Statistical Mechanics

University of Cambridge Part II Natural Sciences Tripos

Yue Wu

Yusuf Hamied Department of Chemistry
Lensfield Road,
Cambridge, CB2 1EW

yw628@cam.ac.uk

Acknowledgements

Contents

1 Introduction

"Anyone who wants to analyze the properties of matter in a real problem might want to start by writing down the fundamental equations and then try to solve them mathematically. Although there are people who try to use such an approach, these people are the failures in this field..."

Richard Feynman

Suppose you know all the fundamental laws of physics, if I give you a box containing 10^{23} particles and tell you their mass, their charge, their interactions, and so on, what can you tell me about the stuff in the box?

Theres one strategy that definitely won't work: writing down the Schrodinger equation for 10^{23} particles and solving it. That's typically not trivial to do for 2 or 3 particles, and generally possible for 23 particles, let alone 10^{23}. What's more, even if you could find the wavefunction of the system, what would you do with it? The positions of individual particles are of little interest to anyone. We want answers to much more basic, almost childish, questions about the contents of the box. Is it wet? Is it hot? What colour is it? Is the box in danger of exploding? What happens if we squeeze it, pull it, heat it up? How can we begin to answer these kind of questions starting from the fundamental laws of physics?

The purpose of this course is to introduce the dictionary that allows you translate from the microscopic world where the laws of Nature are written to the everyday macroscopic world that were familiar with.

A large part of this course will be devoted to figuring out the interesting things that happen when you throw 10^{23} particles together. One of the recurring themes will be that

$$10^{23} \neq 1 \,. \tag{1.1}$$

More is different[1]: there are key concepts that are not visible in the underlying laws of physics but emerge only when we consider a large collection of particles. One very simple example is temperature. This is not a fundamental concept: it doesn't make sense to talk about the temperature of a single electron. But it would be impossible to talk about physics of the everyday world around us without mention of temperature. This illustrates the fact that the language needed to describe physics on one scale is very different from that needed on other scales. We'll see several similar emergent quantities in this course, including the phenomenon of phase transitions where the smooth continuous laws of physics conspire to give abrupt, discontinuous changes in the structure of matter

Statistical mechanics is the art of turning the microscopic laws of physics into a description of Nature on a macroscopic scale.[2]

---

[1] Anderson, P. W. (1972). More Is Different: Broken symmetry and the nature of the hierarchical structure of science. Science, 177(4047), 393-396.

[2] This well-written introduction is from Prof. David Tong.

2 Fundamentals of Statistical Mechanics

2.1 The Microcanonical Ensemble

We will start by stating the most fundamental assumption in statistical mechanics. It is the idea that we should take the most simple minded approach possible and treat all states the same — since we know nothing else about the system, such a democratic approach seems eminently reasonable.

Postulate (Principle of equal *a priori* probabilities). A system with fixed N, V and E is likely to be found in any of its $\Omega(E)$ energy eigenstates.

This is not provable. Some justifications of this assumption includes Boltzmann's H-theorem and Liouville theorem.

Then probability that the system with fixed energy E is in a given state n is simply

$$p(n) = \frac{1}{\Omega(E)} \,. \tag{2.1}$$

The probability that the system is in a state with some different energy $E' \neq E$ is zero. This probability distribution, relevant for systems with fixed energy, is known as the *microcanonical ensemble.*

2.2 Entropy and the Second Law

Consider a system with total energy E that consists of two weakly interacting subsystems, which means that they can only exchange energy, as shown in figure 2.1. The energy is distributed over the two subsystems such that

$$E_1 + E_2 = E \,. \tag{2.2}$$

For a given E_1, the total number of degenerate states of the system is $\Omega_1(E_1) \times \Omega_2(E_2)$. We choose take the logarithm of degeneracy so that the value is extensive. This gives

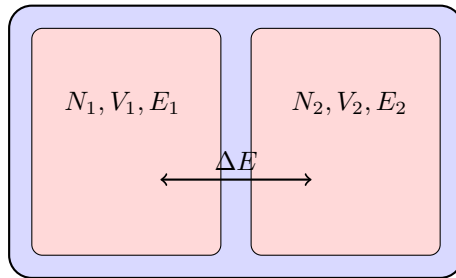$$\ln \Omega(E_1, E - E_1) = \ln \Omega_1(E_1) + \ln \Omega_2(E_2) \,. \tag{2.3}$$

Figure 2.1: Two weakly coupled systems that can exchange energy such that the total energy of the universe is conserved to be E.

By the principle of equal *a priori*, the most likely value of E_1 is the one that maximises $\ln \Omega(E_1, E - E_1)$, i.e.

$$\left(\frac{\partial \ln \Omega(E_1, E - E_1)}{\partial E_1}\right)_{N,V,E} = 0 . \tag{2.4}$$

Using the fact that $\mathrm{d}E_1 = \mathrm{d}E_2$, we obtain the condition

$$\left(\frac{\partial \ln \Omega_1(E_1)}{\partial E_1}\right)_{N_1,V_1} = \left(\frac{\partial \ln \Omega_2(E_2)}{\partial E_2}\right)_{N_2,V_2} . \tag{2.5}$$

We define

$$\beta(E, V, N) := \left(\frac{\partial \ln \Omega(E, V, N)}{\partial E}\right)_{N,V} , \tag{2.6}$$

the the equilibrium condition becomes

$$\beta(E_1, V_1, N_1) = \beta(E_2, V_2, N_2) . \tag{2.7}$$

2.2.1 Boltzmann Entropy

The thermal equilibrium condition of equal β must be equivalent to the condition that the two subsystems having the same temperature (from classical Thermodynamics). Also, just like the thermodynamic entropy S, $\ln \Omega$ is a state function of E, V and N. Moreover, when the thermal equilibrium is reached, $\ln \Omega$ of the total system is maximised, just as S. This suggests that $\ln \Omega$ is a monotonically increasing function of entropy S. Since both S and $\ln \Omega$ are extensive, S is proportional to $\ln \Omega$

$$S(N, V, E) = k_B \ln \Omega(N, V, E) \tag{2.8}$$

for some constant k_B known as the *Boltzmann's constant.*

In the statistical picture, the second law of thermodynamics states that, at thermal equilibrium, the system is most likely to be found in the state that has the largest number of degenerate energy states.

2.2.2 Temperature

The thermodynamic definition of temperature is

$$\frac{1}{T} := \left(\frac{\partial S}{\partial E}\right)_{V,N} , \tag{2.9}$$

which is obtained from the fundamental equation for the internal energy

$$\mathrm{d}E = T\mathrm{d}S - P\mathrm{d}V + \mu\mathrm{d}N . \tag{2.10}$$

Hence β and temperature T is related by

$$\beta = \frac{1}{k_B T} . \tag{2.11}$$

2.2.3 The Third Law

The third law of thermodynamics states that at $T = 0$, the number of accessible states of a pure substance is equal to one (i.e. $\Omega = 1$, so that $\ln \Omega = 0$). In other words: at absolute zero, the system is in its ground state — and this ground state is non-degenerate.

2.3 The Canonical Ensemble

Now let's consider a system that is at thermal equilibrium with a large heat bath. The total system is isolated so that $E_{\text{tot}} = E_{\text{bath}} + E_{\text{sys}}$ is fixed. The system is weakly coupled to the bath so they can exchange energy. This configuration is commonly called the *canonical ensemble*. Suppose the system is prepared in a specific state $|i\rangle$ with energy E_i, then the bath has an energy $E_{\text{bath}} = E_{\text{tot}} - E_i$ with degeneracy $\Omega_{\text{bath}}(E_{\text{tot}} - E_i)$. The degeneracy of the whole system is $\Omega = \Omega_{\text{sys}}\Omega_{\text{bath}} = \Omega_{\text{bath}}$. The principle of equal *a priori* applies only to an isolated system, so all the state of the universe with total energy E_{tot} are equally likely to occur. Hence, the probability of finding the system in state $|i\rangle$ is given by

$$P_i = \frac{\Omega_{\text{bath}}(E_{\text{tot}} - E_i)}{\sum_{|j\rangle} \Omega_{\text{bath}}(E_{\text{tot}} - E_j)} . \tag{2.12}$$
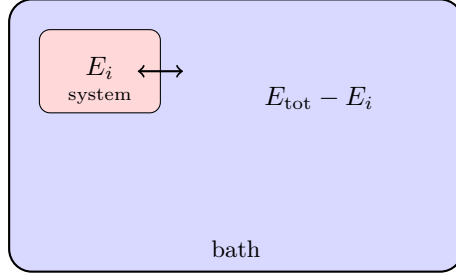
Figure 2.2: The system can exchange energy with the surrounding bath, such that the overall universe has fixed total energy E_{tot}.

Since the bath is very much larger than the system of interest, most of the energy if the universe will be in the bath. To compute $\Omega_{\text{bath}}(E_{\text{tot}} - E_i)$, we expand $\ln \Omega_{\text{bath}}(E_{\text{tot}} - E_i)$ about $E_{\text{bath}} = E_{\text{tot}}$

$$\ln \Omega_{\text{bath}}(E_{\text{tot}} - E_i) = \ln \Omega_{\text{bath}}(E_{\text{tot}}) - E_i \left(\frac{\partial \ln \Omega_{\text{bath}}(E)}{\partial E} \right)_{E=E_{\text{tot}}} + \mathcal{O}\left(\frac{1}{E}\right) , \tag{2.13}$$

or,

$$\ln \Omega_{\text{bath}}(E_{\text{tot}} - E_i) = \ln \Omega_{\text{bath}}(E_{\text{tot}}) - \beta E_i + \mathcal{O}\left(\frac{1}{E}\right) . \tag{2.14}$$

Hence,

$$P_i = \frac{\exp(-\beta E_i)}{\sum_{|j\rangle} \exp(-\beta E_j)} = \frac{\exp(-\beta E_i)}{Q} . \tag{2.15}$$

This is the *canonical distribution*, or *Boltzmann Distribution*. It describes the probability of the whole system being found in a particular microstate. The Boltzmann distribution should not be confused with the Maxwell–Boltzmann distribution (the distribution of the particle speeds in ideal gases) or Maxwell–Boltzmann statistics (the distribution of classical particles over energy states in thermal equilibrium).

In equation (2.13), we have ignored higher order terms in the Taylor expansion and claimed that it is $\mathcal{O}(1/E)$. This is straightforward to prove. The next term in the expansion is

$$\frac{E_i^2}{2}\left(\frac{\partial^2 \ln \Omega_{\text{bath}}(E)}{\partial E^2}\right)_{E=E_{\text{tot}}} = \frac{E_i^2}{2}\left(\frac{\partial \beta}{\partial E}\right)_{E=E_{\text{tot}}} = -\frac{E_i^2}{2k_B T^2 C_{V,\text{bath}}}\,, \qquad (2.16)$$

where we have used

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{N,V} = \left(\frac{\partial E}{\partial \beta}\right)_{N,V}\left(\frac{\partial \beta}{\partial T}\right)_{N,V}. \qquad (2.17)$$

Since both the heat capacity and the total energy are extensive, we can conclude that this term is $\mathcal{O}(1/E)$ as claimed.

2.3.1 Link to Thermodynamics

Knowing the energy distribution (2.15) allows us to compute the average energy of the system at a given temperature

$$\langle E_{\text{sys}}\rangle = \sum_{|i\rangle} E_i P_i = \frac{\sum_{|i\rangle}\exp(-\beta E_i)}{Q}\,. \qquad (2.18)$$

Next we note that the numerator looks very similar to the definition of Q. If we differentiate Q with respect to β, we obtain the negative of the numerator, and so we can write

$$\langle E_{\text{sys}}\rangle = -\frac{1}{Q}\left(\frac{\partial Q}{\partial \beta}\right)_{N,V} = -\left(\frac{\partial \ln Q}{\partial \beta}\right)_{N,V}. \qquad (2.19)$$

We can change the differential variable to T instead

$$\langle E_{\text{sys}}\rangle = -\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V}\frac{\partial T}{\partial \beta} = k_B T^2\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V}. \qquad (2.20)$$

We can compare this to the thermodynamic relation

$$E = -T^2\left(\frac{\partial (A/T)}{\partial T}\right)_{N,V}, \qquad (2.21)$$

which can be easily derived by differentiating A/T, we can see that the Helmholtz free energy A is related to the partition function through the bridge relation

$$A = -k_B T \ln Q\,. \qquad (2.22)$$

2.3.2 Fluctuations

Back to our canonical ensemble — a system of N particles with volume V in the large thermal bath. The probability of finding the system in any of the $\Omega(N, V, E)$ states with energy E is

$$P(E) = \frac{\Omega(N, V, E) \exp(-\beta E)}{\sum_{|i\rangle} \exp(-\beta E_i)} \,. \tag{2.23}$$

Using the definition of entropy, we may rewrite this at

$$P(E) \sim \exp(-\beta E) \exp\left[\frac{S(N, V, E)}{k_B}\right] \,. \tag{2.24}$$

The most likely energy of the system, E^*, is the one for which

$$\left.\frac{\partial P(E)}{\partial E}\right|_{E=E^*} = 0 \quad \Rightarrow \quad \left(\frac{\partial S}{\partial E}\right)_{E=E^*} = \frac{1}{T} \,. \tag{2.25}$$

We can expand S in a Taylor series about E^* and define $\Delta E := E - E^*$, we obtain

$$\frac{S(N, V, E)}{k_B} = \frac{S(N, V, E^*)}{k_B} + \frac{1}{k_B T}\Delta E + \frac{1}{2k_B}\left(\frac{\partial^2 S}{\partial E^2}\right)(\Delta E)^2 + \mathcal{O}((\Delta E)^3) \,. \tag{2.26}$$

Take logarithm of expression (2.24) and use the expansion (2.26), we can see that

$$\ln P(E^* + \Delta E) = c + \frac{1}{2k_B}\left(\frac{\partial^2 S}{\partial E^2}\right)(\Delta E)^2 + \mathcal{O}((\Delta E)^3) \,, \tag{2.27}$$

where c is a constant independent of ΔE. The term linear in ΔE vanishes. In the limit of large N, we can ignore terms of order $(\Delta E)^3$ or above and find that

$$P(E + \Delta E) = c' \exp\left[\frac{1}{2k_B}\left(\frac{\partial^2 S}{\partial E^2}\right)(\Delta E)^2\right] \,, \tag{2.28}$$

where c' is another constant. We now recall the standard thermodynamics relation

$$\left(\frac{\partial^2 S}{\partial E^2}\right) = \left(\frac{\partial(1/T)}{\partial E}\right)_V = -\frac{1}{T^2}\left(\frac{\partial T}{\partial E}\right)_V = -\frac{1}{T^2}\left(\frac{\partial E}{\partial T}\right)_V^{-1} = -\frac{1}{C_V T^2} \,. \tag{2.29}$$

Finally, we normalise the probability density to obtain

$$P(E^* + \Delta E) = (2\pi k_B T^2 C_V)^{-1/2} \exp\left[-\frac{(\Delta E)^2}{2k_B T^2 C_V}\right] \,, \tag{2.30}$$

or

$$P(E) = (2\pi k_B T^2 C_V)^{-1/2} \exp\left[-\frac{(E - E^*)^2}{2k_B T^2 C_V}\right] \,. \tag{2.31}$$

This is a Gaussian distribution. We can see that the mean squared fluctuation in the energy of a system at constant N, V and T is directly related to the heat capacity C_V,

$$\sigma_E^2 = \left\langle(\Delta E)^2\right\rangle = k_B T^2 C_V \,. \tag{2.32}$$

In fact, whilst the result that the probability distribution function is a Gaussian is important, it is possible to compute the variance of the distribution in a much simpler way if we are only interested in it. The thermal average of the energy is

$$\langle E \rangle = \frac{\sum_{|i\rangle} E_i e^{-\beta E_i}}{Q} = -\frac{1}{Q}\frac{\partial Q}{\partial \beta}\,, \tag{2.33}$$

and hence the derivative of $\langle E \rangle$ with respect to β is

$$\frac{\partial \langle E \rangle}{\partial \beta} = -\frac{\sum_{|i\rangle} E_i^2 e^{-\beta E_i}}{Q} + \frac{\sum_{|i\rangle} E_i e^{-\beta E_i}}{Q}\frac{1}{Q}\frac{\partial Q}{\partial \beta}\,. \tag{2.34}$$

This can be simplified to

$$\frac{\partial \langle E \rangle}{\partial \beta} = -\left\langle E^2 \right\rangle + \langle E \rangle^2 \equiv -\sigma_E^2\,. \tag{2.35}$$

We can related this to the heat capacity at constant volume

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{N,V} = \left(\frac{\partial E}{\partial \beta}\right)_{N,V}\frac{\partial \beta}{\partial T} = -\frac{1}{k_B T^2}\left(\frac{\partial E}{\partial \beta}\right)_{N,V}\,. \tag{2.36}$$

Therefore $\sigma_E^2 = k_B T^2 C_V$ as before.

Since both E and C_V are extensive, we can write $\langle E \rangle = N \langle \epsilon \rangle$ and $C_V = N c_V$, where ϵ and c_V are the energy and heat capacity per particle. The relative variance in the energy vanish in the thermodynamic limit as $N \to \infty$, since

$$\frac{\langle (\Delta E)^2 \rangle}{E^2} = k_B T^2 \frac{C_V}{\langle E \rangle^2} = \frac{k_B T^2 c_V}{N \langle \epsilon \rangle^2} \sim \frac{1}{N}\,. \tag{2.37}$$

Similar results apply to other quantities of interest. Fluctuations about the average are negligible in the thermodynamic limit, and the statistical mechanical average of a thermodynamic property will correspond to its measured value.

2.3.3 Helmholtz Energy and Equilibrium

By the second law of thermodynamics, in an isolated system at equilibrium, the entropy of a system is maximised. In the canonical ensemble, a system is in contact with a thermal reservoir. Although heat can be transferred between them, the total energy is conserved. The system plus the thermal bath is isolated, so we may apply the second law of thermodynamics. The total degeneracy is

$$\Omega_{\text{tot}} = \Omega_{\text{sys}}(E_{\text{sys}})\Omega_{\text{bath}}(E_{\text{tot}} - E_{\text{sys}})\,. \tag{2.38}$$

We can Taylor expand $\ln \Omega_{\text{bath}}(E_{\text{tot}} - E_{\text{sys}})$ up to linear order in E_{sys} and find that

$$\ln \Omega_{\text{tot}} = \ln \Omega_{\text{sys}}(E_{\text{sys}}) + \ln \Omega_{\text{bath}}(E_{\text{tot}}) - \beta E_{\text{sys}}\,. \tag{2.39}$$

The equilibrium condition is that the derivative of $\ln \Omega_{\text{tot}}$ with respect to E_{sys} vanishes. Since $\ln \Omega_{\text{sys}}(E_{\text{tot}})$ is independent of E_{sys}, the equilibrium condition is equivalent to minimising $\beta E_{\text{sys}} - \ln \Omega_{\text{sys}}(E_{\text{sys}}) = \beta(E_{\text{sys}} - TS_{\text{sys}})$. The Helmholtz energy A is defined as

$$A := E - TS\,. \tag{2.40}$$

Therefore, at constant temperature and volume, A is at a minimum at equilibrium.

Figure 2.3: A system coupled to the bath such that a piston separating them can change position. The system and the bath can exchange volume.

2.4 Pressure

Now consider a system that can exchange volume with a reservoir such that the total volume of the system plus the bath is fixed. The condition for equilibrium is that the total entropy is maximised, so we need to determine the maximum of

$$\ln \Omega(V_{\text{sys}}, V_{\text{tot}} - V_{\text{sys}}) \ln \Omega_{\text{sys}}(V_{\text{sys}}) + \ln \Omega_{\text{bath}}(V_{\text{tot}} - V_{\text{sys}}) \tag{2.41}$$

with respect to V_{sys}. Using the definition of entropy,

$$\left(\frac{\partial S_{\text{sys}}}{\partial V_{\text{sys}}}\right)_{E,N} + \left(\frac{\partial S_{\text{bath}}}{\partial V_{\text{sys}}}\right)_{E,N} = 0 \,. \tag{2.42}$$

Since $\mathrm{d}V_{\text{sys}} = -\mathrm{d}V_{\text{bath}}$,

$$\left(\frac{\partial S_{\text{sys}}}{\partial V_{\text{sys}}}\right)_{E,N} = \left(\frac{\partial S_{\text{bath}}}{\partial V_{\text{bath}}}\right)_{E,N} \,. \tag{2.43}$$

From thermodynamics, we have

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}E + \frac{P}{T}\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N \tag{2.44}$$

and hence $(\partial S/\partial V)_{E,N} = P/T$. Thus the equilibrium condition becomes

$$\frac{P_{\text{sys}}}{T_{\text{sys}}} = \frac{P_{\text{bath}}}{T_{\text{bath}}} \,. \tag{2.45}$$

If the system can exchange both energy and volume with the bath, then the conditions for equilibrium are $T_{\text{sys}} = T_{\text{bath}}$ and $P_{\text{sys}} = P_{\text{bath}}$.

In practice, it is convenient to use the fundamental relation for the Helmholtz energy, $\mathrm{d}A = -P\mathrm{d}V - S\mathrm{d}T + \mu\mathrm{d}N$, and the corresponding expression for the pressure

$$P = -\left(\frac{\partial A}{\partial V}\right)_{T,N} = k_B T \left(\frac{\partial \ln Q}{\partial V}\right)_{T,N} \,, \tag{2.46}$$

where we have used the bridge relation $A = -k_B T \ln Q$. Later, we shall use the relation to obtain a general expression of the pressure in a system of interacting particles. However, we will need to first derive some further results in classical statistical mechanics.

