

# Physics 708: Statistical Mechanics

Dion O’Neale, Wouter van Zeist

March 21, 2018

## Contents

<b>1</b>	<b>Thermodynamics - A Refresher</b>	<b>2</b>
1.1	The Four Laws of Thermodynamics . . . . .	2
1.2	Zeroth Law . . . . .	3
1.3	First Law . . . . .	3
1.4	Third Law . . . . .	4
1.5	Heat Capacity . . . . .	5
1.6	Intensive & Extensive Variables . . . . .	6
1.7	The Fundamental Hypothesis of Thermodynamics . . . . .	6
1.8	The Central Problem of Thermodynamics . . . . .	6
1.9	Suggested Reading . . . . .	9
<b>2</b>	<b>Probability</b>	<b>9</b>
2.1	Probability generating functions . . . . .	9
2.2	Properties of generating functions . . . . .	10
2.3	The binomial distribution . . . . .	11
2.4	The Poisson distribution . . . . .	11
2.5	The expected long-run outcome of a binary process . . . . .	12
2.6	Brownian motion . . . . .	13
<b>3</b>	<b>Statistical postulates</b>	<b>13</b>
3.1	An ideal gas . . . . .	14
3.2	Maxwell’s postulates . . . . .	14
3.3	Equation of state . . . . .	15
<b>4</b>	<b>Phase space, Hamiltonian dynamics, and the fundamental postulate of statistical mechanics</b>	<b>16</b>
4.1	Phase space . . . . .	16
4.2	Observables . . . . .	17
4.3	The fundamental postulate of statistical mechanics . . . . .	18
4.4	Symplectic transformations and phase space area preservation . . . . .	18
4.5	Liouville’s theorem and preservation of phase space volume . . . . .	19
4.6	Recommended reading . . . . .	20

<b>5</b>	<b>Molecular dynamics and simulating Hamiltonian Systems</b>	<b>20</b>
5.1	Numerical integration . . . . .	22
5.2	Recommended reading . . . . .	25
<b>6</b>	<b>Ensembles</b>	<b>25</b>

## 1 Thermodynamics - A Refresher

We'll start with a quick overview of some of the important concepts from classical thermodynamics. Seeking a microscopic explanation for the macroscopic concepts presented here was an early motivation for the field of statistical mechanics. This section will also recap some content from 315.

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.

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

- **Adiabatic walls/isolated system** — no energy or matter is transferred
- **Diathermal walls/closed system** — no matter can be transferred 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.

### 1.1 The Four Laws of Thermodynamics

Only four laws are required to construct the relationships that control much of classical thermodynamics. We'll see later, that somewhat amazingly, these macroscopic laws arise naturally from the mathematics of the microscope systems.

These are, in brief:

- **Zeroth Law** Defines temperature and thermal equilibrium.
- **First Law** Formulates the principle of conservation of energy for thermodynamic systems. (Energy is conserved)
- **Second Law** Entropy increases; heat spontaneously flows from high to low temperatures.
- **Third Law** The absolute zero of temperature is not attainable.

We'll revisit 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 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, 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 ( $P_1 = P_2$ ) and chemical equilibrium ( $\mu_1 = \mu_2$  — equal rates of reaction).

## 1.3 First Law

Energy remains constant for a (collection of) system(s) isolated from the surroundings. The energy of a system changes if we do work on the system, and/or if we supply heat to the system. We denote work done *on* a system as  $W > 0$ , similarly, heat supplied *to* the system is denoted  $Q > 0$ .

When considering the change in energy  $\Delta E$  of a system it is necessary to consider both work and heat.

E.g. System  $A$  gains energy from system  $B$ , i.e.  $\Delta E_A = -\Delta E_B \implies \Delta E_A + \Delta E_B = 0$ . But, in general,  $\Delta E_A \neq W_{B \rightarrow A}$  since there can also be a heat flow  $Q_{B \rightarrow A}$  due to a temperature difference. So,

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

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

Energy conservation gives

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

In an adiabatic (isolated) system the first law gives a sort of balance sheet for energy/work/heat:  $W_{A \rightarrow B} + W_{B \rightarrow A} = 0$  and  $Q_{A \rightarrow B} + Q_{B \rightarrow A} = 0$ .

Revision: Adiabatic work and heat flow. Quasi-static processes (how slow is slow enough?). Path dependence of work. - see the notes from 315 that are in the repository. You may want to incorporate the relevant content from those notes, or other sources, into this document.

The *internal energy* of a system is the energy associated with the internal degrees of freedom of the system. If a system is at rest and the potential energy of any external field is unimportant, then the internal energy of a system is the total energy of a system.

The equation of state for internal energy is usually written  $E = E(T, V, N)$  or  $E = E(T, P, N)$ . Expressing these as infinitesimals (change in internal energy) gives

$$dE = \left. \frac{\partial E}{\partial T} \right|_{V,N} dT + \left. \frac{\partial E}{\partial V} \right|_{T,N} dV + \left. \frac{\partial E}{\partial N} \right|_{T,V} dN$$

a similarly, for the second formulation.

Note  $\left. \frac{\partial E}{\partial T} \right|_{V,N} \neq \left. \frac{\partial E}{\partial T} \right|_{P,N}$

## 1.4 Third Law

From the second law, we are able to determine the entropy in terms of  $T$  by evaluating the integral:

$$\int_{V_1}^{V_2} \frac{dQ_{\text{rev}}}{T} \Big|_T = S(V_2, T) - S(V_1, T)$$

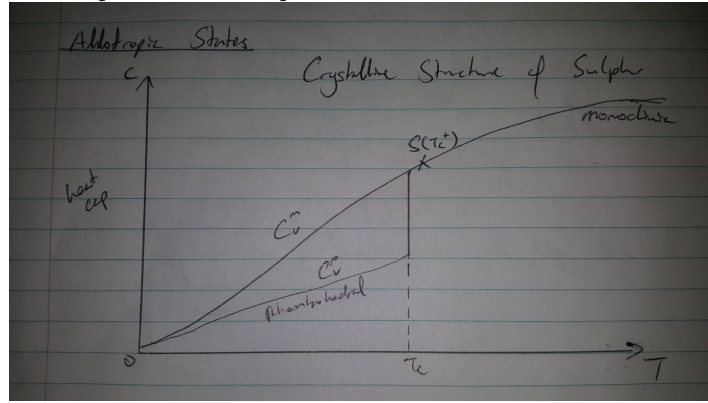
of some system. We repeat this process while constantly lowering the temperature all the way to zero temp such that we get what Nernst postulated:

$$\lim_{T \rightarrow 0} \Delta S_{V_1 \rightarrow V_2}(T) \rightarrow 0$$

As  $T \rightarrow 0$ , entropy becomes more and more independent of its coordinates. We can hypothesize that the entropy of all substances at  $T=0$  is the same universal constant set to zero.

We have small evidence for doing this for different substances.

