Statistical Mechanics

Níckolas de Aguiar Alves

Abstract: This is my study notebook for the course Statistical Mechanics taught by Prof. Carlos E. Fiore at the University of São Paulo's Institute of Physics (IFUSP) on the second semester of 2022. It was written as a way of keeping up with the course and it corresponds to my lecture notes with some additions from extra bibliography and sometimes slightly different notation. These notes are not endorsed by Prof. Fiore or IFUSP.

Contents

1 What is Equilibrium?

1.1 Phenomenological Description

Thermodynamical systems can be in equilibrium or out of equilibrium. While most systems are not in equilibrium, some of them are. Introductory courses on Thermodynamics and Statistical Mechanics will often focus on equilibrium processes, so in order to understand how to describe nonequilibrium process, it is important for us to first understand what are the differences between each situation.

To obtain a general description, let us begin by considering an isolated system composed of two parts, A and B, as illustrated in Fig. 1.1 on the following page.

Let us consider some thermodynamic variable x_k, with x_k^A being associated to the subsystem A and x_k^B being associated to B. x_k can be either the internal energy U, the volume V, or the number of particles N of the constituents of the system[1]. Notice that due to the system being isolated, x_k^A or x_k^B might change, but their sum $x_k^A + x_k^B$ always remains constant (which is just another way of stating the First Law of Thermodynamics). In a differential formulation, we get

$$\mathrm{d}x_k^A + \mathrm{d}x_k^B = 0. \tag{1.1}$$

---

[1] We could be more general and assume many chemical components inside each subsystem, so we'd need to consider the number of particles of each one separately, but our treatment can be easily generalized.
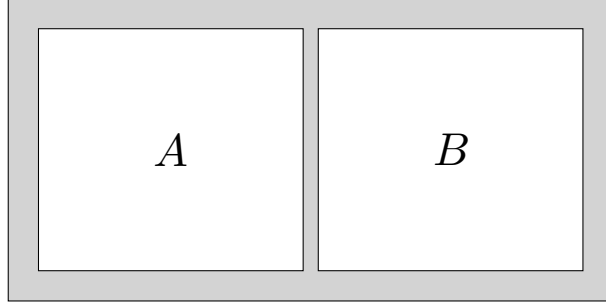
Figure 1.1: *An isolated system composed of two parts, A and B. The wall in between the parts can be removed so that they interact.*

Notice that this expression holds true for the internal energy, volume, and particle number, but might fail for other quantities, such as temperature and entropy.

We know that to each of the system's we can attribute entropy functions $S^A = S^A(U^A, V^A, N^A)$ and $S^B = S^B(U^B, V^B, N^B)$. We know that the total entropy of the system is also given by $S = S^A + S^B$. Hence, we know that

$$dS = dS^A + dS^B \,. \tag{1.2}$$

The Second Law of Thermodynamics ensures that $dS \geq 0$.

Due to the First Law, we know that the extensive quantities associated to A are related to those of B. This leads to the consequence that

$$\left(\frac{\partial S}{\partial x_k^A}\right) = \left(\frac{\partial S^A}{\partial x_k^A}\right) + \left(\frac{\partial S^B}{\partial x_k^A}\right), \tag{1.3a}$$

$$= \left(\frac{\partial S^A}{\partial x_k^A}\right) - \left(\frac{\partial S^B}{\partial x_k^B}\right). \tag{1.3b}$$

This quantity can be interpreted as a sort of "thermodynamical force" (not as a mechanical force). Notice that, if each side is in equilibrium, then this expression leads us to the particular cases

$$\left(\frac{\partial S}{\partial U^A}\right)_{V^A, N^A} = \frac{1}{T^A} - \frac{1}{T^B}, \tag{1.4a}$$

$$\left(\frac{\partial S}{\partial V^A}\right)_{U^A, N^A} = \frac{P^A}{T^A} - \frac{P^B}{T^B}, \tag{1.4b}$$

and

$$\left(\frac{\partial S}{\partial N^A}\right)_{U^A, V^A} = -\frac{\mu^A}{T^A} + \frac{\mu^B}{T^B}, \tag{1.4c}$$

and therefore non-vanishing "thermodynamical forces" will lead to an energy, volume, or particle number flux between the subsystems once the wall is removed.

If we remove the barrier between the subsystems A and B and let them interact, they will then eventually reach equilibrium and a steady state. A different situation is illustrated in Fig. 1.2, where one has two particles, each of them subject to a thermal bath, connected by a spring. Even though the system will eventually evolve to a steady state, it is never in equilibrium, for energy keeps continuously flowing from the hot bath to the cold one.
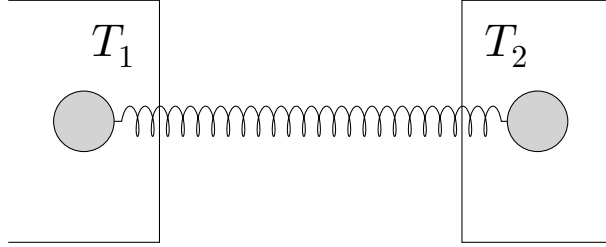

Figure 1.2: *A system composed of two particles connected by a spring. Each particle is subject to a thermal bath. If the temperatures of the baths are unequal, the system is always in nonequilibrium, even though it evolves to a steady state in which energy keeps flowing from the hooter reservoir to the colder one.*

How can we generally distinguish between equilibrium and nonequilibrium then? To see this, let us consider the differential expression

$$\mathrm{d}S = \sum_k \left(\frac{\partial S^A}{\partial x_k^A} - \frac{\partial S^B}{\partial x_k^B} \right) \mathrm{d}x_k^A . \tag{1.5}$$

For thermal equilibrium, we have $\mathrm{d}S = 0$. In this situation, we often have[2]

$$\frac{\partial S^A}{\partial x_k^A} - \frac{\partial S^B}{\partial x_k^B} = 0. \tag{1.6}$$

Since $\mathrm{d}S \geq 0$, if we hold all extensive variables constant but one of them, we'll have that $\frac{\partial S^A}{\partial x_k^A} - \frac{\partial S^B}{\partial x_k^B} > 0$ if, and only if, $\mathrm{d}x_k^A > 0$. Similarly, $\frac{\partial S^A}{\partial x_k^A} - \frac{\partial S^B}{\partial x_k^B} < 0$ if, and only if, $\mathrm{d}x_k^A < 0$ This implies that energy will flow from hot to cold, volume will flow from low pressure to high pressure, particles will flow from large chemical potential to small chemical potential.

Why can't I have $\mathrm{d}x_k^A = 0$?

If $\mathrm{d}S > 0$, then for at least one value of k we have $F_k \equiv \frac{\partial S^A}{\partial x_k^A} - \frac{\partial S^B}{\partial x_k^B} \neq 0$. In this case, notice that we'll have a non-vanishing flux $J_k \equiv \frac{\mathrm{d}x_k^A}{\mathrm{d}t}$. Hence, for $\mathrm{d}S > 0$, we have a flux if, and only if, we have a thermodynamical force.

Is this condition necessary?

This leads us to the following definition: a system is in equilibrium when its entropy production σ, defined as

$$\sigma \equiv \frac{\mathrm{d}S}{\mathrm{d}t} = \sum_k \left(\frac{\partial S^A}{\partial x_k^A} - \frac{\partial S^B}{\partial x_k^B} \right) \frac{\mathrm{d}x_k^A}{\mathrm{d}t} = \sum_k F_k J_k, \tag{1.7}$$

---

[2]Since equilibrium only requires $\mathrm{d}S = 0$, we may still have $\left(\frac{\partial S^A}{\partial x_k^A} - \frac{\partial S^B}{\partial x_k^B} \right) \mathrm{d}x_k^A \neq 0$ for different values of k, as long as their sum vanishes. This can happen, for example, for a Carnot engine.

vanishes, but out of equilibrium when it doesn't.

The importance of entropy production is that it will, for example, generalize to more complicated systems, including small systems (as opposed with systems whose number of particles is comparable to Avogadro's number). However, it will still be positive for equilibrium and strictly positive for nonequilibrium[3].

One can show that in the spring system of Fig. 1.2 on the preceding page the entropy production is proportional to $(T_1 - T_2)^2$, so the system is out of equilibrium whenever the temperatures are different. The direction of the energy flux is determined by the sign of the temperature difference.

In most situations, the entropy production can't be attributed to the subsystem A or B, but rather it is a property of the whole composite system. There is, though, a specific situation in which we can discuss the entropy production due to a subsystem: when one of the subsystems is much larger than the other.

Suppose for example that B is much larger than A. Then we can treat B as a reservoir, with $\frac{\partial S^B}{\partial x_k^B}$ being approximately constant. In this situation, we can write

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \sum_k \left(\frac{\partial S^A}{\partial x_k^A} - \frac{\partial S^B}{\partial x_k^B} \right) \frac{\mathrm{d}x_k^A}{\mathrm{d}t}, \tag{1.8a}$$

$$= \sum_k \frac{\partial S^A}{\partial x_k^A} \frac{\mathrm{d}x_k^A}{\mathrm{d}t} + \phi, \tag{1.8b}$$

$$= \frac{\mathrm{d}S^A}{\mathrm{d}t} + \phi, \tag{1.8c}$$

where ϕ is called the entropy flux and is due only to the bath. It will be given in terms of the heat flux of the reservoir as $\phi = \frac{\dot{Q}}{T}$ (*cf.* the Clausius relation). In this situation, we can interpret

$$\frac{\mathrm{d}S^A}{\mathrm{d}t} = \frac{\mathrm{d}S}{\mathrm{d}t} - \phi = \sigma - \phi \tag{1.9}$$

as the entropy production due to A. Notice that while the Second Law implies $\frac{\mathrm{d}S}{\mathrm{d}t} \geq 0$, there is no restriction on the sign of $\frac{\mathrm{d}S^A}{\mathrm{d}t}$.

At a steady state, we'll have $\frac{\mathrm{d}S^A}{\mathrm{d}t} = 0$, *i.e.*, the system's entropy will no longer depend on time. This then implies $\sigma = \phi$. We can then still have equilibrium or nonequilibrium, depending on whether $\sigma = \phi = 0$ (equilibrium) or $\sigma = \phi > 0$ (nonequilibrium). Notice that $\sigma = 0$ doesn't mean nothing is happening: it means only that on average, the effects cancel out.

Hence, while there are other definitions, the difference between a system being or not in equilibrium is whether the entropy production vanishes or not. As for the difference between equilibrium and steady state, it boils down to the fact that in steady state the quantities no longer change with time on average (*i.e.*, macroscopically).

Should I comment on this?

---

[3]In fact, fluctuations can lead to a measurement of negative entropy production (Crooks 1998, 1999). We shall discuss this later in the course.

Typically, systems will evolve to a steady state, be it in equilibrium or not. To distinguish between them, we can compute the entropy production. But how do we do that?

1.2 Markovian Systems and the Master Equation

In order to be able to do computations, it is interesting for us to reformulate these phenomenological concepts in an stochastic[4] language. This will allow us, for example, to treat more general systems. Some references that might be useful are Tomé and M. J. de Oliveira 2015a,b; Van den Broeck and Esposito 2015.

Since we can't describe large numbers of particles with just Classical or Quantum Mechanics, we often resort to statistical methods to describe large systems. For example, is a system is in thermal equilibrium at fixed inverse temperature β, we assign to its microstates probabilities according to the Gibbs distribution,

$$p_n = \frac{e^{-\beta E_n}}{Z}. \tag{1.10}$$

However, how can we do this in systems that are not in equilibrium?

In order to achieve that, we'll begin by describing the so-called Markovian systems. A Markovian system is a system whose probabilities on a given step depend only on the conditions of the previous step, as opposed to depending on the entire history of the system. Markovian processes are way simpler to describe and they are able to describe a wide range of phenomena, so it is interesting to consider them.

For a Markovian system with l steps, we can write the probability of getting the outcomes $\{n_k\}$ as

$$P_l(n_l, n_{l-1}, \ldots, n_0) = P_l(n_l|n_{l-1}, n_{l-2}, \ldots, n_0)P_{l-1}(n_{l-1}, n_{l-2}, \ldots, n_0), \tag{1.11a}$$

$$= P_l(n_l|n_{l-1})P_{l-1}(n_{l-1}, n_{l-2}, \ldots, n_0) \tag{1.11b}$$

where in the first step we used Bayes' rule and in the second we used the hypothesis that the system is Markovian. If we keep repeating this argument, we get to

$$P_l(n_l, n_{l-1}, \ldots, n_0) = P_l(n_l|n_{l-1})P_{l-1}(n_{l-1}|n_{l-2}) \cdots P_1(n_1|n_0)P_0(n_0). \tag{1.12}$$

Therefore, if we know the transition rates $P_k(n_k|n_{k-1})$ and the initial probability $P_0(n_0)$, we can reconstruct the entire evolution.

Some remarks are in place. For example, due to the properties of conditional probabilities, we have that $P_k(n_k|n_{k-1} \geq 0$ and $\sum_{n_k} P_k(n_k|n_{k-1} = 1$. Furthermore,

$$P_l(n_l) = \sum_{n_0, \ldots, n_{l-1}} P_l(n_l, n_{l-1}, \ldots, n_0), \tag{1.13a}$$

$$= \sum_{n_0, \ldots, n_{l-1}} P_l(n_l|n_{l-1})P_{l-1}(n_{l-1}|n_{l-2}) \cdots P_1(n_1|n_0)P_0(n_0), \tag{1.13b}$$

---

[4]A stochastic system is any system that is not deterministic.

$$
= \sum_{n_0,\dots,n_{l-1}} P_l(n_l|n_{l-1})P_{l-1}(n_{l-1}|n_{l-2})\cdots P_2(n_2|n_1)P_1(n_1,n_0), \tag{1.13c}
$$

$$
= \sum_{n_1,\dots,n_{l-1}} P_l(n_l|n_{l-1})P_{l-1}(n_{l-1}|n_{l-2})\cdots P_2(n_2|n_1)P_1(n_1), \tag{1.13d}
$$

$$
= \sum_{n_{l-1}} P_l(n_l|n_{l-1})P_{l-1}(n_{l-1}). \tag{1.13e}
$$

We see then how in Markov processes the transition rates $P_l(n_l|n_{l-1})$ are important. Hence, we'll describe them in terms of a transition matrix T_{nm} with the properties that

- $T_{nm} \geq 0$,

- $\sum_n T_{nm} = 1$,

- $P_l(n) = \sum_m T_{nm}P_{l-1}(m)$,

which correspond, respectively, to

- $P_k(n_k|n_{k-1}) \geq 0$,

- $\sum_{n_k} P_k(n_k|n_{k-1}) = 1$

- $P_l(n_l) = \sum_{n_{l-1}} P_l(n_l|n_{l-1})P_{l-1}(n_{l-1})$.

So far, we've been describing discrete systems: they have both discrete states and discrete time. Continuous states could require, *e.g.*, the Langevin equation (see, *e.g.*, Salinas 2001; Tomé and M. J. de Oliveira 2015b). However, we will now consider the case with discrete states, but continuous time.

Let us suppose the transitions from state m to n take place along a time $\tau \ll 1$. Then we can represent this in the transition matrix by writing

$$
T_{nm} = \delta_{nm} + \tau W_{nm}, \tag{1.14}
$$

where W_{nm} is assumed to remain finite as we take a limit $\tau \to 0$ later on.

The normalization condition $\sum_n T_{nm} = 1$ now becomes $\sum_n W_{nm} = 0$ (which is possible because the elements of W_{nm} correspond to transition rates, not probabilities). Notice then that $W_{nn} = -\sum_{m \neq n} W_{mn}$. Hence, we find that

$$
P_l(n) = \sum_m T_{nm}P_{l-1}(m), \tag{1.15a}
$$

$$
= \sum_m (\delta_{nm} + \tau W_{nm})P_{l-1}(m), \tag{1.15b}
$$

$$
= \tau \sum_{m \neq n} W_{nm}P_{l-1}(m) + (1 + \tau W_{nn})P_{l-1}(m), \tag{1.15c}
$$

$$
= \tau \sum_{m \neq n} W_{nm}P_{l-1}(m) + \left(1 - \tau \sum_{m \neq n} W_{mn}\right)P_{l-1}(m), \tag{1.15d}
$$

from which we find

$$\frac{P_l(n) - P_{l-1}(n)}{\tau} = \sum_{m \neq n} [W_{nm} P_{l-1}(m) - W_{mn} P_{l-1}(n)]. \tag{1.16}$$

Notice that the left-hand side of this equation describes the probability evolving. The right-hand side has a term that increases probability ($W_{nm} P_{l-1}(m)$), and one that decreases it ($W_{mn} P_{l-1}(n)$). Notice that they are associated with the fact that W_{nm} represents the transition rate for a state m transitioning to n, while W_{mn} represents the transition rate for n to change into m. This can be seen from our previous definition of T_{nm} as $P(n|m)$.

Let now $t = (l-1)\tau$ and take the limit as $\tau \to 0$. We then get to

$$\frac{dP_n}{dt} = \sum_{m \neq n} [W_{nm} P_m(t) - W_{mn} P_n(t)], \tag{1.17}$$

which is known as the master equation. Notice how it resembles a continuity equation with current $J_n = \sum_{m \neq n} [W_{nm} P_m(t) - W_{mn} P_n(t)]$.

Notice that, due to the fact that $\sum_n P_n = 1$ (normalization of probability), we have

$$\sum_n \frac{dP_n}{dt} = \frac{d}{dt} \sum_n P_n = 0, \tag{1.18}$$

implying that

$$\sum_n \sum_{m \neq n} [W_{nm} P_m(t) - W_{mn} P_n(t)] = 0. \tag{1.19}$$

This can also be seen by noticing how

$$\frac{dP_n}{dt} = \sum_{m \neq n} [W_{nm} P_m(t) - W_{mn} P_n(t)], \tag{1.20a}$$

$$= \sum_m W_{nm} P_m(t) \tag{1.20b}$$

and from the previously established fact that $\sum_n W_{nm} = 0$.

In a steady state, we have $\frac{dP_n}{dt} = 0$ for all n. Therefore, we get

$$\sum_{m \neq n} [W_{nm} P_m(t) - W_{mn} P_n(t)] = 0. \tag{1.21}$$

That is, the net current carrying probability around vanishes. However, there are two possibilities: either each term on the sum vanishes independently—which is known as the detailed balance condition—, or at least some of them do not vanish, but their sum does. In the former case, we have equilibrium. In the latter, nonequilibrium. It can be shown that, for equilibrium, the probability distribution tends towards the Gibbs distribution. Reference

1.3 Entropy and Entropy Production in Stochastic Terms

Our goal is now to reformulate thermodynamics concepts in terms of these stochastic ideas. In equilibrium Statistical Mechanics one has to postulate that entropy is given by

$$S = k_B \log \Omega, \tag{1.22}$$

where Ω is the number of accessible microstates of the system. In nonequilibrium Statistical Mechanics, we'll also have to make similar assumptions, that can be motivated, but not proved.

We'll assume that the entropy is given in terms of the probability distribution by

$$S = -k_B \sum_n P_n \log P_n \tag{1.23}$$

(*cf.* the expressions for Gibbs and Shannon entropy).

We'll also make a postulate about the entropy production, which was first made by Schnakenberg (1976). We define

$$\sigma(t) = \frac{k_B}{2} \sum_{n,m} (W_{nm} P_m - W_{mn} P_n) \log \frac{W_{nm} P_m}{W_{mn} P_n}. \tag{1.24}$$

Notice each term of this expression has the form

$$(x - y) \log \frac{x}{y}, \tag{1.25}$$

which is always non-negative, but vanishes if the detailed balance holds.

If we define

$$X_{mn} = \frac{k_B}{2} \log \frac{W_{nm} P_m}{W_{mn} P_n} \tag{1.26}$$

and

$$J_{mn} = W_{nm} P_m - W_{mn} P_n, \tag{1.27}$$

then we can write the entropy production in the form of a sum of products of thermodynamic forces and currents,

$$\sigma(t) = \sum_{n,m} J_{mn} X_{mn}. \tag{1.28}$$

It is worth pointing out that one can describe the currents on a thermodynamic system in a way analogous to electric currents: in terms of Kirchoff laws. See, *e.g.*, Schnakenberg 1976.

We know the time derivative of the entropy and the entropy production are related. It is then interesting to differentiate our prescription for entropy in order to see how exactly

they relate. We find

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -k_B \sum_n \frac{\mathrm{d}P_n}{\mathrm{d}t} \log P_n - k_B \sum_n \frac{P_n}{P_n} \frac{\mathrm{d}P_n}{\mathrm{d}t}, \tag{1.29a}$$

$$= -k_B \sum_n \frac{\mathrm{d}P_n}{\mathrm{d}t} \log P_n - k_B \sum_n \frac{\mathrm{d}P_n}{\mathrm{d}t}, \tag{1.29b}$$

$$= -k_B \sum_n \frac{\mathrm{d}P_n}{\mathrm{d}t} \log P_n, \tag{1.29c}$$

$$= -k_B \sum_{n,m} W_{nm} P_m \log P_n, \tag{1.29d}$$

$$= -k_B \sum_{n,m} (W_{nm} P_m - W_{mn} P_n) \log P_n, \tag{1.29e}$$

$$= \frac{k_B}{2} \sum_{n,m} (W_{nm} P_m - W_{mn} P_n) \log \frac{P_m}{P_n}, \tag{1.29f}$$

where we used $\sum_n P_n = 1$ (a constant) in the third line, $\sum_m W_{mn} = 0$ in the fifth line, and essentially wrote $1 = \frac{1}{2} + \frac{1}{2}$ in the last line.

We see then that we can write

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \sigma(t) - \phi(t), \tag{1.30}$$

where the entropy flux $\phi(t)$ is given by

$$\phi(t) = -\frac{k_B}{2} \sum_{n,m} (W_{nm} P_m - W_{mn} P_n) \log \frac{W_{mn}}{W_{nm}}. \tag{1.31}$$

This expression makes sense as an entropy flux, because there are situations in which one can define thermodynamic quantities[5]—such as temperature and the chemical potential—and in which this expression will reduce to $\phi(t) = \frac{\dot{Q}}{T}$, as we'd expect from the Clausius relation.

It is also interesting to notice that we may write

$$\phi(t) = k_B \sum_{n,m} \left\langle W_{mn} \log \frac{W_{mn}}{W_{nm}} \right\rangle, \tag{1.32}$$

where the angled brackets denote an ensemble average. The interest in this expression is that we can numerically compute ensemble averages without knowledge of the specific probability distribution, and the transition rates W_{nm} are assumed to be given in the problem. Hence, we get an easy to use expression for the entropy flux. If we also have $\frac{\mathrm{d}S}{\mathrm{d}t} = 0$ (steady state), then $\phi(t) = \sigma(t)$ and we have an expression for the entropy production.

---

[5] We do not refer strictly to equilibrium in here, since in that case the detailed balance would hold and the flux would vanish as a consequence. However, there are situations out of the equilibrium which do admit a thermodynamic description. I think Proesmans and Fiore 2019 discusses examples of such systems.

This construction naturally incorporates the Second Law of Thermodynamics (as stated in terms of the entropy production) in its formalism. However, we still haven't discussed the First Law of Thermodynamics.

The internal energy of a system will be given by

$$U = \sum_n E_n(t) P_n(t). \tag{1.33}$$

If we differentiate this expression with respect to time, we find that there are two distinct terms,

$$\frac{\mathrm{d}U}{\mathrm{d}t} = \sum_n \frac{\mathrm{d}E_n}{\mathrm{d}t} P_n + \sum_n E_n \frac{\mathrm{d}P_n}{\mathrm{d}t}. \tag{1.34}$$

The first term corresponds to changing the energy levels of the system, while keeping the probabilities constant. This can be achieved, for example, by varying parameters of the system, such as an external magnetic field, or the system's pressure, and so on. Hence, we can interpret the first term as representing work per time.

The second term represents the change in energy due to variation on the probability of the state n being the system's microstate at a given time. Hence, it corresponds to heat per time.

In consistency with the First Law, we then get

$$\frac{\mathrm{d}U}{\mathrm{d}t} = \dot{W} + \dot{Q}. \tag{1.35}$$

Since we are dealing with stochastic processes, we can actually assign a probability distribution to work. One can show that given a path in phase space, the probabilities for work going forwards and in reverse through such a path will respect

$$\frac{P_F(W)}{P_R(-W)} = e^{W - \Delta F}, \tag{1.36}$$

which is a consequence of the Crooks Fluctuation Theorem[6] (Crooks 1999), which we'll discuss on Section 1.4, where W is the work done along the path and ΔF is the difference in free energy along the transformation.

Notice then that we have

$$\left\langle e^{-W} \right\rangle = \int_{-\infty}^{+\infty} e^{-W} P_F(W) \, \mathrm{d}W, \tag{1.37a}$$

$$= \int_{-\infty}^{+\infty} e^{-W} P_R(-W) e^{W - \Delta F} \, \mathrm{d}W, \tag{1.37b}$$

$$= e^{-\Delta F} \int_{-\infty}^{+\infty} P_R(-W) \, \mathrm{d}W, \tag{1.37c}$$

$$= e^{-\Delta F}, \tag{1.37d}$$

---

[6]Prof. Fiore referred to Eq. (1.36) as the "Jarzynski relation", but I couldn't find references to this. However, Eq. (1.36) seems to be a consequence of what was obtained by Crooks (1999).

which is known as the Jarzynski equality (Jarzynski 1997). Using Jensen's inequality, we have that $\langle e^{-W} \rangle \geq e^{-\langle W \rangle}$, from which it follows that

$$\langle W \rangle \geq \Delta F, \tag{1.38}$$

which is a familiar relation from Thermodynamics (see Fermi 1956, Eq. (112), but beware the sign convention on the definition of work).

1.4 Crooks Fluctuation Theorem

Since we used the Jarzynski relation, which can be seen as a consequence of the Crooks Fluctuation Theorem, we might as well prove the theorem. The result we'll prove is not the same stated on Eq. (1.36) on the previous page, but a related, stronger one. Notice that in Thermodynamics, $W - \Delta F$ is a measure of the reversibility of a process (see Fermi 1956, Sec. 17), vanishing when it is reversible and being positive otherwise. This is extremely similar to entropy production. It is, in fact, entropy production in some sense. For further details, see the original work by Crooks (1999).

> Which sense?

We'll consider a sequence of microscopic states connected by some protocols. In this way, we have a trajectory defined on state space. Our question will be, at first, to find how the probabilities of this trajectory happening forwards and backwards are related. This is illustrated on Fig. 1.3.
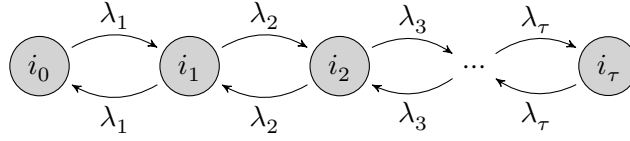

Figure 1.3: *Illustration of the process considered in the Crooks Fluctuation Theorem. Changes from a state i_n to the next are mediated by some protocol λ_{n+1}. We consider the probabilities of both the forward and backwards trajectory happening.*

From each state to the next, there is some difference in energy. We'd like to separately consider the amount due to exchange of heat and the amount due to work. Following our earlier definitions, heat is associated with a change in energy due to change of state, while work is due to a change in protocol. We then define the quantities

$$\Delta W_t \equiv E(i_t, \lambda_{t+1}) - E(i_t, \lambda_t), \tag{1.39}$$

and

> What exactly is $E(i, \lambda)$? Shouldn't it be $E(i)$?

$$\Delta Q_t \equiv E(i_{t+1}, \lambda_t) - E(i_t, \lambda_t). \tag{1.40}$$

Notice that $\Delta W_t + \Delta Q_t = \Delta E_t = E(i_{t+1}, \lambda_{t+1}) - E(i_t, \lambda_t)$, as we would expect. Summing over the discrete times we get to the work and heat along the entire trajectory,

$$W \equiv \sum_{t=0}^{\tau-1} \Delta W_t \quad \text{and} \quad Q \equiv \sum_{t=0}^{\tau-1} \Delta Q_t, \tag{1.41}$$

which satisfy $W + Q = E(i_\tau, \lambda_\tau) - E(i_0, \lambda_0)$. Notice that W and Q are odd functions of the trajectory: if we run the trajectory backwards, the work will be $-W$ and the heat $-Q$, in consistency with $-W - Q = E(i_0, \lambda_0) - E(i_\tau, \lambda_\tau)$.

Let us then compute the probability of each trajectory happening. Assuming the process is Markovian, we get

$$\frac{P_F(i_0, i_1, ..., i_\tau)}{P_R(i_\tau, i_{\tau-1}, ..., i_0)} = \frac{T_{\tau,\tau-1}T_{\tau-1,\tau-2}\cdots T_{1,0}P(i_0)}{T_{\tau-1,\tau}T_{\tau-2,\tau-1}\cdots T_{0,1}P(i_\tau)}, \tag{1.42}$$

where $T_{nm} = P(i_n|i_m)$ are the elements of the transition matrix. To proceed, we'll also assume that the so-called local detailed balance holds:

$$T_{n,n-1}e^{-\beta E(i_{n-1})} = T_{n-1,n}e^{-\beta E(i_n)}. \tag{1.43}$$

Notice that this resembles the expression for the detailed balance $(T_{n,m}P_m = T_{m,n}P_n$, with $n \neq m$ and no sum implied), apart from the facts that we are using the Gibbs distribution for the probability and we're only assuming it holds for "neighbor" states in the chain. Since both sides of Eq. (1.43) are related to the same protocol λ_n, we can write Eq. (1.43) as

$$\frac{T_{n,n-1}}{T_{n-1,n}} = e^{-\beta(E(i_n, \lambda_n) - E(i_{n-1}, \lambda_n))} = e^{-\beta \Delta Q_{n-1}}, \tag{1.44}$$

where we used our previous definition of the heat exchange when going from one state to the other.

Using Eq. (1.44), Eq. (1.42) becomes

$$\frac{P_F(i_0, i_1, ..., i_\tau)}{P_R(i_\tau, i_{\tau-1}, ..., i_0)} = e^{-\beta \Delta Q_{\tau-1}}e^{-\beta \Delta Q_{\tau-2}}...e^{-\beta \Delta Q_0}\frac{P(i_0)}{P(i_\tau)}, \tag{1.45a}$$

$$= e^{-\beta Q}\frac{P(i_0)}{P(i_\tau)}. \tag{1.45b}$$

To proceed, let us now recall that our definition of entropy, Eq. (1.23) on page 9, lets us write[7]

$$S = -k_B \sum_n P_n \log P_n = -k_B \langle \log P_n \rangle, \tag{1.46}$$

which we can interpret as an ensemble average. In particular, it invites us to interpret $-\log P_n$ as a notion of "microscopic entropy". Eq. (1.45) then becomes

$$\frac{P_F(i_0, i_1, ..., i_\tau)}{P_R(i_\tau, i_{\tau-1}, ..., i_0)} = e^{-\beta Q + \log P(i_0) - \log P(i_\tau)}, \tag{1.47a}$$

$$= e^{\Sigma_F}, \tag{1.47b}$$

---

[7]Prof. Fiore actually wrote Eq. (1.46) without the Boltzmann constant. Not sure if it was set to one or something else.

where Σ_F is the entropy production, here defined in analogy with Eq. (1.30) on page 10, which states the entropy production in the variation in entropy added to the incoming heat flux. Notice that Σ_F is odd in trajectory.

Let us now compute the probability of measuring an entropy production Σ when we let the system evolve for a time τ. In the forward trajectory, it will be given by

$$P_F(\Sigma) = \langle \delta(\Sigma - \Sigma_F) \rangle_F, \tag{1.48a}$$

$$= \sum_{i_0, \ldots, i_\tau} \delta(\Sigma - \Sigma_F) P_F(i_0, \ldots, i_\tau). \tag{1.48b}$$

Similarly,

$$P_R(\Sigma) = \langle \delta(\Sigma - \Sigma_R) \rangle_R, \tag{1.49a}$$

$$= \sum_{i_0, \ldots, i_\tau} \delta(\Sigma - \Sigma_R) P_R(i_\tau, \ldots, i_0), \tag{1.49b}$$

$$P_R(-\Sigma) = \sum_{i_0, \ldots, i_\tau} \delta(\Sigma + \Sigma_R) P_R(i_\tau, \ldots, i_0), \tag{1.49c}$$

where we used the fact that the Dirac delta is even. Notice, however, that Eq. (1.47) on the previous page means these expressions imply

$$P_F(\Sigma) = \sum_{i_0, \ldots, i_\tau} \delta(\Sigma - \Sigma_F) P_F(i_0, \ldots, i_\tau), \tag{1.50a}$$

$$= \sum_{i_0, \ldots, i_\tau} \delta(\Sigma - \Sigma_F) P_R(i_\tau, \ldots, i_0) e^{\Sigma_F}, \tag{1.50b}$$

$$= e^\Sigma \sum_{i_0, \ldots, i_\tau} \delta(\Sigma - \Sigma_F) P_R(i_\tau, \ldots, i_0), \tag{1.50c}$$

$$= e^\Sigma \sum_{i_0, \ldots, i_\tau} \delta(\Sigma + \Sigma_R) P_R(i_\tau, \ldots, i_0), \tag{1.50d}$$

$$= e^\Sigma P_R(-\Sigma), \tag{1.50e}$$

where Eq. (1.50d) used the fact that the entropy production is odd in trajectory, and hence $\Sigma_F = -\Sigma_R$. We have thus arrived at the Crooks Fluctuation Theorem,

$$\frac{P_F(\Sigma)}{P_R(-\Sigma)} = e^\Sigma. \tag{1.51}$$

Notice that it means the probability of measuring a negative microscopic entropy production is not zero, but it is exponentially suppressed. This suppression means that it is difficult to explore these effects experimentally. Nevertheless, we see that a bright side of the result is its generality: we only assumed the system to be Markovian and to satisfy the local detailed balance, with no further assumptions on, *e.g.*, being or not in equilibrium.

The Crooks Fluctuation Theorem might seem like a violation of the Second Law of Thermodynamics, but it isn't. Due to the fact that it employs a local form of entropy,

the Second Law does not apply. In fact, notice that if we now deal with averages, we find that

$$\langle e^{-\Sigma} \rangle = \int e^{-\Sigma} P_F(\Sigma) \, d\Sigma \,, \tag{1.52a}$$

$$= \int e^{-\Sigma} e^{+\Sigma} P_R(\Sigma) \, d\Sigma \,, \tag{1.52b}$$

$$= 1. \tag{1.52c}$$

The Jensen inequality then implies that

$$1 = \langle e^{-\Sigma} \rangle \geq e^{-\langle \Sigma \rangle}, \tag{1.53}$$

$$\langle \Sigma \rangle \geq 0, \tag{1.54}$$

meaning the Crooks Fluctuation Theorem doesn't only respect the Second Law, but actually implies it.

2 Equilibrium Statistical Mechanics

Now that we know the differences between equilibrium and nonequilibrium, we'll focus firstly on the simpler case of equilibrium Statistical Mechanics. In this scenario, one usually considers a system in thermal contact with a single reservoir, the properties of which do not depend on time.

When working in equilibrium, the notions of transition rates won't be so useful, for we already know the probability distribution in advance: it is the Gibbs distribution,

$$P_n = \frac{e^{-\beta E_n}}{\sum_m e^{-\beta E_m}}. \tag{2.1}$$

If the probability distribution is known, we don't need to solve the master equation, and as a consequence those ideas become unnecessary. Still, they will be useful when we come back to nonequilibrium.

This will be a lightning review of equilibrium Statistical Mechanics. For further details, hit the books (*e.g.* Kardar 2007b; Pathria and Beale 2022; Reichl 2016; Salinas 2001).

2.1 Canonical Ensemble

In the canonical ensemble, we deal with a fixed temperature T, so the system is assumed to be exchanging heat with a given thermal bath. It is often convenient to define the inverse temperature $\beta = \frac{1}{k_B T}$, in terms of which we may write the partition function,

$$Z = \sum_n e^{-\beta E_n}. \tag{2.2}$$

Notice that we may write the mean energy in terms of the partition function as

$$\langle E \rangle = \sum_n E_n P_n, \tag{2.3a}$$

$$= \sum_n E_n \frac{e^{-\beta E_n}}{Z}, \tag{2.3b}$$

$$= -\frac{1}{Z} \frac{\partial}{\partial \beta} \sum_n e^{-\beta E_n}, \tag{2.3c}$$

$$= -\frac{1}{Z} \frac{\partial Z}{\partial \beta}, \tag{2.3d}$$

$$= -\frac{\partial}{\partial \beta} \log Z, \tag{2.3e}$$

where we employed the Gibbs distribution. Through a similar calculation, we can obtain the second moment of the distribution for the energy by computing

$$\langle E^2 \rangle = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}. \tag{2.4}$$

Notice that these formulae imply that

$$\frac{\partial \langle E \rangle}{\partial \beta} = -\frac{\partial^2}{\partial \beta^2} \log Z, \tag{2.5a}$$

$$= -\frac{\partial}{\partial \beta} \left(\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right), \tag{2.5b}$$

$$= \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2 - \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}, \tag{2.5c}$$

$$= \langle E \rangle^2 - \langle E^2 \rangle. \tag{2.5d}$$

Our interest in this comes from the fact that the specific heat at constant volume can be expressed as

$$c_V = \frac{\partial \langle E \rangle}{\partial T}, \tag{2.6a}$$

$$= \frac{\mathrm{d}\beta}{\mathrm{d}T} \frac{\partial \langle E \rangle}{\partial \beta}, \tag{2.6b}$$

$$= -\frac{1}{k_B T^2} \frac{\partial \langle E \rangle}{\partial \beta}, \tag{2.6c}$$

$$= \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2}. \tag{2.6d}$$

These notions, allied to the hypothesis that the system is extensive (meaning we can write $\langle E \rangle = Nu$ and $c_V = N\tilde{c}_V$, where u and $\tilde{c}_V$ are the energy and specific heat per particle, respectively) let us see the connection of the canonical ensemble with thermodynamics. Firstly, we recall the ergodic hypothesis: time averages of the system correspond to

ensemble averages. This is widely used in Statistical Mechanics, but still lacks a general proof. Assuming it, we notice that, as time passes, the energy of the system will oscillate about $\langle E \rangle$. Nevertheless, the oscillations become negligible in the thermodynamic limit, $N \to +\infty$. Indeed,

$$\lim_{N \to +\infty} \frac{\sqrt{\left\langle (E - \langle E \rangle)^2 \right\rangle}}{\langle E \rangle} = \lim_{N \to +\infty} \frac{\sqrt{\langle E^2 \rangle - \langle E \rangle^2}}{\langle E \rangle}, \tag{2.7a}$$

$$= \lim_{N \to +\infty} \frac{\sqrt{N k_B T^2 \tilde{c}_v}}{N u}, \tag{2.7b}$$

$$= 0. \tag{2.7c}$$

Hence, even though in the canonical ensemble the system is continuously trading energy with the thermal bath, in the thermodynamics limit we can understand $\langle E \rangle$ as the system's internal energy.

Two Level System

As an example, let us briefly consider a two level system, in which each of N non-interacting particles might either be in a ground state with 0 energy or in an excited state with ϵ energy. Since the particles are non-interacting, the partition function can be written as

$$Z = \zeta_1 \cdot \zeta_2 \cdots \zeta_N = \zeta^N, \tag{2.8}$$

where each ζ is understood as a "one-particle partition function". For this problem, we have

$$\zeta = 1 + e^{-\beta\epsilon}. \tag{2.9}$$

Hence, the probability of finding a given particle in the ground state is

$$P_0 = \frac{e^{-\beta \cdot 0}}{\zeta} = \frac{1}{1 + e^{-\beta\epsilon}}, \tag{2.10}$$

while the excited state has

$$P_1 = \frac{e^{-\beta\epsilon}}{\zeta} = \frac{1}{1 + e^{+\beta\epsilon}}. \tag{2.11}$$

These results resemble, but do not match, the Fermi–Dirac distribution. This is expected, since only one fermion can be in each state at a time.

It is interesting to remark that we are considering localized (*i.e.*, distinguishable) particles when making this computation. One can also understand this example as a particular case of the Maxwell–Boltzmann distribution (see, *e.g.*, Reif 2009, Chap. 9).

Care to explain this comment any further?

Let us then use the partition function to compute physical observables. The internal energy per particle is given by

$$u = -\frac{1}{N}\frac{\partial}{\partial\beta}\log Z, \tag{2.12a}$$

$$= -\frac{\partial}{\partial\beta}\log\zeta, \tag{2.12b}$$

$$= \frac{\epsilon e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}}, \tag{2.12c}$$

$$= \frac{\epsilon}{1 + e^{\beta\epsilon}}. \tag{2.12d}$$

It should be remarked that, in the canonical ensemble, the internal energy is not minimized, the Helmholtz free energy is. This is in accordance with the fact from Thermodynamics that a minimum of Helmholtz free energy corresponds to stable thermal equilibrium (see Fermi 1956, Sec. 17).

The Helmholtz free energy per particle will be given by

$$f = -\frac{k_B T}{N}\log Z, \tag{2.13a}$$

$$= -k_B T\log\zeta, \tag{2.13b}$$

$$= -k_B T\log(1 + e^{-\beta\epsilon}). \tag{2.13c}$$

Therefore, we see that the entropy per particle will be given by

$$s = \frac{u - f}{T}, \tag{2.14a}$$

$$= k_B\beta(u - f), \tag{2.14b}$$

$$= \frac{k_B\beta\epsilon}{1 + e^{\beta\epsilon}} + k_B\log(1 + e^{-\beta\epsilon}). \tag{2.14c}$$

Now that we know u, f, and s, we are able to compute any quantities we desire.

In order to better understand the system, we should consider a few "remarkable limits". Namely, how does it behave at low and high temperatures?

The low temperature limit corresponds to picking $k_B T \ll \epsilon$, or, equivalently, to $\beta\epsilon \gg 1$. In this case, the exponential that occurs on the expression for the partition function leads to a high suppression, so that high-energy states become less probable. Hence, we expect for all particles to be on the ground state. Eqs. (2.12) to (2.14) then lead to

$$\lim_{\beta\epsilon \gg 1} u = 0, \quad \lim_{\beta\epsilon \gg 1} f = 0, \quad \lim_{\beta\epsilon \gg 1} s = 0. \tag{2.15}$$

The high temperature limit means $k_B T \gg \epsilon$, which is equivalent to $\beta\epsilon \ll 1$. In this case, there is essentially no suppression on the exponential that goes into the partition function, and hence all states become equally likely. Eqs. (2.12) to (2.14) now imply

$$\lim_{\beta\epsilon \ll 1} u = \frac{\epsilon}{2}, \quad \lim_{\beta\epsilon \ll 1} f = -k_B T\log 2, \quad \lim_{\beta\epsilon \ll 1} s = k_B\log 2. \tag{2.16}$$

Notice that these results means we have essentially $\frac{N}{2}$ in each of the two states. Furthermore, we recover the Boltzmann entropy formula for each particle, since each particle can be in two possible states with equal probabilities.

It is also instructive for us to plot the expressions for the internal energy, entropy, and specific heat as functions of the temperature. These are shown in Fig. 2.1 on the next page. Notice that the specific heat is given by

$$c = \frac{\partial u}{\partial T}, \tag{2.17a}$$

$$= -k_B \beta^2 \frac{\partial u}{\partial \beta}, \tag{2.17b}$$

$$= \frac{k_B \beta^2 \epsilon^2 e^{\beta\epsilon}}{\left(1 + e^{\beta\epsilon}\right)^2}. \tag{2.17c}$$

It is particularly interesting that the specific heat attains a finite maximum. This feature is a general property of systems with finitely many discrete states and it is known as the Schottky anomaly.

Quantum Harmonic Oscillators

Let us next study the case of N quantum harmonic oscillators subject to a thermal bath at inverse temperature $\beta = \frac{1}{k_B T}$. For a single oscillator, we have the energy

$$\epsilon_n = \hbar\omega_0 \left(n + \frac{1}{2}\right), \tag{2.18}$$

where n is a non-negative integer. The partition function is then given by

$$Z = \sum_{\{\text{states}\}} e^{-\beta \sum_{i=0}^{N} \epsilon_i}, \tag{2.19a}$$

$$= \left(\sum_{\{i\}} e^{-\beta\epsilon_i}\right)^N, \tag{2.19b}$$

where we used the assumption that the different oscillators do not interact[1].

