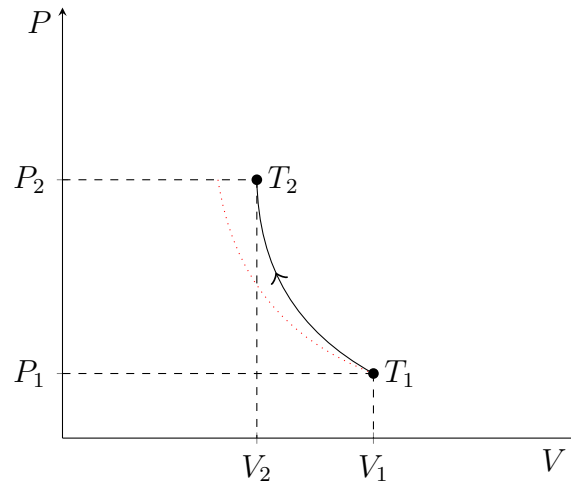

Figure 26: PV diagram for adiabatic compression. The adiabatic curve is pictured in black, and the isothermal curve is dashed and red.

EXAMPLE 37.

In a Diesel engine, atmospheric air (which is mostly nitrogen gas) at temperature $22\,°\text{C}$ is quickly compressed to about $1/25$ of its intial volume. Estimate the temperature of the air after compression and explain (to yourself) why a Diesel engine does not require spark plugs.

We are given that

$$T_1 = 295.15\,\text{K}$$
$$\frac{V_2}{V_1} = \frac{1}{25}$$
$$f = 5$$

In any adiabatic process, $V^{\gamma-1}T = \text{const}$, so

$$V_1^{\gamma-1}T_1 = V_2^{\gamma-1}T_2$$
$$\frac{T_1}{T_2} = \left(\frac{V_2}{V_1}\right)^{\gamma-1}$$

---

[6]It turns out that in general $\gamma = \frac{C_p}{C_v}$. The adiabatic index is just the ratio of heat capacity in constant pressure and heat capacity in constant volume. We see that it is true for ideal gases: $\gamma = \frac{\frac{f+2}{2}Nk_B}{\frac{f}{2}Nk_B} = \frac{f+2}{f}$

$$T_2 = \frac{T_1}{\left(\frac{V_2}{V_1}\right)^{\gamma-1}} = \frac{295.15\,\mathrm{K}}{\left(\frac{1}{25}\right)^{\frac{2}{5}}} = 796.44\,°\mathrm{C}$$

This temperature is very high!

EXAMPLE 38.

Compute work done on a gas that starts with volume V_1 and pressure P_1, and then has its volume adiabaticaly halved.

We have

$$PV^\gamma = P_1 V_1^\gamma$$

The work done in this process is

$$W = -\int_{V_1}^{\frac{V_1}{2}} P dV$$

$$= -\int_{V_1}^{\frac{V_1}{2}} \frac{P_1 V_1^\gamma}{V^\gamma} dV$$

$$= -P_1 V_1^\gamma \left[\frac{V^{1-\gamma}}{1-\gamma}\right]_{V_1}^{\frac{V_1}{2}}$$

$$= \frac{P_1 V_1^\gamma}{1-\gamma} \left[V^{1-\gamma}\right]_{\frac{V_1}{2}}^{V_1}$$

$$= \frac{P_1 V_1^{1+\frac{2}{f}}}{-\frac{2}{f}} \left[\frac{1}{V^{\frac{2}{f}}}\right]_{\frac{V_1}{2}}^{V_1}$$

$$= -\frac{f}{2} P_1 V_1^{1+\frac{2}{f}} \left[\frac{1}{V_1^{\frac{2}{f}}} - \frac{2^{\frac{2}{f}}}{V_1^{\frac{2}{f}}}\right]$$

$$= \frac{f}{2} \frac{P_1 V_1^{1+\frac{2}{f}}}{V_1^{\frac{2}{f}}} \left(2^{\frac{2}{f}} - 1\right)$$

$$= \frac{f}{2} P_1 V_1 \left(2^{\frac{2}{f}} - 1\right)$$

$$= \frac{f}{2} N k_B T_1 \left(2^{\frac{2}{f}} - 1\right)$$

In general, work done on gas in adiabatic process is

$$W = -\int_{V_1}^{V_2} \frac{P_1 V_1^\gamma}{V^\gamma} dV = P_1 V_1^\gamma \frac{V_1^{1-\gamma} - V_2^{1-\gamma}}{1-\gamma} = \frac{P_1 V_1 - P_2 V_2}{1-\gamma}$$

The last term follows because $P_1 V_1^\gamma = P_2 V_2^\gamma$.

2.10   Heat Engines

2.10.1   Introduction

A heat engines is a device that receives heat energy as input and converts it into some form of useful work. We will study engines that deals with gas in this section. The outline of the process goes as follows:

1. The heat engine starts in its original state.

2. The heat engine accepts some heat Q_H from an external hot bath of temperature T_H.

3. The heat engine does work W from the energy given to it in the form of heat.

4. The heat engine throws away some "waste heat" Q_C into a cold bath of temperature T_C.

5. The heat engine returns to its original state, and the cycle begins again.

This is demonstrated in a diagram below. Note that this is just a general outline, but really steps 2-4 can occur in a different order, multiple times, and sometimes concurrently. A diagram illustrates this below:
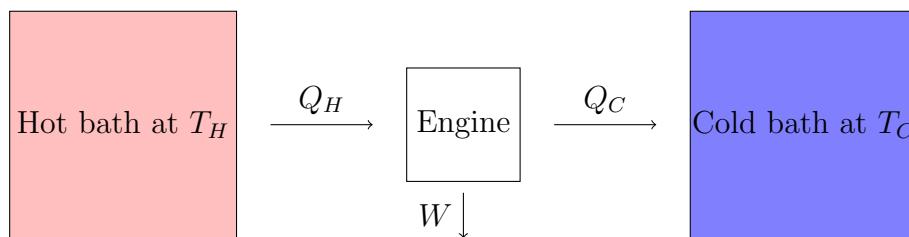

