Physics 742:
Statistical Mechanics & Condensed Matter
Notes on statistical mechanics and complex networks

Dion O'Neale

March 10, 2020

Contents

1 Thermodynamics - a refresher

We'll start with a quick overview of some of the important concepts from thermodynamics. Deriving methods for linking these macroscopic concepts to the microscopic state of a system is the original motivation behind much of statistical mechanics. This section will also recap some content from 315/334.

1.1 Macroscopic laws of thermodynamics

In thermodynamics we study a system — the part of the world that we are interested in — that is separated from its surroundings — the rest of the universe — by some boundary.

A *thermodynamic system* is a macroscopic system whose behaviour is identified thanks to a small and finite number of quantities—the thermodynamic properties. If the state of a thermodynamic system can be fully characterized by the values of the thermodynamic variables, and if these values are invariant over time, one says that it is in a state of *thermodynamic equilibrium*. There are many variables which help characterise a macroscopic system, but only a few are actually important. These variables are not independent, and are related to each other with an *equation of state*. An equation of state guarantees that any partial derivatives (which we will see later) are not independent or random, but have mathematical and physical meaning.

For example: the *ideal gas law*

$$pV = nRT$$
$$pV = Nk_BT$$

can be written an equation of state,

$$p = f(\rho, T)$$
$$p = f(V, N, T)$$

where $f()$ is a function to be determined experimentally or with statistical mechanics.

We typically classify systems into three types based on the properties of walls that define the system boundary:

- **Adiabatic walls/isolated system:** no energy or matter is transferred in or out of the system.

- **Diathermal walls/closed system:** no matter can be transferred in or out of the system but heat can flow through the walls.

- **(Semi-)permeable walls/open system:** in addition to heat, one or more chemical species can be transferred through the walls.
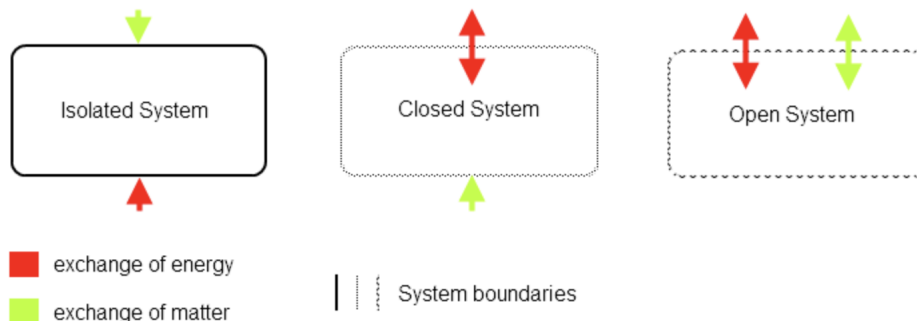

Figure 1: Types of systems, illustrated

The four laws of thermodynamics

Only four laws are required to construct the relationships that control much of classical thermodynamics. Theses are, in brief:

- **Zeroth law** Defines temperature and thermal equilibrium.

- **First law** Formulates the principle of conservation of energy for thermodynamic systems, heat/work are forms of energy; energy is conserved.

- **Second law** Entropy increases and systems involve towards equilibrium; heat spontaneously flows from high to low temperatures.

- **Third law** The absolute zero of temperature is not attainable and entropy $\to$ constant as $T \to 0$.

We'll look the first two laws in a bit more detail and then make use of the second law to derive some familiar properties of heat.

1.2 The zeroth law:

If two systems are each in thermal equilibrium with a third system, then they are also in thermal equilibrium with each other. This implies that they have some property in common. We call this property *temperature*. (In the language of mathematics we would say that thermal equilibrium is a transitive property.)

It is worth noting that thermal equilibrium is not the same as thermodynamic equilibrium. For the latter we also need mechanical equilibrium ($p_1 = p_2$ — there are no unbalanced forces) and chemical equilibrium ($\mu_1 = \mu_2$ — equal rates of reaction).

1.3 The first law

Heat and work are both forms of energy. The first law of thermodynamics says that energy remains constant for a (collection of) system(s) isolated from their surroundings. We denote work done *on* the system by the surroundings as $W > 0$, similarly, heat supplied to the system is denoted $Q > 0$. The opposite sign convention holds when it is by the system. Work is path dependent, i.e. it does not only depend on the initial and final states of the system.