Consider Allotropic states of sulphur:



We have a monoclinic state and a rhombohedral state defined by its heat capacities  $C_V^m$  and  $C_V^\rho$  respectively. By cooling the substance very slowly, we can observe a transition at a temperature  $T_c$ , releasing latent heat  $L$ . By cooling the substance very quickly, we can avoid this transition and maintain in metastable equilibrium. If we desire to obtain the entropy at a temperature slightly above  $T_c$ , we have two possible paths given:

$$\begin{aligned} 1 : S(T_c^+) &= \int_0^{T_c^+} \frac{dT C_v^m(T)}{T} + S_m(0) \\ 2 : S(T_c^+) &= \int_0^{T_c^-} \frac{dT C_v^\rho(T)}{T} + \frac{L}{T_c} + S_\rho(0) \end{aligned}$$

Which from measurements have verified that  $S_m(0) = S_\rho(0) = 0$  Which tells us that entropy is independent at  $T = 0$

Consider the consequences:

$$(1) \lim_{T \rightarrow 0} S(V, T) = 0 \rightarrow \lim_{T \rightarrow 0} \left. \frac{\partial S}{\partial V} \right|_T = 0$$

$$(2) \alpha = \left. \frac{1}{V} \frac{\partial V}{\partial T} \right|_P = - \left. \frac{1}{V} \frac{\partial S}{\partial P} \right|_T \rightarrow 0$$

$$(3) S(T, V) - S(0, V) = \int_0^T \frac{C_v dT'}{T'}$$

$$(4) \text{Unattainability of } T = 0$$

The first consequence represent another expression for entropy leading to zero as the temperature tends to zero.

The second consequence represent the extensivity of a system, manipulated by the maxwell's relation which tends to zero as the temperature tends to zero.

The third consequence represents a finite range of possible entropy values for a given finite temperature value. But if the heat capacity is constant, we will get a logarithmic value of entropy which blows up as the temperature tends to zero. This can only be corrected for  $\lim_{T \rightarrow 0} C_v(T) \rightarrow 0$

The fourth consequence implies that it is impossible to cool any system to absolute zero temperature in a finite number of steps.

## 1.5 Heat Capacity

Heat flowing into a system causes a change in temperature (except in the case of phase transitions).  $T \rightarrow T + \Delta T$ . 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

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

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  $W = PdV$ ). So  $C_V = \left. \frac{\delta Q}{dT} \right|_{V,N} = \left. \frac{\partial E}{\partial T} \right|_{V,N}$  (since  $dV = 0$  for fixed V.) In the constant pressure case  $C_P = \left. \frac{\delta Q}{dT} \right|_{P,N} = \left. \frac{\partial E}{\partial T} \right|_{P,N} + P \left. \frac{\partial V}{\partial T} \right|_{P,N}$ .

Constant pressure heat capacity implies that there is a change in volume, so think about how the volume changes. Define the *volumetric thermal expansivity*  $\alpha_p = \left. \frac{1}{V} \frac{\partial V}{\partial T} \right|_{P,N} \implies V \alpha_p = \left. \frac{\partial V}{\partial T} \right|_{P,N}$ .

We can now write  $C_P = \left. \frac{\partial E}{\partial T} \right|_{P,N} + \alpha_p P V$ .

An alternative way of thinking about  $C_P$ : constant pressure implies that  $PdV = d(PV)$  so that  $\delta Q = dE + \underbrace{dPV}_{=dW} = d(E + PV)$ . We define the composite quantity  $E + PV$  as *enthalpy*,  $H$ , with the equation of state  $H = H(T, P, N)$ .

The infinitesimal for enthalpy is  $dH = \underbrace{dE + PdV}_{\delta Q} + VdP = \delta Q + VdP|_P$ .

The last term is zero, since  $P$  is constant, so we get  $C_P = \frac{\delta Q}{\partial T}|_{P,N} = \frac{\partial H}{\partial T}|_{P,N}$ .

## 1.6 Intensive & Extensive Variables

Intensive variables describe the state of a system but are independent of the system size. E.g.  $T, P$

Extensive, or additive, variables are proportionate to the size of a system (i.e. they depend on  $N$ ).

It is possible to turn an extensive variable into an intensive variable by normalising by the system size. I.e. the variables come in dual pairs, e.g. pressure and volume.

(One important extensive variable is the entropy  $S$  (or disorder) of a system. The dual variable is temperature.)

The intensive variables can be found as derivatives 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. E.g.  $N_{1 \cup 2} = N_1 + N_2$ ,  $V_{1 \cup 2} = V_1 + V_2$ ,  $E_{1 \cup 2} = E_1 + E_2$  ( $E$  = internal energy).

Actually, the case of internal energy is not entirely correct 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  $\frac{E_{int}}{E_1 + E_2} \rightarrow 0$  as the system size gets big. E.g. an oil water interface (insert an image?).

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. (E.g. gravity — it might not work using stat mech to model celestial mechanics.)

## 1.7 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.

## 1.8 The Central Problem of Thermodynamics

Given the initial state of equilibrium of several thermodynamic systems that are allowed to interact, determine the final thermodynamic state of equilibrium.

(The boundaries of the systems — adiabatic, closed, open — determine the types of interactions that are allowed.) We want to pick a final thermodynamic equilibrium state from the space of all possible equilibrium states.

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  $X_1, X_2, \dots, X_r$  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: ( $S$  is an extensive variable) If 1 and 2 are thermodynamic systems then  $S_{1 \cup 2} = S_1 + S_2$ .
2. Convexity: If  $X^1 = (X_0^1, X_1^1, \dots, X_r^1)$  and  $X^2 = (X_0^2, X_1^2, \dots, X_r^2)$  are two thermodynamic states of the same system then for  $0 \leq \alpha \leq 1$

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

. I.e. the entropy of a linear combination of states 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  $\alpha$  at  $\alpha = 0$  we get

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

and

$$\frac{\partial}{\partial \alpha} [(1 - \alpha)S(X^1) + \alpha S(X^2)] = S(X^2) - S(X^1).$$

So,

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

So the entropy surface (as a function of the other extensive variables) is always below the tangent plane of a point on the surface. We'll use this soon.

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

Using these three properties, it is possible to find the final equilibrium thermodynamic state amongst the space of possible states. The equilibrium states is the state with maximum entropy that satisfies the constraints on the system. (I.e. Maximum disorder - can also think about this in terms of it being the state with the greatest possible number of corresponding microstates, but we haven't quite got that far yet.)

Worked example. Consider two 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

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

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

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

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  $S$ .

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

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

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

so  $\frac{\partial S^1}{\partial E^1} \Big|_{E_{eq}^1} = \frac{\partial S^2}{\partial E^2} \Big|_{E_{eq}^2}$ . We already have our first result — the equilibrium state is the one where the change in entropy of the two systems, wrt,  $E$  is equal. (The monotonicity postulate equated this with inverse temperature; hence, the equilibrium state must be the one where the temperatures of the two systems are equal.)

Where does the heat in flow to in the system? 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  $S^1(E_{eq}^1) + S^2(E_{eq}^2) \geq S^1(E_{in}^1) + S^2(E_{in}^2)$ , i.e.  $S^1(E_{in}^1) - S^1(E_{eq}^1) + S^2(E_{in}^2) - S^2(E_{eq}^2) \geq 0$ .

The convexity property of entropy means that both systems have  $S(E_{eq}) - S(E_{in}) \leq \frac{\partial S}{\partial E} \Big|_{E_{in}} (E_{eq} - E_{in})$  and from the previous expression, the LHS of the inequality is bounded below by zero so we have

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

But  $E$  is conserved so  $E = E_{in}^1 + E_{in}^2$  so we have  $E - E_{eq}^1 - E_{in}^2 = -(E_{eq}^1 - E_{in}^1)$  and therefore

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

This implies that energy flows to the system with higher  $\frac{\partial S}{\partial E}$ . From the monotonicity postulate we identified  $\frac{\partial S}{\partial E} = \frac{1}{T}$  so energy flows into the system with lowest temperature, until the temperatures are equal.