We can then obtain the partition function by dealing with a one-particle partition function ζ. It will be given by

$$\zeta = \sum_{n=0}^{+\infty} e^{-\beta\hbar\omega_0\left(n+\frac{1}{2}\right)}, \tag{2.20a}$$

$$= \frac{e^{-\frac{\beta\hbar\omega_0}{2}}}{1 - e^{-\beta\hbar\omega_0}}. \tag{2.20b}$$

---

[1]Interacting systems are more difficult to deal with. Later in the course, we'll learn for example about the transfer matrix approach to deal with them.
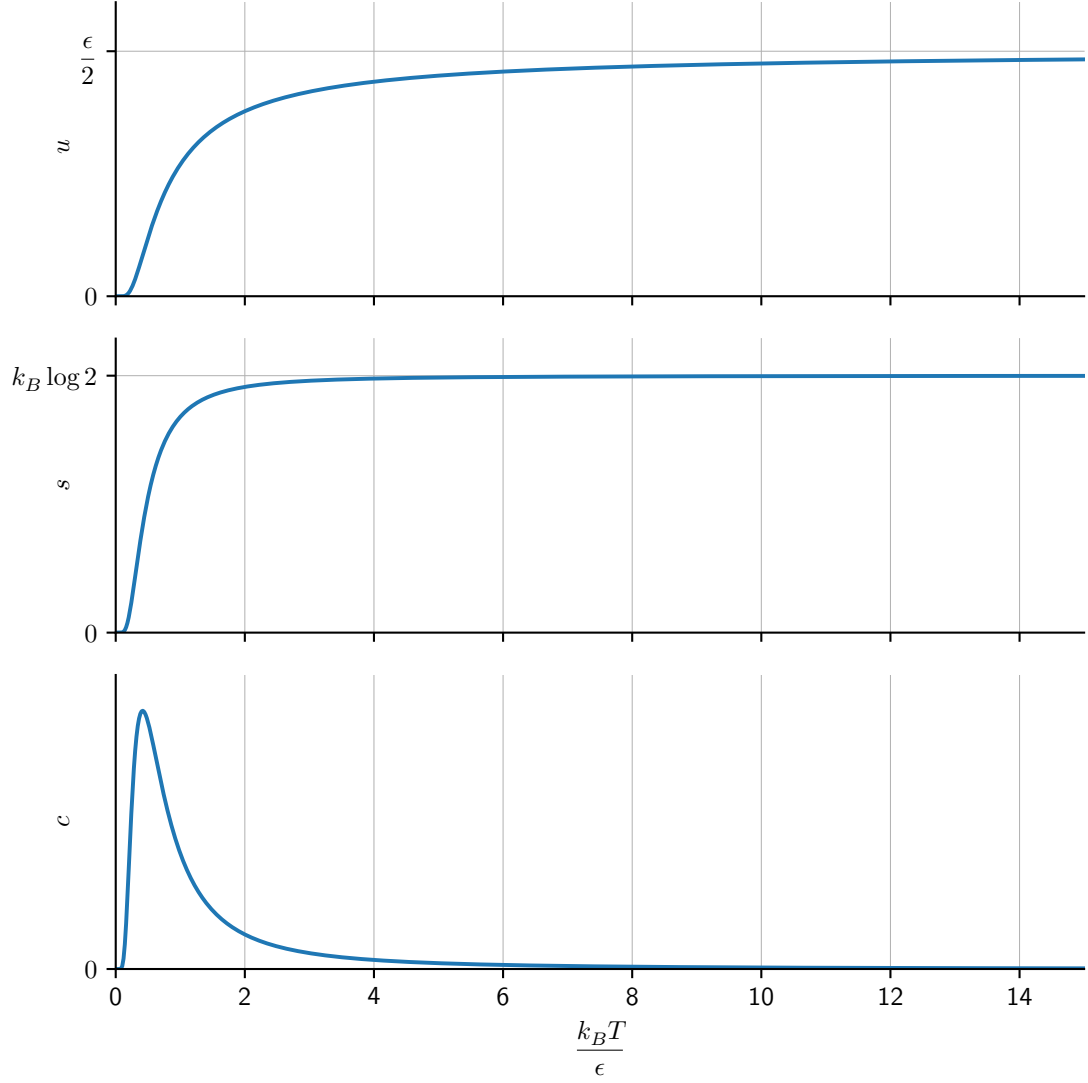
Figure 2.1: *Internal energy per particle u, entropy per particle s, and specific heat per particle c as functions of temperature for the two-level system. Notice that the specific heat attains a finite maximum. This is a general feature of systems with finitely many discrete levels and is known as the Schottky anomaly.*

Therefore,

$$Z = \left(\frac{e^{-\frac{\beta\hbar\omega_0}{2}}}{1 - e^{-\beta\hbar\omega_0}} \right)^N . \tag{2.21}$$

From the partition function we can obtain the internal energy, the Helmholtz free energy, and the entropy just as before. Their expressions per particle will be given by

$$u = \frac{\hbar\omega_0}{2} + \frac{\hbar\omega_0}{e^{\beta\hbar\omega_0} - 1}, \tag{2.22}$$

$$f = \frac{\hbar\omega_0}{2} + k_B T \log(1 - e^{-\beta\hbar\omega_0}), \tag{2.23}$$

and

$$s = \frac{\hbar\omega_0}{T(e^{\beta\hbar\omega_0} - 1)} - k_B \log(1 - e^{-\beta\hbar\omega_0}). \tag{2.24}$$

For low temperatures ($k_B T \ll \hbar\omega_0$ or $\beta\hbar\omega_0 \gg 1$), Eqs. (2.22) to (2.24) lead to the limiting behaviour

$$\lim_{\beta\hbar\omega_0 \gg 1} u = \frac{\hbar\omega_0}{2}, \quad \lim_{\beta\hbar\omega_0 \gg 1} f = \frac{\hbar\omega_0}{2}, \quad \lim_{\beta\hbar\omega_0 \gg 1} s = 0, \tag{2.25}$$

which matches our expectation that all oscillators should be at the ground state.

The specific heat is given by

$$c = \frac{k_B \beta^2 \hbar^2 \omega_0^2 e^{\beta\hbar\omega_0}}{\left(e^{\beta\hbar\omega_0} - 1\right)^2}, \tag{2.26}$$

which implies

$$\lim_{\beta\hbar\omega_0 \gg 1} c = k_B \beta^2 \hbar^2 \omega_0^2 e^{-\beta\hbar\omega_0}. \tag{2.27}$$

This model was originally proposed by Einstein to explain how the specific heat of solids vanishes at low temperatures, and hence we see there is a qualitative agreement between experiment and theory. Nevertheless, there is a quantitative disagreement with experiment, since the correct behaviour at low temperatures should be $\sim T^3$ instead of an exponential decay. This was later corrected by Debye by introducing interactions in the model (for a more detailed account, see Kardar 2007b, Sec. 6.2; Pathria and Beale 2022, Sec. 7.4).

At high temperatures ($k_B T \gg \hbar\omega_0$ or $\beta\hbar\omega_0 \ll 1$), Eqs. (2.22) to (2.24) imply

$$\lim_{\beta\hbar\omega_0 \ll 1} u = k_B T, \quad \lim_{\beta\hbar\omega_0 \ll 1} f = k_B T \log(\beta\hbar\omega_0), \quad \lim_{\beta\hbar\omega_0 \ll 1} s = k_B \left(1 + \log\left(\frac{k_B T}{\hbar\omega_0} \right) \right), \tag{2.28}$$

while Eq. (2.26) leads to the specific heat at high temperatures

$$\lim_{\beta\hbar\omega_0 \ll 1} c = k_B, \tag{2.29}$$

in accordance with the Dulong–Petit law, that states the specific heat of a solid at high temperatures should tend to a constant (this can be understood under the light of the equipartition theorem, which is discussed, *e.g.*, in Pathria and Beale 2022, Sec. 3.7; Salinas 2001, Sec. 6.3).

It is remarkable that, in Eq. (2.28) on the previous page, the expressions for the entropy and Helmholtz free energy still involve $\hbar$, since the high temperature limit corresponds to the classical limit. The occurrence of $\hbar$, however, is a necessity already in classical statistical mechanics in order to ensure that it is recovered by the quantum mechanical calculations.

The plots for u, s, and c for the system of harmonic oscillators are shown in Fig. 2.2 on the following page.

Perhaps discuss this in greater detail

2.2 Classical Statistical Physics

Motivated by the weird occurrence of an $\hbar$ in an expression that should be classical (Eq. (2.28) on the previous page), let us take a moment to discuss classical statistical physics.

In classical systems, the states live in the phase space, which can be described in terms of generalized coordinates and their conjugate momenta, q and p, respectively. The partition function for a system with Hamiltonian $\mathcal{H}$ can then be in principle be written as

$$Z = \sum_{\{\sigma\}} e^{-\beta \mathcal{H}(\sigma)} \stackrel{?}{=} \int e^{-\beta \mathcal{H}(q,p)} \, \mathrm{d}^N q \, \mathrm{d}^N p \,, \tag{2.30}$$

where σ denotes an arbitrary state and denote $q = (q^1, q^2, \ldots, q^N)$ and $p = (p_1, p_2, \ldots, p_N)$, so that our notation allows for an arbitrary number of degrees of freedom. The question mark is in there because there is still an issue: the units on Eq. (2.30) do not match, since the measure of the integral has its own units.

To solve this issue, we'll introduce an arbitrary constant h with dimensions of action such that we can now write

$$Z = \int e^{-\beta \mathcal{H}(q,p)} \frac{\mathrm{d}^N q \, \mathrm{d}^N p}{h^N}. \tag{2.31}$$

This constant can then be fixed by matching the classical predictions with the high temperature limits of the quantum predictions. As the notation suggests, it will indeed be Planck's constant.

A way of interpreting this is that h gives us a measure of "granularity" of phase space in the sense that it converts the volume of phase space and the actual number of states.

Let us now consider a couple of examples.

Classical Harmonic Oscillators

Let us begin with a system of N classical harmonic oscillators, which we'll then expect to reproduce the high energy behavior of our quantum calculation. The Hamiltonian is
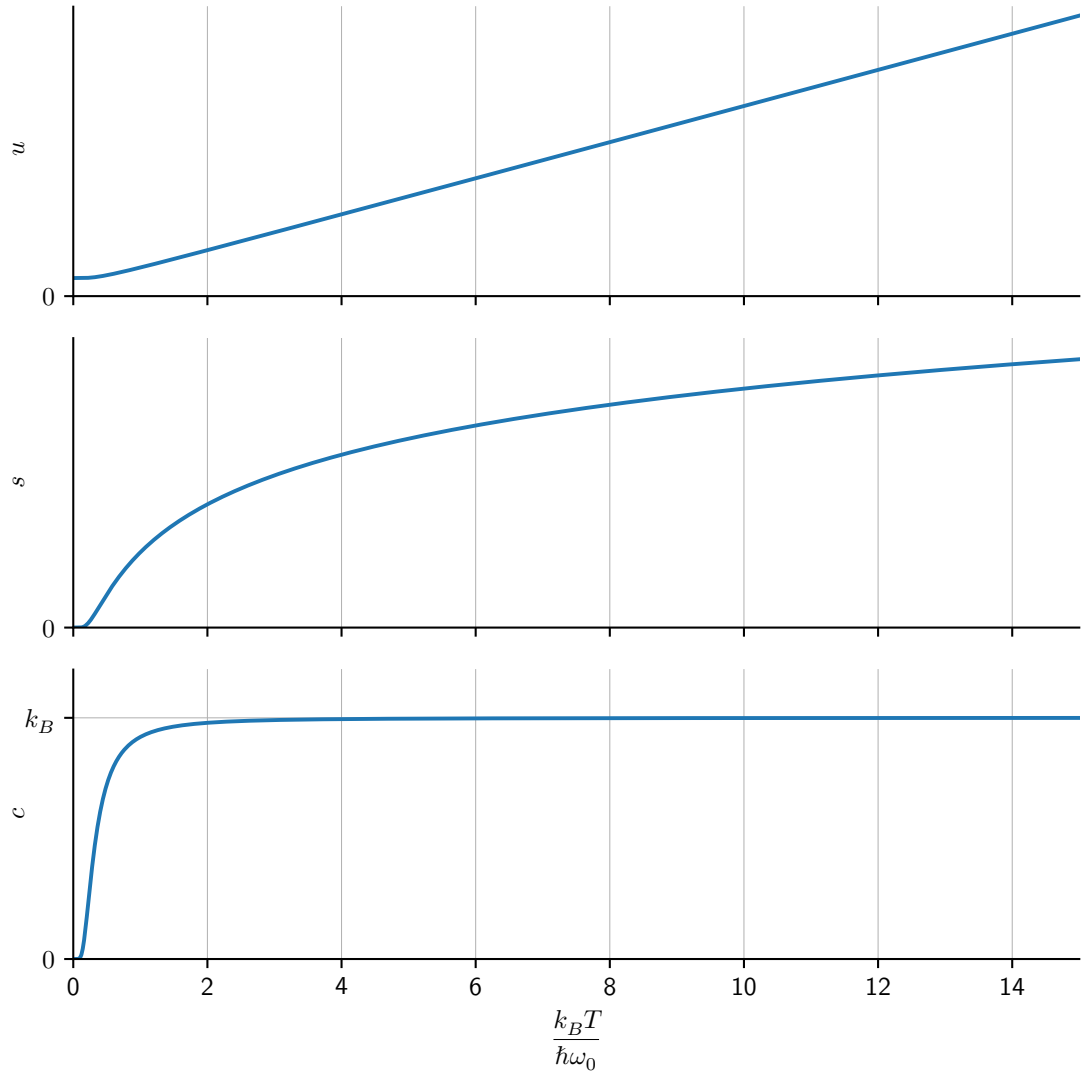
Figure 2.2: *Internal energy per particle u, entropy per particle s, and specific heat per particle c as functions of temperature for a system of N non-interacting quantum harmonic oscillators. As the temperature gets larger, the specific heat tends to a constant, in accordance with the Dulong–Petit law. It also vanishes for small temperatures, in qualitative agreement with experiment.*

given by

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \frac{m\omega_0^2 q_i^2}{2}. \tag{2.32}$$

Therefore, the partition function becomes

$$Z = \int \exp\left(-\beta\left(\sum_{i=1}^{N} \frac{p_i^2}{2m} + \frac{m\omega_0^2 q_i^2}{2}\right)\right) \frac{\mathrm{d}^N q\,\mathrm{d}^N p}{h^N}, \tag{2.33a}$$

$$= \left(\int \exp\left(-\beta\left(\frac{p^2}{2m} + \frac{m\omega_0^2 q^2}{2}\right)\right) \frac{\mathrm{d}q\,\mathrm{d}p}{h}\right)^N, \tag{2.33b}$$

$$= \left(\frac{1}{h}\sqrt{\frac{2\pi m}{\beta}}\sqrt{\frac{2\pi}{\beta m\omega_0^2}}\right)^N, \tag{2.33c}$$

$$= \left(\frac{2\pi}{\beta h\omega_0}\right)^N, \tag{2.33d}$$

$$= \frac{1}{\beta^N \hbar^N \omega_0^N}, \tag{2.33e}$$

which leads to

$$u = k_B T, \tag{2.34}$$

$$f = k_B T \log(\beta\hbar\omega_0), \tag{2.35}$$

and

$$s = k_B\left(1 + \log\left(\frac{k_B T}{\hbar\omega_0}\right)\right), \tag{2.36}$$

which match Eq. (2.28) on page 21.

Classical Ideal Gas

A particularly interesting example for us to consider is that of a classical ideal gas. There are two main reasons for that:

i. we'll be able to recover the we'll known equations of state, and hence understand how to derive them from first principles;

ii. our initial approach will lead to a wrong expression for the entropy that will need to be corrected, hence paving the way for further discussion.

We'll be dealing with a gas comprised of N particles which are assumed not to interact. Furthermore, we assume it to be isotropic (hence, we're ignoring effects due to gravity, for example). The Hamiltonian is then given by

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_{x,i}^2}{2m} + \frac{p_{y,i}^2}{2m} + \frac{p_{z,i}^2}{2m} = \frac{1}{2m}\sum_{i=1}^{N} \|\mathbf{p}_i\|^2. \tag{2.37}$$

Let us begin with the naïve computation. The partition function would then be given by

$$Z = \int \exp\left(-\frac{\beta}{2m}\sum_{i=1}^{N}\|\mathbf{p}_i\|^2\right)\frac{\mathrm{d}^{3N}q\,\mathrm{d}^{3N}p}{h^{3N}}, \tag{2.38a}$$

$$= \frac{1}{h^{3N}}\left[\int \exp\left(-\frac{\beta\|\mathbf{p}\|^2}{2m}\right)\mathrm{d}^3q\,\mathrm{d}^3p\right]^N, \tag{2.38b}$$

$$= \frac{V^N}{h^{3N}}\left[\int \exp\left(-\frac{\beta(p_x^2+p_y^2+p_z^2)}{2m}\right)\mathrm{d}p_x\,\mathrm{d}p_y\,\mathrm{d}p_z\right]^N, \tag{2.38c}$$

$$= \frac{V^N}{h^{3N}}\left[\frac{2\pi m}{\beta}\right]^{\frac{3N}{2}}. \tag{2.38d}$$

Using this wrong partition function, we get to the correct internal energy U,

$$U = -\frac{\partial}{\partial\beta}\log Z = \frac{3}{2}Nk_BT. \tag{2.39}$$

Nevertheless, it leads us to the wrong Helmholtz free energy F. In spite of this, we can get to the correct expression for the pressure by doing

$$p = \left(\frac{\partial F}{\partial V}\right)_T, \tag{2.40a}$$

$$= k_BT\frac{\partial}{\partial V}\log Z, \tag{2.40b}$$

$$= \frac{Nk_BT}{V}. \tag{2.40c}$$

We can see the issue arising by explicitly computing the Helmholtz free energy, which leads us to

$$F = -k_BT\log Z, \tag{2.41a}$$

$$= -Nk_BT\left[\log\left(\frac{V}{h^3}\right) + \frac{3}{2}\log\left(\frac{2\pi m}{\beta}\right)\right], \tag{2.41b}$$

which is not extensive in *V*. Hence, we're doing something wrong.

It is particularly curious that we did get two equations of state correctly, as we can see on Eqs. (2.39) and (2.40). From Thermodynamics, we know this is enough to characterize the system completely. One may then wonder whether the issue in here lies in Statistical Mechanics or was already there in Thermodynamics, putting in risk our model of an ideal gas. Thermodynamics is safe, though: the result that two equations of state are sufficient to characterize a system follows from the assumption that entropy is extensive, and hence it does not hold in our case. If we decide to compute the free energy per unit particle from our equations of state and then assume extensivity explicitly, we'll be able to obtain

the correct expression known from thermodynamics. The question is then: why isn't our expression as taken directly from the partition function matching the correct expression?

Eq. (2.38) on the previous page shows us that we're currently considering the partition function

$$Z = \zeta^N, \tag{2.42}$$

where the one-particle partition function ζ is given by

$$\zeta = \frac{V}{h^3} \left[\frac{2\pi m}{\beta} \right]^{\frac{3}{2}}. \tag{2.43}$$

We'll make an *ad hoc* assumption: the partition function is actually given by

$$Z = \frac{\zeta^N}{N!} = \frac{V^N}{N!} \left[\frac{2\pi m}{\beta h^2} \right]^{\frac{3N}{2}}. \tag{2.44}$$

There are justification for this assumption. For example, it is necessary to recover the classical results from the quantum mechanical results (see Pathria and Beale 2022, Sec. 3.5 and references therein). Another point of view is that it must be introduced to account for the indistinguishability of the gas' particles[2]. At last, we can simply justify it in an *ad hoc* manner: it is needed because without it the result is wrong.

Using this new expression for the partition function, we get to[3]

$$\log Z = N \log \zeta - \log N!, \tag{2.45a}$$
$$= N \log \zeta - N \log N + N \mathcal{O}(\log N), \tag{2.45b}$$
$$\approx N \left[\log \frac{\zeta}{N} + 1 \right], \tag{2.45c}$$

where we employed Stirling's approximation on Eq. (2.45b). As a consequence, the Helmholtz free energy now reads

$$F = -N k_B T \left[\log \frac{\zeta}{N} + 1 \right], \tag{2.46a}$$
$$= -N k_B T \left[\log \left(\frac{V}{N} \frac{1}{h^3} \left(\frac{2\pi m}{\beta} \right)^{\frac{3}{2}} \right) + 1 \right], \tag{2.46b}$$
$$= -N k_B T \left[\log \left(\frac{V}{N} \frac{1}{\lambda^3} \right) + 1 \right], \tag{2.46c}$$

---

[2]Prof. Fiore disagrees with this justification since we can always distinguish classical particles by following their trajectories, a point of view shared by Kardar (2007b, p. 109). Caticha (2019) argues that it doesn't matter whether the particles are or not distinguishable, but rather if their distinguishability interests us.

[3]It is curious that in the case of finite N the results will lead to a non-extensive entropy, due to the remaining terms in the Stirling approximation. In the literature, there seems to be at least one claim that entropy will actually not be exactly extensive (see Peters 2014).

where in the last line we defined the thermal wavelength λ by

$$\lambda = h\left(\frac{\beta}{2\pi m}\right)^{\frac{1}{2}} \tag{2.47}$$

Notice that Eq. (2.46) on the previous page is indeed extensive on volume and number of particles, as it should be.

We may then compute the entropy for an ideal gas. It will be given by

$$S = \frac{U-F}{T}, \tag{2.48a}$$

$$= \frac{3}{2}Nk_B + Nk_B\left[\log\left(\frac{V}{N\lambda^3}\right) + 1\right], \tag{2.48b}$$

$$= Nk_B\left[\log\left(\frac{V}{N\lambda^3}\right) + \frac{5}{2}\right], \tag{2.48c}$$

$$= Nk_B\left[\log\frac{V}{N} + \log\frac{1}{h^3} + \frac{3}{2}\log(2\pi mk_BT) + \frac{5}{2}\right], \tag{2.48d}$$

$$= Nk_B\left[\frac{3}{2}\log k_BT + \log\frac{V}{N} - 3\log h + \frac{3}{2}\log(2\pi m) + \frac{5}{2}\right]. \tag{2.48e}$$

Let us now perform the same calculation with the equations of state we previously obtained, Eqs. (2.39) and (2.40) on page 25. They yield

$$\frac{1}{T} = \frac{3Nk_B}{2U} = \frac{3k_B}{2u} = \left(\frac{\partial s}{\partial u}\right)_v, \tag{2.49a}$$

$$\frac{p}{T} = \frac{Nk_B}{V} = \frac{k_B}{v} = \left(\frac{\partial s}{\partial v}\right)_u, \tag{2.49b}$$

which can be integrated to obtain

$$s(u,v) = \frac{3}{2}k_B\log u + k_B\log v + s_0, \tag{2.50}$$

where s_0 is some integration constant. If we explicitly impose extensivity of entropy, we'll obtain

$$S(U,V,N) = Ns(u,v), \tag{2.51a}$$

$$= \frac{3}{2}Nk_B\log\frac{U}{N} + Nk_B\log\frac{V}{N} + Ns_0, \tag{2.51b}$$

$$= Nk_B\left[\frac{3}{2}\log\left(\frac{3}{2}k_BT\right) + \log\frac{V}{N} + \frac{s_0}{k_B}\right], \tag{2.51c}$$

$$= Nk_B\left[\frac{3}{2}\log k_BT + \log\frac{V}{N} + \log\frac{3}{2} + \frac{s_0}{k_B}\right]. \tag{2.51d}$$

2.3 Gases with Diatomic Molecules

Classical Theory

For classical systems, the mean energy always has the form[4]

$$U = \frac{f}{2} N k_B T, \tag{2.52}$$

where f is the number of quadratic terms on the Hamiltonian (*i.e.*, essentially the number of degrees of freedom). Hence, for harmonic oscillators, we have $U = N k_B T$. For a monoatomic ideal gas, $U = \frac{3}{2} N k_B T$. For an ideal gas made of rigid diatomic molecules, $U = \frac{5}{2} N k_B T$. If the molecules of this diatomic ideal gas can vibrate, $U = \frac{7}{2} N k_B T$.

For all ideal gases, however, the pressure is the same: since they are non-interacting by definition, the dependence on the positional degrees of freedom is the same, so the volume appears on the calculation in the same way it did for the monoatomic gas. Since the pressure is deduced from the volume dependency, the equation of state for the pressure is the same for all ideal gases. On Section 2.4 we'll consider interactions among different molecules in the Van der Waals gas, which will lead to a different expression for the pressure.

To figure out the thermodynamic properties of the diatomic gas, let us start by noticing the kinetic energy of a single molecule is given by

$$K = \frac{m_1 \|\dot{\mathbf{r}}_1\|^2}{2} + \frac{m_2 \|\dot{\mathbf{r}}_2\|^2}{2}, \tag{2.53}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the positions of the atoms composing the molecule as measured from the laboratory reference frame. We may then define the position of the center of mass, $\mathbf{R}$, and the separation between the particles, $\mathbf{r}$, by

$$\mathbf{R} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2} \quad \text{and} \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2. \tag{2.54}$$

In terms of these variables the kinetic energy can be written as

$$K = \frac{M \|\dot{\mathbf{R}}\|^2}{2} + \frac{\mu \|\dot{\mathbf{r}}\|^2}{2}, \tag{2.55}$$

where we also defined

$$M = m_1 + m_2 \quad \text{and} \quad \mu = \frac{m_1 m_2}{M}. \tag{2.56}$$

It is convenient to express the terms related to $\mathbf{r}$ in spherical coordinates. We then get

$$\frac{\mu r^2}{2} = \frac{\mu}{2} \left(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta + \dot{\phi}^2 \right), \tag{2.57}$$

---

[4]This is know as the equipartition theorem, discussed, *e.g.*, by Pathria and Beale (2022, Sec. 3.7) and Salinas (2001, Sec. 6.3).

and find the Lagrangian

$$L = K - V, \tag{2.58a}$$

$$= \frac{M\|\dot{\mathbf{R}}\|^2}{2} + \frac{\mu}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta + \dot{\phi}^2\right) - V, \tag{2.58b}$$

where V is some potential that might later model the oscillations of the molecule. Notice that the first term of Eq. (2.58b) corresponds to the translational degrees of freedom, while the three following terms correspond to rotational and vibrational degrees of freedom. Translation is already present for monoatomic gases, but rotations and vibrations are a new feature.

The Hamiltonian for a single molecule is then

$$H = \frac{\|\mathbf{p}_R\|^2}{2M} + \frac{p_r^2}{2\mu} + \frac{p_\theta^2}{2\mu r^2} + \frac{p_\phi^2}{2\mu r^2\sin^2\theta} + V. \tag{2.59}$$

For a rigid molecule, similar to a dumbbell, we'll have $p_r = 0$ and $V = 0$ (non-interacting gas, no internal interactions). Let us denote the fixed distance between the two atoms by b. In this case, the Hamiltonian simplifies to

$$H = \frac{\|\mathbf{p}_R\|^2}{2M} + \frac{p_\theta^2}{2\mu b^2} + \frac{p_\phi^2}{2\mu b^2\sin^2\theta}, \tag{2.60}$$

meaning it has five quadratic terms. As a consequence, the internal energy will be

$$U = \frac{5}{2}Nk_BT. \tag{2.61}$$

For a vibrating molecule, the potential has to be that of a harmonic oscillator, so we'll have $p_r \neq 0$ and $V \propto r^2 0$. This time, the Hamiltonian becomes

$$H = \frac{\|\mathbf{p}_R\|^2}{2M} + \frac{p_r^2}{2\mu} + \frac{p_\theta^2}{2\mu r^2} + \frac{p_\phi^2}{2\mu r^2\sin^2\theta} + \frac{m\omega^2 r^2}{2}, \tag{2.62}$$

with seven quadratic terms. Therefore, the internal energy is

$$U = \frac{7}{2}Nk_BT. \tag{2.63}$$

For completeness, let us compute the partition function for the rigid, dumbbell-like molecule. It will be given by

$$Z = \frac{\zeta^N}{N!}, \tag{2.64}$$

where the one-molecule partition function ζ is given by

$$\zeta = \int e^{-\beta H} \frac{\mathrm{d}^3 R \, \mathrm{d}^3 p_R \, \mathrm{d}\theta \, \mathrm{d}p_\theta \, \mathrm{d}\phi \, \mathrm{d}p_\phi}{h^5}, \tag{2.65a}$$

$$= \frac{V}{h^5} \int e^{-\beta \left[\frac{\|\mathbf{p}_R\|^2}{2M} + \frac{p_\theta^2}{2\mu b^2} + \frac{p_\phi^2}{2\mu b^2 \sin^2 \theta} \right]} \mathrm{d}^3 p_R \, \mathrm{d}\theta \, \mathrm{d}p_\theta \, \mathrm{d}\phi \, \mathrm{d}p_\phi, \tag{2.65b}$$

$$= \frac{V}{h^5} \int \left(\frac{2\pi M}{\beta} \right)^{\frac{3}{2}} \left(\frac{2\pi \mu b^2}{\beta} \right)^{\frac{1}{2}} \left(\frac{2\pi \mu b^2 \sin^2 \theta}{\beta} \right)^{\frac{1}{2}} \mathrm{d}\theta \, \mathrm{d}\phi, \tag{2.65c}$$

$$= \frac{(2\pi)^{\frac{7}{2}} V}{h^5 \beta^{\frac{5}{2}}} M^{\frac{3}{2}} \mu b^2 \int_0^\pi \sin \theta \, \mathrm{d}\theta, \tag{2.65d}$$

$$= \frac{2(2\pi)^{\frac{7}{2}} V}{h^5 \beta^{\frac{5}{2}}} M^{\frac{3}{2}} \mu b^2 \int_0^\pi, \tag{2.65e}$$

$$= \frac{4\pi V}{h^5 \beta^{\frac{5}{2}}} (2\pi M)^{\frac{3}{2}} (2\pi \mu b^2). \tag{2.65f}$$

Do notice that on Eq. (2.65a) the integral is being taken in phase space, so there is no Jacobian due to the choice of spherical coordinates.

From Eq. (2.65), it is then straightforward to obtain that

$$U = \frac{5}{2} N k_B T, \tag{2.66}$$

$$c_V = \frac{5}{2} N k_B, \tag{2.67}$$

$$pV = N k_B T, \tag{2.68}$$

as previously claimed.

Quantum Theory

Eq. (2.67), obtained with the classical theory, claims that the specific heat of a gas of diatomic molecules is constant. However, that is not true. Experimentally, a minimum amount of energy is needed to activate the rotational and vibrational degrees of freedom, which leads to the graph of specific heat as a function of temperature having a few plateaus, corresponding to the temperatures in which the new degrees of freedom become relevant (for graphs and experimental data, see Pathria and Beale 2022, Sec. 6.5.B; Salinas 2001, Sec. 8.4; Wannier 1987, Chap. 11). The temperature dependency of the specific heat is sketched on Fig. 2.3 on the next page.

While classical theory fails to explain why the diatomic gas behaves as if it was monoatomic, quantum theory succeeds. To understand how, let us analyse the rotational degrees of freedom in particular. This is possible because the partition function can be decomposed as a product of simpler partition functions, (see Salinas 2001, Sec. 8.4).

Quantum mechanically, the rotational Hamiltonian of a single molecule will be

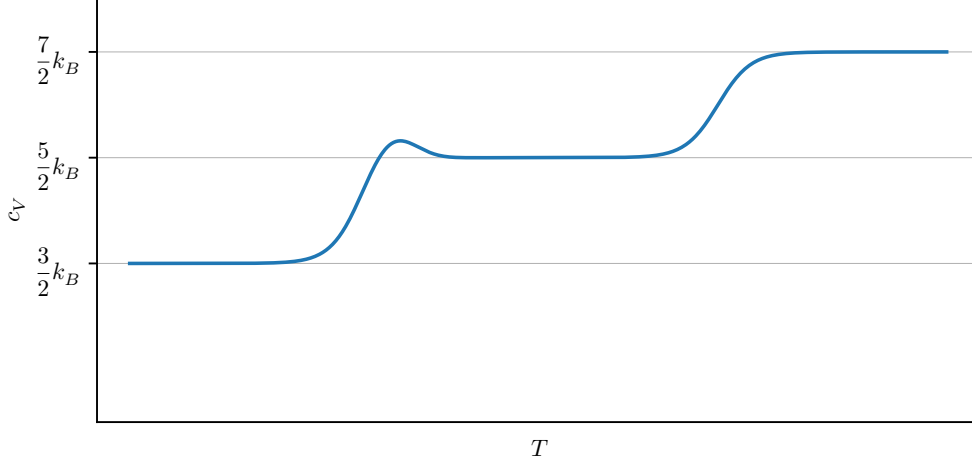$$\hat{H}_{\text{rot}} = \frac{\hat{J}^2}{2I}, \tag{2.69}$$

Figure 2.3: *Sketch of the specific heat per molecule of a diatomic ideal gas of heteromolecules (*i.e.*, each molecule is composed of two different nuclei). The characteristic bump before the specific heat reaches $\frac{5}{2}k_B$ is typical of diatomic gases composed of heteromolecules. As higher temperatures are achieved, new degrees of freedom get activated and the specific heat increases. The figure is adapted from the figures given by Chabay and Sherwood (2015, Sec. 12.9), Pathria and Beale (2022, Sec. 6.5.B), Salinas (2001, Sec. 8.4), and Wannier (1987, Chap. 11)*

where I is the molecule's moment of inertia. We know that $\hat{J}^2$ has the eigenstates

$$\hat{J}^2 \left| l, m \right\rangle = \hbar^2 l(l+1) \left| l, m \right\rangle, \tag{2.70}$$

where $l \in \mathbb{N}_0$, $-l \leq m \leq l$, and each energy level has a $(2l+1)$-fold degeneracy. Therefore, the "one-molecule rotational partition function" is

$$\zeta_{\text{rot}} = \sum_{l=0}^{+\infty} \sum_{m=-l}^{l} e^{-\frac{\beta \hbar^2 l(l+1)}{2I}}, \tag{2.71a}$$

$$= \sum_{l=0}^{+\infty} (2l+1) e^{-\frac{\beta \hbar^2 l(l+1)}{2I}}. \tag{2.71b}$$

This expression can be computed numerically, but not analytically. In spite of this difficulty, we can obtain analytical expressions for the remarkable limits of high and low temperatures.

At low temperatures, we have $\frac{\beta \hbar^2}{2I} \gg 1$, so rotational energy is much lower than thermal energy. As a consequence, high values of l are suppressed and we can approximate ζ_{rot} by the first two terms of the sum. We get

$$\zeta_{\text{rot}} \approx \sum_{l=0}^{1} (2l+1) e^{-\frac{\beta \hbar^2 l(l+1)}{2I}}, \tag{2.72a}$$

$$= 1 + 3 e^{-\frac{\beta \hbar^2}{I}}. \tag{2.72b}$$

– 31 –

From Eq. (2.72) on the preceding page, we see the "rotational Helmholtz free energy per molecule" is given by

$$f_{\text{rot}} = -k_B T \log \zeta_{\text{rot}}, \tag{2.73a}$$

$$= -k_B T \log\left(1 + 3e^{-\frac{\beta \hbar^2}{I}}\right), \tag{2.73b}$$

$$\approx -3k_B T e^{-\frac{\beta \hbar^2}{I}}. \tag{2.73c}$$

Eq. (2.72) on the previous page also leads to

$$u_{\text{rot}} = -\frac{\partial}{\partial \beta} \log \zeta_{\text{rot}}, \tag{2.74a}$$

$$\approx \frac{3\hbar^2}{I} e^{-\frac{\beta \hbar^2}{I}}, \tag{2.74b}$$

and hence

$$c_{\text{rot}} = \frac{\partial u}{\partial T} \approx 3k_B \left(\frac{\beta \hbar^2}{I}\right)^2 e^{-\frac{\beta \hbar^2}{I}}, \tag{2.75}$$

which tends to zero for temperatures much smaller than the "rotational temperature" $T_R = \frac{\hbar^2}{2k_B I}$, showing the rotational degrees of freedom "freeze" at low temperatures. The rotational (and vibrational) temperatures for some gases are shown on Table 2.1.

Table 2.1: *Rotational and vibrational temperatures for some diatomic gases. Taken from Salinas 2001, p. 156.*

Gas	T_R (K)	T_V (10^3 K)
H_2	85.4	6.10
N_2	2.86	3.34
O_2	2.07	2.23
CO	2.77	3.07
NO	2.42	2.69
HCl	15.2	4.14

For high temperatures ($\frac{\beta \hbar^2}{2I} \ll 1$), there is a very small spacing between different energy levels, which allows us to approximate the sum on Eq. (2.71) on the previous page by an integral, which leads us to

$$\zeta_{\text{rot}} = \sum_{l=0}^{+\infty} (2l+1) e^{-\frac{\beta \hbar^2 l(l+1)}{2I}}, \tag{2.76a}$$

$$= \int_0^{+\infty} (2\xi + 1) e^{-\frac{\beta \hbar^2 \xi(\xi+1)}{2I}} \, d\xi, \tag{2.76b}$$

$$= \int_0^{+\infty} e^{-\frac{\beta \hbar^2 u}{2I}} \, du, \tag{2.76c}$$

$$= \frac{2k_B T I}{\hbar^2}, \tag{2.76d}$$

where on Eq. (2.76c) on the preceding page we performed the substitution $u = \xi(\xi + 1)$.

Notice that Eq. (2.76) on the previous page will then lead to the conclusion that $c_{\text{rot}} = k_B$, corresponding to all rotational degrees of freedom being activated and matching the classical prediction, as expected.

Instead of just approximating the sum by an integral on Eq. (2.76) on the preceding page, we could be careful and use the Euler–MacLaurin expansion formula (see Abramowitz and Stegun 1972; Arfken, Weber, and Harris 2013, Sec. 12.3), as Salinas (2001, Sec. 8.4) does, to get the asymptotic expression

$$\zeta_{\text{rot}} = \frac{T}{T_R}\left[1 + \frac{1}{3}\left(\frac{T_R}{T}\right) + \frac{1}{15}\left(\frac{T_R}{T}\right)^2 + \cdots\right], \tag{2.77}$$

which leads to the specific heat

$$c_{\text{rot}} = k_B\left[1 + \frac{1}{45}\left(\frac{T_R}{T}\right)^2 + \cdots\right], \tag{2.78}$$

from which we can see that the specific heat will tend to k_B for $T \gg T_R$. For hetero-molecules, the dots on Eq. (2.78) predict the bump of Fig. 2.3 on page 31, matching experiment.

A similar analysis can be carried out for the vibrational degree of freedom, which can be modeled as a harmonic oscillator. We'll get a similar result that the degree of freedom "freezes" at low temperatures. The specific heat will be given by Eq. (2.26) on page 21, which we previously obtained when studying an ensemble of harmonic oscillators.

Our main lesson is then that, in polyatomic molecules, we are often able to deal with each degree of freedom separately. For some more details on polyatomic molecules, Pathria and Beale (see 2022, Sec. 6.5.C).

2.4 Interactions and the Virial Expansion

So far, we have only considered non-interacting systems, which, albeit simple, are not realistic. We won't be able to deal exactly with interactions, but it is possible to obtain approximate expressions. In this section, we'll see one such method and illustrate it with the Van der Waals gas, the most famous model of a "real" gas. For some more information, one can see, *e.g.*, the texts by Kardar (2007b, Chap. 5), Pathria and Beale (2022, Chap. 10), and Salinas (2001, Sec. 6.4).

For an ideal gas, the equation of state for pressure reads

$$\frac{p}{k_B T} = \frac{1}{v}, \tag{2.79}$$

and we have $p = p(T, v)$ in general. Hence, it is reasonable to expect we can write the equation of state for more real gases in the form

$$\frac{p}{k_B T} = \frac{1}{v} + \frac{A(T)}{v^2} + \frac{B(T)}{v^3} + \cdots, \tag{2.80}$$

which is know as the virial expansion. We'd like then to compute the virial coefficients A, B, and so on from first principles. We'll focus on only the first coefficient, A.

The actual equation of state we'll try to understand from first principles is that of a Van der Waals gas,

$$p = \frac{k_B T}{v - b} - \frac{a}{v^2}, \tag{2.81}$$

where a and b are constants. As we shall see, we can interpret b as representing the finite volume of the gas' molecules and a as representing the intermolecular interactions, which are assumed to behave as hard spheres with a small attractive potential. As illustrated in Fig. 2.4 on the following page, this model provides a simplified version of the Lennard-Jones potential, which is the typical potential for intermolecular interactions, but is harder to treat.

We'll begin with a general approach and impose the hard sphere potential only at the end, so we'll also derive a more general expression that applies to a wide class of interactions.

We assume the interaction potential to depend solely on the distances between molecules, $V = V(\|\mathbf{r}_i - \mathbf{r}_j\|)$. This time, we can't use a one-particle partition function, so we write Z as

$$Z = \frac{1}{N!} \int e^{-\beta H(q,p)} \frac{\mathrm{d}^{3N} q \, \mathrm{d}^{3N} p}{h^{3N}}, \tag{2.82a}$$

$$= \frac{1}{N!} \int e^{-\beta \left(\sum_{i=1}^{N} \frac{\|\mathbf{p}_i\|^2}{2m} + \sum_{i<j} V(\|\mathbf{r}_i - \mathbf{r}_j\|) \right)} \frac{\mathrm{d}^{3N} q \, \mathrm{d}^{3N} p}{h^{3N}}, \tag{2.82b}$$

where the sum over i and j with the condition $i < j$ prevents us from overcounting the interactions. While the integrals over generalized coordinates got more complicated, the integrals over momenta work just like the ideal gas, leading us to

$$Z = \frac{1}{N! h^{3N}} \left(\frac{2\pi m}{\beta} \right)^{\frac{3N}{2}} \int e^{-\beta \sum_{i<j} V(\|\mathbf{r}_i - \mathbf{r}_j\|)} \mathrm{d}^{3N} q, \tag{2.83a}$$

$$= \frac{1}{N! h^{3N}} \left(\frac{2\pi m}{\beta} \right)^{\frac{3N}{2}} Q_N, \tag{2.83b}$$
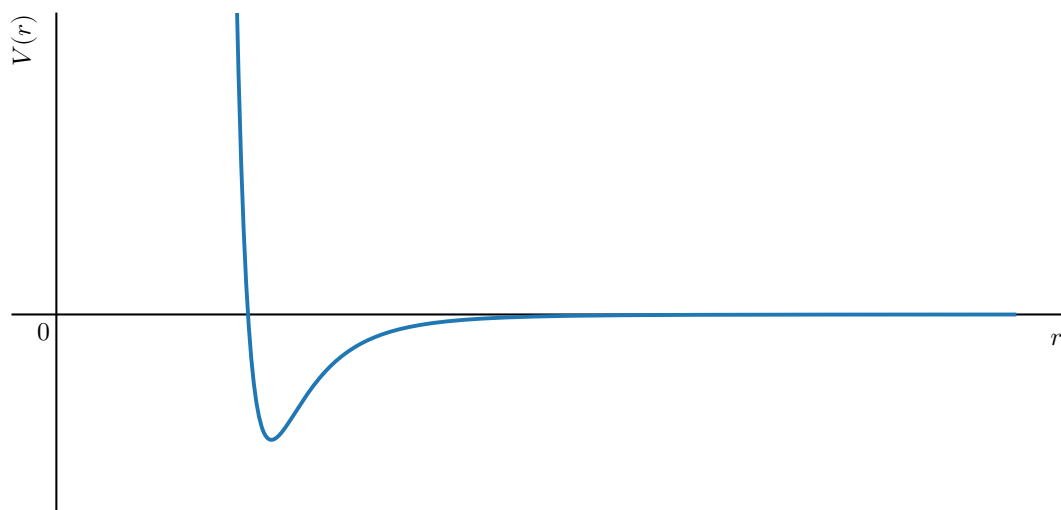
where Q_N is given by

$$Q_N \equiv \int e^{-\beta \sum_{i<j} V(\|\mathbf{r}_i - \mathbf{r}_j\|)} \mathrm{d}^{3N} q, \tag{2.84a}$$

$$= \int \prod_{i<j} e^{-\beta V(\|\mathbf{r}_i - \mathbf{r}_j\|)} \prod_{k=1}^{N} \mathrm{d}^3 r_k. \tag{2.84b}$$
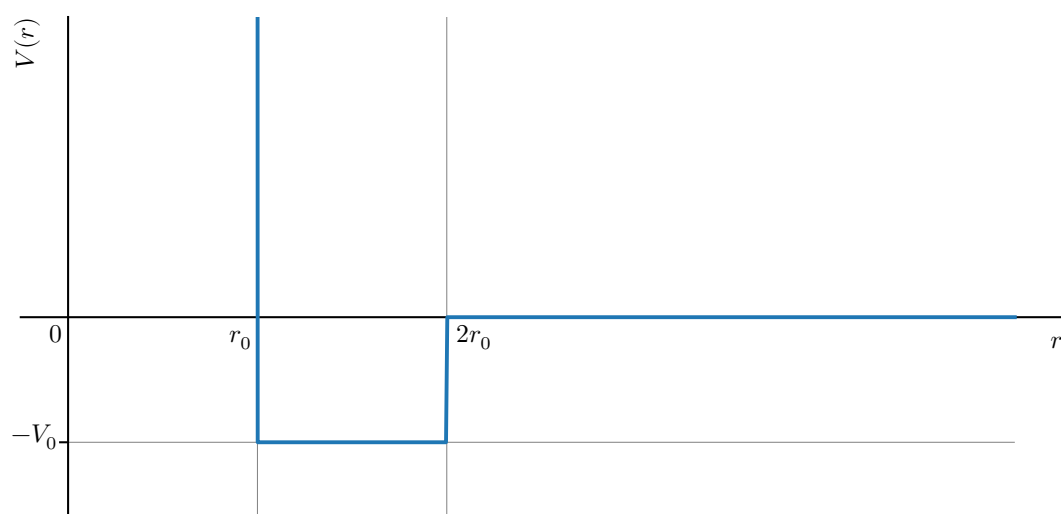
For simplicity, let us denote $V_{ij} \equiv V(\|\mathbf{r}_i - \mathbf{r}_j\|)$. It will also be convenient to define

$$f_{ij} = e^{-\beta V_{ij}} - 1, \tag{2.85}$$

for f_{ij} will remain finite when V_{ij} diverges (see Fig. 2.5 on page 36) and overall make the calculations simpler.