When considering the change in energy ΔE of a system it is necessary to consider both work and heat. For example, if system A gains energy from system B, $\Delta E_A = -\Delta E_B \implies \Delta E_A + \Delta E_B = 0$. Since, in general, $\Delta E_A \neq W_{B \to A}$ there will also be a heat flow $Q_{B \to A}$, due to a temperature difference between systems A and B. So,

$$\Delta E_A = W_{B \to A} + Q_{B \to A}$$

$$\Delta E_B = W_{A \to B} + Q_{A \to B}.$$

Energy conservation gives

$$\underbrace{(W_{A \to B} + W_{B \to A})}_{\text{Work done by the composite system}} + \underbrace{(Q_{A \to B} + Q_{B \to A})}_{\text{Heat flow in the composite system}} = 0$$

In an isolated system the first law gives a sort of balance sheet for energy: $W_{A \to B} + W_{B \to A} = 0$ and $Q_{A \to B} + Q_{B \to A}$.

Suppose work $dW = -pdV$ is done on a closed system and heat is also introduced, but thermodynamic equilibrium is maintained. Then, the change in the internal energy U is $dU = dW + dQ$.

If work is done adiabatically (no energy or matter transfer), then entropy S is constant, so $dQ = TdS = 0$. Therefore, $dU = dW = -pdV$.

When N is constant: If heat is introduced with no change in volume or other such variable, $dW = 0$ so $dU = dQ = TdS$.

When N is not constant: If heat is introduced with no change in volume or other such variable, $dW = 0$ so $dU = dQ = \mu dN$.

1.4 The second law

The approach to equilibrium has to be characterized by an entropy that increases steadily up to the asymptotic value corresponding to equilibrium. This defines an *irreversible process*.

In a *reversible* process with $dW = 0$ the entropy necessarily remains fixed at whatever maximal value it had.

Heat flows from hot to cold. Systems in thermal contact will experience this until the temperatures are equal.

If particles can flow from one system to another, they do so with chemical potential from high to low until equal.

1.5 Internal energy

The *internal energy* of a system is the energy associated with the internal degrees of freedom of the system. This includes energy that might come from the motion of particles within the system, along with energy due to internal degrees of freedom that might be associated with the molecules that comprise those particles, for example rotational and vibrational degrees of freedom.

The equation of state for internal energy is often written as $E = E(S, V, N)$ where S is entropy, V is the volume of the system and N is the number of particles. (Note - all of these are extensive variables.) Looking at the first of these, we can calculate the total derivative of the E. We get

$$dE = \frac{\partial E}{\partial S}dS + \frac{\partial E}{\partial V}dV + \frac{\partial E}{\partial N}dN.$$

The partial derivatives above must be physically relevant. We identify these as

$$T = \frac{\partial E}{\partial S}, \quad p = -\frac{\partial E}{\partial V}, \quad \text{and} \quad \mu = \frac{\partial E}{\partial N};$$

temperature, pressure and chemical potential, respectively. (Note the sign on the term for pressure.) This also suggests that we could write the equation of state for internal energy using the intensive variables that we found as the dual variables of S and V; namely $E = E(T, p, N)$.

1.6 Intensive and extensive variables

Intensive variables control the state of a system but are independent of the system size. For example: temperature T, pressure p, density ρ, chemical potential μ

Extensive, or additive, variables are proportionate to the size of a system (i.e. they depend on N). For example: energy E, entropy S, heat capacity C, particle number N, volume V. One important extensive variable is the entropy (or disorder) of a system. As we just showed above, entropy is related to temperature by the internal energy E.

It is often helpful to think of intensive and extensive variables as coming in dual pairs with the intensive variable being the derivative of the internal energy with respect to the corresponding extensive variable given all other extensive variables are held constant. For example, $T = \frac{\partial E}{\partial S}|_{V,N}$.

Given two systems 1 and 2, an extensive variable for the composite system $1 \cup 2$ can be found by simply adding the individual extensive variables. For example, $N_{1\cup 2} = N_1 + N_2$, $V_{1\cup 2} = V_1 + V_2$, $E_{1\cup 2} = E_1 + E_2$ (E = internal energy). But this assumption clearly depends on the structural details of the interface. Imagine the oil-water interface in an emulsion like mayonnaise — in this case the ratio $\frac{E_{int}}{E_1 + E_2}$ is more like $\mathcal{O}(1)$. Similarly, if there are significant long range interactions, it may not be true that we can treat the extensive variables as being truly additive.

Figure 2: Oil-water interface

The internal energy is not entirely an extensive variable. Since there is often an interaction term between the two systems at the interface. i.e. $E_{1 \cup 2} = E_1 + E_2 + E_{int}$. Since E_{int} depends on the interface between the two systems it scales like an area as a function of system size, while E_1 and E_2 will scale like a volume, so typically $\frac{E_{int}}{E_1 + E_2} \to 0$ as the system size gets big. For example, an oil water interface.