## 1.9 Suggested Reading

Chapter 3 of *Statistical Mechanics made Simple* (It's on the recommended reading list) covers the elements of thermodynamics nicely. Sections 3.1 to 3.5 in particular go over most of the concepts we have looked at here and introduce a few additional ideas.

## 2 Probability

This section is mostly an introduction (or a refresher?) 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  $\mathbf{r} = (x, y, z)$ ;  $x \rightarrow x + dx$  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 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 an example of the issue of approximating a discrete distribution with a continuous one, see the article Clauset, Shalizi, and Newman. "Power-Law Distributions in Empirical Data." SIAM Review (2009) doi:10.1137/070710111) [It is probably a good idea to add a bibliography to the notes and use LaTeX citations, (along with naming the reference in the text to make it more readable.)]

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* (or generating functions for short). Although PGFs can be somewhat cumbersome, they provide a very powerful general tool that can be 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  $\mathcal{G}'(x) = \mathcal{G}(x)\mathcal{G}(x) = 0x^0 + 0x^1 + \frac{1}{36}x^1x^1 + \frac{2}{36}x^1x^2 + \dots + \frac{1}{36}x^6x^6$ .

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.

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

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

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

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.  $\sum (r^2 - r)P(X = r) = \sum r^2 P(X = r) - \sum r P(X = r)$  to get

$$V(X) = \langle X^2 \rangle - \langle X \rangle^2 \tag{1}$$

$$= \underbrace{\langle X^2 \rangle - \langle X \rangle}_{\mathcal{G}''(1)} + \underbrace{\langle X \rangle}_{\mathcal{G}'(1)} - \underbrace{\langle X \rangle^2}_{\mathcal{G}'(1)^2} \tag{2}$$

$$= \mathcal{G}''(1) + \mathcal{G}'(1) - \mathcal{G}'(1)^2 \tag{3}$$

## 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

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

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:

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

Evaluating the above at  $x = 1$  gives:

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

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

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

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 Poisson distribution

Since the Poisson distribution shows up quite often in stat mech, we will have a look at how its properties can be derived via PGFs. The Poisson distribution

describes the probability of  $r$  distinct events occur in a time interval, given that the long-run average rate is  $\lambda$  per time interval. It has probabilities given by  $P(X = r) = \frac{\lambda^r}{r!} \exp(-\lambda)$ . The PGF for the Poisson distribution is given by

$$G(x) = \sum_r = 0^\infty \frac{\lambda^r}{r!} \exp(-\lambda) x^r = \exp(-\lambda) \sum \frac{(\lambda x)^r}{r!}.$$

Note that the sum in the last term of the expression above is the power series expansion for the exponential function. Handy — we can now write:

$$\begin{aligned} \mathcal{G}(x) &= \exp(-\lambda) \exp(\lambda x) \\ \mathcal{G}'(x) &= \lambda \exp(-\lambda) \exp(\lambda x) \\ \mathcal{G}''(x) &= \lambda^2 \exp(-\lambda) \exp(\lambda x) \end{aligned}$$

So we have  $\langle X \rangle = \mathcal{G}'(1) = \lambda$ , as per the definition, and

$$V(X) = \mathcal{G}''(1) + \mathcal{G}'(1) - \mathcal{G}'(1)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

## 2.5 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 deviation

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

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.

[Draw/add a figure of this with  $N$  as the independent variable.]

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  $\alpha$  and the bounds for the rms distance from the mean are  $\alpha N \pm \sqrt{N}$ .

[Draw/add this too.]

## 2.6 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  $\lambda^2 = \langle \mathbf{s}^2 \rangle$  — the average square step-length. We write  $\mathbf{r}_N$  for the position of the walker (or the particle) after  $N$  steps, then

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

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  $\langle \mathbf{r}_N^2 \rangle = N\lambda^2$  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

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 Equation of state

Maxwell's postulates allow us to derive the equation of state for an ideal gas (the ideal gas law). They provide a microscopic interpretation of absolute temperature in terms of kinetic energy.

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  $\Delta t$  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  $\rho Av_x\Delta t f_v(\mathbf{v})$ , 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  $\Delta P$  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 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

$$p = m\rho \langle v_x^2 \rangle. \quad (4)$$

Compare this with the classical equation of state for an ideal gas  $pV = nRT$  where  $n = N/N_A$ ,  $T$  is temperature and  $R = 8.3J/K$  is a constant. If we introduce the Boltzman constant  $k_B = R/N_A$  and the particle density  $\rho = n/V$ . The ideal gas law becomes  $p = \rho k_B T$ . Comparing the ideal gas law with equation 4 we find that we can relate the absolute temperature  $T$  to the mean square velocities of the particles.

## 4 Phase space, Hamiltonian dynamics, and the fundamental postulate of statistical mechanics

### 4.1 Phase space

At a microscopic level, a thermodynamic system is comprised of  $N$  particles, each with some mass  $m_i$ , which can interact with each other and with the boundaries of the system. (Although we are going to motivate this discussion with a thermodynamic system, everything that follows is also true for other systems where the total energy of the system is conserved — i.e. anywhere that a Hamiltonian could show up.) If we know the positions and momenta for each of the particles then we can completely describe the state of the system. In three dimensions, this means we must know the values for each of the  $6N$  degrees of freedom (3 for position and 3 for momentum); more generally, we have  $2d$  d.o.f. where  $d$  is the number of d.o.f. for the positions. That is the state of the system is described by a point in a  $6N$  (or  $2d$ ) dimensional phase space. Once a point in the microscopic phase space is known, the evolution of the system is completely governed by the equations of motion for the system. These in turn are prescribed by the Hamiltonian of the system

$$H = T + U = H(\mathbf{r}, \mathbf{v}) = H(\mathbf{q}, \mathbf{p}),$$

$T$  = kinetic energy,  $U$  = potential energy. In general,  $H$ ,  $T$ , and  $U$  may all depend on both the positions and the momenta of the particles, however, for many systems we find that, conveniently,  $T = \sum_{i=1}^N T_i$ , where  $T_i = \frac{\mathbf{p}_i^2}{2m_i}$ , and the potential  $U$  depends only on the positions:  $U = U(\mathbf{q})$ .

If the Hamiltonian is known for a system, then the equations of motion for that system are given by the *Hamiltonian equations of motion*:

$$\begin{aligned} \frac{d\mathbf{q}_i}{dt} &= \frac{\partial H}{\partial \mathbf{p}_i}, \\ \frac{d\mathbf{p}_i}{dt} &= -\frac{\partial H}{\partial \mathbf{q}_i}. \end{aligned}$$

If we write  $\mathbf{x} = [\mathbf{q}^\top, \mathbf{p}^\top]^\top$ , then we can compactly express the equations of motion as

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}) = J^{-1} \nabla H(\mathbf{x}), \quad (5)$$

where

$$J = \begin{bmatrix} 0 & -Id \\ Id & 0 \end{bmatrix}.$$

( $Id$  is the identity matrix of dimension  $d$ .)

Given a Hamiltonian,  $H$ , and a suitable set of initial conditions  $\mathbf{x}_0$ , the evolution of the solution  $\mathbf{x}(t)$  can be quite complicated. (Hamiltonian systems are one of the archetypal systems for studying chaotic dynamics). However, the fact that the system is Hamiltonian provides us with some additional information



that can help us characterise the state of the system. For example, it is not hard to prove that a solution  $\mathbf{x}(t)$  that satisfies  $\dot{\mathbf{x}} = f(\mathbf{x}) = J^{-1}\nabla H(\mathbf{x})$ ,  $\mathbf{x}_0 = \mathbf{x}(0)$  has  $H(\mathbf{x}(t)) = H(\mathbf{x}_0) = \text{const.}$  That is, Hamiltonian systems conserve energy along their solutions. We identify the Hamiltonian  $H$  with the internal energy of the system.