(a) *Lennard-Jones potential. Typical shape of the potential for intermolecular interactions.*

(b) *Potential for hard sphere scattering with a small attraction.*

Figure 2.4: *A simplification of the Lennard-Jones potential is a potential for hard sphere scattering with a small attractive region.*
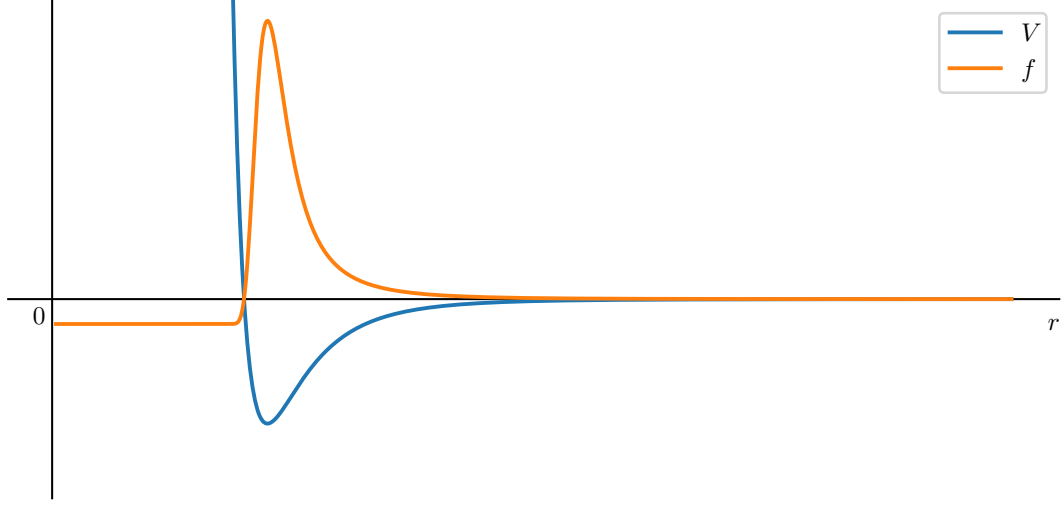
Figure 2.5: *While the Lennard-Jones potential diverges at the origin, the function f defined by Eq. (2.85) on page 34 is well-behaved everywhere.*

In terms of this new function, we can write Q_N as

$$Q_N = \int \prod_{i<j}(1 + f_{ij}) \prod_{k=1}^{N} \mathrm{d}^3 r_k \,. \tag{2.86}$$

Notice that the product $\prod_{i<j}(1 + f_{ij})$ has the form

$$\prod_{i<j}(1 + f_{ij}) = (1 + f_{12})(1 + f_{13})(1 + f_{14}) \cdots (1 + f_{1N})(1 + f_{23}) \cdots (1 + f_{N-1,N}), \tag{2.87a}$$

$$= 1 + \sum_{i<j} f_{ij} + \cdots, \tag{2.87b}$$

where the dots include terms quadratic or higher on the f_{ij}. Our first approximation will consist on neglecting such terms. Hence, we'll write

$$\prod_{i<j}(1 + f_{ij}) \approx 1 + \sum_{i<j} f_{ij}. \tag{2.88}$$

With this approximation, Q_N can be written as

$$Q_N = \int \left(1 + \sum_{i<j} f_{ij}\right) \prod_{k=1}^{N} \mathrm{d}^3 r_k \,, \tag{2.89a}$$

$$= \int \prod_{k=1}^{N} \mathrm{d}^3 r_k + \sum_{i<j} \int f_{ij} \prod_{k=1}^{N} \mathrm{d}^3 r_k \,, \tag{2.89b}$$

$$= V^N + \sum_{i<j} \int f_{ij} \, \mathrm{d}^3 r_i \, \mathrm{d}^3 r_j \prod_{k \neq i,j} \mathrm{d}^3 r_k \,, \tag{2.89c}$$

$$= V^N + V^{N-2} \sum_{i<j} \int f_{ij} \, \mathrm{d}^3 r_i \, \mathrm{d}^3 r_j \,, \tag{2.89d}$$

$$= V^N + \frac{V^{N-2} N(N-1)}{2} \int f_{12} \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,, \tag{2.89e}$$

where we used the fact that, since we're integrating over $\mathbf{r}_i$ and $\mathbf{r}_j$, all terms in the sum are equal. If we now introduce coordinates $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ and[5] $\mathbf{R} = \mathbf{r}_1 + \mathbf{r}_2$, the integral becomes

$$Q_N = V^N + \frac{V^{N-2} N(N-1)}{2} \int f(r) \, \mathrm{d}^3 R \, \mathrm{d}^3 r \,, \tag{2.90a}$$

$$= V^N + \frac{V^{N-1} N(N-1)}{2} \int f(r) \, \mathrm{d}^3 r \,, \tag{2.90b}$$

$$= V^N + 2\pi V^{N-1} N(N-1) \int f(r) r^2 \, \mathrm{d}r \,, \tag{2.90c}$$

where in the last step we chose to work in spherical coordinates and used the assumption that the interaction depends only on the distance between the molecules, and hence it is isotropic.

To compute thermodynamics quantities we'll ultimately be interested in $\log Z$. Hence, we notice now that Eq. (2.90) leads us to

$$\log Q_N = N \log V + \log \left[1 + \frac{2\pi N(N-1)}{V} \int f(r) r^2 \, \mathrm{d}r \right]. \tag{2.91}$$

Now comes our second approximation: we'll assume $\frac{N}{V} \ll 1$, i.e., that the gas is in a low density regime[6]. Under this hypothesis, we can write

$$\log Q_N \approx N \log V + \frac{2\pi N(N-1)}{V} \int f(r) r^2 \, \mathrm{d}r \,. \tag{2.92}$$

Using Eqs. (2.83) and (2.92) on page 34 and on the current page we get to

$$\frac{1}{N} \log Z = -\frac{1}{N} \log N! + \frac{3}{2} \log \frac{2\pi m}{\beta h^2} + \frac{1}{N} \log Q_N, \tag{2.93a}$$

$$\approx -\log N + 1 + \frac{3}{2} \log \frac{2\pi m}{\beta h^2} + \log V + \frac{2\pi(N-1)}{V} \int f(r) r^2 \, \mathrm{d}r \,, \tag{2.93b}$$

$$\approx -\log N + 1 + \frac{3}{2} \log \frac{2\pi m}{\beta h^2} + \log V + \frac{2\pi N}{V} \int f(r) r^2 \, \mathrm{d}r \,, \tag{2.93c}$$

$$= \log \frac{V}{N} + \frac{3}{2} \log \frac{2\pi m}{\beta h^2} + \frac{2\pi N}{V} \int f(r) r^2 \, \mathrm{d}r + 1 \,, \tag{2.93d}$$

$$= \log v + \frac{3}{2} \log \frac{2\pi m}{\beta h^2} + \frac{2\pi}{v} \int f(r) r^2 \, \mathrm{d}r + 1 \,. \tag{2.93e}$$

---

[5]One could choose $\mathbf{R}$ to be the position of the center of mass. I'm using $\mathbf{R} = \mathbf{r}_1 + \mathbf{r}_2$ to avoid having to think about the Jacobians that go into the integral, but either choice should work out equally well.

[6]$\frac{N}{V} \ll 1$ is not dimensionally correct, but it should be understood in comparison with the terms that go along with it on Eq. (2.91).

With Eq. (2.93) on the preceding page we can now compute thermodynamic quantities. The Helmholtz free energy per molecule, for example, is given by

$$f(T, v) = -\frac{1}{\beta N} \log Z, \tag{2.94a}$$

$$= -\frac{1}{\beta}\left[\log v + \frac{3}{2}\log\frac{2\pi m}{\beta h^2} + \frac{2\pi}{v}\int f(r)r^2\,\mathrm{d}r + 1\right]. \tag{2.94b}$$

Therefore, the pressure is given by

$$p(T, v) = -\left(\frac{\partial f}{\partial v}\right)_T, \tag{2.95a}$$

$$= \frac{k_B T}{v} - \frac{2\pi k_B T}{v^2}\int f(r)r^2\,\mathrm{d}r. \tag{2.95b}$$

By comparing Eqs. (2.80) and (2.95) on page 33 and on this page, we see we've found the first virial term to be

$$A(T) = -2\pi\int_0^{+\infty} f(r)r^2\,\mathrm{d}r. \tag{2.96}$$

Let us now consider the specific case of the Van der Waals gas. We go back to the hard sphere potential of Fig. 2.4b on page 35 and notice that it leads us to

$$A(T) = -2\pi\left[\int_0^{r_0} f(r)r^2\,\mathrm{d}r + \int_{r_0}^{2r_0} f(r)r^2\,\mathrm{d}r + \int_{2r_0}^{+\infty} f(r)r^2\,\mathrm{d}r\right], \tag{2.97a}$$

$$= -2\pi\left[-\int_0^{r_0} r^2\,\mathrm{d}r + \int_{r_0}^{2r_0}(e^{\beta V_0}-1)r^2\,\mathrm{d}r + 0\right], \tag{2.97b}$$

$$= -2\pi\left[-\frac{r_0^3}{3} + \frac{7r_0^3}{3}(e^{\beta V_0}-1)\right], \tag{2.97c}$$

$$= \frac{2\pi r_0^3}{3} - \frac{14\pi r_0^3}{3}(e^{\beta V_0}-1). \tag{2.97d}$$

For a weakly attractive potential, in which V_0 can be assumed to be very small, we can then write

$$A(T) = \frac{2\pi r_0^3}{3} - \frac{14\pi r_0^3 V_0}{3k_B T}. \tag{2.98}$$

If we now recall the equation of state for the Van der Waals gas, Eq. (2.81) on page 34, we can see that

$$\frac{p}{k_B T} = \frac{1}{v-b} - \frac{a}{k_B T v^2}, \tag{2.99a}$$

$$= \frac{1}{v\left(1-\frac{b}{v}\right)} - \frac{a}{k_B T v^2}, \tag{2.99b}$$

$$= \frac{1}{v} + \frac{b}{v^2} + \frac{b^2}{v^3} + \cdots - \frac{a}{k_B T v^2}, \tag{2.99c}$$

$$= \frac{1}{v} + \left(b - \frac{a}{k_B T}\right)\frac{1}{v^2} + \frac{b^2}{v^3} + \cdots. \tag{2.99d}$$

Comparing Eqs. (2.98) and (2.99) on the previous page lets us identify

$$a = \frac{14\pi r_0^3 V_0}{3} \quad \text{and} \quad b = \frac{2\pi r_0^3}{3}, \tag{2.100}$$

which justifies interpreting the Van der Waals gas as a gas of weakly attractive hard spheres. For further discussion on the Van der Waals equation see, *e.g.*, the text by Kardar (2007b, Sec. 5.3).

As last, it is worth mentioning that for high temperatures the Van der Waals gas behaves just like an ideal gas, as depicted on Fig. 2.6 on the following page. However, for small temperatures, the isotherms do not match experiment and there's even the presence of thermodynamic instabilities ($\left(\frac{\partial p}{\partial v}\right)_T > 0$). On Section 3 we'll take a close a look at these issues and notice there is a phase transition happening with the same critical exponents of the Curie–Weiss model for a ferromagnet.

2.5 Grand Canonical Ensemble

The canonical ensemble is frequently useful, but the canonical partition function doesn't factorize for quantum gases due to particle indistinguishability. For these cases, it will be useful to use the grand canonical ensemble.

Let us begin by discussing when to use the different ensembles. While we overlooked the microcanonical ensemble at jumped straight at the canonical ensemble, a traditional way of deriving the canonical partition function is by considering the physical situation corresponding to each ensemble.

The so-called microcanonical ensemble (discussed, *e.g.*, by Kardar 2007b, Sec. 4.2; Pathria and Beale 2022, Sec. 2.3; Salinas 2001, Chap. 4) corresponds to an isolated system, which means we hold the internal energy U, the volume V, and the number of particles N fixed. In this situation, the fundamental principle of Statistical Mechanics states that all microstates are equally likely and, as a consequence, the problem boils down to counting the number of accessible microstates. Denoting said number of microstates by $\Omega(U, V, N)$, the entropy is obtained through Boltzmann's formula $S(U, V, N) = k_B \log \Omega(U, V, N)$, from which we can obtain all other thermodynamic properties.

The canonical ensemble corresponds to a different physical situation. When using it, we assume to be describing a subsystem S of an isolated system. S itself is not isolated, but rather is subject to a thermal bath R at temperature T. Under these conditions, one can show (see, *e.g.*, Salinas 2001, Chap. 5), that the probability distribution for the microstates of the subsystem will be given by the Gibbs distribution,

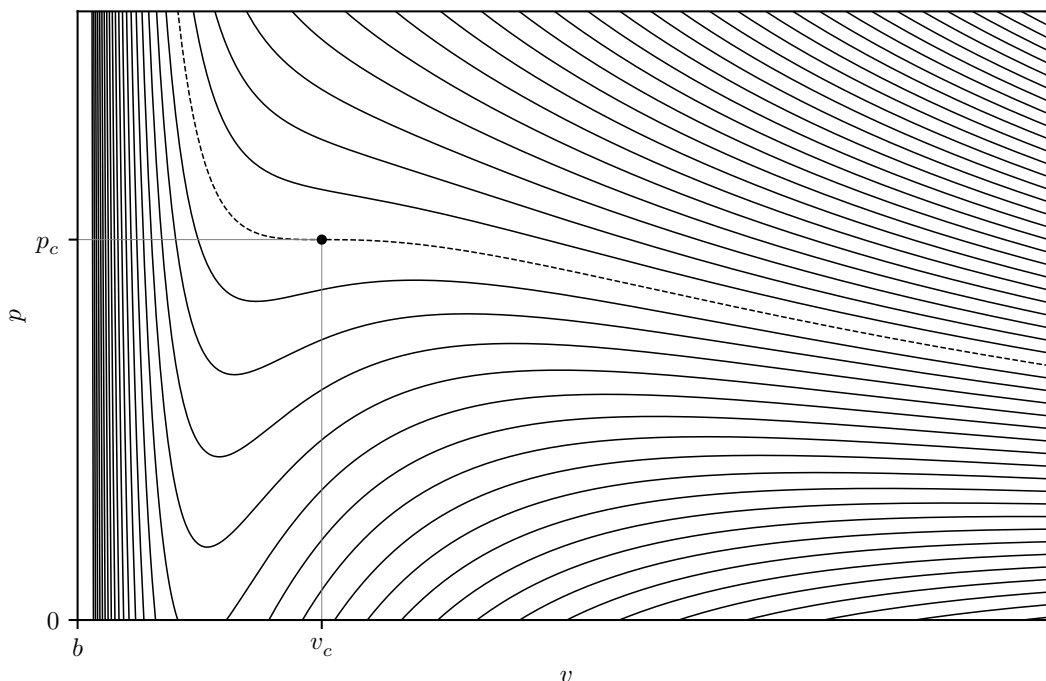$$p_j = \frac{e^{-\beta E_j}}{Z}. \tag{2.101}$$

Figure 2.6: *pv diagram for the Van der Waals gas. Each line is an isotherm. Notice that for high temperatures, we recover the behavior of an ideal gas, while small temperatures (namely, those below the dashed line) present new, problematic behaviors that hint at what we'll later learn to be a phase transition. The dashed line correspond to the so-called critical temperature, and the highlighted point is the so-called critical point. This graph was based on the code by christian (2016).*

Notice, in particular, that Z is a function of temperature, volume, and number of particles. Hence, the system and bath exchange energy in order to keep the temperature fixed, but the volume and number of particles of the subsystem remain fixed.

The canonical ensemble is one of the most used ones. Computing the number of microstates in the microcanonical ensemble is often a cumbersome task, while computing the canonical partition function is typically fair easier.

In spite of that, there are physical situations in which the canonical ensemble is not convenient. In some of these, the grand canonical ensemble might be a better choice. It corresponds to the situation in which the subsystem S is subject not only to a thermal bath, but also to a particle bath. Hence, volume is held fixed, but energy and number of particles are not. Instead, temperature and chemical potential are kept constant while energy and number of particles fluctuate.

In this last physical situation, using the facts known from the microcanonical ensemble and the hypothesis that the energy and number of particles of the subsystem are much less than those of the composite system involving the bath, one can show (Salinas 2001,

Sec. 7.2) that the probability distribution for the microstates of the subsystem will be

$$p_j = \frac{e^{-\beta E_j + \beta \mu N_j}}{\sum_i e^{-\beta E_i + \beta \mu N_i}}, \tag{2.102}$$

where β is the (fixed) inverse temperature, μ is the (fixed) chemical potential, and the sum is carried over states of fixed temperature and chemical potential. We define

$$\Xi(T, V, \mu) = \sum_i e^{-\beta E_i + \beta \mu N_i} \tag{2.103}$$
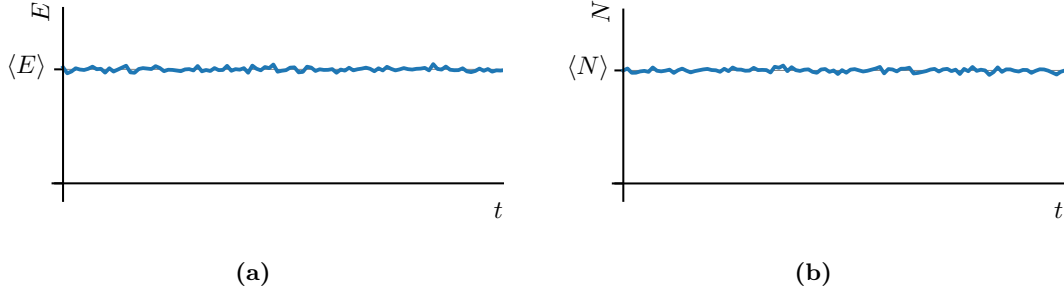
to be the grand canonical partition function.

Figure 2.7: *In the grand canonical ensemble, the energy and number of particles on the subsystem under consideration fluctuate about an average. In the thermodynamic limit, these fluctuations become negligible.*

The fact that we are dealing with systems in equilibrium means, in this context, that while the energy and number of particles of the subsystem do fluctuate with time, they fluctuate about average values, as illustrated on Fig. 2.7. In the thermodynamic limit, these fluctuations will die off and become negligible, just like we had in the canonical ensemble.

One then might inquire about how we'll compute the grand canonical partition function. There are two main approaches:

i. our first option is to list all states, their respective energies E_i and particle numbers N_i, and then perform the sum on Eq. (2.103);

ii. an alternative, but equivalent, approach is to regroup the sum in terms of fixed values of N, writing

$$\Xi(T, V, \mu) = \sum_{N=0}^{+\infty} e^{\beta \mu N} \sum_i e^{-\beta E_i(N)}, \tag{2.104a}$$

$$= \sum_{N=0}^{+\infty} e^{\beta \mu N} Z(T, V, N), \tag{2.104b}$$

where $E_i(N)$ denotes the energy of state i at fixed N.

Notice that the second method allows us to understand the grand canonical partition function in terms of the more familiar canonical partition function.

We are able to relate partition functions to thermodynamic potentials. For example, for the canonical partition function we have

$$Z(T, V, N) = e^{-\beta F(T,V,N)}, \tag{2.105}$$

where $F(T, V, N)$ is the Helmholtz free energy. To find out the expression for the grand canonical partition function, we begin by noticing that

$$\Xi(T, V, \mu) = \sum_{N=0}^{+\infty} e^{\beta\mu N - \beta F(T,V,N)}, \tag{2.106a}$$

$$= \sum_{N=0}^{+\infty} e^{-\beta[F(T,V,N)-\mu N]}. \tag{2.106b}$$

To make the connection with thermodynamics, we may then approximate this sum by its maximum term, which is the one that dominates in the thermodynamic limit. In this situation, we'll get to

$$\Xi(T, V, \mu) \approx e^{-\beta \min_N \{F(T,V,N)-\mu N\}}, \tag{2.107a}$$

$$= e^{-\beta\Phi(T,V,\mu)}, \tag{2.107b}$$

where we used the fact that $\min_N \{F(T, V, N) - \mu N\}$ is just a Legendre transformation from the Helmholtz free energy to the grand potential[7] Φ. Therefore, we can write

$$\Phi(T, V, \mu) = -k_B T \log \Xi(T, V, \mu). \tag{2.108}$$

Our next step is then to compute the ensemble averages, with particular interest on the expressions for the mean energy and mean number of particles. For the mean number of particles we have

$$\langle N \rangle = \sum_i N_i p_i, \tag{2.109a}$$

$$= \frac{1}{\Xi} \sum_i N_i e^{-\beta E_i + \beta\mu E_i}, \tag{2.109b}$$

$$= \frac{1}{\beta\Xi} \frac{\partial}{\partial\mu} \sum_i e^{-\beta E_i + \beta\mu E_i}, \tag{2.109c}$$

$$= \frac{1}{\beta\Xi} \frac{\partial\Xi}{\partial\mu}, \tag{2.109d}$$

$$= \frac{1}{\beta} \frac{\partial}{\partial\mu} \log \Xi. \tag{2.109e}$$

---

[7]Kardar (2007b, Sec. 119) and Reichl (2016, Sec. 3.5.5) call it grand potential. Salinas (2001, p. 52) calls it "grand thermodynamic potentiall". Pathria and Beale (2022, Sec. 4.3) prefers "q-potential". Prof. Fiore referred to it as "grand canonical potential".

Notice next that

$$-\frac{1}{\Xi}\frac{\partial\Xi}{\partial\beta} = -\frac{1}{\Xi}\frac{\partial}{\partial\beta}\sum_i e^{-\beta E_i + \beta\mu E_i}, \tag{2.110a}$$

$$= -\frac{1}{\Xi}\sum_i (E_i - \mu N_i)e^{-\beta E_i + \beta\mu E_i}, \tag{2.110b}$$

$$= -\sum_i (E_i - \mu N_i)p_i, \tag{2.110c}$$

$$= -\langle E\rangle + \mu\langle N\rangle. \tag{2.110d}$$

Eqs. (2.109) and (2.110) on the previous page and on this page imply that

$$\langle E\rangle = -\frac{1}{\Xi}\frac{\partial\Xi}{\partial\beta} - \mu\langle N\rangle, \tag{2.111a}$$

$$= -\frac{\partial}{\partial\beta}\log\Xi - \frac{\mu}{\beta}\frac{\partial}{\partial\mu}\log\Xi. \tag{2.111b}$$

To simplify this expression for $\langle E\rangle$, we can introduce the fugacity[8],

$$z = e^{\beta\mu}, \tag{2.112}$$

in terms of which we may write

$$p_i = \frac{z^{N_i}e^{-\beta E_i}}{\Xi} \tag{2.113}$$

and

$$\Xi(T, V, z) = \sum_{N=0}^{+\infty} z^N Z(T, V, N). \tag{2.114}$$

In terms of the fugacity, the mean energy can be written as

$$\langle E\rangle = \frac{1}{\Xi}\sum_i E_i z^{N_i}e^{-\beta E_i}, \tag{2.115a}$$

$$= -\frac{1}{\Xi}\left(\frac{\partial}{\partial\beta}\sum_i z^{N_i}e^{-\beta E_i}\right)_z, \tag{2.115b}$$

$$= -\left(\frac{\partial}{\partial\beta}\log\Xi\right)_z. \tag{2.115c}$$

Eqs. (2.111) and (2.115) are consistent with each other because different variables are kept constant in each of them.

---

[8]Salinas (2001, p. 130) mentions the fugacity is sometimes called the "activity".

As for the number of particles, we have

$$\langle N \rangle = \frac{1}{\Xi} \sum_i N_i z^{N_i} e^{-\beta E_i}, \tag{2.116a}$$

$$= \frac{z}{\Xi} \left(\frac{\partial \Xi}{\partial z} \right)_\beta, \tag{2.116b}$$

$$= z \left(\frac{\partial}{\partial z} \log \Xi \right)_\beta. \tag{2.116c}$$

Using reduction of derivatives (see Callen 1985, Sec. 7.3) one can show that, for some quantity q,

$$\left(\frac{\partial q}{\partial \beta} \right)_z = \left(\frac{\partial q}{\partial \beta} \right)_\mu + \left(\frac{\partial q}{\partial \mu} \right)_\beta \left(\frac{\partial \mu}{\partial \beta} \right)_z, \tag{2.117a}$$

$$= \left(\frac{\partial q}{\partial \beta} \right)_\mu - \frac{\mu}{\beta} \left(\frac{\partial q}{\partial \mu} \right)_\beta, \tag{2.117b}$$

where we computed $\left(\frac{\partial \mu}{\partial \beta} \right)_z = -\frac{\mu}{\beta}$ from $z = e^{\beta \mu}$. Notice Eq. (2.117) is precisely the relation necessary for consistency between Eqs. (2.111) and (2.115) on the previous page, as one can see from plugging in $q = -\log \Xi$ into Eq. (2.117).

It remains for us to study the quantity

$$\left\langle (N - \langle N \rangle)^2 \right\rangle = \langle N^2 \rangle - \langle N \rangle^2, \tag{2.118}$$

which will be relevant to the grand canonical ensemble for a couple of reasons.

i. $\left\langle (N - \langle N \rangle)^2 \right\rangle$ is related to $\frac{\partial \langle N \rangle}{\partial \mu}$, and hence the positivity of one will imply the positivity of the other, ensuring that the number of particles increases with the chemical potential.

ii. This will allow us to notice that the fluctuations of the number of particles about the expected value become negligible in the thermodynamic limit, with the expected value matching the fixed value of N one would get from the canonical ensemble, hence showing the ensembles are equivalent.

Through the usual arguments, one can show that

$$\langle N^2 \rangle = \frac{1}{\Xi} \frac{1}{\beta^2} \frac{\partial^2 \Xi}{\partial \mu^2}. \tag{2.119}$$

A calculation analogous to the one we did in Eq. (2.5) on page 16 also lets us see that

$$\langle N^2 \rangle - \langle N \rangle^2 = \frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} \log \Xi, \tag{2.120a}$$

$$= \frac{1}{\beta} \frac{\partial \langle N \rangle}{\partial \mu}. \tag{2.120b}$$

Since $\langle N^2 \rangle - \langle N \rangle^2 \geq 0$ and $\beta > 0$, we conclude $\frac{\partial \langle N \rangle}{\partial \mu} \geq 0$.

In some situations, when dealing with phase transitions, one might get $\frac{\partial \langle N \rangle}{\partial \mu} < 0$. This is an incorrect result and must be fixed, but it might also be used as a hint of the occurrence of a phase transition.

To deal with the fluctuations, we notice that the thermodynamic number of particles N is the expected value $\langle N \rangle$. One can then show that (Salinas 2001, Sec. 7.2B)

$$\frac{\sqrt{\langle N^2 \rangle - \langle N \rangle^2}}{\langle N \rangle} = \frac{1}{N} \sqrt{\frac{N^2 \kappa_T}{\beta V}}, \tag{2.121a}$$

$$= \left(\frac{k_B T \kappa_T}{v} \right)^{\frac{1}{2}} \frac{1}{\sqrt{N}}, \tag{2.121b}$$

$$\to 0, \tag{2.121c}$$

where in the last line we took the thermodynamic limit and κ_T is the isothermal compressibility. Hence, we see that in the thermodynamic limit the fluctuations will vanish. This is in fact what allows us to understand the ensemble average $\langle N \rangle$ as the actual thermodynamic number of particles N, leading us also to an equivalence of the canonical and grand canonical ensembles.

We typically expect ensembles to be equivalent in the thermodynamic limit for homogeneous systems with short-range interactions. Near phase transitions, this might fail, as we can have $\kappa_T \to +\infty$, which renders the previous argument inconsistent (Eq. (2.121c) only holds if the coefficient on Eq. (2.121b) remains finite in the thermodynamic limit).

For long-range interactions, there is no consensus in the literature on whether there is equivalence of ensembles.

Maybe add diverging references?

Monoatomic Ideal Gas

As an example of how the grand canonical potential works, let us consider the monoatomic ideal gas one more time.

From Eq. (2.44) on page 26 we know the canonical partition function for the ideal gas is given by

$$Z(T, V, N) = \frac{V^N}{N!} \left[\frac{2\pi m}{\beta h^2} \right]^{\frac{3N}{2}}. \tag{2.122}$$

Therefore, the grand canonical partition function will be

$$\Xi(T, V, \mu) = \sum_{N=0}^{+\infty} z^N Z(T, V, N), \tag{2.123a}$$

$$= \sum_{N=0}^{+\infty} \frac{1}{N!} \left[\frac{2\pi m}{\beta h^2} \right]^{\frac{3N}{2}} z^N V^N, \tag{2.123b}$$

$$= \exp \left(\left[\frac{2\pi m}{\beta h^2} \right]^{\frac{3}{2}} z V \right). \tag{2.123c}$$

Hence, the internal energy as a function of fugacity (not number of particles) is given by

$$U(T, V, z) = \langle E \rangle, \tag{2.124a}$$

$$= -\frac{\partial}{\partial \beta} \log \Xi(T, V, z), \tag{2.124b}$$

$$= \frac{3}{2} \left[\frac{2\pi m}{\beta h^2} \right]^{\frac{3}{2}} \frac{zV}{\beta}. \tag{2.124c}$$

As for the expected number of particles, we get

$$N(T, V, z) = \langle N \rangle, \tag{2.125a}$$

$$= z\frac{\partial}{\partial z} \log \Xi(T, V, z), \tag{2.125b}$$

$$= \left[\frac{2\pi m}{\beta h^2} \right]^{\frac{3}{2}} zV. \tag{2.125c}$$

Notice that Eqs. (2.124) and (2.125) imply

$$\frac{\langle E \rangle}{\langle N \rangle} = \frac{3}{2} k_B T, \tag{2.126}$$

in agreement with the canonical ensemble.

2.6 Quantum Gases

So far, we have dealt with quantum systems in the sense we were able to compute the thermodynamic properties of some systems with discrete energy levels and which obeyed the Third Law of Thermodynamics. Nevertheless, we still haven't tackled one of the most important features of quantum systems: undistinguishability. So far, we were able to avoid it by assuming we were looking at localized particles, but this can't be done for a gas.
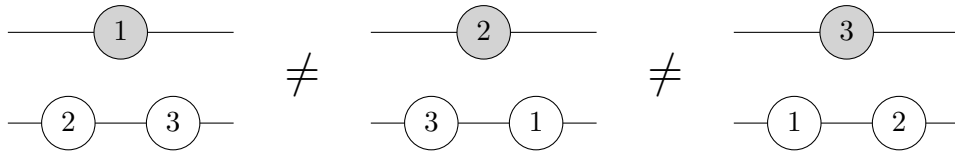

Figure 2.8: *When dealing with distinguishable particles, we have to account for permutations of particles counting as different states. In the figure, each picture represents a different particle of a set of three particles being on an excited state. If the particles were undistinguishable, the three states would be a single one.*

For distinguishable particles, different permutations of particles can lead to different states, as illustrated on Fig. 2.8. A gas with discrete states and such properties (such as the Boltzmann gas discussed by Salinas 2001, pp. 77–79) is said to obey Maxwell–Boltzmann

statistics. For quantum gases, particles will be undistinguishable, and they will obey other statistics.

The canonical partition function for the Boltzmann gas, which is simply a gas made of distinguishable particles with $\tilde{n}$ discrete energy levels, is

$$Z(T, V, N) = \left(e^{-\beta\epsilon_1} + e^{-\beta\epsilon_2} + \cdots + e^{-\beta\epsilon_{\tilde{n}}} \right)^N, \tag{2.127a}$$

$$= \sum_{\substack{n_1, \ldots, n_{\tilde{n}} \\ n_1 + \cdots + n_{\tilde{n}} = N}} \frac{N!}{n_1! \cdots n_{\tilde{n}}!} e^{-\beta(\epsilon_1 n_1 + \cdots + \epsilon_{\tilde{n}} n_{\tilde{n}})}, \tag{2.127b}$$

where both lines are related by the multinomial theorem. The combinatorial factors are introduced to account for the fact that permutations of particles lead to different states. When this is no longer true, we won't be able to use the multinomial theorem to compute the sum, and hence the canonical partition function will no longer be a convenient way of performing the calculations. Nevertheless, we will be able to obtain closed expressions by working with the grand canonical ensemble.

Identical Particles

In order to discuss quantum gases, we'll need to review some concepts about identical particles in Quantum Mechanics. While this review will be brief, more information can be found in the literature, such as in the Quantum Mechanics books by Griffiths (2005, Chap. 5), Sakurai and Napolitano (2017, Chap. 7), Shankar (1994, Sec. 10.3), and Weinberg (2015, Sec. 4.5, 2021, Sec. 5.5), or the Statistical Mechanics books by Kardar (2007b, Sec. 7.1), Pathria and Beale (2022, Sec. 5.4), and Salinas (2001, Chap. 8).

In classical mechanics, permuting two particles leads to a different state. However, this is often not the case in Quantum Mechanics. Let us define a permutation operator $\hat{P}$ by

$$\hat{P}\psi(1, 2) = \psi(2, 1), \tag{2.128}$$

i.e., it "switches" the states of two particles. For identical particles, this doesn't change the physical state.

Permuting twice should bring us back to our original wavefunction, $\hat{P}^2\psi = \psi$. Hence, we identify two cases

- $\hat{P}\psi = +\psi$, corresponding to bosons, which are particles of integer spin;

- $\hat{P}\psi = -\psi$, corresponding to fermions, which are particles of half-integer.

This connection between spin and parity (or, more appropriately, spin and statistics) can be understood under the light of Quantum Field Theory (Streater and Wightman 2000, Sec. 4.4).

It is interesting to notice that imposing statistics—*i.e.*, imposing that wavefunctions should always be an eigenstate of the permutation operator with eigenvalue ± 1—leads to correlations between the particles even in the absence of interactions. For example, two

fermions with the same spin can't be on the same position simultaneously, leading to a "repulsion effect".

The totally antisymmetric wavefunction of fermions is given, in the absence of spin[9], by the Slater determinant,

$$\Psi_A(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{1}{\sqrt{n!}} \det \begin{pmatrix} \psi_1(\mathbf{r}_1) & \cdots & \psi_1(\mathbf{r}_n) \\ \vdots & \ddots & \vdots \\ \psi_n(\mathbf{r}_1) & \cdots & \psi_n(\mathbf{r}_n) \end{pmatrix}. \tag{2.129}$$

Notice that it implies that $\Psi_A = 0$ if any two fermions have the same state. In the particular case of $n = 2$ we find

$$\Psi_A(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}(\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) - \psi_1(\mathbf{r}_2)\psi_2(\mathbf{r}_1)). \tag{2.130}$$

Notice that $\hat{P}\Psi_A = -\Psi_A$.

For bosons, the totally symmetric wavefunction can be written in terms of a permanent, which is similar to a determinant, but with only positive signs. Hence,

$$\Psi_S(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{1}{\sqrt{n!}} \operatorname{perm} \begin{pmatrix} \psi_1(\mathbf{r}_1) & \cdots & \psi_1(\mathbf{r}_n) \\ \vdots & \ddots & \vdots \\ \psi_n(\mathbf{r}_1) & \cdots & \psi_n(\mathbf{r}_n) \end{pmatrix}. \tag{2.131}$$

This leads to $\hat{P}\Psi_S = +\Psi_S$ and to two bosons being able to be on the same state.

It should be noticed that these symmetry properties are not mere theoretical simplifications, but rather experimental facts about nature with physical consequences. Chemistry is built upon Pauli's exclusion principle, which states two fermions (such as electrons) can't be on the same quantum state simultaneously. Bosons are capable of forming condensate states in which all particles are at the ground state, but fermions are not. Fermions can have positive chemical potential, while bosons can't. And so on.

Statistical Mechanics of Free Bosons and Fermions

Let us denote by n_j the number of particles on an energy level ϵ_j. It is common to borrow from Chemistry the name "orbital" for each energy level in this context of Quantum Statistical Mechanics, hence so shall we do. The total energy would then be

$$E = \sum_j n_j \epsilon_j. \tag{2.132}$$

For bosons, n_j can take any non-negative integer values. For fermions, n_j can take only the values 0 and 1, since two or more fermions can't occupy the same orbital.

---

[9]Fermions are always particles of half-integer spin, so they can't really be spinless, but we can treat this as an approximation to the more complicated problem that considers the spin as well. For our purposes, spin will only be relevant by means of degeneracy effects or interactions with an external magnetic field, so we can ignore it for now.

Since spin would only contribute to the problem through interaction with an external magnetic field or by adjusting some factors due to degeneracy, we'll ignore it for now.

In the canonical ensemble, we have

$$Z(T,V,N) = \sum_{\substack{n_1,\dots,n_{\tilde{n}} \\ n_1+\cdots+n_{\tilde{n}}=N}} e^{-\beta(\epsilon_1 n_1 + \cdots + \epsilon_{\tilde{n}} n_{\tilde{n}})}. \tag{2.133}$$

There is nothing wrong with this expression, but we are not able to express the sum in a closed form. Hence, this approach it not particularly convenient.

On the other hand, we can write the grand canonical partition function as

$$\Xi(T,V,\mu) = \sum_{N=0}^{+\infty} e^{\beta\mu N} \sum_{\substack{n_1,\dots,n_{\tilde{n}} \\ n_1+\cdots+n_{\tilde{n}}=N}} e^{-\beta(\epsilon_1 n_1 + \cdots + \epsilon_{\tilde{n}} n_{\tilde{n}})}. \tag{2.134}$$

The trick is to now notice that summing over n_1, ..., $n_{\tilde{n}}$ with the restriction of $n_1+\cdots+n_{\tilde{n}} = N$ and later summing over N is identical to just summing over n_1, ..., $n_{\tilde{n}}$ with no restrictions at all (*cf.* the illustration on Fig. 2.9 on the next page). Hence,

$$\Xi(T,V,\mu) = \sum_{N=0}^{+\infty} e^{\beta\mu N} \sum_{\substack{n_1,\dots,n_{\tilde{n}} \\ n_1+\cdots+n_{\tilde{n}}=N}} e^{-\beta(\epsilon_1 n_1 + \cdots + \epsilon_{\tilde{n}} n_{\tilde{n}})}, \tag{2.135a}$$

$$= \sum_{N=0}^{+\infty} \sum_{\substack{n_1,\dots,n_{\tilde{n}} \\ n_1+\cdots+n_{\tilde{n}}=N}} e^{\beta\mu N} e^{-\beta(\epsilon_1 n_1 + \cdots + \epsilon_{\tilde{n}} n_{\tilde{n}})}, \tag{2.135b}$$

$$= \sum_{N=0}^{+\infty} \sum_{\substack{n_1,\dots,n_{\tilde{n}} \\ n_1+\cdots+n_{\tilde{n}}=N}} e^{\beta\mu(n_1+\cdots+n_{\tilde{n}})} e^{-\beta(\epsilon_1 n_1 + \cdots + \epsilon_{\tilde{n}} n_{\tilde{n}})}, \tag{2.135c}$$

$$= \sum_{n_1,\dots,n_{\tilde{n}}} e^{\beta\mu(n_1+\cdots+n_{\tilde{n}})} e^{-\beta(\epsilon_1 n_1 + \cdots + \epsilon_{\tilde{n}} n_{\tilde{n}})}, \tag{2.135d}$$

$$= \left(\sum_{n_1} e^{-\beta(\epsilon_1-\mu)n_1} \right) \cdots \left(\sum_{n_{\tilde{n}}} e^{-\beta(\epsilon_{n_{\tilde{n}}}-\mu)n_{n_{\tilde{n}}}} \right), \tag{2.135e}$$

$$= \prod_j \left(\sum_n e^{-\beta(\epsilon_j-\mu)n} \right). \tag{2.135f}$$

Notice that the absence of interactions allowed us to decompose the grand canonical partition function in a product of orbitals (not of particles).

In possession of Eq. (2.135), we now need to treat bosons and fermions differently. For bosons, we get

$$\Xi_{\text{BE}}(T,V,\mu) = \prod_j \left(1 + e^{-\beta(\epsilon_j-\mu)} + e^{-2\beta(\epsilon_j-\mu)} + \cdots \right), \tag{2.136a}$$

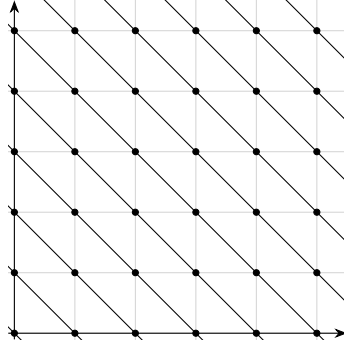$$= \prod_j \frac{1}{1 - e^{-\beta(\epsilon_j-\mu)}}, \tag{2.136b}$$

Figure 2.9: *To sum the points on each individual line (which are defined by $x + y = $ constant) and then sum over the lines is the same thing as just summing over all dots. A similar idea applies to Eq. (2.135) on the previous page.*

where "BE" stands for Bose–Einstein statistics, the name given for this case.

For fermions, we get Fermi–Dirac statistics, which has

$$\Xi_{\text{FD}}(T, V, \mu) = \prod_j \left(1 + e^{-\beta(\epsilon_j - \mu)}\right). \tag{2.137}$$

Hence, we can write both cases in a single line by writing

$$\log \Xi_{\text{BE/FD}} = \mp \sum_j \log\left(1 \mp e^{-\beta(\epsilon_j - \mu)}\right), \tag{2.138}$$

where the upper sign stands for Bose–Einstein statistics and the lower sign for Fermi–Dirac statistics.

From Eq. (2.138) we can obtain the grand potential,

$$\Phi_{\text{BE/FD}}(T, V, \mu) = \pm k_B T \log\left(1 \pm e^{-\beta(\epsilon_j - \mu)}\right), \tag{2.139}$$

which now allows us to compute the mean number of particles, $N = \sum_j \langle n_j \rangle = -\left(\frac{\partial \Phi}{\partial \mu}\right)_{T,V}$. We find

$$N_{\text{BE/FD}} = \mp k_B T \sum_j \frac{(\mp \beta) e^{-\beta(\epsilon_j - \mu)}}{1 \pm e^{-\beta(\epsilon_j - \mu)}}, \tag{2.140a}$$

$$= \sum_j \frac{e^{-\beta(\epsilon_j - \mu)}}{1 \pm e^{-\beta(\epsilon_j - \mu)}}, \tag{2.140b}$$

$$= \sum_j \frac{1}{e^{\beta(\epsilon_j - \mu)} \pm 1}. \tag{2.140c}$$

Hence, we can also write

$$\langle n_j \rangle_{\text{BE/FD}} = \frac{1}{e^{\beta(\epsilon_j - \mu)} \pm 1}. \tag{2.141}$$

Firstly, we notice that, as expected, Eq. (2.141) on the preceding page predicts $0 \leq \langle n_j \rangle_{\text{FD}} \leq 1$. For bosons, we notice that we must have $\langle n_j \rangle_{\text{BE}} \geq 0$, which implies

$e^{\beta \epsilon_j} \geq e^{\beta \mu}$. Hence, $\epsilon_j \geq \mu$ for all j, in particular for the ground state. This shows that the chemical potential of a gas of bosons is bounded from above by ϵ_0. If $\epsilon_0 = 0$ (as we can achieve by redefining the energy with a constant), the chemical potential can never be positive.

Classical Limit

We should point out that the difference between bosons and fermions ends up lying on the sign of Eq. (2.141) on the previous page. Nevertheless, in the classical limit the exponential dominates over the constant 1, and hence we obtain

$$\langle n_j \rangle \approx e^{-\beta(\epsilon_j - \mu)}. \tag{2.142}$$

Notice that the classical limit $e^{\beta(\epsilon_j - \mu)} \gg 1$ can also be written as $e^{\beta \epsilon_j} \gg e^{\beta \mu}$. Since this last expression should hold for any j, this implies $1 \gg e^{\beta \mu}$.

For the grand canonical partition function, the classical limit yields (*cf.* Eq. (2.138) on the preceding page)

$$\log \Xi \approx \sum_j e^{-\beta(\epsilon_j - \mu)}. \tag{2.143}$$

To proceed, let us choose the particular case of an ideal gas. The spectrum of such a gas is given by

$$\epsilon_{\mathbf{k}} = \frac{\hbar^2 \|\mathbf{k}\|^2}{2m}, \tag{2.144}$$

where the wave vector $\mathbf{k}$ is given by

$$\mathbf{k} = \frac{2\pi}{L}\mathbf{n}, \tag{2.145}$$

with $\mathbf{n} \in \mathbb{Z}^3$. This quantization condition is enforced by imposing boundary conditions on the problem (we can choose for the wavefunctions to vanish on the walls of a cubic box, or choose periodic boundary conditions). In the thermodynamic limit, the dimensions L of the box will go to infinity, and hence these boundary conditions won't lead to loss of generality.

In fact, in the thermodynamic limit, we can approximate the discrete sum defining the grand canonical partition function by an integral, and hence write

$$\log \Xi \approx \sum_{\mathbf{k}} e^{-\beta(\epsilon_{\mathbf{k}} - \mu)}, \tag{2.146a}$$

$$\approx \frac{V}{(2\pi)^3} \int e^{-\beta(\epsilon_{\mathbf{k}} - \mu)} \, \mathrm{d}^3 k, \tag{2.146b}$$

$$= \frac{V}{(2\pi)^3} \int e^{-\beta\left(\frac{\hbar^2 \|\mathbf{k}\|^2}{2m} - \mu\right)} \, \mathrm{d}^3 k, \tag{2.146c}$$

$$= \frac{4\pi V}{(2\pi)^3} e^{\beta \mu} \int e^{-\beta \frac{\hbar^2 k^2}{2m}} k^2 \, \mathrm{d}k, \tag{2.146d}$$

$$= \frac{V}{(2\pi)^3} e^{\beta \mu} \left(\frac{2\pi m}{\beta \hbar^2}\right)^{\frac{3}{2}}. \tag{2.146e}$$

Hence, we find the grand potential

$$\Phi(T, V, \mu) = -(k_B T)^{\frac{5}{2}} V \left(\frac{2\pi m}{\hbar^2}\right)^{\frac{3}{2}} e^{\frac{\mu}{k_B T}}. \qquad (2.147)$$

Notice that Eq. (2.123) on page 45, computed for the monoatomic classical ideal gas, and Eq. (2.146), obtained as the classical limit of the ideal quantum gas, are the same expression (recall that $h = 2\pi\hbar$). Hence, we are recovering the classical gas as a limiting case of the quantum gas. Furthermore, notice that this time we did not need to introduce the ad hoc factor of $N!$ needed for the classical gas. Hence, we can understand that factor as a requirement that the classical gas be the classical limit of the quantum gas.