1.7   Heat capacity:

Heat flowing into a system causes a change in temperature (except in the case of phase transitions). The amount of change in temperature for a given about of heat is the *heat capacity* of the system. The heat capacity of a system depends, in part on the experimental conditions of the system under consideration. Two important cases are constant volume and constant pressure.

For constant volume we have

$$C_V = \left. \frac{\delta Q}{dT} \right|_V$$

and similarly, for constant pressure,

$$C_p = \left. \frac{\delta Q}{dT} \right|_p .$$

The first law of thermodynamics (conservation of energy) implies that $\delta Q = dE - \delta W = dE + pdV$ (since $W = Fdx$ and $F = p \times Area \implies W = pdV$). So, since $dV = 0$ for constant volume,

$$C_V = \left. \frac{\delta Q}{\partial T} \right|_V = \left. \frac{\partial E}{\partial T} \right|_V .$$

In the constant pressure case

$$C_p = \frac{\delta Q}{dT}|_p = \frac{\partial E}{\partial T}|_p + p \frac{\partial V}{\partial T}|_p .$$

If we have constant pressure with a change in heat, it implies that there is a change in volume. We define the *volumetric thermal expansivity* as $\alpha_p = \frac{1}{V} \frac{\partial V}{\partial T}|_p$ and hence $V \alpha_p = \frac{\partial V}{\partial T}|_p$.

We can now write $C_p = \frac{\partial E}{\partial T}|_p + \alpha_p p V$.

We can also find an expression for C_p without resorting to introducing thermal expansivity. Constant pressure implies that $d(pV) = pdV$ so that $\delta Q = dE - dW = dE + pdV = d(E + pV)$. We define the composite quantity $E + pV$ as *enthalpy, H*.

The infinitesimal for enthalpy is $dH = \underbrace{dE + pdV}_{\delta Q} + Vdp = \delta Q + Vdp|_p$. In the constant pressure case, the last term is zero, so we get

$$C_p = \left. \frac{\delta Q}{\partial T} \right|_p = \left. \frac{\partial H}{\partial T} \right|_p .$$

1.8   Important thermodynamic quantities

Entropy is discussed in section 1.10.

Enthalpy H governs the equilibrium of adiabatic processes that occur while pressure is constant. Among all states that have the same entropy and pressure values, the state of equilibrium is the one that corresponds to the minimum value of enthalpy. Enthalpy is important for chemical reactions in open systems. If a system relaxes toward equilibrium while the pressure is kept constant, the maximum heat that can be produced by the system is equal to its variation in enthalpy. For this reason, enthalpy it is also called free heat.

$$H(S, p, ..., X_r) = E + pV$$

Among all states that have the same temperature and pressure values, the state of equilibrium is that in which the *Gibbs free energy* G assumes the minimum value.

$$G(T, p, ..., X_r) = E - TS + pV$$

The value of the *Helmholtz free energy* F is minimal for the equilibrium state among all virtual states at the given temperature T.

$$F(T, X) = E(S(T, X), X) - TS(T, X)$$

1.9   The fundamental hypothesis of thermodynamics

It is possible to characterise the state of a thermodynamic system by specifying the values of a set of extensive variables.

1.10   The central problem of thermodynamics

Given the initial state of equilibrium for several thermodynamic systems that are allowed to interact, we want to be able to determine the final thermodynamic state of equilibrium. The boundaries of the systems — adiabatic, closed, open — determine the types of interactions that are allowed, while the four laws of thermodynamics (and the first and second law in particular) determine how the composite system evolves. We want to be able to describe a final thermodynamic equilibrium state from amongst the space of all possible states for the composite system.

Entropy plays a special role in this problem due to the entropy postulate — the second law of thermodynamics.

The entropy postulate

There exists a function S of the extensive variables $X_1, X_2, \ldots, X_r$, called the entropy, that assumes a maximum value for a state of equilibrium among the space of possible states.

Entropy has the following properties:

1. Extensivity: If 1 and 2 are thermodynamic systems then $S_{1 \cup 2} = S_1 + S_2$.

2. Convexity: If $X^1 = (X_0^1, X_1^1, \ldots, X_r^1)$ and $X^2 = (X_0^2, X_1^2, \ldots, X_r^2)$ are two thermodynamic states of the same system then for $0 \leq \alpha \leq 1$

$$S((1 - \alpha)X^1 + \alpha X^2) \geq (1 - \alpha)S(X^1) + \alpha S(X^2).$$

   That is, the entropy of a linear combination of states for a single system is greater than or equal to the same linear combination of entropies of the individual states. A consequence of this is that if we take derivatives with respect to α and then evaluate at $\alpha = 0$ we get

$$\text{LHS:} \quad \frac{\partial}{\partial \alpha} S((1 - \alpha)X^1 + \alpha X^2) = \sum_{i=0}^{r} \frac{\partial S}{\partial X_i}(X_i^2 - X_i^1)$$

and
$$\text{RHS:} \quad \frac{\partial}{\partial \alpha}\left[(1-\alpha)S(X^1) + \alpha S(X^2)\right] = S(X^2) - S(X^1).$$

$$\text{Hence,} \quad \sum_{i=0}^{r} \frac{\partial S}{\partial X_i}(X_i^2 - X_i^1) \geq S(X^2) - S(X^1).$$

Mathematically, this means that the entropy surface (as a function of the other extensive variables) is always below the tangent plane of a point on the surface.