2.5 Classical Statistical Mechanics

The thermal average of the expectation value of some observable X is

$$\langle X \rangle = \frac{\sum_{|i\rangle} \exp(-E_i/k_B T) \langle i|X|i \rangle}{\sum_{|i\rangle} \exp(-E_i/k_B T)} \, . \tag{2.47}$$

This suggests that in order to compute the thermal average, we first need to solve the Schrödinger equation of the many-body system of interest, then we need to compute the expectation value of the operator X for all those quantum states that have non-negligible statistical weight. This approach is doomed for all but the simplest system. Fortunately, this can be simplified to a more workable expression in the classical limit $h \to 0$. For a d-dimensional, N-particle system, this is given by

$$\langle X \rangle = \frac{\int \prod_{i=1}^{N} \mathrm{d}^d \mathbf{p}_i \, \mathrm{d}^d \mathbf{r}_i \, \exp\left(-\beta \left[\sum_i p_i^2/2m_i + U(\{\mathbf{r}_i\})\right]\right) X(\{\mathbf{p}_i\}, \{\mathbf{r}_i\})}{\int \prod_{i=1}^{N} \mathrm{d}^d \mathbf{p}_i \, \mathrm{d}^d \mathbf{r}_i \, \exp\left(-\beta \left[\sum_i p_i^2/2m_i + U(\{\mathbf{r}_i\})\right]\right)} \, , \tag{2.48}$$

where $\mathbf{r}_i$ and $\mathbf{p}_i$ are the coordinate and momentum of the ith particle. The classical partition function is

$$Q = \frac{1}{h^{dN} N!} \int \prod_{i=1}^{N} \mathrm{d}^3 \mathbf{p}_i \, \mathrm{d}^3 \mathbf{r}_i \, \exp\left(-\beta \left[\sum_i^N \frac{p_i^2}{2m_i} + U(\{\mathbf{r}_i\})\right]\right) \, , \tag{2.49}$$

where the $1/N!$ factor is to account for the indistinguishability of the particles[3] and $1/h^{dN}$ is to make the partition function dimensionless.

2.5.1 Integration over the Momenta

From now on assume the space is three dimensional. Note that the integration over momenta and spatial coordinates can be separated

$$Q(E, V, T) = \frac{1}{h^{3N} N!} \int \prod_{i=1}^{N} \mathrm{d}\mathbf{p}_i \, \exp\left(-\beta \sum_{i=1}^{N} \frac{p_i^2}{2m}\right) \int \prod_{i=1}^{N} \mathrm{d}\mathbf{r}_i \, \exp(-\beta U(\{\mathbf{r}_i\})) \, . \tag{2.50}$$

Let's focus on the first integral. We can resolve p_i into its cartesian components $p_i^2 = p_{ix}^2 + p_{iy}^2 + p_{iz}^2$, so it can be further factorised as

$$\int \prod_{i=1}^{N} \mathrm{d}\mathbf{p}_i \, \exp\left(-\beta \sum_{i=1}^{N} \frac{p_i^2}{2m}\right)$$
$$= \int \prod_{i=1}^{N} \mathrm{d}p_{ix} \, \mathrm{d}p_{iy} \, \mathrm{d}p_{iz} \, \exp\left(-\frac{\beta p_{ix}^2}{2m_i}\right) \exp\left(-\frac{\beta p_{iy}^2}{2m_i}\right) \exp\left(-\frac{\beta p_{iz}^2}{2m_i}\right) \, . \tag{2.51}$$

These are just $3N$ identical Gaussian integrals, each evaluated as

$$\int_{-\infty}^{\infty} \mathrm{d}p \, \exp\left(-\frac{\beta p}{2m}\right) = \sqrt{\frac{2m\pi}{\beta}} \, . \tag{2.52}$$

---

[3]The situation is actually a bit subtle...See *Gibbs' Paradox*.

Hence, the partition function can be written as

$$Q(N,V,T) = \frac{1}{h^{3N}N!}\left(\frac{2m\pi}{\beta}\right)^{3N/2}\int\prod_{i=1}^{N}\mathrm{d}^3\mathbf{r}_i\,\exp[-U(\{\mathbf{r}_i\})]$$

$$=: \frac{1}{\Lambda^{3N}N!}\int\prod_{i=1}^{N}\mathrm{d}\mathbf{r}_i\,\exp[-\beta U(\{\mathbf{r}_i\})]\,, \tag{2.53}$$

where $\Lambda := h/\sqrt{2\pi m k_B T}$ is the *de Broglie wavelength*. If the system consists of a mixture of two components A and B, then the classical partition function becomes

$$Q(N_A,N_B,V,T) = \frac{1}{\Lambda_A^{3N_A}\Lambda_B^{3N_B}N_A!N_B!}\int\prod_{i=1}^{N_A}\prod_{j=1}^{N_B}\mathrm{d}^3\mathbf{r}_i\,\mathrm{d}^3\mathbf{r}_j\,\exp[-\beta U(\{\mathbf{r}_i\},\{\mathbf{r}_j\})]\,. \tag{2.54}$$

In molecular systems, a purely classical description is inappropriate. In particular, molecular vibrations and sometimes rotations have to be treated quantum mechanically. Then the classical partition function has the form

$$Q(N,V,T) = \frac{q_{\text{intra}}^N}{\Lambda^{3N}N!}\int\prod_{i=1}^{N}\mathrm{d}^3\mathbf{r}_i\,\exp[-\beta U(\{\mathbf{r}_i\})]\,, \tag{2.55}$$

where q_{intra} is the partition sum over all molecular energy levels

$$q_{\text{intra}} = \sum_i\exp(-\beta\epsilon_i)\,. \tag{2.56}$$

It often has the form

$$q_{\text{intra}} = q_{\text{vib}}q_{\text{rot}}q_{\text{elec}}\,. \tag{2.57}$$

2.5.2 Classical Ideal Gas

Consider classical ideal gas, where there is no interaction between particles. The Hamiltonian of the system is given only by the kinetic energy contribution

$$H_{\text{ideal}}(\{\mathbf{r}_i\},\{\mathbf{p}_i\}) = \sum_{i=1}^{N}\frac{p_i^2}{2m}\,. \tag{2.58}$$

Since there is no interaction between particles, $U = 0$, so the partition function is given by

$$Q_{\text{ideal}}(N,V,T) = \frac{1}{\Lambda^{3N}N!}\int\prod_{i=1}^{N}\mathrm{d}\mathbf{r}_i = \frac{V^N}{\Lambda^{3N}N!}\,. \tag{2.59}$$

Hence,

$$A = -k_B T\ln Q_{\text{ideal}} = -k_B T\ln\left(\frac{V^N}{\Lambda^{3N}N!}\right)\,, \tag{2.60}$$

and from the fundamental equation $\mathrm{d}A = -P\mathrm{d}V - S\mathrm{d}T + \mu\mathrm{d}N$, the expression for the pressure is

$$P_{\text{ideal}} = -\left(\frac{\partial A}{\partial V}\right)_{N,T} = \frac{Nk_B T}{V}\,. \tag{2.61}$$

This is the well-known *ideal gas equation*, which can be rewritten in a perhaps more familiar form

$$pV = nRT \,. \tag{2.62}$$

2.6 Other Ensembles

There are five common ensembles that we usually consider

(i) *Microcanonical (NVE) ensemble*: an isolated system;

(ii) *Canonical (NVT) ensemble*: can exchange energy with the bath;

(iii) *Isothermal-isobaric (NPT) ensemble*: can exchange both energy and volume with the bath;

(iv) *Isoenthalpic-isobaric (NPH) ensemble*: can exchange volume with the bath;

(v) *Grand canonical (μVT) ensemble*: can exchange particles and energy with the bath.

We have already considered microcanonical ensemble and canonical ensemble. Now let's consider an isothermal-isobaric ensemble: a system with N particles that can exchange energy and volume with a large reservoir. The probability finding the system at a given microstate $|i\rangle$ with energy E_i and volume V_i is determined by $\Omega_{\text{bath}}(E_{\text{tot}} - E_i, V_{\text{tot}} - V_i)$. By Taylor expansion as before, and using

$$\left(\frac{\partial S}{\partial E} \right)_{N,V} = \frac{1}{T} \,, \quad \left(\frac{\partial S}{\partial V} \right)_{E,N} = \frac{P}{T} \,, \tag{2.63}$$

we get

$$\ln \Omega_{\text{bath}}(E_{\text{tot}} - E_i, V_{\text{tot}} - V_i) = \ln \Omega_{\text{bath}}(E_{\text{tot}}, V_{\text{tot}}) - \frac{E_i + PV_i}{k_B T} \,. \tag{2.64}$$

The probability of finding the system with volume V is therefore given by

$$P(V) = \frac{Q(N,V,T)e^{-\beta PV}}{\Delta} \,, \tag{2.65}$$

where Δ is the *isothermal-isobaric partition function* quantum mechanically defined as

$$\Delta := \sum_{|i\rangle} Q(N, V_i, T)e^{-\beta PV_i} \,. \tag{2.66}$$

Classically, it is defined as

$$\Delta := \beta P \int_0^\infty \mathrm{d}V \, Q(N,V,T)e^{-\beta PV} \,, \tag{2.67}$$

where βP is to make the quantity dimensionless, and correspondingly

$$P(V) = \frac{\beta P Q(N,V,T)e^{-\beta PV}}{\Delta} \,. \tag{2.68}$$

Then we have the bridge relationship

$$G = -k_B T \ln \Delta \,. \tag{2.69}$$

Another ensemble of great importance is the grand canonical ensemble. Consider a system of volume V that can exchange particles and energy with the reservoir. The probability finding N particles in the system in a state with energy E_i is determined by $\Omega_{\text{bath}}(E_{\text{tot}} - E_i, N_{\text{tot}} - N_i)$. Using

$$\left(\frac{\partial S}{\partial N}\right)_{E,V} = -\frac{\mu}{T} \,, \tag{2.70}$$

we get

$$P(N) = \frac{Q(N,V,T)\exp(\beta\mu N)}{\Xi} \,, \tag{2.71}$$

where

$$\Xi := \sum_{N=0}^{\infty} Q(N,V,T)\exp(\beta\mu N) \tag{2.72}$$

is the *grand partition function*. Then the *grand potential*, defined as[4]

$$\Phi = E - TS - N\mu = -PV \,, \tag{2.73}$$

is related to the grand partition function by the bridge relation

$$\Phi = -k_B T \ln \Xi \,. \tag{2.74}$$

The average number of particles in the system is

$$\langle N \rangle = \left(\frac{\partial \ln \Xi}{\partial \beta\mu}\right)_{V,T} \,. \tag{2.75}$$

It is possible to generate a lot more different ensembles by different combinations of Legendre transforms; however, the microcanonical, canonical and grand canonical ensembles are the most important ones.

2.6.1 Langmuir Adsorption

Why do we need different ensembles? In fact we will show that in the thermodynamic limit, all the ensembles are equivalent. However, some calculations may be more difficult in one ensemble than another. The choice of ensemble is primarily a matter of convenience.

As an illustration, consider the adsorption of gas molecules on the metal surface. Assume the metal has M sites, each can accommodate one molecule at most, and if a particle is adsorbed, its energy is lowered by ϵ. The gas has chemical potential μ.

---

[4]The second equality below follows from the fact that Φ is extensive, while among the natural variables T, V, N, only V is extensive, so Φ must be proportional to V. If you don't trust this relation, you can check appendix section A for a formal proof.

microcanonical ensemble constant N, V, E	canonical ensemble constant N, V, T
$S = k_B \ln \Omega(N, V, E)$	$A = E - TS = -k_B T \ln Q(N, V, T)$
$\mathrm{d}E = -P\mathrm{d}V + T\mathrm{d}S + \mu\mathrm{d}N$	$\mathrm{d}A = -P\mathrm{d}V - S\mathrm{d}T + \mu\mathrm{d}N$
$(1/T) = (\partial S/\partial E)_{N,V}$	$E = (\partial(\beta A)/\partial\beta)_{N,V}$
$P = T(\partial S/\partial V)_{E,N}$	$P = -(\partial A/\partial V)_{N,T}$
$\mu = -T(\partial S/\partial N)_{E,V}$	$\mu = (\partial A/\partial N)_{V,T}$

isothermal-isobaric ensemble constant N, P, T	grand canonical ensemble constant μ, V, T
$G = A + PV = N\mu = -k_B T \ln \Delta(N, P, T)$	$\Phi = A - \mu N = -PV = -k_B T \ln \Xi(\mu, V, T)$
$\mathrm{d}G = V\mathrm{d}P - S\mathrm{d}T + \mu\mathrm{d}N$	$\mathrm{d}\Phi = -P\mathrm{d}V - S\mathrm{d}T - N\mathrm{d}\mu$
$H = E + PV = (\partial(\beta G)/\partial\beta)_{N,P}$	$N = -(\partial\Phi/\partial\mu)_{VT}$
$V = (\partial G/\partial P)_{N,T}$	$P = -(\partial\Phi/\partial V)_{T,\mu}$
$\mu = (\partial G/\partial N)_{P,T}$	$E = N\mu + (\partial(\beta\Phi)/\partial\beta)_{\mu,V}$

Table 1: Common ensembles and related relations.

First let's solve this question in canonical ensemble. If N particles are absorbed, the energy of the system is

$$E(N) = -N\epsilon, \tag{2.76}$$

and the degeneracy of this energy of the system is

$$\Omega(N) = \frac{M!}{N!(M-N)!}. \tag{2.77}$$

The canonical partition function is therefore

$$Q(N, M, T) = \frac{M!}{N!(M-N)!} \exp(\beta N\epsilon). \tag{2.78}$$

The chemical potential is

$$\mu = -k_B T \frac{\partial \ln Q}{\partial N} = -\epsilon + k_B T(\ln N - \ln(M - N)) \tag{2.79}$$

by Stirling's approximation. Define the density of adsorbed particles as $\rho := N/M$, then

$$\mu + \epsilon = k_B T \ln\left(\frac{\rho}{1-\rho}\right). \tag{2.80}$$

By rearranging, we get

$$\rho = \frac{1}{1 + \exp[-\beta(\mu + \epsilon)]}. \tag{2.81}$$

Now let's derive the same result in the grand canonical ensemble.

$$\Xi = \sum_{N=0}^{M} \frac{M!}{N!(M-N)!} \exp[N\beta(\mu+\epsilon)]$$

$$= (1 + \exp[\beta(\mu+\epsilon)])^M \tag{2.82}$$

by the binomial theorem. The average number of particles is therefore

$$\langle N \rangle = \left(\frac{\partial \ln \Xi}{\partial \beta \mu}\right)_{M,T} = \frac{M \exp[\beta(\mu+\epsilon)]}{1 + \exp[\beta(\mu+\epsilon)]} . \tag{2.83}$$

It immediately follows that

$$\langle \rho \rangle = \frac{1}{1 + \exp[-\beta(\mu+\epsilon)]} \tag{2.84}$$

as calculated before.

2.7 Pressure of Interacting Particles

2.7.1 Pressure from the Partition Function

Consider a system of interacting gas particles in a cubic box with edge length $L = V^{1/3}$. Define the scaled coordinate

$$\mathbf{s}_i = \frac{\mathbf{r}_i}{L}, \tag{2.85}$$

so that each component of $\mathbf{s}_i$ ranges between zero and unity. We can then write

$$Q(N,V,T) = \frac{V^N}{\Lambda^{3N} N!} \int_0^1 \cdots \int_0^1 \prod_{i=1}^N \mathrm{d}^3 \mathbf{s}_i \, \exp[-\beta U(\{\mathbf{s}_i\};L)], \tag{2.86}$$

where we have included L as the parameter of U to indicate that U depends on the real rather than the scaled distances between the particles. The pressure is therefore

$$P = k_B T \left(\frac{\partial \ln V^N}{\partial V}\right)_{T,N} + k_B T \left(\frac{\partial}{\partial V} \ln \int_0^1 \cdots \int_0^1 \prod_{i=1}^N \mathrm{d}^3 \mathbf{s}_i \, \exp[-\beta U(\{\mathbf{s}_i\};L)]\right)_{T,N}$$

$$= \frac{N k_B T}{V} - \frac{1}{\int \prod_{i=1}^N \mathrm{d}^3 \mathbf{s}_i \, \exp[-\beta U(\{\mathbf{s}_i\};L)]} \int \prod_{i=1}^N \mathrm{d}^3 \mathbf{s}_i \, \exp[-\beta U(\{\mathbf{s}_i\};L)] \frac{\partial U(\{\mathbf{s}_i\};L)}{\partial V}$$

$$= \underbrace{\frac{N k_B T}{V}}_{\text{ideal}} - \underbrace{\left\langle \frac{\partial U(\{\mathbf{s}_i\};L)}{\partial V} \right\rangle}_{\text{non-ideal}} . \tag{2.87}$$

We now use the chain rule

$$\frac{\partial U(\{\mathbf{s}_i\};L)}{\partial V} = \sum_{i=1}^N \sum_{j=1}^3 \frac{\partial U(\{\mathbf{s}_i\};L)}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial V}, \tag{2.88}$$

where $j \in \{1, 2, 3\}$. We have

$$\frac{\partial r_{ij}}{\partial V} = L\frac{\partial s_{ij}}{\partial V} = s_{ij}\frac{\partial V^{1/3}}{\partial V} = \frac{s_{ij}}{3V^{2/3}} = \frac{r_{ij}}{3V}. \tag{2.89}$$

Hence

$$\frac{\partial U(\{\mathbf{s}_i\}; L)}{\partial V} = \sum_{i=1}^{N}\sum_{j=1}^{3}\frac{\partial U(\{\mathbf{s}_i\}; L)}{\partial r_{ij}}\frac{r_{ij}}{3V}. \tag{2.90}$$

For conservative forces, $\mathbf{f} = -\boldsymbol{\nabla}U$, so

$$\frac{\partial U(\{\mathbf{r}_i\})}{\partial r_{ij}} = -f_{ij}. \tag{2.91}$$

Therefore,

$$P = \frac{Nk_BT}{V} + \frac{1}{3V}\left\langle\sum_{i=1}^{N}\mathbf{f}_i\cdot\mathbf{r}_i\right\rangle. \tag{2.92}$$

2.7.2 Pressure for Pairwise-additive Forces

Consider the special case that the intermolecular forces are pairwise-additive such that

$$\mathbf{f}_i = \sum_{j\neq i}\mathbf{f}_{ij}, \tag{2.93}$$

where $\mathbf{f}_{ij}$ if the force on particle i exerted by particle j. Then

$$P = \frac{Nk_BT}{V} + \left\langle\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{1}{3V}\mathbf{r}_i\cdot\mathbf{f}_{ij}\right\rangle. \tag{2.94}$$

If we permute the dummy indices i and j, we get

$$P = \frac{Nk_BT}{V} + \left\langle\sum_{j=1}^{N}\sum_{i\neq j}^{N}\frac{1}{3V}\mathbf{r}_j\cdot\mathbf{f}_{ji}\right\rangle. \tag{2.95}$$

By Newton's third law, we get

$$P = \frac{Nk_BT}{V} + \frac{1}{6V}\left\langle\sum_{j=1}^{N}\sum_{i\neq j}^{N}\mathbf{r}_{ij}\cdot\mathbf{f}_{ij}\right\rangle, \tag{2.96}$$

where $\mathbf{r}_{ij} := \mathbf{r}_i - \mathbf{r}_j$. This can be rewritten in the dimensionless form as

$$Z := \frac{PV}{Nk_BT} = 1 + \frac{1}{6k_BT}\left\langle\sum_{i\neq j}\mathbf{r}_{ij}\cdot\mathbf{f}_{ij}\right\rangle, \tag{2.97}$$

where Z is the *compressibility factor*. For an ideal gas, $Z = 1$.
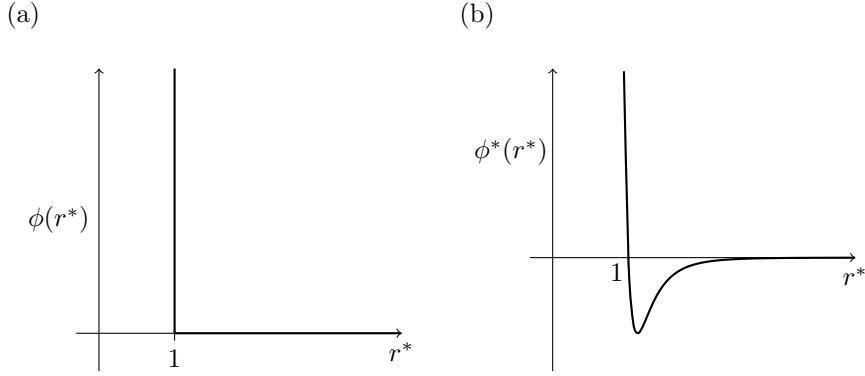
(a)   (b)

Figure 2.4: Two common pair potentials: (a) Hard-sphere potential (b) Lennard-Jones potential.

2.7.3   Law of Corresponding States

If the interactions between the molecules are pairwise-additive, then the potential energy $U(\{\mathbf{r}_i\})$ can be written as

$$U(\{\mathbf{r}_i\}) = \frac{1}{2}\sum_{j=1}^{N}\sum_{i\neq j}\phi(r_{ij})\,. \tag{2.98}$$

The figure shows two simple examples of pair potentials. The hard-sphere potential is usually used to describe the interactions between uncharged colloids, and the Lennard-Jones potential is often used to describe the interaction between atoms or simple (nearly spherical) molecules. At distance less than effective molecular diameter σ, the intermolecular pair potential is harshly repulsive. This is a consequence of Pauli principle. At large distances $r \gg \sigma$, the London dispersion forces take over, the strength of which decays as $1/r^6$. The Lennard-Jones potential provides a convenient interpolation between the short-range repulsion and the long-range attraction. It is of the form

$$\phi(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]\,. \tag{2.99}$$

The depth of the potential is ϵ, whilst the range of the potential is determined by the effective diameter σ. Define $r^* := r/\sigma$ and $\phi^* := \phi/\epsilon$, then

$$\phi^*(r^*) = u_{LJ}(r^*) := 4\left[r^{*12} - r^{*6}\right]\,. \tag{2.100}$$

Therefore, if we express all energies in units of ϵ and distances in units of σ, then all Lennard-Jones potentials are identical.