Knowing that  $\frac{dH}{dt} = 0$ , we can reduce the size of the phase space where  $\mathbf{x}(t)$  can lie, from  $2d$  d.o.f. down to  $2d - 1$ . If we know other quantities that remain invariant over time (e.g. the total momentum of the system) then we can further reduce the size of the phase space. Strictly speaking, there are  $2d - 1$  invariants (AKA first integrals or constants of motion) for a Hamiltonian system. This defines a curve in the remaining 1 d.o.f. along which the solution must lie (that is, a line in the phase space). This can be seen as a consequence of the uniqueness theorem for differential equations (think back to the 2nd year maths course on D.E.s). In reality, we typically can't find a closed form expression for most of these conserved quantities and have no idea what they might be.

It is possible that a (vanishingly) small change in the positions of the particles in phase space can lead to a large change in the state of the system and in the subsequent trajectories of the particles. This is similar to saying that the dynamics of a Hamiltonian system can switch abruptly from regular to chaotic motion, as some parameter of the system is varied. Such quantities are not helpful in characterising the thermodynamics of the system. In statistical mechanics, we are interested in the invariant quantities that only change by a little when there is a small change in the microscopic variables of the system. E.g.  $\mathbf{x} \rightarrow \mathbf{x} + \delta\mathbf{x} \implies H \rightarrow H + \delta H$ .

One way to phrase this is to say that we want to consider systems where the region in phase space where the system is evolving can be (almost) completely identified, for thermodynamic purposes, by the values of a (not too large) number of observable invariant quantities.

## 4.2 Observables

In statistical mechanics, we say that an observable is a function defined on the phase space of the system that varies smoothly with changes in the representative point of the system. This definition is different to the definition in quantum mechanics.

The kinetic energy  $\sum_i \frac{p_i^2}{2m_i}$  is an observable, while the function  $\Theta = \prod_i \theta(r_0^2 - r_i^2)$ , where  $\theta(x)$  is the Heaviside function, is not, since an arbitrarily small change in a position  $r_i$  can lead to a finite change in  $\Theta$ . This behaviour is not smooth which means  $\Theta$  is not an observable.

Other examples of (thermodynamic) observables include the internal energy  $E$  (i.e. the Hamiltonian) and the number of particles in the system ( $N = \int d\mathbf{r} \sum \delta(\mathbf{r} - \mathbf{r}_i)$ ). It turns out that the entropy  $S = S(X_0 = E = H, X_1, \dots, X_r)$  is also an observable. However, we don't yet have a way to express the entropy — the extensive variable of a thermodynamic system — as a function of the microscopic state of the system and its equations of motion. To do so, we are going to invoke the fundamental postulate of statistical mechanics.

### 4.3 The fundamental postulate of statistical mechanics

The fundamental postulate expresses the entropy  $S$  as a function of the accessible volume of the phase space of a dynamical system. I.e. that part of the phase space where the observables have values that are compatible with a specific thermodynamic state.

Assume that the thermodynamic state of a system is determined by the extensive variables  $X_1, \dots, X_r$  and that each of the  $X_i$  can be expressed as a function over the microscopic phase space; I.e.  $X_i = X_i(\mathbf{x})$ . The region of phase space that is accessible is defined by the value of  $X = (X_1, \dots, X_r)$ . Well, almost; we actually also allow for small discrepancies  $\epsilon_i$ ,  $i = 1, \dots, r$  to account for the effect of any non-extensive terms.

We denote the region of phase space in which the observable  $X_i$  have the specific values that match the thermodynamic state as  $\Gamma$  and we denote the corresponding phase space volume  $\text{vol}(\Gamma) = \Omega$ . Then the fundamental postulate of thermodynamics states that

$$S = k_B \ln(\Omega), \quad (6)$$

where  $S$  is the thermodynamic entropy,  $K_B$  is Boltzmann's constant ( $K_B = 1.384 \times 10^{-23} \text{J/K}$ ) and the equality should be interpreted as meaning “is equal, up to terms of size  $\epsilon_i$ ”. The accessible phase space volume  $\Omega$  measures the amount of the phase space available to the microscopic system while it is in this macroscopic state. If the system had only discrete micro-states, then we could simply count the number of micro-states that correspond to the same macro-state.

### 4.4 Symplectic transformations and phase space area preservation

Why is phase space so important in the formulation of entropy? To understand this we are going to look closer at some of the properties that being Hamiltonian impose on a system, including the choice of coordinate system.

We saw earlier that the equations of motion for a time-independent Hamiltonian system can be written as  $\frac{d\mathbf{x}}{dt} = f(\mathbf{x}) = J^{-1} \nabla H(\mathbf{x})$ . This Hamiltonian form is not preserved by arbitrary coordinate transformations; it *is* preserved by *symplectic* transformations.

**Definition:** A linear transformation  $A : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$  is said to be symplectic if it satisfies

$$A^\top J A = J.$$

For a non-linear transformation  $\mathcal{A}$ , we replace  $A$  above with its jacobian  $D\mathcal{A}$ .

An alternative, but equivalent, definition of a symplectic transformation — and one that will turn out to be useful in a bit — borrows some tools from differential geometry. (No prior experience with differential geometry is assumed.)

**Definitions:** A symplectic transformation is a diffeomorphism (a smooth, invertible mapping between manifolds)  $f : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$ , with coordinates  $q$  and  $p$ , which preserves the symplectic two-form  $\omega$  (an anti-symmetric, bilinear form) that acts on pairs of vectors and gives the sum of the oriented areas that they define. In two dimensions, this is the area of the parallelogram spanned by the two vectors.

More explicitly,

$$\omega(u, v) = u^\top J v = \sum_{i=1}^d u_{p_i} v_{q_i} - u_{q_i} v_{p_i}.$$

At this point it is necessary to introduce a (potentially) new concept — the flow map of a vector field. The flow  $\varphi_{t,f}(\mathbf{x})$  of a Hamiltonian system is a function that maps a point  $\mathbf{x}^*$  in the phase space of the system with equations of motion  $\dot{\mathbf{x}} = f(\mathbf{x}) = J^{-1} \nabla H(\mathbf{x})$  forward by time  $t$ . I.e. for any  $f(\mathbf{x})$  and  $\mathbf{x}_0$ ,  $\varphi_{t,f}(\mathbf{x}_0)$  gives the solution to  $\dot{\mathbf{x}} = f(\mathbf{x})$  at time  $t$ . That is, it shows the time evolution of  $\mathbf{x}$  and  $\mathbf{p}$  on the phase space.

**Theorem:** If the Hamiltonian  $H(\mathbf{x})$  is a twice differentiable function on some neighbourhood of  $\mathbb{R}^{2d}$ , then the time- $t$  flow map  $\varphi_{t,f}(\mathbf{x}_0)$  of the Hamiltonian vector field is a symplectic map for all  $t$ . **Proof:** We will show that  $\omega(u, v) = u^\top J v$  is constant along solutions of  $\dot{\mathbf{x}} = J^{-1} \nabla H(\mathbf{x})$ .