3. Monotonicity: $S(E, X_1, \ldots, X_r)$ is a monotonically increasing function of the internal energy E. That is, $\frac{\partial S}{\partial E}|_{X_1,\ldots,X_r} = \frac{1}{T} > 0$.

Using these three properties, it is possible to find the final equilibrium thermodynamic state amongst the space of possible states of a system. The equilibrium state is the state with maximum entropy that satisfies the constraints on the system.

An example: Consider two closed systems, 1 and 2, in thermal contact such that they can exchange energy, but nothing else (i.e. no other extensive quantities change). The space of possible states is defined by

$$E^1 + E^2 = X_0^1 + X_0^2 = E = \text{const.}$$

$$X_i^1 = \text{const.} \quad i = 1, 2, \ldots, r$$

$$X_i^2 = \text{const.} \quad i = 1, 2, \ldots, r.$$

We want to find the maximum of S as a function of E^1 (we could just as well use E^2). Start by taking the derivative of S.

$$\frac{\partial S}{\partial E^1} = \frac{\partial}{\partial E^1}\left(S^1(E^1, X_1^1, X_2^1, \ldots, X_r^1) + S^2(\underbrace{E - E^1}_{E = E^1 + E^2}, X_1^2, X_2^2, \ldots, X_r^2)\right)$$

$$= \frac{\partial S^1}{\partial E^1}\bigg|_{E^1} - \frac{\partial S^2}{\partial E^2}\bigg|_{E^2 = E - E^1}.$$

For E^1 at equilibrium we will write E_{eq}^1 (sim. for E^2). Then to maximise S we must have

$$\frac{\partial S}{\partial E^1}\bigg|_{E_{eq}^1} = \frac{\partial S^1}{\partial E^1}\bigg|_{E_{eq}^1} - \frac{\partial S^2}{\partial E^2}\bigg|_{E - E_{eq}^1} = 0$$

hence

$$\frac{\partial S^1}{\partial E^1}\bigg|_{E_{eq}^1} = \frac{\partial S^2}{\partial E^2}\bigg|_{E_{eq}^2}.$$

We already have our first result — the equilibrium state of the composite system is the one where the change in entropy of the two systems, with respect to, E is equal. Since $\frac{\partial E}{\partial S} = \frac{1}{T}$ (by monotonicity) this means that the equilibrium occurs when the temperatures of the two sub-systems are equal.

Where does the heat flow to in the system in order to reach this equilibrium? The system started in an initial state with $E = E_{in}^1 + E_{in}^2$. Since entropy increases to reach the maximum value at equilibrium we have

$$S^1(E_{eq}^1) + S^2(E_{eq}^2) \geq S^1(E_{in}^1) + S^2(E_{in}^2),$$

so,

$$S^1(E_{in}^1) - S^1(E_{eq}^1) + S^2(E_{in}^2) - S^2(E_{eq}^2) \geq 0.$$

The convexity property of entropy means that both systems 1 and 2 have

$$\frac{\partial S}{\partial E}\big|_{E_{in}}(E_{eq} - E_{in}) \geq S(E_{eq}) - S(E_{in})$$

8

and from the previous expression, the RHS of this inequality is bounded below by zero so we have

$$\left.\frac{\partial S^1}{\partial E^1}\right|_{E_{in}^1} (E_{eq}^1 - E_{in}^1) + \left.\frac{\partial S^2}{\partial E^2}\right|_{E_{in}^2} \underbrace{(E - E_{eq}^1}_{=E_{eq}^2} - E_{in}^2) \geq 0$$

But E is conserved so $E = E_{in}^1 + E_{in}^2$ and hence $E - E_{eq}^1 - E_{in}^2 = -(E_{eq}^1 - E_{in}^1)$. Therefore

$$\left[\left.\frac{\partial S^1}{\partial E^1}\right|_{E_{in}^1} - \left.\frac{\partial S^2}{\partial E^2}\right|_{E_{in}^2}\right] \left(E_{eq}^1 - E_{in}^1\right) \geq 0.$$

This implies that energy flows to the system with higher $\frac{\partial S}{\partial E}$ and since $\frac{\partial S}{\partial E} = \frac{1}{T}$ this means that energy flows into the system with lowest temperature, until the temperatures are equal.