Figure 27: Heat Engine Cartoon

Of course, we could also flip the direction in which we do things to give the opposite result, which results in a heat pump (or refrigerators). Instead of putting heat at a high temperature T_H to get work W out, we can instead supply the heat engine with heat Q_C at a low temperature T_C (which is very cheap) and then supply work to the heat engine to get heat Q_H out at a high temperature.

To measure how good a heat engine is, we define efficiency:

$$e = \frac{W}{Q_H}$$

It can range from 0 (where $W = 0$ and we have a broken heat engine) to 1 (where $W = Q_H$ and we have a perfect conversion of the energy we put in to the work we get out). However, we will soon see that a 100% efficient heat engine is not realistic at all; this would require that no waste heat Q_C is produced at all, which is realistically impossible.

The first law of thermodynamics, which is the statement of conservation of energy, for heat engines can be stated as follows:

$$Q_H = W + Q_C$$

No energy gets lost or created spontaneously during the whole process. We can rewrite the heat engine efficiency as:

$$e = \frac{Q_H - Q_C}{Q_H} = 1 - \frac{Q_C}{Q_H}$$

If a heat engine is very efficient ($e = 1$), then there is no heat wasted (ie. $Q_C = 0$). But in the real world, e is always less than 1.

We will study a simple heat engine, shown in the below PV diagram. The gaseous medium used in this engine is $1\,\mathrm{mol}$ of ideal gas with f quadratic degrees of freedom.
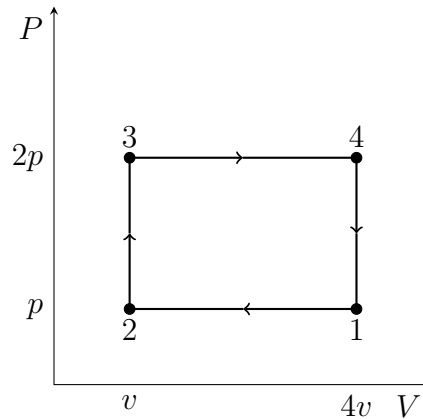

Figure 28: Simple Heat Engine

In this engine, the gas in the engine starts at state 1, at pressure p and volume $4v$. We isobarically compress the gas (by doing work on it), the gas gives off more heat in the process, until the gas reaches state 2 at volume v. Then, we heat the gas isochorically, giving it heat, until it reaches state 3 at pressure $2p$. Then, we add more heat and let it isobarically expand until it reaches state 4 at volume $4v$. Then, we let the gas isochorically cool and give off heat until it returns to its initial state at pressure p.

We start by calculating temperatures at each state. Since we know there is $1\,\mathrm{mol}$ of ideal gas, we can apply the ideal gas law: $PV = RT$ to find temperatures. We have:

$$T_1 = \frac{4pv}{R}$$

$$T_2 = \frac{pv}{R}$$

$$T_3 = \frac{2pv}{R}$$

$$T_4 = \frac{8pv}{R}$$

We see that the temperature is minimized at state 2 but is maximized at state 4. The ratio of maximum and minimum temperatures is:

$$\frac{T_4}{T_2} = 8$$

Next, we calculate the amount of work produced by the engine to us. Note that here we want to calculate work done **by** the system (engine) **to** us, unlike the previous sections where we calculate the work done **on** the system. For example, when the engine moves from states 2 to 4, it does positive work to us, represented by the red area below.

50
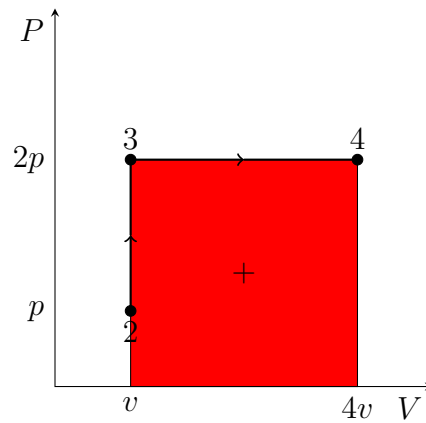
The net work done by the engine is the area of the rectangular region enclosed by the cycle, as illustrated below:
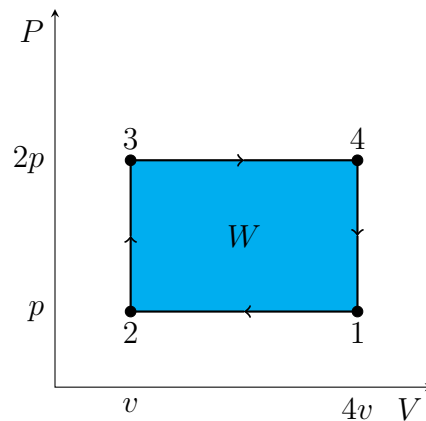
So the net work done is therefore

$$W = (4v - v)(2p - p) = 3pv$$

Applying the first law of thermodynamics:

$$\Delta U = Q - W$$

[7] Since $\Delta U = 0$ (we will show why later in this section),

$$Q = W$$

This means all of the heat flowing into the system is converted into work done by the engine; The total heat that flows in is $Q = 3pv$. Now, we will explicitly calculate heat produced/released at each state transition. Recall that heat produced at fixed volume is

$$Q = \frac{f}{2} N k_B \Delta T$$

and for fixed pressure is

$$Q = \frac{f+2}{2} N k_B \Delta T$$

Applying these formulas gives

---
[7]$-W$ since we redefined work to be the one done **by** the engine.