If in appropriate units, the interactions between any kind of gas molecules are equal, their physics should also be the same. Using this definition, we can rewrite the compressibility in terms of the reduced quantities r^* and $f_{ij}^* := f_{ij}\sigma/\epsilon$, giving

$$Z = 1 + \frac{\epsilon}{6k_BT}\left\langle\sum_{i\neq j}\mathbf{r}_{ij}^*\cdot\mathbf{f}_{ij}^*\right\rangle \tag{2.101}$$

Now define the reduced temperature $T^* := k_BT/\epsilon$ and the reduced density $\rho^* :=$
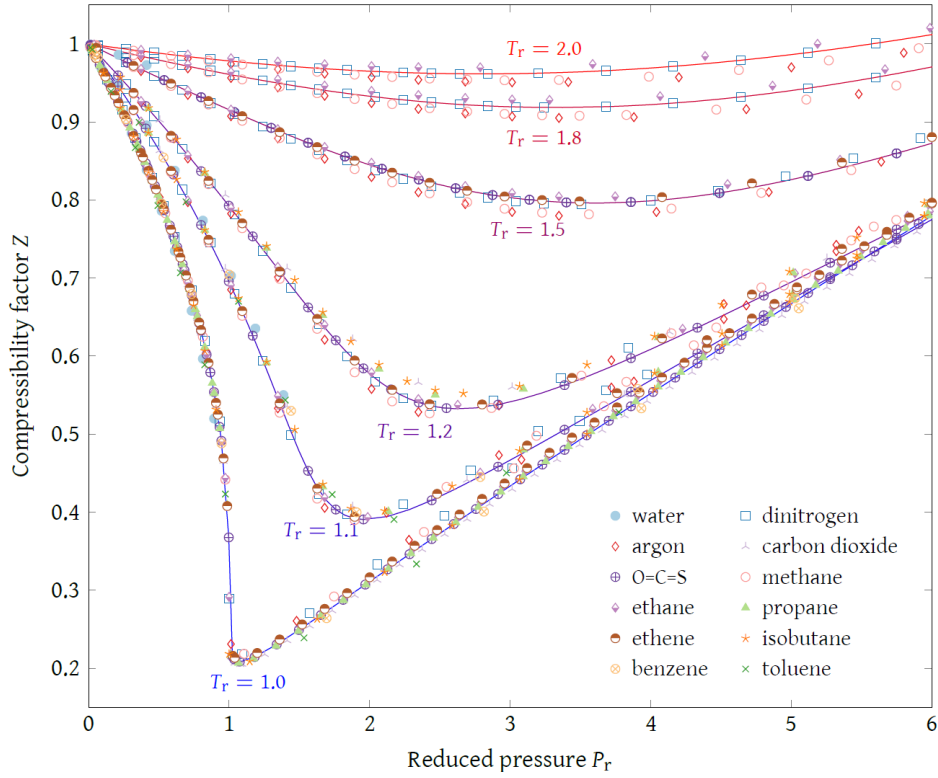
20

Figure 2.5: This beautiful figure is adapted from the official course notes by Dr. Aleks Reinhardt. The compressibilities of a range of gases are plotted in reduced quantities, and they overlap nicely. This implies that the functional form of their pair potentials must be very similar.

$\sigma^3 N/V$, then

$$Z(\rho^*, T^*) = 1 + \frac{1}{6T^*} \left\langle \sum_{i \neq j} \mathbf{r}_{ij}^* \cdot \mathbf{f}_{ij}^* \right\rangle. \tag{2.102}$$

This equation implies that if we express the temperature in units of ϵ/k_B and density in units of σ^{-3}, then the compressibility factors of all substances that can be described by a pair potential in the same functional form collapse onto the same set of curves. This is known as *the principle of corresponding states*.

However, it should be stressed that not all substances obey the same law of corresponding states. For instance, the compressibility factor of dipolar molecules would not collapse onto any of the curves in the figure, because the functional form of the pair potential is too different.

We can rewrite the above expression explicitly in terms of ρ^*. We have

$$P = \frac{N k_B T}{V} + \frac{N}{6V} \left\langle \sum_{i \neq j} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} \right\rangle, \tag{2.103}$$

which can be written in reduced quantity as

$$P^* = \rho^* T^* + \frac{\rho^*}{6} \left\langle \sum_{i \neq j} \mathbf{r}_{ij}^* \cdot \mathbf{f}_{ij}^* \right\rangle. \tag{2.104}$$

21

We then have

$$\left\langle \sum_{i \neq j} \mathbf{r}_{ij}^* \cdot \mathbf{f}_{ij}^* \right\rangle = \frac{6(P^* - \rho^* T^*)}{\rho^*} \, . \tag{2.105}$$

Hence the compressibility factor is

$$Z = 1 + \frac{1}{6T^*} \frac{P^* - \rho^* T^*}{\rho^*}$$
$$= \frac{P^*}{\rho^* T^*} \, . \tag{2.106}$$

3 Phase Behaviour

3.1 Equilibrium between Phases

Consider an isolated system of N particles in a volume V with a total energy E. The system consists of multiple distinct phases between which particles, volume and energy can be exchanged. We ignore any contributions of the interfaces to the extensive properties, as their contribution scales as $N^{2/3}$, so their ratio compared with bulk contribution scales as $N^{-1/3}$, vanishing in the thermodynamic limit as $N \to \infty$. Then the equilibrium conditions of coexistence of phases are that the pressures of all coexisting phases must be equal ($P_1 = P_2 = \cdots = P$), as must be the temperature ($T_1 = T_2 = \cdots = T$) and the chemical potentials of all species ($\mu_1^\alpha = \mu_2^\alpha = \cdots = \mu^\alpha$).

3.1.1 Stability

We use the thermodynamic arguments to find conditions for the stability of a single phase. Consider a homogeneous phase in an isolated system with constant N, V and E that is divided into two equal halves. At equilibrium, the entropy of the total system is at a maximum. Now consider transferring a small amount of energy ΔE from one half of the system to the other. At equilibrium, the temperature in the two halves of the system is the same, and hence to linear order of ΔE, the entropy does not change. Consider the second order variation,

$$\Delta S = \frac{1}{2}\left(\frac{\partial(1/T)}{\partial E}(\Delta E)^2\right)_1 + \frac{1}{2}\left(\frac{\partial(1/T)}{\partial E}(\Delta E)^2\right)_2 . \tag{3.1}$$

Since the two halves of the system are identical, these two terms are the same. We can write

$$\Delta S = \frac{\partial(1/T)}{\partial E}(\Delta E)^2 . \tag{3.2}$$

As S is a maximum, $\Delta S \leq 0$, and hence

$$\frac{\partial E}{\partial(1/T)} \leq 0 , \tag{3.3}$$

which implies that

$$-T^2\frac{\partial E}{\partial T} = -T^2 C_V \leq 0 , \tag{3.4}$$

and hence the heat capacity $C_V \geq 0$. Thermodynamic stability implies that the heat capacity is never negative.

A similar argument can be used to show that the compressibility of a system is non-negative. Consider the condition for equilibrium at constant N, V and T. Under these conditions, the Helmholtz energy must be a minimum. We divide the system into two equal parts, both containing $N/2$ particles in a volume $V/2$ at constant T. We now vary the volume of one half of the system by an amount ΔV, and the other by $-\Delta V$. To first order in ΔV, the total variation of the free energy is

$$\Delta A = \frac{\partial^2 A}{\partial V^2}(\Delta V)^2 . \tag{3.5}$$

This can be rewritten as

$$\Delta A = -\frac{\partial P}{\partial V}(\Delta V)^2 = \frac{1}{\kappa V}(\Delta V)^2, \tag{3.6}$$

where κ is the *isothermal compressibility* of the system. As A is minimised at equilibrium, this implies that $\kappa \geq 0$.

3.2 Coexistence

Consider a con-component system that can exist in two phases, with free energies $A_1(N, V, T)$ and $A_2(N, V, T)$. We would like to determine the condition that the two phase coexists. It turns out that there is a easy graphical method to do this. For a given temperature, we plot $A_1(N, V, T)$ and $A_2(N, V, T)$ as a function of V while keeping N and T constant. We now have to check if it is possible to draw a line that is tangent of both curves. If this common tangent exists, then this line touches A_1 at volume V_1 and it touches A_2 at a volume V_2. As the curves A_1 and A_2 has a common tangent, the derivatives of A_1 and A_2 at V_1 and V_2, respectively, are the same. As $(\partial A/\partial V)_T = -P$, this implies that the pressure of phase 1 at V_1 is the same of that of phase 2 at V_2, i.e. $P_1 = P_2$. The tangents also have a common intercept at $V = 0$. This intercept is

$$A_1 - \frac{\partial A_1}{\partial V}V_1 = A_1 + PV_1 = G_1 = N\mu_1, \tag{3.7}$$

and this is equal to

$$A_2 - \frac{\partial A_2}{\partial V}V_2 = A_2 + PV_2 = G_2 = N\mu_2. \tag{3.8}$$

Therefore $\mu_1 = \mu_2$ and the volumes V_1 and V_2 are the volumes of the coexisting phases. A completely analogous analysis can be performed if we plot A/V against N/V, in which case the slope is related to the chemical potential and the intercept is related to the pressure.
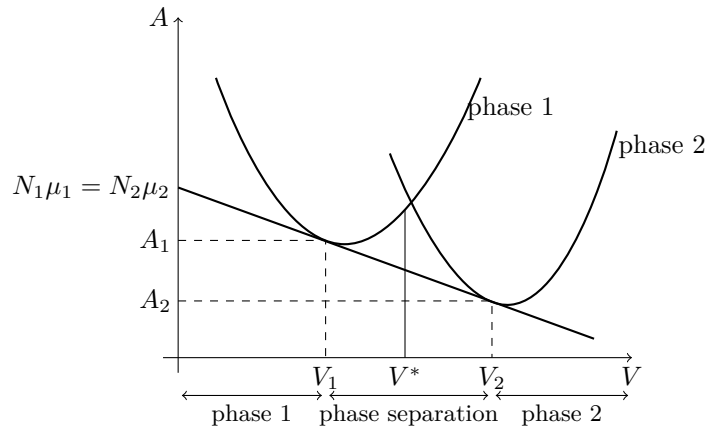

If the system is at some volume V^* with $V_1 \leq V^* \leq V_2$, the system would have lower Helmholtz energy if it is separated into the two phases than if it exists in either of the single phase. Define $v := V/N$, then

$$Nv^* = N_1 v_1 + (N - N_1)v_2. \tag{3.9}$$

24

We can then derive the proportion of particles being in phase 1 at equilibrium

$$x_1 := \frac{N_1}{N} = \frac{v^* - v_2}{v_1 - v_2}.$$

(3.10)

Suppose now we have a free energy curve for some phase. We have shown that the free energy must be a convex function of V because the compressibility κ must be non-negative at equilibrium. Hence, the non-convex part in curve A versus V does not correspond to an equilibrium situation. We can still work out the point of coexistence by common tangent construction. But now we have a region where the phase separation is thermodynamically favourable, but the local second derivative of A with respect to V is still positive, meaning that any perturbation of such system will increase the free energy. Such system is described as meta-stable as phase separation will involve a free-energy barrier. Between the inflection points, the free energy curve is non-convex and hence in this range a homogeneous phase would be absolutely unstable, with a negative compressibility. Such a system will phase separate spontaneously under any infinitesimal fluctuation. The boundary between the metastable and unstable regions are called *spinodal*.
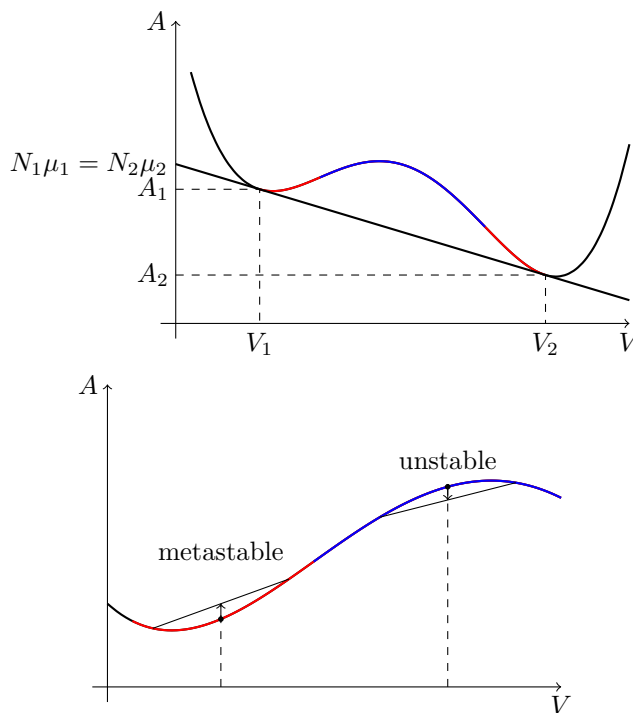
It should be stressed that the spinodal, although a useful qualitative concept, is not well defined. It appears naturally when we use approximate expressions for the free energy. However, an exact expression for the equilibrium free energy is necessarily convex and therefore has no spinodal.

3.3 Phase Diagrams

A phase diagrams can tell us at a glance the conditions under which each phase is thermodynamically stable. For a one-component system, the thermodynamic variables

P, T and V, which are usually used to describe the thermodynamic state of the matter, are not independent of one another. For example, if we specify the volume and the temperature of a system, the pressure will be defined as well through the equation of states. We can therefore think of the phase diagram as a surface in the PVT space. Such a three-dimensional phase diagram can be projected onto the $P - T$, $V - T$, $P - V$ or $T - \rho$ planes to give the more familiar two-dimensional phase diagrams.

There are many other parameters that we could vary; for example we could change the composition of a mixture or the strength of an external magnetic field, and so other types of phase diagrams can be constructed, some of which we shall see in the chapters that follow.

3.4 Thermodynamic Integration

The second law of thermodynamics states that for an isolated system with energy E, volume V and number of particles N, the entropy S is maximised when the system is at equilibrium. From this formulation of the second law, it is straightforward to derive the corresponding equilibrium conditions for systems that can exchange heat, particles or volume with a reservoir, as we have already shown. In particular, if a system is in contact with a heat bath, such that its temperature T, volume V and number of particles N are fixed, then the Helmholtz energy $A := E - TS$ is at a minimum in equilibrium. Analogously, for a system of N particles at constant pressure P and temperature T, the Gibbs energy $G := A + PV$ is at a minimum.

It we wish to know which of two phases, denoted 1 and 2, is stable at a given temperature and density, we simply need to compare the Helmholtz energies A_1 and A_2. Unfortunately, it is not possible to measure the free energy or entropy directly. What we can measure are averages of *mechanical quantities*, i.e. averages of functions of the coordinates and momenta of the molecules in the system, such as the pressure or the dielectric constant. If we denote such a mechanical quantity by $X(\{\mathbf{p}_i\}, \{\mathbf{r}_i\})$, then the average of X that can be measure in an experiment at constant N, V and T is

$$\langle X \rangle_{NVT} = \frac{\int \prod_{i=1}^{N} \mathrm{d}^3\mathbf{p}_i \, \mathrm{d}^3\mathbf{r}_i \, X(\{\mathbf{p}_i\}, \{\mathbf{r}_i\}) \exp[-\beta H(\{\mathbf{p}_i\}, \{\mathbf{r}_i\})]}{\int \prod_{i=1}^{N} \mathrm{d}^3\mathbf{p}_i \, \mathrm{d}^3\mathbf{r}_i \, \exp[-\beta H(\{\mathbf{p}_i\}, \{\mathbf{r}_i\})]}, \tag{3.11}$$

where $H(\{\mathbf{p}_i\}, \{\mathbf{r}_i\})$ is the system's Hamiltonian.

However, the entropy, the free energy and related quantities are not simply averages of functions of the phase-space coordinates of the system. Rather, they are directly related to the volume in phase space that is accessible to a system. For instance, in classical statistical mechanics, the Helmholtz energy A is directly related to the canonical partition function $Q(N, V, T)$ through

$$A = -k_B T \ln Q(N, V, T)$$

$$= -k_B T \ln \left[\frac{1}{h^{3N} N!} \int \prod_{i=1}^{N} \mathrm{d}^3\mathbf{p} \, \mathrm{d}^3\mathbf{r} \, \exp[-\beta H(\{\mathbf{p}_i\}, \{\mathbf{r}_i\})] \right]. \tag{3.12}$$

Unlike quantities such as the internal energy the pressure or the polarisation,

$Q(N, V, T)$ itself is not a canonical average over the phase space. This is why A and, for that matter, S or G, cannot be measured directly. We call quantities that depend directly on the available volume in the phase space *thermal quantities*. However, derivatives of the free energy with respect to the volume V or the temperature T are mechanical quantities and can be measured: namely, they are

$$\left(\frac{\partial A}{\partial V}\right)_{N,T} = -P \text{ and } \left(\frac{\partial(A/T)}{\partial(1/T)}\right)_{V,T} = E\,. \tag{3.13}$$

In order to compute the free energy of a system at a given temperature and density, we must find a reversible path in the volume-temperature plane that links the state under consideration to a state of known free energy. Then change in A along that path can then be simply be evaluated by integrating (3.13). Only a few thermodynamic states exist whose free energy is known exactly: one such state in the ideal gas; another may be perfectly ordered ground state at $T = 0\,\text{K}$.

4 Thermodynamic Perturbation Theory

4.1 Gibbs–Bogoliubov Relation

The aim of thermodynamic perturbation theory is to arrive at an estimate of the free energy (and all derived properties) of a many-body system, using as input information about the free energy and structure of a simpler reference system. We assume that the potential energy function of this reference system is denoted by U_0, while the potential energy function of the system of interest is denoted by U_1. In order to compute the free-energy difference between the known reference system and the system of interest, we use a linear parameterisation of the potential energy function

$$U(\lambda) \coloneqq \lambda U_1 + (1-\lambda)U_0 \,. \tag{4.1}$$

The free energy of a system with this generalised potential energy is denoted $A(\lambda)$ and can be computed from $A(\lambda) = -k_B T \ln Q(\lambda)$. In particular, we noted when introducing (3.13) that derivatives of the free energy are often mechanical observables, and the same applies in this case. The derivative of $A(\lambda)$ with respect to λ is

$$
\begin{aligned}
\left(\frac{\partial A}{\partial \lambda}\right)_{N,V,T} &= \frac{\int \prod_{i=1}^{N} \mathrm{d}^3\mathbf{r}_i \left(\frac{\partial U(\{\mathbf{r}_i\};\lambda)}{\partial \lambda}\right) \exp[-\beta U(\{\mathbf{r}_i\};\lambda)]}{\int \prod_{i=1}^{N} \mathrm{d}^3\mathbf{r}_i \, \exp[-\beta U(\{\mathbf{r}_i\};\lambda)]} \\
&= \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \\
&= \langle U_1 - U_0 \rangle_\lambda \,,
\end{aligned}
\tag{4.2}
$$

and hence we can express the free-energy difference $A_1 - A_0$ as

$$A_1 - A_0 = \int_0^1 \mathrm{d}\lambda \, \langle U_1 - U_0 \rangle_\lambda \,, \tag{4.3}$$

where the subscript λ indicates that we evaluate the thermal average of the value of λ. This expression, first derived by John Kirkwood in 1935, allows us to compute the free energy of a system of interest by relating it to a reference system. This approach has extensively been used in computer simulations since the mid-1980s. It is often referred to as 'artificial' or 'Hamiltonian' thermodynamic integration because it entails changing the underlying interactions in the system's Hamiltonian.

Moreover, it allows us to find bounds on the free energy A_1. It is straightforward to show that

$$\left(\frac{\partial^2 A}{\partial \lambda^2}\right)_{NVT} = -\beta \left(\left\langle (U_1 - U_0)^2 \right\rangle_\lambda - \langle U_1 - U_0 \rangle_\lambda^2\right) \leq 0 \,. \tag{4.4}$$

Since the right-hand side of the equation is the negative of a variance, we can note that the second derivative of A with respect to λ is always non-positive. This implies that

$$\left(\frac{\partial A}{\partial \lambda}\right)_{\lambda=0} \geq \left(\frac{\partial A}{\partial \lambda}\right)_{\lambda>0} \,, \tag{4.5}$$

and hence

$$A_1 \leq A_0 + \langle U_1 - U_0 \rangle_\lambda \ . \tag{4.6}$$

This variational principle for the free energy is known as the *Gibbs–Bogoliubov inequality*: we compute an upper bound to the free energy of the system of interest from a knowledge of the average of $U_1 - U_0$ evaluated for the reference system. The latter is something that can often be computed because the reference system is by construction a system which we can solve.
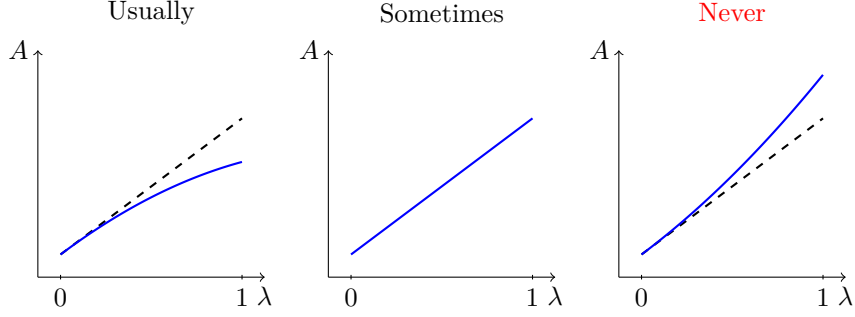

Figure 4.1: The Gibbs–Bogoliubov inequality shows that the first order perturbation estimation of the free energy is never lower than the true free energy.

We may let of reference system depend on some parameters and use the variational principle the get our best estimation of the upper bound of Helmholtz energy.

It may seem at first glance that we are doing something slightly off here: what does it actually mean that we compute $\langle U_1 \rangle_{\lambda=0}$ for the reference potential, given that when $\lambda = 0$, the system is governed by U_0 only? In this case, U_1 does not contribute to the Boltzmann factor, but that does bot mean we are not able to compute what it is for a given microstate, since it depend only on the momenta of a given microstate. This is similar to computing any other mechanical observable, e.g. the pressure, which also would not appear in the Boltzmann factor. It is evaluated as

$$\langle U_1 - U_0 \rangle_{\lambda=0} = \frac{\int \prod_{i=1}^N \mathrm{d}^3 \mathbf{r}_i \, (U_1 - U_0) \exp(-\beta U_0)}{\int \prod_{i=1}^N \mathrm{d}^3 \mathbf{r}_i \, \exp(-\beta U_0)} \ . \tag{4.7}$$

The usefulness of (4.6) depends crucially on the quality of the choice of reference system. A good reference system is not necessarily close in free energy to the system of interest, but one for which the fluctuation in the potential energy difference $U_1 - U_0$ are small. Thermodynamic perturbation theory for simple liquids has been very successfully precise because the structure of the reference system (hard-sphere fluid) and the liquid under consideration (e.g. Lennard-Jones) are very similar. As a result, $\langle U_1 - U_0 \rangle_\lambda$ hardly depends on λ and, as a consequence, its derivative is small.