At last, we should also ask ourselves how to know when the classical limit is applicable in terms of macroscopic variables. Let us notice that

$$N = -\left(\frac{\partial \Phi}{\partial \mu}\right)_{T,V}, \qquad (2.148a)$$

$$= (k_B T)^{\frac{3}{2}} V \left(\frac{2\pi m}{\hbar^2}\right)^{\frac{3}{2}} e^{\frac{\mu}{k_B T}}. \qquad (2.148b)$$

Hence, the classical limit is characterized by

$$e^{\frac{\mu}{k_B T}} = \frac{\hbar^3}{(2\pi m k_B T)^{\frac{3}{2}}} \frac{N}{V}, \qquad (2.149a)$$

$$\ll 1, \qquad (2.149b)$$

and hence the classical limit can be understood as a limit of high temperatures and/or low densities.

2.7 Ultracold Fermi Gases

At low temperatures, we typically expect particles to be at the ground state. Nevertheless, due to the Pauli exclusion principle, this is not possible for fermions. Instead, they "pile up" on the lower energy states, with no two or more fermions ever occupying the same state. Fig. 2.10 shows graphs of the occupation numbers of state as a function of the state's energy for finite and vanishing temperatures. We notice that the lower energy states will always be excited, while the states with energies closer to the so-called Fermi energy ϵ_F can get excited or unexcited as a consequence of thermal fluctuations.

Since the Pauli exclusion principle prevents the lower energy states from feeling effects due to temperature, they won't contribute to the specific heat. The states with energies close to the Fermi energies, though, will contribute.

The Fermi energy ϵ_F is define as the chemical potential at vanishing temperature[10]. This is motivated by the behavior of the number of particles per state at vanishing

---

[10]While we were working at the grand canonical potential to derive the formulae we'll be working with, it is experimentally more feasible to control the particle density than the chemical potential. Hence, we need to invert a few relations and eventually obtain the chemical potential as a function determined by temperature and density.
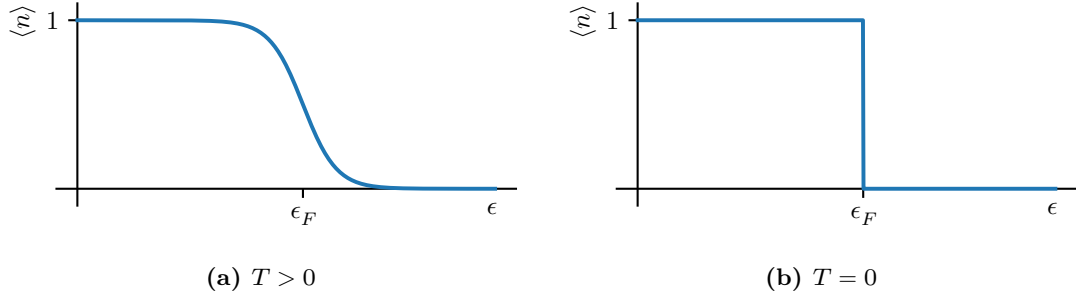
(a) $T > 0$ | **(b)** $T = 0$

Figure 2.10: *Plot of the occupation numbers of fermions as a function of orbital energy, obtained by plotting Eq. (2.141) on page 50 with the fermionic choice of sign while keeping temperature and chemical potential constant. Notice that lower energy states are filled. The relevance of temperature fluctuations is larger for the states close to the Fermi energy ϵ_F, for they can be excited or unexcited with temperature variations. Hence, they are the ones that contribute the most to the specific heat.*

temperature. Notice that Eq. (2.141) on page 50 implies

$$\lim_{T \to 0} \langle n_j \rangle_{\text{FD}} = \lim_{T \to 0} \frac{1}{e^{\beta(\epsilon_j - \mu)} \pm 1}, \tag{2.150a}$$

$$= \Theta\big(\mu|_{T=0} - \epsilon_j\big). \tag{2.150b}$$

Hence, it is convenient to define $\epsilon_F \equiv \mu|_{T=0}$.

Completely Degenerate Fermi Gas

Let us then obtain some thermodynamic quantities for the ideal Fermi gas. Since the computations involved will be pretty difficult, we'll focus on the particular cases of vanishing and small temperatures, although other cases can be dealt with, *e.g.*, numerical methods. For vanishing temperature, we say the gas is completely degenerate.

The internal energy and number of particles are given by

$$U = \sum_j \langle n_j \rangle \, \epsilon_j, \tag{2.151}$$

$$N = \sum_j \langle n_j \rangle \,. \tag{2.152}$$

For an ideal gas in the thermodynamic limit we can write

$$N = \sum_{\mathbf{k}} \frac{1}{\exp\!\left(\frac{\hbar^2 \|\mathbf{k}\|^2}{2m} - \mu\right) + 1}, \tag{2.153a}$$

$$\approx \left(\frac{L}{2\pi}\right)^3 \int \frac{1}{\exp\!\left(\frac{\hbar^2 \|\mathbf{k}\|^2}{2m} - \mu\right) + 1} \, \mathrm{d}^3 k \,, \tag{2.153b}$$

$$= 4\pi \left(\frac{L}{2\pi}\right)^3 \int_0^{+\infty} \frac{k^2}{\exp\!\left(\frac{\hbar^2 k^2}{2m} - \mu\right) + 1} \, \mathrm{d}k \,. \tag{2.153c}$$

At zero temperature we can then write

$$N = 4\pi\left(\frac{L}{2\pi}\right)^3 \int_0^{k_F} k^2\,\mathrm{d}k\,, \tag{2.154a}$$

$$= \frac{4\pi V k_F^3}{3(2\pi^3)}, \tag{2.154b}$$

where k_F is defined in terms of the Fermi energy through

$$\epsilon_F = \frac{\hbar^2 k_F^2}{2m}. \tag{2.155}$$

Notice, however, that this result does not take into account eventual degeneracies of the energy spectrum. If all states have a γ-fold degeneracy, then the correct number of particles will be

$$N = \frac{4\pi\gamma V k_F^3}{3(2\pi^3)}. \tag{2.156}$$

Electrons have twofold degeneracy due to being spin $\frac{1}{2}$ particles. Hence, for electrons, we have

$$N = \frac{8\pi V k_F^3}{3(2\pi^3)}. \tag{2.157}$$

The internal energy can be dealt with in a similar fashion. We'll have, by means of the same steps,

$$U = \gamma \sum_{\mathbf{k}} \frac{1}{\exp\left(\frac{\hbar^2\|\mathbf{k}\|^2}{2m} - \mu\right) + 1} \frac{\hbar^2\|\mathbf{k}\|^2}{2m}, \tag{2.158a}$$

$$= 4\pi\gamma\left(\frac{L}{2\pi}\right)^3 \int_0^{+\infty} \frac{k^4}{\exp\left(\frac{\hbar^2 k^2}{2m} - \mu\right) + 1}\,\mathrm{d}k\,, \tag{2.158b}$$

$$= \frac{\gamma\hbar^2 V k_F^5}{20\pi^2 m}. \tag{2.158c}$$

Therefore, we learn that, for any value of γ,

$$\frac{U}{N} = \frac{3}{5}\epsilon_F \tag{2.159}$$

From Eqs. (2.155), (2.158) and (2.159) on pages 54–55 we can obtain

$$\epsilon_F = \frac{\hbar^2}{2m}\left(\frac{6\pi^2 N}{\gamma V}\right)^{\frac{2}{3}}. \tag{2.160}$$

We can define the Fermi temperature T_F by

$$\epsilon_F = k_B T_F. \tag{2.161}$$

Using Eqs. (2.160) and (2.161) we can obtain the Fermi temperature of a given material from properties such as its density, molecular mass, and so on. For sodium, one has $T_F \approx 10^4\,\mathrm{K}$. For copper, $T_F \approx 8 \times 10^4\,\mathrm{K}$. For many metals, the Fermi temperature is then way higher than ambient temperature, $T \approx 300\,\mathrm{K}$. Hence, we can treat these metals at ambient temperature in a low-temperature expansion.

Degenerate Fermi Gas

Since for metals the Fermi temperature is often very high, we can obtain appropriate descriptions for low temperatures. In this case, we say the gas is degenerate. However, before we dive into the theory, it is worth recalling what we are ignoring.

In a metal, we do need to describe the behavior of the electrons on the conduction band, but there are also phonons that should be accounted for. Phonons are what we obtain once we quantize the vibrations of the atoms in the lattice. If we consider each of them as free, the specific heat will have the form

$$c_v = \underbrace{c_v^{\text{phonons}}}_{\sim T^3} + \underbrace{c_v^{\text{electrons}}}_{\sim T^2}. \tag{2.162}$$

Electrons interact with phonons, so we could take these interactions into account. We won't. This simplification might be justified under the light of Eq. (2.162), which suggests such contributions would be negligible.

Electrons could also collide with each other, but we'll also ignore this effect. Since each collision must respect energy and momentum conservation in addition to Pauli's exclusion principle, few energy levels would be admissible, and hence these contributions are suppressed.

Our procedure to study the Fermi gas will be to perform a low temperature expansion, but another possibility would be to do as Kardar (2007b, Sec. 7.4) and Pathria and Beale (2022, Sec. 8.1) and show that the pressure and density of the gas can be written as

$$\frac{p}{k_B T} = \frac{g}{\lambda^3} f_{\frac{5}{2}}(z) \quad \text{and} \quad \frac{N}{V} = \frac{g}{\lambda^3} f_{\frac{3}{2}}(z), \tag{2.163}$$

where g is a weight factor (determined, *e.g.*, by spin or other internal structure of the particles), λ is the thermal wavelength, and the Fermi–Dirac functions $f_\nu(z)$ are given by

$$f_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^{+\infty} \frac{x^{\nu-1}}{z^{-1}e^x + 1} \, dx. \tag{2.164}$$

Nevertheless, we'll follow a different approach, favored, *e.g.*, by Salinas (2001, Chap. 9). Kardar (2007b, Sec. 7.5) seems to discuss how both approaches relate to each other.

Let us denote the density of states as a function of energy by $\mathcal{D}(\epsilon)$. Then a change of variables from $\mathbf{k}$ to ϵ would lead us to

$$\mathcal{D}(\epsilon) = \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{\frac{3}{2}} \epsilon^{\frac{1}{2}}. \tag{2.165}$$

For generality, we'll keep γ unspecified. In this case, the expressions for the grand canonical

Does this make sense?

partition function, internal energy, and number of particles become

$$\log \Xi = \gamma V \int_0^{+\infty} \mathcal{D}(\epsilon) \log\big(1 + e^{-\beta(\epsilon-\mu)}\big) \, \mathrm{d}\epsilon \,, \tag{2.166}$$

$$U = \gamma V \int_0^{+\infty} \epsilon \mathcal{D}(\epsilon) f(\epsilon) \, \mathrm{d}\epsilon \,, \tag{2.167}$$

$$N = \gamma V \int_0^{+\infty} \mathcal{D}(\epsilon) f(\epsilon) \, \mathrm{d}\epsilon \,, \tag{2.168}$$

where we denoted

$$f(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)} + 1}. \tag{2.169}$$

Hence, our present goal is to understand integrals of the form

$$I(T) = \int_0^{+\infty} g(\epsilon) f(\epsilon) \, \mathrm{d}\epsilon \,, \tag{2.170a}$$

$$= \int_0^{+\infty} \frac{g(\epsilon)}{e^{\beta(\epsilon-\mu)} + 1} \, \mathrm{d}\epsilon \,. \tag{2.170b}$$

Let us then study $I(T)$ for small temperatures. Once we obtain general results for these sorts of integrals, we'll be able to apply them to the specific cases we're interested in.

We can write

$$I(0) = \int_0^{\mu} g(\epsilon) \, \mathrm{d}\epsilon \,, \tag{2.171}$$

due to the behavior of the fermion occupation number at zero temperature. As a consequence, notice that

$$I(T) - I(0) = \int_0^{\mu} \frac{g(\epsilon)}{e^{\beta(\epsilon-\mu)} + 1} \, \mathrm{d}\epsilon + \int_\mu^{+\infty} \frac{g(\epsilon)}{e^{\beta(\epsilon-\mu)} + 1} \, \mathrm{d}\epsilon - \int_0^{\mu} g(\epsilon) \, \mathrm{d}\epsilon \,, \tag{2.172a}$$

$$= \int_\mu^{+\infty} \frac{g(\epsilon)}{e^{\beta(\epsilon-\mu)} + 1} \, \mathrm{d}\epsilon - \int_0^{\mu} \frac{g(\epsilon)}{1 + e^{-\beta(\epsilon-\mu)}} \, \mathrm{d}\epsilon \,, \tag{2.172b}$$

$$= k_B T \int_0^{+\infty} \frac{g(\mu + k_B T \xi)}{e^\xi + 1} \, \mathrm{d}\xi - \int_{\beta\mu}^0 \frac{g(\mu - k_B T \xi)}{e^\xi + 1} \, \mathrm{d}\xi \,, \tag{2.172c}$$

where in the last line we changed variables according to $\xi = \beta(\epsilon - \mu)$ in the first integral and $\xi = -\beta(\epsilon - \mu)$ in the second.

At this stage, since we're assuming low temperatures, we can make two approximations. The first of them is to consider the Taylor expansions

$$g(\mu \pm k_B T \xi) = g(\mu) \pm k_B T \xi g'(\mu) + \frac{(k_B T)^2 \xi^2}{2} g''(\mu) + \cdots. \tag{2.173}$$

The second of them is to take $\beta\mu$, the lower limit of the second integral of Eq. (2.172) on page 57, to infinity. Under these two approximations, Eq. (2.172) leads to

$$I(T) - I(0) \approx 2g'(\mu)(k_B T)^2 \int_0^{+\infty} \frac{\xi}{e^\xi + 1} \, \mathrm{d}\xi \,. \tag{2.174}$$

This integral on Eq. (2.174) can be computed, and yields $\frac{\pi^2}{12}$. Hence, at last we obtain the approximation

$$\int_0^{+\infty} \frac{g(\epsilon)}{e^{\beta(\epsilon-\mu)} + 1} \, d\epsilon \approx \int_0^\mu g(\epsilon) \, d\epsilon + \frac{\pi^2}{6} g'(\mu)(k_B T)^2. \tag{2.175}$$

With this expression at hands, we can now start computing the expressions for the grand canonical partition function, internal energy and number of particles in the low temperature limit. This asymptotic approximation is known as Sommerfeld expansion.

Using the Sommerfeld expansion, we can find from Eqs. (2.165), (2.167) and (2.168) on the preceding page that

$$U = \frac{\gamma V}{10\pi^2} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \mu^{\frac{5}{2}} \left[1 + \frac{5\pi^2}{8} \left(\frac{k_B T}{\mu}\right)^2\right], \tag{2.176}$$

$$N = \frac{\gamma V}{6\pi^2} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \mu^{\frac{3}{2}} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu}\right)^2\right]. \tag{2.177}$$

In experimental situations, it is easier to control the number of particles than the chemical potential, and hence we'd rather work with $U(T, V, N)$ instead of $U(T, V, \mu)$. The grand canonical ensemble has served its purpose for us in allowing us to obtain analytical expression, but we now would like to turn away from it in favor of quantities that are easier to control in practice.

Let us then invert the expression $N = N(T, V, \mu)$ to obtain μ as a function of temperature, volume, and number of particles. We'll begin by noticing that if we plug Eq. (2.177) on the previous page on Eq. (2.160) on page 55 we'll get to

$$\epsilon_F = \frac{\hbar^2}{2m} \left(\frac{6\pi^2}{\gamma V} \frac{\gamma V}{6\pi^2} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \mu^{\frac{3}{2}} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu}\right)^2\right]\right)^{\frac{2}{3}}, \tag{2.178a}$$

$$= \mu \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu}\right)^2\right]^{\frac{2}{3}}. \tag{2.178b}$$

Using this result, we see that, at second order in temperature,

$$\mu = \epsilon_F \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu}\right)^2\right]^{-\frac{2}{3}}, \tag{2.179a}$$

$$\approx \epsilon_F \left[1 - \frac{2}{3} \frac{\pi^2}{8} \left(\frac{k_B T}{\mu}\right)^2\right], \tag{2.179b}$$

$$\approx \epsilon_F \left[1 - \frac{2}{3} \frac{\pi^2}{8} \left(\frac{k_B T}{\epsilon_F}\right)^2\right], \tag{2.179c}$$

$$= \epsilon_F \left[1 - \frac{2}{3} \frac{\pi^2}{8} \left(\frac{T}{T_F}\right)^2\right], \tag{2.179d}$$

where we neglected higher order corrections[11]. Notice that these expressions imply that the chemical potential for the fermion gas is always positive and that it diminishes as the temperature increases.

Using Eqs. (2.176), (2.177) and (2.179) on pages 57–58 we find

$$\frac{U}{N} = \frac{3\mu}{5}\left[1 + \frac{5\pi^2}{8}\left(\frac{k_B T}{\mu}\right)^2\right]\left[1 + \frac{\pi^2}{8}\left(\frac{k_B T}{\mu}\right)^2\right]^{-1}, \tag{2.180a}$$

$$\approx \frac{3\mu}{5}\left[1 + \frac{5\pi^2}{8}\left(\frac{k_B T}{\mu}\right)^2\right]\left[1 - \frac{\pi^2}{8}\left(\frac{k_B T}{\mu}\right)^2\right], \tag{2.180b}$$

$$\approx \frac{3\mu}{5}\left[1 + \frac{5\pi^2}{8}\left(\frac{T}{T_F}\right)^2\right]\left[1 - \frac{\pi^2}{8}\left(\frac{T}{T_F}\right)^2\right], \tag{2.180c}$$

$$\approx \frac{3\epsilon_F}{5}\left[1 - \frac{\pi^2}{12}\left(\frac{T}{T_F}\right)^2\right]\left[1 + \frac{5\pi^2}{8}\left(\frac{T}{T_F}\right)^2\right]\left[1 - \frac{\pi^2}{8}\left(\frac{T}{T_F}\right)^2\right], \tag{2.180d}$$

$$\approx \frac{3\epsilon_F}{5}\left[1 + \frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2\right], \tag{2.180e}$$

where once again we've kept the terms only up to second order.

The specific heat at constant volume is then given by

$$c_v = \left(\frac{\partial U}{\partial T}\right)_{V,N}, \tag{2.181a}$$

$$= \frac{3}{5}N\epsilon_F\frac{5}{6}\pi^2\frac{T}{T_F^2}, \tag{2.181b}$$

$$= \frac{3}{2}Nk_B\left(\frac{\pi^2}{3}\frac{T}{T_F}\right). \tag{2.181c}$$

The final factor is often of the order of 10^{-2}, which means the specific heat is way smaller than the expected classical value of $\frac{3}{2}Nk_B$. This can be understood from the fact that, at low temperatures, most electrons are filling the lower energy levels and won't contribute to conduction Hence, few electrons are actually contributing to the specific heat.

At last, it is interesting to point out that fermions have a non-vanishing pressure (and even a non-vanishing isothermal compressibility) at vanishing temperature, as a consequence of the Pauli exclusion principle.

Show it!

2.8 Bose–Einstein Condensation

Since bosons do not obey the Pauli exclusion principle, their behavior at low temperatures will be fairly different from the fermionic situations.

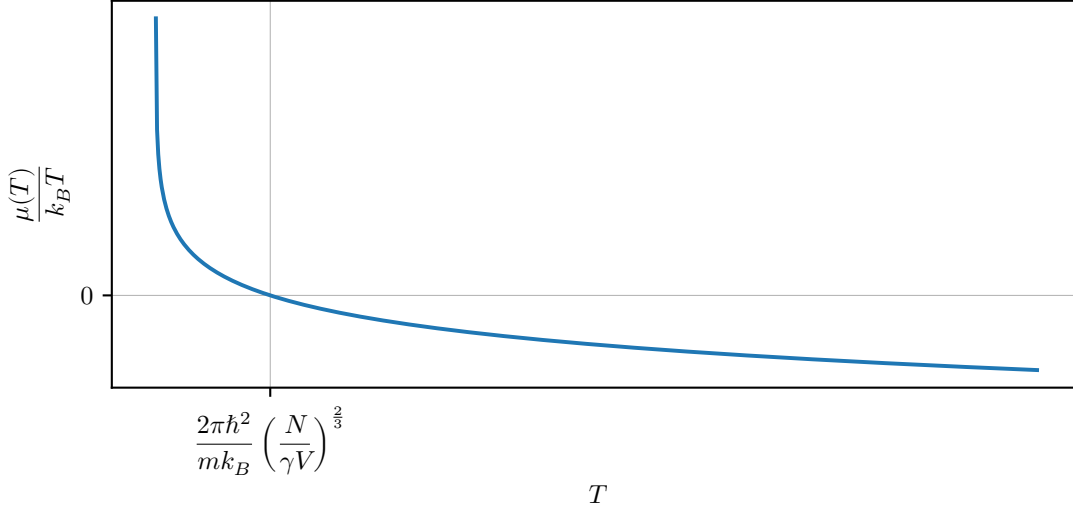[11]To consider them would be inconsistent with our use of the Sommerfeld expansion.

Figure 2.11: *Chemical potential as a function of temperature for the classical ideal gas. Notice that, as expected, the chemical potential vanishes at a scale in which the quantum effects are no longer negligible.*

First and foremost, let us recall that, in accordance with Eq. (2.149) on page 52, quantum effects will be relevant when

$$\frac{\hbar^3}{(2\pi m k_B T)^{\frac{3}{2}}} \frac{N}{V} \gtrsim 1, \tag{2.182}$$

which characterizes the condition we'll be working in. Furthermore, we recall that, as argued just after Eq. (2.141) on page 50, the chemical potential for bosonic systems is never positive: $\mu \leq 0$.

We then ask ourselves: if we hold N constant (as is usually the case in practical situations), how will $\mu(T)$ behave?

For large temperatures, it should resemble the classical ideal gas, since we're back in the classical limit. For a classical ideal gas, the chemical potential is given by (see Salinas 2001, Eq. (10.6))

$$\frac{\mu(T)}{k_B T} = \log\left(\frac{N}{\gamma V}\left(\frac{2\pi\hbar^2}{m k_B T}\right)^{\frac{3}{2}}\right), \tag{2.183}$$

which is plotted on Fig. 2.11 on the next page. Notice that μ vanishes for

$$T = \frac{2\pi\hbar^2}{m k_B}\left(\frac{N}{\gamma V}\right)^{\frac{2}{3}}. \tag{2.184}$$

As we mentioned, quantum effects kick in at scales in which Eq. (2.184) hold, and hence the boson gas will not follow the classical curve exactly up to the point in which the chemical potential vanishes. This is not unexpected, since we know $\mu \leq 0$ must hold.

Nevertheless, there will still be a value of the temperature for which the chemical potential will vanish. Below this critical temperature, the chemical potential will remain at zero.

As a preview of our next results, once that critical temperature is achieved, we'll find that there will be a maximum limit on the occupation number of excited states as bosons condense on the ground state. Notice this behavior is fairly different from what we get in a "normal phase" (as opposed to a condensed or coexisting phase), in which the occupation density of each individual orbital tends to zero in the thermodynamic limit.

This is then the summary of the phenomenon known as Bose–Einstein condensation: a macroscopic accumulation of particles at the ground state below a threshold temperature at which the chemical potential vanishes.

Let us then proceed to the calculations. In the grand canonical ensemble for bosons, the gas' pressure and density are given by

$$\frac{p}{k_B T} = -\frac{1}{V} \sum_{\epsilon} \log\big(1 - e^{-\beta(\epsilon - \mu)}\big), \tag{2.185}$$

and

$$\frac{N}{V} = \frac{1}{V} \sum_{\epsilon} \frac{1}{e^{\beta(\epsilon - \mu)} - 1}. \tag{2.186}$$

For a free boson gas ($\epsilon_{\mathbf{k}} = \frac{\hbar^2 \|\mathbf{k}\|^2}{2m}$) in the normal phase, we can write these sums as integrals according to

$$\sum_{\mathbf{k}} \log\Big(1 - z e^{-\beta \frac{\hbar^2 \|\mathbf{k}\|^2}{2m}}\Big) = \frac{4\pi V}{(2\pi)^3} \int_0^{+\infty} \log\Big(1 - z e^{-\beta \frac{\hbar^2 k^2}{2m}}\Big) k^2 \, \mathrm{d}k, \tag{2.187}$$

and

$$\sum_{\mathbf{k}} \frac{1}{z^{-1} e^{\beta \frac{\hbar^2 \|\mathbf{k}\|^2}{2m}} - 1} = \frac{4\pi V}{(2\pi)^3} \int_0^{+\infty} \frac{k^2}{z^{-1} e^{\beta \frac{\hbar^2 k^2}{2m}} - 1} \, \mathrm{d}k. \tag{2.188}$$

However, these expressions will not be as easily adapted for the condensed and coexisting phases. For $\epsilon = 0$ and $\mu \to 0$, the occupation density $\frac{1}{e^{\beta(\epsilon - \mu)} - 1}$ will blow up, and so will the ground's state contribution to the pressure in the form of $\log\big(1 - e^{-\beta(\epsilon - \mu)}\big)$. As a consequence, these quantities must be considered in the thermodynamic limit.

In the thermodynamic limit, the logarithmic divergence due to the ground state will vanish when compared to the linear divergence of the infinite volume. Hence, no problems will arise with the pressure.

As for the density, things are a bit more complicated. The thermodynamic limit is taken at constant density, and hence the occupation densities must remain finite. We can then write

$$\lim_{\mu \to 0} \frac{1}{V} \frac{1}{e^{-\beta\mu} - 1} = \frac{N_0}{V}, \tag{2.189}$$

where N_0 is the (diverging) occupation number of the ground state. The ratio $\frac{N_0}{V}$ will then remain finite. Removing a single point from the integral doesn't change its result, so

the integral on Eq. (2.188) on page 61 can still be used to compute the number of states with $k > 0$ (however, notice that in the coexistence region and in the condensed phase the fugacity z evaluates to 1, since $\mu = 0$).

An alternative approach to treating the ground state separately is given by Pathria and Beale (2022, App. F).

Under the light of these facts, we see that at the coexistence region we'll get the expression

$$\frac{N}{V} = \frac{N_0}{V} + \frac{4\pi}{(2\pi)^3} \int_0^{+\infty} \frac{k^2}{e^{\beta \frac{\hbar^2 k^2}{2m}} - 1} \, \mathrm{d}k \,. \tag{2.190}$$

Let us compute this integral. Firstly we notice that a change of variables $\xi = \frac{\beta \hbar^2 k^2}{2m}$ leads us to

$$\int_0^{+\infty} \frac{k^2}{e^{\beta \frac{\hbar^2 k^2}{2m}} - 1} \, \mathrm{d}k = \int_0^{+\infty} \frac{2m\xi}{\beta \hbar^2} \frac{1}{e^\xi - 1} \, \mathrm{d}\!\left(\sqrt{\frac{2m\xi}{\beta \hbar^2}} \right), \tag{2.191a}$$

$$= \frac{1}{2} \left(\frac{2m}{\beta \hbar^2} \right)^{\frac{3}{2}} \int_0^{+\infty} \frac{\xi^{\frac{1}{2}}}{e^\xi - 1} \, \mathrm{d}\xi \,. \tag{2.191b}$$

However, we should notice that, for $s > 1$,

$$\int_0^{+\infty} \frac{\xi^{s-1}}{e^\xi - 1} \, \mathrm{d}\xi = \int_0^{+\infty} \frac{\xi^{s-1} e^{-\xi}}{1 - e^{-\xi}} \, \mathrm{d}\xi \,, \tag{2.192a}$$

$$= \int_0^{+\infty} \xi^{s-1} \sum_{n=1}^{+\infty} e^{-n\xi} \, \mathrm{d}\xi \,, \tag{2.192b}$$

$$= \sum_{n=1}^{+\infty} \int_0^{+\infty} \xi^{s-1} e^{-n\xi} \, \mathrm{d}\xi \,, \tag{2.192c}$$

$$= \sum_{n=1}^{+\infty} \frac{1}{n^s} \int_0^{+\infty} \rho^{s-1} e^{-\rho} \, \mathrm{d}\rho \,, \tag{2.192d}$$

$$= \sum_{n=1}^{+\infty} \frac{\Gamma(s)}{n^s} \,, \tag{2.192e}$$

$$= \Gamma(s) \sum_{n=1}^{+\infty} \frac{1}{n^s} \,, \tag{2.192f}$$

$$= \Gamma(s)\zeta(s), \tag{2.192g}$$

where on Eq. (2.192d) on the following page we made the change of variables $\rho = n\xi$. In the previous expressions, $\Gamma(s)$ is Euler's gamma function (see Arfken, Weber, and Harris 2013, Chap. 13) and $\zeta(s)$ is Riemann's zeta function (Arfken, Weber, and Harris 2013, Sec. 13.5).

Hence, our integral of interest evaluates to

$$\int_0^{+\infty} \frac{k^2}{e^{\beta \frac{\hbar^2 k^2}{2m}} - 1} \, \mathrm{d}k = \frac{1}{2} \left(\frac{2m}{\beta \hbar^2} \right)^{\frac{3}{2}} \Gamma\!\left(\frac{3}{2} \right) \zeta\!\left(\frac{3}{2} \right). \tag{2.193}$$

Using the facts that $\Gamma(z+1) = z\Gamma(z)$ (Arfken, Weber, and Harris 2013, Eq. (13.2)) and that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ (Arfken, Weber, and Harris 2013, Eq. (13.8)), we find that $\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}$, and hence

$$\int_0^{+\infty} \frac{k^2}{e^{\beta\frac{\hbar^2 k^2}{2m}} - 1}\,\mathrm{d}k = \sqrt{\frac{\pi}{2}}\left(\frac{m}{\beta\hbar^2}\right)^{\frac{3}{2}}\zeta\left(\frac{3}{2}\right). \tag{2.194}$$

Therefore, we see that

$$\frac{N}{V} = \frac{N_0}{V} + \left(\frac{m}{2\pi\beta\hbar^2}\right)^{\frac{3}{2}}\zeta\left(\frac{3}{2}\right). \tag{2.195}$$

If we now impose $N_0 = 0$, we'll be considering the normal phase[12], and can then use this expression to compute the critical temperature T_0 at which the system starts to condense. Solving for the temperature on Eq. (2.195) at $N_0 = 0$ leads us to

$$T_0 = \frac{2\pi\hbar^2}{mk_B\zeta\left(\frac{3}{2}\right)^{\frac{2}{3}}}\left(\frac{N}{V}\right)^{\frac{2}{3}}, \tag{2.196}$$

which is the Bose–Einstein temperature for a gas of free bosons.

Notice that we can now write the density of excited states in the form

$$\frac{N_e}{V} = \left(\frac{m}{2\pi\beta\hbar^2}\right)^{\frac{3}{2}}\zeta\left(\frac{3}{2}\right), \tag{2.197a}$$

$$= T^{\frac{3}{2}}\left(\frac{mk_B}{2\pi\hbar^2}\right)^{\frac{3}{2}}\zeta\left(\frac{3}{2}\right), \tag{2.197b}$$

$$= \left(\frac{T}{T_0}\right)^{\frac{3}{2}}\frac{N}{V}. \tag{2.197c}$$

Therefore, since $N = N_0 + N_e$, we find that

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_0}\right)^{\frac{3}{2}}. \tag{2.198}$$

The behavior of N_0 and N_e with temperature is plotted on Fig. 2.12 on the following page.

Superfluidity of Helium-4

Helium-4 has a behavior that resembles that of a free boson gas at low temperatures. Namely, just as a free boson gas has a normal and a condensed phase, ^{4}He has a normal and a superfluid phase. The second-order phase transition happens at the so-called "λ point", with temperature $T_\lambda \approx 2.17\,\mathrm{K}$, pressure $p_\lambda \approx 1\,\mathrm{atm}$, and specific volume $v_\lambda \approx 46.2\,\text{Å}^3$ (data from Salinas 2001, p. 193).

---

[12]In the normal phase, the individual occupation densities of all states tend to zero in the thermodynamic limit, including the ground state. Hence, we can compute all of the occupation numbers by considering only the integral.
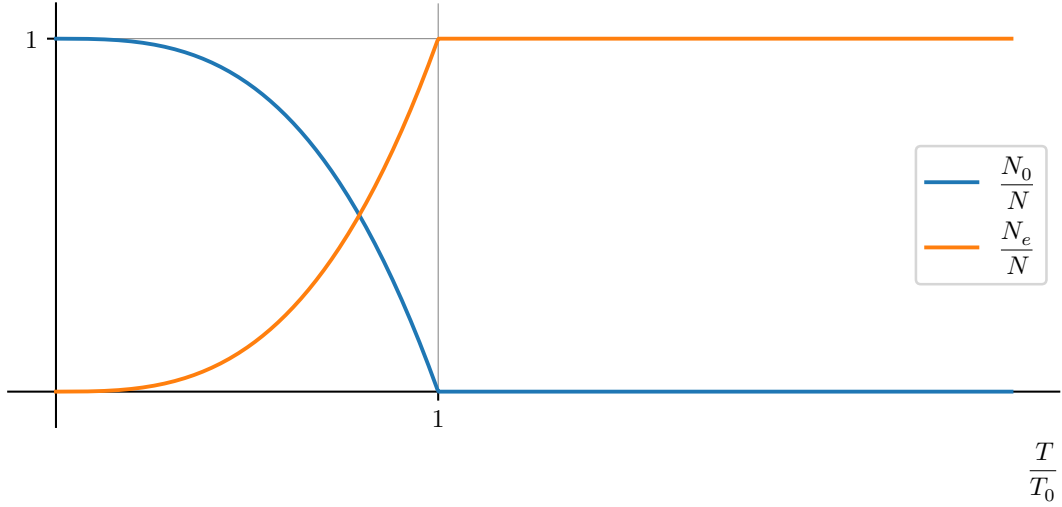
Figure 2.12: *Behavior of the occupation numbers of ground and excited states as functions of temperature. Notice how the states condense at the ground state below the Bose–Einstein temperature.*

We can compute the Bose–Einstein temperature of helium-4 by treating it as if it were a free gas and using the value of the specific volume at the λ point, together with the fact that the mass of a helium-4 atom is roughly $4m_p$, where m_p is the mass of a proton, having the value of $m_p \approx 1.67 \times 10^{-27}$ kg (Particle Data Group et al. 2022). Using these values, one gets to $T_0 \approx 3.14\,\text{K}$. This is quite close to the experimental value $T_\lambda \approx 2.17\,\text{K}$, hinting at the fact that the behaviour of helium-4 is related to its bosonic character. Still, we should recall that it is a bad approximation to treat a liquid as a gas of free particles.

For more on superfluids and helium-4, see the texts by Altland and Simons (2010, Sec. 6.3), Kardar (2007b, Sec. 7.7), Lancaster and S. Blundell (2014, Cap. 42), and Salinas (2001, Sec. 10.1).

Normal Phase

Let us next consider how the gas behaves in the normal phase, when $T > T_0$ and $\mu < 0$. In this case, the amount of particles condensed in the ground state is negligible, and hence we won't have to bother about treating the ground state separately (we'll do it either way, just to show its individual contribution vanishes). Hence, we can compute the grand canonical partition function in the following manner.

From Eq. (2.138) on page 50 we know that

$$\frac{1}{V}\log\Xi(\beta, V, z) = -\frac{1}{V}\log(1-z) - \frac{1}{V}\sum_{j\neq 0}\log\left(1 - ze^{-\beta\epsilon_j}\right). \qquad (2.199)$$

We can then approximate the sum with an integral by writing

$$\frac{1}{V}\log \Xi(\beta, V, z) = -\frac{1}{V}\log(1-z) - \frac{1}{V}\frac{V}{(2\pi)^3}\int \log\left(1 - ze^{-\beta\frac{\hbar^2\|\mathbf{k}\|^2}{2m}}\right)\mathrm{d}^3k, \qquad (2.200a)$$

$$= -\frac{1}{V}\log(1-z) - \frac{4\pi}{8\pi^3}\int_0^{+\infty}\log\left(1 - ze^{-\beta\frac{\hbar^2 k^2}{2m}}\right)k^2\,\mathrm{d}k. \qquad (2.200b)$$

In the thermodynamic limit, we'll take $V \to +\infty$, but z will remain bounded away from 1. Hence, we'll obtain

$$\frac{1}{V}\log \Xi(\beta, V, z) = -\frac{1}{2\pi^2}\int_0^{+\infty}\log\left(1 - ze^{-\beta\frac{\hbar^2 k^2}{2m}}\right)k^2\,\mathrm{d}k. \qquad (2.201)$$

If we now change integration variables to $\epsilon = \frac{\hbar^2 k^2}{2m}$, we get

$$\frac{1}{V}\log \Xi(\beta, V, z) = -\frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\int_0^{+\infty}\log(1 - ze^{-\beta\epsilon})\epsilon^{\frac{1}{2}}\,\mathrm{d}\epsilon. \qquad (2.202)$$

Since z won't be 1, but is bounded above by 1, we can use it to expand the integrand in a Taylor series. We find

$$\frac{1}{V}\log \Xi(\beta, V, z) = \frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\int_0^{+\infty}\epsilon^{\frac{1}{2}}\sum_{n=1}^{+\infty}\frac{z^n}{n}e^{-n\beta\epsilon}\,\mathrm{d}\epsilon, \qquad (2.203a)$$

$$= \frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\sum_{n=1}^{+\infty}\frac{z^n}{n}\int_0^{+\infty}\epsilon^{\frac{1}{2}}e^{-n\beta\epsilon}\,\mathrm{d}\epsilon, \qquad (2.203b)$$

$$= \frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\sum_{n=1}^{+\infty}\frac{z^n}{n}\frac{\sqrt{\pi}}{2}\frac{1}{(n\beta)^{\frac{3}{2}}}, \qquad (2.203c)$$

$$= \left(\frac{m}{2\pi\beta\hbar^2}\right)^{\frac{3}{2}}\sum_{n=1}^{+\infty}\frac{z^n}{n^{\frac{5}{2}}}, \qquad (2.203d)$$

$$= \frac{1}{\lambda^3}\sum_{n=1}^{+\infty}\frac{z^n}{n^{\frac{5}{2}}}, \qquad (2.203e)$$

where λ is the thermal wavelength.

If we define the Bose–Einstein function $g_\nu(z)$ through[13]

$$g_\nu(z) \equiv \sum_{n=1}^{+\infty}\frac{z^n}{n^\nu}, \qquad (2.204)$$

---

[13]Bose–Einstein functions are discussed in detail, for example, by Pathria and Beale (2022, App. D). These functions are "dual" to the ones that we mentioned on Eq. (2.163) on page 56. Notice that writing $g_\nu(z)$ for the Bose–Einstein functions is the choice made by Pathria and Beale (2022) and Salinas (2001), but Kardar (2007b) uses $f_\nu^\pm(z)$ for the Bose–Einstein and Fermi–Dirac functions, with the sign determining the choice (+ for bosons, − for fermions).

then we may write

$$\frac{1}{V}\log\Xi(\beta, V, z) = \frac{1}{\lambda^3}g_{\frac{5}{2}}(z). \tag{2.205}$$

It then follows from Eqs. (2.115) and (2.116) on page 43 and on page 44 that

$$N = z\left(\frac{\partial}{\partial z}\log\Xi\right)_\beta = \frac{V}{\lambda^3}g_{\frac{3}{2}}(z), \tag{2.206}$$

$$U = -\left(\frac{\partial}{\partial\beta}\log\Xi\right)_z = \frac{3V}{2\beta\lambda^3}g_{\frac{5}{2}}(z). \tag{2.207}$$

We can also obtain the pressure through

$$\frac{p}{k_BT} = \left(\frac{\partial}{\partial V}\log\Xi\right)_{T,\mu} = \frac{1}{\lambda^3}g_{\frac{5}{2}}(z). \tag{2.208}$$

If we want to obtain the more convenient form $U(T, V, N)$, we'll need to face the laborious task of inverting Eq. (2.206), which might need to be performed numerically.

Eqs. (2.206) and (2.208) should be compared to Eq. (2.163) on page 56 (notice that we're doing the bosonic computations with $\gamma = 1$, which explains why there is no weight factor in our currency results). Both results can actually be obtained at once, as done by Kardar (2007b, Sec. 7.4).

As an example, let us compute the specific heat at constant volume of a gas of free bosons at the normal phase. It is given by

$$c_v = \frac{1}{N}\left(\frac{\partial U}{\partial T}\right)_{V,N}, \tag{2.209a}$$

$$= -\frac{k_B\beta^2}{N}\left(\frac{\partial U}{\partial\beta}\right)_{V,N}. \tag{2.209b}$$

Unfortunately, we do not know $U(T, V, N)$. However, we can bypass this difficulty by employing Jacobian methods (Salinas 2001, App. A.5). We notice that[14]

$$\left(\frac{\partial U}{\partial\beta}\right)_N = \frac{\partial(U, N)}{\partial(\beta, N)}, \tag{2.210a}$$

$$= \frac{\partial(U, N)}{\partial(\beta, z)}\frac{\partial(\beta, z)}{\partial(\beta, N)}, \tag{2.210b}$$

$$= \left[\left(\frac{\partial U}{\partial\beta}\right)_z\left(\frac{\partial N}{\partial z}\right)_\beta - \left(\frac{\partial U}{\partial z}\right)_\beta\left(\frac{\partial N}{\partial\beta}\right)_z\right]\left[\frac{\partial(N, \beta)}{\partial(z, \beta)}\right]^{-1}, \tag{2.210c}$$

$$= \left[\left(\frac{\partial U}{\partial\beta}\right)_z\left(\frac{\partial N}{\partial z}\right)_\beta - \left(\frac{\partial U}{\partial z}\right)_\beta\left(\frac{\partial N}{\partial\beta}\right)_z\right]\left[\left(\frac{\partial N}{\partial z}\right)_\beta\right]^{-1}, \tag{2.210d}$$

$$= \left(\frac{\partial U}{\partial\beta}\right)_z - \left(\frac{\partial U}{\partial z}\right)_\beta\left(\frac{\partial N}{\partial\beta}\right)_z\left[\left(\frac{\partial N}{\partial z}\right)_\beta\right]^{-1}. \tag{2.210e}$$

---

[14]Since the volume is always constant in this computation, we'll omit it.

Eqs. (2.206) and (2.207) allow us to compute all of these expressions. We have

$$\left(\frac{\partial U}{\partial \beta}\right)_z = -\frac{15V}{4\beta^2\lambda^3}g_{\frac{5}{2}}(z), \tag{2.211}$$

$$\left(\frac{\partial U}{\partial z}\right)_\beta = \frac{3V}{2\beta\lambda^3 z}g_{\frac{3}{2}}(z), \tag{2.212}$$

$$\left(\frac{\partial N}{\partial \beta}\right)_z = -\frac{3V}{2\beta\lambda^3}g_{\frac{3}{2}}(z), \tag{2.213}$$

$$\left(\frac{\partial N}{\partial z}\right)_\beta = \frac{V}{\lambda^3 z}g_{\frac{1}{2}}(z). \tag{2.214}$$

We can now use these derivatives to compute the specific heat. It will be given by

$$c_v = -\frac{k_B\beta^2}{N}\left[-\frac{15V}{4\beta^2\lambda^3}g_{\frac{5}{2}}(z) - \left(\frac{3V}{2\beta\lambda^3 z}g_{\frac{3}{2}}(z)\right)\left(-\frac{3V}{2\beta\lambda^3}g_{\frac{3}{2}}(z)\right)\left(\frac{V}{\lambda^3 z}g_{\frac{1}{2}}(z)\right)^{-1}\right], \tag{2.215a}$$

$$= \frac{k_B\beta^2 V}{N}\left[\frac{15}{4\beta^2\lambda^3}g_{\frac{5}{2}}(z) - \left(\frac{3}{2\beta\lambda^3 z}g_{\frac{3}{2}}(z)\right)\left(\frac{3V}{2\beta\lambda^3}g_{\frac{3}{2}}(z)\right)\left(\frac{V}{\lambda^3 z}g_{\frac{1}{2}}(z)\right)^{-1}\right], \tag{2.215b}$$

$$= \frac{k_B\beta^2 V}{N}\left[\frac{15}{4\beta^2\lambda^3}g_{\frac{5}{2}}(z) - \frac{9g_{\frac{3}{2}}(z)^2}{4\beta^2\lambda^3 g_{\frac{1}{2}}(z)}\right], \tag{2.215c}$$

$$= \frac{3k_B}{2}\frac{V}{N}\frac{1}{\lambda^3}\left[\frac{5}{2}g_{\frac{5}{2}}(z) - \frac{3}{2}\frac{g_{\frac{3}{2}}(z)^2}{g_{\frac{1}{2}}(z)}\right], \tag{2.215d}$$

$$= \frac{3k_B}{2}\left[\frac{5}{2}\frac{g_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)} - \frac{3}{2}\frac{g_{\frac{3}{2}}(z)}{g_{\frac{1}{2}}(z)}\right]\frac{V}{N}\frac{g_{\frac{3}{2}}(z)}{\lambda^3}, \tag{2.215e}$$

$$= \frac{3k_B}{2}\left[\frac{5}{2}\frac{g_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)} - \frac{3}{2}\frac{g_{\frac{3}{2}}(z)}{g_{\frac{1}{2}}(z)}\right], \tag{2.215f}$$

where the last step employed Eq. (2.206) on the previous page. We then would need to eliminate the fugacity in favor of the number of particles which, as we previously said, is a laborious task.