$$Q_{1\to2} = \frac{f+2}{2}Nk_B(T_2 - T_1) = \frac{f+2}{2}Nk_B\left(\frac{pv}{R} - \frac{4pv}{R}\right) = -\frac{f+2}{2}3pv$$

$$Q_{2\to3} = \frac{f}{2}Nk_B(T_3 - T_2) = \frac{f}{2}Nk_B\left(\frac{2pv}{R} - \frac{pv}{R}\right) = \frac{f}{2}pv$$

$$Q_{3\to4} = \frac{f+2}{2}Nk_B(T_4 - T_3) = \frac{f+2}{2}Nk_B\left(\frac{8pv}{R} - \frac{2pv}{R}\right) = \frac{f+2}{2}6pv$$

$$Q_{4\to1} = \frac{f}{2}Nk_B(T_1 - T_4) = \frac{f}{2}Nk_B\left(\frac{4pv}{R} - \frac{8pv}{R}\right) = -\frac{f}{2}4pv$$

Note that $\frac{Nk_B}{R} = \frac{N}{N_A} = n = 1$. One can show that

$$Q = Q_{1\to2} + Q_{2\to3} + Q_{3\to4} + Q_{4\to1} = W$$

The overall effciency of the engine is the net work divided by the amount of heat absorbed:

$$e = \frac{W}{Q_H} = \frac{W}{Q_{2\to3} + Q_{3\to4}} = \frac{3pv}{\frac{f}{2}pv + \frac{f+2}{2}6pv} = \frac{6}{7f + 12}$$

The below table summarizes the energy flow in each state transistion.

Segment	Heat	Work **by** Engine	Temperature
$1 \to 2$	release	negative	decrease
$2 \to 3$	absorb	none	increase
$3 \to 4$	absorb	positive	increase
$4 \to 1$	release	none	decrease

Lastly, we show that the net change of internal energy of the whole cycle is zero. Recall that

$$U = \frac{f}{2}Nk_BT = \frac{f}{2}PV$$

The net change of the internal energy is

$$\Delta U_{net} = \Delta U_{1\to2} + \Delta U_{2\to3} + \Delta U_{3\to4} + \Delta U_{4\to1}$$
$$= \frac{f}{2}p(v - 4v) + \frac{f}{2}(2p - p)v + \frac{f}{2}2p(4v - v) + \frac{f}{2}(p - 2p)4v$$
$$= \frac{f}{2}(pv - 4pv + 2pv - pv + 8pv - 2pv + 4pv - 8pv)$$
$$= 0$$

as expected.

2.10.2 Carnot Cycle

The Carnot engine is more than just an engine. It is a limit on the efficiency of real-world thermodynamic processes, and was used to develop the first notions of entropy.
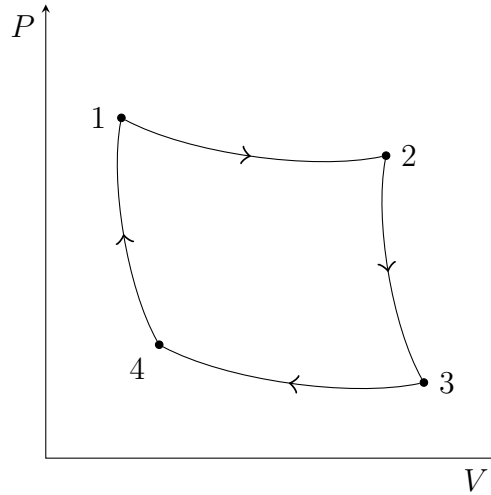
Figure 29: Carnot Cycle

If we start at state 1 in the diagram, the engine first undergoes an isothermal expansion to state 2, while in contact with the hot heat bath. It then undergoes an adiabatic expansion to state 3 in contact with no heat bath. This is followed by an isothermal compression to state 4, which occurs while in contact with a cool heat bath. Finally, the engine undergoes an adiabatic compression to return to state 1. Note that since the processes $1 \rightarrow 2$ and $3 \rightarrow 4$ are isothermal, so the engine has only two temperatures. States 1 and 2 are at a high temperature, and states 3 and 4 are at a low temperature:

$$T_1 = T_2 = T_H$$

$$T_3 = T_4 = T_C$$

We will calculate the efficiency for the Carnot engine, assuming it is filled with ideal gas medium. We start by calculating work done by the engine on each segment.

