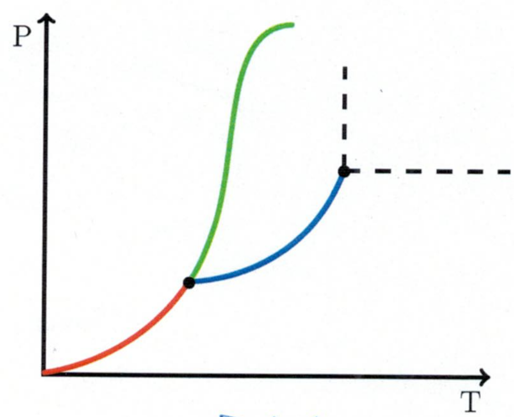
8.-10.7.17

SCRIPT

Statistical Mechanics and Thermodynamics

Professor: Roland Netz
Author: Martin Borchert
Date: Winter Semester 2017

- add summaries and all figures from lecture



- right bar in $\left.\frac{\partial U}{\partial S}\right|_{V,N}$ always long!

Please look out for errors and send them to
martin.b@fu-berlin.de

- Do not ~~cite~~ number each equation!
- Cite equations and make connections between sections
- add schematic figures (as in Lecture)
- more text and explanation
- some formula are missing!

$$m v_z^2 - \tilde{m} \tilde{v}'^2 = (m v_z - \tilde{m} \tilde{v}')^2 \frac{1}{m}$$

$$= m v_z^2 - 2 \tilde{m} v_z \tilde{v}' + \frac{\tilde{m}^2}{m} \tilde{v}'^2$$

$$\to \tilde{v}' \left(\tilde{m} + \frac{\tilde{m}^2}{m}\right) = 2 \tilde{m} v_z$$

$$\cancel{\tilde{v}' = \frac{2 \tilde{m} v_z}{\tilde{m} + \tilde{m}^2/m}} \qquad \tilde{m} \tilde{v}' = \frac{2 v_z \tilde{m}}{1 + \tilde{m}/m} = \frac{2 m v_z}{1 + m/\tilde{m}}$$

Contents

many-particle

1 Introduction

Calculate the properties

1.1 Objective of Statistical Mechanics and Thermodynamics

Systems

Consisting of a large

Statistical mechanics (SM) and thermodynamics (ThD) describe ~~systems with Avogadro~~ number of particles. Examples are gases, liquids, condensed matter, photon gas and many more; ~~so~~ almost all systems are made of many particles. True one particle systems are very rare. The objective of statistical mechanics is to ~~make assumptions about the behaviour~~ of a ~~total~~ system based on the properties and interactions between individual particles. In principle it is possible to write down all equations of motion for an N-particle system and solve them, which might get infinitely complicated if not impossible; but it is not very interesting anyway as ~~they~~ are not experimentally ~~feasible~~. measurable Experimentally only macroscopic quantities and properties are of interest, like temperature, pressure, heat capacity and so on. So one of the main objectives is to find those parameters based on behaviour of individual particles. The objective of thermodynamics is to ~~get the~~ relations between macroscopic parameters without knowing the microscopic details.

dense the theory individual particle trajectories

1.2 Systems and Equilibrium

There exist the following three systems ~~to distinguish between~~.

1.2.1 ~~Insulated /~~ Isolated system

This system can not interact with its environment, neither by exchanging energy nor by exchanging particles (example: closed Thermos jug).

1.2.2 Closed System

This system can exchange energy ~~in~~ with its environment but no particles (example: hot coffee in a closed glass jug).

1.2.3 Open System

This system can exchange particles and energy with its environment (example: hot coffee in an open cup).

1.2.4 Remarks

Of course the first two systems are idealised as there exists no real jug which does not exchange particles or energy with its environment. If all macroscopic parameters of a system are invariant in time, the system is in equilibrium. So if the system is not in equilibrium, the parameters will change in time until the system is in equilibrium. (In a stationary non-equilibrium state the macroscopic parameters do not change either, but energy flows through the system. An example would be a heat ~~sink~~ conductor between two ~~differently tempered~~ reservoirs. at different temperatures.

We will not consider non-equilibrium effects in this lecture.)

1.3 History

1.3.1 Thermodynamics

Thermodynamics (thermostatics would be a better name) describes macroscopic~~al~~ systems with the help of a few ~~phenomenological rules~~ heuristic laws which are called the laws of thermodynamics. These rules are not mathematically derived, but rather are generalisations or idealisations of experimental results. With ~~this a big~~ these laws a large number of predictions become possible. The advantage of this is the generality of the predictions, the disadvantage is that material specific properties like the heat capacity of gases are not deducible.

1.3.2 Statistical Mechanics

Statistical mechanics derives predictions with the help of statistical methods based on physical, microscopic~~al~~ laws.

1.3.3 Historical Origins

Historically, thermodynamics was created before the idea of statistical mechanics and also before the atomic~~al~~ structure of matter was understood. Important steps were

- Equivalence of heat and energy (Mayer 1842, Joule 1849)
- Formulation of thermodynamics (Clausius and Kelvin 1850, Gibbs 1878)

- Development of statistical mechanics of which the basics still hold today o.n. (Gibbs and Boltzmann 1860-1900)
- Improvements of statistical mechanics with quantum mechanics (>1900)

This lecture does not recreate the historic development, but will rather show the derivation of thermodynamics based on statistical mechanics.