Let us consider how the specific heat behaves at high and low temperatures. Close to the phase transition at the Bose–Einstein temperature, we have $z = 1$. Notice that for $\nu > 1$

$$g_\nu(1) = \sum_{n=1}^{+\infty}\frac{1}{n^\nu}, \tag{2.216a}$$

$$= \zeta(\nu), \tag{2.216b}$$

i.e., the Bose–Einstein functions reduce to Riemann's zeta function. Hence, we can use known values of $\zeta(\nu)$ to obtain $g_{\frac{3}{2}}(1) \approx 2.612\,38$ and $g_{\frac{5}{2}}(1) \approx 1.341\,49$. For $\nu \leq 1$, the series defining $g_\nu(1)$ diverges[15], and therefore $g_{\frac{1}{2}}(1) \to \infty$. Hence, close to the Bose–Einstein

---

[15]$\zeta(\nu)$ doesn't diverge for $\nu < 1$ because it is defined as an analytic continuation. $g_\nu(z)$, however, is not an analytic continuation, but the sum itself.

phase transition we have

$$\lim_{T \to T_0^+} c_v = \frac{15 k_B}{4} \frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)} < \infty. \tag{2.217}$$

For high temperatures, we have $z = e^{\beta\mu} \ll 1$, since $\mu < 0$. Hence, we can make the approximation

$$g_\nu(z) = \sum_{n=1}^{+\infty} \frac{z^n}{n^\nu}, \tag{2.218a}$$

$$\approx z + \mathcal{O}(z^2). \tag{2.218b}$$

Hence, for high temperatures, the specific heat on Eq. (2.215) on the preceding page will behave according to

$$\lim_{T \to +\infty} c_v = \frac{3 k_B}{2} \left[\frac{5}{2} \frac{z}{z} - \frac{3}{2} \frac{z}{z} \right] = \frac{3 k_B}{2}, \tag{2.219}$$

matching the equipartition theorem. Notice that $c_v(T_0) > c_v(+\infty)$. This will lead to a "cusp" in the graph for the specific heat as a function of temperature (see Kardar 2007b, Fig. 7.11; Pathria and Beale 2022, Fig. 7.4).

Coexistence Region

For $T \leq T_0$ (and $\mu < 0$) we know from Eqs. (2.206) and (2.207) on page 65 that the mean energy and number of particles in the excited states for the condensed region will be

$$U = \frac{3}{2} \frac{V}{\beta \lambda^3} \zeta\left(\frac{5}{2}\right), \tag{2.220}$$

and

$$N_e = \frac{V}{\lambda^3} \zeta\left(\frac{3}{2}\right), \tag{2.221}$$

where we used the fact that $\zeta(\nu) = g_\nu(1)$ for $\nu > 1$. Furthermore the number of states in the ground state will be given by

$$N_0 = N - N_e. \tag{2.222}$$

Using the expression for the internal energy, we can find that

$$c_v = \frac{1}{N} \left(\frac{\partial U}{\partial T} \right)_{V,N}, \tag{2.223a}$$

$$= \frac{15 V}{4 N} \left(\frac{k_B m}{2\pi \hbar^2} \right)^{\frac{3}{2}} T^{\frac{3}{2}}. \tag{2.223b}$$

This expression vanishes at $T = 0$, but matches our previous expression at $T = T_0$. Hence, the specific heat is continuous with temperature, even though it is not differentiable at the

Bose–Einstein temperature[16]. The plot of the specific heat as a function of temperature can be found, *e.g.*, in the books by Kardar (2007b, Fig. 7.11) and Pathria and Beale (2022, Fig. 7.4) and has a finite, but non-differentiable, peak at the Bose–Einstein temperature.

We could compute the pressure as well and find that it vanishes in the limit of low temperatures.

3   Phase Transitions

Knowing how to work with equilibrium Statistical Mechanics, we'll start to investigate more complicated systems displaying phenomena known as phase transitions. Notice that a phase transition does not imply nonequilibrium: in fact, we'll firstly consider phase transitions while employing equilibrium methods. Later, we shall move on to nonequilibrium phenomena.

3.1   Main Notions

Before we begin with specific studies of examples of phase transitions, we'll first discuss some general notions concerning them. This section is based partially on Prof. Fiore's lectures and slideshow, and partially on further literature. Notably the books by S. Blundell and K. M. Blundell (2010, Chap. 28) and Kardar (2007a, Sec. 1.3).

So far, we've been focusing mainly on non-interacting systems. Nevertheless, once interactions are turned on very interesting phenomena might occur. Among them, there is, for example, the fact that a single substance can have many different macroscopic properties depending on the particular state it is in—*e.g.*, depending on the pressure and temperature, water might be a liquid, a solid, a vapor, or even some more diverse phases.

Mathematically, all of these phases can be described in terms of a fundamental equation, *i.e.*, we can compute their properties starting from a free energy or from a partition function. Hence, since features of the system change drastically from a phase to another (the density, for example), these phase transitions correspond to singularities in the free energy, *i.e.*, to discontinuities or divergences in the derivatives of the free energy[1]. Notice that said derivatives are exactly the thermodynamic properties of the system, and hence discontinuities in them means exactly drastic changes between different phases. Furthermore, note also that these phenomena can happen in many systems, not only fluids: magnetic systems, metallic alloys, liquid crystals, and many other examples also feature such properties.

From a technical point of view, it is important to notice that phase transitions only occur at the thermodynamic limit, in the sense that we can't have singularities on the free energy for finite N, since in this case the partition function will always be analytic. Hence, we'll study the systems in the thermodynamic limit with the goal of finding and

---

[16]In the opinion of Prof. Fiore, the continuity of specific heat is what makes Bose–Einstein condensation resemble a first-order phase transition, instead of a second-order phase transition. For a second-order transition, the specific heat would typically diverge.

[1]Of the Gibbs' free energy, to be more precise, as we shall see later on.

understanding the origin of discontinuities in the derivatives of the free energy, hence obtaining information about what happens in real physical systems.

At this point, it is natural to ask what is meant by a phase. "A homogenous system, that is, completely uniform with regard to specific properties, constitutes a thermodynamic phase" (M. J. de Oliveira 2013, p. 104). Notice this isn't always the case. When boiling water on a stove, the system (water) is heterogeneous, presenting two coexistent phases.

As for a phase transition, we can understand it in many systems[2] as an analytical singularity in the system's Gibbs free energy. The specific choice of the Gibbs free energy is due to the fact that the Gibbs free energy per particle is a function of intensive quantities only, and we need to have a singularity as a function of the intensive quantities, since they have to match even for two phases in coexistence. On the other hand, the volume, for example, can be different for two phases in coexistence.

We may then classify different types of phase transitions. Ehrenfest gave such a classification according to the rule that a phase transition is said to be of nth order if the nth derivative of the Gibbs free energy presents a discontinuity. Nevertheless, there are important examples of phase transitions that fall outside of this scheme, such as the phase transitions that occur on the two-dimensional Ising model and on liquid helium. The Ising model, for example, has a derivative becoming infinite rather than discontinuous. Hence, eventually the Ehrenfest model became insufficient (for more historical details, see Jaeger 1998).

A more modern classification scheme is to classify as first-order phase transitions (or discontinuous phase transitions) those that have a latent heat (see p. Section 3.1 on page 72). Second-order phase transitions (or continuous phase transitions) are then the remaining ones[3] (S. Blundell and K. M. Blundell 2010, Sec. 28.7). Liquid-gas phase transitions are often examples of first-order phase transitions[4], while a piece of iron losing its ferromagnetic properties at high temperatures is an example of a second-order phase transition.

Notice that discontinuities in the derivatives of the free energy are a quite physical effect. For a simple fluid, the first derivatives of the Gibbs free energy are the entropy, the volume, and the chemical potential. Hence, an example of a first-order phase transition is a sudden "jump" in the density of a fluid.

First-Order Phase Transitions for Fluids

In order to have some concreteness, let us discuss first-order phase transitions in fluid systems. In this case, let us recall that when pressure and temperature are held fixed, the Second Law of Thermodynamics implies the Gibbs free energy is minimized (Fermi 1956,

---

[2]I am not fully sure if there is a completely general definition of phase transition, but singularities in the free energy do encompass a wide class due to our previous arguments.

[3]The classification given by Prof. Fiore was that first-order phase transitions are those with discontinuities on a first derivative of the free energy, while second-order transitions are those with a divergence on a second derivarive of the free energy.

[4]The use of the word "often" is because at the so-called critical point the transition does not involve a latent heat, and hence it becomes a continuous transition.

Sec. 18). Therefore, given a pressure and temperature, the system will be in the state with a minimum of Gibbs free energy.

In the situation in which the Gibbs free energy might have more than one minimum we will have coexistence of phases, though. If we then follow a coexistence curve (such as the ones shown on Fig. 3.1 on the next page), we'll often see the two minima of the Gibbs free energy coming closer together until they merge. This final point where the two minima merge is known as a critical point. Notice that when a coexistence line terminates, we can go from a phase to another by going around the critical point, without ever having a phase transition, since we haven't found any singularities of the free energy along the way. Hence, this illustrates that, fundamentally, there is no difference between the liquid and gas phases. Furthermore, at the critical point, the transition actually becomes continuous, and hence it is now a second-order phase transition.
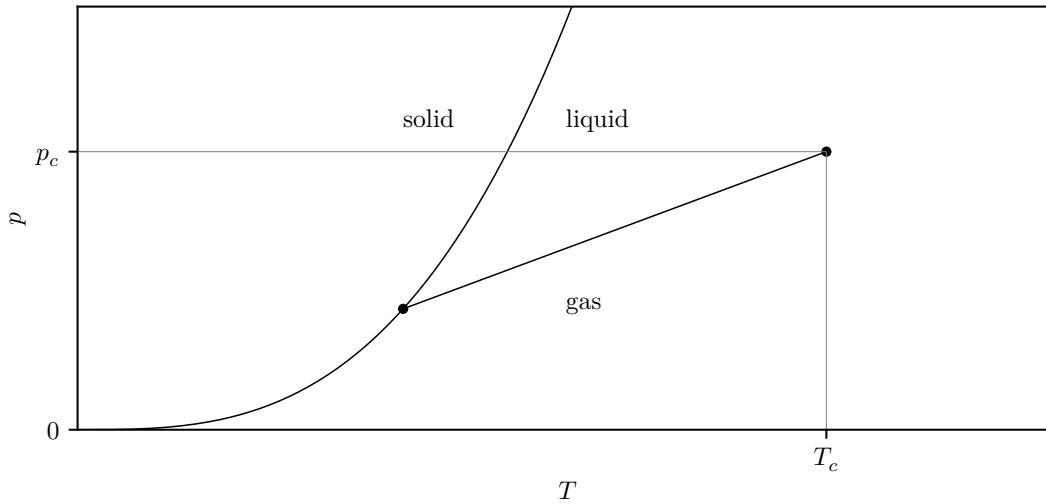
Figure 3.1: *Typical form of the pressure-temperature phase diagram for a fluid. The critical points and the triple point are highlighted. This graph was based on the figure by Kardar (2007a, Fig. 1.3).*

It should be noted that not all coexistence lines end in a critical point. S. Blundell and K. M. Blundell (2010, p. 337) point out that the liquid-gas transition doesn't involve any symmetry breaking, which is why it is possible to go around a fixed point. A liquid-solid transition, however, does involve symmetry breaking: while the liquid has no preferred directions, a solid is often more organized. As a consequence, there is a fundamental difference between both phases and it is not possible to "cheat" by going around a critical point.

It is also interesting to wonder what happens when we discuss the thermodynamics of phase transitions in terms of pressure and volume. A typical *p-v* diagram is shown on Fig. 3.2 on page 72. We see that, if we consider volume as a function of pressure (*i.e.*, if we recall that $V = \left(\frac{\partial G}{\partial p}\right)_{T,N}$, we notice that there are discontinuities in volume below the critical temperature T_c. The transition can be seen not as lines on the phase diagram,

but as a coexistence region involving mixtures of liquid and gas. Of course, we can still "bypass" the phase transition by going around the critical point, as shown on Fig. 3.3 on page 73.
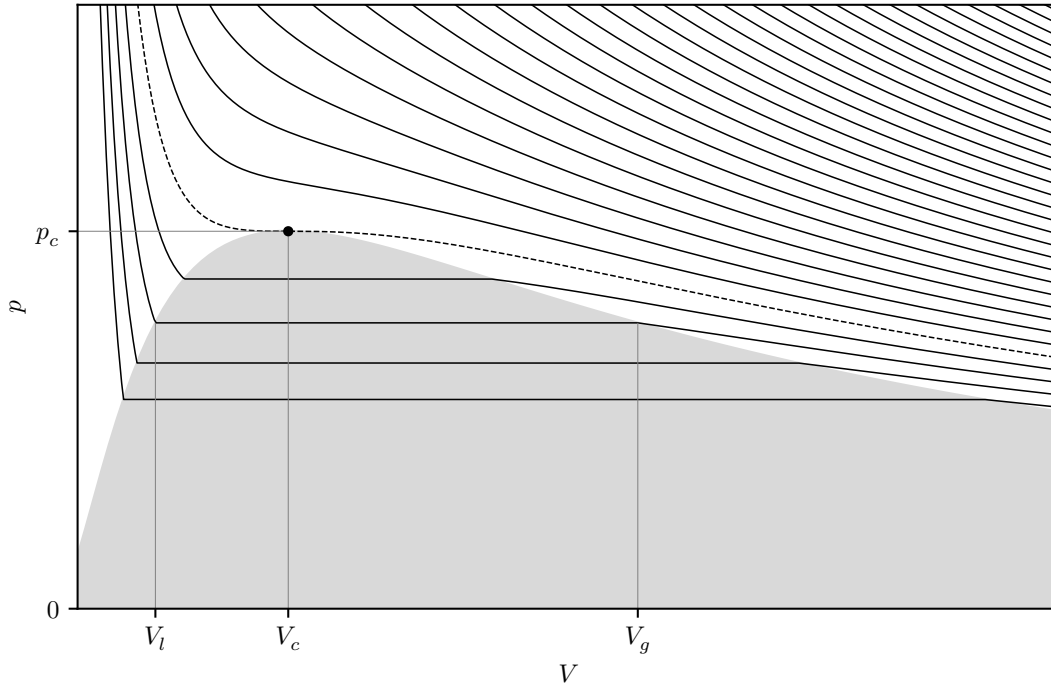

Figure 3.2: *Typical form of the isotherms for a fluid. The grey region is the coexistence region between liquid and gas phases. The dashed line corresponds to the critical temperature, while the highlighted point is the critical point. Notice the isotherms are flat on the coexistence region. V_l stands for the "liquid volume", V_g for the "gaseous volume". This graph was based on the code by christian (2016) and on the figure by Kardar (2007a, Fig. 1.3).*

Figs. 3.2 and 3.3 on the following page and on page 73 are built upon the equations for a Van der Waals gas, which doesn't have a solid phase. Nevertheless, for other fluids there often is a solid phase, and even a point where the three phases coexist, known as a triple point and shown on Fig. 3.1.

Notice that Fig. 3.2 on the next page allows us to understand some of the fluid's behavior as $T \to T_c$. For high temperatures, the isotherms tend to become flatter as the temperature is lowered, and hence we see that the isothermal compressibility κ_T will diverge as $T \to T_c^+$. For low temperatures, we see that the coexistence region at a fixed temperature gets smaller as $T \to T_c^-$. The differences in density between a gas and a liquid will vanish at the critical temperature.

An experimental observation that is also worth mentioning is that, close to criticality, the fluid will acquire a "milky" appearance. This phenomenon, known as critical opalescence, is due to the fact that there are large density fluctuations close to the critical point,
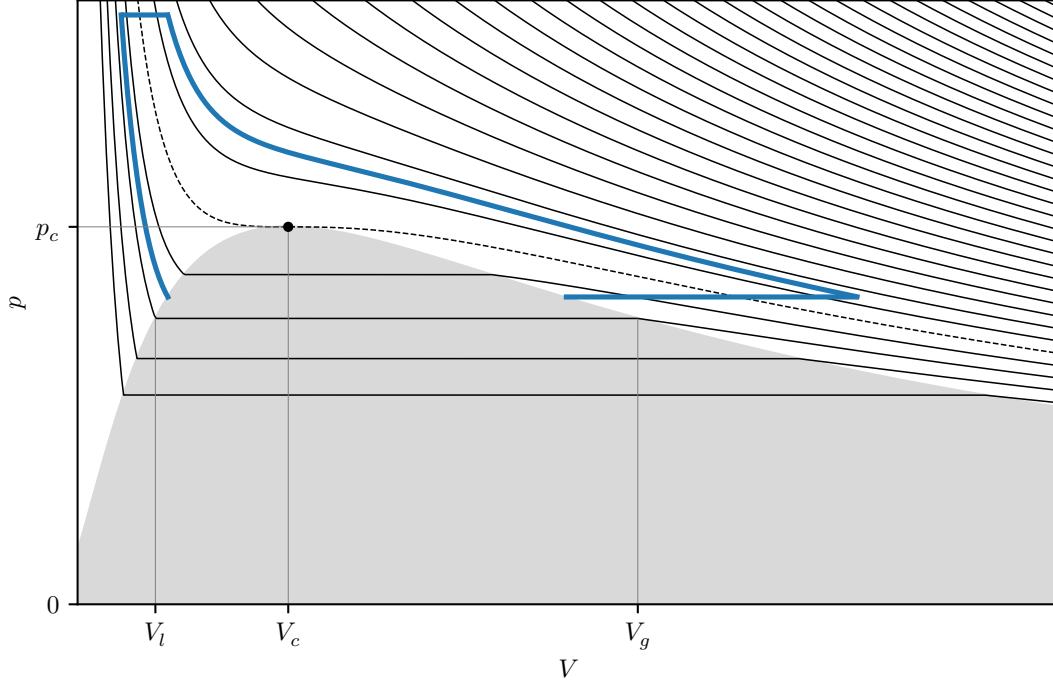
Figure 3.3: *Example of how to "bypass" a phase transition in a p-v diagram by using isobaric and isothermal transformations.*

which also leads to large fluctuations in the refractive index. Another example of large variations of density is the bubbling of boiling water on a saucepan. These effects mean that our usual approach of assuming thermodynamic quantities are well-defined breaks down close to criticality, and we might need to resort to other techniques (S. Blundell and K. M. Blundell 2010, p. 336; Kardar 2007a, p. 10).

While we are focusing a lot on how volume signals phase transitions, it is worth recalling that entropy is also a derivative of the Gibbs free energy, for $S = -\left(\frac{\partial G}{\partial T}\right)_{p,N}$. When two phases have different entropies, we need to supply extra heat to change one into the other. This heat, known as latent heat, will then be given by

$$L = T\Delta S = T(S_2 - S_1). \tag{3.1}$$

Hence, we'll get a spike on the heat capacity to account for this discontinuity on entropy.

At the phase transition, both phases (say, liquid and gas) have the same value for the Gibbs free energy. Hence, we can write

$$g_g(T, p) = g_l(T, p), \tag{3.2a}$$

$$-s_g\, \mathrm{d}T + v_g\, \mathrm{d}p = -s_l\, \mathrm{d}T + v_l\, \mathrm{d}p\,, \tag{3.2b}$$

from which we get

$$\frac{\mathrm{d}p}{\mathrm{d}T} = \frac{s_g - s_l}{v_g - v_l}, \tag{3.3a}$$

$$= \frac{l}{T(v_g - v_l)}, \tag{3.3b}$$

where $l = \frac{L}{N}$ is the latent heat per particle (or per mole).

Eq. (3.3) on page 73 is known as the Clausius–Clapeyron equation, and it is particularly useful to understand and construct p-T diagrams. For example, the coexistence line between water's solid and liquid phases has a negative slope, as illustrated on the sketch of Fig. 3.4 on the next page. Why is that so?
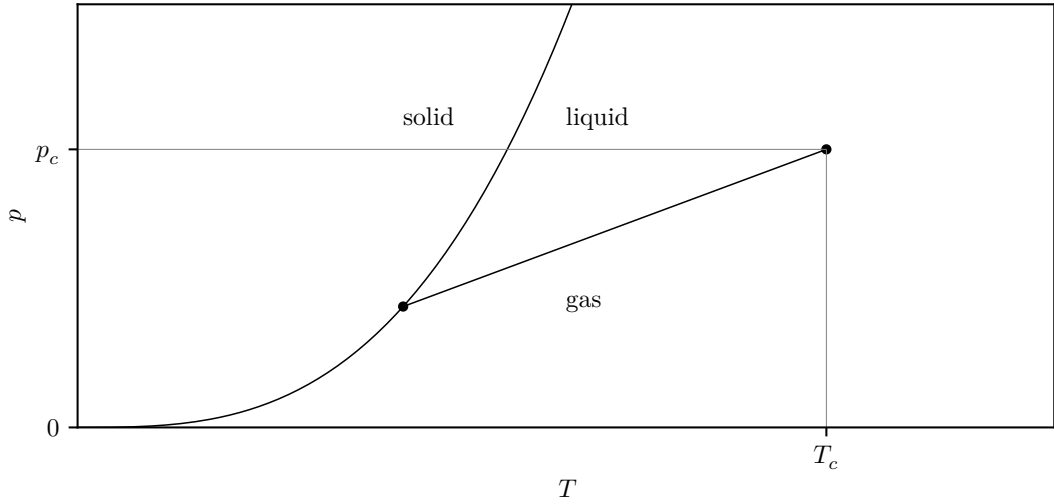
Figure 3.4: *Sketch of the pressure-temperature phase diagram for water. The critical points and the triple point are highlighted. Notice that the coexistence line between the solid and liquid phases has a negative slope. This is just a sketch: real water actually has many different ice phases and it is believed it might even have a second critical point.*

This can be understood with the Clausius–Clapeyron equation. For the liquid-vapour phase transition, we have $l_v > 0$ and $v_v > v_l$. Hence,

$$\frac{\mathrm{d}p}{\mathrm{d}T} = \frac{l_v}{T(v_v - v_l)} > 0. \tag{3.4}$$

For the solid-liquid transition, we also have $l_l > 0$, but $v_s > v_l$, for ice floats on water. Hence,

$$\frac{\mathrm{d}p}{\mathrm{d}T} = \frac{l_l}{T(v_l - v_s)} < 0, \tag{3.5}$$

which explains the diagram.

Thermodynamic Instabilities and First-Order Phase Transitions

As we mentioned at the end of Section 2.4, the Van der Waals gas presents instabilities at low temperatures. As we can see on Fig. 2.6 on page 40, there are isotherms in which at which the isothermal compressibility $\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{T,N}$ becomes negative. This is a problem, because it means that the more we compress the gas, the more it expands, making it mechanically unstable. In many situations, and in this in particular, that problematic behavior is actually signaling that the gas undergoes a phase transition and becomes a liquid. The issue actually arises from our mean field approximation when deriving the Van der Waals equation, and now we must fix the expression to recover stability.

To solve this issue, let us look at what is happening with the Helmholtz free energy. Notice that the Helmholtz free energy will be given by (*cf.* Eq. (2.81) on page 34)

$$\left(\frac{\partial f}{\partial v}\right)_T = -p, \tag{3.6a}$$

$$= -\frac{k_B T}{v - b} + \frac{a}{v^2}, \tag{3.6b}$$

$$f(T, v) = -k_B T \log(v - b) - \frac{a}{v} + f_0(T), \tag{3.6c}$$

where $f_0(T)$ is some arbitrary function depending on temperature, but not on the volume. Let us plot the pressure and Helmholtz free energy for some convenient temperature[5]. These plots are shown on Fig. 3.5 on the following page.

Fig. 3.5b makes it clear that the instability region corresponds to a region in which the Helmholtz free energy fails to be a convex function of the volume. This means the Gibbs free energy is ill-defined, since we need the Helmholtz free energy to be a convex function to perform a Legendre transformation. The so-called Maxwell construction consists in replacing the concave region of the Helmholtz free energy with a straight line, which is doubly tangent to the free energy's graph. In this way, we can obtain an expression for the Helmholtz free energy that is differentiable and convex, which fix our problems, at least from a theoretical perspective[6].

Suppose the points highlighted on Fig. 3.5b on the preceding page correspond to the points (v_A, f_A) and (v_B, f_B) ($v_B > v_A$, for concreteness). Then the slope of the dashed line is given by

$$-p^* = \frac{f_B - f_A}{v_B - v_A}, \tag{3.7}$$

where the sign was chosen so that we can understand this quantity as a pressure. We now notice that, using of the Van der Waals equation of state, we may write

$$p^*(v_B - v_A) = f_A - f_B, \tag{3.8a}$$

---

[5]I picked $T = 0.84575 T_c$, where T_c, known as the critical temperature, is given by $T_c = \frac{8a}{27bk_B}$. This choice is simply because this isotherm displays quite clearly the problematic features we're interested in.

[6]For a different point of view on justifying the Maxwell construction, see the text by Kardar (2007b, Sec. 5.4).
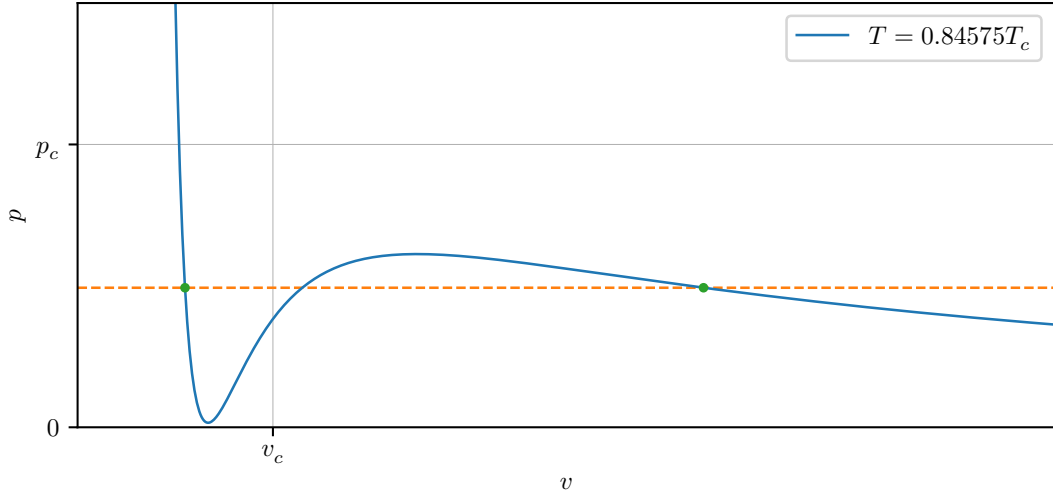
(a) *Plot of the pressure against specific volume for the chosen isotherm. Notice there is an unstable region with $\left(\frac{\partial p}{\partial v}\right)_T > 0$. The dashed line corresponds to the dashed line shown in the Helmholtz free energy plot.*
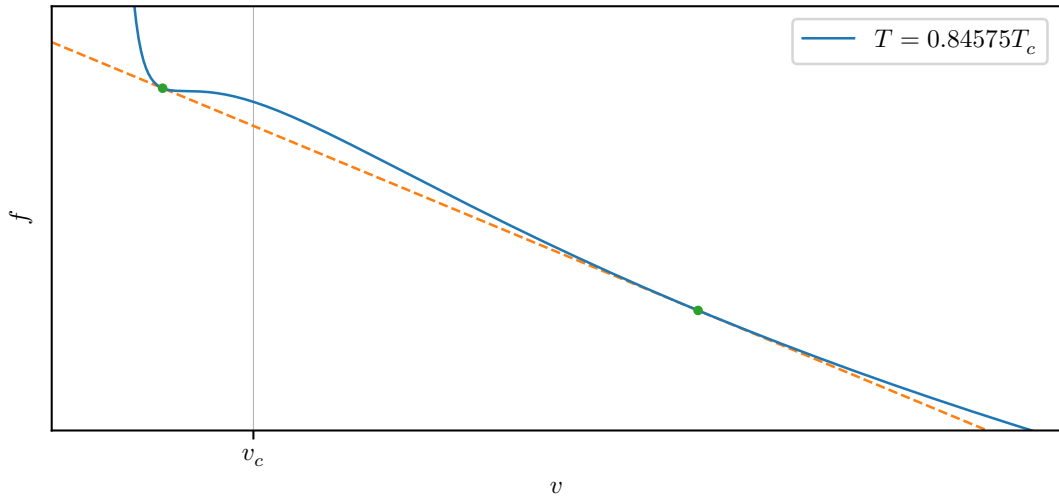
(b) *Plot of the Helmholtz free energy against specific volume for the chosen isotherm. Notice there is an unstable region where the function fails to be convex. To fix it, we might replace that piece of the function by the dashed line.*

Figure 3.5: *Pressure and Helmholtz free energy of the Van der Waals gas for the isotherm $T = 0.84575T_c$, where $T_c = \frac{8a}{27bk_B}$ is known as the critical temperature.*

$$= \int_{v_A}^{v_B} p(v)\, \mathrm{d}v, \tag{3.8b}$$

which can be rewritten as

$$\int_{v_A}^{v_B} p(v) - p^*\, \mathrm{d}v. \tag{3.9}$$

Notice this equation corresponds to imposing that the areas in between the isotherm and the dashed line on Fig. 3.5a on the previous page are the same. This fact can be exploited to actually perform the Maxwell construction.

Before we actually discuss how to perform the construction, it is useful to find the Van der Waals' gas critical point, so we can express the quantities as ratios between the physical values and the critical values. The critical point must satisfy three criteria:

- it solves the Van der Waals equation of state,

$$p_c = \frac{k_B T_c}{v_c - b} - \frac{a}{v_c^2}; \tag{3.10}$$

- it satisfies

$$\left(\frac{\partial p}{\partial v}\right)_T = -\frac{k_B T_c}{(v_c - b)^2} + \frac{2a}{v_c^3} = 0, \tag{3.11}$$

since it is the limit of corrected isotherms which are simply plateaus;

- it satisfies

$$\left(\frac{\partial^2 p}{\partial v^2}\right)_T = \frac{2k_B T_c}{(v_c - b)^3} - \frac{6a}{v_c^4} = 0, \tag{3.12}$$

due to the requirement of thermodynamic stability (see Kardar 2007b, Sec. 1.9).

These are three equations for three unknowns, T_c, p_c, and v_c. Solving them is an algebraic exercise, and at the end of it one gets to

$$p_c = \frac{a}{27b^2}, \quad T_c = \frac{8a}{27bk_B}, \quad \text{and} \quad v_c = 3b. \tag{3.13}$$

If we now define

$$\tilde{p} \equiv \frac{p}{p_c}, \quad \tilde{T} \equiv \frac{T}{T_c}, \quad \text{and} \quad \tilde{v} \equiv \frac{v}{v_c}, \tag{3.14}$$

one can show that the Van der Waals equation reduces to

$$\tilde{p} = \frac{8\tilde{T}}{3\tilde{v} - 1} - \frac{3}{\tilde{v}^2}. \tag{3.15}$$

The equal area imposition is now written as

$$\int_{\tilde{v}_l}^{\tilde{v}_g} \tilde{p}(\tilde{v})\, \mathrm{d}v = \tilde{p}_l(\tilde{v}_g - \tilde{v}_l), \tag{3.16}$$

where $\tilde{v}_l$ is the liquid phase volume, $\tilde{v}_g$ is the gas phase volume, and $\tilde{p}_l = \tilde{p}_g$ is the pressure at both the liquid and gas phases (since it is constant during the phase transition). If we compute the integral and impose the Van der Waals equation at both the liquid and gas phases (in addition to $\tilde{p}_l = \tilde{p}_g$), we get to the system of equations

$$\begin{cases} \log(3\tilde{v}_g - 1) + \dfrac{9}{4\tilde{T}\tilde{v}_g} - \dfrac{3\tilde{v}_g}{3\tilde{v}_g - 1} = \log(3\tilde{v}_l - 1) + \dfrac{9}{4\tilde{T}\tilde{v}_l} - \dfrac{3\tilde{v}_l}{3\tilde{v}_l - 1}, \\[2mm] \tilde{p}_l = \dfrac{8\tilde{T}}{3\tilde{v}_l - 1} - \dfrac{3}{\tilde{v}_l^2}, \\[2mm] \tilde{p}_l = \dfrac{8\tilde{T}}{3\tilde{v}_g - 1} - \dfrac{3}{\tilde{v}_g^2}. \end{cases} \tag{3.17}$$

Given $\tilde{T}$, one can then—at least in principle—solve for $\tilde{p}_l$, $\tilde{v}_l$, $\tilde{v}_g$. In practice, numerical methods might be needed. In the figures throughout this text, I have been using a Python code based on the one by christian (2016).

We should point out clearly what is the idea behind the Maxwell construction: the Van der Waals gas was derived as an approximate equation of state for a gas of particles that behave as hard spheres with a small attractive potential. Being an approximation, the equation does not need to work at every situation. The Maxwell construction is a manner of obtaining a better approximation for the isotherms of a real fluid in a situation in which the Van der Waals equation is no longer working. In other words, it is a manner of using the failure of the Van der Waals equation to obtain a better description. Furthermore, while this construction can also be used for other equations of state, it will usually not be available on nonequilibrium thermodynamics.

It is also worth noticing that, as one can show, for temperatures below $\tilde{T} = 0.843\,75$ the Van der Waals equation will predict negative pressures.

Second-Order Phase Transitions

To get a different grasp of how convexity has to do with the breakdown of the Van der Waals equation and to build a bridge with second-order phase transitions, it is interesting for us consider how one would usually compute the Gibbs free energy[7]. By definition, the Gibbs free energy is the Legendre transformation of the Helmholtz free energy with respect to volume and pressure, meaning it is given by

$$g(T, p) = \inf_v \{f(T, v) + pv\}. \tag{3.18}$$

We're focusing here at only the intensive quantities, but we could also consider their extensive versions and just hold N fixed.

For fixed values of T and p, how does the functions $f(T, v) + pv$ look like as a function of v? The answer to this question will let us know what is the actual physical volume

---

[7]This section is inspired mostly by Prof. Fiore's lecture and slideshow, but also draws a bit from the book by Callen (1985, Chap. 10). More information on convex functions and Legendre transformations in the context of Thermodynamics can be found on the book by Wreszinski (2018).

of the system at fixed temperature and pressure, for it will be that which minimizes $f(T, v) + pv$. The values of T and p that should be the most interesting are those on the coexistence region, depicted as a solid line on the diagram of Fig. 3.6 on page 79. We'd also be interested in following the coexistence region after it no longer exists, so we extrapolated the curve on Fig. 3.6 on the next page with a dashed line.

Here's the algorithm for drawing Fig. 3.6 on the following page: for each fixed value of T you want to plot, use the Maxwell construction to find the coexistence pressure. Draw it on the graph. After you have drawn the solid line, pick some points on it (I used fifteen) and use them to fit a convenient model (I used $c_0 + c_1 \exp(c_2 T)$). With this model, you can plot the dashed line[8].
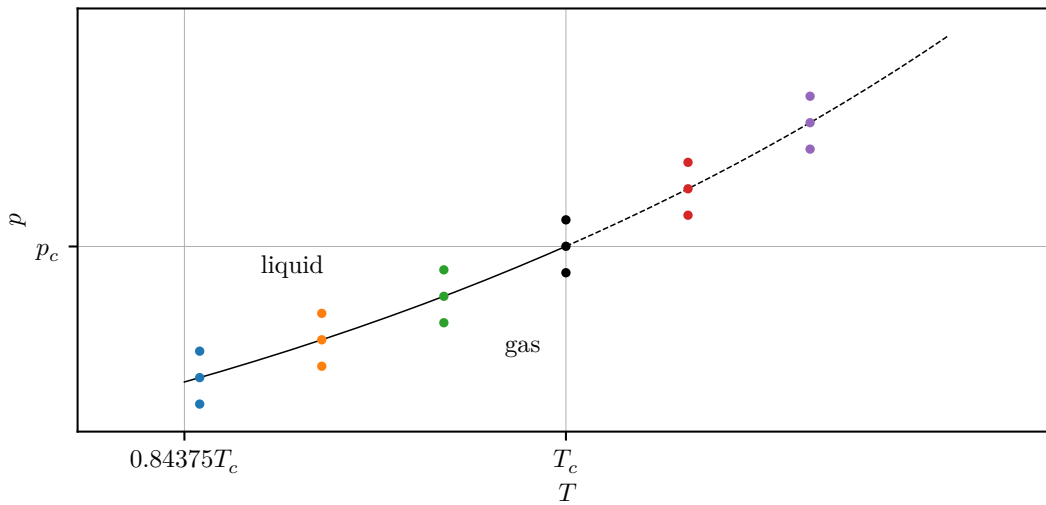
Figure 3.6: *Pressure and temperature phase diagram for the Van der Waals gas. For $T \leq 0.84375T_c$ the equation of state predicts negative pressures, so it was disregarded in that region. The solid line corresponds to the coexistence region, and the black point in the middle is the critical point. The dashed line is an extrapolation of the coexistence line made by fitting the curve $c_0 + c_1 \exp(c_2 T)$ to points from the coexistence line, so we can later check the behavior of the gas as we "follow the coexistence line" in the region without phase transitions. We also highlight some points of interest that we'll be analyzing in more detail.*

From Eq. (3.6) on page 74 we already know the expression for the Helmholtz free energy of the Van der Waals gas without the Maxwell correction. Hence, we know that

$$f(T, v) + pv = -k_B T \log(v - b) - \frac{a}{v} + f_0(T) + pV. \tag{3.19}$$

We do not know yet the function $f_0(T)$, but we'll just ignore it. We are interested in the behavior as a function of volume, so we can live with isotherm plots that could need to

---

[8]To be fair, I did use the fitted model to plot the solid line as well, since plotting a curve is faster than calculating all of the points using numerical root-finding algorithms. Still, the principle is the same.

be shifted up and own due to an extra constant. The resulting plots of Eq. (3.19) for the points we highlighted on Fig. 3.6 on the next page are shown on Figs. 3.7 to 3.9 on pages 79–80
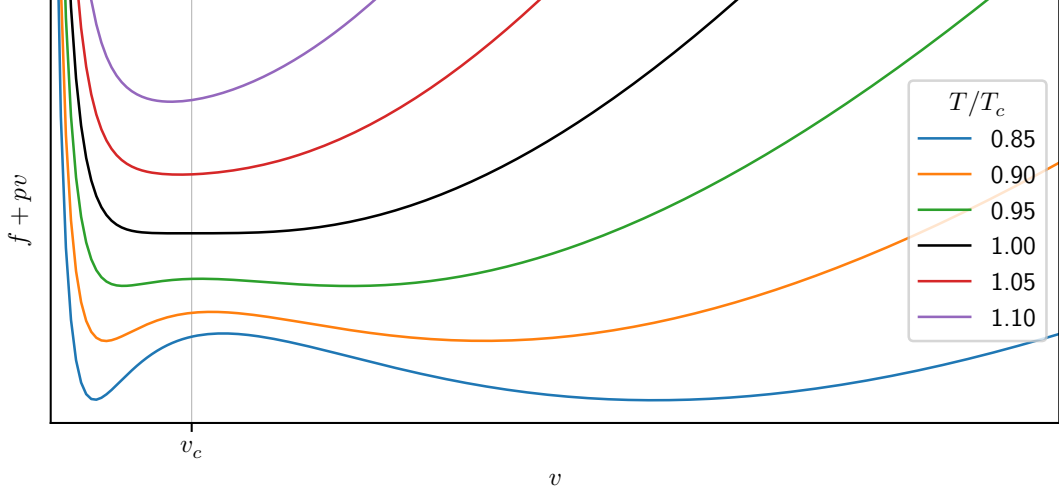

Figure 3.7: *Plots of the function $f(T, v) + pv$ as a function of v for the points highlighted on Fig. 3.6 on page 79 that lie on the coexistence line.*
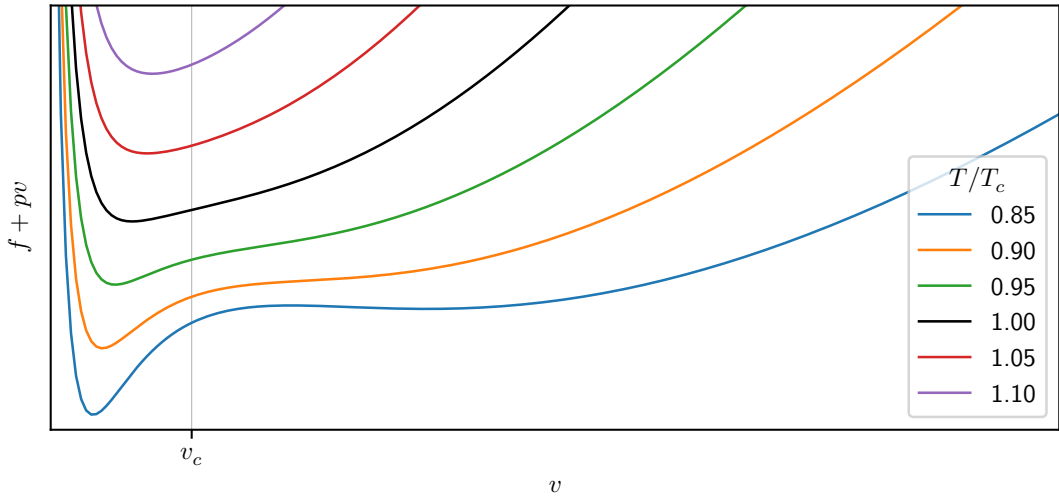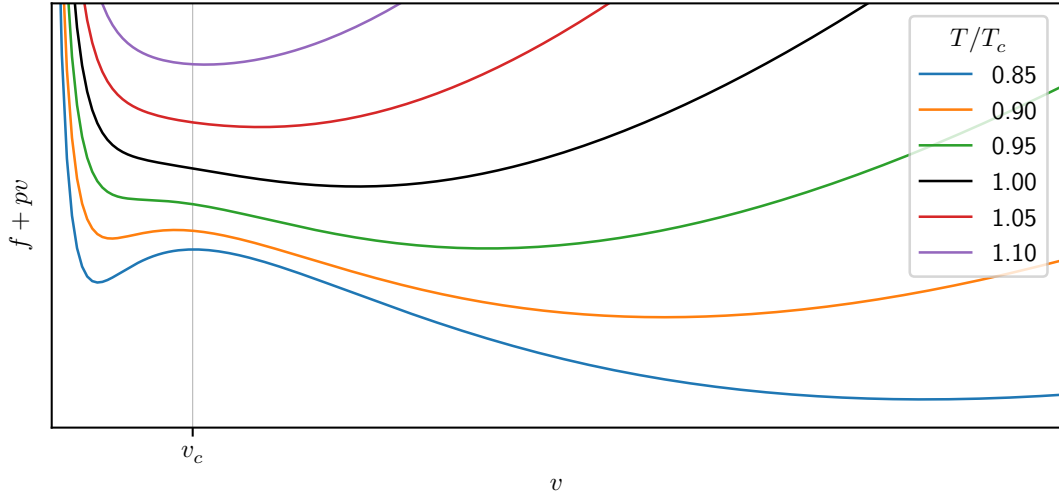

Figure 3.8: *Plots of the function $f(T, v) + pv$ as a function of v for the points highlighted on Fig. 3.6 on page 79 that lie above the coexistence line.*

The first thing we notice on Figs. 3.7 to 3.9 on pages 79–80 is the existence of two minima on the curves with $T < T_c$. Under constant temperature and pressure, the Gibbs free energy is minimized. Hence, what we see when both minima are the same (the case

Figure 3.9: *Plots of the function $f(T, v) + pv$ as a function of v for the points highlighted on Fig. 3.6 on page 79 that lie below the coexistence line.*

of Fig. 3.7 on the preceding page) is exactly that there are two possible phases. Hence, on the coexistence line, we see the existence of two phases. Below and above the coexistence line, we see that one of the minima becomes a global minimum, and hence there is a single phase.

Furthermore, notice that if we keep the temperature constant and slowly increase the pressure with the goal of crossing the coexistence line, we'll see a discontinuous phase transition. The global minimum discontinuously changes from a value to another.

However, notice now what happens for temperatures $T \geq T_c$. In this case, the transition between different values of the global minimum is continuous. For $T > T_c$, this is because there is always a single minimum. For $T = T_c$, we have a limiting, or critical, behavior in which the transition starts to be continuous. Hence, in the critical point we no longer have a first-order, discontinuous phase transition, but rather a continuous transition. The two minima present at $T < T_c$ merge into one exactly at the critical temperature.

Another way of seeing this is by looking at the isotherms as expressions for the volume as a function of pressure. These plots are shown for the Van der Waals gas on Fig. 3.10 on the next page. At the critical temperature, the curve for the volume becomes continuous, but it has an infinite derivative at the critical point.

In summary, for first-order phase transitions we'll find discontinuities on the first derivatives of the Gibbs free energy. This means we find discontinuities on quantities such as volume, entropy, of the magnetization of a ferromagnet.

For second-order phase transitions, we'll find divergences on the second derivatives of the Gibbs free energy. These are, for example, the specific heat, or the isothermal compressibility.
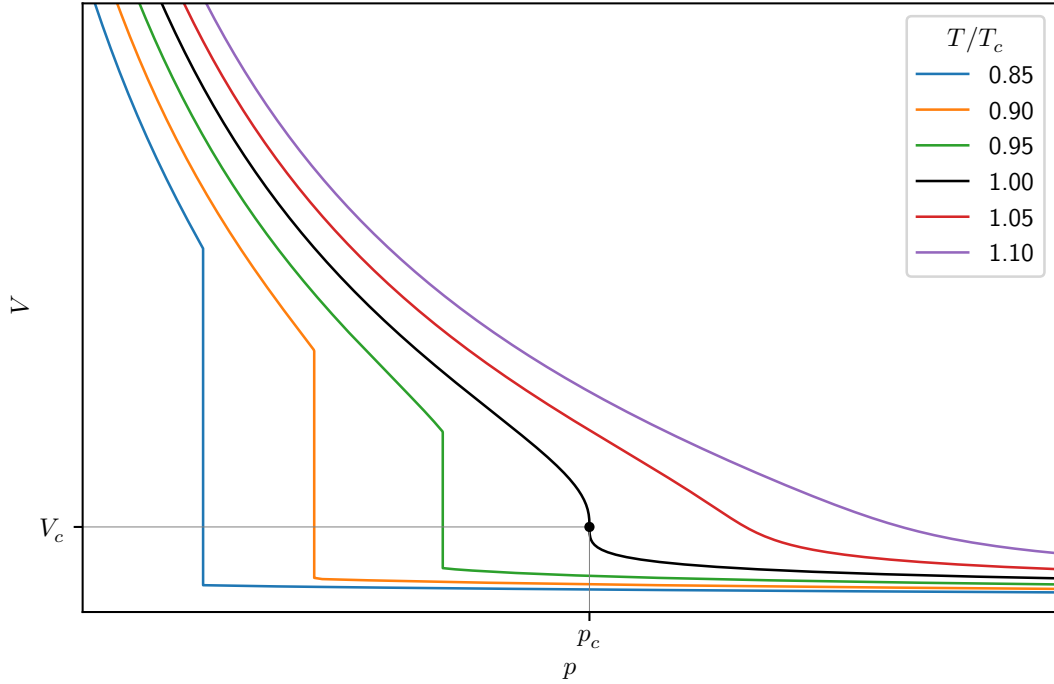
Figure 3.10: *Volume as a function of pressure for a few of the Van der Waals isotherms. At the critical temperature, the curve becomes continuous, but it has an infinite derivative at the critical point.*

Order Parameters

An alternative way of characterizing a phase transition is in terms of a so-called order parameter. This is a parameter ψ that assumes different values on each phase, allowing us to distinguish them. Often we may want to choose an order parameter in such a way that is vanishes in one of the phases, but not on the other. Furthermore, on some cases, we might need to use a vector or a tensor as an order parameter.