Work done **by** the engine in segment $1 \rightarrow 2$, when the gas is undergoes isothermal expansion, is:

$$W_{1\rightarrow 2} = Nk_B T_H \ln\left(\frac{V_2}{V_1}\right)$$

Since $V_2 > V_1$, the work is positive. According to the first law of thermodynamics, the heat absorbed is

$$Q_{1\rightarrow 2} = W_{1\rightarrow 2}$$

Now, the work done by the engine in segment $2 \rightarrow 3$, when the gas undergoes adiabatic expansion, is

$$W_{2\rightarrow 3} = \frac{P_3 V_3 - P_2 V_2}{1 - \gamma} = \frac{Nk_B}{1 - \gamma}(T_3 - T_2) = \frac{Nk_B}{1 - \gamma}(T_C - T_H)$$

which is positive. Similarly, for segment $3 \rightarrow 4$, the work done by the engine is

$$W_{3\rightarrow 4} = Nk_B T_C \ln\left(\frac{V_4}{V_3}\right)$$

which is negative. Finally, for segment $4 \rightarrow 1$, the work done by the engine is

$$W_{4\to1} = \frac{P_1 V_1 - P_4 V_4}{1 - \gamma} = \frac{N k_B}{1 - \gamma}(T_H - T_C)$$

which is negative. We see that $W_{2\to3} = -W_{4\to1}$. The efficiency is

$$e = \frac{W_{1\to2} + W_{2\to3} + W_{3\to4} + W_{4\to1}}{Q_{1\to2}}$$

Since $Q_{1\to2} = W_{1\to2}$ and $W_{2\to3} = -W_{4\to1}$,

$$= 1 + \frac{W_{3\to4}}{Q_{1\to2}}$$

$$= 1 + \frac{N k_B T_C \ln\left(\frac{V_4}{V_3}\right)}{N k_B T_H \ln\left(\frac{V_2}{V_1}\right)}$$

$$= 1 + \frac{T_C \ln\left(\frac{V_4}{V_3}\right)}{T_H \ln\left(\frac{V_2}{V_1}\right)}$$

Recall for adiabatic processes, $V^{\gamma-1}T = \text{const}$, so we have

$$V_2^{\gamma-1} T_H = V_3^{\gamma-1} T_C$$

$$V_4^{\gamma-1} T_C = V_1^{\gamma-1} T_H$$

which gives

$$\frac{T_H}{T_C} = \left(\frac{V_3}{V_2}\right)^{\gamma-1} = \left(\frac{V_4}{V_1}\right)^{\gamma-1}$$

It indicates that $\frac{V_3}{V_4} = \frac{V_2}{V_1}$. We are now allowed to eliminate ln terms in our efficiency which simplifies to:

$$\boxed{e = 1 - \frac{T_C}{T_H}}$$

The Carnot efficiency is the maximum possible efficiency attainable by a heat engine. This means a heat engine can only be 100% efficient if it's a Carnot engine, and the lowest temperature is $T_C = 0$, or the maximum temperature is $T_H = \infty$. Even though Carnot engines are efficient, they are very unrealistic to have practicial applications. The heat flows very slowly during isothermal processes, so it will take long time to get useful work out of it.

The below table summarizes the energy flow in each state transistion in the Carnot engine.

Segment	Heat	Work **by** Engine	Temperature
$1 \to 2$	absorb	positive	no change
$2 \to 3$	none	positive	decrease
$3 \to 4$	release	negative	no change
$4 \to 1$	none	negative	increase

2.10.3 Otto Cycle

We will study an internal combustion engine found in most automobiles: Otto engine.
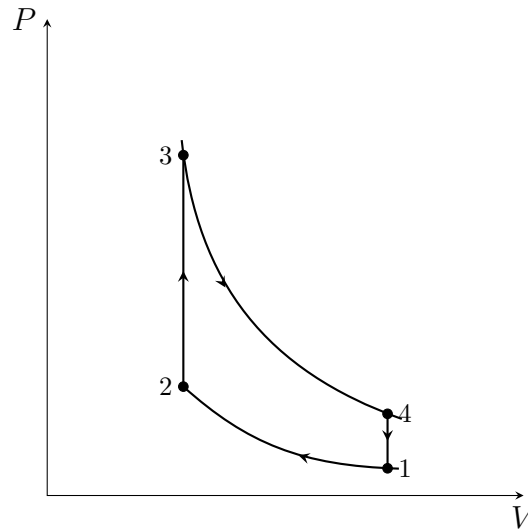

Figure 30: Ideal Otto Cycle

Gas is first injected into a cylindrical piston at state 1. It is first adiabatically compressed by the piston, transistioning from state 1 to state 2. It is then ignited by a spark which raises its pressure and temperature while its volume does not change, represented by the isochoric process $2 \to 3$. The high pressure gas then expands adiabatically, pushing the piston outwards, and produces mechanical work $(3 \to 4)$. Finally, the hot gas gets released and is replaced by the new gas mixture $(4 \to 1)$. The cycle continues. There is no heat bath in this engine. All the thermal energy is produced internally by burning the gas.

Let's find the efficiency of the Otto cycle, assuming that the gas mixture is ideal. We start with the definition of efficiency:

$$e = 1 - \frac{Q_C}{Q_H}$$

From the PV diagram, we see that Q_H is obtained when the gas transistion from state 2 to state 3 and Q_C is obtained when the gas transistion from state 4 to state 1. From the first law of thermodynamics (recall that $W = 0$ in isochoric processes), they can be calculated as follows:

$$Q_H = \Delta U_{2\to3} = C_v \Delta T_{2\to3} = C_v(T_3 - T_2)$$
$$Q_C = \Delta U_{4\to1} = -C_v \Delta T_{4\to1} = C_v(T_4 - T_1)$$