The tangent vector  $u(t) = \frac{\partial \varphi_t}{\partial \mathbf{x}_0}(\mathbf{x}_0) u(0)$  satisfies  $\frac{du}{dt} = J^{-1} \nabla^2 H(\mathbf{x}(t)) u(t)$  and the tangent vector  $v(t)$  satisfies a similar expression. Now we can calculate the time derivative of  $\omega(u, v)$ .

$$\begin{aligned} \frac{d}{dt} \omega(u, v) &= \dot{u}^\top J v + u^\top J \dot{v} \\ &= (J^{-1} \nabla^2 H u)^\top J v + u^\top J J^{-1} \nabla^2 H v \\ &= 0 \end{aligned}$$

since  $J^{-\top} = (J^{-1})^\top$ ,  $J^{-\top} J = -I$  and  $J J^{-1} = I$ .

I.e. as the system evolves according to its equations of motion, it always preserves the area for  $d = 1$ , or the sum of the oriented areas for  $d \geq 2$ .

## 4.5 Liouville's theorem and preservation of phase space volume

We are now in a position to prove that the flow of a Hamiltonian system preserves phase space volume. That is, for every bounded, open set  $A \subset \mathcal{R}^{2d}$  and every  $t$  for which  $\varphi_{t,f}(\mathbf{x})$  exists

$$\text{Vol}(\varphi_{t,f}(\Gamma)) = \text{Vol}(\Gamma)$$

where  $f = J^{-1} \nabla H$  and  $\text{Vol}(\Gamma) = \int_\Gamma d\mathbf{x}$ .

To prove this, we will actually prove a slightly more general case: the flow of a differential equation  $\dot{\mathbf{x}} = f(\mathbf{x})$  in  $\mathbb{R}^n$  is volume preserving if and only if it is divergence-free; i.e.  $\nabla \cdot f = 0$ .

But first, why do we care about preservation of phase space volume? The fundamental postulate of statistical mechanics,  $S = k_B \ln(\Omega)$  related the entropy of a system to the volume of the part of the phase space accessible to that system when it is in that particular thermodynamic state. We would like to be sure that this quantity doesn't vary too much over time as the microscopic system evolves, according to the Hamiltonian equations of motion. Although the microscopic system might be in continuous motion, the trajectories of the particles it is comprised of are such that the extensive variables which characterise the macroscopic state of the system are (almost) constant. I.e. there is no change in entropy, no heat flow, no work done while the system is at equilibrium. This is reassuring — it says that disorder doesn't increase unless something happens.

**Proof: volume preservation:** The Jacobian of the flow map  $\varphi_{t,f}(\mathbf{x}_0)$  defines the matrix function  $X(t) = D\varphi_{t,f}(\mathbf{x}_0)$ . The volume of the shape spanned by the columns of  $X$  is given by  $\det(X)$ , and  $X(t)$  is a solution of the equation

$$\dot{X} = A(t)X(t), \quad X(0) = I,$$

where  $A(t) := Df(\mathbf{x}(t))$  denotes the Jacobian of  $f(\mathbf{x}(t))$  evaluated at  $\mathbf{x}(t) = \varphi_{t,f}(\mathbf{x}_0)$ . To show that the volume defined by  $X$  is constant we use the Abel-Liouville-Jacobi-Ostrogradskii identity to find the time derivative of  $\det X$ :

$$\frac{d}{dt} \det X(t) = \text{trace} A(t) \cdot \det X(t).$$

Since  $A(t) = Df(\mathbf{x}(t))$  we have  $\text{trace} A(t) = \sum_i \frac{\partial f_i}{\partial \mathbf{x}_i}(\mathbf{x}(t)) = \nabla \cdot f(\mathbf{x}(t))$  and therefore  $\det X(t) = \det X(0) = 1$  if and only if  $\nabla \cdot f(\mathbf{x}) = 0$ .

## 4.6 Recommended reading

This section has mostly followed the first half of chapter 3 of *Statistical Mechanics in a Nutshell*. However, some of the significance of Hamiltonian systems is lost in SMiaN, so I've supplemented it with more substantive (mathematically oriented) details from *Geometric Numerical Integration* by E. Hairer, Ch. Lubich and G. Wanner (Springer, 2006) — one of the best books on (the numerics of) Hamiltonian dynamical systems and a personal favourite.

## 5 Molecular dynamics and simulating Hamiltonian Systems

Much of the point of statistical mechanics is to *avoid* dealing with the dynamics of microscopic systems by studying their aggregate (or extensive) properties in a statistical fashion. Given this, and the computational challenges of dealing with the dynamics of very large numbers of (probably interacting) particles, you might think that direct simulations don't have much of a role to play in statistical mechanics. But:

- It isn't always tractable (or it's just plain difficult) to find analytical expressions or sufficiently accurate approximations to all stat mech systems.
- Computing power is increasingly cheap and the size/complexity of the systems that we can study via simulation are always improving.
- Simulations let us directly test assumptions about what effects are important to a system by, for example, simulating a different interaction term or different geometries.
- There are many simulation approaches where clever heuristics allow us to deal with much larger systems than might be expected, while still getting the physics correct. (E.g. Monte Carlo methods — see Shaun's content.) In some cases physical laws, such as energy-conservation, or preservation of phase space volume can be “baked in” to the simulation, such that the simulation automatically preserves the appropriate behaviour.

We're going to look in more detail at this last point — how we might ensure that we can run a simulation that preserves physical laws. The context that we are going to look at is molecular dynamics.

As you might expect from the previous section, such systems can be described by specifying a Hamiltonian where the kinetic energy component takes the familiar form

$$T = \sum_i \frac{\mathbf{p}_i^2}{2m_i},$$

where the  $p_i$  are the momenta of the molecules and  $m_i$  are their masses.

Interactions between particles are described by  $U$ , the potential term of the Hamiltonian. About the simplest model would be to treat each particle as a hard sphere, with no interactions between them, except during collisions. An interaction term that is only slightly more complicated, but is much more realistic, is given by the Lennard-Jones potential:

$$U(\mathbf{q}) = 4\epsilon_{ij} \left( \left( \frac{\sigma_{ij}}{q_{ij}} \right)^{12} - 2 \left( \frac{\sigma_{ij}}{q_{ij}} \right)^6 \right),$$

where  $|q_{ij}| = |q_i - q_j|$  are the inter-atomic separations. It's not hard to see that the potential has an absolute minimum at  $q = \sigma$  — this corresponds to the equilibrium separation between atoms. (Omitting the factor of 2 in the potential means the location of the minimum picks up a factor of  $2^{-6}$ .) The  $\epsilon_{ij}$  terms gives the depth of the potential well — the strength of the inter-atomic attraction.

As a brief historical aside, the  $q^{-6}$  term in the potential can be derived mathematically as the attraction due to induced-dipole – induced dipole interaction. This was known before Lennard-Jones, whose contribution was to include a term for the repulsion when the inter-atomic distance gets very small (Pauli repulsion due to overlapping wave-functions for electrons that are closer

than the preferred bonding distance — but that is getting into condensed matter physics). The repulsion term is actually exponential, rather than  $q^{-12}$ , but Lennard-Jones decided that  $q^{-12}$  is pretty close to an exponential, and is much easier to calculate. Once you already know  $q^{-6}$  you can just square it and save yourself a (relatively) expensive calculation of an exponential.