For the liquid-gas transition, possible choices of order parameters are $v_l - v_c$ (the liquid volume minus the critical volume), $v_g - v_c$ (the gas volume minus the critical volume), and $v_g - v_l$. Alternatively, we could also work with densities $\rho = \frac{1}{v}$ instead of the specific volumes. For a magnetic system, as we shall see, the spontaneous magnetization plays the role of an order parameter.

It is then possible to discuss phase transitions in terms of their order parameters. If the parameter vanishes (dis)continuously, the transition is (dis)continuous. To work with order parameter is always interesting in out-of-equilibrium problems, when we might not be able to find a thermodynamic potential, but can still find order parameters.

Critical Exponents

Since we characterized phase transitions in terms of discontinuities and divergences of the derivatives of the Gibbs free energy, one might wonder on whether these singularities occur on all derivatives, or only on some of them. For example, is it possible to have a system with discontinuous volume, but continuous entropy? The answer to this question is in principle "no", but it actually boils down to the so-called critical exponents.

Close to the critical point, the thermodynamic quantities will vanish or diverge according to a power law. More specifically, they will diverge or vanish following some power of the quantity

$$t \equiv \frac{T - T_c}{T_c}, \tag{3.20}$$

where T_c is the critical temperature. For a general quantity $F(t)$, the critical exponent is defined by[9]

$$\lambda = \lim_{t \to 0} \frac{\log |F(t)|}{\log |t|}. \tag{3.21}$$

The reason for this algebraic behavior is subtle, but it has to do with the inexistence of a characteristic length at the critical point, for in criticality, clusters of all sizes can form. We'll discuss this later, when discussing the critical exponent related to the correlation length. Notice that critical exponents are defined only on the critical point, not at coexistence regions.

Let us then follow Kardar (2007a, Sec. 1.4) to discuss the most encountered critical exponents.

The first interesting case is that of the order parameter, which will typically vanish on one of the phases, but not on the other. Hence, it has the behavior

$$\psi \propto \begin{cases} 0, & \text{for } T > T_c, \\ |t|^\beta, & \text{for } T < T_c, \end{cases} \tag{3.22}$$

or the other way around (*i.e.*, exchanging the $>$ and $<$ signs).

For concreteness, this could be, for example, the spontaneuous magnetization of a magnetic system. For high temperatures, the material behaves as a paramagnet, and there is no spontaneuous magnetization. For low temperatures, the material is ferromagnetic and does have a non-vanishing spontaneuous magnetization. The way in which this spontaneous magnetization vanishes as one increases the temperature close to the critical point is described by the critical exponent β. Similarly, for a liquid-gas transition, $\rho_g - \rho_l$, $\rho_g - \rho_c$, and $\rho_l - \rho_c$ also behave according to β. See Fig. 3.11a on page 84.

If we now hold the temperature of a magnetic system fixed at the critical temperature and consider slight deviations of an external magnetic field from zero, the magnetization

---

[9]Eq. (3.21) on the following page is the definition given by Prof. Fiore. In some situations, however, we might be interested in critical exponents when some quantity other than temperature is being varied. Furthermore, $\lambda = 0$ might lead to less information than desired. Reichl (2016, Sec. 4.9.1) discusses situations such as these and admits more general definitions (in addition to Eq. (3.21)).

will also vanish for external field $H \to 0$. This behavior is also described by a critical exponent, δ, according to

$$m(T = T_c, H) \propto H^{\frac{1}{\delta}}. \tag{3.23}$$

For a liquid-gas system, this critical exponent will yield how the density difference vanishes with variations of pressure, and hence it is leads us to understanding how to draw the critical isotherm on Fig. 3.2 on page 72. See Fig. 3.11b on the following page.

Other important thermodynamic functions are the response functions, which exhibit the clear response of the system to external perturbations. These are, for example, the (divergent) isothermal compressibility of a liquid-gas system or the magnetic susceptibility of a magnetic system. For this cases, we write

$$\chi_{\pm}(T, h = 0) \propto |t|^{-\gamma_{\pm}}, \tag{3.24}$$

where the $\pm$ signs correspond to whether we are approaching the critical point from high or low temperatures. While γ_+ and γ_- do not need to be equal, in most cases they are indeed. See Fig. 3.11c on the next page.

The response function associated with temperature is the specific heat, and it deserves a critical exponent of its own, α. We have

$$c_{\pm}(T, h = 0) \propto |t|^{-\alpha_{\pm}}. \tag{3.25}$$

Depending on the signs of the amplitudes of the specific heat on each sign (*i.e.*, depending on the signs of $C_{\pm}$, where $c_{\pm} = C_{\pm}|t|^{-\alpha_{\pm}}$), there might or not be a cusp in the plot for c when $\alpha < 0$. See Figs. 3.11d and 3.11e on the following page.

As mentioned earlier, critical behavior is related to the divergence of correlation lenghts at the critical temperature. It is interesting to get a better understanding of this and to assign a critical exponent to this divergence, so let us consider it as well. For concreteness, let us consider as an example a magnetic system, whose partition function is given generically by

$$Z(T, H) = \sum_{\text{microstates}} e^{-\beta \mathcal{H}_0 + \beta H M}, \tag{3.26}$$

where M is the magnetization of each state, $\mathcal{H}_0$ is the Hamiltonian in each state, and H is the external magnetic field. One can show (see Kardar 2007a, Sec. 1.4) that the expected value of the magnetization is

$$\langle M \rangle = \frac{\partial \log Z}{\partial \beta H} \tag{3.27}$$

and that a consequence of this is that the susceptibility is

$$\chi = \frac{1}{k_B T} \big(\langle M^2 \rangle - \langle M \rangle^2 \big). \tag{3.28}$$

However, the total magnetization can be written as a sum over the contributions from each part of the magnet, *i.e.*,

$$M = \int m(\mathbf{r}) \, \mathrm{d}^3 r \,. \tag{3.29}$$

(a) $\beta = \frac{1}{2}$ *near* $T = T_c$.

(b) $\delta = 3$ *near* $T = T_c$.

(c) $\gamma_- = \gamma_+ = 1$, *with amplitudes* **(d)** $\alpha_- = \alpha_+ = -\frac{1}{3}$, *with ampli-* **(e)** $\alpha_- = \alpha_+ = -\frac{1}{3}$, *with ampli-*
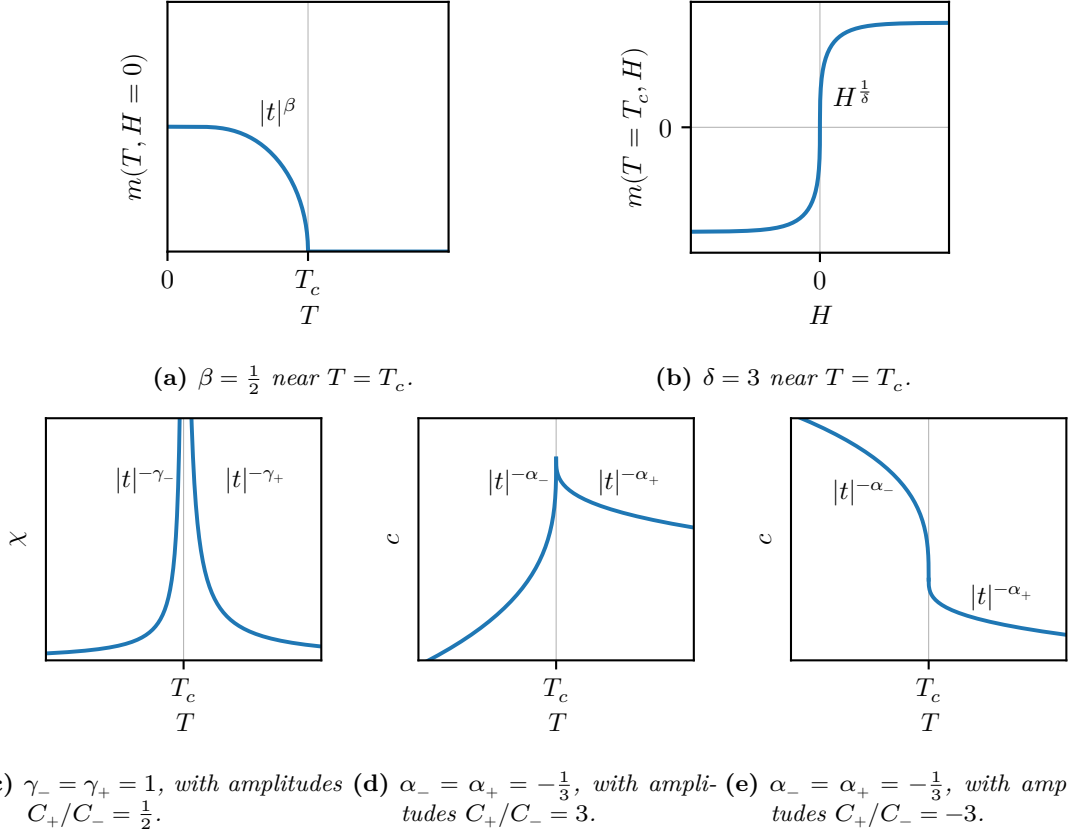$C_+/C_- = \frac{1}{2}$. *tudes* $C_+/C_- = 3$. *tudes* $C_+/C_- = -3$.

Figure 3.11: *Examples of singular behavior near the critical point for a magnetic system. The graphs for the magnetization m and magnetic susceptibility χ as functions of temperature T anfd external magnetic field H correspond to the actual behavior of the Curie–Weiss model for a ferromagnet. The graphs for specific heat c are just examples of possible behaviors. Notice that, for negative α, the signs of the amplitudes determine whether there is a cusp or not.*

We can then write the susceptibility as

$$\chi = \beta \int \langle m(\mathbf{r}_1)m(\mathbf{r}_2)\rangle - \langle m(\mathbf{r}_1)\rangle \langle m(\mathbf{r}_2)\rangle \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,. \tag{3.30}$$

For a homogeneous system, we have translational symmetry, which means $\langle m(\mathbf{r})\rangle = m$ is a constant and $\langle m(\mathbf{r}_1)m(\mathbf{r}_2)\rangle$ has the form $\langle m(\mathbf{r}_1)m(\mathbf{r}_2)\rangle = G(\mathbf{r}_1 - \mathbf{r}_2)$, meaning it depends only on the separation of the two points. Therefore, the connected correlation function

$$\langle m(\mathbf{r}_1)m(\mathbf{r}_2)\rangle_c \equiv \langle m(\mathbf{r}_1)m(\mathbf{r}_2)\rangle - \langle m(\mathbf{r}_1)\rangle \langle m(\mathbf{r}_2)\rangle \tag{3.31}$$

depends only on the separation of the two points. Let us then write

$$\chi = \beta \int \langle m(\mathbf{r}_1)m(\mathbf{r}_2)\rangle - \langle m(\mathbf{r}_1)\rangle \langle m(\mathbf{r}_2)\rangle \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,, \tag{3.32a}$$

$$= \beta \int \langle m(\mathbf{r}_1)m(\mathbf{r}_2)\rangle_c \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,, \tag{3.32b}$$

$$= \beta \int \langle m(\mathbf{r}_1 - \mathbf{r}_2)m(\mathbf{0})\rangle_c \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,, \tag{3.32c}$$

$$= \beta \int \langle m(\mathbf{r})m(\mathbf{0})\rangle_c \, \mathrm{d}^3 r \, \mathrm{d}^3 R \,, \tag{3.32d}$$

$$= \beta V \int \langle m(\mathbf{r})m(\mathbf{0})\rangle_c \, \mathrm{d}^3 r \,, \tag{3.32e}$$

where we defined $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ and $\mathbf{R} = \mathbf{r}_1 - \mathbf{r}_2$ to perform the integration.

Eq. (3.32) on the next page exhibits how a bulk response function (the susceptibility) can depend on microscopic correlations. These correlation functions typically decay for distances larger than a correlation length ξ, often exponentially. These correlations can be probed experimentally: for the phenomenon of critical opalescence mentioned on p. 71, the correlations much be at a length scale comparable to the wavelength of light, which is much larger than the typical atomic distance.

If g is a typical value for the connected correlation function for distances $\|\mathbf{r}\| < \xi$, Eq. (3.32) implies

$$\frac{k_B T \chi}{V} < g\xi^3, \tag{3.33}$$

and hence the divergence of the susceptibility ensures the correlation function must diverge as well, explaining, for example, why we observe critical opalescence. The critical exponents associated with the divergence of the correlation length are $\nu_\pm$, defined through

$$\xi_\pm(T, h = 0) \propto |t|^{\nu_\pm}. \tag{3.34}$$

Universality Classes

A remarkable fact we'll notice in the following discussions is that many different physical systems describing completely different Physics have the same sets of critical exponents. The reason for that is that the critical exponents don't really depend on the details of the system, but rather on fairly general properties, such as dimensionality, symmetries, interaction ranges, and similar features.

A notable example we'll consider is that the Van der Waals gas and the Curie–Weiss model for a ferromagnet present the same set of critical exponents. In other words, they belong to the same universality class.

An example of universality is provided by the experimental data collected by Guggenheim (1945, Fig. 2) and also displayed on the book by Salinas (reproduced on 2001, Fig. 12.3), which shows how the densities of different gases all behave close to the critical point with a critical exponent $\beta \approx \frac{1}{3}$. This means the Van der Waals gas fails to describe the behavior near the critical point appropriately, since it predicts $\beta = \frac{1}{2}$, as we shall see.

Cross reference

A modern approach to understanding these universality classes is in terms of the so-called renormalization group, discussed, *e.g.*, in the texts by Kardar (2007a, Sec. 4.4 and 4.5), Pathria and Beale (2022, Chap. 14), Salinas (2001, Chap. 14), and Zinn-Justin (2007).

3.2 Curie–Weiss Model for a Ferromagnet

At this point, it is interesting for us to pick a more concrete model to do calculations with. While discussing the main notions behind phase transitions, we went back and forth between fluids and ferromagnetic systems because while fluids are more familiar when it comes to phenomenology, the theory of ferromagnetic systems can often be easier to deal with due to the presence of symmetries.

Phenomenology of Ferromagnetism

Let us begin by discussing the phenomenology of ferromagnetism. We'll introduce an *ad hoc* equation that will be adequate to describe a ferromagnet, and which later we'll derive from a statistical model on Section 3.4. Some familiarity with the basic notions concerning magnetic materials will be assumed[10], and more information on this theme can be found on books on Electromagnetism, such as Griffiths (2017, Chap. 6), Wald (2022, Sec. 4.4), and Zangwill (2013, Chap. 13).

Using the canonical ensemble, one can show that a model for an ideal paramagnet of spin $\frac{1}{2}$ obeys the equation (Kardar 2007b, Eq. (4.97); Salinas 2001, Eq. (5.36))

$$m = \mu_0 \tanh\left(\frac{\mu_0 H}{k_B T}\right), \tag{3.35}$$

where m is the magnetization per site. In this case, we note that $H \to 0$ leads to $m \to 0$.

However, there are materials that behave as ferromagnets at low temperatures, but as paramagnets at high temperatures. How could we describe this sort of behavior?

A possibility is to make an *ad hoc* modification of Eq. (3.35) on page 87 and write[11]

$$m = \tanh(\beta\lambda m + \beta H). \tag{3.36}$$

The phenomenological justification for this equation, known as the Curie–Weiss equation, is that the ions themselves present in the magnet have a magnetic field, and each of them will interact with the fields of ions close to them. This interaction with nearest neighbors should then lead to an effective mean field proportional to the local magnetization. For this, this is an *ad hoc* choice, but as mentioned before we'll justify Eq. (3.36) with a statistical model on Section 3.4.

We can solve for H on Eq. (3.36) and plot the resulting equation. One can see that

$$H = \frac{1}{\beta}(\operatorname{artanh} m - \beta\lambda m). \tag{3.37}$$

---

[10]Here's a lightning nomenclature summary: diamagnetic materials have negative susceptibility and hence their magnetization is anti-aligned with an external magnetic field **H**, paramagnetic materials have positive susceptibility and their magnetization is aligned with **H**, ferromagnetic materials have permanent magnetization.

[11]In going from Eq. (3.35) to Eq. (3.36) we changed the meaning of m. On Eq. (3.35) it is simply the magnetization per site, bu on Eq. (3.36) it is the dimensionless magnetization per site, meaning a division by μ_0 is understood. A similar comment is in place for H. Another way of thinking would be simply that we are working in units with $\mu_0 = 1$.
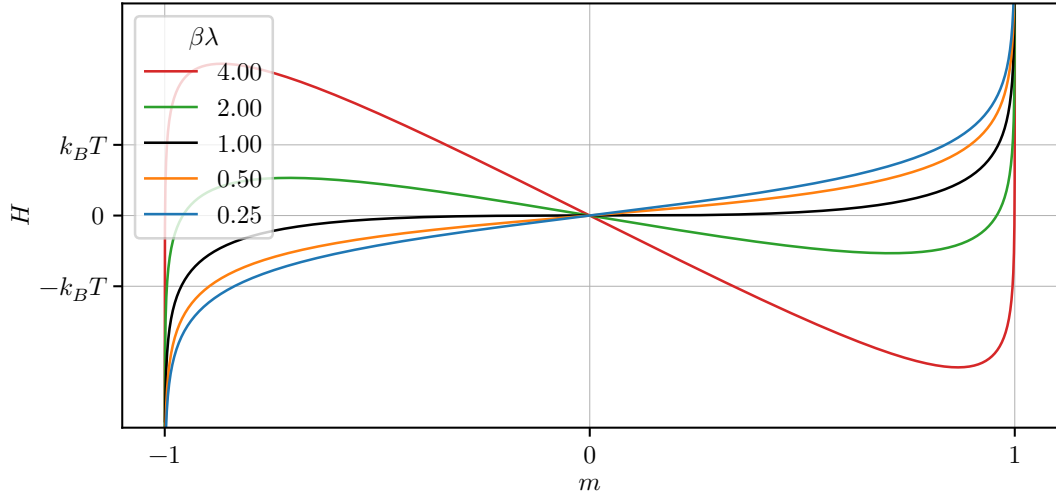
(a) *Isotherms of the Curie–Weiss equation. Notice the "handles" on isotherms with temperature lower than the critical temperature $k_B T_c = \lambda$.*
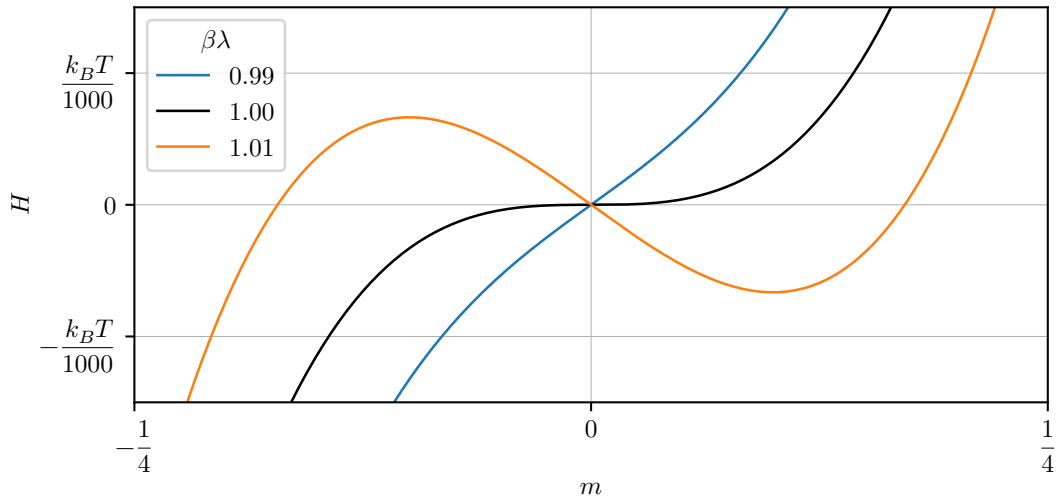
(b) *Isotherms of the Curie–Weiss equation. The "handles" of low-temperature isotherms disappear exactly at the critical temperature.*

Figure 3.12: *Isotherms of the Curie–Weiss equation. The "handles" occurring for low-temperature signal a phase transition, just as with the Van der Waals.*

From Fig. [3.12b] on the following page we can anticipate that $m(H)$ has a divergent derivative at $H = 0$. However, the susceptibility is related to $\frac{\partial m}{\partial H}$, and hence we can see it will blow up as well.

Notice the handles on Fig. [3.12] on the next page: just as in the Van der Waals gas, they are signaling a failure of the mean field approach and should be dealt with by means of the Maxwell construction. Another way of seeing this is by building the Helmholtz free energy of the system (Salinas [2001], Sec. 12.2). The dimensionless magnetization satisfies

$$m = -\left(\frac{\partial g}{\partial H}\right)_T, \tag{3.38}$$

where g plays the role of Gibbs free energy per site. Hence, we can define the magnetic Helmholtz free energy per site as a Legendre transformation and get

$$H = \left(\frac{\partial f}{\partial m}\right)_T. \tag{3.39}$$

Hence, for the Curie–Weiss model,

$$f(T, m) = \frac{1}{\beta}\left(\int \operatorname{artanh}(m)\,\mathrm{d}m - \frac{\beta\lambda}{2}m^2\right) + f_0(T), \tag{3.40a}$$

$$= \frac{1}{\beta}\left(m\operatorname{artanh} m + \frac{1}{2}\log(1 - m^2) - \frac{\beta\lambda}{2}m^2\right) + f_0(T), \tag{3.40b}$$

where $f_0(T)$ is some unknown function. Plotting the Helmholtz free energy for different isotherms leads us to Fig. [3.13] on page [89]. Notice how the Helmholtz free energy fails to be convex for the low-temperature isotherms. Furthermore, to recover the Gibbs free energy we'd do

$$g(T, H) = \inf_m \{f(T, m) - Hm\}, \tag{3.41}$$

but at $H = 0$ this means

$$g(T, H = 0) = \inf_m \{f(T, m)\}, \tag{3.42}$$

and hence the presence of two minima at low-temperature isotherms on Fig. [3.13] signals two possible phases.

There is, however, something different from the Van der Waals case: we now have a $H \to -H$ symmetry, for the system treats equally magnetic fields in either direction. This means the constant field plateaus of the Maxwell construction will always be at $H = 0$, making this model much easier to deal with. At $H = 0$ and low-temperatures, we'll have coexistence of two ferromagnetic phases ($m < 0$ and $m > 0$), but the coexistence vanishes at the critical temperature, $\beta_c\lambda = 1$. The isotherms are plotted on Fig. [3.14] on the following page.

We can also take a different point of view on why there are two phases at low temperatures and zero field. Notice that at $H = 0$ the Curie–Weiss equation, Eq. ([3.36]) on page [87], can be written as

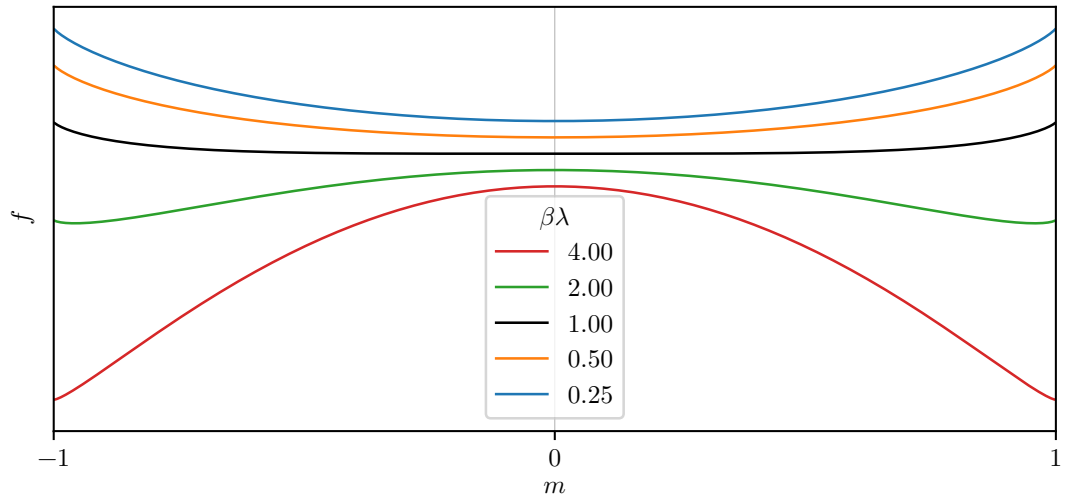$$m = \tanh(\beta\lambda m). \tag{3.43}$$

Figure 3.13: *Helmholtz free energy for the Curie–Weiss model. Notice that low-temperature isotherms do not have a convex Helmholtz free energy. Notice also that this same graph corresponds to $f(T, m) - Hm$ at $H = 0$, and hence the existence of two minima at the same isotherm means there are two possible values of the magnetization.*
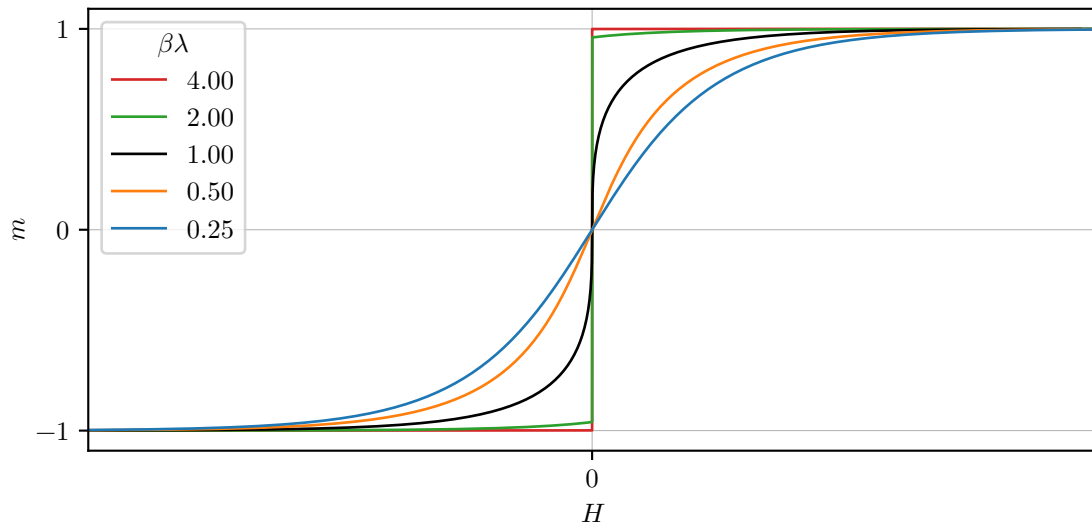

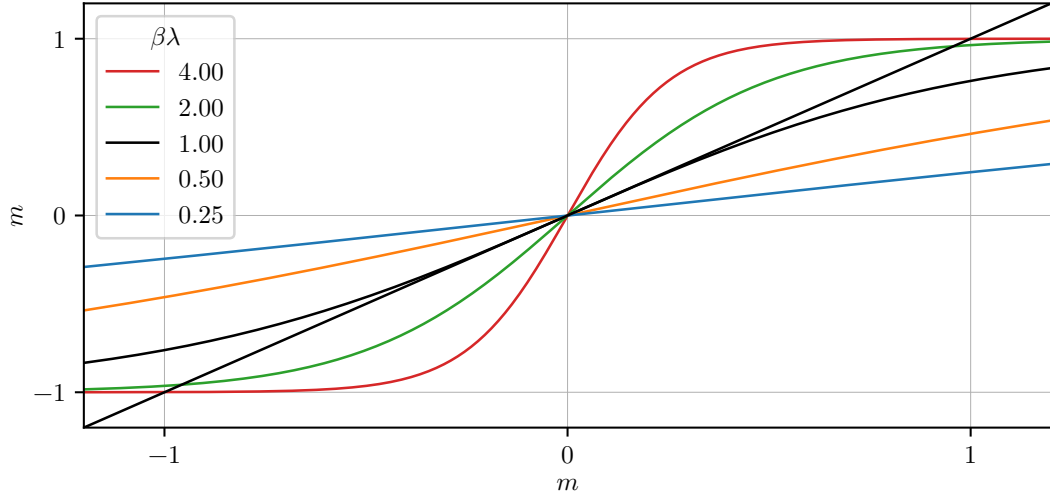**Figure 3.14:** *Isotherms for the Curie–Weiss model after the Maxwell construction has been performed.*

Figure 3.15: *Graphical method for finding the zeros of Eq. (3.43) on page 90. Notice that there are three solutions for $\beta\lambda > 1$. We know that $m = 0$ is an unphysical solution because it is a maximum of the free energy, as seen on Fig. 3.13 on page 89.*

Let us find the zeros of this expression. This can be done graphically by plotting the functions on each side of the equation. This is shown on Fig. 3.15 on the following page. We see that there are three roots for $\beta\lambda > 1$, but a single one ($m = 0$) otherwise. For $\beta\lambda > 1$, $m = 0$ is unphysical for it corresponds to a maximum of the free energy, as seen on Fig. 3.13 on the previous page. The (physical) solutions of Eq. (3.43) as a function of temperature are shown on Fig. 3.16 on the next page.

Critical Exponents

Let us then find the critical exponents of the Curie–Weiss model. We begin with the temperature dependence of the order parameter, β.

Close to the critical point, the magnetization m is small, and hence we can approximate Eq. (3.36) on page 87 by

$$\beta H + \beta\lambda m = \operatorname{artanh} m, \tag{3.44a}$$

$$= \sum_{n=0}^{+\infty} \frac{m^{2n+1}}{2n+1}, \tag{3.44b}$$

$$\approx m + \frac{m^3}{3}. \tag{3.44c}$$

The phase transition happens at $H = 0$, and hence we are left with

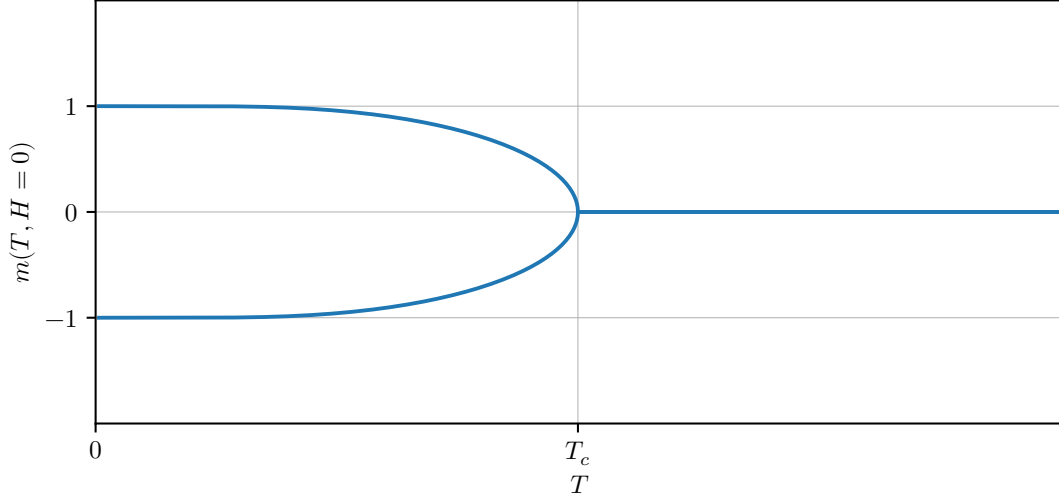$$(\beta\lambda - 1)m \approx \frac{m^3}{3}. \tag{3.45}$$

Figure 3.16: *Magnetization as a function of temperature for the Curie–Weiss model. For low temperatures, we have ferromagnetic domains, but from the critical temperature $k_B T_c = \lambda$ onward the spontaneous magnetization vanishes.*

One of the solutions is $m = 0$, which is not interesting to us because it is not a minimum of the free energy. The non-trivial solutions will be

$$m^2 = 3(\beta\lambda - 1), \tag{3.46a}$$

$$= 3\lambda(\beta - \beta_c), \tag{3.46b}$$

$$= 3\frac{\lambda}{k_B}\left(\frac{1}{T} - \frac{1}{T_c}\right), \tag{3.46c}$$

$$= 3\frac{\lambda}{k_B}\frac{T_c - T}{TT_c}, \tag{3.46d}$$

$$= -3\frac{\lambda t}{k_B T}, \tag{3.46e}$$

where we used t as defined on Eq. (3.20) on page 82. Notice, though, that we can write

$$m^2 = -3\frac{\lambda t}{k_B}\frac{1}{T}, \tag{3.46f}$$

$$= -3\frac{\lambda t}{k_B}\frac{1}{T_c + (T - T_c)}, \tag{3.46g}$$

$$= -3\frac{\lambda t}{k_B T_c}\sum_{n=0}^{+\infty}(-1)^n t^n, \tag{3.46h}$$

$$\approx -3\frac{\lambda t}{k_B T_c}, \tag{3.46i}$$

$$= -3t. \tag{3.46j}$$

Therefore, we see that

$$m \approx \pm\sqrt{3}(-t)^{\frac{1}{2}}, \tag{3.47a}$$

$$\tag{3.47b}$$

meaning $\beta = \frac{1}{2}$. Later, we'll see the Van der Waals gas shares the same critical exponent (and the same goes for the other exponents). Notice that this expression only works for $T < T_c$, but for $T > T_c$ we already know m is set to zero.

Next we work out the isothermal susceptibility, $\chi_T = \left(\frac{\partial m}{\partial H}\right)_T$. Since the susceptibility doesn't need to vanish on either side of the critical temperature, we'll have to compute two exponents.

We start again from

$$\beta H + \beta\lambda m \approx m + \frac{m^3}{3}. \tag{3.48}$$

We know $\chi_T = \left(\frac{\partial m}{\partial H}\right)_T = \left(\frac{\partial H}{\partial m}\right)_T^{-1}$. Hence, we notice that

$$\beta\left(\frac{\partial H}{\partial m}\right)_T + \beta\lambda = 1 + m^2, \tag{3.49a}$$

$$\beta\chi_T^{-1} + \beta\lambda = 1 + m^2. \tag{3.49b}$$

Now we check the behavior on each side.

For $T > T_c$, $m = 0$ and we get

$$\beta\chi_T^{-1} + \beta\lambda = 1, \tag{3.50a}$$

$$\chi_T^{-1} + \lambda = k_B T, \tag{3.50b}$$

$$\chi_T^{-1} = k_B T - \lambda, \tag{3.50c}$$

$$= k_B(T - T_c), \tag{3.50d}$$

$$= k_B T_c t, \tag{3.50e}$$

$$\chi_T = \frac{t^{-1}}{k_B T}, \tag{3.50f}$$

and hence $\gamma_+ = 1$. Notice the amplitude $C_+ = \frac{1}{k_B T}$.

For $T < T_c$, $m^2 = -3t$ and we get

$$\beta\chi_T^{-1} + \beta\lambda = 1 - 3t, \tag{3.51a}$$

$$\beta(\chi_T^{-1} + k_B T_c) = 1 - 3t, \tag{3.51b}$$

$$\chi_T^{-1} = k_B T(1 - 3t) - k_B T_c, \tag{3.51c}$$

$$= k_B T(1 - 3t) - k_B T_c, \tag{3.51d}$$

$$= t k_B T_c(1 - 3t) + k_B T_c(1 - 3t) - k_B T_c, \tag{3.51e}$$

$$= t k_B T_c(1 - 3t) - 3t k_B T_c, \tag{3.51f}$$

$$\approx -2t k_B T_c, \tag{3.51g}$$

$$\chi_T = \frac{(-t)^{-1}}{2k_B T}. \tag{3.51h}$$

and hence $\gamma_- = 1$ as well. Notice the amplitude $C_- = \frac{1}{2k_B T}$. We have the ratio

$$\frac{C_+}{C_-} = 2. \tag{3.52}$$

This ratio is also universal and will be found in other phase transitions, even though the amplitudes themselves aren't.

The computation of the specific heat will be left for the problem set. For the Curie–Weiss model, the associated critical exponent is given by $\alpha_- = \alpha_+ = 0$. There is no divergence, but rather only a jump. This is an issue of the mean field approximation.

At last, let us get the critical exponent for the dependence of the susceptibility with the magnetic field at the critical isotherm. We have

$$\beta_c(H + \lambda m) = m + \frac{m^3}{3}, \tag{3.53a}$$

$$\beta_c H + m = m + \frac{m^3}{3}, \tag{3.53b}$$

$$\beta_c H = \frac{m^3}{3}, \tag{3.53c}$$

and hence $\delta = 3$.

Let us then see how these critical exponents compare to the ones measured experimentally and to other magnetic models (namely, the Ising model in two and three dimensions). The data is compiled on Table 3.1 on page 94.

Table 3.1: *Critical exponents for mean-field approximations, for the Ising model in $d = 2$ and in $d = 3$ dimensions, and experimental data for magnetic systems (these are taken from Pathria and Beale 2022, Tab. 12.1, who cite Stierstadt et al. 1990). This table is based on the table by Salinas (2001, p. 260), from which the data for the Ising model is taken. "0 (log)" means that there is a logarithmic divergence in the specific heat instead of a discontinuity.*

	mean-field	Ising ($d = 2$)	Ising ($d = 3$)	experiment
α_+, α_-	0	0 (log)	$\approx 1/8$	0.0–0.2
β	1/2	1/8	$\approx 5/16$	0.30–0.36
γ_+	1	7/4	$\approx 5/4$	1.2–1.4
γ_-	1	7/4	$\approx 5/4$	1.0–1.2
δ	3	15	≈ 5	4.2–4.8

The Ising model is a lattice model for a ferromagnet that considers only next-neighbors interactions. Hence, a way of understanding why the $d = 3$ results are closer to the mean-field approximation (a path to which is to consider that all sites interact with all sites) is by noticing that in higher dimensions one has more next-neighbors than in lower dimensions. Namely, in $d = 2$ there are four next-neighbors, while in $d = 3$ there are six. Of course, this is not a proof, just a way of getting more intuition.

3.3 Landau's Phenomenological Theory

We can recover the free energy from the order parameter. This is interesting not only for us to later study phase coexistence, but also in the study of Landau's phenomenological theory of phase transitions. Now, we'll obtain an expansion for the free energy in terms of the order parameter both for the Curie–Weiss model and for the Van der Waals model, so we can work first with simple examples and build our way up.

Curie–Weiss Model

In terms of the Gibbs free energy per site $g(T, H)$, we can write

$$m(T, H) = -\left(\frac{\partial g}{\partial H}\right)_T. \tag{3.54}$$

We can then define the Helmholtz free energy $f(T, m)$ of the magnetic system as the Legendre transform of $g(T, H)$ and hence get

$$H = \left(\frac{\partial f}{\partial m}\right)_T, \tag{3.55}$$

which can be solved to get the Helmholtz free energy, as we did on Eq. (3.40) on page 89. This was useful back then for us to find the phases of the system (and, performing the Maxwell construction, we can also see the coexistence regions).

This time, our interest is more focused on foreshadowing Landau's phenomenological theory, so we'll take a different path and write the free energy in a series expansion. From Eq. (3.40) on page 89 one can show that

$$f(T, m) = f_0(T) + \frac{1}{2\beta}(1 - \beta\lambda)m^2 + \frac{m^4}{12\beta} + \frac{m^6}{30\beta} + \mathcal{O}(m^8), \tag{3.56}$$

where all terms are even on m. This is a consequence of the $H \to -H$ symmetry of the magnetic system, and it won't occur for the liquid-gas transition, as we'll see soon.

The Gibbs free energy per site is recovered through the Legendre transformation

$$g(T, H) = \inf_m \{f(T, m) - Hm\} \equiv \inf_m \{g(T, H; m)\}, \tag{3.57}$$

where

$$g(T, H; m) = f_0(T) - Hm + \frac{1}{2\beta}(1 - \beta\lambda)m^2 + \frac{m^4}{12\beta} + \frac{m^6}{30\beta} + \mathcal{O}(m^8). \tag{3.58}$$

If we hadn't plotted Fig. 3.13 on page 89 to check for the minima, we'd be able to use this expansion to check that for low temperatures the Gibbs free energy has two minima, but for high temperatures it has a single one. Of course, Fig. 3.13 on page 89 was possible this time due to the model's simplicity.

Van der Waals gas

Let us now perform a similar calculation for the more complicated Van der Waals gas.

We know that

$$p = -\left(\frac{\partial f}{\partial v}\right)_T, \tag{3.59}$$

and using this expression we have already obtained that the Helmholtz free energy per particle for the Van der Waals gas is (Eq. (3.6) on page 74)

$$f(T, v) = -k_B T \log(v - b) - \frac{a}{v} + f_0(T). \tag{3.60}$$

As usual, the Gibbs free energy per site is obtained with a Legendre transformation

$$g(T, p) = \inf_v \{f(T, v) + pv\} \equiv \inf_v \{g(T, p; v)\}, \tag{3.61}$$

where

$$g(T, p; v) = -k_B T \log(v - b) - \frac{a}{v} + f_0(T) + pv. \tag{3.62}$$

Let us expand this expression in terms of an order parameter. Close to the critical point, the volume is close to the critical volume $v_c = 3b$. Hence, let us make a Taylor expansion of $g(T, p; v)$ about $v = v_c$. Or, equivalently, let us expand in terms of $\psi = v - v_c$. We find that

$$g(T, p; v) = \sum_{n=0}^{+\infty} g_n(T, p)\psi^n, \tag{3.63}$$

where

$$g_0 = f_0(T) + pv_c - k_B T \log(v_c - b) - \frac{a}{v_c}, \tag{3.64a}$$

$$g_1 = p - \frac{k_B T}{v_c - b} + \frac{a}{v_c^2}, \tag{3.64b}$$

$$g_2 = \frac{k_B T}{2(v_c - b)^2} - \frac{a}{v_c^3}, \tag{3.64c}$$

$$g_3 = -\frac{k_B T}{3(v_c - b)^3} + \frac{a}{v_c^4}, \tag{3.64d}$$

$$g_4 = \frac{k_B T}{4(v_c - b)^4} - \frac{a}{v_c^5}, \tag{3.64e}$$

and so on. If we use the values of the critical parameters Eq. (3.13) on page 77 and define

$$t \equiv \frac{T - T_c}{T_c} \quad \text{and} \quad \Delta p \equiv \frac{p - p_c}{p_c}, \tag{3.65}$$

then we can write

$$g_0 = f_0(T) + 3bp - k_B T \log(2b) - \frac{a}{3b}, \tag{3.66a}$$

$$g_1 = -\frac{4a}{27b^2}t + \frac{a}{27b^2}\Delta p, \tag{3.66b}$$

$$g_2 = \frac{a}{27b^3}t, \tag{3.66c}$$

$$g_3 = -\frac{a}{81b^4}t, \tag{3.66d}$$

$$g_4 = \frac{a}{1944b^5} + \frac{a}{216b^5}t. \tag{3.66e}$$

We should compare these expressions to Eq. (3.58) on the previous page: this time, there are odd terms. The reason is that we no longer have a symmetry, which makes the system more complicated to deal with.

At the critical point ($t = 0$, $\Delta p = 0$), we have $g_1 = g_2 = g_3 = 0$, while $g > 4$. Since ψ is small in the critical region, the existence of a minimum is ensured at fourth order, and hence it is possible to analyze the critical region considering only the approximation

$$g(T, p; v) = g_0 + g_1\psi + g_2\psi^2 + g_3\psi^3 + g_4\psi^4. \tag{3.67}$$

At this point, it is convenient to make a different definition of the order parameter with the goal of simplifying this expression. If we define a shifted parameter $\psi' = \psi - c$, we'll find that $c = -\frac{g_3}{4g_4}$ leads us to

$$g(T, p; v) = A_0 + A_1\psi' + A_2\psi'^2 + A_4\psi'^4, \tag{3.68}$$

with

$$A_0 = g_0 - \frac{g_1 g_3}{4g_4} + \frac{g_2 g_3^2}{16g_4^2} - \frac{3g_3^4}{256g_4^3}, \tag{3.69a}$$

$$A_1 = g_1 - \frac{g_2 g_3}{2g_4} + \frac{g_3^3}{8g_4^2}, \tag{3.69b}$$

$$A_2 = g_2 - \frac{3g_3^2}{8g_4}, \tag{3.69c}$$

$$A_4 = g_4. \tag{3.69d}$$

Notice A_1 and A_2 vanish at the critical point, while $A_4 > 0$. Furthermore, ψ' still vanishes at the critical point, since g_3 does as well. Hence, we can also analyze the minima of Eq. (3.68) on the following page to understand the critical behavior. Notice this provides an example of the freedom of choice we have with the order parameter ψ, since we were able to redefine it (with the constraint that it should still vanish at the critical point).