1.3.4 Motivation

A mole (English name: mole, international unit: mol) of any substance is made of (by definition) $N_A = 6.0221 \cdot 10^{23}$ molecules. Here N_A is the Avogadro constant (or: Avogadro's constant, not to be confused with the historical, closely related term Avogadro's number). One mole of gas has a mass of 2 g-100 g and a volume of (at standard conditions (0°C and 10^5 Pa)) around 22.7 l. It seems to be infinitely complicated to calculate any parameters of a system this size, but it will become clear that statistical methods are especially good for larger systems. With this, accurate descriptions will actually become ~~a lot easier~~ possible (as long as one only tries to only calculate macroscopic properties of the system).

1.3.5 The Boyle–Mariotte Law

To understand how ~~thermodynamics~~ statistical mechanics works, a simple example will be solved using only Newton's laws and some basic statistics: An ideal gas made from N non-interacting particles is confined in a container with volume V. It is sealed with a lid of area A which can move up or down (due to an outside force F), while keeping the container sealed, effectively changing the volume inside the container. In equilibrium the pressure $P = F/A$ is compensated by the particles inside the container; here m is the mass of the particles, $\tilde{m}$ is the mass of the lid, v is the velocity of the particles, $\tilde{v}$ is the velocity of the lid, v' and $\tilde{v}'$ are the velocities after collisions between the lid and a particle.

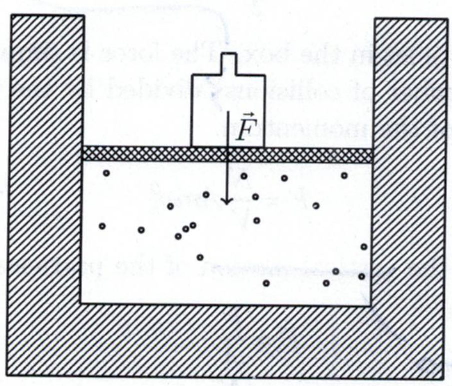


Figure 1.1: Container with N particles and volume V

With every elastic scattering between a particle and the lid the conservation of momentum and energy needs to be satisfied. Before the collision the lid is at rest. The conservation of momentum is the following.

picture of P colliding particle?

$$mv_z = mv'_z + \tilde{m}\tilde{v}' \tag{1.1}$$
$$mv_z - \tilde{m}\tilde{v}' = mv'_z \tag{1.2}$$
$$(mv_z - \tilde{m}\tilde{v}')^2 = m^2 v_z'^2 \tag{1.3}$$
$$mv_z'^2 = \frac{1}{m}(mv_z - \tilde{m}\tilde{v}')^2 \tag{1.4}$$

Here v'_z is the vertical velocity after the collision; all other components are not relevant. The conservation of energy is the following.

$\frac{1}{2}$

$$\frac{1}{2}mv_z^2 = \frac{1}{2}mv_z'^2 + \tilde{m}\tilde{v}'^2 \tag{1.5}$$

Rearranged and combined with equation (1.4) this becomes

$$0 = -2\tilde{m}\tilde{v}'v_z + \frac{\tilde{m}^2}{m}v_z'^2 + \tilde{m}v_z'^2 \tag{1.6}$$
$$\tilde{m}\tilde{v}' = \frac{2mv_z}{1 + {}^m/_{\cancel{M}}\ \tilde{m}}. \tag{1.7}$$

For a very heavy lid ($\frac{m}{\tilde{m}} \to 0$) the momentum transfer $\tilde{p}'$ becomes

$$\tilde{p}' = \tilde{m}\tilde{v}' \tag{1.8}$$
$$= 2mv_z. \tag{1.9}$$

The vertical distance Δz travelled by a particle in the interval Δt is $\Delta z = \Delta t v_z$. The probability p that a particle hits the lid (or any wall) in a volume of $V = Av_z\Delta t$ is $p = 0.5$. With that the number N' of particles hitting the box within Δt is

$$N' = \frac{1}{2}\frac{N}{V}Av_z\Delta t \tag{1.10}$$
$$= \frac{1}{2}\rho Av_z\Delta t. \tag{1.11}$$

$=N/V$ $\tilde{p}'$ N' Δt

Here ρ is the density of particles in the box. The force is simply the momentum transfer per collision times the number of collisions divided by the time interval, as the force is just the time derivative of the momentum.