It is worth noting that the pairwise interactions in the Lennard-Jones potential mean that for  $N$  particles we must calculate  $N^2$  interaction terms. Clearly, this can cause issues once  $N$  gets large. One common approach is to use a *long-range cut-off* and only calculate the interaction terms for atoms that are less than some distance  $L$  away, or perhaps, only for those inside a unit cell, with sides of length  $L$ .

Regardless of the exact nature of the interaction term, we end up dealing with a Hamiltonian of the form

$$H(\mathbf{q}, \mathbf{p}) = T(\mathbf{p}) + U(\mathbf{q}).$$

## 5.1 Numerical integration

Given the equations of motion (or the Hamiltonian) that describes a system, and some compatible initial conditions  $\mathbf{q}(0)$ , we want to find the state of the system at some time  $T$  in the future (or the solution along some interval  $[0, T]$ ). We do this by breaking the interval into small enough time steps of size  $h$  and discretizing the solution  $\mathbf{q}(t)$  as  $\mathbf{q}_n = \mathbf{q}(t_n)$  where  $t_n = nh$ .

The simplest approach to approximating the discrete solution is the *explicit Euler method*. It approximates the solution at  $\mathbf{q}(t_{n+1}) \simeq \mathbf{q}_{n+1}$  with

$$\mathbf{q}_{n+1} = \mathbf{q}_n + h\dot{\mathbf{q}}_n.$$

That is, it evaluates the differential equation at the current time step and approximates the solution with a straight line of slope  $\dot{\mathbf{q}}_n$ .

[You can add a diagram to illustrate this - even a hand drawn example.]

The error in the approximation of the solution by the explicit Euler method grows in proportion to the step size  $h$ ; that is,  $\|\mathbf{q}_N - \mathbf{q}(Nh)\| = \mathcal{O}(h)$ . We therefore say that the method is first order accurate.

While the explicit Euler method is very easy to implement, it has a number of issues (you'll see one of them in the assignment). These include the fact that it becomes unstable (the error in its approximation starts to grow exponentially) when the step size is too large. The *implicit Euler* (like the explicit Euler method, but the derivative is evaluated at the end of the time step i.e.  $f(q_{n+1})$ ) method avoids this issue, but at the cost of needing to iteratively solve an implicit equation  $q_{n+1} = q_n + hf(q_{n+1})$ . It is still only first order.

Since  $M\dot{\mathbf{q}} = \mathbf{p}$  and  $M\ddot{\mathbf{q}} = \dot{\mathbf{p}}$  the equations of motion for the system can be written as

$$M\ddot{\mathbf{q}} = f(\mathbf{q}) = -\nabla U(\mathbf{q}). \quad (7)$$

That is, we can turn the set of  $2d$  first-order differential equations in  $\mathbf{q}$  and  $\mathbf{p}$  into  $d$  second order differential equation in  $\mathbf{q}$ . In much of what follows, I'm going

to omit the mass matrix and write  $\ddot{\mathbf{q}} = f(\mathbf{q})$ . This is purely for convenience. Everything still holds when  $M$  takes other values for the masses — just make sure to include a factor of  $M^{-1}$  in the  $f(\mathbf{q})$  term.

A better approach to numerical integration is to use the Verlet (AKA Störmer, AKA leap-frog) method. This is most easily understood in its *two-step formulation* (i.e. the method requires information from two previously known time steps ( $t_{n-1}$  and  $t_n$ ) in order to compute the approximation at the future time step  $t_{n+1}$ ). The approximation is given by

$$\frac{\mathbf{q}_{n+1} - 2\mathbf{q}_n + \mathbf{q}_{n-1}}{h^2} = f(\mathbf{q}_n), \quad (8)$$

where  $f(\mathbf{q})$  is the second order differential equation in equation 7. It's not hard to see (you should verify it yourself) that the LHS of 8 is an approximation of the second derivative at the point  $\mathbf{q}_n$ .

[add a picture to illustrate this]

The two-step formulation of the leapfrog method can therefore be seen as determining a parabola that interpolates the points  $\mathbf{q}_{n-1}$ ,  $\mathbf{q}_n$ , and  $\mathbf{q}_{n+1}$  and matches the second derivative at the midpoint.

It turns out that, for a number of reasons, it is better to work with the Verlet method in its one-step formulation. This can be derived by introducing the velocity  $\mathbf{p} = \dot{\mathbf{q}}$  and turning equation 7 back into a first-order system of twice the dimension

$$\dot{\mathbf{q}} = \mathbf{p}, \quad \dot{\mathbf{p}} = f(\mathbf{q}).$$

We then introduce discrete approximations of  $q$  and  $p$  as follows:

$$p_n = \frac{q_{n+1} - q_{n-1}}{2h}, \quad p_{n-\frac{1}{2}} = \frac{q_n - q_{n-1}}{h}, \quad q_{n-\frac{1}{2}} = \frac{q_n + q_{n-1}}{2},$$

where some of the values have been calculated on a staggered grid. We can then use these to formulate a one-step version of the Verlet method.

$$\begin{aligned} \mathbf{p}_{n+\frac{1}{2}} &= \mathbf{p}_n + \frac{h}{2}f(\mathbf{q}_n), \\ \mathbf{q}_{n+1} &= \mathbf{q}_n + h\mathbf{p}_{n+\frac{1}{2}}, \\ \mathbf{p}_{n+1} &= \mathbf{p}_{n+\frac{1}{2}} + \frac{h}{2}f(\mathbf{q}_{n+1}). \end{aligned}$$

[Add a picture to illustrate this]

There is also a dual formulation of the method that gives the values of  $\mathbf{q}$  and  $\mathbf{p}$  on the staggered grid points  $t_{n-\frac{1}{2}}$ ,  $t_{n+\frac{1}{2}}$ , etc. (It makes sense to think of the one-step method as a discrete map  $\Phi_h$ , similar to the flow map  $\varphi_{f,t}$  that maps from the known solution, to its approximation one time step forward.)

If the velocity values are not needed at the end of each step, then the last line of one step can be combined with the first line of the next step to avoid one evaluation of  $f$  per step by only calculating  $\mathbf{p}$  on the staggered grid:

$$\begin{aligned} \mathbf{p}_{n+\frac{1}{2}} &= \mathbf{p}_{n-\frac{1}{2}} + hf(\mathbf{q}_n), \\ \mathbf{p}_{n+1} &= \mathbf{q}_n + h\mathbf{p}_{n+\frac{1}{2}}. \end{aligned}$$

This is the most computationally efficient form of the Verlet method. It is also numerically more stable than the two-step formulation.

It is not too hard to show (you should prove it for yourself) that the Verlet method is symplectic. That is, the discrete map that it defines preserves the symplectic form that we saw earlier. This is almost, but not quite the same as preserving the value of the associated Hamiltonian along solutions of the vector field. In fact, it is possible to show (but trickier than proving symplecticity) that the difference between the value of the Hamiltonian evaluated along the discrete approximation from the Verlet method differs only exponentially slowly from the exact value of the Hamiltonian for the true solution. This makes the Verlet method a popular choice for molecular dynamics. In addition to having good stability and being easy to implement, it is automatically almost energy-preserving.

The Verlet method is *symmetric* (replacing  $h$  with  $-h$  and swapping the subscripts  $n$  and  $n + 1$  gives the same method again) and *reversible* (changing the sign on the velocity for the initial condition is the same as running the method in reverse). It is also second order accurate; that is the error in the solution  $\|\mathbf{x}_{Nh} - \mathbf{x}(Nh)\| = \mathcal{O}(h^2)$ .