Landau's Phenomenological Theory

To discuss Landau's phenomenological theory, we'll follow the discussions by Landau and Lifshitz (1980, Chap. XIV), Pathria and Beale (2022, Sec. 12.9), and Salinas (2001, Sec. 12.3).

Second-order phase transitions happen continuously. In particular, the order parameter varies continuously until it vanishes in a (typically) high-temperature, (typically) more symmetric phase. Since the order parameter can take arbitrarily small values close to the critical point (this is the meaning of the phase transition being continuous), in the vicinity of the critical point we can write the Gibbs free energy in the form

$$g(T,p;\psi) = g_0(T,p) + g_1(T,p)\psi + g_2(T,p)\psi^2 + g_3(T,p)\psi^3 + g_4(T,p)\psi^4 + \cdots, \quad (3.70)$$

where ψ is the order parameter and we are neglecting higher-order contributions. Notice that not all variables on Eq. (3.70) are on equal footing: while T and p (which can stand also for, *e.g.*, an external magnetic field, not only pressure) are chosen arbitrarily, ψ is to be obtained by minimizing $g(T,p;\psi)$.

The expansion on Eq. (3.70), referred to as the Landau expansion, is similar to what we just did for the Van der Waals gas and for the Curie–Weiss model, but it works in far more generality, since it now applies to way more general continuous transitions[12]. Furthermore, fairly general properties of Eq. (3.70) on the previous page can be obtained in a straightforward manner:

- if $\psi \neq 0$ and $\psi = 0$ correspond to phases with different symmetries, then it is necessary[13] that $g_1(T,p) = 0$, just like we had on the Curie–Weiss model;

- $g_2(T_c, p_c) = 0$, for it must be negative when $\psi \neq 0$ (or there wouldn't be minima in the unsymmetrical phase), and it must be positive when $\psi = 0$ (or $\psi = 0$ would be unstable);

- since $\psi = 0$ should be stable at the critical point, we must have $g_1(T_c, p_c) = 0$, $g_3(T_c, p_c) = 0$, and $g_4(T_c, p_c) > 0$.

Due to the freedom in the definition of ψ, we can also often redefine it to get to the form

$$g(T,p;\psi) = A_0(T,p) + A_1(T,p)\psi + A_2(T,p)\psi^2 + \psi^4. \quad (3.71)$$

Notice we still need to have $A_1(T_c, p_c) = 0$, and $A_2(T_c, p_c) = 0$ from stability arguments.

If we now take the derivatives of Eq. (3.71), we find

$$\frac{\partial g}{\partial \psi} = A_1 + 2A_2\psi + 4\psi^3, \quad (3.72)$$

---

[12]But not to all of them! Salinas (2001, p. 252) mentions that the exact solution to the two-dimensional Ising model is a famous example of a system that can't be put in the form of the Landau expansion.

[13]This is essentially because a linear term will always "lift" one of the minima of the Gibbs free energy, hence favoring one minimum over the other and spoiling the symmetry of the $\psi = 0$ phase. For a more detailed discussion, see the text by Landau and Lifshitz (1980, Sec. 144).

and

$$\frac{\partial^2 g}{\partial \psi^2} = 2A_2 + 12\psi^2. \tag{3.73}$$

Since the physical value of ψ must be a minimum of the free energy, we know $\frac{\partial g}{\partial \psi} = 0$. For the case with $A_1 = 0$, this means either $\psi = 0$ or $\psi^2 = -\frac{A_2}{2}$. Stability is dictated by the second derivative. $\psi = 0$ will be stable for $A_2 > 0$ and $\psi \neq 0$ for $A_2 < 0$.

Let us consider a system with a symmetry, so that only even powers occur on the expression for the free energy. In this case, the most general expansion will be of the form

$$g(T, H; m) = g_0(T) - Hm + A(T)m^2 + B(T)m^4 + \cdots, \tag{3.74}$$

where we adopted the notation usual to ferromagnetic systems, which will typically present this symmetry. Since $A(T_c) = 0$ and $B(T_c) > 0$, we can make an expansion about the critical point to write

$$f_0(T) \approx f_0(T_c), \quad A(T) \approx \frac{a(T - T_c)}{T_c}, \quad B(T) \approx b, \tag{3.75}$$

for constants $a > 0$ and $b > 0$. Hence, about the critical point we have

$$g(T, H; m) = f_0(T_c) - Hm + atm^2 + bm^4, \tag{3.76}$$

where $t = \frac{T - T_c}{T_c}$, as usual. Hence, at $H = 0$, the stable minima of the potential occur at

$$m^2 = -\frac{at}{2b}, \tag{3.77}$$

which immediately yields the critical exponent $\beta = \frac{1}{2}$. The remaining critical exponents and the amplitude ratio C_+/C_- can be computed similarly (see Pathria and Beale 2022, Sec. 12.9).

It is interesting to notice how general the Landau theory is. We were able to obtain the critical exponents without ever picking too many details about the system. It might seem that the Landau phenomenology comprehends all phenomena and enclose then within a single universality class, but that is incorrect. In fact, as pointed out by Pathria and Beale (2022, p. 461) and Peliti (2011, pp. 155–156), the lesson is that the critical exponents are arising from assumptions such as analyticity of the free energy (allowing us to write it as a Taylor series), dimensionality (which in the particular cases of the Landau theory does become irrelevant), and symmetries of the system.

Notice also that more complicated results can be signaled by the Landau phenomenology. As pointed out by Landau and Lifshitz (1980, pp. 452–453) and Salinas (2001, p. 253), assuming the linear term vanishes, there are two sorts of systems: those with an identically vanishing cubic term and those whose cubic term vanishes only at the critical point. Assuming the quadratic coefficient is $A(T, p)$ and the cubic coefficient is $C(T, p)$, the case $C(T, p) = 0$ means the critical point is at $A(T_c, p_c) = 0$, which is a line in a p-T

diagram. On the other hand, if $C(T, p)$ doesn't vanish identically, we have second-order transitions in isolated points only, for the critical point must respect both the conditions $A(T_c, p_c) = 0$ and $C(T_c, p_c) = 0$.

It is curious that we do a Taylor expansion to deal with a problem that is, by its very nature, singular. Pathria and Beale (2022, p. 461) explains that we don't really use a smooth expansion in the process, but technically use the Maxwell corrected free energy. Nevertheless, since the original free energy was obtained by mean field methods, the results still remember the mean field approximation and the critical exponents are those of mean field theory.

A possibility to improve on the Landau phenomenology mentioned by Pathria and Beale (2022, p. 461) is to proceed on the direction of understanding the scaling approach. Salinas (2001, p. 254) mentions the Landau free energy behaves according to a simple scaling law, which we shall now describe. Let g_s be defined (say, for a ferromagnetic system) through

$$g_s(t, H) = -Hm + atm^2 + bm^4, \tag{3.78}$$

where m is defined as a solution to

$$-H + 2atm + 4bm^3 = 0. \tag{3.79}$$

From these two equations, it can be shown that

$$g_s(t, H) = \lambda g_s\left(\lambda^{-\frac{1}{2}} t, \lambda^{-\frac{3}{4}} H\right), \tag{3.80}$$

for any λ. Hence, if we pick λ such that $\lambda^{-\frac{1}{2}} t = 1$ we'll find that

$$g_s(t, H) = t^2 g_s\left(1, \frac{H}{t^{\frac{3}{2}}}\right) = t^2 F\left(\frac{H}{t^{\frac{3}{2}}}\right). \tag{3.81}$$

Salinas (2001, p. 254) also mentions the phenomenological scaling hypotheses can be way more general than the Landau theory, typically assuming only the form

$$g_s(t, H) = t^{2-\alpha} F\left(\frac{H}{t^{\beta\delta}}\right), \tag{3.82}$$

where α and the product $\beta\delta$ should be determined experimentally.

3.4   The Ising Model

To obtain a microscopic model of ferromagnetism, we'll consider a lattice made of N spins with magnetic dipole moments μ, each with strength $g\mu_B\sqrt{j(j+1)}$ and degeneracy $2j+1$. The atoms in the lattice are allowed to interact with their q nearest neighbors and are under the effect of a magnetic field $\mathbf{H}$. For certain dimensions, this model will lead to a ferromagnetic-paramagnetic phase transition at a critical temperature T_c. Most such systems will have a $m(0, T)$ curve better fitted for the choice $j = \frac{1}{2}$, which is another way of saying ferromagnetism is related to the electron's spin, rather than to orbital angular momentum.

Exchange Interaction

The interaction actually responsible for ferromagnetism is known as "exchange interaction", and we shall discuss it in the following. Briefly, when electrons are close to each other, the Pauli exclusion principle keeps their spins from pointing in the same direction, but they can align if they are a bit further apart. Therefore, spatial separation affects the possible electrostatic interactions and, as a consequence, the possible spin orientations.

To see this mathematically, let us notice that, since electrons are fermions, the possible wavefunctions for a pair of electrons are of the form

$$\psi_1(\mathbf{r}_1, s_1; \mathbf{r}_2, s_2) = \phi_S(\mathbf{r}_1, \mathbf{r}_2)\chi_A(s_1, s_2) \tag{3.83}$$

or of the form

$$\psi_2(\mathbf{r}_1, s_1; \mathbf{r}_2, s_2) = \phi_A(\mathbf{r}_1, \mathbf{r}_2)\chi_S(s_1, s_2), \tag{3.84}$$

where ϕ stands for the spatial wavefunction, χ for the spin state, and the subindices A and S stand for "antisymmetric" and "symmetric", respectively. The options for ϕ are

$$\phi_A(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}[\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) - \phi_1(\mathbf{r}_2)\phi_2(\mathbf{r}_1)], \tag{3.85}$$

and

$$\phi_S(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}[\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) + \phi_1(\mathbf{r}_2)\phi_2(\mathbf{r}_1)]. \tag{3.86}$$

$\chi_A(s_1, s_2)$ corresponds to the singlet state

$$|\chi_A\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{3.87}$$

$\chi_S(s_1, s_2)$ can correspond to any of the triplet states,

$$|\chi_S\rangle = |\uparrow\uparrow\rangle, \quad \text{or} \quad |\chi_S\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), \quad \text{or} \quad |\chi_S\rangle = |\downarrow\downarrow\rangle. \tag{3.88}$$

Let us estimate the difference in energy between the states $|\psi_1\rangle$ and $|\psi_2\rangle$. This can be done by considering the expression

$$\left\langle \psi_1 \left| -\frac{e^2}{r_{12}} \right| \psi_1 \right\rangle - \left\langle \psi_2 \left| -\frac{e^2}{r_{12}} \right| \psi_2 \right\rangle. \tag{3.89}$$

However, we know that we can write

$$|\psi_{1,2}\rangle = \frac{1}{\sqrt{2}}(|\phi_1, \phi_2\rangle \pm |\phi_2, \phi_1\rangle) \otimes |\chi_{A,S}\rangle, \tag{3.90}$$

with the negative sign holding for the spatially antisymmetric combination. Using this we find that

$$\left\langle \psi_1 \left| -\frac{e^2}{r_{12}} \right| \psi_1 \right\rangle = \left\langle \phi_S \left| -\frac{e^2}{r_{12}} \right| \phi_S \right\rangle, \tag{3.91a}$$

$$= \int \langle \phi_S | \mathbf{r}_1, \mathbf{r}_2 \rangle \, \langle \mathbf{r}_1, \mathbf{r}_2 | -\frac{e^2}{r_{12}} | \phi_S \rangle \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,, \qquad (3.91\text{b})$$

$$= A + B + C + D, \qquad (3.91\text{c})$$

where A, B, C, and D are given by

$$A = \frac{1}{2} \int |\phi_1(\mathbf{r}_1)|^2 |\phi_2(\mathbf{r}_2)|^2 \left(-\frac{e^2}{r_{12}} \right) \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,, \qquad (3.92\text{a})$$

$$B = \frac{1}{2} \int |\phi_1(\mathbf{r}_2)|^2 |\phi_2(\mathbf{r}_1)|^2 \left(-\frac{e^2}{r_{12}} \right) \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,, \qquad (3.92\text{b})$$

$$C = \frac{1}{2} \int \phi_1^*(\mathbf{r}_1) \phi_2^*(\mathbf{r}_2) \left(-\frac{e^2}{r_{12}} \right) \phi_1(\mathbf{r}_2) \phi_2(\mathbf{r}_1) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,, \qquad (3.92\text{c})$$

$$D = \frac{1}{2} \int \phi_1^*(\mathbf{r}_2) \phi_2^*(\mathbf{r}_1) \left(-\frac{e^2}{r_{12}} \right) \phi_1(\mathbf{r}_1) \phi_2(\mathbf{r}_2) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \,. \qquad (3.92\text{d})$$

Similarly,

$$\left\langle \psi_2 \left| -\frac{e^2}{r_{12}} \right| \psi_2 \right\rangle = A + B - C - D, \qquad (3.93)$$

and hence the difference in energy between the singlet and triplet states is $2(C + D)$, which corresponds only to the "exchange" terms. Such an energy is closely dependent on the Coulombian interaction.

Hamiltonian for a Lattice Model

Our goal will now be to model the exchange interaction by means of a simplified Hamiltonian, of the form

$$\mathcal{H} = c \mathbf{S}_1 \cdot \mathbf{S}_2 + d. \qquad (3.94)$$

There are other options for lattice models (see Kardar 2007a, Sec. 6.1), but we shall start with this Hamiltonian and walk towards the simpler Ising model, which already presents a lot of interesting Physics.

compute

$$\langle s | \mathbf{S}_1 \cdot \mathbf{S}_2 | s \rangle - \langle t | \mathbf{S}_1 \cdot \mathbf{S}_2 | t \rangle \,, \qquad (3.95)$$

where $|s\rangle$ stands for the singlet state and $|t\rangle$ for the triplet.

Recall that

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = S_{1x} S_{2x} + S_{1y} S_{2y} + S_{1z} S_{2z}, \qquad (3.96)$$

where each of the spin matrices is given in terms of Pauli matrices by

$$S_{\alpha,i} = \frac{\hbar}{2} \sigma_i, \qquad (3.97)$$

with $\alpha = 1, 2$ and $i = 1, 2, 3$.

Using the properties of the Pauli matrices and considering the singlet and triplet states written in terms of the σ_z eigenvectors as

$$|s\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle), \quad \text{and} \quad |t\rangle = \frac{1}{\sqrt{2}}(|+-\rangle + |-+\rangle), \tag{3.98}$$

one can show that

$$\langle s|c\mathbf{S}_1 \cdot \mathbf{S}_2|s\rangle - \langle t|c\mathbf{S}_1 \cdot \mathbf{S}_2|t\rangle = -\frac{c3\hbar^2}{4} - \frac{c\hbar^2}{4} = c\hbar^2. \tag{3.99}$$

From a qualitative point of view, this resembles the exchange interaction term. The coefficient c can then be identified as representing the inner products of the exchange interaction. For $c > 0$, the model will favor ferromagnetism, while $c < 0$ leads to antiferromagnetism.

Therefore, we shall consider the Hamiltonian

$$\mathcal{H} = -\tilde{J}\sum_{(i,j)} \mathbf{S}_i \cdot \mathbf{S}_j - \tilde{\mathbf{H}} \cdot \sum_i \mathbf{S}_i, \tag{3.100}$$

where (i,j) denotes a sum over nearest neighbors. Kardar (2007a, p. 99) refers to this as the O(n) model (n being the number of entries in each of the vectors $\mathbf{S}_i$), and points out that different choices of n are known particular cases. $n = 3$ yields the Heisenberg model, $n = 2$ the *XY*-model, and $n = 1$ the Ising model, which is the one we're interested in.

Hence, from now on we'll focus on the subcase with $\mathbf{S}_i \cdot \mathbf{S}_j = S_{i,z}S_{j,z}$, where $S_{i,z} = \pm\frac{1}{2}$. A simple variable redefinition allows us to write

$$\mathcal{H} = -J\sum_{(i,j)} \sigma_i\sigma_j - h\sum_i \sigma_i, \tag{3.101}$$

where $\sigma_i = \pm 1$.

Bragg–Williams Mean Field Approximation

Our first approach to deal with the Ising model will be to perform a form of mean field approximation known as the Bragg–Williams approximation. It consists in neglecting fluctuations in the correlation functions. In other words, we write

$$\langle \sigma_i\sigma_j \rangle \approx \langle \sigma_i \rangle \langle \sigma_j \rangle. \tag{3.102}$$

Physically, this means that the spins care only about the overall magnetization, rather than the values of their nearest neighbors.

If we employ Eq. (3.102) on the following page on the mean energy obtained from Eq. (3.101) on the previous page, we find

$$U = \langle \mathcal{H} \rangle, \tag{3.103a}$$

$$= -J\sum_{(i,j)} \langle \sigma_i\sigma_j \rangle - h\sum_i \langle \sigma_i \rangle, \tag{3.103b}$$

$$\approx -J \sum_{(i,j)} \langle \sigma_i \rangle \langle \sigma_j \rangle - h \sum_i \langle \sigma_i \rangle, \tag{3.103c}$$

$$= -JdN \langle \sigma \rangle \langle \sigma \rangle - hN \langle \sigma \rangle, \tag{3.103d}$$

where N is the number of spins and d is the system's dimensionality.

$\langle \sigma \rangle$ is nothing but the system's magnetization, m. If we have N_+ spins pointing up and N_- pointing down, we'll have

$$m = \langle \sigma \rangle = \frac{N_+ - N_-}{N}. \tag{3.104}$$

Hence, the mean energy is

$$U = -JdNm^2 - hNm. \tag{3.105}$$

The entropy can be found in terms of N_+ and N_-. It is given by

$$S = k_B \log \left(\frac{N!}{N_+! N_-!} \right), \tag{3.106a}$$

$$= k_B \log \left[\frac{N!}{\left(\frac{N+M}{2} \right)! \left(\frac{N-M}{2} \right)!} \right], \tag{3.106b}$$

where we used the facts that $N_+ + N_- = N$ and $N_+ - N_- = Nm \equiv M$. With the Stirling approximation, we find that

$$S = k_B \log N! - k_B \log \left(\frac{N+M}{2} \right)! - k_B \log \left(\frac{N-M}{2} \right)!, \tag{3.107a}$$

$$= k_B N \log N - k_B N - k_B \left(\frac{N+M}{2} \right) \log \left(\frac{N+M}{2} \right) + k_B \left(\frac{N+M}{2} \right)$$
$$- k_B \left(\frac{N-M}{2} \right) \log \left(\frac{N-M}{2} \right) + k_B \left(\frac{N-M}{2} \right), \tag{3.107b}$$

$$= k_B N \log N - k_B \left(\frac{N+M}{2} \right) \log \left(\frac{N+M}{2} \right) - k_B \left(\frac{N-M}{2} \right) \log \left(\frac{N-M}{2} \right), \tag{3.107c}$$

$$s = k_B \log N - k_B \left(\frac{1+m}{2} \right) \log \left(\frac{N+M}{2} \right) - k_B \left(\frac{1-m}{2} \right) \log \left(\frac{N-M}{2} \right), \tag{3.107d}$$

$$= k_B \log N - k_B \left(\frac{1+m}{2} \right) \log \left(\frac{1+m}{2} \right) - k_B \left(\frac{1-m}{2} \right) \log \left(\frac{1-m}{2} \right)$$
$$- k_B \left(\frac{1+m}{2} \right) \log N - k_B \left(\frac{1-m}{2} \right) \log N, \tag{3.107e}$$

$$= -k_B \left(\frac{1+m}{2} \right) \log \left(\frac{1+m}{2} \right) - k_B \left(\frac{1-m}{2} \right) \log \left(\frac{1-m}{2} \right). \tag{3.107f}$$

The Gibbs free energy will be given by

$$g(T, h) = \min_m \left\{ g(T, h; m) \right\}, \tag{3.108}$$

where

$$g(T, h; m) = u(T, h; m) - TS(m), \tag{3.109a}$$

$$= -Jdm^2 - hm + k_BT\left[\left(\frac{1+m}{2}\right)\log\left(\frac{1+m}{2}\right) + \left(\frac{1-m}{2}\right)\log\left(\frac{1-m}{2}\right)\right]. \tag{3.109b}$$

Let us then find these minima. One can show that

$$0 = \left(\frac{\partial g}{\partial m}\right), \tag{3.110a}$$

$$= -2Jdm - h + \frac{k_BT}{2}\log\left(\frac{1+m}{1-m}\right), \tag{3.110b}$$

$$= -2Jdm - h + k_BT\,\mathrm{artanh}\,m, \tag{3.110c}$$

from which we find that

$$m = \tanh(2\beta Jdm + \beta h), \tag{3.111}$$

which is just the Curie–Weiss equation. Hence, as promised, we have derived our previously phenomenological prescription from a statistical model.

Hence, the Bragg–Williams approximation for the Ising model will recover the results we previously obtained from phenomenological considerations. In particular, it predicts a phase transition even at $d = 1$, despite the fact that the exact Ising model does not have such a feature.

Curie–Weiss Model

A different road to the Curie–Weiss equation is by means of the Curie–Weiss model. While in the Bragg–Williams approximation we discarded the correlations, this time we'll keep them, but we'll modify the model to keep it exactly solvable. Namely, instead of Eq. (3.101) on the preceding page, we'll work with

$$\mathcal{H} = -\frac{J}{2N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j - h\sum_{i=1}^{N}\sigma_i, \tag{3.112}$$

and hence we now consider a model in which all spins interact with all spins. The division by N in the first term is to ensure the thermodynamic limit is still well defined.

This time, we'll compute the partition function for the theory. We know it will be given by

$$Z = \sum_{\{\sigma\}}\exp\left(\frac{\beta J}{2N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i\sigma_j + \beta h\sum_{i=1}^{N}\sigma_i\right), \tag{3.113a}$$

$$= \sum_{\{\sigma\}}\exp\left[\frac{\beta J}{2N}\left(\sum_{i=1}^{N}\sigma_i\right)^2 + \beta h\sum_{i=1}^{N}\sigma_i\right]. \tag{3.113b}$$

To compute this sum, we'll notice that

$$\int_{-\infty}^{+\infty} e^{-x^2+2ax}\,\mathrm{d}x = \sqrt{\pi}e^{a^2}, \qquad (3.114)$$

which allows us to write

$$\exp\left[\frac{\beta J}{2N}\left(\sum_{i=1}^{N}\sigma_i\right)^2\right] = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\exp\left[-x^2 + \sqrt{\frac{2\beta J}{N}}\left(\sum_{i=1}^{N}\sigma_i\right)x\right]\mathrm{d}x. \qquad (3.115)$$

The advantage of this trick is that now we can put the partition function in a form in which it decomposes and resembles that of a non-interacting system. More specifically, we can write

$$Z = \sum_{\{\sigma\}}\exp\left[\frac{\beta J}{2N}\left(\sum_{i=1}^{N}\sigma_i\right)^2 + \beta h\sum_{i=1}^{N}\sigma_i\right], \qquad (3.116a)$$

$$= \frac{1}{\sqrt{\pi}}\sum_{\{\sigma\}}\int_{-\infty}^{+\infty}\exp\left[-x^2 + \left(\sqrt{\frac{2\beta J}{N}}x + \beta h\right)\left(\sum_{i=1}^{N}\sigma_i\right)\right]\mathrm{d}x, \qquad (3.116b)$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-x^2}\sum_{\{\sigma\}}\exp\left[\left(\sqrt{\frac{2\beta J}{N}}x + \beta h\right)\left(\sum_{i=1}^{N}\sigma_i\right)\right]\mathrm{d}x, \qquad (3.116c)$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-x^2}\sum_{\{\sigma\}}\prod_{i=1}^{N}\exp\left[\left(\sqrt{\frac{2\beta J}{N}}x + \beta h\right)\sigma_i\right]\mathrm{d}x, \qquad (3.116d)$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-x^2}\left(\sum_{\sigma}\exp\left[\left(\sqrt{\frac{2\beta J}{N}}x + \beta h\right)\sigma\right]\right)^N\mathrm{d}x, \qquad (3.116e)$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-x^2}\left(e^{-\left(\sqrt{\frac{2\beta J}{N}}x+\beta h\right)} + e^{+\left(\sqrt{\frac{2\beta J}{N}}x+\beta h\right)}\right)^N\mathrm{d}x, \qquad (3.116f)$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-x^2}\left[2\cosh\left(\sqrt{\frac{2\beta J}{N}}x + \beta h\right)\right]^N\mathrm{d}x, \qquad (3.116g)$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-x^2+N\log\left[2\cosh\left(\sqrt{\frac{2\beta J}{N}}x+\beta h\right)\right]}\mathrm{d}x. \qquad (3.116h)$$

Let us change variables by defining m through

$$\sqrt{\frac{2\beta J}{N}}x = \beta J m. \qquad (3.117)$$

This allows us to write

$$Z = \sqrt{\frac{\beta J N}{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{\beta J N}{2}m^2+N\log\left[2\cosh\left(\beta J m+\beta h\right)\right]}\mathrm{d}m. \qquad (3.118)$$

Or, alternatively,

$$Z = \sqrt{\frac{\beta J N}{2\pi}} \int_{-\infty}^{+\infty} e^{-\beta N g(T,h;m)} \, \mathrm{d}m \,, \tag{3.119}$$

where

$$g(T,h;m) = \frac{Jm^2}{2} - \frac{1}{\beta} \log \left[2 \cosh \left(\beta Jm + \beta h\right)\right]. \tag{3.120}$$

To obtain the Gibbs free energy in the thermodynamic limit, we can use Laplace's method (see Erdélyi 1956, Sec. 2.4) to compute Eq. (3.119) on the next page. We'll end up with

$$g(T,h) = \lim_{N \to +\infty} \left(-\frac{1}{\beta N} \log Z\right) = \min_{m} \{g(T,h;m)\}. \tag{3.121}$$

We can use Eq. (3.120) to find this minimum. We notice that

$$\frac{\partial g}{\partial m} = Jm - J \tanh(\beta Jm + \beta h) = 0, \tag{3.122}$$

from which we find

$$m = \tanh(\beta Jm + \beta h), \tag{3.123}$$

which is the Curie–Weiss equation.

Notice Eqs. (3.109) and (3.120) on page 104 and on this page are particularly different, and hence they have different expansions in the Landau phenomenology. In spite of that, they lead to the very same critical exponents corresponding to mean field theory.

Bogoliubov Inequality

Yet another mean field theoretic approach we can employ to simplify the Ising model is the so-called Bogoliubov inequality. This time, we'll consider the system we can't solve exactly (the Ising model) as a modification of some other system we can solve. If we denote the Hamiltonian of the latter by $\mathcal{H}_0$ and that of the former by $\mathcal{H}$, we can write

$$\mathcal{H}(\lambda) = \mathcal{H}_0 + \lambda \mathcal{H}_1, \tag{3.124}$$

where λ is a parameter that allows us to "turn on and off" the effects of $\mathcal{H}_1$. For $\lambda = 1$, we get the Ising model. For $\lambda = 0$, we get the simplified model we can solve exactly.

The so-called Bogoliubov inequality is an estimate on the free energy associated to each of the systems being considered. More specifically, the free energy for the system with parameter λ respects

$$F(\lambda) \leq F(0) + \lambda \langle \mathcal{H}_1 \rangle_0 \,, \tag{3.125}$$

where $\langle \cdot \rangle_0$ denotes an ensemble average under the Hamiltonian $\mathcal{H}_0$.

To show this estimate, let us begin by noticing the free energy[14] is given by

$$-\beta F(\lambda) = \log\!\left(\sum_\sigma e^{-\beta E_\sigma(\lambda)}\right) \equiv \log\!\left(\operatorname{tr} e^{-\beta\mathcal{H}(\lambda)}\right), \tag{3.126}$$

where the logarithm's argument is the partition function and the sum runs over the possible configurations (states) of the system. The trace defined by the second equality corresponds simply to a sum of the operator's eigenvalues (in the quantum theory), or to a sum over the possible energies (in the classical theory). We introduce this notation to keep the expressions a bit more compact and to keep them in the same notation used by Callen (1985, Sec. 20.1).

If we differentiate Eq. (3.126) on the next page with respect to λ, we get to

$$-\beta\frac{\mathrm{d}F}{\mathrm{d}\lambda} = -\beta\frac{\operatorname{tr}\mathcal{H}_1 e^{-\beta(\mathcal{H}_0+\lambda\mathcal{H}_1)}}{\operatorname{tr} e^{-\beta(\mathcal{H}_0+\lambda\mathcal{H}_1)}}, \tag{3.127a}$$

$$\frac{\mathrm{d}F}{\mathrm{d}\lambda} = \langle\mathcal{H}_1\rangle_\lambda, \tag{3.127b}$$

where $\langle\cdot\rangle_\lambda$ denotes an ensemble average under the Hamiltonian $\mathcal{H}(\lambda)$.

Similarly, a second derivative yields

$$\frac{\mathrm{d}^2 F}{\mathrm{d}\lambda^2} = \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{\operatorname{tr}\mathcal{H}_1 e^{-\beta(\mathcal{H}_0+\lambda\mathcal{H}_1)}}{\operatorname{tr} e^{-\beta(\mathcal{H}_0+\lambda\mathcal{H}_1)}}\right], \tag{3.128a}$$

$$= -\beta\left[\frac{\operatorname{tr}\mathcal{H}_1^2 e^{-\beta(\mathcal{H}_0+\lambda\mathcal{H}_1)}}{\operatorname{tr} e^{-\beta(\mathcal{H}_0+\lambda\mathcal{H}_1)}} - \left(\frac{\operatorname{tr}\mathcal{H}_1 e^{-\beta(\mathcal{H}_0+\lambda\mathcal{H}_1)}}{\operatorname{tr} e^{-\beta(\mathcal{H}_0+\lambda\mathcal{H}_1)}}\right)^2\right], \tag{3.128b}$$

$$= -\beta\left[\langle\mathcal{H}_1^2\rangle_\lambda - \langle\mathcal{H}_1\rangle_\lambda^2\right], \tag{3.128c}$$

$$= -\beta\left\langle\mathcal{H}_1 - \langle\mathcal{H}_1\rangle_\lambda\right\rangle_\lambda^2. \tag{3.128d}$$

Notice Eq. (3.128) on the following page implies $\frac{\mathrm{d}^2 F}{\mathrm{d}\lambda^2} \le 0$ for any values of λ, *i.e.*, $F(\lambda)$ is an everywhere concave function of λ. As a consequence, it must lie beneath the line tangent to itself at λ. Mathematically,

$$F(\lambda) \le F(0) + \lambda\langle\mathcal{H}_1\rangle_0, \tag{3.129}$$

where we used Eq. (3.127) on the previous page. This inequality is represented graphically on Fig. 3.17.

Let us now apply this result to the Ising model. The Hamiltonian we're interested in studying is Eq. (3.101) on page 102. The exactly solvable Hamiltonian $\mathcal{H}_0$ we'll be considering is

$$\mathcal{H}_0 = -\eta\sum_i \sigma_i. \tag{3.130}$$

---

[14]The free energy might be the Helmholtz or the Gibbs free energy, depending on the particular problem we're interested in and on the ensemble we're working with. Callen (1985, Sec. 20.1) states it in terms of the Helmholtz free energy, while Prof. Fiore's lecture notes state the result in terms of the Gibbs free energy. The difference, as far as I can tell, is only on whether one is working in the canonical ensemble or in the Gibbs canonical ensemble (see Kardar 2007b, Sec. 4.8).
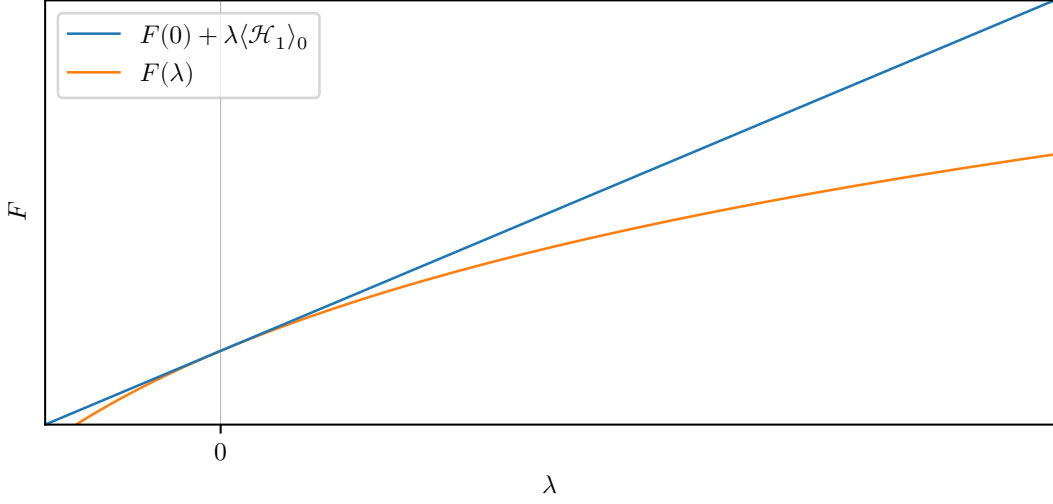
Figure 3.17: *Grafical representation of the Bogoliubov inequality, Eq. (3.129) on page 108.*

Hence, we get

$$\mathcal{H}_1 = \mathcal{H} - \mathcal{H}_0 = -J \sum_{(i,j)} \sigma_i \sigma_j - (h - \eta) \sum_i \sigma_i. \tag{3.131}$$

For $\mathcal{H}_0$ we have

$$Z_0 = \sum_{\{\sigma\}} e^{\beta \eta \sum_{i=1}^N \sigma_i}, \tag{3.132a}$$

$$= \sum_{\{\sigma\}} \prod_{i=1}^N e^{\beta \eta \sigma_i}, \tag{3.132b}$$

$$= \left(\sum_\sigma e^{\beta \eta \sigma} \right)^N, \tag{3.132c}$$

$$= (2 \cosh \beta \eta)^N, \tag{3.132d}$$

and therefore

$$F(0) = -N k_B T \log(2 \cosh \beta \eta). \tag{3.133}$$

Hence,

$$\langle \sigma \rangle_0 = m, \tag{3.134a}$$

$$= -\frac{1}{N} \left(\frac{\partial F}{\partial h} \right)_T, \tag{3.134b}$$

$$= \tanh \beta \eta. \tag{3.134c}$$

Furthermore,

$$\langle \mathcal{H}_1 \rangle_0 = -J \sum_{(i,j)} \langle \sigma_i \sigma_j \rangle_0 - (h - \eta) \sum_i \langle \sigma_i \rangle_0, \tag{3.135a}$$

– 108 –

$$= -J \sum_{(i,j)} \langle \sigma_i \rangle_0 \langle \sigma_j \rangle_0 - (h - \eta) \sum_i \langle \sigma_i \rangle_0, \tag{3.135b}$$

$$= -JNd \langle \sigma \rangle_0 \langle \sigma \rangle_0 - (h - \eta)N \langle \sigma \rangle_0, \tag{3.135c}$$

$$= -JNd \tanh^2(\beta\eta) - (h - \eta)N \tanh\beta\eta. \tag{3.135d}$$

Hence, we learn that the free energy for the problem we're interested in satisfies

$$f(T, h) \leq \frac{F(0)}{N} + \frac{\langle \mathcal{H}_1 \rangle_0}{N}, \tag{3.136a}$$

$$= -k_B T \log(2 \cosh\beta\eta) - Jd \tanh^2(\beta\eta) - (h - \eta) \tanh\beta\eta. \tag{3.136b}$$

Since this holds for any η, we can minimize the expression on the right to get the most stringent estimate. Therefore, we write

$$f(T, h) \leq -\max_\eta \left\{ k_B T \log(2 \cosh\beta\eta) + Jd \tanh^2(\beta\eta) + (h - \eta) \tanh\beta\eta \right\} \tag{3.137a}$$

(notice we pulled a minus sign out of the maximum, which is why we wrote a maximum instead of a minimum). If we define

$$f(T, h; \eta) = k_B T \log(2 \cosh\beta\eta) + Jd \tanh^2(\beta\eta) + (h - \eta) \tanh\beta\eta, \tag{3.138}$$

we can find the point we are interested in by noticing that

$$\frac{\partial f}{\partial \eta} = 2Jd\beta \operatorname{sech}^2(\beta\eta) \tanh\beta\eta + \beta(h - \eta) \operatorname{sech}^2(\beta\eta), \tag{3.139}$$

and hence $\frac{\partial f}{\partial \eta} = 0$ translates to

$$2Jd \tanh\beta\eta = (\eta - h). \tag{3.140}$$

If we now recall Eq. (3.134) on the following page and simplify the resulting expressions, we see that

$$2Jd \tanh\beta\eta = (\eta - h), \tag{3.141a}$$

$$2Jdm + h = \eta, \tag{3.141b}$$

$$2\beta Jdm + \beta h = \beta\eta, \tag{3.141c}$$

$$\tanh(2\beta Jdm + \beta h) = \tanh(\beta\eta), \tag{3.141d}$$

$$\tanh(2\beta Jdm + \beta h) = m, \tag{3.141e}$$

which is, for the third time, the Curie–Weiss equation.

Transfer Matrix Method

While mean field approximations are interesting and provide us with important information, we also wonder whether there is a way to solve the Ising model exactly. In fact, for $d = 1$ and $d = 2$ there is. In this section, we'll use the $d = 1$ case to illustrate the transfer

matrix method, which can be generalized for other systems and dimensionalities to obtain information about interacting systems. For $d = 2$, see, *e.g.*, the texts by Kardar (2007a, Chap. 7) and Pathria and Beale (2022, Chap. 13). For the $d = 1$ case, we'll follow Prof. Fiore's lectures, supplemented by the texts by Kardar (2007a, Sec. 6.2) and Salinas (2001, Sec. 13.1).

In one dimensions, the Hamiltonian for the Ising model, Eq. (3.101) on page 102, can be written as

$$\mathcal{H} = -J \sum_{i=1}^{N} \sigma_i \sigma_{i+1} - h \sum_{i=1}^{N} \sigma_i, \tag{3.142}$$

where we are assuming periodic boundary conditions: $\sigma_{N+1} = \sigma_1$. Hence, the Ising chain is arranged in a circle.

Our goal will be to write the partition function for the model as the trace of a power of a matrix, known as the transfer matrix. Since the trace is basis independent, we can compute it by diagonalizing the transfer matrix. Furthermore, we'll see that the partition function in the thermodynamic limit is dominated by the largest eigenvalue, and hence we don't even need to find all eigenvalues. This will allows us to obtain an exact solution to the model.

The partition function can be written as

$$Z = \sum_{\{\sigma\}} e^{\beta J \sum_{i=1}^{N} \sigma_i \sigma_{i+1} + \beta h \sum_{i=1}^{N} \sigma_i}, \tag{3.143a}$$

$$= \sum_{\{\sigma\}} e^{\beta J \sum_{i=1}^{N} \sigma_i \sigma_{i+1} + \frac{\beta h}{2} \sum_{i=1}^{N} (\sigma_i + \sigma_{i+1})}, \tag{3.143b}$$

$$= \sum_{\{\sigma\}} \prod_{i=1}^{N} e^{\beta J \sigma_i \sigma_{i+1} + \frac{\beta h}{2} (\sigma_i + \sigma_{i+1})}. \tag{3.143c}$$

We can now understand the quantity inside the product as the elements of a matrix. More specifically, we define the transfer matrix $\hat{T}$ as having the elements

$$\langle \sigma_i | \hat{T} | \sigma_j \rangle = e^{\beta J \sigma_i \sigma_j + \frac{\beta h}{2} (\sigma_i + \sigma_j)}. \tag{3.144}$$

Since $\sigma_i = \pm 1$, this means $\hat{T}$ is given by

$$\hat{T} = \begin{pmatrix} e^{\beta J + \beta h} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J - \beta h} \end{pmatrix}. \tag{3.145}$$

The advantage of this definition is that we can now write

$$Z = \sum_{\{\sigma\}} \prod_{i=1}^{N} \langle \sigma_i | \hat{T} | \sigma_{i+1} \rangle, \tag{3.146a}$$

$$= \sum_{\{\sigma\}} \langle \sigma_1 | \hat{T} | \sigma_2 \rangle \langle \sigma_2 | \hat{T} | \sigma_3 \rangle \cdots \langle \sigma_N | \hat{T} | \sigma_1 \rangle, \tag{3.146b}$$

$$= \sum_{\sigma_1, \sigma_2, \dots, \sigma_N} \langle \sigma_1 | \hat{T} | \sigma_2 \rangle \langle \sigma_2 | \hat{T} | \sigma_3 \rangle \cdots \langle \sigma_N | \hat{T} | \sigma_1 \rangle, \tag{3.146c}$$

$$= \sum_{\sigma_1} \langle \sigma_1 | \hat{T}^N | \sigma_1 \rangle, \tag{3.146d}$$

$$= \mathrm{Tr}\left[\hat{T}^N\right], \tag{3.146e}$$

where we repeatedly used the resolution of the identity and then the definition of trace.

Since the trace is independent of basis, we can compute it in the eigenbasis of $\hat{T}$. As we can see on Eq. (3.145) on the following page, $\hat{T}$ is symmetric, and hence diagonalizable. If λ_1 and λ_2 are the eigenvalues of $\hat{T}$, we'll then have

$$Z = \lambda_1^N + \lambda_2^N. \tag{3.147}$$

Suppose, without any loss of generality, that $\lambda_1 \geq \lambda_2$. Then we can write

$$Z = \lambda_1^N \left[1 + \left(\frac{\lambda_2}{\lambda_1}\right)^N\right]. \tag{3.148}$$

Diagonalizing Eq. (3.145), we find the eigenvalues

$$\lambda_{1,2} = e^{\beta J} \cosh \beta h \pm \sqrt{e^{2\beta J} \cosh^2 \beta h - 2 \sinh(2\beta J)}. \tag{3.149}$$

At $h = 0$, notice we get

$$\lambda_1 = 2 \cosh \beta J \quad \text{and} \quad \lambda = 2 \sinh \beta J. \tag{3.150}$$

Hence, we have $\lambda_1 > \lambda_2$ for $T > 0$, but $\lambda_1 = \lambda_2$ for $T = 0$. In situations with $\lambda_1 > \lambda_2$, the thermodynamic limit and Eq. (3.148) lead us to

$$g(T, h) = -k_B T \log \lambda_1, \tag{3.151}$$

for the term $\left(\frac{\lambda_2}{\lambda_1}\right)^N$ ends up vanishing in the limit.

For $T > 0$, we get from the previous expressions the Gibbs free energy per site

$$g(T, h) = -k_B T \log\left(e^{\beta J} \cosh \beta h + \sqrt{e^{2\beta J} \cosh^2 \beta h - 2 \sinh(2\beta J)}\right). \tag{3.152}$$

Differentiating this expression leads us to the magnetization

$$m(T, h) = -\left(\frac{\partial g}{\partial h}\right)_T = \frac{\sinh \beta h}{\sqrt{\sinh^2 \beta h + e^{-4\beta J}}}. \tag{3.153}$$

Notice this expression vanishes for $h = 0$ regardless of the temperature $T > 0$. Hence, we get only a paramagnetic phase, and the model fails to explain ferromagnetism. We can still get a trivial phase transition at $T = 0$, but since this is a semiclassical model, this isn't such an interesting result.

Kardar (2007a, p. 103) points out that the absence of phase transitions at finite temperature is a fairly general property of one-dimensional spin chains. Salinas (2001, p. 262) mentions there is an argument attributed to Landau to understand why that is

so, at least in the Ising model. Suppose we have an ordered phase of up spins. Suppose further that we are to flip the spins up from some point onward. The energy cost for doing this is merely $\Delta U = 2J > 0$ due to antialigning a pair of spins. Nevertheless, since there are N possible places for the flipping to occur, the entropy change is $\Delta S = k_B \log N$. Hence, the change in Gibbs free energy is

$$\Delta G = 2J - k_B T \log N, \tag{3.154}$$

which, for $T > 0$ and sufficiently large N, is negative. Hence, the spontaneous creation of new domains is favored, and ordered domains are unstable.

4 Nonequilibrium Thermodynamics

In possession of the concepts and techniques we've covered so far, we're ready to tackle nonequilibrium systems. We'll begin by considering phase transitions in nonequilibrium systems.

4.1 Phase Transitions in Nonequilibrium Systems

When dealing with phase transitions in equilibrium systems on Section 3, we were able to obtain physical information by analyzing partition functions and the free energy. Nevertheless, those techniques will not be available in out-of-equilibrium situations. Instead, we shall use the master equation introduced on Section 1.

To understand the physical significance of what we are about to study, let us first notice that phase transitions can occur even in systems that are not clearly characterized by quantities such as temperature, pressure, magnetic fields, and so on. In fact, we might be interested in studying systems that are not even describable in terms of a Hamiltonian.

For example, consider a shoal of fish. A "shoal" is a group of fish swimming close to each other for social reasons. A "school" is a type of shoal in which the fish are swimming in an orderly manner (Pitcher 1986, pp. 294–296), often with all individuals swimming in the same direction with the same speed. Notice that the fact of whether the shoal is schooling or not defines two phases in the system: a disordered phase (shoaling, but not schooling) and an ordered phase (schooling). How can we characterize these two phases in such a system?

Shoals and Schools of Fish

For simplicity, let us assume the fish can swim along only one direction, but heading towards either way. For example, they can swim only north or south. Notice this simplification is similar to the one we made on the O(N) model to obtain the Ising model. Under this hypothesis, we can let m denote the fraction of fishes swimming towards some direction, *i.e.*,

$$m = \frac{N_{\text{north}} - N_{\text{south}}}{N}. \tag{4.1}$$

While we don't have a Hamiltonian for this system, it is symmetric in exchanging north and south. Hence, it is reasonable to guess m should evolve in time according to a differential equation of the form

$$\frac{\mathrm{d}m}{\mathrm{d}t} = (a_c - a)m - bm^3 + \mathcal{O}(m^5), \tag{4.2}$$

where a is some control parameter, analogous to temperature. Symmetry forbids any even terms from occurring on the right-hand side. The terms m^5 onward can be neglected if we are interested in modelling a second-order phase transition, but they are relevant for first-order transitions.