$$F = \frac{N}{V}Amv_z^2 \tag{1.12}$$

The pressure P caused by the ~~vertical motion~~ collisions of the particles ~~on~~ with the lid is given by

$$P = \frac{F}{A} \tag{1.13}$$
$$= \frac{N}{V}mv_z^2. \tag{1.14}$$

where $\langle E_{kin} \rangle = \frac{m\langle \vec{v}^2 \rangle}{2}$

Actually, a ~~real~~ gas has rather a continuous distribution of velocities and the pressure should depend on the average square of the velocity $\langle v_z^2 \rangle$ (more on statistical distributions later). For symmetry reason $\langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle$ and $\langle \vec{v}^2 \rangle = \langle v_x^2 + v_y^2 + v_z^2 \rangle$. With that the pressure p can be written as

$$P = \frac{N}{V}\frac{2}{3}\langle E_{\mathrm{kin}} \rangle \cancel{= \frac{N}{V}\frac{m\langle \vec{v}^2 \rangle}{3}} \tag{1.15}$$

which ~~is by experimental evidence~~ agrees with the experimentally known relation

$$\boxed{P \propto \frac{1}{V}}. \tag{1.16}$$

From experimental data is is further known that $P \cdot V$ is constant for a fixed temperature, even for different atomic/ molecular masses. However it is clear that $\langle E_{\mathrm{kin}} \rangle$ becomes larger with higher temperatures. Actually with $\langle E_{\mathrm{kin}} \rangle$ and the use of the Boltzmann constant $k_B = 1.381 \cdot 10^{-23}\,\mathrm{JK}^{-1}$ the temperature T can be defined via

$$\langle E_{\mathrm{kin}} \rangle = \frac{3}{2} k_B T \tag{1.17}$$

$$= \frac{3}{2}\frac{PV}{N} \tag{1.18}$$

such that for 273.15 K water freezes and for 373.12 K water boils. k_B is not a fundamental constant of nature, but chosen in such a way that the temperature scale of thermodynamics matches the temperature scale of Celsius/ Kelvin. All this can summarised into the ideal gas law.

$$PV = N k_B T \tag{1.19}$$

1.3.6 Velocity of Gas Molecules

From the relationship between average kinetic energy and temperature results the average velocity to be

$$\sqrt{\langle v^2 \rangle} = \sqrt{\frac{3 k_B T}{m}}\,. \tag{1.20}$$

~~for T = 237 K.~~

For T=300K we obtain

For hydrogen ($m \approx 2 \cdot 1.661 \cdot 10^{-27}\,\mathrm{kg}$): $\sqrt{\langle v^2 \rangle} = 1800 \frac{m}{s}$ and
for oxygen ($m \approx 32 \cdot 1.661 \cdot 10^{-27}\,\mathrm{kg}$): $\sqrt{\langle v^2 \rangle} = 460 \frac{m}{s}$, so in fact molecules are moving quite fast. Furthermore $\langle E_{\mathrm{kin}} \rangle = \frac{3}{2} k_B T$ shows that there is a temperature point of absolute zero, at which all molecules stop moving ~~in a classical way~~ according to classical statistical mechanics (~~exception:~~ in disagreement with quantum mechanics).

2 Mathematical Statistics

This chapter is about the the mathematical aspects of statistics and how to make statistical predictions based on ~~microscopical, individual phenomena~~.
Large number of events.

2.1 Probabilities

An experiment ~~gets~~ is conducted N times. Each experiment results in an integer number m. After the experiments there will be N numbers m_i with $i \in \{1, \ldots, N\}$. The absolute (statistical) frequency of outcome m is called $n(m)$, which is the number of results m. The relative frequency of event m is called $h(m) = \frac{n(m)}{N}$, which ~~can be normalised~~ is as

$$\sum_m h(m) = 1. \tag{2.1}$$

In the limit $n \to \infty$ the relative frequency approaches the probability distribution.

$$p(m) = \lim_{N \to \infty} h(m), \tag{2.2}$$

where

$$\sum_m p(m) = 1. \tag{2.3}$$

2.1.1 Law of Additivity

For mutually excluding events the probability to observe one of a few events in an experiment is the sum of the individual probabilities.

$$p(m_1 \vee m_2 \vee m_3) = \cancel{p(m_1) +} p(m_1) + p(m_2) + p(m_3) + \cancel{\ldots} \tag{2.4}$$

Here $\vee$ is the logical or operator.

Example

~~**Experiment**~~

Now a ~~singular~~ single die is thrown. The probability to roll a (1) or a (2) is the sum of the individual probabilities.

$$p(1 \vee 2) = p(1) + p(2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} \tag{2.5}$$

2.1.2 Multiplication Theorem

simultaneously

Now independent experiments are considered. The probability of observing two events is the ~~probability~~ product of the individual probabilities.

$$p(m_1 \wedge m_2) = p(m_1) \cdot p(m_2) \tag{2.6}$$

Here $\wedge$ stands for logical and.

Example

~~Experiment~~ 1

Now two dice are thrown. The probability to roll two (1)s is the product of the individual probabilities.

$$p(1 \wedge 1) = p(1) \cdot p(1) \tag{2.7}$$

$$= \frac{1}{6} \cdot \frac{1}{6} \tag{2.8}$$

$$= \frac{1}{36} \tag{2.9}$$

The number of different orderings is given by the factorial 4!.