4.2 Mean-Field Theory and the Ising Model

The Gibbs–Bogoliubov inequality can be used as a starting point from which we can derive mean-field theory, which provides a systematic approximation to the free energy
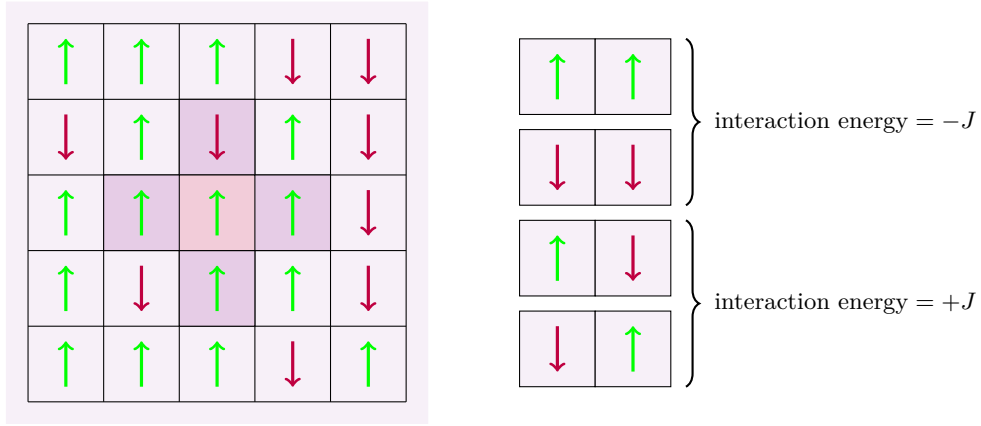
Figure 4.2: A schematic representation of the Ising model in two dimensions.

of a many-body system. For the sake of simplicity, we consider the Ising model, for which a many-body Hamiltonian of N interacting spins is given by

$$U_1 = -\frac{J}{2} \sum_{i=1}^{N} \sum_{\langle i,j \rangle} s_i s_j \,, \tag{4.8}$$

where $\langle i,j \rangle$ indicates that the particles j must be nearest neighbours of i and the factor of $1/2$ ensures that we do not double count the interactions. The Ising model is not just relevant for magnetic systems; in section D, we show that a lattice-gas model that can be used to describe the liquid-vapour transition is, in fact, equivalent to the Ising model.

We wish to approximate this model system using a reference system with a much simpler Hamiltonian, namely one that consists only of a sum of one-particle contributions. For example,

$$U_0 = -\sum_{i=1}^{N} h s_i \,, \tag{4.9}$$

where h denotes the effective field that replaces the interaction with the other particles. The molecular partition function counting all states of a single spin (with energy $-hs$) in the reference system is

$$q_0 = \int \mathrm{d}s \, \exp(\beta h s) \tag{4.10}$$

and the free energy per spin is

$$a_0(h) = -k_B T \ln q_0 = -k_B T \ln \int \mathrm{d}s \, \exp(\beta h s) \,. \tag{4.11}$$

Here, we use the shorthand a for A/N. If the spins are quantised and we only have two spin states ($+1$ and -1), the partition function reduces to

$$q_0 = \exp(\beta h) + \exp(-\beta h) = 2 \cosh(\beta h) \tag{4.12}$$

and the free energy becomes

$$a_0(h) = -k_B T \ln(2 \cosh(\beta h)) \,. \tag{4.13}$$

We can easily compute the average value of s in the reference system,

$$\langle s \rangle_0 = \frac{\int \mathrm{d}s\, s \exp(\beta h s)}{q_0} = \frac{1}{q_0} \frac{\partial}{\partial \beta h} \int \mathrm{d}s\, \exp(\beta h s) = k_B T \frac{\partial \ln q_0}{\partial h}\,, \qquad (4.14)$$

or simply

$$\langle s \rangle_0 = -\frac{\partial a_0(h)}{\partial h}\,. \qquad (4.15)$$

In the case of spins ± 1,

$$\langle s \rangle_0 = \tanh(\beta h)\,, \qquad (4.16)$$

which we could have obtained directly from the definition of the mean magnetisation, but the expression involving the derivative of a_0 will come in useful shortly. Now, we consider the Gibbs–Bogoliubov inequality, where

$$a_{\mathrm{MF}} = a_0 + \left\langle -\frac{J}{2} \sum_{\langle i,j \rangle} s_i s_j + h s_i \right\rangle_0$$

$$= a_0 - \frac{J}{2} z \langle s \rangle_0^2 + h \langle s \rangle_0\,. \qquad (4.17)$$

In the last line, we have introduced z, the coordination number of particle i. Moreover, we have used the fact that, in the reference system, different spins are uncorrelated, $\langle s_i s_j \rangle_0 = \langle s_i \rangle_0 \langle s_j \rangle_0$. We now look for the optimum value fo h, i.e. the one that minimises our estimate of a_{MF}. Carrying out the differentiation with respect to h, we find that

$$0 = \frac{\partial}{\partial h} \left(a_0 - \frac{J}{2} z \langle s \rangle_0^2 + h \langle s \rangle_0 \right)$$

$$= -\langle s \rangle_0 - (Jz \langle s \rangle_0 - h) \frac{\partial \langle s \rangle_0}{\partial h} + \langle s \rangle_0$$

$$= -(Jz \langle s \rangle_0 - h) \frac{\partial \langle s \rangle_0}{\partial h}\,. \qquad (4.18)$$

Since $\frac{\partial \langle s \rangle_0}{\partial h} \neq 0$, we can conclude that

$$h = Jz \langle s \rangle_0\,. \qquad (4.19)$$

Finally, we can insert this expression for h into (4.16) to obtain an implicit equation for $\langle s \rangle_0$,

$$\langle s \rangle_0 = \tanh(\beta Jz \langle s \rangle_0)\,, \qquad (4.20)$$

which can be solved to yield $\langle s \rangle_0$ as a function of T.

Above the *critical* (or *Curie*) *temperature* $k_B T = Jz$, there is only one solution: at high temperatures, the entropy dominates the free energy and so the spins are randomly distributed. The significance of the entropy as a driving force is diminished as the temperature is decreased, and below the Curie temperature, the net alignment of spins, which is energetically favourable, becomes dominant, and the mean magnetisation per spin is non-zero. In fact, as can be seen from the figure, there are three solutions to the mean magnetisation. However, the $\langle s \rangle = 0$ solution is actually a maximum, rather than a minimum, so we will ignore it. In the absence of an external magnetic field, the system is equally likely to be magnetised in the up or the
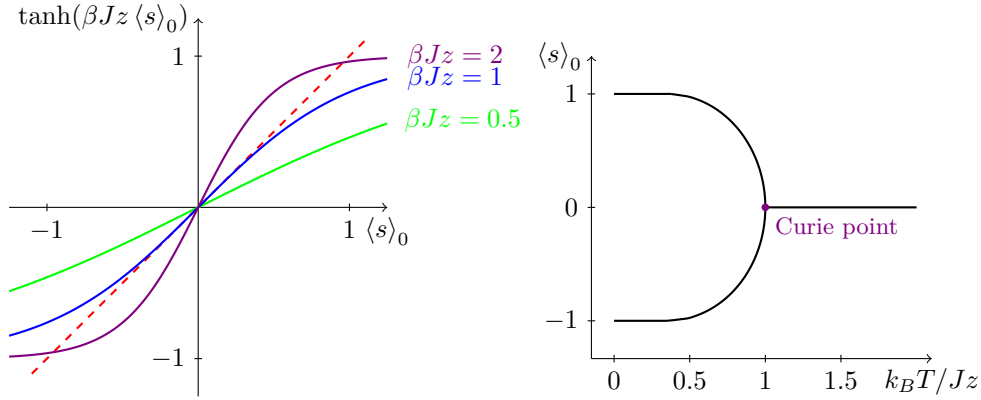
31

Figure 4.3: (a) A graphical method for the solution of the mean magnetisation in Ising model. (b) The phase diagram of an Ising model system.

down spin configurations, but the fluctuations causing the transitions between the two are extremely unlikely. The ergodic hypothesis is thus formally violated, and this phenomenon is sometimes described as *spontaneous symmetry breaking*.

The free energy estimate that we obtain when inserting (4.19) into (4.17) is

$$a_{\mathrm{MF}} = a_0 + \frac{J}{2} z \langle s \rangle_0^2 \, . \tag{4.21}$$

The subscript 'MF' in this expression stands for the mean field approximation. It is very important to bear in mind that the free energy that results from the mean-field approximation is in general not simply the free energy of the reference system with the effective field, but has an additional term that depends on the difference between the two potentials measured for the system with the effective field.

4.2.1 Validity of Mean Field Theory

Having solved the Ising model using mean field theory, a question that arises is: are our results correct.

There is actually a version of the Ising model for which the mean field theory is exact: it is the $d = \infty$ dimensional lattice. This is unphysical. Roughly speaking, mean field theory works for large d because each spin has a large number of neighbours and so indeed sees something close to the average spin.

But what about dimensions of interest? Mean field theory gets things most dramatically wrong in $d = 1$. In that case, no phase transition actually occurs. We will solve the 1D Ising model exactly in the section F. There is a general lesson here: in low dimensions, both thermal and quantum fluctuations are more important and invariably stop systems forming ordered phases.

In higher dimensions, $d \geq 2$, the crude features of the phase diagram, including the existence of a phase transition, given by mean field theory are essentially correct. In fact, the very existence of a phase transition is already worthy of comment. The
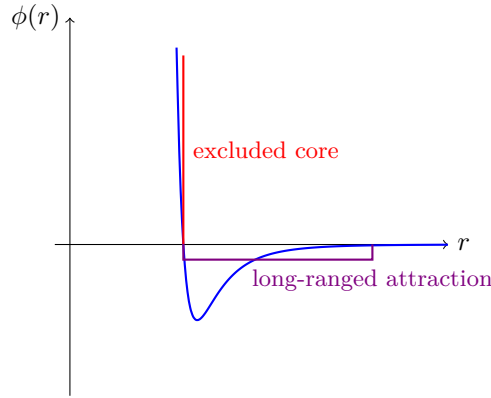
Figure 4.4: A typical interatomic potential (shown in blue) can be approximated by a hard-core repulsion and a very weak attractive interaction $-\epsilon$.

defining feature of a phase transition is behaviour that jumps discontinuously as we vary β. Mathematically, the functions must be non-analytic. Yet all properties of the theory can be extracted from the partition function Q which is a sum of smooth, analytic exponential functions

$$Q = \sum_{|i\rangle} e^{-\beta E_i} \,. \tag{4.22}$$

How can we get a phase transition? The loophole is that Q is only necessarily analytic if the sum is finite. But there is no such guarantee when the number of lattice sites $N \to \infty$. We reach a conclusion: phase transitions only strictly happen in the thermodynamic limit. There are no phase transitions in finite systems.

The 2D Ising model in a square lattice has been solved exactly, first by Onsager — his method was famously complicated. In appendix section F, we will use a clever trick to obtain the exact result of critical temperature, which is

$$k_B T = \frac{2J}{\log(\sqrt{2} + 1)} \approx 2.269 J \,. \tag{4.23}$$

This is actually not hugely different from the mean field approximation, which predicts

$$k_B T = Jz = 4J \,. \tag{4.24}$$

4.3 Van der Waals Equation

In some cases, the fluctuation term in the Gibbs–Bogoliubov approach can be made to vanish altogether. A famous example of this occurring is the mean-field treatment of the Van der Waals fluid, where we take a hard-sphere reference system and add a very weak, attractive interaction $-\epsilon$ that extends to a very large distance R_c. In order to compute the Van der Waals free energy, we first need to work out the equation of state of a hard-sphere gas. Unfortunately this proves to be the more challenging part of the calculation in this case.

4.3.1 Hard-Sphere Gas

The hard-sphere potential is given by

$$U_0 = \sum_{i<j} \phi(r_{ij}) , \tag{4.25}$$

where

$$\phi(r_{ij}) = \begin{cases} 0 & \text{if } r_{ij} > \sigma \\ \infty & \text{if } r_{ij} \leq \sigma , \end{cases} \tag{4.26}$$

where σ is the diameter of the particle.

Provided the hard-sphere gas is sufficiently dilute, we can estimate that the addition of each additional hard sphere reduced the volume available to the next one by $v_{\text{ex}} = 4\pi\sigma^3/3$, the volume excluded by a single hard sphere. In the dilute limit, we estimate that no two excluded volumes will overlap. An illustration of this set-up is given in the figure below. In this case, we can simply modify the perfect gas partition function to give

$$\begin{aligned} Q_{\text{HS}} &\approx \frac{1}{\Lambda^{3N} N!} V(V - v_{\text{ex}})(V - 2v_{\text{ex}}) \ldots (V - (N-1)v_{\text{ex}}) \\ &= \frac{V^N}{\Lambda^{3N} N!} \left(1 - \frac{v_{\text{ex}}}{V}\right) \left(1 - 2\frac{v_{\text{ex}}}{V}\right) \ldots \left(1 - (N-1)\frac{v_{\text{ex}}}{V}\right) . \end{aligned} \tag{4.27}$$
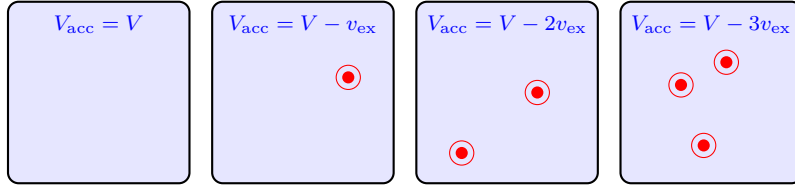

Figure 4.5: Ignoring any volume excluded by the walls and any excluded volume overlaps, each particle inserted into the system excludes a volume v_{ex}, reducing the accessible volume.

In the dilute regime, $(N-1)v_{\text{ex}}/V \ll 1$, so we can use the power series $\exp(-x) = 1 - x$, giving

$$\begin{aligned} Q_{\text{HS}} &= \frac{V^N}{\Lambda^{3N} N!} \exp\left(-\frac{v_{\text{ex}}}{V}\right) \exp\left(-\frac{2v_{\text{ex}}}{V}\right) \ldots \exp\left(-\frac{(N-1)v_{\text{ex}}}{V}\right) \\ &= \frac{V^N}{\Lambda^{3N} N!} \exp\left(-\frac{v_{\text{ex}}}{V} \sum_{i=1}^{N-1} i\right) . \end{aligned} \tag{4.28}$$

The summation is an arithmetic series that we can evaluate as $N(N-1)/2$. Since $Nv_{\text{ex}}/2V \ll 1$ by construction, we can ignore the linear term in the exponential and write

$$Q_{\text{HS}} \approx \frac{V^N}{\Lambda^{3N} N!} \exp\left(-\frac{N^2 v_{\text{ex}}}{2V}\right) = \frac{V^N}{\Lambda^{3N} N!} \left[\exp\left(-\frac{N v_{\text{ex}}}{2V}\right)\right]^N . \tag{4.29}$$

If we use the same expansion again and if we note that $v_{\text{ex}} = 8v_0$, where v_0 is the volume of a single hard sphere particle, we obtain

$$Q_{\text{HS}} \approx \frac{(V - 4Nv_0)^N}{\Lambda^{3N} N!} \, . \tag{4.30}$$

Using $A = -k_B T \ln Q$, Stirling's approximation and $P = -(\frac{\partial A}{\partial V})_{N,T}$, we can write

$$P_{\text{HS}} \approx \frac{Nk_B T}{V - 4Nv_0} = \frac{k_B T}{v_0} \frac{\phi}{1 - 4\phi} \, , \tag{4.31}$$

where in the last step we rewrote the volume in terms of the overall fraction of volume occupied, $\phi := Nv_0/V$. We note that this function diverges when $\phi \to 1/4$. In reality, hard spheres can be compressed to the close-packed volume fraction $\phi = \pi/\sqrt{18} \approx 0.7405$, and so the approximate partition function above is not appropriate at higher densities. This hard-sphere equation of state, and the corresponding Van der Waals equation of state that we will derive from it, are thus approximations that are only reasonable at low density.

4.3.2 Mean-Field Attractions

We have accounted for, in an approximated way, the hard-core repulsion between particles. We would now like to add the weak pairwise attraction $-\epsilon$. In particular, we consider the limit of very weak attractions ($\epsilon \to 0$) that are very long-range ($R_c \to \infty$). In this limit, the potential energy of the perturbed fluid can be computed directly. Within a shell of radius R_c, there will be, on average, $N_c = (4/3)\pi R_c^3 \rho$ other particles, where ρ is the number density of the fluid. All these particles contribute $-\epsilon/2$ to the potential energy of the fluid, with the factor of $1/2$ accounting for the double-counting of pairwise interaction. For $R_c \to \infty$, the number of neighbours N_c also tends to infinity and hence the relative fluctuation in N_c becomes negligible, as it is of order $1/\sqrt{N_c}$. It then follows that the fluctuation in the perturbation (i.e. $U_1 - U_0$) also become negligible and, since we have shown that the second derivative of the free energy is proportional to the variance in $U_1 - U_0$, the first-order Maclaurin expansion of the free energy in λ becomes exact. Therefore the Gibbs–Bogoliubov inequality becomes an identity. The free energy per particle of the Van der Waals fluid is therefore

$$\left(\frac{A}{N}\right)_{\text{VdW}} (\rho, T) = \left(\frac{A}{N}\right)_{\text{HS}} (\rho, T) - \rho a \, , \tag{4.32}$$

where we have defined $a = (2\pi/3)R_c^3 \epsilon$. The corresponding pressure is

$$P_{\text{VdW}} = -\left(\frac{\partial A_{\text{VdW}}}{\partial V}\right)_{N,T} = \rho^2 \frac{\partial (A/N)_{\text{VdW}}}{\partial \rho} = P_{\text{HS}}(\rho, T) - a^2 \rho \, , \tag{4.33}$$

where P_{HS} denotes the pressure of the hard-sphere reference system. We therefore have the well-known Van der Waals equation

$$(P_{\text{VdW}} + a\rho^2)(V - Nb) = Nk_B T \, , \tag{4.34}$$

where $b = 4v_0$. This equation is very interesting to study because it is probably the simplest model system which exhibits vapour-liquid coexistence. A simple way of computing the phase diagram involves the use of Maxwell's equal-area construction.

4.4   Maxwell Equal-Area Construction

We can sketch the behaviour of the pressure with volume above and below the critical point. Above the critical point, a single fluid phase exists; below the critical point, a vapour (low-density fluid) and a liquid (high-density fluid) are predicted. Below the critical point, there are regions where the equation of state describes a mechanically unstable state where the isothermal compressibility is negative, and where a single pressure corresponds to multiple possible volumes. Such regions are known as Van der Waals loops. The critical point itself occurs at the point at which the minimum and the maximum of the Van der Waals loop merge, i.e. the first and second derivatives of the pressure with respect to volume vanish.

In regions where there are Van der Waals loops, we can immediately find the points corresponding to limits of mechanical stability: these spinodals correspond to the points at which $(\partial P/\partial V)$ changes sign. Finding the points of coexistence (the *binodals*) is slightly more challenging. Various approaches can be used; here, we will use a method called Maxwell construction.

We have already shown that the condition for phase equilibrium — in this case between the liquid and the vapour — is the equality of pressures, temperatures and chemical potentials. From our discussion of the common tangent construction approach, we know that the condition that the chemical potentials are equal is given by

$$N\mu_1 = N\mu_2 = A_1 + P_{\text{comm}}V_1 = A_2 + P_{\text{comm}}V_2\,, \tag{4.35}$$

where P is the common pressure. This is fulfilled if $P_{\text{comm}}(V_1 - V_2) = A_1 - A_2$.

We can also express the pressure as $P = -(\partial A/\partial V)_{N,T}$, and therefore, if we use thermodynamic integration, we find that

$$A_1 - A_2 = -\int_0^{V_1} dV\, P + \int_0^{V_2} dV\, P\,. \tag{4.36}$$

Hence the condition for coexistence is

$$P_{\text{comm}}(V_2 - V_1) = \int_{V_1}^{V_2} dV\, P\,. \tag{4.37}$$

This condition means that the area of the rectangle spanned by V_1 and V_2 along the volume axis and 0 and P_{comm} along the pressure axis is equal to the area under the P curve between V_1 and V_2. As can be seen in the figure, the only difference between the two areas is the region where area 1 carves an area into the rectangle and area 2 protrudes from the rectangle: to satisfy (4.37), we simply require that area 1 should equal area 2. In practice, we need to vary P_{comm} until the two areas match. A phase diagram corresponding to the Van der Waals equation of state is shown in the figure.

We used an approximate equation of state for a hard-sphere gas in our derivation, and so, as we argued above, the Van der Waals equation of state is also only reasonable at low density. If, on the other hand, we use the exact equation of state of hard spheres as deduced from computer simulations, then we can compute the 'exact' equation of state of the Van der Waals model (i.e. the equation of state that Van der Waals would
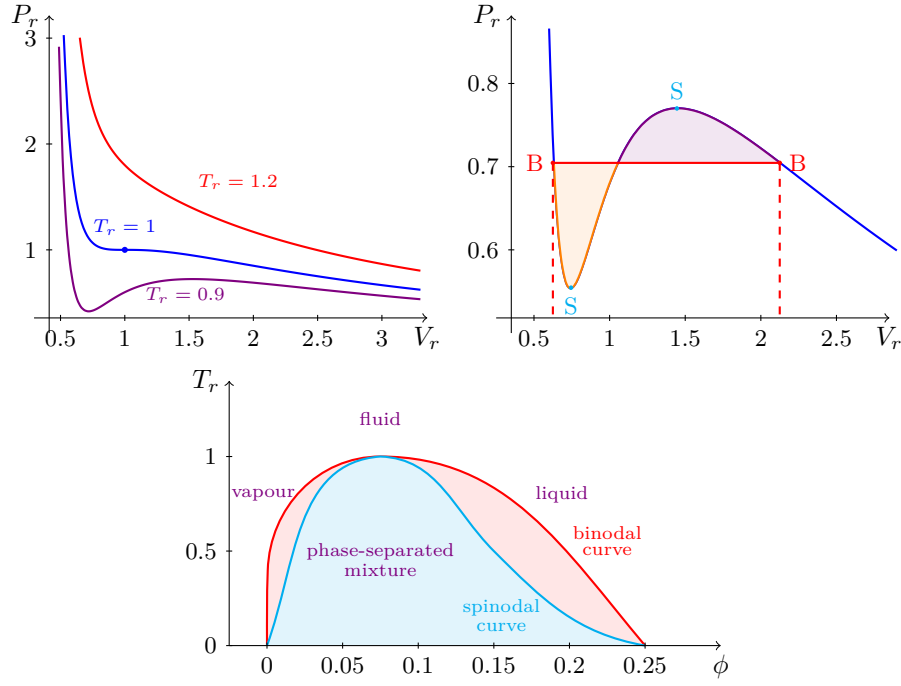
36

Figure 4.6: (a) The pressure as a function of the volume above, at and below the critical temperature. All quantities are scaled by their values at the critical point. (b) Below the critical point, the pressure has a 'Van der Waals loop'. A Maxwell equal-area construction, shown here, is used to find the point of coexistence (labelled 'B'). Spinodals, where the system becomes mechanically unstable, are labelled 'S'. (c) Phase diagram for the Van der Waals model determined using Maxwell constructions.

have given an arm and a leg for). Using this approach, Longuet–Higgins and Widom were the first to compute the true phase diagram of the Van der Waals model.
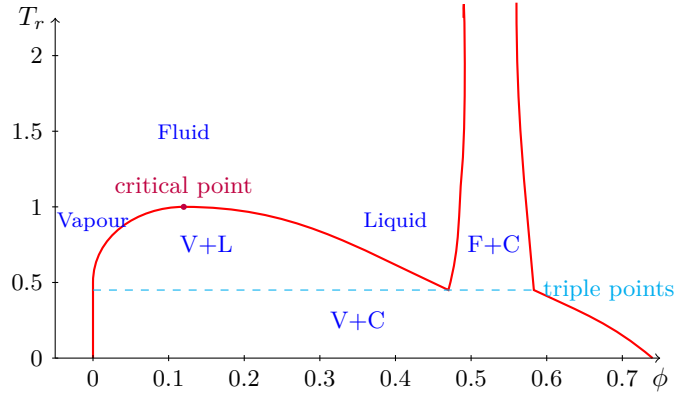

Figure 4.7: Longuet–Higgins–Widom-style phase diagram for the 'exact' Van der Waals model. This phase diagram was computed by using the hard-sphere system as a reference system and adding a weak, long-ranged attractive interaction.