2   Probability

This section gives an introduction to some tools and ideas from probability theory. One of the main reasons for including it is that it lets us give background to the final sub-section on Brownian motion, but we'll also use some of the methods from here in the final section of the course on complex networks.

Physical variables can be either continuous or discrete. Continuous variables, like the spatial position of a particle $\mathbf{r} = (x, y, z)$; $x \to x + dx$ allow us to apply tools from differential calculus. Discrete variables (like the spin of an electron, the face of a playing card, or the energy level of a quantum particle) require a different approach. As the number of discrete states gets gets large, it is possible to approximate a discrete system by a continuous one. However, there are often statistical reasons why this is not always the best approach. (For one example of why it is not always good to approximate a discrete distribution with a continuous one, see the article "Power-Law Distributions in Empirical Data" by Clauset, Shalizi, and Newman, SIAM Review (2009) doi:10.1137/070710111)

Here we are going to look at some simple discrete distributions, including one of the simplest — the binomial distribution which describes the probability of events that can take one of only two possible outcomes. The tool we will employ to study these distributions is *probability generating functions* (PGFs) (or generating functions for short). Although PGFs can be somewhat cumbersome, they provide a powerful general tool that can applied to a wide range of situations involving discrete probabilities. In this sense they are like the Taylor series of discrete probability.

2.1   Probability generating functions

A probability generating function (PGF) is a polynomial whose coefficients are the probabilities associated with each outcome of some random process. For example, the PGF for an equal probability die is

$$\mathcal{G}(x) = 0x^0 + \frac{1}{6}x^1 + \frac{1}{6}x^2 + \frac{1}{6}x^3 + \frac{1}{6}x^4 + \frac{1}{6}x^5 + \frac{1}{6}x^6.$$

As an example of how PGFs are a handy general tool, say we now wanted to find the PGF for the sum when we roll two such dice. We can just multiply the PGFs to get $\mathcal{G}'(x) = \mathcal{G}(x)\mathcal{G}(x) = 0x^0 + 0x^1 + \frac{1}{36}x^1x^1 + \frac{2}{36}x^1x^2 + \cdots + \frac{1}{36}x^6x^6$.

More generally, the set of probabilities associated with any discrete distribution can be used as coefficients of $\mathcal{G}(x)$:

$$\mathcal{G}(x) = P(X = r_0)x^0 + P(X = r_1)x^1 + \cdots = \sum_{n=0}^{\infty} P(X = r_n)x^n.$$

2.2   Properties of generating functions

$$\begin{aligned} \mathcal{G}(0) &= P(X = 0) \\ \mathcal{G}(1) &= \text{sum over all probabilities} = 1 \end{aligned}$$

The two properties above are easy to see and are not unexpected. More interesting is the fact that derivatives of PGFs can be used to obtain expressions for the mean, and higher order moments, of the distribution associated with the PGF.

$$\begin{aligned} \frac{d}{dx}\mathcal{G}(x) = \mathcal{G}'(x) &= \sum_r r P_r x^{r-1}, \qquad \text{and} \\ \mathcal{G}''(x) &= \sum_r r(r-1)P_r x^{r-2}. \end{aligned}$$

If we evaluate the first derivative of the PGF at 1 we get

$$\mathcal{G}'(1) = \sum_r r P_r.$$

The expected value of an arbitrary function $f(r)$ is given by $\langle f(r) \rangle = \sum_r f(r)P(X = r)$, therefore $\mathcal{G}'(1) = \langle r \rangle$ — the expected value, or mean, of the distribution.

Similarly, Evaluating $\mathcal{G}''(x)$ at 1 gives

$$\mathcal{G}''(1) = \sum_r r(r-1)P(X = r) = \langle X(X-1) \rangle.$$

Why would we care about being able to find $\langle X(X-1) \rangle = \langle X^2 - X \rangle$? Because the variance of X, is given by

$$V(X) = \langle X^2 \rangle - \langle X \rangle^2.$$

We are therefore able to construct $V(X)$ from combinations of $\mathcal{G}$, $\mathcal{G}'$, and $\mathcal{G}''$. We can use the fact that $\langle X^2 - X \rangle = \langle X^2 \rangle - \langle X \rangle$ I.e. $\sum(r^2 - r)P(X = r) = \sum r^2 P(X = r) - \sum rP(X = r)$ to get

$$
\begin{align}
V(X) &= \langle X^2 \rangle - \langle X \rangle^2 \tag{1} \\
&= \underbrace{\langle X^2 \rangle - \langle X \rangle}_{\mathcal{G}''(1)} + \underbrace{\langle X \rangle}_{\mathcal{G}'(1)} - \underbrace{\langle X \rangle^2}_{\mathcal{G}'(1)^2} \tag{2} \\
&= \mathcal{G}''(1) + \mathcal{G}'(1) - \mathcal{G}'(1)^2 \tag{3}
\end{align}
$$

Generating functions may not always be the shortest way to find the properties of some distribution, but they are a powerful way.