Experiment 2

Now four dice are thrown. The result shall be a straight (1,2,3,4). The first guess would be that the first dice needs to show a (1), the second one needs to show a (2) and so on. Then the probability would be $\frac{1}{6^4} = \frac{1}{1296} = 0.00077$. But the ordering of dice does not matter because when they are rolled together, they are indistinguishable: getting a (1)(2)(3)(4) is the same as (4)(3)(1)(2). So the probability becomes

$$p([1 \wedge 2 \wedge 3 \wedge 4] \vee [2 \wedge 3 \wedge 4 \wedge 1] \vee \ldots) = \frac{4!}{6^4} \tag{2.10}$$

$$= \frac{1 \cdot 2 \cdot 3 \cdot 4}{1296} \tag{2.11}$$

$$\approx 0.019. \tag{2.12}$$

2.1.3 Expectation Values and Variance

denoted

We

~~One usually~~ define~~s~~ an observable $x(m)$. The expectation value of x(m) is, using the normalised probability distribution $p(m)$, ~~given~~ by $\langle x \rangle$.

$$\langle x \rangle = \sum_m x(m)p(m) = \sum_{m=1} \frac{m}{6} \tag{2.13}$$

$$= \frac{1+2+3+4+5+6}{6} \tag{2.14}$$

$$= \frac{7}{2} \tag{2.15}$$

For the specific example
$x(m)=m$ we
obtain $\langle m \rangle =$

To estimate derivations from mean the variance Δx^{2} is defined as

$$\Delta x^2 = \langle (x - \langle x\rangle)^2\rangle \tag{2.16}$$
$$= \langle x^2 - 2x\langle x\rangle + \langle x\rangle^2\rangle \tag{2.17}$$
$$= \langle x^2\rangle - 2\langle x\rangle\langle x\rangle + \langle x\rangle^2 \tag{2.18}$$
$$= \langle x^2\rangle - \langle x\rangle^2 \tag{2.19}$$

The deviation is the the square root of the variance.

$$\Delta x = \sqrt{\langle x^2\rangle - \langle x\rangle^2} \tag{2.20}$$

Example

~~Experiment~~

Now one die is rolled. Again we consider $X(m) = m$.

$$\langle x^2\rangle = \langle m^2\rangle \tag{2.21}$$
$$= \sum_{m=1}^{6} m^2 p(m) \tag{2.22}$$
$$= \sum_{m=1}^{6} \frac{m^2}{6} \tag{2.23}$$
$$= \frac{1+4+9+16+25+36}{6} \tag{2.24}$$
$$= \frac{91}{6} \tag{2.25}$$
$$\Delta x^2 = \langle x^2\rangle - \langle x\rangle^2 \tag{2.26}$$
$$= \frac{91}{6} - \frac{7^2}{2^2} \tag{2.27}$$
$$= \frac{91}{6} - \frac{49}{4} \tag{2.28}$$
$$= \frac{182 - 147}{12} \tag{2.29}$$
$$= \frac{35}{12} \tag{2.30}$$
$$\Delta X = \sqrt{\frac{35}{12}} \tag{2.31}$$
$$\approx 1.7 \tag{2.32}$$

2.2 Binomial Distribution, Random Walk Example

A one dimensional random walker (in old literature sometimes called the drunk walker) makes a step up with probability p or a step down with probability $q = 1-p$ per time unit.

looking for food fluctuations

Examples of random walkers are animal ~~behaviour~~, protein configuration or diffusion processes. What is the probability that after $N = 5$ steps the random walker is at position $x = +1$? The walker ~~needs~~ to go up three steps and down two steps. The probability is (from section 2.1.2) $P = p^3q^2$. But there are ten different paths to achieve this.

$$\frac{5!}{3!2!} = 10 \tag{2.33}$$

If the up and down motion are equally probable with $p = \frac{1}{2}$, the total probability is

$$P_5(3) = \frac{5!}{3!2!}\cancel{p^5}\left(\frac{1}{2}\right)^5 \tag{2.34}$$
$$\approx 0.3. \tag{2.35}$$

In general if a walker does N steps, where m steps go up and $N-m$ go down, the number of different paths is

$$\frac{N!}{m!(N-m)!}. \tag{2.36}$$

The probability is

$$P_N(m) = \frac{N!}{m!(N-m)!}p^m q^{N-m} \tag{2.37}$$
$$= \frac{N!}{m!(N-m)!}p^m (1-p)^{N-m}. \tag{2.38}$$

This is called the binomial distribution.

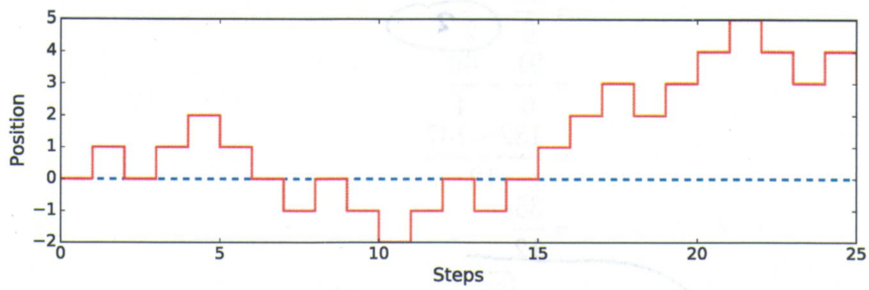


Figure 2.1: Random walker[1]

[1]Seed for numpy.random: 10

2.2.1 Moments of Binomial Distribution

Definition 1. *The first moment $\langle m \rangle$ of $P_N(m)$ is ~~denoted~~ given by the following expression, since $P_N(m)$ is normalised.*

$$\langle m \rangle = \sum_{m=0}^{N} m P_N(m) \tag{2.39}$$

$$= \sum_{m=0}^{N} \frac{mN!}{m!(N-m)!} p^m (1-p)^{N-m} \tag{2.40}$$

Evaluating the sum can be done explicitly with a trick.

$$1 = (p+q)^N \tag{2.41}$$

$$= \sum_{m=0}^{N} \binom{N}{m} p^m q^{N-m} \tag{2.42}$$

$$= \sum_{m=0}^{N} \frac{N!}{m!(N-m)!} p^m q^{N-m} \tag{2.43}$$

Here, ~~of course~~ $\binom{N}{m}$ is the binomial factor which tells the distribution of m objects to N places.