Appendices

A   Thermodynamics

Thermodynamics is a remarkable discipline. It provides us with relations between measurable quantities such as $(\partial S/\partial V)_T = (\partial P/\partial T)_V$. These relations are valid for any substance. But, precisely for this reason, thermodynamics contains no information whatsoever about the underlying microscopic structure of a substance. Thermodynamics preceded statistical mechanics and quantum mechanics, yet not a single thermodynamic relation had to be modified in light of these later developments. The reason is simple: thermodynamics is a phenomenological science. It is based on the properties of matter as we observe them, not upon any theoretical ideas that we may have about matter.

A.1   The First Law

The first law of thermodynamics express the empirical observation that energy is conserved, even though it can be converted into various forms. The internal energy of a system can be changed by either performing work on the system or by transferring an amount of heat.

Law (First law of thermodynamics). Let q be the heat transferred to a system and let w be the work done to the system. Then the change in the internal energy of the system is given by

$$\Delta E = q + w \, . \tag{A.1}$$

A.2   The Second Law

The second law is based on experimental observations.

Law (Kelvin's second law). It is impossible to make an engine that works by converting heat from a single heat bath (i.e. a large reservoir at equilibrium) into work.

This observation is equivalent to another equally empirical observation.

Law (Clausius' second law). Heat can never flow spontaneously (i.e. without work being performed) from a cold reservoir to a warmer reservoir.

This statement is actually a bit more subtle than it seems because, before we have defined temperature, we can only distinguish hotter and colder by looking at the direction of heat flow. What the second law says is that it is never possible to make heat flow spontaneously in the "wrong" direction. How do we get from such a seemingly trivial statement to something as abstract as entropy? This is most easily achieved by introducing the concept of a reversible heat engine.
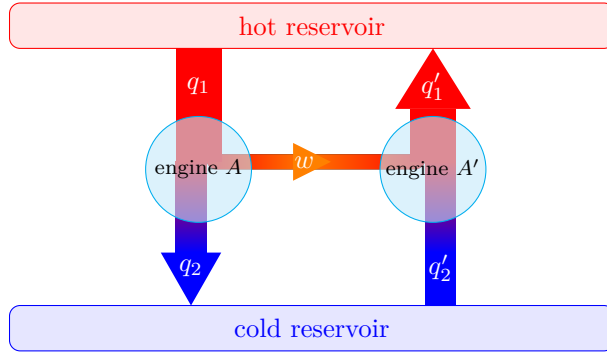
Figure A.1: If the efficiency of the two heat engines are not equal, then this setup violates the second law of thermodynamics.

A.2.1 Heat Engines

A reversible engine is, as the word suggests, an engine that can be operated in reverse. During one cycle (i.e. a sequence of steps that is completed when the engine is returned into its original state), this engine takes in an amount of heat q_1 from a hot reservoir, converts part of it into work w and delivers the remaining amount of heat q_2 to a cold reservoir. The reverse process is that, by performing an amount of work w, we can take an amount of heat q_2 from the cold reservoir and deliver an amount of heat q_1 to the hot reservoir. Reversible engines are an idealisation because in any real engine, there will be additional heat losses. However, the ideal reversible engine can be approximated arbitrarily closely by a real engine if, at every stage, the real engine is sufficiently close to equilibrium. As the engine is returned to its original state at the end of one cycle, its internal energy E has not changed. Hence, the first law tells us that $\Delta E = q_1 - (w + q_2) = 0$ or $q_1 = w + q_2$.

Now consider the efficiency of the engine $\eta := w/q_1$, i.e. the amount of work delivered per amount of heat taken in. At first, one might think that η depends on the precise design of our reversible engine, but this turns out not to be the case: η is the same for all reversible engines operating between the same two reservoirs. To demonstrate this, suppose that we have another reversible engine with efficiency η' that takes in an amount of heat q_1' from the hot reservoir, delivers the same amount of work w, and then delivers an amount of heat q_2' to the cold reservoir, as shown in figure A.1. Suppose $\eta \neq \eta'$, then we use the work generated by the engine with the higher efficiency (say η) to drive the second engine (with efficiency η') in reverse. The amount of heat delivered to the hot reservoir by the second engine is

$$q_1' = \frac{w}{\eta'} = \frac{q_1 \eta}{\eta'} \, . \tag{A.2}$$

As $\eta' < \eta$ by construction, it follows that $q_1' > q_1$. Hence there is a net heat flow from the cold reservoir into the hot reservoir, which contradicts the second law of thermodynamics (in the form "heat can never spontaneously flow from a cold to a hot reservoir"). We have therefore proved by contradiction that the efficiency of all reversible heat engines operating between the same reservoirs is identical.

A.2.2 Absolute Temperatures

Since η does not depend on the specific design of a reversible engine, the only variables it can depend on are the temperatures t_1 and t_2 of the reservoirs it connects, i.e. $\eta = \eta(t_1, t_2)$. We are denoting these temperatures with a lowercase t for the time being because we have not yet defined precisely what we mean by the absolute temperature; the "temperatures" defined here merely need to be well-ordered (i.e. monotonic) and serve to identify which reservoir is colder and which is hotter.

For further convenience, we will also consider the quantity

$$R_{12}(t_1, t_2) := 1 - \eta(t_1, t_2) = \frac{q_1 - w}{q_1} = \frac{q_1}{q_2}\,. \tag{A.3}$$

Suppose now that we have a reversible engine that consists of two stages: engine 1 working between reservoirs A and B, and engine 2 working between reservoirs B and C, as shown in the figure. The three reservoirs are at temperatures t_A, t_B and t_C, respectively. We have R_{AB} and R_{BC}

$$R_{AB} = \frac{q_{1B}}{q_{1A}}\,, \quad \text{and} \quad R_{BC} = \frac{q_{2C}}{q_{2B}}\,.$$

Since by construction, the combined engine $1 + 2$ is a reversible engine working in two stages, reservoir B must be neither a source nor a sink of energy, otherwise the overall efficiency of the engine would depend on how many such stages it is made up of, but we saw earlier that the efficiency of every reversible engine is the same. We must therefore conclude that $q_{1B} = q_{2B}$. The overall value of $R_{\text{tot}} = 1 - \eta_{\text{tot}}$ for the combined reversible engine is thus given by

$$R_{\text{tot}} = \frac{q_{2C}}{q_{1A}} = R_{AB}R_{BC}\,. \tag{A.4}$$

Suppose that in addition, we have another reversible engine, engine 3, that works directly between reservoirs A and C, with

$$R_{AC} = \frac{q_{3C}}{q_{3A}}\,. \tag{A.5}$$

Since the engines are reversible, engine 3 and the combined engine $1 + 2$ must be equally efficient, as we showed above. This means that

$$R_{AB}R_{BC} = R_{AC}\,. \tag{A.6}$$

However, recall that these ratios are functions solely of the temperatures of the reservoirs. The right-hand side does not depend on t_2, and so the product of R_{AB} and R_{BC} must therefore result in the cancellation of this functional dependence. The above equality can only hold in general provided that

$$R_{AB} = \frac{f(t_B)}{f(t_A)}\,, \qquad R_{BC} = \frac{f(t_C)}{f(t_B)} \quad \text{and} \quad R_{AC} = \frac{f(t_C)}{f(t_A)}\,, \tag{A.7}$$

where $f(t)$ is some function of our previously defined temperature t.
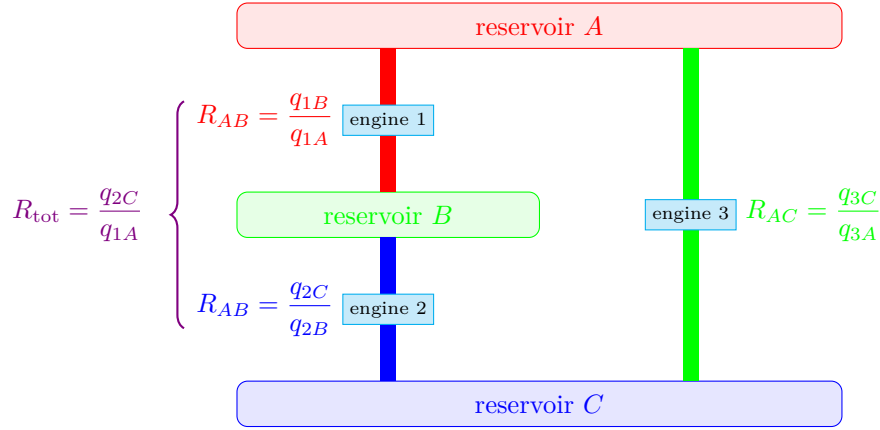
What we do next is to introduce an *absolute* (or *thermodynamic*) *temperature* T given by $T = f(t)$ so that

$$R_{\mathrm{AB}} = \frac{q_B}{q_A} = \frac{T_B}{T_A} \,. \tag{A.8}$$

This well defines the temperatures! We can first claim a certain object have a temperature T_0. Then to measure the temperature of another object, we only need to construct a reversible heat engine between them and measure the efficiency R. Then the temperature ratio of the two objects would be R.

The absolute temperature is defined up to a overall scaling constant. In practice, this is fixed such that 1 unit in the absolute (Kelvin) scale is equal to 1 degree Celsius. But that choice is of course purely historical.

A.2.3 Entropy

The reason we derived everything above is to introduce the entropy. To do so, we rewrite (A.8) as

$$\frac{q_A}{T_A} = \frac{q_B}{T_B} \,, \tag{A.9}$$

where q_A is the heat that flows in reversibly at the high temperature T_A, and q_B is the heat that flows out reversibly at the low temperature T_B. We see therefore that, during a complete cycle, the difference between q_A/T_A and q_B/T_B is zero. Recall that, at the end of a cycle, the internal energy of the system has not changed. Now (A.9) tells us that there is also another quantity that we call *entropy* and that we denote by S that is unchanged when we restore the system to its original state. In the language of thermodynamics, we call S a *state function*. We do not know what S is, but we do know how to compute its change. In the above example, the change in S was given by $\Delta S = (q_A/T_A) - (q_B/T_B) = 0$. In general, the change in entropy of a system due to the reversible addition of an infinitesimal amount of heat δq_{rev} from a reservoir at temperature T is

$$\delta S = \frac{\delta q_{\mathrm{rev}}}{T} \,. \tag{A.10}$$

We also note that S is *extensive*. That means that the total entropy of two non-interacting systems is equal to the sum of the entropies of the individual systems.
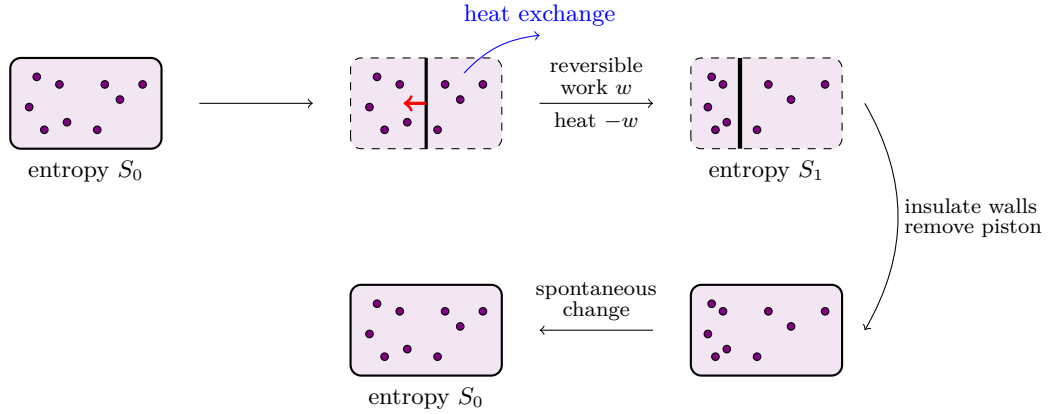
Figure A.2: To drive a system away from equilibrium, e.g. by changing the density, we need to do work. If we do this work reversibly and maintain the total energy of the system, this entails a loss of heat from the system. Hence $S_1 < S_0$. When we remove the constraint, the system spontaneously returns to equilibrium and its entropy is once again S_0. For spontaneous changes, $\Delta S > 0$.

Consider a system with a fixed number of particles N and a fixed volume V. If we transfer an infinitesimal amount of heat δq to this system, then the change in the internal energy of the system, dE, is equal to δq. Therefore

$$\left(\frac{\partial S}{\partial E}\right)_{V,N} = \frac{1}{T}. \tag{A.11}$$

From these, we can get the most famous, albeit not most intuitively obvious, statement of the second law of thermodynamics:

Law (Second law of thermodynamics). Any spontaneous change in a system that exchanges neither heat nor particles with its environment can never lead to a decrease of the entropy.

Hence, at equilibrium, the entropy of an isolated system is at a maximum.

As an example, let us consider a system with an energy E, volume V and number of particles N that is at equilibrium. The entropy of this system is $S_0(E, V, N)$. Suppose we wish to change something in this system; for instance, we increase the density in one half and decrease it in the other. As the system was at equilibrium, this change does not occur spontaneously; in order to effect this change, we must perform a certain amount of work w, for instance by placing a piston in the system and moving it. Let us perform this work reversibly in such a way that E, the total energy of the system, stays constant, as do V and N. The first law tells us that we can only keep E constant if, while we do the work, we allow an amount of heat q to flow out of the system, such that $-q = w$. However, when an amount of heat $-q$ flows out of the system, (A.10) tells us that the entropy S of the system must decrease: $S_1(E, V, N) < S_0(E, V, N)$.

Having completed the change in the system, we insulate the system thermally from the rest of the world, and we remove the constraint that kept the system in its special constrained state. In our piston example, this might entail making a hole in the piston separating the two parts of the system. Now the system goes back spontaneously and

irreversibly to equilibrium. However, since the volume is fixed, no work is done, and since the system is thermally insulated, no heat is transferred. Hence the final energy E is equal to the original energy, and V and N are also unchanged. This means that the system is now back in its original equilibrium state and its entropy is once more equal to $S_0(E, V, N)$.

The entropy change during this spontaneous change is equal to $\Delta S = S_0 - S_1$. But, as $S_1 < S_0$, it follows that $\Delta S > 0$.

A.3   The Third Law

From this point on, we can derive all of thermodynamics, except one law — the so-called third law of thermodynamics.

Law (Third law of thermodynamics). At $T = 0$, the entropy of the equilibrium state of a pure perfectly crystalline substance equals zero.

The third law is not nearly as fundamental as the first two, and in any case, as discussed in the main text, statistical mechanics gives us a more direct interpretation of its meaning.

A.4   Fundamental Equations

From the first law of thermodynamics, we write, in infinitesimal form

$$\mathrm{d}E_{\mathrm{rev}} = đq_{\mathrm{rev}} + đw_{\mathrm{rev}} \,. \tag{A.12}$$

If the system only does reversible PV work, then $đw_{\mathrm{rev}} = -P\mathrm{d}V$. For reversible processes, the *Clausius inequality*, $\mathrm{d}S \geq đq/T$ is an equality (A.10); this is a statement of the second law of thermodynamics. Thus we can write $\mathrm{d}E_{\mathrm{rev}} = T\mathrm{d}S - P\mathrm{d}V$. Since the internal energy is a state function, this equation applies between any equilibrium state, not just along reversible pathways. The equation is known as the *fundamental equation* for the internal energy.

A.4.1   Legendre Transforms

Consider a general function $f(x, y)$, whose total differential is

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x}\right)_y \mathrm{d}x + \left(\frac{\partial f}{\partial y}\right)_x \mathrm{d}y = u\mathrm{d}x + v\mathrm{d}y \,. \tag{A.13}$$

Since u is the derivative of f with respect to x, we say that u and x are *conjugate variables*. Then, consider the function $F = f - (ux)$, whose differential is

$$\mathrm{d}F = \mathrm{d}f - \mathrm{d}(ux) = v\mathrm{d}y - x\mathrm{d}u \,. \tag{A.14}$$

The function F is known as the *Legendre transformation* of f. Evidently, F remains a function of y, but is now a function of u, the conjugate pair of x, rather than of x itself.

To see how this formalism can be applied to thermodynamics, consider the *enthalpy*. When you first introduced it in Part IA, you said that you wanted a state function which corresponded to the heat flow at constant pressure. But how do we know how to define it? The enthalpy is the appropriate measure of "energy" for a system at constant pressure rather than constant volume, so we would like to convert $E = E(S, V)$ into $H = H(S, P)$ so that P is a natural variable of our function H that we can control, and V is a derived property. The enthalpy is the Legendre transform of E, defined as

$$H = E + PV \tag{A.15}$$

so that

$$\mathrm{d}H = T\mathrm{d}S + V\mathrm{d}P\,. \tag{A.16}$$

Similarly, we may change the natural variable from S to T for internal energy and enthalpy, giving *Helmholtz free energy* $A = A(T, V)$ and *Gibbs free energy* $G = G(T, P)$ defined by Legendre transforms

$$A = E - TS\,, \tag{A.17}$$
$$G = H - TS\,, \tag{A.18}$$

so that we have the fundamental equations

$$\mathrm{d}A = -S\mathrm{d}T - P\mathrm{d}V \tag{A.19}$$
$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}P\,. \tag{A.20}$$

The above relations are for the cases where the number of particles N is fixed. If the number of particles can also change, we need to add a further term $+\mu\mathrm{d}N$ to the above fundamental equations, where μ is the *chemical potential*. This allows us to define further Legendre transformations to change the natural variables from N to μ. The one we are interested in is the *Grand potential*

$$\Phi = E - TS - N\mu\,, \tag{A.21}$$

so that

$$\mathrm{d}\Phi = -S\mathrm{d}T - P\mathrm{d}V - N\mathrm{d}\mu\,. \tag{A.22}$$

This completes the definition of the five most common *thermodynamic potentials*: E, H, A, G, Φ. More thermodynamic potentials can be defined by doing further Legendre transforms, but they are of less interests.

A.4.2 Maxwell Relations

Schwarz's theorem guarantees that if $f = f(x, y, \dots)$ have continuous second partial derivatives, then

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\,. \tag{A.23}$$

Since the fundamental are exact differentials, we can apply the Schwarz's theorem to our thermodynamic potentials. For example, from the fundamental equation of the internal energy

$$\mathrm{d}E = T\mathrm{d}S - P\mathrm{d}V + \mu\mathrm{d}N\,, \tag{A.24}$$

we have

$$\left[\frac{\partial}{\partial V}\left(\frac{\partial E}{\partial S}\right)_V\right]_S = \left[\frac{\partial}{\partial S}\left(\frac{\partial E}{\partial V}\right)_S\right]_V\,, \tag{A.25}$$

which gives the relation

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V\,. \tag{A.26}$$

We can get a lot more relations like this from the fundamental equations of different thermodynamic potentials, using the symmetry of partial derivatives with respect to different variables. They are called the *Maxwell relations*. They are not worth memorising as they can be easily derived.

A.4.3 Euler Relations

Since all natural variables of the internal energy E are extensive quantities,

$$\alpha E(N, V, S) = E(\alpha N, \alpha V, \alpha S) \tag{A.27}$$

for $\alpha \in \mathbb{R}$. The total differential is now

$$\mathrm{d}(\alpha E) = \left(\frac{\partial(\alpha E)}{\partial(\alpha N)}\right)_{\alpha V, \alpha S}\mathrm{d}(\alpha N) + \left(\frac{\partial(\alpha E)}{\partial(\alpha V)}\right)_{\alpha N, \alpha S}\mathrm{d}(\alpha V) + \left(\frac{\partial(\alpha E)}{\partial(\alpha S)}\right)_{\alpha N, \alpha V}\mathrm{d}(\alpha S)\,. \tag{A.28}$$

Differentiating both sides with respect to α gives

$$\begin{aligned}
\left(\frac{\partial(\alpha E)}{\partial\alpha}\right)_{N,V,S} &= \left(\frac{\partial(\alpha E)}{\partial(\alpha N)}\right)_{\alpha V, \alpha S}\left(\frac{\partial(\alpha N)}{\partial\alpha}\right)_{N,V,S}\\
&+ \left(\frac{\partial(\alpha E)}{\partial(\alpha V)}\right)_{\alpha N, \alpha S}\left(\frac{\partial(\alpha V)}{\partial\alpha}\right)_{N,V,S}\\
&+ \left(\frac{\partial(\alpha E)}{\partial(\alpha S)}\right)_{\alpha N, \alpha V}\left(\frac{\partial(\alpha S)}{\partial\alpha}\right)_{N,V,S}\,. \tag{A.29}
\end{aligned}$$

Simplify this derivative and set $\alpha = 1$ gives

$$E = \left(\frac{\partial E}{\partial N}\right)_{V,S} N + \left(\frac{\partial E}{\partial V}\right)_{N,S} V + \left(\frac{\partial E}{\partial S}\right)_{N,V} S \tag{A.30}$$

$$= \mu N - PV + TS\,. \tag{A.31}$$

This is the *Euler relation*.[5]

---

[5]It was not discovered by Euler in an investigation of thermodynamics, which did not exist in his day. It is called Euler relation just because Euler's theorem on homogeneous functions leads to it.