[8]. The efficiency is now:

$$e = 1 - \frac{T_4 - T_1}{T_3 - T_2}$$

Since transistions $1 \to 2$ and $4 \to 1$ are adiabatic, we apply the relation $V^{\gamma-1}T = \text{const}$ which gives

$$V_1^{\gamma-1}T_1 = V_2^{\gamma-1}T_2$$

---

[8]There is a minus sign in Q_C because we defined it to be the heat that flows **out** the engine.

$$V_2^{\gamma-1}T_3 = V_1^{\gamma-1}T_4$$

Subtracting these two equations,

$$V_2^{\gamma-1}(T_3 - T_2) = V_1^{\gamma-1}(T_4 - T_1)$$

$$\left(\frac{V_2}{V_1}\right)^{\gamma-1} = \frac{T_4 - T_1}{T_3 - T_2}$$

Substituting this result to our efficiency gives:

$$\boxed{e = 1 - \left(\frac{V_2}{V_1}\right)^{\gamma-1}}$$

The ratio V_1/V_2 is often called "compression ratio." The higher compression ratio is the better the efficiency is. However if the ratio is too large, the gas will become very hot and it will preignitation before being fully compressed (state 2).

The below table summarizes the energy flow in each state transistion in the Otto cycle.

Segment	Heat	Work **by** Engine	Temperature
$1 \to 2$	none	negative	increase
$2 \to 3$	absorb	none	increase
$3 \to 4$	none	positive	decrease
$4 \to 1$	release	none	decrease

2.11   Heat Capacity and Latent Heat

Recall that heat capacity is defined as

$$C = \frac{Q}{\Delta T}$$

It is the heat needed to change the temperature of a system by one degree. There is also specific heat capacity which is heat capacity per unit mass which is defined as:

$$c = \frac{C}{m}$$

The amount of heat required actually depends on what kind physical situation. If we rewrite the heat capacity definition in terms of first law of thermodynamics:

$$C = \frac{\Delta U - W}{\Delta T}$$

We see that the value W can vary. For isochoric process where $W = 0$, we get heat capacity for constant volume:

$$C_v = \left(\frac{\Delta U}{\Delta T}\right)_V = \left(\frac{\partial U(V,T)}{\partial T}\right)_V$$

The V subscript here indicates we treat volume V as a constant quantity when differentiating. For isobaric process where pressure is kept constant, we get heat capacity for constant pressure:

$$C_p = \left(\frac{\Delta U - (-P\Delta V)}{\Delta T}\right)_P = \left(\frac{\partial U(P,T)}{\partial T}\right)_P + P\left(\frac{\partial V(P,T)}{\partial T}\right)_P$$

The last term on the right is the additional heat needed to compensate for the energy lost as work. Notice that the more the volume increases, the larger this term is.

In some situations, it is possible to add energy to a system without increasing its temperature. This happens when the system is undergoing a first order phase transition: for example, melting (solid to liquid), evaporating (liquid to gas), sublimating (solid to gas, without melting in between). The heat capacity is infinite during transistion since there is no change in temperature:

$$C = \frac{Q}{0} = \infty$$

However, there is energy associated with this process. It is called latent heat (from Latin, 'latent' means hidden), the amount of heat required to melt or boil the substance completely. It is denoted by L. Latent heat is usually given per amount of substance m which is defined as the specific latent heat:

$$l = \frac{L}{m}$$

The value of latent heat actually depends on amount of work done, like heat capacity. But for most of the time, we can assume that pressure is kept constant at $1\,\text{atm}$.

EXAMPLE 39.

Consider a gas whose equation of state is $PV = Nk_BT$, but whose internal energy is $U = \frac{f}{2}Nk_BT + aP^2V$. What are its C_v and C_p?

$$
\begin{aligned}
C_v &= \left(\frac{\partial U}{\partial T}\right)_V \\
&= \frac{\partial}{\partial T}\bigg|_V \left(\frac{f}{2}Nk_BT + a\left(\frac{Nk_BT}{V}\right)^2 V\right) \\
&= \frac{f}{2}Nk_B + \frac{2a(Nk_B)^2 T}{V}
\end{aligned}
$$

$$
\begin{aligned}
C_p &= \left(\frac{\partial U}{\partial T}\right)_P + P\left(\frac{\partial V}{\partial T}\right)_P \\
&= \frac{\partial}{\partial T}\bigg|_P \left(\frac{f}{2}Nk_BT + aPNk_BT\right) + P\frac{\partial}{\partial T}\bigg|_P \left(\frac{Nk_BT}{P}\right) \\
&= \frac{f}{2}Nk_B + aPNk_B + Nk_B
\end{aligned}
$$

EXAMPLE 40.

Estimate how long (in seconds) it should take to bring a cup of water to boiling temperature in a typical $1000\,\text{W}$ microwave oven, assuming that all the energy ends up in the water. Something to think about: why is no heat involved in this process?

The (molar) heat capacity of water is $75.29\,\text{J}\,\text{K}^{-1}\,\text{mol}^{-1}$. A cup of water is about $250\,\text{mL}$. Assume the water is at $25\,°\text{C}$ to start with (typical room temperature).

We have

$$V = 0.000\,25\,\text{m}^3$$

$$\rho = 1\text{e}3\,\text{kg}\,\text{m}^{-3}$$

$$M_{H_2O} = 18\text{e}{-}3\,\text{kg}\,\text{mol}^{-1}$$

$$\Delta T = 75\,\text{K}$$

$$P = 1000\,\text{W}$$

The number of moles of water is given by

$$n = \frac{\rho V}{M_{H_2O}}$$

The heat necessary to boil a cup of water is:

$$Q = nc_m \Delta T = 78\,427.1\,\text{J}$$

The waiting time is given by

$$t = \frac{Q}{P} = 78.43\,\text{s}$$

There is heat from electromagnetic radiation.