Proof that the Binomial Distribution is Normalised

For this set $q = 1-p$ and

$$\sum_{m=0}^{N} P_N(m) = (p+1-p)^N = 1^N = 1 \tag{2.44}$$

2.2.2 Calculation of the First Moment

With that the first moment can be calculated.

$$\sum_{m=0}^{N} \frac{N!}{m!(N-m)!} m p^m q^{N-m} = p \frac{\partial}{p} \sum_{m=0}^{N} \frac{N!}{m!(N-m)!} p^m q^{N-m} \tag{2.45}$$

with the use of the previous trick (equation (2.43)) this yields

$$= p \frac{\partial}{p} (p+q)^N \tag{2.46}$$

$$= pN(p+q)^{N-1}. \tag{2.47}$$

This holds true for all q, so it is also true for $q = 1-p$.

$$\sum_{m=0}^{N} \frac{N!}{m!(N-m)!} m p^m (1-p)^{N-m} = pN \tag{2.48}$$

$$= \langle m \rangle \tag{2.49}$$

2.2.3 Second Moment

Now for the second moment.

$$\sum_{m=0}^{N} \frac{N!}{m!(N-m)!} m^2 p^m q^{N-m} = p\frac{\partial}{\partial p} p \frac{\partial}{\partial p} \sum_{m=0}^{N} \frac{N!}{m!(N-m)!} p^m q^{N-m} \tag{2.50}$$

$$= p\frac{\partial}{\partial p} p \frac{\partial}{\partial p} (p+q)^N \tag{2.51}$$

$$= p\frac{\partial}{\partial p} \left[pN(p+q)^{N-1} \right] \tag{2.52}$$

$$= p\left[N(p+q)^{N-1} + pN(N-1)(p+q)^{N-2} \right] \tag{2.53}$$

Now q is set to $q = 1-p$ again.

$$= \langle m^2 \rangle \tag{2.54}$$

$$= pN + p^2 N(N-1) \tag{2.55}$$

This can be generalised to higher moments by doing this trick k times to reach the kth moment. From ~~this~~ the results for $\langle m \rangle$ and $\langle m^2 \rangle$ one can calculate the variance or mean-squared-deviation

$$\Delta m^2 = \langle (m - \langle m \rangle)^2 \rangle \tag{2.56}$$

$$= \langle m^2 \rangle - \langle m \rangle^2 \tag{2.57}$$

$$= pN + p^2 N(N-1) - p^2 N^2 \tag{2.58}$$

$$= Np(1-p) \tag{2.59}$$

$$\Delta m = \sqrt{Np(1-p)} \tag{2.60}$$

Here Δm is the ~~absolute~~ deviation from mean. The relative deviation is

$$\frac{\Delta m}{\langle m \rangle} = \cancel{\sqrt{\frac{Np(1-p)}{pN}}} \quad \frac{\sqrt{Np(1-p)}}{Np} \tag{2.61}$$

$$= \sqrt{\frac{1-p}{pN}} \tag{2.62}$$

Definition 2. *As N gets very large $\frac{\Delta m}{\langle m \rangle} = \sqrt{\frac{1-p}{pN}}$ goes to Zero. This is called the law of large numbers. Maybe the most important law in statistics.*

Example

~~Experiment~~ 1

N gas molecules are put into a box of volume V. N_A is Avogadro's constant (as described in section 1.3.4). Now the box is divided into two equal sub volumes. The probability to find m out of N gas molecules in one sub volume is given by the binomial distribution. The occupation probabilities are $p = q = \frac{1}{2}$. If $N = 10^{24}$ particles are put into the box the mean number in one of the boxes is $\langle m \rangle = Np = \frac{N}{2}$. The mean deviation $\Delta m = \sqrt{Np(1-p)} = \sqrt{\frac{N}{4}} = \frac{N^{12}}{2}$. The relative deviation though is only 10^{-12} which is absolutely negligible. This is the reason why thermodynamics and statistics work.

Experiment 2

With rare events it is the opposite. The relative deviations are large, statistics of course still works, but these problems need to be treated with extreme care. This is different from thermodynamics and statistical mechanics.

write down example from lecture!