Let us solve Eq. (4.2) on the next page considering only up to third order terms. We can write

$$\frac{\mathrm{d}m}{\mathrm{d}t} = (a_c - a)m - bm^3, \tag{4.3a}$$

$$m\frac{\mathrm{d}m}{\mathrm{d}t} = (a_c - a)m^2 - bm^4, \tag{4.3b}$$

$$\frac{1}{2}\frac{\mathrm{d}m^2}{\mathrm{d}t} = (a_c - a)m^2 - bm^4. \tag{4.3c}$$

This is now a separable differential equation. Integrating it, one finds

$$m^2(t) = \frac{a_c - a}{b\big[1 - Ae^{-2(a_c-a)t}\big]}, \tag{4.4}$$

where the integration constant A can be related to $m^2(0)$ through

$$A = 1 - \frac{a_c - a}{bm^2(0)}. \tag{4.5}$$

This prescription fails for $m^2(0) = 0$ (it leaves A undefined), in which case the solution is the trivial solution $m^2(t) = 0$.

A steady state of Eq. (4.2) will be characterized by $\frac{\mathrm{d}m}{\mathrm{d}t} = 0$. To order m^3, the stable solutions to this equation are

$$m = \begin{cases} \pm\sqrt{\frac{a_c-a}{b}}, & \text{if } a \leq a_c, \\ 0, & \text{if } a \geq a_c. \end{cases} \tag{4.6}$$

This expression is plotted on Fig. 4.1. Since the order parameter vanishes m vanishes continuously with the control parameter a, we are describing a continuous transition. Notice that schools are formed for $a < a_c$, since fish start to swim towards the same direction. For $a > a_c$, we have disorder. $a = a_c$ is a critical value (for $b > 0$) and the critical exponent β is $\beta = \frac{1}{2}$ (just as is mean field theory, being hence an example of universality). Note also that the cases of Eq. (4.6) turn out to be the late time behavior of Eq. (4.4) on the previous page for $m(0) \neq 0$ (which means they are the stable steady states).
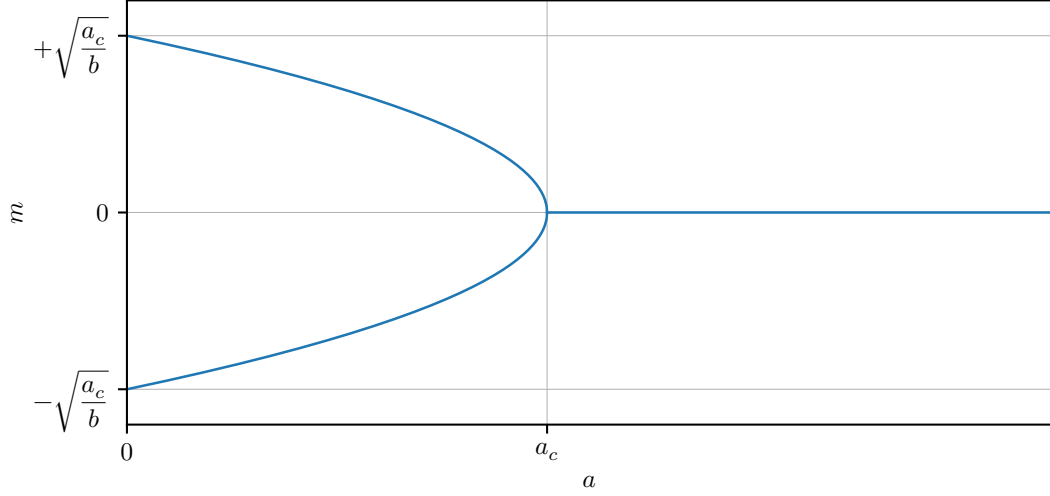
Figure 4.1: *Nonequilibrium steady states (NESS) for Eq. (4.2) on page 113 to order m^3. Notice that the order parameter vanishes continuously, and hence this is a continuous (i.e., second-order) transition.*

Notice the model of Eq. (4.2) on the preceding page does not apply specifically to fish shoals: its assumptions were as simple as a $\mathbb{Z}_2$ symmetry. Hence, a vast number of systems with completely different characteristics will present the very same critical exponent $\beta = \frac{1}{2}$, despite each one of them having different interpretations for m and different values for a, a_c, and b. This is what is meant by "universality".

Discontinuous Phase Transitions

To model a discontinuous phase transition, we can keep the m^5 term on Eq. (4.2) on the previous page and write[1]

$$\frac{\mathrm{d}m}{\mathrm{d}t} = (a_b - a)m + bm^3 - cm^5. \tag{4.7}$$

Notice the difference in sign between the second and third terms.

The steady states are given by $m = 0$ (as usual) and by the solutions to

$$(a_b - a) + bm^2 - cm^4 = 0, \tag{4.8}$$

which are

$$m^2 = \frac{b \pm \sqrt{b^2 - 4(a - a_b)c}}{2c}. \tag{4.9}$$

---

[1]Eq. (4.7) differs from Prof. Fiore's version by two signs (an overall sign and the sign of the linear term). I believe this difference is necessary in order to reproduce the discontinuous phase transition, but look out for possible mistakes on my calculations.
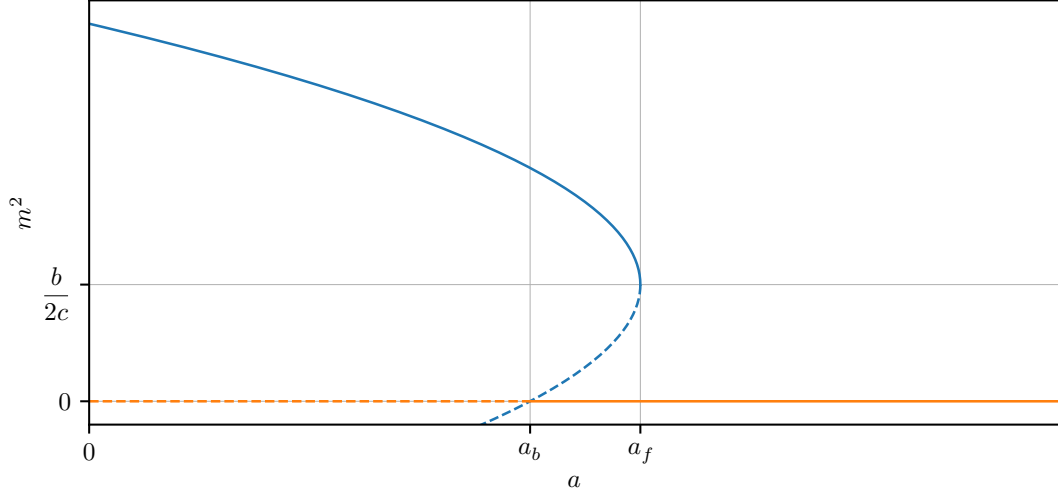
Figure 4.2: *Nonequilibrium steady states (NESS) for Eq. (4.7) on page 115. The dashed lines are unstable, while solid lines are stable. Notice that for $a_b < a < a_f$ there are two possible phases (bistability).*

Notice these solutions are only physical while they are real. Hence, at some value $a = a_f$ (f stands for ferromagnetic) we have a discontinuous jump from $m^2 = \frac{b}{2c}$ to $m^2 = 0$. a_f is given by when the argument of the square root vanishes, *i.e.*,

$$a_f = a_b + \frac{b^2}{4c}. \tag{4.10}$$

The nonequilibrium steady states (NESS) as given in terms of m^2 are shown on Fig. 4.2 on the next page. As the graph suggests, we could also find a_f by writing a as a quadratic function of m^2 and finding its peak.

To figure out the stability of each steady state, we could try to resort to the same methods we used previously (solve the differential equation for $m(t)$), but in this case we are dealing with a far more complex equation. A different possibility is to perform a linear stability analysis (Strogatz 2018, see).

For small m (say $m(t) \ll 1$), we can approximate Eq. (4.7) on the preceding page by

$$\frac{\mathrm{d}m}{\mathrm{d}t} \approx (a_b - a)m, \tag{4.11}$$

the solution to which is

$$m(t) \approx m(0)e^{(a_b - a)t}, \tag{4.12}$$

which means $m(0)$ is stable for $a > a_b$, as indicated on Fig. 4.2 on the next page.

For the other solutions, it is more convenient to write the differential equation as

$$\frac{1}{2}\frac{\mathrm{d}m^2}{\mathrm{d}t} = -cm^2(m^2 - m_-^2)(m^2 - m_+^2), \tag{4.13}$$

where $m^2_\pm$ stand for the different nontrivial solutions $m^2_+ \geq m^2_-$. Using this expression, one can get a linear approximation near $m^2_\pm$ and use it to check whether the solutions get closer or farther away from $m^2_\pm$ as time passes. One learns that m^2_+ is always unstable and m^2_- is unstable for $a_b < a < a_f$ (which is the region of interest, since it becomes negative for $a < a_b$). This is depicted on Fig. 4.2 on the following page.

Notice that the region $a_b < a < a_f$ corresponds to a bistable hysteretic branch, *i.e.*, there are two possible coexisting phases and the actual phase of the system will depend on the history of the system. If its initial state was below $m^2 = m^2_-$, it evolves towards $m = 0$, but if its initial state was above $m^2 = m^2_-$ it will evolve towards $m^2 = m^2_+$. This is a characteristic feature of discontinuous phase transitions.

Majority Vote Model

Another interesting example we can mention is the majority-vote model. It models the evolution of individual opinions in a community, where each particular individual decides their opinion based on the opinion of the majority of their neighbors. We can consider hesitant and receptive individuals—the former tend to act against the majority of their neighbors, while the latter act in favor of the majority.

The model consists in labelling opinions in favor or against some affirmation as being valued as ± 1. $+1$ if in favor, -1 if against. There is a probability $f \leq 0.5$ that an individual will not follow the opinion of the majority of its neighbors, and a probability $1 - f$ that they will (see Fig. 4.3 on the next page).

Later, we'll see that this model, as simple as it may be, presents a phase transition Cross reference similar to the ones we've found in the previous equilibrium models. However, this model has positive entropy production, and hence it is not in equilibrium.

As with the fish, we can define the fraction of individuals in favor of some opinion as

$$m = \frac{N_+ - N_-}{N},\tag{4.14}$$

in analogy with the ferromagnetic systems. For very small f, we expect a dominant opinion to predominate. Nevertheless, for $f \approx 0.5$, both opinions should be coexisting as individuals change their mind nearly randomly. Hence, we get a disordered phase. Somewhere in between, we get a phase transition at f_c. This critical value f_c will depend on aspects such as dimensionality and topology. Different neighbor arrangements might lead to different results (M. J. de Oliveira 1992, compare; Pereira and Moreira 2005).

Since now we're working out-of-equilibrium, we'll need to write transition rates and evolve the system with a master equation. Notice that the dynamics shown on Fig. 4.3 on the following page is described by the flipping transition rate

$$w_i(\sigma) = \frac{1}{2}\left[1 - (1 - 2f)\sigma_i \,\text{sign}\left(\sum_\delta \sigma_{i+\delta}\right)\right],\tag{4.15}$$

where the sum runs over neighbors of i. Hence, the individual i will change opinion with probability $w_i(\sigma)$.
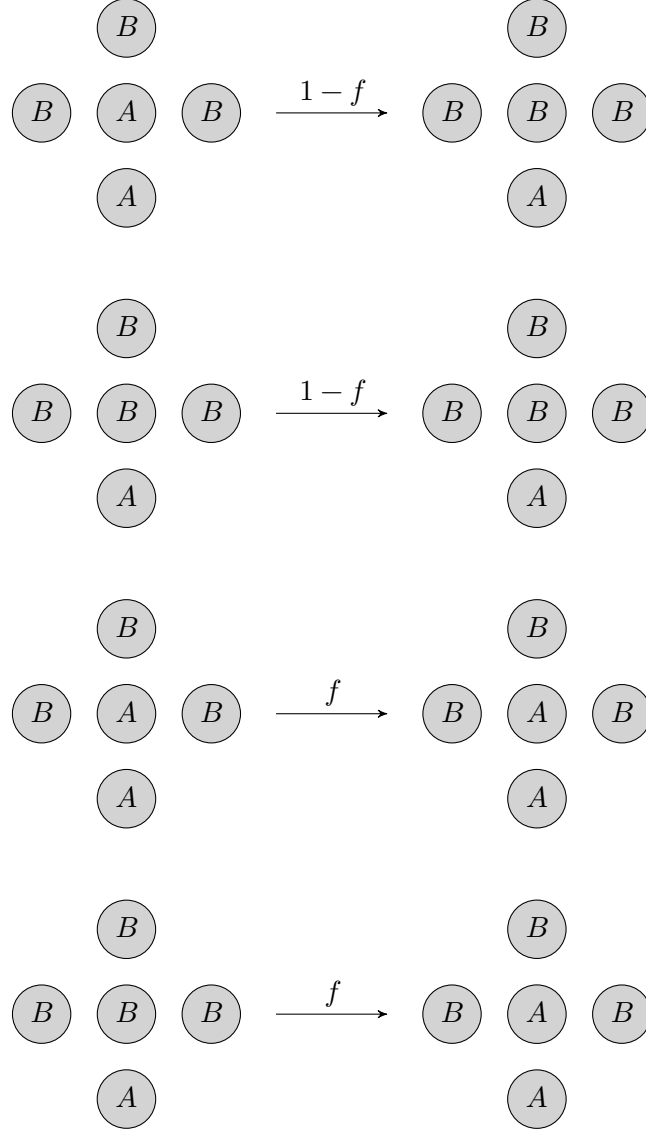
Figure 4.3: *Transition probabilities for the majority vote model. A and B denote different opinions (if one is +1, the other is −1 and vice-versa). f ≤ 0.5 is the transition probability for the individual to not follow the opinion of the majority of its neighbors.*

We then have main two ways of approaching the problem: either with mathematical approximations or numerically. Numerical simulations on a lattice have been done, *e.g.*, by M. J. de Oliveira (1992) and Pereira and Moreira (2005), and found that the critical exponents are the same as for the Ising model. A difficulty with numerical simulations is that the divergences of phase transitions only occur in the thermodynamic limit, so one won't really see divergences on a lattice. Nevertheless, one can notice a phase transition occurring by increasing the lattice's size and noticing how a peak rises more and more in the variance of the quantity whose average plays the role of order parameter[2].

Mean Field Treatment of the Majority Vote Model

A way of approximating the majority vote model is to use the master equation and ignore correlations to obtain the expression

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \frac{(1-6f)}{2}m - \frac{(1-2f)}{2}m^3, \tag{4.16}$$

which holds for a system in which each individual has four neighbors. Notice that this has the same form of Eq. (4.2) on page 113, and hence the steady state solution will have

$$m = \begin{cases} \pm\sqrt{\frac{1-6f}{1-2f}}, & \text{if } f \leq \frac{1}{6}, \\ 0, & \text{if } f \geq \frac{1}{6}. \end{cases} \tag{4.17}$$

Not only does this provide more information on the behavior of the majority vote model, but it also provides an example of a statistical system that leads to Eq. (4.2) on page 113. Let us then derive it.

Quite naturally, we'll start with the master equation,

$$\frac{\mathrm{d}P_\sigma}{\mathrm{d}t} = \sum_{\sigma' \neq \sigma} W_{\sigma\sigma'} P_{\sigma'} - W_{\sigma'\sigma} P_\sigma, \tag{4.18}$$

where $\sigma = (\sigma_1, \dots, \sigma_N)$ is the state, given as a collection of the values of all spins in the model.

We'll make a simplification known as "one-site dynamics". It consists in assuming the system can only flip a single spin at a time, so that all of the possible neighbor states σ' are of the form $\sigma' = (\sigma_1, \dots, \sigma_{i-1}, -\sigma_i, \sigma_{i+1}, \dots, \sigma_N)$. Under this simplification, we can write the transition rates $W_{\sigma'\sigma}$ as

$$W_{\sigma'\sigma} = \sum_{i=1}^N \delta_{\sigma'_1,\sigma_1} \delta_{\sigma'_2,\sigma_2} \cdots \delta_{\sigma'_{i-1},\sigma_{i-1}} \delta_{\sigma'_i,-\sigma_i} \delta_{\sigma'_{i+1},\sigma_{i+1}} \cdots \delta_{\sigma'_N,\sigma_N} w_i(\sigma), \tag{4.19}$$

Add a comment on the simplification? Maybe check references mentioned by Encinas et al. (2018, MFT results).

---

[2]Notice that the variance of such a quantity plays a role similar to that of the susceptibility in the ferromagnetic models. Furthermore, just a peak is not enough: remember the Schottky anomaly has a peak in the specific heat, but no phase transition. Another possibility to be sure of the phase transition is to compute a third quantity in addition to the mean and variance.

with $w_i(\sigma)$ being given by Eq. (4.15) on the previous page. Using this expression in the master equation leads us to

$$\frac{\mathrm{d}P_\sigma}{\mathrm{d}t} = \sum_{i=1}^{N} w_i(\sigma^i)P_{\sigma^i} - w_i(\sigma)P_\sigma, \tag{4.20}$$

where we wrote $\sigma^i \equiv (\sigma_1, \dots, \sigma_{i-1}, -\sigma_i, \sigma_{i+1}, \dots, \sigma_N)$.

Let us now suppose we want to compute the ensemble average of a quantity $f(\sigma)$, which depends only on the system's state, but not on time. Then

$$\langle f(\sigma) \rangle = \sum_\sigma f(\sigma)P_\sigma(t). \tag{4.21}$$

Hence, the master equation implies[3]

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle f(\sigma) \rangle = \sum_\sigma f(\sigma) \frac{\mathrm{d}}{\mathrm{d}t} P_\sigma(t), \tag{4.22a}$$

$$= \sum_\sigma f(\sigma) \left[\sum_{i=1}^{N} w_i(\sigma^i)P_{\sigma^i} - w_i(\sigma)P_\sigma \right]. \tag{4.22b}$$

Nevertheless, we notice that

$$\sum_\sigma f(\sigma) \left[\sum_{i=1}^{N} w_i(\sigma)P_\sigma \right] = \sum_{i=1}^{N} \sum_\sigma f(\sigma)w_i(\sigma)P_\sigma, \tag{4.23a}$$

$$= \sum_{i=1}^{N} \langle f(\sigma)w_i(\sigma) \rangle. \tag{4.23b}$$

Furthermore,

$$\sum_\sigma f(\sigma) \left[\sum_{i=1}^{N} w_i(\sigma^i)P_{\sigma^i} \right] = \sum_{i=1}^{N} \sum_\sigma f(\sigma)w_i(\sigma^i)P_{\sigma^i}, \tag{4.24a}$$

$$= \sum_{i=1}^{N} \sum_{\sigma^i} f(\sigma^i)w_i(\sigma)P_\sigma, \tag{4.24b}$$

$$= \sum_{i=1}^{N} \sum_\sigma f(\sigma^i)w_i(\sigma)P_\sigma, \tag{4.24c}$$

$$= \sum_{i=1}^{N} \langle f(\sigma^i)w_i(\sigma) \rangle, \tag{4.24d}$$

where we used the fact that, if we are summing over all configurations, it doesn't matter whether we tag them as σ or σ^i.

Through this procedure, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle f(\sigma) \rangle = \sum_{i=1}^{N} \langle (f(\sigma^i) - f(\sigma))w_i(\sigma) \rangle. \tag{4.25}$$

---

[3]Notice the allowed states of the system do not change with time. The probabilities do.

For concreteness, let us consider, say, $f(\sigma) = \sigma_k$. Then $f(\sigma^i) = -\sigma_k$ if $k = i$ and $f(\sigma^i) = +\sigma_k$ otherwise. Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \sigma_k \rangle = \sum_{i=1}^{N} \langle (f(\sigma^i) - f(\sigma))w_i(\sigma) \rangle, \tag{4.26a}$$

$$= \langle (-\sigma_k - \sigma_k)w_k(\sigma) \rangle + \sum_{i \neq k} \langle (\sigma_k - \sigma_k)w_i(\sigma) \rangle, \tag{4.26b}$$

$$= -2\langle \sigma_k w_k(\sigma) \rangle. \tag{4.26c}$$

However, for a homogeneous system, any site can be used to compute the magnetization per site, and hence we can simply write $m = \langle \sigma_k \rangle$. Therefore,

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -2\langle \sigma_k w_k(\sigma) \rangle. \tag{4.27}$$

So far our derivation has been very general. We can now specify to the majority vote model by imposing Eq. (4.15) on the preceding page. This allows us to write Eq. (4.27) as

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -\langle \sigma_k \rangle + (1 - 2f)\left\langle \sigma_k^2 \, \mathrm{sign}\left(\sum_{\delta} \sigma_{k+\delta} \right) \right\rangle, \tag{4.28a}$$

$$= -m + (1 - 2f)\left\langle \mathrm{sign}\left(\sum_{\delta} \sigma_{k+\delta} \right) \right\rangle, \tag{4.28b}$$

for $\sigma_k^2 = 1$.

We can't compute the expectation value on Eq. (4.28) exactly, which is why we'll need an approximation. The same interaction that makes the problem interesting by giving it a phase transition also prevents it from being solved exactly. Nevertheless, we can already tell that for the NESS we'll have

$$m = (1 - 2f)\left\langle \mathrm{sign}\left(\sum_{\delta} \sigma_{k+\delta} \right) \right\rangle. \tag{4.29}$$

To approximate the expression we shall write it as

$$\left\langle \mathrm{sign}\left(\sum_{\delta} \sigma_{k+\delta} \right) \right\rangle = \sum_{\sigma_{k+\delta}} \mathrm{sign}\left(\sum_{\delta} \sigma_{k+\delta} \right) P(\sigma_{k+1}, \sigma_{k+2}, \sigma_{k+3}, \sigma_{k+4}), \tag{4.30}$$

where we assumed the system to be on a regular square lattice with periodic boundary conditions (just like the one used by M. J. de Oliveira 1992), implying each site has four neighbors. If we ignore correlations, we can write

$$P(\sigma_{k+1}, \sigma_{k+2}, \sigma_{k+3}, \sigma_{k+4}) \approx P(\sigma_{k+1})P(\sigma_{k+2})P(\sigma_{k+3})P(\sigma_{k+4}), \tag{4.31}$$

from which it follows that

$$\left\langle \mathrm{sign}\left(\sum_{\delta} \sigma_{k+\delta} \right) \right\rangle \approx \sum_{\sigma_{k+\delta}} \mathrm{sign}\left(\sum_{\delta} \sigma_{k+\delta} \right) P(\sigma_{k+1})P(\sigma_{k+2})P(\sigma_{k+3})P(\sigma_{k+4}). \tag{4.32}$$

Since the system is homogeneous, the problem has been reduced to counting how many options lead to each possible value of sign $\left(\sum_\delta \sigma_{k+\delta}\right)$. This is just a combinatorics problem, and we find that

$$\left\langle \text{sign}\left(\sum_\delta \sigma_{k+\delta}\right)\right\rangle \approx \sum_{n=3}^{4}\binom{4}{n}P_+^n P_-^{4-n} - \sum_{n=3}^{4}\binom{4}{n}P_-^n P_+^{4-n}. \tag{4.33}$$

If we now recall that

$$P_+ + P_- = 1 \quad \text{and} \quad P_+ - P_- = m, \tag{4.34}$$

we'll see that we are able to express $\left\langle \text{sign}\left(\sum_\delta \sigma_{k+\delta}\right)\right\rangle$ in terms of m only. If we bring together all of the previous expressions, we'll find Eq. (4.16) on page 118.

Notice that the same computation could be used for lattices more complex than the square lattice. Furthermore, the same procedure could be applied for other transition rates and lead to similar expressions, which is a trait of universality. It is also important to notice that, as often happens with mean field theories, the predictions for the critical point and critical exponents do not match the actual values (obtained numerically by M. J. de Oliveira 1992). In particular, for systems with up-down symmetry, we always find $\beta = \frac{1}{2}$.

We can modify Eq. (4.33) for the more general case in which the site i has k neighbors instead of only 4. Then we get

$$\left\langle \text{sign}\left(\sum_\delta \sigma_{i+\delta}\right)\right\rangle \approx \sum_{n=\lceil\frac{k+1}{2}\rceil}^{k}\binom{k}{n}P_+^n P_-^{k-n} - \sum_{n=\lceil\frac{k+1}{2}\rceil}^{k}\binom{k}{n}P_-^n P_+^{k-n}, \tag{4.35}$$

where $\lceil x \rceil$ is the ceiling function, which yields the smallest integer larger than or equal to x ($\lceil 2 \rceil = 2$, $\lceil 2.1 \rceil = 3$, etc).

If we use Eq. (4.34) on the previous page on Eqs. (4.28) and (4.35) on page 120 and on the current page we find that[4]

$$\frac{dm}{dt} = -m + (1-2f)\sum_{n=\lceil\frac{k+1}{2}\rceil}^{k}\binom{k}{n}\left[\left(\frac{1+m}{2}\right)^n\left(\frac{1-m}{2}\right)^{k-n} - \left(\frac{1-m}{2}\right)^n\left(\frac{1+m}{2}\right)^{k-n}\right]. \tag{4.36}$$

For arbitrary k we'll have the general behavior with $\beta = \frac{1}{2}$, since near the critical point m is small and we'll be able to neglect higher order terms. For this model, varying k can't change the critical exponents (they are always those of mean field theory) nor the order of the transition (higher powers of m might lead to first-order transitions under particular choices of signals on each term that won't occur in here). However, the critical point f_c does depend on k. Away from the critical point, k also changes the behavior of the system, since it leads to higher powers of m in Eq. (4.36).

---

[4]I believe Eq. (4.34) on the preceding page only holds if the lattice is regular. Otherwise, it might be harder to relate the magnetization to the probabilities on each site. Hence, I believe Eq. (4.36) also only holds for regular lattices.

Dependence of the Critical Point with the Number of Neighbors

To see a bit of the dependence of the critical value f_c with k, we start by noticing from Eq. (4.17) on page 118 that $f_c = \frac{1}{6}$ for $k = 4$. For $k = 8$ one can show that

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \frac{(19 - 70f)}{16}m - \frac{35(1 - 2f)}{16}m^3 + \frac{21(1 - 2f)}{16}m^5 - \frac{5(1 - 2f)}{16}m^7. \qquad (4.37)$$

Close to the critical point ($|m| \ll 1$), we can neglect the higher order terms and find the steady state solutions

$$m \approx \begin{cases} \pm\sqrt{\frac{19 - 70f}{35(1 - 2f)}}, & \text{if } f \leq \frac{19}{70}, \\ 0, & \text{if } f \geq \frac{19}{70}, \end{cases} \qquad (4.38)$$

and hence $f_c = \frac{19}{70}$.

Notice then that f_c seems to increase with k. f_c does typically grow monotonically with k. Furthermore, notice this means f_c is not universal.

For large k, we can obtain approximate expressions for f_c as a function of k. To do so, we'll start again from Eq. (4.35). Notice it is composed of two binomial distributions. For large values of k, we can approximate these distributions by Gaussian distributions (see Salinas 2001, Sec. 1.5). Hence, we get

$$\left\langle \mathrm{sign}\left(\sum_\delta \sigma_{i+\delta}\right)\right\rangle \approx \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\frac{k}{2}}^k e^{-\frac{(\xi - kP_+)}{2\sigma^2}}\,\mathrm{d}\xi - \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\frac{k}{2}}^k e^{-\frac{(\xi - kP_-)}{2\sigma^2}}\,\mathrm{d}\xi, \qquad (4.39)$$

where the standard deviation σ is such that $\sigma^2 = kP_+P_-$.

We can express the results of such integrals in terms of the error function, defined by

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-\xi^2}\,\mathrm{d}\xi. \qquad (4.40)$$

We find that

$$\left\langle \mathrm{sign}\left(\sum_\delta \sigma_{i+\delta}\right)\right\rangle \approx \frac{1}{2}\left[\mathrm{erf}\left(\frac{k(1 - P_+)}{\sqrt{2\sigma^2}}\right) - \mathrm{erf}\left(\frac{k(\frac{1}{2} - P_+)}{\sqrt{2\sigma^2}}\right)\right.$$
$$\left. - \mathrm{erf}\left(\frac{k(1 - P_-)}{\sqrt{2\sigma^2}}\right) + \mathrm{erf}\left(\frac{k(\frac{1}{2} - P_-)}{\sqrt{2\sigma^2}}\right)\right]. \qquad (4.41)$$

Let us define $y = P_+ - \frac{1}{2}$. Since $0 \leq P_+ \leq 1$, $-\frac{1}{2} \leq y \leq +\frac{1}{2}$. Notice that

$$\frac{k(1 - P_+)}{\sqrt{2\sigma^2}} = \frac{k(1 - P_+)}{\sqrt{2kP_+P_-}}, \qquad (4.42\mathrm{a})$$

$$= \frac{k(\frac{1}{2} - y)}{\sqrt{2k(\frac{1}{2} + y)(\frac{1}{2} - y)}}, \qquad (4.42\mathrm{b})$$

$$= \sqrt{2k}\left(\frac{1}{2} - y\right)(1 + 2y^2) + \mathcal{O}(y^4). \tag{4.42c}$$

Similarly,

$$\frac{k\left(\frac{1}{2} - P_+\right)}{\sqrt{2\sigma^2}} = -\sqrt{2k}y(1 + 2y^2) + \mathcal{O}(y^4), \tag{4.43}$$

$$\frac{k(1 - P_-)}{\sqrt{2\sigma^2}} = \sqrt{2k}\left(\frac{1}{2} + y\right)(1 + 2y^2) + \mathcal{O}(y^4), \tag{4.44}$$

and

$$\frac{k\left(\frac{1}{2} - P_-\right)}{\sqrt{2\sigma^2}} = \sqrt{2k}y(1 + 2y^2) + \mathcal{O}(y^4). \tag{4.45}$$

Using these approximations, we find that

$$\left\langle \text{sign}\left(\sum_\delta \sigma_{i+\delta}\right)\right\rangle \approx \frac{1}{2}\left[\text{erf}\left(\sqrt{2k}\left(\frac{1}{2} - y\right)\right) - \text{erf}\left(\sqrt{2k}\left(\frac{1}{2} + y\right)\right) + 2\,\text{erf}\left(\sqrt{2k}y\right)\right] \tag{4.46}$$

where we used the fact that the error function is odd, and hence

$$\text{erf}\left(-\sqrt{2k}y\right) = -\text{erf}\left(\sqrt{2k}y\right). \tag{4.47}$$

Notice now that our values of interest for y are typically such that $|y| < \frac{1}{2}$, since the extrema correspond to the trivial cases with $P_\pm = 1$ and $P_\mp = 0$, leading to $m = \pm 1$. Near the critical point, which is what we want to compute, $|m| \ll 1$. Due to this, notice that

$$\text{erf}\left(\sqrt{2k}\left(\frac{1}{2} - y\right)\right) - \text{erf}\left(\sqrt{2k}\left(\frac{1}{2} + y\right)\right) = \frac{2}{\sqrt{\pi}}\int_{\sqrt{2k}\left(\frac{1}{2} + y\right)}^{\sqrt{2k}\left(\frac{1}{2} - y\right)} e^{-\xi^2}\, \mathrm{d}\xi. \tag{4.48}$$

Since we are assuming $|y| < \frac{1}{2}$, this integration region is always bounded away from the origin. Since k is large, this means we're integrating over just a piece of the Gaussian's tail, and avoiding the center, which corresponds to most of the Gaussian's area. Notice that if $|y| = \frac{1}{2}$, then the integral necessarily goes over the origin.

$\text{erf}\left(\sqrt{2k}y\right)$, on the other hand, always includes the center, because

$$\text{erf}\left(\sqrt{2k}y\right) = \frac{2}{\sqrt{\pi}}\int_0^{\sqrt{2k}y} e^{-\xi^2}\, \mathrm{d}\xi. \tag{4.49}$$

Hence, for large k and $|y| < \frac{1}{2}$, we have

$$\text{erf}\left(\sqrt{2k}\left(\frac{1}{2} - y\right)\right) - \text{erf}\left(\sqrt{2k}\left(\frac{1}{2} + y\right)\right) + 2\,\text{erf}\left(\sqrt{2k}y\right) \approx 2\,\text{erf}\left(\sqrt{2k}y\right). \tag{4.50}$$

One can also check this graphically by plotting both sides of the equation as functions of k and varying y.

The upshot of this discussion is that, for large k and $|y| < \frac{1}{2}$,

$$\left\langle \operatorname{sign}\left(\sum_\delta \sigma_{i+\delta}\right) \right\rangle \approx \operatorname{erf}\left(\sqrt{2k}\,y\right). \tag{4.51}$$

If we recover the definition of $y = P_+ - \frac{1}{2}$ and remember Eqs. (4.28) and (4.34) on page 120 and on page 121, then we find that

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -m + (1 - 2f)\operatorname{erf}\left(\sqrt{\frac{k}{2}}m\right). \tag{4.52}$$

Expanding the error function in a Taylor series up to order m^3 allows one to compute the critical value f_c and find

$$f_c = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\pi}{2k}}. \tag{4.53}$$

Notice f_c grows with k for large k.

References

The course's official bibliography is comprised of the books by Callen (1985), Pathria and Beale (2022), Salinas (2001), Sander (2013), and Yeomans (1992). Papers by Crooks (1998, 1999), Noa et al. (2019), M. M. de Oliveira, Luz, and Fiore (2018), Proesmans and Fiore (2019), Tomé and M. J. de Oliveira (2015a), and Van den Broeck and Esposito (2015) were suggested as well. The other references were added by me. References with open access sources are indicated by a ᔕ next to them, which is in fact a link to somewhere in which the reference can be retrieved legally and for free.

Abramowitz, Milton and Irene A. Stegun (1972). *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Washington: National Bureau of Standards (cit. on p. 33).

Altland, Alexander and Ben D. Simons (2010). *Condensed Matter Field Theory*. Cambridge: Cambridge University Press (cit. on p. 64).

Arfken, George B., Hans J. Weber, and Frank E. Harris (2013). *Mathematical Methods for Physicists: A Comprehensive Guide*. Oxford: Academic Press (cit. on pp. 33, 62).

Blundell, Stephen and Katherine M. Blundell (2010). *Concepts in Thermal Physics*. 2nd ed. Oxford: Oxford University Press (cit. on pp. 69, 70, 72).

Callen, Herbert B. (1985). *Thermodynamics and an Introduction to Thermostatistics*. 2nd ed. New York: Wiley (cit. on pp. 44, 78, 107, 125).

Caticha, Nestor (Feb. 19, 2019). *Mecânica Estatística e Informação*. São Paulo (cit. on p. 26).

Chabay, Ruth W. and Bruce A. Sherwood (2015). *Matter & Interactions*. Vol. 1: *Modern Mechanics*. 4th ed. Hoboken: Wiley (cit. on p. 31).

christian (Dec. 16, 2016). *The Maxwell Construction*. Learning Scientific Programming with Python. url: https://scipython.com/blog/the-maxwell-construction/ (visited on 10/06/2022) (cit. on pp. 40, 72, 77).

Crooks, Gavin E. (1998). "Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems". *Journal of Statistical Physics* **90**, pp. 1481–1487. doi: 10.1023/A:1023208217925 (cit. on pp. 5, 125).

——— (Sept. 1, 1999). "Entropy Production Fluctuation Theorem and the Nonequilibrium Work Relation for Free Energy Differences". *Physical Review E* **60**.3, pp. 2721–2726. doi: 10.1103/PhysRevE.60.2721. arXiv: cond-mat/9901352 (cit. on pp. 5, 11, 12, 125).

Encinas, Jesus M. et al. (Dec. 2018). "Fundamental Ingredients for Discontinuous Phase Transitions in the Inertial Majority Vote Model". *Scientific Reports* **8**.1, 9338. doi: 10.1038/s41598-018-27240-4 (cit. on p. 119).

Erdélyi, Arthur (1956). *Asymptotic Expansions*. New York, NY: Dover (cit. on p. 106).

Fermi, Enrico (1956). *Thermodynamics*. New York: Dover (cit. on pp. 12, 18, 70).

Griffiths, David J. (2005). *Introduction to Quantum Mechanics*. 2nd ed. Cambridge: Cambridge University Press (cit. on p. 47).

——— (2017). *Introduction to Electrodynamics*. Cambridge: Cambridge University Press. doi: 10.1017/9781108333511 (cit. on p. 87).

Guggenheim, E. A. (July 1945). "The Principle of Corresponding States". *The Journal of Chemical Physics* **13**.7, pp. 253–261. doi: 10.1063/1.1724033 (cit. on p. 86).

Jaeger, Gregg (May 1, 1998). "The Ehrenfest Classification of Phase Transitions: Introduction and Evolution". *Archive for History of Exact Sciences* **53**.1, pp. 51–81. doi: 10.1007/s004070050021 (cit. on p. 70).

Jarzynski, C. (Apr. 7, 1997). "Nonequilibrium Equality for Free Energy Differences". *Physical Review Letters* **78**.14, pp. 2690–2693. doi: 10.1103/PhysRevLett.78.2690. arXiv: cond-mat/9610209 (cit. on p. 12).

Kardar, Mehran (2007a). *Statistical Physics of Fields*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511815881 (cit. on pp. 69, 71, 72, 83, 84, 86, 102, 110, 112).

——— (2007b). *Statistical Physics of Particles*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511815898 (cit. on pp. 15, 21, 26, 33, 39, 42, 47, 56, 64, 65, 67, 68, 76, 77, 87, 107).

Lancaster, Tom and Stephen Blundell (2014). *Quantum Field Theory for the Gifted Amateur*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199699322.001.0001 (cit. on p. 64).

Landau, Lev Davidovič and Evgenij M. Lifshitz (1980). *Course of Theoretical Physics*. Vol. 5: *Statistical Physics, Part 1*. 3rd ed. Oxford: Pergamon Press (cit. on pp. 97–99).

Noa, C. E. Fernández et al. (July 3, 2019). "Entropy Production as a Tool for Characterizing Nonequilibrium Phase Transitions". *Physical Review E* **100**.1, 012104. doi: 10.1103/PhysRevE.100.012104. arXiv: 1811.06310 [cond-mat.stat-mech] (cit. on p. 125).

Oliveira, Marcelo M. de, M. G. E. da Luz, and Carlos E. Fiore (June 8, 2018). "Finite-Size Scaling for Discontinuous Nonequilibrium Phase Transitions". *Physical Review E* **97**.6, 060101. doi: 10.1103/PhysRevE.97.060101. arXiv: 1804.00467 [cond-mat.stat-mech] (cit. on p. 125).

Oliveira, Mário José de (Jan. 1992). "Isotropic Majority-Vote Model on a Square Lattice". *Journal of Statistical Physics* **66**, pp. 273–281. doi: 10.1007/BF01060069 (cit. on pp. 118, 121).

——— (2013). *Equilibrium Thermodynamics*. Graduate Texts in Physics. Berlin: Springer. doi: 10.1007/978-3-642-36549-2 (cit. on p. 69).

Particle Data Group et al. (Aug. 8, 2022). "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* **2022**.8, 083C01. doi: 10.1093/ptep/ptac097 (cit. on p. 64).

Pathria, Raj K. and Paul D. Beale (2022). *Statistical Mechanics*. 4th ed. London: Academic Press (cit. on pp. 15, 21, 22, 26, 28, 30, 31, 33, 39, 42, 47, 56, 61, 65, 67, 68, 86, 94, 97, 99, 110, 125).

Peliti, Luca (2011). *Statistical Mechanics in a Nutshell*. Trans. by Mark Epstein. In a Nutshell. Princeton, NJ: Princeton University Press (cit. on p. 99). Trans. of *Appunti di Meccanica Statistica*. Torino: Bollati Boringhieri, 2003.

Pereira, Luiz F. C. and F. G. Brady Moreira (Jan. 18, 2005). "Majority-Vote Model on Random Graphs". *Physical Review E* **71**.1, 016123. doi: 10.1103/PhysRevE.71.016123. arXiv: cond-mat/0408706 (cit. on p. 118).

Peters, Hjalmar (Jan. 1, 2014). "Demonstration and Resolution of the Gibbs Paradox of the First Kind". *European Journal of Physics* **35**.1, 015023. doi: 10.1088/0143-0807/35/1/015023. arXiv: 1306.4638 [cond-mat.stat-mech] (cit. on p. 26).

Pitcher, Tony J. (1986). "Functions of Shoaling Behaviour in Teleosts". In: *The Behaviour of Teleost Fishes*. Ed. by Tony J. Pitcher. Boston, MA: Springer, pp. 294–337. doi: 10.1007/978-1-4684-8261-4_12 (cit. on p. 113).

Proesmans, Karel and Carlos E. Fiore (Aug. 28, 2019). "General Linear Thermodynamics for Periodically Driven Systems with Multiple Reservoirs". *Physical Review E* **100**.2, 022141. doi: 10.1103/PhysRevE.100.022141. arXiv: 1906.10752 [cond-mat.stat-mech] (cit. on pp. 10, 125).

Reichl, Linda E. (2016). *A Modern Course in Statistical Physics*. 4th ed. Weinheim: Wiley-VCH (cit. on pp. 15, 42, 82).

Reif, Frederick (2009). *Fundamentals of Statistical and Thermal Physics*. Long Grove, IL: Waveland Press (cit. on p. 17). Repr. of *Fundamentals of Statistical and Thermal Physics*. McGraw-Hill Series in Fundamentals of Physics: An Undergraduate Textbook Program. New York: McGraw-Hill, 1965.

Sakurai, Jun John and Jim Napolitano (2017). *Modern Quantum Mechanics*. 2nd ed. Cambridge: Cambridge University Press. doi: 10.1017/9781108499996 (cit. on p. 47).

Salinas, Silvio R. A. (2001). *Introduction to Statistical Physics*. Graduate Texts in Contemporary Physics. New York: Springer. doi: 10.1007/978-1-4757-3508-6 (cit. on pp. 7, 15, 22, 28, 30–33, 39, 40, 42, 43, 45–47, 56, 59, 63–66, 86, 87, 94, 97–100, 110, 112, 123, 125). Trans. of *Introdução à Física Estatística*. Acadêmica 9. São Paulo: EdUSP, 1997.

Sander, Leonard M. (July 27, 2013). *Equilibrium Statistical Physics: With Computer Simulations in Python.* United States: Createspace Independent Publishing Platform (cit. on p. 125).

Schnakenberg, J. (Oct. 1, 1976). "Network Theory of Microscopic and Macroscopic Behavior of Master Equation Systems". *Reviews of Modern Physics* **48**.4, pp. 571–585. doi: 10.1103/RevModPhys.48.571 (cit. on p. 9).

Shankar, Ramamurti (1994). *Principles of Quantum Mechanics.* 2nd ed. New York: Springer. doi: 10.1007/978-1-4757-0576-8 (cit. on p. 47).

Stierstadt, K. et al. (1990). *Physics Data: Experimental Values of Critical Exponents and Amplitude Ratios of Magnetic Phase Transitions.* Karlsruhe, Germany: Fachinformationszentrum (cit. on p. 94).

Streater, R. F. and A. S. Wightman (2000). *PCT, Spin and Statistics, and All That.* 1st pbk. print., with rev. pref. and corrections. Princeton Landmarks in Physics. Princeton, N.J: Princeton University Press. 207 pp. (cit. on p. 47).

Strogatz, Steven H. (2018). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering.* Boca Raton: CRC Press (cit. on p. 115).

Tomé, Tânia and Mário José de Oliveira (Apr. 29, 2015a). "Stochastic Approach to Equilibrium and Nonequilibrium Thermodynamics". *Physical Review E* **91**.4, 042140. doi: 10.1103/PhysRevE.91.042140. arXiv: 1503.04342 [cond-mat.stat-mech] (cit. on pp. 6, 125).

Tomé, Tânia and Mário José de Oliveira (2015b). *Stochastic Dynamics and Irreversibility.* Graduate Texts in Physics. Cham: Springer. doi: 10.1007/978-3-319-11770-6 (cit. on pp. 6, 7).

Van den Broeck, C. and M. Esposito (Jan. 2015). "Ensemble and Trajectory Thermodynamics: A Brief Introduction". *Physica A: Statistical Mechanics and its Applications* **418**, pp. 6–16. doi: 10.1016/j.physa.2014.04.035. arXiv: 1403.1777 [cond-mat.stat-mech] (cit. on pp. 6, 125).

Wald, Robert M. (2022). *Advanced Classical Electromagnetism.* Princeton: Princeton University Press (cit. on p. 87).

Wannier, Gregory H. (1987). *Statistical Physics.* New York: Dover (cit. on pp. 30, 31). Repr. of *Statistical Physics.* New York: Wiley, 1966.

Weinberg, Steven (2015). *Lectures on Quantum Mechanics.* 2nd ed. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781316276105 (cit. on p. 47).

——— (2021). *Foundations of Modern Physics.* Cambridge: Cambridge University Press. doi: 10.1017/9781108894845 (cit. on p. 47).

Wreszinski, Walter F. (2018). *Termodinâmica.* Acadêmica 50. São Paulo: EdUSP (cit. on p. 78).

Yeomans, J. M. (1992). *Statistical Mechanics of Phase Transitions.* Oxford: Clarendon Press (cit. on p. 125).

Zangwill, Andrew (2013). *Modern Electrodynamics.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139034777 (cit. on p. 87).

Zinn-Justin, Jean (July 5, 2007). *Phase Transitions and Renormalization Group.* Oxford University Press. doi: 10.1093/acprof:oso/9780199227198.001.0001 (cit. on p. 86).