EXAMPLE 41.

0.5 kg of ice, at $-20\,°\text{C}$, is taken out of the freezer and placed in a pot on a stove. If the stove's burner produces 1750 W of power, and the efficiency of heat transmission to the water is 65%, how many minutes will it take before the water in the pot starts to boil? You can ignore the evaporation of water while it's heating up?

The heat capacity of ice is $c_{ice} = 2.03\,\text{J}\,\text{g}^{-1}\,\text{K}^{-1}$, the latent heat for melting ice is $l = 333\,\text{J}\,\text{g}^{-1}$. The heat capacity of water is $c_{m,water} = 75.29\,\text{J}\,\text{K}^{-1}\,\text{mol}^{-1}$.

We have

$$m = 500\,\text{g}$$

$$n = \frac{0.5\,\text{kg}}{18\text{e}{-}3\,\text{kg}\,\text{mol}^{-1}} = 27.7778\,\text{mol}$$

$$P = 1750\,\text{W}(0.65) = 1137.5\,\text{W}$$

We need to find the total heat necessary to warm up ice from $-20\,°\text{C}$ to $0\,°\text{C}$, turn ice into liquid state, then warm up water from $0\,°\text{C}$ to $100\,°\text{C}$. The heat is:

$$Q = mc_{ice}(20\,\text{K}) + ml + nc_{m,water}(100\,\text{K}) = 395\,938.8889\,\text{J}$$

The total time it takes to see the pot to boil is:

$$t = \frac{Q}{P} = 5.8\,\text{min}$$

2.12   Intensive and Extensive Properties

Many properties of a system like pressure, volume, internal energy, etc. can be categorized either as **intensive** property or **extensive** property.

Intensive properties are properties whose magnitude that does not change when the system size changes. They are local (ie. can be defined at a point).

Extensive properties, in contrast, refer to properties whose magnitude changes with system size linearly.

If an extensive property is multiplied by an intensive property (keeping it fixed), it becomes another extensive property. When an extensive property is divided by another extensive property, it becomes intensive property. But if two extensive properties are multiplied together, it becomes neither.

Examples of intensive properties:

1. Temperature T (Since $U = \frac{f}{2}Nk_BT$, temperature is proportional to $\frac{U}{N}$. Extensive property divided by another becomes intensive property.)

2. Pressure P (According to the ideal gas law $P = \frac{Nk_BT}{V}$, so pressure is proportional to $\frac{N}{V}$ which is an extensive property.)

3. Density ρ (Proportional to $\frac{m}{V}$.)

Examples of extensive properties:

1. Mass m

2. Volume V

3. Number of particles N

4. Internal energy U (Depends on number of particles: $U = \frac{f}{2}Nk_BT$ and also the product of an extensive property and an intensive property is another extensive property.)

EXAMPLE 42.

Determine whether these quantities are extensive or intensive: $\frac{N}{P}$, PT, $\frac{U^2}{N}$, number of microstates Ω, PV and $\frac{V}{N}$.

Intensive:

1. $\frac{V}{N}$ (If we double the size of the system, it becomes $\frac{2V}{2N} = \frac{V}{N}$. The value does not change, so it is intensive.)

2. PT (Both P and T are intensive.)

Extensive:

1. $\frac{N}{P}$ (Extensive property divided by intensive property still remain extensive.)

2. $\frac{U^2}{N}$ (If we double the size of the system, $(2U)\frac{2U}{2N} = 2\frac{U^2}{N}$. The magnitude changes linearly, so it is extensive.)

---

3. PV

Note that Ω is neither intensive nor extensive. Consider a simple two state system with N particles and q energy quantas, the number of microstates is calculated by

$$\Omega(N, q) = \binom{N}{q}$$

If we double the size of the system, we get

$$\Omega(2N, 2q) = \binom{2N}{2q} > 2\binom{N}{q}$$

It does not change linearly at all. We call Ω a "super extensive" property.

EXAMPLE 43.

Rewrite the ideal gas law in terms of intensive quantities: P, T and mass density ρ (or number density, if you prefer).

$$PV = Nk_BT$$

Dividing volume both sides,

$$P = nk_BT$$

where $n = \frac{N}{V}$ is number density. Multiplying both sides by mass of the single particle m gives

$$mP = \rho k_BT$$

where $\rho = mn$ is mass density. Finally, we multiply both sides by Avogadro's number N_A,

$$MP = \rho N_A k_B T = \rho RT$$

where $M = mN_A$ is molar mass.

$$\boxed{PM = \rho RT}$$

2.13 Review of Basic Statistical Mechanics Concepts

Macrostate is the thermodynamic state of a system, fully described by macroscopic quantities such as energy (U), temperature, pressure, volume, etc...

Microstate is an exact microscopic description of the system, including, for example, all the positions and velocities of all the particles in a gas.

Each macrostate has many microstates associated with it. We called the number of these microstates Ω, the multiplicity. Ω is a function of state (the macrostate).

For example, for a gas of N particles, at volume V and energy U, we should be able to compute

$\Omega(N, V, U)$.

Fundamental assumption of statistical mechanics states that in a system whose internal energy and volume is fixed (isolated), all microstates are equally likely, and the system spends, on average, the same amount of time in each microstate.

It implies that an isolated system almost always evolves towards a macrostate with the largest multiplicity.