Substituting the Euler relation into the expressions for the other main potentials gives

$$A = -PV + \mu N \tag{A.32}$$

$$H = TS + \mu N \tag{A.33}$$

$$G = \mu N \tag{A.34}$$

$$\Phi = -PV \,. \tag{A.35}$$

This is essentially why we said the chemical potential μ is equal to the Gibbs energy per particle.

We may differentiate $G = \mu N$ to get $\mathrm{d}G = N\mathrm{d}\mu + \mu\mathrm{d}N$. From the fundamental equation $\mathrm{d}G = V\mathrm{d}P - S\mathrm{d}P + \mu\mathrm{d}N$, we can get the *Gibbs–Duhem relation*

$$N\mathrm{d}\mu = V\mathrm{d}P - S\mathrm{d}T \,. \tag{A.36}$$

B From Quantum to Classical

It is possible to derive the classical partition function directly from the quantum partition function without resorting to hand-waving. It will also show us why the factor of $1/h$ sits outside the partition function. The derivation is a little tedious, but worth seeing. (Similar techniques are useful in later courses when you first meet the path integral). To make life easier, let's consider a single particle moving in one spatial dimension. It has position operator $\hat{x}$, momentum operator $\hat{p}$ and Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \,. \tag{B.1}$$

If $|n\rangle$ is the energy eigenstate with energy E_n, the quantum partition function is

$$Q = \sum_n e^{-\beta E_n} = \sum_n \left\langle n \middle| e^{-\beta\hat{H}} \middle| n \right\rangle \,, \tag{B.2}$$

where the exponential of an operator $\hat{X}$ is defined via the expansion

$$e^{\hat{X}} = \sum_{i=1}^{\infty} \frac{\hat{X}^n}{n!} \,. \tag{B.3}$$

In what follows, well make liberal use of the fact that we can insert the identity operator anywhere in this expression. We have the resolution of the identities in the position eigenstates $|x\rangle$ and the momentum eigenstates $|p\rangle$:

$$1 = \int \mathrm{d}x \, |x\rangle \langle x| \qquad 1 = \int \mathrm{d}p \, |p\rangle \langle p| \,. \tag{B.4}$$

We insert two copies of the identity built from position eigenstates into our partition function to get

$$Q = \sum_n \left\langle n \middle| \int \mathrm{d}x \, \middle| x \right\rangle \left\langle x \middle| e^{-\beta\hat{H}} \int \mathrm{d}x' \, \middle| x' \right\rangle \langle x'|n\rangle$$

$$= \int \mathrm{d}x \, \mathrm{d}x' \, \left\langle x \middle| e^{-\beta\hat{H}} \middle| x' \right\rangle \sum_n \langle x'|n\rangle \langle n|x\rangle \,. \tag{B.5}$$

And now $\sum_n |n\rangle \langle n|$ is the identity, while $|x'\rangle \langle x| = \delta(x - x')$, so

$$Q = \int \mathrm{d}x \, \langle x | e^{-\beta \hat{H}} | x \rangle . \tag{B.6}$$

We see that the result is to replace the sum over energy eigenstates with a sum (or integral) over position eigenstates. If you wanted, you could play the same game and get the sum over any complete basis of eigenstates of your choosing. This means that we can write the partition function in a basis independent fashion as

$$Q = \operatorname{tr} e^{-\beta \hat{H}} . \tag{B.7}$$

So far, our manipulations could have been done for any quantum system. Now we want to use the fact that we are taking the classical limit. This comes about when we try to factorize $e^{-\beta H}$ into a momentum term and a position term — the trouble is that this isn't always possible when there are matrices (or operators) in the exponent. The Hamiltonian $\hat{H}$ is the sum of a kinetic part $\hat{T}$ and a potential part $\hat{V}$. The kinetic energy eigenstates are also the momentum eigenstates, but potential energy eigenstates are the position eigenstates, so $\hat{H} = \hat{T} + \hat{V}$ is not diagonal in either basis i.e. $\hat{T}$ and $\hat{V}$ does not commute. In fact, we have

$$e^{\hat{A}} e^{\hat{B}} = e^{\hat{A} + \hat{B} + \frac{1}{2}[\hat{A}, \hat{B}] + \cdots} . \tag{B.8}$$

For us $[x, p] = i\hbar$. This means that if we're willing to neglect terms of order $\hbar$ — which is the meaning of taking the classical limit — then we can write

$$e^{-\beta \hat{H}} = e^{-\beta \hat{T}} e^{-\beta \hat{V}} = e^{-\beta \hat{p}^2/2m} e^{-\beta V(\hat{x})} . \tag{B.9}$$

We can now start to replace some of the operators in the exponent, like $V(\hat{x})$, with functions $V(x)$. (Note the subtle but important notational difference!)

$$\begin{aligned}
Q &= \int \mathrm{d}x \, \langle x | e^{-\beta V(\hat{x})} e^{-\beta \hat{p}^2/2m} | x \rangle \\
&= \int \mathrm{d}x \, \mathrm{d}x' \, \mathrm{d}r \, \mathrm{d}r' \, \langle x | e^{-\beta \hat{V}(\hat{x})} | x' \rangle \langle x' | p \rangle \langle p | e^{-\beta \hat{p}^2/2m} | p' \rangle \langle p' | x \rangle \\
&= \int \mathrm{d}x \, \mathrm{d}p \, \langle x | e^{-\beta \hat{V}(\hat{x})} | x \rangle \langle x | p \rangle \langle p | e^{-\beta \hat{p}^2/2m} | p \rangle \langle p | x \rangle \\
&= \int \mathrm{d}x \, \mathrm{d}p \, |\langle x | p \rangle|^2 \, e^{-\beta p^2/2m} e^{-\beta V(x)} , \tag{B.10}
\end{aligned}$$

where we have used the fact that all matrix elements are diagonal. We have that

$$\langle p | q \rangle = \frac{1}{\sqrt{h}} e^{ipx/\hbar} \tag{B.11}$$

is the wavefunction of the momentum wavefunction, so

$$Q = \frac{1}{h} \int \mathrm{d}x \, \mathrm{d}p \, e^{-\beta H(x,p)} . \tag{B.12}$$

C Pressure from Clausius Virial

Very often, there is not just one way of deriving a particular result. In the main text, we derived an expression for the pressure of interacting particles in terms of the inter-particle forces directly from the partition function by using a scaling argument. We show here a derivation using a completely different line of reasoning, originally proposed by Clausius.

Consider a system of N particles in a volume V at total energy E. Define the *virial* of the system to be

$$\mathcal{V} := \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{p}_1 \,. \tag{C.1}$$

This is one half of the time derivate of the moment of inertia I of the system since

$$\frac{1}{2}\frac{\partial I}{\partial t} = \frac{1}{2}\sum_{i=1}^{N} m_i \frac{\partial}{\partial t}\mathbf{r}_i \cdot \mathbf{r}_i = \sum_{i=1}^{N} m_i \mathbf{r}_i \cdot \dot{\mathbf{r}}_i = \mathcal{V} \,. \tag{C.2}$$

Since the particles are constrained in a container, and the total kinetic energy is bounded by E, $\mathcal{V}$ is also bounded both above and below, by $\mathcal{V}_{\text{max}}$ and $\mathcal{V}_{\text{min}}$. At long times $\tau \to \infty$, the time average of the time derivative of $\mathcal{V}$, denoted by $\dot{\mathcal{V}}$, vanishes:

$$\overline{\dot{\mathcal{V}}} = \frac{1}{\tau}\int_0^\tau \dot{\mathcal{V}}(t)\mathrm{d}t = \frac{\mathcal{V}(\tau) - \mathcal{V}(0)}{\tau} \to 0 \text{ as } \tau \to \infty \,. \tag{C.3}$$

We can also evaluate this time derivative explicitly:

$$\overline{\dot{\mathcal{V}}} = \sum_{i=1}^{N} \overline{\dot{\mathbf{r}}_i \cdot \mathbf{p}_i} + \overline{\mathbf{r}_i \cdot \dot{\mathbf{p}}_i} = 0 \,. \tag{C.4}$$

We have $\dot{\mathbf{r}}_i = \mathbf{p}_i/m_i$ and $\dot{\mathbf{p}}_i = \mathbf{f}_i$ by Newton's second law, where $\mathbf{f}_i$ if the force on the particle i, so we can write

$$\sum_{i=1}^{N} \overline{\frac{p_i^2}{m_i}} + \overline{\mathbf{r}_i \cdot \mathbf{f}_i} = 0 \,. \tag{C.5}$$

The first term on the left is simply twice the kinetic energy of the system. Using the equipartition theorem, we can write this as $3Nk_BT$. We thus find that

$$3Nk_BT + \sum_{i=1}^{N} \overline{\mathbf{r}_i \cdot \mathbf{f}_i} = 0 \,. \tag{C.6}$$

The force $\mathbf{f}_i$ acting on a particle can originate either from the intermolecular interactions or interactions with the wall of the container. Let us consider the contribution to the virial due to the interaction with a surface element $\mathrm{d}\mathbf{S}$ of the wall of the container. We denote the position of this surface element by $\mathbf{r}_S$. Whenever a particle collides with this surface element, its co-ordinate must be given by $\mathbf{r}_S$. The contribution to the force virial arising from collisions with this surface element is thus given by

$$\overline{\left(\sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{f}_i\right)}^{(\mathrm{d}\mathbf{S})} = \sum_{i=1}^{N} \overline{\mathbf{r}_S \cdot \mathbf{f}_i^{(\mathrm{d}\mathbf{S})}} = \mathbf{r}_S \cdot \sum_{i=1}^{N} \overline{\mathbf{f}_i^{(\mathrm{d}\mathbf{S})}} = \mathbf{r}_S \cdot \overline{\mathbf{F}_{\text{tot}}^{\mathrm{d}\mathbf{S}}} \,. \tag{C.7}$$

$\overline{\mathbf{F}^{\mathrm{dS}}_{\mathrm{tot}}}$ is the average total force exerted on the particles by the wall of the container — by Newton's third law, this is also the force exerted on the container wall by the particles. This is exactly the force causing the pressure! Therefore we have

$$\overline{\left(\sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{f}_i\right)}^{(\mathrm{d}\mathbf{S})} = \mathbf{r}_S \cdot (-P)\mathrm{d}\mathbf{S}\,. \tag{C.8}$$

We can now integrate over the entire surface of the container to obtain

$$\sum_{i=1}^{N} \overline{\mathbf{r}_i \cdot \mathbf{f}_i} = -P \int \mathrm{d}\mathbf{S} \cdot \mathbf{r}_S\,. \tag{C.9}$$

Using the divergence theorem, we can write

$$-P \int \mathrm{d}\mathbf{S} \cdot \mathbf{r}_S = -P \int \mathrm{d}V \left(\boldsymbol{\nabla} \cdot \mathbf{r}\right) = -3PV\,. \tag{C.10}$$

Therefore,

$$3Nk_BT + \sum_{i=1}^{N} \overline{\mathbf{r}_i \cdot \mathbf{f}_i^{\mathrm{inter}}} - 3PV = 0\,. \tag{C.11}$$

We get the same result

$$P = \frac{Nk_BT}{V} + \frac{1}{3V} \left\langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{f}_i \right\rangle\,, \tag{C.12}$$

where notice that we have replaced the time average by a Boltzmann-weighted average over states at constant N, V and T.

D   Ising and Lattice-Gas Model Equivalence

The Ising model for a lattice of interacting magnetic spins is amongst the most widely used models in physics because many other models can be mapped onto it. As an example, we consider the relation between the Ising model and a lattice gas. The lattice gas model is a simple model used to describe the liquid-vapour transition. Let us first consider the Ising model, in which neighbouring parallel spins have an interaction energy $-J$, whilst anti-parallel spins have an interaction energy $+J$. We denote the value of the spin on lattice site i by $s_i \in \{\pm 1\}$. The Ising Hamiltonian in the presence of a magnetic field B is given by

$$H_{\mathrm{Ising}} = -\frac{J}{2} \sum_{i=1}^{N} \sum_{\langle i,j \rangle} s_i s_j - B \sum_{i=1}^{N} s_i\,. \tag{D.1}$$

The partition function of the Ising model can be written as

$$Q_{\mathrm{Ising}} = \sum_{(s_k)_{k=1}^{N}} \exp\left[-\beta\left(-\frac{J}{2} \sum_{i=1}^{N} \sum_{\langle i,j \rangle} s_i s_j - B \sum_{i=1}^{N} s_i\right)\right]\,, \tag{D.2}$$

where $(s_k)_{k=1}^N$ denote all possible arrangements of spin values $(s_1, \ldots, s_N)$.

Now consider a gas on a lattice with N sites, where there can be at most one particle per lattice site. Neighbouring particles have an interaction energy $-\epsilon$. The Hamiltonian is

$$H_{\mathrm{LG}} = -\frac{\epsilon}{2} \sum_{i=1}^{N} \sum_{\langle i,j \rangle} n_i n_j \,, \tag{D.3}$$

where $n_i \in \{0, 1\}$ denotes the number of particle on site i. The partition function is

$$Q_{\mathrm{LG}}(M) = \sum_{(n_k)_{k=1}^N | M} \exp \left[-\beta \left(-\frac{\epsilon}{2} \sum_{i=1}^{N} \sum_{\langle i,j \rangle} n_i n_j \right) \right] \,, \tag{D.4}$$

where M is the number of particles and $(n_k)_{k=1}^N | M$ is the arrangement of occupation numbers that satisfy $\sum_i n_i = M$. It turns out to be more convenient to consider the grand partition function Ξ instead to remove this awkward constraint.

$$\begin{aligned}
\Xi &= \sum_{M=0}^{N} \sum_{(n_k)_{k=1}^N | M} \exp \left[-\beta \left(-\frac{\epsilon}{2} \sum_{i=1}^{N} \sum_{\langle i,j \rangle} n_i n_j \right) \right] e^{\beta \mu M} \\
&= \sum_{(n_k)_{k=1}^N} \exp \left[-\beta \left(-\frac{\epsilon}{2} \sum_{i=1}^{N} \sum_{\langle i,j \rangle} n_i n_j - \mu \sum_{i=1}^{N} n_i \right) \right] .
\end{aligned} \tag{D.5}$$

We can recognise that, if we identify $s = 2n - 1$, $J = \epsilon$ and $B = m$, the grand partition function of this lattice gas is identical in form to the canonical partition function of an Ising model in a magnetic field.[6]

E   More on Interacting Gases

[7]In this section, we will re-do some of our previous calculation in a more uniformed approach, and extend the results to the beautiful and important cluster expansions.

E.1   Mayer f Functions and Second Virial Coefficients

The Hamiltonian of the interacting gas is

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i>j} U(r_{ij}) \,. \tag{E.1}$$

---

[6]There are a lot of similarities between Ising model and liquid-gas phase transitions on e.g. *critical exponents*. There are beautiful theories behind this, but it is too large to fit in the margin... See Prof. David Tong's notes for something on this.

[7]This section is EXTREMELY mathematically demanding. This is included in my notes because of my personal interests. This is not in the official notes.

We would like to find a general method calculating the coefficient for the *virial expansion*

$$\frac{p}{k_B T} = \frac{N}{V} + B_2(T)\frac{N^2}{V^2} + B_3(T)\frac{N^3}{V^3} + \dots \tag{E.2}$$

As we've seen before, the partition function is

$$Q(N, V, T) = \frac{1}{h^{3N} N!} \int \prod_{i=1}^{N} \mathrm{d}^3\mathbf{p}_i \, \mathrm{d}^3\mathbf{r}_i \, e^{-\beta H}$$

$$= \frac{1}{\Lambda^{3N} N!} \int \prod_i \mathrm{d}^3\mathbf{r}_i \, e^{-\beta \sum_{j<k} U(r_{jk})} \,. \tag{E.3}$$

We still need to do the integral over positions, which looks hard. The interactions mean that the integrals don't factor in any obvious way.

An obvious thing to try is the Taylor expansion, which should be called as *cumulant expansion* in this context.

$$e^{-\beta \sum_{j<k} U(r_{jk})} = 1 - \beta \sum_{j<k} U(r_{jk}) + \frac{\beta^2}{2} \sum_{j<k, l<m} U(r_{jk}) U(r_{lm}) + \dots \tag{E.4}$$

Unfortunately, this is not very useful. We want each term to be smaller than the preceding one. But as $r_{ij} \to 0$, the potential $U(r_{ij}) \to \infty$, which does not look promising for an expansion parameter.

We will instead work with the following quantity called the *Mayer f function*.

$$f(r) := e^{-\beta U(r)} - 1 \,. \tag{E.5}$$

This is a nicer expansion parameter. When the particles are far, $f(r) \to 0$ and when particles are close, $f(r) \to -1$.

Let us define

$$f_{ij} = f(r_{ij}) \,, \tag{E.6}$$

then we can write the partition function as

$$Q(N, V, T) = \frac{1}{\Lambda^{3N} N!} \int \prod_i \mathrm{d}^3\mathbf{r}_i \prod_{j>k}(1 + f_{jk})$$

$$= \frac{1}{\Lambda^{3N} N!} \int \prod_i \mathrm{d}^3\mathbf{r}_i \left(1 + \sum_{j>k} f_{jk} + \sum_{j>k, l>m} f_{jk} f_{lm} + \dots \right) \tag{E.7}$$

The first term simply gives a factor of the volume V for each integral, so we get V^N. The second term has a sum, each element of which is the same. They all look like

$$\int \prod_{i=1}^{N} \mathrm{d}^3\mathbf{r}_i \, f_{12} = V^{N-2} \int \mathrm{d}^3\mathbf{r}_1 \, \mathrm{d}^3\mathbf{r}_2 \, f(r_{12}) = V^{N-1} \int \mathrm{d}^3\mathbf{r} \, f(r) \,, \tag{E.8}$$

where, in the last equality, we have changed the integration variables from $\mathbf{r}_1$ and $\mathbf{r}_2$ to the centre of mass $\mathbf{R} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)$ and the separation $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. You might

worry that the limits of integration change in the integral over $\mathbf{r}$, but the integral over $f(r)$ only picks up contributions from atomic size distances and this is only actually a problem close to the boundaries of the system where it is negligible.

There is a term like this for each pair of particles, so there are in total $\frac{1}{2}N(N-1) \approx \frac{1}{2}N^2$ terms. Then ignoring terms $O(f^2)$, the partition function is given approximately by

$$Q(N, V, T) = \frac{V^N}{\Lambda^{3N} N!} \left(1 + \frac{N^2}{2V} \int \mathrm{d}^3\mathbf{r}\, f(r) + \dots \right) \tag{E.9}$$

$$= Q_{\text{ideal}} \left(1 + \frac{N}{2V} \int \mathrm{d}^3\mathbf{r}\, f(r) + \dots \right)^N, \tag{E.10}$$

where we promoted one power of N from the front of the integral to an overall exponent. Massaging the expression in this way ensures that the free energy is proportional to the number of particles as one would expect:

$$A = -k_B T \ln Q = A_{\text{ideal}} - N k_B T \ln \left(1 + \frac{N}{2V} \int \mathrm{d}^3\mathbf{r}\, f(r) \right). \tag{E.11}$$

From this expression for the free energy, it is clear that we are indeed performing an expansion in density of the gas since the correction term is proportional to N/V. This form of the free energy will give us the second virial coefficient $B_2(T)$.

We can be somewhat more precise about what it means to be at low density in which our expression is valid. The exact form of the integral $\int \mathrm{d}^3\mathbf{r}\, f(r)$ depends on the potential, but for both the Lennard Jones potential and the hard-core repulsion the integral is approximately $\int \mathrm{d}^3\mathbf{r}\, f(r) \sim r_0^3$, where r_0 some measure of the length scale of the potential (We will compute this later). For the expansion to be valid, we want each term with an extra power of f to be smaller than the preceding one. That means that the second term in the argument of the log should be smaller than 1. In other words,

$$\frac{N}{V} \ll \frac{1}{r_0^3}. \tag{E.12}$$

We have a name for the substance packed as closely as in the right hand side — we call them *liquids*. Our expansion is valid for densities of the gas that are much lower than that of the liquid state.

E.2   van der Waals Equation of State

We can use the free energy equation (E.11) to compute the pressure of the gas. Expanding the log gives

$$p = -\frac{\partial F}{\partial V} = \frac{N k_B T}{V} \left(1 - \frac{N}{2V} \int \mathrm{d}^3\mathbf{r}\, f(r) + \dots \right). \tag{E.13}$$

As expected, the pressure deviated from that of an ideal gas. We can characterise this by writing

$$\frac{pV}{N k_B T} = 1 - \frac{N}{2V} \int \mathrm{d}^3\mathbf{r}\, f(r). \tag{E.14}$$

To understand this, we need to calculated the integral. But first, let us look at two trivial examples.

- Repulsion. Suppose that $U(r) > 0$ for all r with $U(\infty) = 0$. Then $f < 0$ and so the pressure will increase.

- Attraction. If $U(r) < 0$ then $f > 0$ and so the pressure will decrease.

Let's try this with some different potentials. We can define a hard core potential with van der Waals attraction

$$U(r) = \begin{cases} \infty & r < r_0 \\ -U_0 \left(\frac{r_0}{r}^6\right) & r \geq r_0 . \end{cases} \tag{E.15}$$

The integral of the Mayer f function is

$$\int \mathrm{d}^3\mathbf{r}\, f(r) = \int_0^{r_0} \mathrm{d}^3 r\, (-1) + \int_{r_0}^{\infty} \mathrm{d}^3\mathbf{r}\, \left(e^{\beta U_0(r_0/r)^6} - 1\right) . \tag{E.16}$$

We will approximate the second integral in the high temperature limit with $\beta U_0 \ll 1$, where $e^{\beta U_0(r_0/r)^6} \approx 1 + \beta U_0(r_0/r)^6$. Then

$$\int \mathrm{d}^3\mathbf{r}\, f(r) = -4\pi \int_0^r \mathrm{d}r\, r^2 + \frac{4\pi U_0}{k_B T} \int_{r_0}^{\infty} \mathrm{d}r\, \frac{r_0^6}{r^4} \tag{E.17}$$

$$= \frac{4\pi r_0^3}{3} \left(\frac{U_0}{k_B T} - 1\right) . \tag{E.18}$$

Insert this into equation (E.14), we have

$$\frac{pV}{N k_B T} = 1 - \frac{N}{V} \left(\frac{a}{k_B T} - b\right) , \tag{E.19}$$

where a and b are constants defined by

$$a = \frac{2\pi r_0^3 U_0}{3} \quad , \quad b = \frac{2\pi r_0^3}{3} . \tag{E.20}$$

We can recognise that this is indeed the second viral coefficient.