2.3 Normal Distribution

For large N and finite p the expectation value $pN = \langle m \rangle$ gets very large and the binomial distribution simplifies.

$$\begin{aligned} ln[P_N(m)] &= \ln\left[\frac{N!}{m!(N-m)!}p^m q^{N-m}\right] &(2.63)\\ &= m\ln(p) + (N-m)\ln(q) + \ln(N!) - \ln(m!) - \ln[(N-m)!] &(2.64)\end{aligned}$$

The Stirling formula is

$$\begin{aligned} ln(N!) &= \ln\left(\prod_{j=1}^{N}\ln(j)\right) &(2.65)\\ &= \sum_{j=1}^{N}\ln(j) &(2.66)\\ &\approx \int_1^N dx\ln(x) &(2.67)\\ &= [x\ln(x) - x]_1^N &(2.68)\\ &= N\ln(N) - N + 1 &(2.69)\\ &= N\ln(N) - N + O(ln(N)) &(2.70)\end{aligned}$$

$$\ln(P_N(m)) = m\ln(p) + (N-m)\ln(q) + N\ln(N) - m\ln(m) - (N-m)\ln(N-m) \tag{2.71}$$

$P_N(m)$ is sharply peaked for large N. Now the Taylor expansion can be done ~~at~~ around m'.

$$\ln P_N(m) \equiv \ln(P_N(m')) + \frac{1}{2}(m-m')^2\frac{d^2}{dm^2}\ln P_N(m)|_{m=m'} + \dots \tag{2.72}$$

Here m' is defined by

$$\begin{aligned} \frac{d\ln P_N(m)}{dm}|_{m=m'} &= 0 &(2.73)\\ &= \ln p - \ln q - \ln m - 1 + \ln(N-m) + 1 &(2.74)\\ \ln\left(\frac{N-m}{m}\right) &= ln\left(\frac{q}{p}\right) &(2.75)\end{aligned}$$

where is the rest?

2.4 Poisson Distribution

In the limit that $N >> 1$ and $p << 1$ such that pN is finite the following holds.

$$\frac{N!}{m!(N-m)!}p(1-p)^{N-m} = P_N(m) \tag{2.76}$$

$$(1-p)^{N-m} = e^{(n-m)ln(1-p)} \tag{2.77}$$

$$\approx e^{-(N-m)p} \tag{2.78}$$

In the limit that $mp \to 0$

$$\approx e^{-Np} \tag{2.79}$$

$$\frac{N!}{(N-m)!} = e^{\ln N! - \ln(N-m)!}, \tag{2.80}$$

using the ~~useful~~ Stirling approximation ~~of second order~~ ($\ln N! = N\ln(N) - N$) this yields

$$\frac{N!}{(N-m)!} \approx e^{N\ln(N)-N-(N-m)\ln(N-m)+N-m} \tag{2.81}$$

$$= e^{N\ln(\frac{N}{N-m})+m\ln(N-m)-m} \tag{2.82}$$

$$= e^{-N\ln(\frac{N-m}{N})+m\ln(\frac{N-m}{N})+m\ln(N)-m} \tag{2.83}$$

$$= e^{-N\ln(1-\frac{m}{N})+mln(1-\frac{m}{N})+mln(N)-m} \tag{2.84}$$

$$\approx e^{m+O(\frac{m^2}{N})+O(\frac{m^2}{N})+m\ln(N)-m} \tag{2.85}$$

this

$$\approx e^{mln(N)} \qquad = N^m. \tag{2.86}$$

With all ~~of~~ the probability distribution becomes

$$P_N(m) = \frac{N^m}{m!}p^m e^{-Np} \tag{2.87}$$

$$= W(m). \tag{2.88}$$

And with $\lambda = Np$ the (already normalised) Poisson distribution can be written as

$$W(m) = \frac{\lambda^m}{m!}e^{-\lambda}. \tag{2.89}$$

The expectation value is $\langle m \rangle = Np$. The expectation value of the Poisson and binomial distribution are identical!

2.5 Continuous Distribution

Here x is a random variable, controlled by distribution $p(x)$ such that moments are given by

$$\langle x^n \rangle = \int_{-\infty}^{\infty} dx x^n p(x). \tag{2.90}$$

By this method all moments can be calculated from $G(k)$. In turn moments can be used to construct the Taylor expansion of G.

$$G(k) = \sum_{n=0}^{\infty} \frac{k^n}{n!} \frac{d^n G(k)}{dk^n}\Big|_{k=0} \tag{2.105}$$

$$= \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \cancel{f(x)} \rangle \tag{2.106}$$

Definition 4. *To know $G(k)$ is nice, but there is an even nicer distribution, which can be found by just ~~simply~~ taking the logarithm $\ln(G(k))$. With this new distribution cumulant moments can be defined as*

$$\langle x^n \rangle_c \equiv i^n \frac{d^n ln(G(k))}{dk^n}\Big|_{k=0}, \tag{2.107}$$

with

$$\ln(G(k)) = \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle_c. \tag{2.108}$$

2.6.1 Characteristic Function of the Normal Distribution

$$W(x) = \frac{1}{\sqrt{2\pi\Delta^2}} e^{\frac{-(x-x')^2}{2\Delta^2}} \tag{2.109}$$

$$G(k) = \int_{-\infty}^{\infty} dx e^{-ikx} W(x) \tag{2.110}$$

$$= (2\pi\Delta^2)^{-\frac{1}{2}} \int_{-\infty}^{\infty} dx e^{-ikx - \frac{(x-x')^2}{2\Delta^2}} \tag{2.111}$$