The larger the system is, the more the macrostate with the largest multiplicity dominates and the less likely any fluctuations away from it are sampled: this is the thermodynamic limit.

Once a large system reaches a macrostate close to one with the largest multiplicity, it will never spontaneously fluctuate away from it.

EXAMPLE 44.

Consider an isolated box of gas expanding into vacuum:
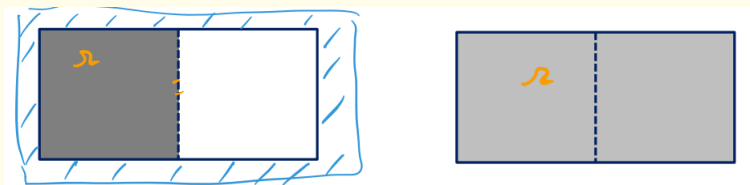
What can you conclude from this system?

Since the system is isolated, total heat flowing in is $Q = 0$ and work on the system is $W = 0$. According to the first law of thermodynamics, $\Delta U = 0$. Also, the volume of the whole system does not change, so $\Delta V = 0$.

Note that the pressure of the gas is changing, since it is expanding throughout the whole box.

---

3   Entropy

3.1   Types of Systems

Just before we define our notion of this mysterious "entropy", it will help to define what different kinds of systems mean. An **Open system** is one in which both matter and energy can be exchanged with its environment. A **Closed system** is a system where energy can be exchanged with the environment, but no matter. Finally, an **isolated** system is where neither matter nor energy can be exchanged with the environment.

As a quick example, suppose the system is a half cup of water. If we leave the top of the cup open, it's an open system as the water molecules can evaporate and leave the cup. If we close the top of the cup, then it's a closed system as we can still give energy to the cup (say, by heating it in a microwave) even if the water molecules can't escape from it. If the cup is perfectly insulated and closed off, then its an isolated system as neither matter nor energy can enter or exit the cup.

3.2   Entropy and Second Law of Thermodynamics

We saw that interacting systems, after long time, reach thermal equilibrium when they are in the macrostate with the greatest multiplicity. This is the result of the **second law of thermodynamics**: multiplicity tends to increase.

In order to quantify multiplicity, we define entropy:

$$S = k_B \ln \Omega$$

It is nothing but the scaled natural logarithm of the number of microstates of the system in some particular macrostate Ω. The reason we take logarithm of Ω is that Ω tends to be very large. Logarithm helps us to easily perform calculations involving large numbers.

The total entropy of a composite system containing two parts A and B is the sum of entropies of A and B:

$$S_{total} = k_B \ln \Omega_{total} = k_B \ln(\Omega_A \Omega_B) = k_B \ln \Omega_A + k_B \ln \Omega_B = S_A + S_B$$

This result can be generalized to system containing more than two parts. We also see the another reason to take logarithm of Ω. Recall from last section that Ω is a "super extensive" quantity. We take logarithm of Ω in order to turn it into a quantity which can be added up linearly. This makes the logarithm of Ω (a.k.a entropy) an extensive quantity!

In general, more particles there are in a system, the bigger the system is, and the more energy it contains, the greater its entropy and multiplicity.

There is more precise statement for the second law of thermodynamics. The total entropy of an *isolated* system either increases or remains constant in any spontaneous process[9], but it never decreases. The second law explains why heat can only flow spontaneously from the hot object to the cold object, but not not the other way around because it will decrease entropy.

For open or closed systems, it is more preferrable to rewrite the second law mathematically as follows:

$$\Delta S_{\text{total}} = \Delta S_{\text{system}} + \Delta S_{\text{surroundings}} \geq 0$$

The total entropy of the combined system and surroundings must increase or stay constant. A decrease in system entropy can only occur spontaneously if the entropy change of the surroundings is both positive in sign and has a larger magnitude than the entropy change of the system:

$$\Delta S_{\text{surroundings}} > 0$$

and

$$|\Delta S_{\text{surroundings}}| > |\Delta S_{\text{system}}|$$

One possible situation when entropy decreases is the freezing of a cup of water. Note that systems that can have its entropy decreased are referred to reversible processes which we will

---

[9]Spontaneous processes refers to a process in which a system evolute in time without any external input. The evolution makes the system reach to its thermal equilibrium. An example of such process is a dispersion of gas molecules in a sealed container.

discuss later.