Now that we have a suitable method for calculating the micro-state of a system at some point in the future, we can look for ways to measure macroscopic quantities from the solution. It is worth bearing in mind that for any simulations, the initial conditions used are unlikely to correspond to a representative configuration for the system. It is therefore good practise to let the simulation run for a period of time (perhaps a few thousand time steps) before attempting to measure any properties. This process is known as equilibration or thermalization.

Using simulated particle trajectories usually means we implicitly invoke an assumption of *ergodicity*. This essentially says that averaging an observable along a solution, for a long enough time period for a system, is equivalent to averaging that observable over a large number of initial conditions for the same system, where the initial conditions correspond to the same macro-state of the system. While the hypothesis of ergodicity is useful, there are many real world cases where it doesn't hold — for example any system that exhibits *hysteresis* in which properties of the system are path dependent. Magnet systems, in particular, tend to display such properties. For more discussion of this, see section 4.0 of *Statistical Mechanics in a Nutshell*.

We have seen earlier that we identify the temperature of a thermodynamic system with the average square velocities of the particles in the system with some suitable normalisation. The temperature of our simulated system, with  $N$  particles is therefore given by

$$T(\mathbf{p}) = \frac{1}{Nk_B} \sum_{i=1}^N \frac{1}{2m_i} \|p_i\|^2.$$

Another easily computable quantity is the virial function (we haven't mentioned this yet, but it is used to relate the total potential energy of a system to



the total kinetic energy and can be used to calculate the pressure of a system — see pp 255–256 of *Statistical Mechanics in a Nutshell*):

$$C(\mathbf{r}) = \sum_{i=2}^N \sum_{j=1}^{i-1} \nabla U(r_{ij}) r_{ij},$$

Both of these quantities give instantaneous values of the observables. Other quantities can be calculated as time averages of a solution of the system. For example, the constant volume specific heat:

$$C_v = k_B \left[ \frac{2}{3N} - \frac{4}{9} \frac{\langle (T - \langle T \rangle)^2 \rangle}{\langle T \rangle^2} \right]^{-1}$$

## 5.2 Recommended reading

- Sections 8.1 and 8.2 of *Statistical Mechanics in a Nutshell*
- *Geometric numerical integration illustrated by the Störmer–Verlet method*, E. Hairer, Ch. Lubich, and G. Wanner, Acta Numerica (2003)
- Chapter 11 of *Simulating Hamiltonian Dynamics*, B. Leimkuhler and S. Reich, Cambridge University Press (2004)

## 6 Ensembles

In the previous section, we considered an isolated system where we could keep track of the dynamics of every particle and use that to calculate the values of extensive, macroscopic properties of the system. An important aspect of this was the conservation of energy for the system. We refer to such systems as a *micro-canonical ensemble* — the system is isolated with no heat flux and no change in the number of particles, hence the internal energy of the system is constant and it is described entirely by the Hamiltonian dynamics

We also saw (a couple of sections earlier) that we can equate the entropy of a system with the accessible volume of the phase space of that system. This was part of what motivated us to study the microscopic dynamics of the system using molecular dynamics. We’d now like to consider some slightly more realistic cases where, for example, we may want to allow for multiple systems in contact.

We’ll start by considering the simplest case: two (isolated) systems in contact, without any exchange of energy. If the state of system 1 corresponds to the region of phase space  $\Gamma^1$  and similarly, the state of system 2 to  $\Gamma^2$  then the state of the composite system  $1 \cup 2$  corresponds to the phase space regions given by the Cartesian product  $\Gamma^{1 \cup 2} = \Gamma^1 \times \Gamma^2$  and the volume of this accessible volume of phase space is given by  $|\Gamma^{1 \cup 2}| = |\Gamma^1| |\Gamma^2|$ . From this, it’s easy to see that the entropy of the composite system is given by

$$\begin{aligned} S &= k_B \ln |\Gamma^{1 \cup 2}| = k_B \ln (|\Gamma^1| |\Gamma^2|) \\ &= k_B \ln |\Gamma^1| + k_B \ln |\Gamma^2| = S^1 + S^2, \end{aligned}$$

which is fortunate, since entropy is an extensive variable and we therefore expect it to be additive. In this example the two sub-system were completely isolated from each other; the dynamics of one system had no influence on the dynamics of the other. This condition of dynamic independence corresponds to the independence of the observables that pertain to these sub-systems.

Now, let's look at how the entropy of a composite system changes if we allow for the exchange of energy between the two sub-systems. This has the effect of increasing the accessible volume of the phase space, since exchange of energy means that there are more possible configurations for the overall system.

Without energy exchange, the volume of the accessible phase space to the total system is given by

$$|\Gamma_0| = |\Gamma^1| |\Gamma^2|.$$

Once we allow for the exchange of energy, this becomes

$$|\Gamma| = \sum_{E^1} |\Gamma(E^1)| |\Gamma^2(E_{\text{tot}} - E^1)|.$$

That is, we must now consider all possible configurations where the total energy of the composite system is partitioned between the two sub-systems. This volume is bounded below by  $|\Gamma_0|$  since the expression for  $|\Gamma_0|$  is just one of the terms in the sum. At first glance, it may look like this increase in the volume of the accessible phase space is enormous (we have added many more possible configurations), however, the volume of the accessible phase space for the composite system corresponds almost entirely to the states where  $E^1 = E_{\text{eqm}}^1$  and  $E^2 = E_{\text{eqm}}^2$ . As a consequence, for large enough  $N$ , the difference between  $\Gamma$  and  $|\Gamma^1(E_{\text{eqm}}^1)| |\Gamma^2(E_{\text{eqm}}^2)|$  is negligible. It's not too hard to show (see section 3.7 of *Statistical Mechanics in a Nutshell*) that the contribution to the accessible phase space volume due to the exchange of energy between the two systems is of order  $\sqrt{N}$  compared with the total system size  $N$  — small enough to be negligible when  $N$  is large.

## The canonical ensemble

Rather than considering a perfectly isolated system, for which energy is conserved, a more realistic experimental situation may be to consider a system  $S$  in thermal contact with some much larger reservoir  $R$ . This has the effect of holding the total system at constant temperature. In such a situation we want to be able to calculate the average value  $\langle A \rangle$  of an observable  $A$  for the system  $S$ ; we are not interested in the state of the reservoir  $R$ , except to the extent that it helps us determine the state of  $S$ .

If, for the composite system  $S \cup R$ , we write  $x_S, x_R$  for the state, then the average value of the observable  $A$  is given by

$$\langle A \rangle = \frac{1}{|\Gamma|} \int_{\Gamma} dx_S dx_R A(x_s),$$

where  $\Gamma$  is the region of the phase space for the composite system, when it has total internal energy of  $E$ .

In order to make explicit the parts of the total phase space that is accessible to the composite system, we write the above expression as the Cartesian product of the phase space for the system and the reservoir, i.e.

$$\langle A \rangle = \frac{1}{|\Gamma|} \int dx_S A(x_S) \times \int dx_R \delta(H^R(x_R) - (E - H^S(x_S))).$$

The delta function in the last term is zero, except when  $x_S$  and  $x_R$  in the two sub-systems take values such that  $H^S + H^R = E$ . That is the delta function defines the accessible phase space volume when the two sub-systems can exchange internal energy between them, but the total internal energy of the composite system is conserved.