To solve this x is shifted to be $x = \tilde{x} + x'$ (with x' being finite, of course).

$$= (2\pi\Delta^2)^{-\frac{1}{2}} \int_{-\infty}^{\infty} d\tilde{x} e^{-ikx' - ik\tilde{x} - \frac{\tilde{x}^2}{2\Delta^2}} \tag{2.112}$$

By completing the square this yields

$$-ik\tilde{x} - \frac{\tilde{x}^2}{2\Delta^2} = -\frac{1}{2\Delta^2}(\tilde{x} + ik\Delta^2)^2 - \frac{k^2\Delta^2}{2} \tag{2.113}$$

$$G(k) = \frac{1}{\sqrt{2\pi\Delta^2}} e^{-ikx'} \int_{-\infty}^{\infty} d\tilde{x} e^{-\frac{k^2\Delta^2}{2} - \left(\frac{\tilde{x}+ik\Delta^2}{2\Delta^2}\right)^2}. \tag{2.114}$$

The expectation value of a function $\cancel{\langle f(x)\rangle}$ $f(x)$ is denoted by

$$\langle f\rangle = \int_{-\infty}^{\infty} dx f(x)p(x). \tag{2.91}$$

The discrete distribution $\cancel{p(x)}$ p_m can be written as

$$p(x) = \sum_m p_m \delta(x - x_m) \tag{2.92}$$

$$\langle x\rangle = \int_{-\infty}^{\infty} dx x p(x) \tag{2.93}$$

$$= \int_{-\infty}^{\infty} dx x \sum_m p_m \delta(x - x_m) \tag{2.94}$$

$= \sum x_m p_m$

As a reminder the delta function is defined by

$$\int_{-\infty}^{\infty} dx f(x)\delta(x - y) \equiv f(y) \tag{2.95}$$

$$\langle x\rangle = \sum_m p_m \int_{-\infty}^{\infty} dx \delta(x - x_m)\, x \tag{2.96}$$

$$= \sum_m p_m x_m \tag{2.97}$$

$$\cancel{= \langle x\rangle} \tag{2.98}$$

2.6 Characteristic Function

Definition 3. *The characteristic function is defined by*

$$G(k) = \langle e^{-ikx}\rangle \tag{2.99}$$

$$= \int_{-\infty}^{\infty} dx e^{-ikx} p(x). \tag{2.100}$$

This is the Fourier transform of the probability distribution. It can be used for the calculation of moments.

The derivatives of $G(k)$ are

$$\frac{d^n G(k)}{dk^n} = \int_{-\infty}^{\infty} (-ix)^n e^{ikx} p(x) \tag{2.101}$$

$$\frac{d^n G(k)}{dk^n}|_{k=0} = \int_{-\infty}^{\infty} dx(-ix)^n p(x) \tag{2.102}$$

$$= \langle(-ix)^n\rangle \tag{2.103}$$

$$= (-i)^n \langle x^n\rangle. \tag{2.104}$$

Again a shift needs to be done.

$$\tilde{x} = \tilde{\tilde{x}} - ik\Delta^2 \tag{2.115}$$

$$G(k) = \frac{1}{\sqrt{2\pi\Delta^2}} e^{-ikx' - \frac{k^2\Delta^2}{2}} \int_{-\infty}^{\infty} d\tilde{\tilde{x}} e^{-\frac{\tilde{\tilde{x}}^2}{2\Delta^2}} \tag{2.116}$$

$$= e^{-ikx' - \frac{k^2\Delta^2}{2}} \tag{2.117}$$

$$I = \cancel{\int_{-\infty}^{\infty} e^{\frac{x^2}{2\Delta^2}}} \quad ? \tag{2.118}$$

$$G(k) \cancel{= e^{ikx' - k^2\frac{\Delta^2}{2}}} \quad ? \tag{2.119}$$

$$\langle x^n \rangle = i^n \frac{d^n G(k)}{dk^n}|_{k=0} \tag{2.120}$$

$$\langle x^n \rangle_C = i^n \frac{d^n lnG(k)}{dk^n}|_{k=0} \tag{2.121}$$

The first two derivatives of $\ln(G(k))$are the following.

$$\frac{d\ln(G(k))}{dk} = -i\cancel{k}x' - k\Delta^2 \qquad \rightarrow \langle x \rangle_C = x' \tag{2.122}$$

$$\frac{d^2\ln(G(k))}{dk} = -\Delta^2 \qquad \rightarrow \langle x^2 \rangle_C = \Delta^2 \tag{2.123}$$

All ~~No~~ higher derivatives are zero: $\qquad \langle x^n \rangle_C = 0$ for $n \geq 3$ (2.124)

The first two cumulants (equations 2.122 and 2.123) characterise the normal distribution. Higher cumulants (equation 2.124) describe ~~the~~ deviation of a distribution from the normal distribution. ~~from it~~.