2.3 The binomial distribution

The binomial distribution describes the probability of getting r identical outcomes from a sequence of N events where each event has only two possible states, e.g. number of heads from a sequence of coin tosses, number of Au atoms from deposition of Au and Ag atoms in some sputtering process, or number of spin up electrons in an uncorrelated system.

The probabilities for the binomial distribution are given by

$$P(X = r) = \binom{N}{r} p^r q^{N-r}, \quad \text{where} \quad \binom{N}{r} = \frac{N!}{r!(N-r)!}, \quad r = 0, 1, \ldots, N$$

and where (mostly) $p + q = 1$. The PGF for the binomial distribution is therefore given by $\mathcal{G}(x) = \sum_{r=0}^{N} \binom{N}{r} p^r q^{N-r} x^r$. The coefficients in the binomial distribution are (not surprisingly) the coefficients from the binomial expansion theorem. This means that instead of writing the PGF as a sum, we can easily express it (and its derivatives) as a polynomial:

$$
\begin{align*}
\mathcal{G}(x) &= (xp + q)^N \\
\mathcal{G}'(x) &= N(xp + q)^{N-1} p \\
\mathcal{G}''(x) &= N(N-1)(xp + q)^{N-2} p^2.
\end{align*}
$$

Evaluating the above at $x = 1$ gives:

$$
\begin{align*}
\mathcal{G}(1) &= (p + x)^N = 1 \ \text{if } p + q = \\
\mathcal{G}'(1) &= N(p + q)^{N-1} P = Np \ \text{if } p + q = 1 \\
\mathcal{G}''(1) &= N(N-1)(p + q)^{N-2} p^2 \\
&= N(N-1)p^2 \ \text{if } p + q = 1.
\end{align*}
$$

As expected, we have $\langle X \rangle = Np$ for the event associated with the probability p. Similarly, we can calculate the variance of the binomial distribution using equation 3

$$
\begin{align*}
V(X) &= \mathcal{G}''(1) - \mathcal{G}'(1) - \mathcal{G}'(1)^2 \\
&= N^2 p^2 - Np^2 + Np - N^2 p^2 \\
&= N(p - p^2) \\
&= Np(1 - p) = Npq.
\end{align*}
$$

If you are feeling enthusiastic, compare the calculation above with the usual method for finding the variation of the binomial distribution without using a PGF.

2.4   The expected long-run outcome of a binary process

What is the net outcome we can expect from a sequence of N binary events? What is the variance in this net outcome? How long can we expect the net outcome to remain positive?

Think of a sequence of coin tosses where heads implies $+1$ and tails implies -1. We will denote the number of heads from N events as n_h and the number of tails as $n_t = N - n_h$, then the net outcome is $d = n_t - n_h = N - 2n_h$.

If $p = q = \frac{1}{2}$ then the expected value of d is $\langle d \rangle = 0$. (It's easy to check this: $\langle d \rangle = \langle N - 2n \rangle = N - 2\langle n \rangle = N - \frac{2}{2}N = 0$.)

We want to find the variation in d, or rather, the root-mean-squared (rms) deviation

$$d_{rms} = \left(\langle d^2 \rangle - \langle d \rangle^2 \right)^{\frac{1}{2}} = \langle d^2 \rangle^{\frac{1}{2}}.$$

At this point we'll drop the subscript in n since we only have n_h and since the probabilities are equal. From $d = N - 2n$ we have

$$\langle d^2 \rangle = \langle (N - 2n)^2 \rangle = \langle N^2 - 4Nn + 4n^2 \rangle.$$

We know that $\langle n \rangle = \mathcal{G}'(1) = Np = N/2$ and from the expressions for $\mathcal{G}''(1)$ and $\mathcal{G}'(1)$ we get $\langle n^2 \rangle = \mathcal{G}''(1) + \mathcal{G}'(1)$ where $\mathcal{G}''(1) = N^2 p^2 = N^2/4 - N/4$. We can now calculate $\langle d^2 \rangle$.

$$
\begin{aligned}
\langle d^2 \rangle &= N^2 - 4N\frac{N}{2} + 4\left(\frac{N^2}{4} - \frac{N}{4} + \frac{N}{2} \right) \\
&= N^2 - 2N^2 + N^2 + N \\
&= N.
\end{aligned}
$$

So, $d_{rms} = \sqrt{N}$ and the variance in the net expected outcome scales like $\sqrt{N}$ as N increases. I.e. after N steps we expect to be a rms distance of $\pm\sqrt{N}$ away from the mean of zero.