Let's study an isolated system made up of two components: A and B. The total energy $U_{tot} + U_A + U_B$ is constant. The only macroscopic internal adjustment the system can make is to move energy between A and B. The macrostates are described by U_A alone.
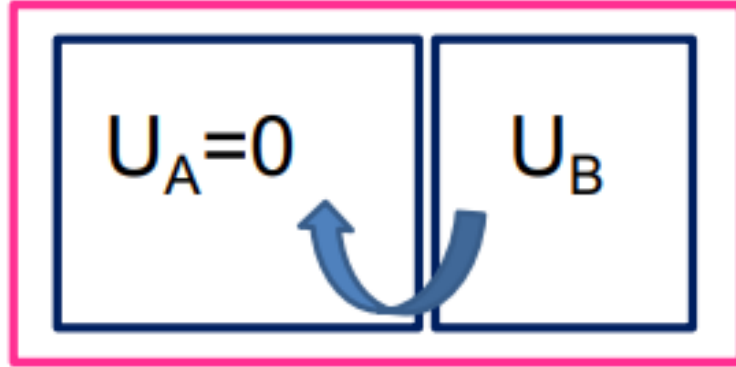
Suppose we start with $U_A = 0$. Since system A has zero energy, it has one possible microstate, so that means its initial entropy is zero: $S = k_B \ln 1 = 0$. On the other hand, system B has energy $U_B = U_{tot} - U_A$. It has higher entropy than system A. This means energy will spontaneously flow from system B to system. By the second law of thermodynamics, entropy of system A will increase while the entropy of system B will decrease.
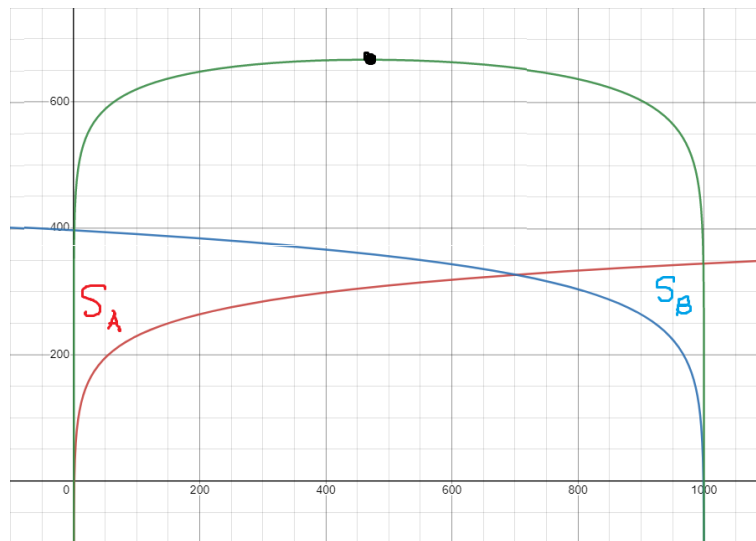

Figure 31: Graphs of S, S_A and S_B when $U_{tot} = 1000$ (Desmos link). Black dot is the equilibrium point

From the graph above, we see that the increasing of system A's entropy is faster than the decreasing of system B's entropy:

$$|\Delta S_A| > |\Delta S_B|$$

since the slope for the S_A curve is steeper than the slope for the S_B curve. The total entropy of the isolated system S will spontaneously keep increasing until thermal equilibrium (at $U_A = 465$) is reached where it will remain constant. At thermal equilibrium,

$$\frac{\partial S}{\partial U_A} = 0$$

$$\frac{\partial S_A}{\partial U_A} = \frac{\partial S_B}{\partial U_B}$$

Recall from last unit that at equilibrium, temperatures of A and B are equal. Also, we defined inverse temperature β as:

$$\beta = \frac{\partial \ln \Omega}{\partial U}$$

Let's redefine temperature in terms of entropy. Since $\beta = \frac{1}{k_B T}$,

$$\frac{1}{k_B T} = \frac{\partial \ln \Omega}{\partial U}$$

$$\frac{1}{T} = \frac{\partial k_B \ln \Omega}{\partial U} = \frac{\partial S}{\partial U}$$

Finally,

$$\boxed{\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{N,V}}$$

Temperature of a system is the reciprocal of the slope of its entropy vs. energy graph. The partial derivative is computed with the system's volume and number of particles treated as constant.

This new definition of temperature allow us to find a formula for measuring entropy changes. Suppose we add a bit of heat Q to a system while holding its volume V and number of particles N constant and doing no other forms of work, the first law of thermodynamics says:

$$\Delta U = Q$$

The change in entropy can be approximated using:

$$\Delta S = \left(\frac{\partial S}{\partial U} \right)_{N,V} \Delta U = \frac{Q}{T}$$

assuming that temperature is kept constant during the process.

EXAMPLE 45.

An object with temperature $T_H = 400\,\text{K}$ is placed briefly in thermal contact with a cooler object at temperature $T_C = 300\,\text{K}$. A small amount of heat $1200\,\text{J}$ flows between the two objects. How much has the entropy of the world changed?

The heat will flow from the hotter object to the colder object. The total entropy change is the sum of entropy changes for both objects:

$$\Delta S = \Delta S_H + \Delta S_C = \frac{Q_H}{T_H} + \frac{Q_C}{T_C}$$

where Q_H is the amount of heat flowed *in* the hot object which is negative in this question and Q_C is the amount of the heat flowed in the cold object.

$$\Delta S = \frac{-1200\,\text{J}}{400\,\text{K}} + \frac{+1200\,\text{J}}{300\,\text{K}} = 1\,\text{J}\,\text{K}^{-1}$$

If the temperature is instead varying, the total entropy should take account of varying temperature. Also, we use $Q = C_v dT$ to express the relationship in terms of constant volume heat capacity:

$$dS = \frac{C_v dT}{T}$$

The total change in entropy can be calculated by integration:

$$\Delta S = \int dS = \int_{T_i}^{T_f} \frac{C_v}{T} dT$$

EXAMPLE 46.

A steel ball with heat capacity $C_v = 1000 \, \text{J K}^{-1}$ is being warmed by the heat of the Sun. Its temperature increases from $300 \, \text{K}$ to $350 \, \text{K}$. How much does its entropy change? Assume C_v is constant.

$$\begin{aligned}
\Delta S &= \int_{T_i}^{T_f} \frac{C_v}{T} dT \\
&= C_v \ln\left(\frac{T_2}{T_1}\right) \\
&= 1000 \, \text{J K}^{-1} \ln\left(\frac{350}{300}\right) \\
&= 154 \, \text{J K}^{-1}
\end{aligned}$$