Explicit Expressions for Cumulants

$$G(k) = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle \tag{2.125}$$

$$= 1 - ik\langle x \rangle - \frac{k^2}{2}\langle x^2 \rangle + \frac{ik^3}{6}\langle x^3 \rangle + \frac{k^4}{24}\langle x^4 \rangle - \frac{ik^5}{120}\langle x^5 \rangle \ldots \tag{2.126}$$

$$= 1 - z \tag{2.127}$$

$$\ln(G(k)) = \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle \tag{2.128}$$

~~Here we can now use what we defined as z from equations 2.126 and 2.127~~

by expansion we obtain

$$= -ik\langle x \rangle - \frac{k^2}{2}\langle x^2 \rangle + \frac{k^2}{2}\langle x \rangle^2 + O(k^3) \tag{2.129}$$

and ~~with~~ by equating coefficients we get

$$\langle x\rangle_c = \langle x\rangle \tag{2.130}$$
$$\langle x\rangle_c = \langle x^2\rangle - \langle x\rangle^2 \tag{2.131}$$
$$\langle x^3\rangle_c = \langle x^3\rangle - e\langle x^2\rangle\langle x\rangle + 2\langle x\rangle^3 \tag{2.132}$$
$$= \langle (x-\langle x\rangle)^3\rangle \tag{2.133}$$
$$\langle x^4\rangle_c = \langle x^4\rangle - 4\langle x\rangle\langle x^3\rangle - 3\langle x^2\rangle^2 + 12\langle x^2\rangle\langle x\rangle^2 - 6\langle x\rangle^4 \tag{2.134}$$
$$\neq \langle (x-\langle x\rangle)^4\rangle . \tag{2.135}$$

2.7 Multi-Dimensional Probability Distributions

n random variables $x_1, x_2, \ldots, x_n$ are described by the joint distribution $P(x_1, x_2, \ldots, x_n)$ which is normalised. ~~If it is not normalised already, it usually can be normalised easily.~~ Often, but not always, $P(x_1, x_2, \ldots, x_n)$ factorises to

$$P(x_1, x_2, \ldots, x_n) = P(x_1)P(x_2), \ldots, P(x_n). \tag{2.136}$$

Definition 5. *Projection is the process of integrating out ~~all but one~~ random variables,* for example

$$P(x_1) = \int dx_2 \ldots x_n P(x_1, x_2, \ldots, x_n) \tag{2.137}$$

2.8 Central Limit Theorem

For m random variables x_i with mean value

$$y = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{2.138}$$

with all x_i following the probability distribution $p(x_i)$, what is the distribution of y? To make this more clear, an example.

2.8.1 Example

$p(x_i)$ is the weight distribution of a single person. What is the distribution of the total weight of $m = 100$ people? Here y obeys the probability distribution W.

$$W(y) = \int dx_1 \ldots dx_m \underbrace{p(x_1)\ldots p(x_m)}_{\text{multiplication theorem}} \delta\Big(y - \sum_{i=1}^{m}\frac{x_i}{m}\Big) \tag{2.139}$$

The moments $\langle y^n \rangle$ are

$$\langle y^n \rangle = \int dy y^n W(y) \tag{2.140}$$

$$= \int dx_1 \dots dx_m p(x_1) \dots p(x_m) \left[\sum_{i=1}^{m} \frac{x_i}{m} \right]^n \tag{2.141}$$

$$= \left\langle \left[\sum_{i=1}^{m} \frac{x_i}{m} \right]^n \right\rangle \tag{2.142}$$

Characteristic function of W

$$G(k) = \int dy e^{-iyk} W(y) \tag{2.143}$$

$$= \int dx_1 \dots dx_m p(x_1) \dots p(x_m) \int dy e^{-iyk} \delta(y - \sum_{i=1}^{m} \frac{x_i}{m}) \tag{2.144}$$

$$= \int dx_1 \dots dx_m p(x_1) \dots p(x_m) e^{-ik \sum_{i=1}^{m} \frac{x_i}{m}} \tag{2.145}$$

$$= \int dx_1 \dots dx_m p(x_1) \dots p(x_m) e^{\frac{-ikx_1}{m}} e^{\frac{-ikx_2}{m}} \dots e^{\frac{-ikx_m}{m}} \tag{2.146}$$

$$= \int dx_1 p(x_1) e^{\frac{-ikx_1}{m}} \int dx_2 p(x_2) e^{\frac{-ikx_2}{m}} \dots \int dx_m p(x,m) e^{\frac{-ikx_m}{m}} \tag{2.147}$$

$$= \left[\int dx_1 p(x_1) e^{\frac{-ikx_1}{m}} \right]^m \tag{2.148}$$

$$= \left[g\left(\frac{k}{m}\right) \right]^m \tag{2.149}$$

$$g(q) = \int dx p(x) e^{-iqx} \tag{2.150}$$

~~Now as a reminder~~ $ln(G(k)) = mln[g(\frac{k}{m})]$ and

We see that

$$\langle y^n \rangle_c \equiv i^n \frac{d^n ln(G(k))}{dk^n}|_{k=0} \tag{2.151}$$

$$= i^n m \frac{d^n ln g(\frac{k}{m})}{dk^n}|_{k=0} \tag{2.152}$$

$$= i^n m^{1-m} \frac{d^n ln g(q)}{dq^n}|_{q=0} \tag{2.153}$$

Here $k = qm$.

$$\langle y^n \rangle_c = m^{1-n} \langle x^n \rangle_c \tag{2.154}$$

This is called the central limit theorem and it is a very important concept in statistics.

back to example: discuss scaling of moments!