We can rearrange this equation to a more useful form

$$k_B T = \frac{V}{N} \left(p + \frac{N^2}{V^2} a\right) \left(1 + \frac{N}{V} b\right)^{-1} . \tag{E.21}$$

As we are working in an expansion in density, we can do another Taylor expansion to get

$$k_B T = \left(p + \frac{N^2}{V^2} a\right) \left(\frac{V}{N} - b\right) . \tag{E.22}$$

This again get the famous *van der Waals equation of states*. This equation is valid at low densities and high temperatures.

It seems that the equation we obtained (E.14) can be used to compute the equation of state for any potential between atoms. However, there are limitations. The integral only converges for the long-range potentials in the form $1/r^n$ when $r \geq 4$. This means that the techniques described above do not work for long-range potentials with fall-off $1/r^3$ or slower. This includes the important case of $1/r$ Coulomb interactions.

E.3 The Cluster Expansion

We will now develop the full expansion and explain how to compute the higher viral coefficients.

Let us go back to our first expression of the partition function in terms of f.

$$Q(N, V, T) = \frac{1}{\Lambda^{3N} N!} \int \prod_i d^3 \mathbf{r}_i \prod_{j>k} (1 + f_{jk}) \tag{E.23}$$

$$= \frac{1}{\Lambda^{3N} N!} \int \prod_i d^3 \mathbf{r}_i \left(1 + \sum_{j>k} f_{jk} + \sum_{j>k, l>m} f_{jk} f_{lm} + \dots \right) \tag{E.24}$$

Above we effectively related the second virial coefficient to the term linear in f: this is the essence of the equation of state equation (E.14). One might think that terms quadratic in f give rise to the third virial coefficient and so on. But we will see that the expansion is more subtle than that.

The expansion in equation (E.24) includes terms of the form $f_{ij} f_{kl} f_{mn} \dots$, where the indices denote pairs of atoms, (i, j) and (k, l) and so on. These pairs may have atoms in common or they may all be different. However, the same pair never appears twice in a given term as you may check by going back to the first line in equation (E.24).

We will introduce a diagrammatic method to keep track of all the terms in the sum. To each term of the form $f_{ij} f_{kl} f_{mn} \dots$ we associate a picture using the following rules:

- Draw N atoms (We actually need pictures with much less than $N \sim 10^{23}$ atoms).

- Draw a line to connect each pair of atoms that appear as indices.

This allow us to never repeat the same pair of indicies.

For example, if $N = 4$, we may have the following pictures for different terms in the expansion:
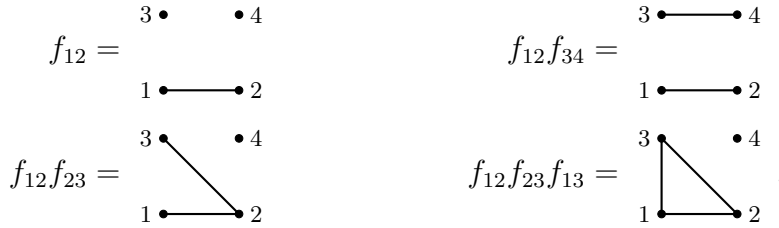
We call these diagrams *graphs*. Each possible graph appears exactly once in the partition function equation (E.24). In other words, the partition function is a sum over all graphs. We still have to do the integrals over all positions r_i. We will denote the integral over graph G to be $W[G]$. Then the partition function is

$$Q(N, V, T) = \frac{1}{\Lambda^{3N} N!} \sum_G W[G]. \tag{E.25}$$

Nearly all the graphs that we can draw will have disconnected components. For example, those graphs that correspond to just a single f_{ij} will have two atoms connected

and the remaining $N - 2$ sitting alone. Those graphs that correspond to $f_{ij}f_{kl}$ fall into two categories: either they consist of two pairs of atoms (like the second example above) or, if (i, j) shares an atom with (k, l), there are three linked atoms (like the third example above). Importantly, the integral over positions $\mathbf{r}_i$ then factorises into a product of integrals over the positions of atoms in disconnected components. This is illustrated by an example with $N = 5$ atoms

$$W \left[\begin{array}{c} \\ \end{array} \right] = \left(\int d^3\mathbf{r}_1 \, d^3\mathbf{r}_2 \, d^3\mathbf{r}_3 \, f_{12}f_{23}f_{31} \right) \left(\int d^3\mathbf{r}_4 \, d^3\mathbf{r}_5 \, f_{45} \right). \quad \text{(E.26)}$$

We call the disconnected components of the graph *clusters*. If a cluster has l atoms, we will call it an *l-cluster*. The $N = 5$ example above has a single 3-cluster and a single 2-cluster. In general, a graph G will split into m_l l-clusters. Clearly, we must have

$$\sum_{l=1}^{N} m_l l = N. \quad \text{(E.27)}$$

The key idea is that we will organise the expansion in such a way that the $(l + 1)$-clusters are much less important than the l-clusters. Then we do not need to draw all $N \sim 10^{23}$ atoms in the graph.

To see how this works, let us focus on the 3-clusters for now. There are 4 different ways that we can have a 3-cluster

Each of these 3-clusters will appear in a graph with any other combination of clusters among the remaining $N-3$ atoms. But since clusters factorise in the partition function, we know that Z must include a factor

$$U_3 = \int d^3\mathbf{r}_1 \, d^3\mathbf{r}_2 \, d^3\mathbf{r}_3 \left(+ + + \right). \quad \text{(E.28)}$$

U_3 contains terms of order f^2 and f^3. It turns out that this is the correct way to arrange the expansion: not in terms of the number of lines in the diagram, which is equal to the power of f, but instead the number of atoms that they connect. The partition function will similarly contain factors associated to all other l-clusters. We define the corresponding integrals as

$$U_l \equiv \int \prod_{i=1}^{l} d^3\mathbf{r}_i \sum_{G \in \{l\text{-cluster}\}} G. \quad \text{(E.29)}$$

Notice that U_1 is simply the integral over space, so $U_1 = V$. The full partition function must be a product of U_l's. The tricky part is to get all the combinatoric factors right

to make sure that you count each graph exactly once. The sum over graphs G that appears in the partition function turns out to be

$$\sum_G W[G] = N! \sum_{\{m_l\}} \prod_l \frac{U_l^{m_l}}{(l!)^{m_l} m_l!} \,. \tag{E.30}$$

The product $N! / \prod_l m_l! (l!)^{m_l}$ counts the number of ways to split the particles into m_l l-clusters, while ignoring the different ways to internally connect each cluster: this is taken into account in the integral U_l.

Let us make a couple checks to make sure that we have got the right answer. First consider $N = 4$ atoms splitting into two 2-clusters (i.e. $m_2 = 2$). There are three such diagrams.

$$f_{12}f_{34} = \begin{array}{c} 3 \bullet\!\!-\!\!\!-\!\!\!-\!\!\bullet 4 \\ \\ 1 \bullet\!\!-\!\!\!-\!\!\!-\!\!\bullet 2 \end{array} \;,\quad f_{13}f_{24} = \begin{array}{cc} 3\bullet & \bullet 4 \\ | & | \\ 1\bullet & \bullet 2 \end{array} \;,\quad f_{14}f_{23} = \begin{array}{c} 3\bullet \quad \bullet 4 \\ \times \\ 1\bullet \quad \bullet 2 \end{array} \;. \tag{E.31}$$

Each of these gives the same answer when integrated, namely U_2^2 so the final result should be $3U_2^2$. We can check this against the relevant terms in equation (E.30), which are $4! U_2^2 / 2!^2 2! = 3U_2^2$ as expected.

Another check: $N = 5$ atoms with $m_2 = m_3 = 1$. All diagrams come in the combinations

$$U_3 U_2 = \int \prod_{i=1}^5 \mathrm{d}^3 \mathbf{r}_i \left(\begin{array}{c}\diagdown\end{array} \; | \; + \; \begin{array}{c}\diagup\end{array} \; | \; + \; \begin{array}{c}\bigwedge\end{array} \; | \; + \; \begin{array}{c}\triangle\end{array} \; | \; \right) \tag{E.32}$$

together with graphs that are related by permutations. The permutations are fully determined by the choice of the two atoms that sit in the pair: there are 10 such choices. The answer should therefore be $10U_3 U_2$. Comparing to equation (E.30), we have $5! U_3 U_2 / 3! 2! = 10 U_3 U_2$ as required.

The end result for the partition function is therefore

$$Q(N, V, T) = \frac{1}{\Lambda^{3N}} \sum_{\{m_l\}} \prod_l \frac{U_l^{m_l}}{(l!)^{m_l} m_l!} \,. \tag{E.33}$$

The problem with computing this sum is that we still have to work out the different ways that we can split N atoms into different clusters. In other words, we still have to obey the constraint equation (E.27). It will be much easier if we do not have to worry about this. Then we could just sum over any m_l. This is exactly what we can do if we work in the grand canonical ensemble where N is not fixed. The grand partition function is

$$\Xi(\mu, V, T) = \sum_N z^N Q(N, V, T) \,, \tag{E.34}$$

where $z := e^{\beta \mu}$ is the *fugacity*. Then we can write

$$\Xi(\mu, V, T) = \sum_N z^N Q(N, V, T) \tag{E.35}$$

$$= \sum_{m_l=0}^{\infty} \prod_{l=1}^{\infty} \left(\frac{z}{\Lambda^3}\right)^{m_l l} \frac{1}{m_l!} \left(\frac{U_l}{l!}\right)^{m_1} \tag{E.36}$$

$$= \prod_{l=1}^{\infty} \exp\left(\frac{U_l z^l}{\Lambda^{3l} l!}\right). \tag{E.37}$$

Define

$$b_l = \frac{\Lambda^3}{V} \frac{U_l}{l! \Lambda^{3l}}. \tag{E.38}$$

Notice in particular that $U_1 = V$ so this definition gives $b_1 = 1$. Then we can write the grand partition function as

$$\Xi(\mu, V, T) = \prod_{l=1}^{\infty} \exp\left(\frac{V}{\Lambda^3} b_l z^l\right) = \exp\left(\frac{V}{\Lambda^3} \sum_{l=1}^{\infty} b_l z^l\right). \tag{E.39}$$

Something interesting happened here. The sum over all the diagrams got rewritten as the exponential over the sum of all the clusters. This is a general lesson which also carries over to many places — like when we are dealing with the Ising model later, or even in fancy stuffs like quantum field theory when dealing with the Feynman diagrams.

Back to the main plot of the story, we can now compute the pressure

$$\frac{pV}{k_B T} = \ln \Xi = \frac{V}{\Lambda^3} \sum_{l=1}^{\infty} b_l z^l, \tag{E.40}$$

and the number of particles

$$\frac{N}{V} = \frac{z}{V} \frac{\partial}{\partial z} \ln \Xi = \frac{1}{\Lambda^3} \sum_{l=1}^{\infty} l b_l z^l. \tag{E.41}$$

Dividing the two gives us the equation of state

$$\frac{pV}{N k_B T} = \frac{\sum_l b_l z^l}{\sum_l l b_l z^l}. \tag{E.42}$$

The only downside is that the equation of state is expressed in terms of z. To massage it into the form of the virial expansion, we need to invert equation (E.41) to get z in terms of the particle density N/V. Equating equation (E.42) with the viral equation

(and defining $B_1 = 1$), we have

$$\sum_{l=1}^{\infty} b_l z^l = \sum_{l=1}^{\infty} B_l \left(\frac{N}{V}\right)^{l-1} \sum_{m=1}^{\infty} m b_m z^m \tag{E.43}$$

$$= \sum_{l=1}^{\infty} \frac{B_l}{\Lambda^{3(l-1)}} \left(\sum_{n=1}^{\infty} n b_n z^n\right)^{l-1} \sum_{m=1}^{\infty} m b_m z^m \tag{E.44}$$

$$= \left[1 + \frac{B_2}{\Lambda^3}(z + 2b_2 z^2 + 3b_3 z^3 + \dots) + \frac{B_3}{\Lambda^6}(z + 2b_2 z^2 + 3b_3 z^3 + \dots) + \dots\right] \tag{E.45}$$

$$\times \left[z + 2b_2 z^2 + 3b_3 z^3 + \dots\right], \tag{E.46}$$

where we have used both $B_1 = 1$ and $b_1 = 1$. Expanding out the left- and right-hand sides to order z^3 gives

$$z + b_2 z^2 + b_3 z^3 = z + \left(\frac{B_2}{\Lambda^3} + 2b_2\right) z^2 + \left(3b_3 + \frac{4b_2 B_2}{\Lambda^3} + \frac{B_3}{\Lambda^3}\right) z^3 + \dots \tag{E.47}$$

Comparing the terms and recollecting the definitions of b_l in terms of U_l gives the second viral coefficient

$$B_2 = -\Lambda^3 b_2 = -\frac{U_2}{2V} = -\frac{1}{2V} \int d^3\mathbf{r}_1 \, d^3\mathbf{r}_2 \, f(\mathbf{r}_1 - \mathbf{r}_2) = -\frac{1}{2} \int d^3\mathbf{r} \, f(r) \,. \tag{E.48}$$

This reproduces the result equation (E.14) we found before.

We now also have an expression for the third coefficient

$$B_3 = \Lambda^6 (4b_2^2 - 2b_3) \tag{E.49}$$

where we still have nasty integrals to do before we have a concrete result.

F   Some Exact Results on the Ising Model

[8]In this section, we will describe some exact results for the Ising model using techniques that do not rely on the mean field approximation. Many of the results that we derive have broader implications for systems beyond the Ising model.

There is an exact solution for the Ising model in $d = 1$ dimension and, when $B = 0$, in $d = 2$ dimensions. Here we will describe the $d = 1$ solution but not the full $d = 2$ solution because it is infamously complicated. We will, however, derive a number of results for the $d = 2$ Ising model which, while falling short of the full solution, nonetheless provide important insights into the physics.

F.1   The Ising Model in $d = 1$ Dimension

We start with the *Ising chain*, the Ising model on a one dimensional line. Here we will see that the mean field approximation fails miserably, giving qualitatively incorrect

---

[8]Again, this section is EXTREMELY mathematically demanding. It is included in my notes because of my personal interests. It is not in the official notes.

results: the exact results shows that there are no phase transitions in the Ising chain. We will write the energy slightly weirdly as

$$E = -J \sum_{i=1}^{N} s_i s_{i+1} - \frac{B}{2} \sum_{i=1}^{N} (s_i + s_{i+1}) \,. \tag{F.1}$$

We will impose periodic boundary conditions, so the spins live on a circular lattice with $s_{N+1} \equiv s_1$. The partition function is then

$$Q = \sum_{s_1 = \pm 1} \cdots \sum_{s_N = \pm 1} \prod_{i=1}^{N} \exp\left(\beta J s_i s_{i+1} + \frac{\beta B}{2} (s_i + s_{i+1}) \right) \,. \tag{F.2}$$

The crucial observation that allows us to solve the problem is that this partition function can be written as a product of matrices. We adopt notation from quantum mechanics and define the 2×2 matrix,

$$\langle s_i | T | s_{i+1} \rangle := \exp\left(\beta J s_i s_{i+1} + \frac{\beta B}{2} (s_i + s_{i+1}) \right) \,. \tag{F.3}$$

The row of the matrix is specified by the value of $s_i = \pm 1$ and the column by $s_{i+1} = \pm 1$. T is known as the *transfer matrix* and, in more conventional notation, is given by

$$\mathsf{T} = \begin{pmatrix} e^{\beta J + \beta B} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J - \beta B} \end{pmatrix} \,. \tag{F.4}$$

The sums over the spins $\sum s_i$ and product over lattice sites $\prod_i$ in the partition function (F.2) simply tell us to multiply the matrices defined in (F.3) and the partition function becomes

$$Q = \langle s_1 | T | s_2 \rangle \langle s_2 | T | s_3 \rangle \ldots \langle s_N | T | s_1 \rangle = \operatorname{tr} T^N \,, \tag{F.5}$$

where the trace arises because we have imposed periodic boundary conditions. To complete the story, we need only compute the eigenvalues of T to determine the partition function. A quick calculation shows that the two eigenvalues of T are

$$\lambda_\pm = e^{\beta J} \cosh \beta B \pm \sqrt{e^{2\beta J} \cosh^2 \beta B - 2 \sinh 2\beta J} \,, \tag{F.6}$$

where, clearly, $\lambda_- < \lambda_+$. The partition function is then

$$Q = \lambda_+^N + \lambda_-^N = \lambda_+^N \left(1 + \frac{\lambda_-^N}{\lambda_+^N} \right) \approx \lambda_+^N \tag{F.7}$$

for very large N.

The partition function Q contains many quantities of interest. In particular, we can use it to compute the magnetisation as a function of temperature when $B = 0$. This, recall, is the quantity which is predicted to undergo a phase transition in the mean field approximation, going abruptly to zero at some critical temperature. In the $d = 1$ Ising model, the magnetisation is given by

$$\langle s \rangle_0 = \frac{1}{N\beta} \frac{\partial \ln Q}{\partial B} \bigg|_{B=0} = \frac{1}{\lambda_+ \beta} \frac{\partial \lambda_+}{\partial B} \bigg|_{B=0} = 0 \,. \tag{F.8}$$

We see that the true physics for $d = 1$ is very different than that suggested by the mean field approximation. When $B = 0$, there is no magnetisation! While the J term in the energy encourages the spins to align, this is completely overwhelmed by thermal fluctuations for any value of the temperature.

There is a general lesson in this calculation: thermal fluctuations always win in one dimensional systems. They never exhibit ordered phases and, for this reason, never exhibit phase transitions. The mean field approximation is bad in one dimension.

F.2 2D Ising Model

Let's now turn to the Ising model in $d = 2$ dimensions. We'll work on a square lattice and set $B = 0$. Rather than trying to solve the model exactly, we'll have more modest goals. We will compute the partition function in two different limits: high temperature and low temperature.

F.2.1 Low Temperature Limit and Peierls Droplets

The partition function is given by the sum over all states, weighted by $e^{-\beta E}$. At low temperatures, this is always dominated by the lowest lying states. For the Ising model, we have[9]

$$Q = \sum_{(s_i)} \exp\left(\beta J \sum_{\langle i,j \rangle} s_i s_j\right) . \tag{F.9}$$

The low temperature limit is $\beta J \to \infty$, where the partition function can be approximated by the sum over the first few lowest energy states. All we need to do is list these states.

The ground states are easy. There are two of them: spins all up or spins all down. For example, the ground state with spins all up looks like

$$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow$$
$$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow$$
$$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow$$

Each of these ground states has energy $E = E_0 = -2NJ$.

The first excited states arise by flipping a single spin. Each spin has $q = 4$ nearest neighbours — denoted by red lines in the example below — each of which leads to an energy cost of $J - (-J) = 2J$. The energy of the first excited states are therefore $E_1 = E_0 + 8J$.

---

[9]From now on, we will directly use $\sum_{\langle i,j \rangle}$ to denote the sum over all neighbouring pairs (which we used to denote as $\sum_i \sum_{\langle i,j \rangle}$) to reduce the cluttering of notations.
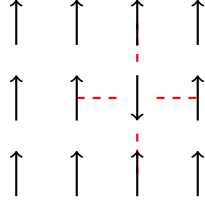
There are, of course, N different spins that we we can flip and, correspondingly, the first energy level has a degeneracy of N.

To proceed, we introduce a diagrammatic method to list the different states. We draw only the broken bonds which connect two spins with opposite orientation and, as in the diagram above, denote these by red lines. We further draw the flipped spins as red dots, the unflipped spins as blue dots. The energy of the state is determined simply by the number of red lines in the diagram. Pictorially, we write the first excited state as

$$E_1 = E_0 + 8J$$
$$\text{Degeneracy} = N$$
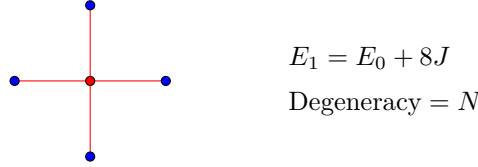
The next lowest state has six broken bonds. It takes the form where the extra

$$E_2 = E_0 + 12J$$
$$\text{Degeneracy} = 2N$$
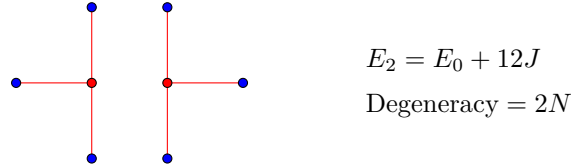
factor of 2 in the degeneracy comes from the two possible orientations (vertical and horizontal) of the graph.

Things are more interesting for the states which sit at the third excited level. These have 8 broken bonds. The simplest configuration consists of two, disconnected, flipped spins

$$E_3 = E_0 + 16J$$
$$\text{Degeneracy} = \tfrac{1}{2}N(N-5)$$

(F.10)
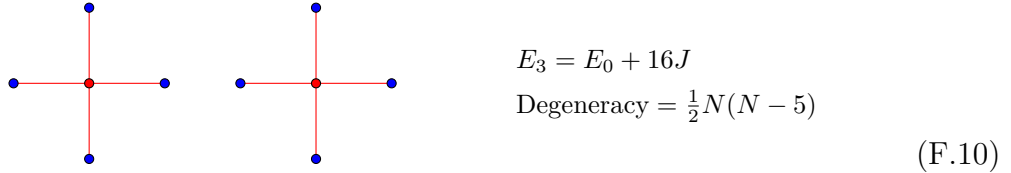
The factor of N in the degeneracy comes from placing the first graph; the factor of $N-5$ arises because the flipped spin in the second graph can sit anywhere apart from on the five vertices used in the first graph. Finally, the factor of $1/2$ arises from the interchange of the two vertices.