Recalling the fundamental postulate of statistical mechanics ( $S = k_B \ln |\Gamma|$ ), we rewrite the last expression to replace the phase space volume with the corresponding expression for entropy:

$$\int dx_R \delta(H^R(x_R) - (E - H^S(x_S))) \simeq \exp\left(\frac{1}{k_B} S^R(E - H^S)\right).$$

Since  $H^S$  is much less than  $E$  it makes sense to now expand the exponential in a Taylor series about  $E$ :

$$\exp\left(\frac{1}{k_B} S^R(E - H^S)\right) \simeq \exp\left(\frac{1}{k_B} S^R(E)\right) \exp\left(\frac{-1}{k_B} \frac{\partial S^R}{\partial E} \Big|_E H^S(x_S)\right) \dots$$

Earlier in the course, we identified  $\frac{\partial S}{\partial E}$  with  $\frac{1}{T}$ , and, for a canonical ensemble,  $T$  is constant (that's the point of the reservoir) so we write the expected value of the observable as

$$\langle A \rangle \simeq \frac{1}{\Gamma} \int dx_S A(x_S) \times \exp\left(\frac{1}{k_B} S^R(E)\right) \exp\left(\frac{-H^S(x_S)}{k_B T}\right).$$

It remains to make the normalisation precise. When we do this, the factors of  $s \exp\left(\frac{1}{k_B} S^R(E)\right)$  cancel from the integral and its normalisation and we get

$$\langle A \rangle = \frac{1}{Z} \int dx_S A(x_S) \exp\left(\frac{-H^S(x_S)}{k_B T}\right), \quad (9)$$

where  $Z$  is known as the *partition function* and is given by

$$Z = \int dx_S \exp\left(\frac{-H^S(x_S)}{k_B T}\right). \quad (10)$$

One way to think about equations 9 and 10 is that we no longer need to keep track of which parts of phase space are accessible. Instead we can integrate over the entire phase space and each region is weighted appropriately, according to  $\exp\left(\frac{-H}{k_B T}\right)$  — the so-called *Boltzmann factor*, a probability density in the phase space.

Similar to the case of the micro-canonical ensemble, for the canonical ensemble, the contributions to  $A$  are dominated by the part of phase space that corresponds to when the system's internal energy is at the equilibrium value.

You should read through, and understand, section 3.12 of *Statistical Mechanics in a Nutshell* and sections 4.1–4.3 of *Statistical Mechanics Made Simple* for some extra details and explanation.

## Other ensembles

We can follow a similar approach to what we did for the canonical ensemble and generalise to other types of ensembles. If  $f_i$  is some variable that we want to hold fixed and  $X_i$  is the conjugate extensive variable, then putting the system in contact with a reservoir with which it can exchange  $X_i$  gives

$$\langle A \rangle = \frac{1}{Z} \int dx_S A(x_S) \exp \left( \frac{f_i X_i^S(x_S)}{k_B T} \right),$$

where we have used the relation  $\left. \frac{\partial S^R}{\partial X_i^R} \right|_E = -\frac{f_i}{T}$ .

Calculating the partition function proceeds similarly too, (see SMiaN, section 3.13) and gives

$$Z \simeq \exp \left( \frac{T^S(X_i^*) + f_i X_i^*}{k_B T} \right),$$

where  $X_i^*$  is the value of  $X_i$  which maximises the value of the exponential.

The  $p$ – $T$  ensemble is one specific example of a generalised ensemble. In this case the pressure and temperature are fixed, while the internal energy and volume (their conjugate variables) are allowed to fluctuate. Using the generalised formula above, and dropping the subscripts that we were previously using to denote the system, the  $p$ – $T$  ensemble is given by

$$\langle A \rangle = \frac{1}{Z} \int dx A(x) \exp \left( -\frac{E(x) - pV(x)}{k_B T} \right),$$

while the partition function is given by

$$\ln Z = -\frac{E - TS + pV}{k_B T}.$$

The quantity on the top of the fraction is the *Gibbs free energy*. (See section 3.14 of SMiaN)

Another important ensemble is the so-called *grand ensemble* in which the number of particles in the system is also allowed to fluctuate due to, for example, chemical reactions, while the *chemical potential*  $\mu$  is held fixed. In this case the expression for the expected value of an observable is

$$\langle A \rangle = \frac{1}{Z} \sum_{N=1}^{\infty} \int dx A(x) \exp \left( -\frac{H_N(x) - \mu N}{k_B T} \right),$$

while the corresponding partition function is given by

$$\ln Z = -\frac{E - TS - \mu N}{k_B T}.$$

(See section 3.15 of SMiaN and section 4.4 of SMMS for more discussion of the grand canonical ensemble.)

## Information theory and the Gibbs Formula for entropy

To finish off this section, we'll look at one last application of entropy; not because it is particularly useful to physical systems in statistical mechanics, but because it gives a result that forms the basis of information theory.

We'll return to a considering a generalised ensemble, like the one we looked at a couple of sections earlier. However, in this case we'll assume that the phase space is discretized and that the index  $n$  runs over all of the microstates of the system. If we consider an intensive variable  $f$  and its corresponding extensive variable  $X$  then the expression for the expected value of any observable  $A$  is

$$\langle A \rangle = \frac{1}{Z} \sum_n A_n \exp\left(\frac{fX_n}{k_B T}\right),$$

while the partition function  $Z$  is given by

$$Z = \sum_n \exp\left(\frac{fX_n}{k_B T}\right).$$

The partition function is related to the thermodynamic potentials (see the previous section on generalised ensembles and SMiaN sections 3.12 and 3.13) via

$$\ln Z = \frac{TS + f\langle X \rangle}{k_B T}. \quad (11)$$

Now, for any individual microstate  $n$  the probability is therefore given by

$$p_n = \frac{1}{Z} \exp\left(\frac{fX_n}{k_B T}\right).$$

Taking the log of both sides of this expression gives

$$\ln p_n = \frac{fX_n}{k_B T} - \ln Z,$$

which after substituting 11 for  $\ln Z$  gives

$$\ln p_n = \frac{1}{k_B T}(fX_n - TS - f\langle X \rangle).$$

Now we can calculate the expected value of both side of the above equation

$$\langle \ln p_n \rangle = \sum_n p_n \ln p_n = \frac{1}{k_B T}(f\langle X \rangle - TS - \langle X \rangle) = \frac{-S}{k_B}.$$

After rearranging this for  $S$  (and making explicit the sum for the expected value of  $p_n$ ) we arrive at the *Gibbs formula for entropy*:

$$S = -k_B \sum_n p_n \ln p_n.$$

Although this is elegant, it's generally useless in the context of physical systems since the number of microstates it would be necessary to sum over is far too large to be practical and, in any case, we generally don't know the probability distribution  $p_n$  for each of the microstates. The value of this expression lies in its application to other systems, particularly information theory, where it can be used to quantify amount of information in, for example a digital signal, in which case the  $p_n$  represent the probability of receiving the  $n$ -th possible value in the list of signals.

## Recommended reading

Most of the notes in this section follow closely the second half of chapter 3 in *Statistical Mechanics in a Nutshell*; specifically, section 3.6–3.18. Chapter four of *Statistical Mechanics made Simple* covers the same material in sections 4.0–4.4. It gives some intuitive and succinct explanations, but I find it to be of more use *after* you've already looked at SMiaN. Also useful, and with a slightly different presentation (perhaps with more traditional notation), is *Entropy, Order Parameters, and Complexity*. Here the content is spread around a bit over chapters three through six. Much of the useful content, including some relevant to early sections of these notes, is in sections 3.1, 3.5, 6.1, 6.2, 6.3, and 5.3.