We can now look at how this would change if the binary process had some small bias $\alpha = |p - q| > 0$. In this case, at each step, the expected outcome is α and the bounds for the rms distance from the mean are $\alpha N \pm \sqrt{N}$.

2.5   Brownian motion

Brownian motion is a model that describes the behaviour of particles in suspension (e.g. of dust in air, fine silt in dirty water) that have some velocity but which are constantly changing direction as they are buffeted by other particles. Their motion can be understood through a calculation similar to that for the net outcome of the binary process above, but in 3D.

The conceptual model that is often used when discussing Brownian motion is that of a *random walk*. Consider a random walker who after n steps is at the position $\mathbf{r} = (x, y, z)$. Each step is given by $\mathbf{s}_{n+1} = \mathbf{r}_{n+1} - \mathbf{r}_n$ and has step length s_n in a random and uncorrelated direction $\hat{\mathbf{s}}_n$. Again (by symmetry) $\langle \mathbf{s} \rangle = 0$. We want to find how far a particle might have gone after N steps. We will define $\lambda^2 = \langle s^2 \rangle$ — the average square step-length. We write $\mathbf{r}_N$ for the position of the walker (or the particle) after N steps, then

$$
\begin{aligned}
\langle \mathbf{r}_N^2 \rangle &= \left\langle \left(\sum_{n=1}^{N} \mathbf{s}_n \right)^2 \right\rangle \\
&= \left\langle \sum_{n=1}^{N} \mathbf{s}_n^2 + 2\sum_{n=1}^{N-1} \sum_{m=n+1}^{N} \mathbf{s}_n \cdot \mathbf{s}_m \right\rangle \\
&= \sum \underbrace{\langle \mathbf{s}_n \rangle}_{=\lambda^2} + 2 \sum \sum \underbrace{\langle \mathbf{s}_n \cdot \mathbf{s}_m \rangle}_{=\langle \mathbf{s}_n \rangle \cdot \langle \mathbf{s}_m \rangle = 0}
\end{aligned}
$$

The first term in the expression is the mean squared step-length while the second term is zero since there is no correlation between the directions of each of the steps. Hence $\langle \mathbf{r}_N \rangle = N\lambda^2$ and the rms displacement is $\sqrt{N}\lambda$. If the number of steps is proportionate to the length of time that the particles move for, then we have that the mean-square displacement scales like $\sqrt{t}$. This is one of the key features of diffusive processes.

3   Statistical postulates

So far we have looked at the macroscopic properties of a thermodynamical system and at some ways of calculating properties of random processes that obey some given probability distribution. Now it is time to combine these ideas and have a first attempt at linking the microscopic behaviour of a thermodynamic system (an idealised gas) with some of its macroscopic properties.

The behaviour of a given (mechanical) system depends on both the structure of the system — described by its equations of motion — and on the *initial conditions* of the system. Therefore, in order to describe the behaviour of the system, we need both the laws of mechanics for the system along with some statistical postulates about the initial conditions of the system. (We could also require that we know the exact initial conditions for all the particles in the system, but this is not realistic for $\mathcal{O}(10^{23})$ particles.) Different choices of the statistical postulates can lead to different behaviour (not just different states) of the system. We will look at Maxwell's postulates for the initial positions and velocities of a dilute collection of gas particles. We will see that they allow us to derive the ideal gas law from a microscopic basis.

3.1   An ideal gas

We want to keep the equations of motion for our system of particles as simple as possible. We are going to assume that our system has the following properties:

- We have N identical point particles, each with mass m, constrained in a volume V.

- There are no mutual interactions — no van der Waals effects or inter-particle forces for us!

- The walls of the container constraining the particles are perfectly reflecting.

- The mechanical state of the system is known when the position $\mathbf{r}$ and velocity $\mathbf{v}$ is known for each particle and these variables evolve according to Newton's laws of motion.

3.2   Maxwell's postulates

Maxwell's postulates translate the ideal gas assumptions from the previous section into conditions that we can use with Newton's equations of motion in order to, in principle, describe everything about the behaviour of an ideal gas. From these we can infer macroscopic properties by simply integrating over the microscopic variables of the system.