There are also three further graphs with the same energy E_3. These are and where the degeneracy comes from the two orientations, and finally where the degeneracy comes from the four orientations.

$$E_3 = E_0 + 16J$$
Degeneracy $= N$

$$E_3 = E_0 + 16J$$
Degeneracy $= 2N$

$$E_3 = E_0 + 16J$$
Degeneracy $= 4N$
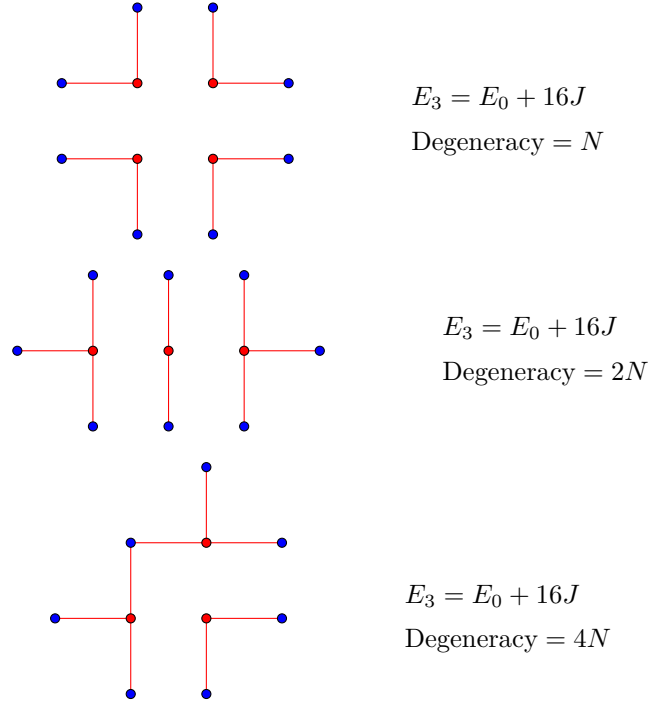
We can of course calculate more excited states but we will stop at here. Adding all the graphs above gives us an expansion of the partition function in power of $e^{-\beta J} \ll 1$. This is

$$Q = 2e^{2N\beta J}\left(1 + Ne^{-8\beta J} + 2Ne^{-12\beta J} + \frac{1}{2}(N^2 + 9N)e^{-16\beta J} + \ldots\right), \qquad \text{(F.11)}$$

where the overall factor of 2 originates from the two ground states of the system. We'll make use of the specific coefficients in this expansion in section F.2.3. Before we focus on the physics hiding in the low temperature expansion, it's worth making a quick comment that something quite nice happens if we take the log of the partition function and expand $\ln(1+x) = x - \frac{1}{2}x^2 + \ldots$, we get

$$\ln Q = \ln 2 + 2N\beta J + Ne^{-8\beta J} + 2Ne^{-12\beta J} + \frac{9}{2}Ne^{-16\beta J} + \ldots \qquad \text{(F.12)}$$

The thing to notice is that the N^2 term in the partition function has precisely cancelled out and $\ln Z$ is proportional to N, which is to be expected since the free energy of the system is extensive. Looking back, we see that the N^2 term was associated to the disconnected diagrams in (F.10). There is actually a general lesson hiding here: the partition function can be written as the exponential of the sum of connected diagrams. We saw exactly the same issue arise in the cluster expansion in (E.39).

Peierls Droplets

Continuing the low temperature expansion provides a heuristic, but physically intuitive, explanation for why phase transitions happen in $d \geq 2$ dimensions but not in $d = 1$. As we flip more and more spins, the low energy states become *droplets*, consisting of a region of space in which all the spins are flipped, surrounded by a larger sea

in which the spins have their original alignment. The energy cost of such a droplet is roughly

$$E \sim 2JL \,, \tag{F.13}$$

where L is the perimeter of the droplet. Notice that the energy does not scale as the area of the droplet since all spins inside are aligned with their neighbours. It is only those on the edge which are misaligned and this is the reason for the perimeter scaling. To understand how these droplets contribute to the partition function, we also need to know their degeneracy. We will now argue that the degeneracy of droplets scales as

$$\text{Degeneracy} \sim e^{\alpha L} \tag{F.14}$$

for some value of α. To see this, consider firstly the problem of a random walk on a 2D square lattice. At each step, we can move in one of four directions. So the number of paths of length L is

$$\#\text{paths} \sim 4^L = e^{L \ln 4} \,. \tag{F.15}$$

Of course, the perimeter of a droplet is more constrained that a random walk. Firstly, the perimeter can't go back on itself, so it really only has three directions that it can move in at each step. Secondly, the perimeter must return to its starting point after L steps. And, finally, the perimeter cannot self-intersect. One can show that the number of paths that obey these conditions is

$$\#\text{paths} \sim e^{\alpha L} \,, \tag{F.16}$$

where $\ln 2 < \alpha < \ln 3$. Since the degeneracy scales as $e^{\alpha L}$, the entropy of the droplets is proportional to L.

The fact that both energy and entropy scale with L means that there is an interesting competition between them. At temperatures where the droplets are important, the partition function is schematically of the form

$$Q \sim \sum_L e^{\alpha L} e^{-2\beta JL} \,. \tag{F.17}$$

For large β (i.e. low temperature) the partition function converges. However, as the temperature increases, one reaches the critical temperature

$$k_B T_c \approx \frac{2J}{\alpha} \tag{F.18}$$

where the partition function no longer converges. At this point, the entropy wins over the energy cost and it is favourable to populate the system with droplets of arbitrary sizes. This is the how one sees the phase transition in the partition function. For temperature above T_c, the low-temperature expansion breaks down and the ordered magnetic phase is destroyed.

We can also use the droplet argument to see why phase transitions don't occur in $d = 1$ dimension. On a line, the boundary of any droplet always consists of just two points. This means that the energy cost to forming a droplet is always $E = 2J$, regardless of the size of the droplet. But, since the droplet can exist anywhere along the line, its degeneracy is N. The net result is that the free energy associated to creating a droplet scales as

$$A = 2J - k_B T \ln N \tag{F.19}$$

and, as $N \to \infty$, the free energy is negative for any $T > 0$. This means that the system will prefer to create droplets of arbitrary length, randomizing the spins. This is the intuitive reason why there is no magnetic ordered phase in the $d = 1$ Ising model.

F.2.2 High Temperature Limit

We now turn to the 2D Ising model in the opposite limit of high temperature. Here we expect the partition function to be dominated by the completely random, disordered configurations of maximum entropy. Our goal is to find a way to expand the partition function in $\beta J \ll 1$.

We again work with zero magnetic field, $B = 0$ and write the partition function as

$$ Q = \sum_{(s_i)} \exp\left(\beta J \sum_{\langle i,j \rangle} s_i s_j \right) = \sum_{(s_i)} \prod_{\langle i,j \rangle} e^{\beta J s_i s_j} . \tag{F.20} $$

There is a useful way to rewrite $e^{\beta J s_i s_j}$ which relies on the fact that the product $s_i s_j$ only takes ± 1. It doesn't take long to check the following identity:

$$ \begin{aligned} e^{\beta J s_i s_j} &= \cosh \beta J + s_i s_j \sinh \beta J \\ &= \cosh \beta J (1 + s_i s_j \tanh \beta J) . \end{aligned} \tag{F.21} $$

Using this, the partition function becomes

$$ \begin{aligned} Q &= \sum_{(s_i)} \prod_{\langle i,j \rangle} \cosh \beta J (1 + s_i s_j \tanh \beta J) \\ &= (\cosh \beta J)^{qN/2} \sum_{(s_i)} \prod_{\langle i,j \rangle} (1 + s_i s_j \tanh \beta J) , \end{aligned} \tag{F.22} $$

where the number of nearest neighbours is $q = 4$ for the 2D square lattice.

With the partition function in this form, there is a natural expansion which suggests itself. At high temperatures $\beta J \ll 1$ which, of course, means that $\tanh \beta J \ll 1$. But the partition function is now naturally a product of powers of $\tanh \beta J$. This is again somewhat analogous to the cluster expansion for the interacting gas. As in the cluster expansion, we will represent the expansion graphically.

We need no graphics for the leading order term. It has no factors of $\tanh \beta J$ and is simply

$$ Q \approx (\cosh \beta J)^{2N} \sum_{(s_i)} 1 = 2^N (\cosh \beta J)^{2N} . \tag{F.23} $$

Let's now turn to the leading correction. Expanding the partition function (F.22), each power of $\tanh \beta J$ is associated to a nearest neighbour pair $\langle i, j \rangle$. We'll represent this by drawing a line on the lattice:

$$ i \circ\!\!-\!\!-\!\!-\!\!\circ\, j \; = s_i s_j \tanh \beta J . \tag{F.24} $$

But a sum over graphs of these kinds vanishes: each factor of $\tanh \beta J$ in (F.22) also comes with a sum over all spins s_i and s_j. And these are $+1$ and -1 which means that they simply sum to zero

$$\sum_{s_i, s_j} s_i s_j = +1 - 1 - 1 + 1 = 0 \,. \tag{F.25}$$

How can we avoid this? The only way is to make sure that we're summing over an even number of spins on each site, since then we get factors of $s_i^2 = 1$ and no cancellations. Graphically, this means that every site must have an even number of lines attached to it. The first correction is then of the form

$$\begin{array}{c} {}^{1} \circ \!\!\!-\!\!\!-\!\!\! \circ\, {}^{2} \\ \Big| \qquad \Big| \\ {}_{4} \circ \!\!\!-\!\!\!-\!\!\! \circ\, {}_{3} \end{array} = (\tanh \beta J)^4 \sum_{(s_i)} s_1 s_2 \, s_2 s_3 \, s_3 s_4 \, s_4 s_1 = 2^4 (\tanh \beta J)^4 \,. \tag{F.26}$$

There are N such terms since the upper left corner of the square can be on any one of the N lattice sites (assuming periodic boundary conditions for the lattice). So including the leading term and first correction, we have

$$Q = 2^N (\cosh \beta J)^{2N} \big(1 + N(\tanh \beta J)^4 + \dots \big) \,. \tag{F.27}$$

We can go further. The next terms arise from graphs of length 6 and the only possibilities are rectangles, oriented as either landscape or portrait. Each of them can sit on one of N sites, giving a contribution

$$\begin{array}{c} \text{landscape} \end{array} + \begin{array}{c} \text{portrait} \end{array} = 2N(\tanh \beta J)^6 \,. \tag{F.28}$$

Things get more interesting when we look at graphs of length 8. We have four different types of graphs. Firstly, there are the trivial, disconnected pair of squares
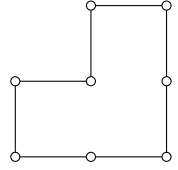
$$\begin{array}{cc} \square & \square \end{array} = \frac{1}{2} N(N - 5)(\tanh \beta J)^8 \,. \tag{F.29}$$

Here the first factor of N is the possible positions of the first square; the factor of $N - 5$ arises because there are 5 locations the second square can't be: overlapping or neighbouring with the first square. Finally, the factor of $1/2$ comes because the two squares are identical. The other graphs of length 8 are a large square, a rectangle and a corner.

$$\square = N(\tanh \beta J)^8 \,, \tag{F.30}$$

$$\square = 2N(\tanh \beta J)^8 \tag{F.31}$$

and

$$\text{(diagram)} = 4N(\tanh \beta J)^8 \,. \tag{F.32}$$

Adding all contributions together gives us the first few terms in high temperature expansion of the partition function

$$Q = 2^N(\cosh \beta J)^{2N}\Big(1 + N(\tanh \beta J)^4 + 2N(\tanh \beta J)^6$$
$$+ \frac{1}{2}(N^2 + 9N)(\tanh \beta J)^8 + \dots\Big) \,. \tag{F.33}$$

You you might have already noticed, everything above are weirdly familiar to everything in the low temperature expansion.

F.2.3 Kramers–Wannier Duality

We have computed the partition function perturbatively in two extreme regimes of low temperature and high temperature. The physics in the two cases is, of course, very different. At low temperatures, the partition function is dominated by the lowest energy states; at high temperatures it is dominated by maximally disordered states. Yet comparing the partition functions at low temperature (F.11) and high temperature (F.33) reveals an extraordinary symmetry between them! The two series agree if we exchange

$$e^{-2\beta J} \longleftrightarrow \tanh \beta J \,. \tag{F.34}$$

Of course, we've only checked the agreement to the first few orders in perturbation. Below we shall prove that this miracle continues to all orders in perturbation theory. The symmetry of the partition function under the interchange (F.34) is known as *Kramers–Wannier duality*. Before we prove this duality, we will first just assume that it is true and extract some consequences.

We can express the statement of the duality more clearly. The Ising model at temperature β is related to the same model at temperature $\tilde{\beta}$, defined as

$$e^{-2\tilde{\beta} J} = \tanh \beta J \,. \tag{F.35}$$

This way of writing things hides the symmetry of the transformation. A little algebra shows that this is equivalent to

$$\sinh 2\tilde{\beta} J = \frac{1}{\sinh 2\beta J} \,. \tag{F.36}$$

This is a hot-cold duality — when $\tilde{\beta}$ is large, $\tilde{\beta}$ is small. Kramers–Wannier duality is the statement that, when $B = 0$, the partition functions of the Ising model at two temperatures are related by

$$Q(\beta) = \frac{2^N(\cosh \beta J)^{2N}}{2e^{2N\tilde{\beta} J}}Q(\tilde{\beta})$$
$$= 2^{N-1}(\cosh \beta J \sinh \beta J)^N Q(\tilde{\beta}) \,. \tag{F.37}$$

This means that if you know the thermodynamics of the Ising model at one temperature, then you also know the thermodynamics at the other temperature. Notice however, that it does not say that all the physics of the two models is equivalent. In particular, when one system is in the ordered phase, the other typically lies in the disordered phase.

One immediate consequence of the duality is that we can use it to compute the exact critical temperature T_c. This is the temperature at which the partition function in singular in the $N \to \infty$ limit. If we further assume that there is just a single phase transition as we vary the temperature, then it must happen at the special self-dual point $\beta = \tilde{\beta}$. This is

$$k_B T = \frac{2J}{\ln(\sqrt{2}+1)} \approx 2.269 J \,. \tag{F.38}$$

The exact solution of Onsager confirms that this is indeed the transition temperature. It's also worth noting that it's fully consistent with the more heuristic Peierls droplet argument (F.18) since $\ln 2 < \ln(1 + \sqrt{2}) < \ln 3$.

Proving the Kramers–Wannier Duality

The key idea that we need can actually be found in the various graphs that we have drawn. You may have realised that the graphs we have drawn in the two expansions are actually the same, albeit drawn differently. For example, consider the two "corner" diagrams
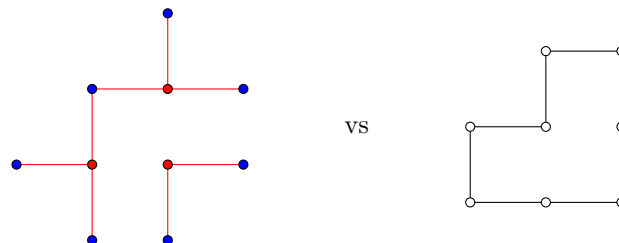
The two graphs are *dual*. The red lines in the first graph intersect the black lines in the second as can be seen by placing them on top of each other:
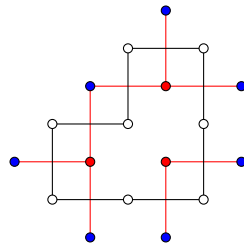
The same pattern occurs more generally: the graphs appearing in the low temperature expansion are in one-to-one correspondence with the dual graphs of the high temperature expansion. Here we will show how this occurs and how one can map the partition functions onto each other.

Let's start by writing the partition function in the form (F.22) that we met in the

high temperature expansion and presenting it in a slightly different way,

$$Q(\beta) = \sum_{(s_i)} \prod_{\langle i,j \rangle} (\cosh \beta J + s_i s_j \sinh \beta J) \tag{F.39}$$

$$= \sum_{(s_i)} \prod_{\langle i,j \rangle} \sum_{k_{ij}=0,1} C_{k_{ij}}(\beta J)(s_i s_j)^{k_{ij}}, \tag{F.40}$$

where we have introduced the rather strange variable k_{ij} associated to each nearest neighbour pair that takes values 0 and 1, together with the functions

$$C_0(\beta J) = \cosh \beta J \qquad \text{and} \qquad C_1(\beta J) = \sinh \beta J. \tag{F.41}$$

The variables in the original Ising model were spins on the lattice sites. The observation that the graphs which appear in the two expansions are dual suggests that it might be profitable to focus attention on the links between lattice sites. Clearly, we have one link for every nearest neighbour pair. If we label these links by l, we can trivially rewrite the partition function as

$$Q = \sum_{k_l=0,1} \prod_l \sum_{(s_i)} C_{k_l}(\beta H)(s_i s_j)^{k_l}. \tag{F.42}$$

Notice that the strange label k_{ij} has now become a variable that lives on the links l rather than the original lattice sites i.

At this stage, we do the sum over the spins s_i. We've already seen that if a given spin, say s_i, appears in a term an odd number of times, then that term will vanish when we sum over the spin. Alternatively, if the spin s_i appears an even number of times, then the sum will give 2. We'll say that a given link l is turned on in configurations with $k_l = 1$ and turned off when $k_l = 0$. In this language, a term in the sum over spin s_i contributes only if an even number of links attached to site i are turned on. The partition function then becomes

$$Q = 2^N \sum_{k_l} \prod_l C_{k_l}(\beta H) \Bigg|_{\text{constrained}}. \tag{F.43}$$

Now we have something interesting. Rather than summing over spins on lattice sites, we're now summing over the new variables k_l living on links. This looks like the partition function of a totally different physical system, where the degrees of freedom live on the links of the original lattice. But there's a catch — that "Constrained" label on the sum. This is there to remind us that we don't sum over all k_l configurations; only those for which an even number of links are turned on for every lattice site. And that's annoying. It's telling us that the k_l aren't really independent variables. There are some constraints that must be imposed.

Fortunately, for the 2D square lattice, there is a simple way to solve the constraint. We introduce yet more variables, $\tilde{s}_i$ which, like the original spin variables, take values ± 1. However, the $\tilde{s}_i$ do not live on the original lattice sites. Instead, they live on the vertices of the dual lattice. For the 2D square lattice, the dual vertices are drawn in the figure. The original lattice sites are in white; the dual lattice sites in black.
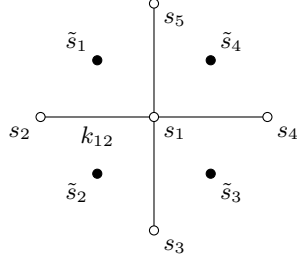
The link variables k_l are related to the two nearest (dual) spin variables $\tilde{s}_i$ as follows:

$$k_{12} = \frac{1}{2}(1 - \tilde{s}_1\tilde{s}_2) \tag{F.44}$$

$$k_{13} = \frac{1}{2}(1 - \tilde{s}_2\tilde{s}_3) \tag{F.45}$$

$$k_{14} = \frac{1}{2}(1 - \tilde{s}_3\tilde{s}_4) \tag{F.46}$$

$$k_{15} = \frac{1}{2}(1 - \tilde{s}_1\tilde{s}_4) \tag{F.47}$$

Notice that we've replaced four variables k_l taking values $0, 1$ with four variables $\tilde{s}_i$ taking values ± 1. Each set of variables gives 2^4 possibilities. However, the map is not one-to-one. It is not possible to construct for all values of k_l using the parameterization in terms of $\tilde{s}_i$. To see this, we need only look at

$$
\begin{aligned}
k_{12} + k_{13} + k_{14} + k_{15} &= 2 - \frac{1}{2}(\tilde{s}_1\tilde{s}_2 + \tilde{s}_2\tilde{s}_3 + \tilde{s}_3\tilde{s}_4 + \tilde{s}_1\tilde{s}_4) \\
&= 2 - \frac{1}{2}(\tilde{s}_1 + \tilde{s}_3)(\tilde{s}_2 + \tilde{s}_4) \\
&= 0, 2 \text{ or } 4 \,.
\end{aligned}
\tag{F.48}
$$

In other words, the number of links that are turned on must be even. But that's exactly what we want! Writing the k_l in terms of the auxiliary spins $\tilde{s}_i$ automatically solves the constraint that is imposed on the sum in (F.43). Moreover, it is simple to check that for every configuration $\{k_l\}$ obeying the constraint, there are two configurations of $\{\tilde{s}_i\}$. This means that we can replace the constrained sum over $\{k_l\}$ with an unconstrained sum over $\{\tilde{s}_i\}$. The only price we pay is an additional factor of $1/2$.

$$Q(\beta) = \frac{1}{2}2^N \sum_{(\tilde{s}_i)} \prod_{\langle i,j\rangle} C_{\frac{1}{2}(1-\tilde{s}_i\tilde{s}_j)}(\beta J) \,. \tag{F.49}$$

Finally, we'd like to find a simple expression for C_0 and C_1 in terms of $\tilde{s}_i$. That's easy enough. We can write

$$C_k(\beta J) = \cosh\beta J \exp(k \ln\tanh\beta J) \tag{F.50}$$

$$= (\sinh\beta J \cosh\beta J)^{1/2} \exp\left(-\frac{1}{2}\tilde{s}_i\tilde{s}_j \ln\tanh\beta J\right) \,. \tag{F.51}$$

Substituting this into our newly re-written partition function gives

$$Q(\beta) = 2^{N-1} \sum_{(\tilde{s}_i)} \prod_{\langle i,j \rangle} (\sinh \beta J \cosh \beta J)^{1/2} \exp\left(-\frac{1}{2}\tilde{s}_i \tilde{s}_j \ln \tanh \beta J\right)$$

$$= 2^{N-1}(\sinh \beta J \cosh \beta J)^N \sum_{(\tilde{s}_i)} \exp\left(-\frac{1}{2}\ln(\tanh \beta J) \sum_{\langle i,j \rangle} \tilde{s}_i \tilde{s}_j\right). \qquad (F.52)$$

This final form of the partition function in terms of the dual spins $\tilde{s}_i$ has exactly the same functional form as the original partition function in terms of the spins s_i! More precisely, we can write

$$Q(\beta) = 2^{N-1}(\sinh 2\beta J)^N Z(\tilde{\beta}), \qquad (F.53)$$

where

$$e^{-2\tilde{\beta}J} = \tanh \beta J \qquad (F.54)$$

as advertised. This is the Kramers–Wannier duality.

This seems to be too out of a tangent from our main theme, but the concept of *duality* is a major feature in theoretical researches. The key idea is that when the temperature gets large there may be a different set of variables in which a theory can be written where it appears to live at low temperature. The same idea often holds in quantum theories, where duality maps strong coupling problems to weak coupling problems.

The duality in the Ising model is special for two reasons: firstly, the new variables $\tilde{s}_i$ are governed by the same Hamiltonian as the original variables s_i. We say that the Ising model is *self-dual*. In general, this need not be the case — the high temperature limit of one system could look like the low-temperature limit of a very different system. Secondly, the duality in the Ising model can be proven explicitly. For most systems, we have no such luck. Nonetheless, the idea that there may be dual variables in other, more difficult theories, is compelling. Commonly studied examples include the exchange particles and vortices in two dimensions, and electrons and magnetic monopoles in three dimensions.

G   An Introduction to Linear-Response Theory