We assume that the vectors describing the initial conditions of the system are randomly distributed. More specifically:

1. The vectors relating to each particle are independent from each other. (e.g. there is no interaction between particles when they are close.) This is generally not true for any but the most dilute systems. For such a non-interacting system, the system's state is determined when $dN = f(\mathbf{r}, \mathbf{v})$ is known, where dN is the number of particles in a box with sides $d\mathbf{r} = (dx, dy, dz)$ centered on the point $\mathbf{r} = (x, y, z)$ and where the corresponding velocities of the particles lie within a box $d\mathbf{v} = (dv_x, dv_y, dv_z)$ centered on $\mathbf{v} = (v_x, v_y, v_z)$. This relationship defines the *single-particle distribution* $f(\mathbf{r}, \mathbf{v})$.

2. The positions $\mathbf{r}$ are independent of the velocities $\mathbf{v}$. This means we can factorise $f(\mathbf{r}, \mathbf{v})$ as $f(\mathbf{r}, \mathbf{v})$ $= f_r(\mathbf{r})\, f_v(\mathbf{v})$.

3. The density of the gas is uniform within the volume so we can write $f(\mathbf{r}) = N/V = \rho = \text{constant}$ inside the volume (and zero outside it).

4. The velocity components are independent of each other so we can factorise f_v as $f_v(\mathbf{v}) = f_x(v_x)f_y(v_y)f_z(v_z)$.

5. The distribution $f_v(\mathbf{v})$ is isotropic in velocity space so that f_v depends only on the magnitude of the velocity $|\mathbf{v}| = v$

3.3 Micro to Macro

Using the postulates above, we can derive a macroscopic equation of state. We will prove that Maxwell's postulates will allow us to derive the equation of state for ideal gases.
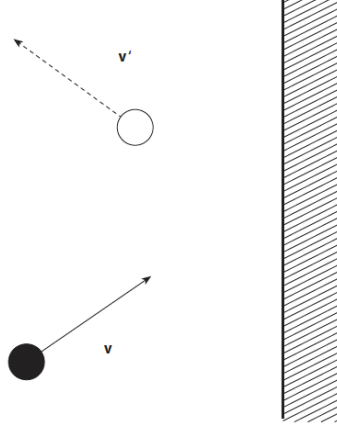

Figure 3: Particle impacting a wall

Consider a particle with velocity $\mathbf{v} = (v_x, v_y, v_z)$, moving from the left, that hits a wall parallel to the yz-plane. After impact its velocity is $\mathbf{v}' = (-v_x, v_y, v_z)$. Hence, the change in momentum is $\Delta \mathbf{p} = \mathbf{p}' - \mathbf{p} = m(\mathbf{v}' - \mathbf{v}) = m(-2v_x, 0, 0)$.

How many such impacts occur in a given time interval Δt on a region of the wall with area A? Take a box with a side of area A on the wall and perpendicular sides of length $v_x \Delta t$. This contains the particles of velocity $\mathbf{v}$ that can hit the wall. The volume of the box is $Av_x \Delta t$ and the number of particles it contains is $\rho A v_x \Delta t f_v(\mathbf{v})$, where $\rho = N/V$ is the density of all particles and f_v is the probability that a particle has velocity $\mathbf{v}$. The total momentum ΔP transmitted to the wall is therefore

$$\Delta P = \int_0^\infty dv_x \int_{-\infty}^\infty dv_y \int_{-\infty}^\infty dv_z f_v(\mathbf{v}) \rho A \Delta t (-2m) v_x^2 \mathbf{i}$$

where $\mathbf{i} = (1,0,0)$ picks out the x-component. i.e. integrate over the number of particles times the momentum per particle. The first integral only runs from zero since we are only interested in particles traveling towards the right, i.e. $\mathbf{v} > 0$. The factor of $\frac{1}{2}$ comes from integration over v_x, which runs only from $v_x > 0$.

The integral above gives

$$\Delta P = -2m\rho A \Delta t \langle v_x^2 \rangle \frac{1}{2}$$

where the average square velocities come from integrating over the velocity distribution and the factor of half comes from the integral over v_x, i.e. particles moving towards the right.

The force exerted by the gas on the wall is $F = \Delta P / \Delta t$ and pressure is force/area $= |F|/A$. We therefore have

$$p = m\rho \langle v_x^2 \rangle. \tag{4}$$

Compare this with the classical equation of state for an ideal gas $pV = NK_BT$ where T is the absolute temperature and $k_B = R/N_A$ is Boltzmann's constant ($R \simeq 8.3 J K^{-1} \text{mol}^{-1}$ is the ideal gas constant and $N_A \simeq 6 \times 10^{23} mol^{-1}$ is Avogadro's number). Since $\rho = N/V$, equation 4 suggests that we can relate the absolute temperature T to the mean square velocities of the particles (times a constant).

$$p = \rho k_B T$$