# MATH327: Statistical Physics

**David Schaich** **Spring 2021**

## LECTURE NOTES

Last modified 7 May 2021

# Module information and logistics

## Coordinator

Dr David Schaich
Mathematics Building Room 124 (Theoretical Physics Wing)
david.schaich@liverpool.ac.uk
www.davidschaich.net

## Weekly schedule

**Synchronous sessions** are timetabled at 9:00 on Mondays and 13:00 on Thursdays.

**Office hours** will take place immediately afterwards, at 10:00 on Mondays and 14:00 on Thursdays. If these times do not work with your schedule, you can also make an appointment through calendly.com/daschaich.

## Delivery plans and module resources

Redesigning this module for remote delivery requires some experimentation and improvisation, so you should feel free to make requests and suggestions regarding how the module may best be delivered. In order to ensure that the amount of material remains manageable, we will cover the same content as last year's edition of the module, rewriting it for remote learning in place of in-person lectures and computer labs. All resources for the module will be collected at its Canvas site,

liverpool.instructure.com/courses/19478

I will aim to minimize the number of announcements sent to you through Canvas. You have the option to receive email notification of these announcements. If you choose to turn off these email notifications, you should make sure to check Canvas regularly for announcements. I will assume announcements have been read within 24 hours.

### Lecture notes

We will use the *Notes-driven Asynchronous Delivery* framework designed by the Department of Mathematical Sciences, in which the primary learning materials are the lecture notes you are currently reading. As you read further, you will encounter **gaps** in the notes where you will be able to check your understanding by completing some exercises that are intended to be bite-sized. Feel free to complete these exercises in whatever format you find most convenient, for instance on separate pieces of paper. While the size of the gap is intended to indicate the rough scale of the exercise, your work may or may not fit depending on your handwriting and solution strategy.

You will also see many numbered equations in the lecture notes. In general, equations are numbered so that they can easily be referred to in subsequent discussions, derivations or examples. Therefore a number is not necessarily an indication of importance. Some particularly important results, definitions and concepts will be highlighted by coloured boxes. You should ensure you understand the key learning objectives in these coloured boxes. There is no need to memorize any equations.

### Synchronous sessions

These gaps will guide our twice-weekly synchronous sessions using Zoom on Monday mornings and Thursday afternoons. In addition to discussing any general questions you may have, we will aim to complement the asynchronous material by focusing our attention on any gaps that proved challenging to fill. In order for these synchronous sessions to remain in **active learning** mode, you may be asked to describe your own approach to any given gap. Thinking about them in advance will make this easier. If you get stuck during the synchronous session, I will aim to ask leading questions rather than lecturing.

### Video recordings and asynchronous discussions

To supplement these lecture notes, the Department's delivery framework also involves regular video recordings, expected to add up to around 20 minutes per week. For each week I plan to provide both an introductory video presenting big-picture context and motivation for the week's topic, followed by a concluding summary of the key results to take away as learning outcomes. ~~Since **synchronous sessions will not be recorded**, to encourage your participation,~~ By popular request, **synchronous sessions will be recorded**. I also expect to add supplemental videos addressing any particularly tricky issues that arise in the course of those synchronous sessions (for example, working through gaps that prove challenging). Supplemental exercises may also be released to respond to particular challenges we encounter. The Canvas site also provides a *Discussion* utility where questions can be asked, discussed and answered asynchronously.

### Term schedule and organization

The planned schedule for the module is summarized on the first page of these notes. If necessary to respond to unanticipated challenges, there should be some flexibility to adapt this schedule. The amount and difficulty of material necessarily fluctuate from week to week, and different material may prove more or less difficult to different students. Sticking with the same content as last year should ensure that these fluctuations average out over the course of the term.

This content is organized around the concept of *statistical ensembles* introduced by J. Willard Gibbs in the early 1900s. In essence, a statistical ensemble is a mathematical framework for concisely describing the properties of (idealized) physical systems with a large number of independent objects, subject to certain

constraints. After a week studying the mathematical foundations underlying these frameworks, we will learn about the *micro-canonical ensemble* in week 2 and the *canonical ensemble* in week 3. We will spend weeks 4 and 5 on applications of the canonical ensemble, first considering ideal gases and then investigating thermodynamic cycles. In week 6 we will see our third and final statistical ensemble, the *grand-canonical ensemble*, which we will apply to various types of quantum gases in weeks 7–9. Those first nine weeks will all consider systems whose constituent objects do not interact with each other. We will explore effects of interactions in week 10, which opens up a much broader landscape of applications that we will survey for the remainder of the term (along with some overall synthesis and review to help prepare for the final assessment).

The asynchronous materials for each week will be available through Canvas at least a week in advance. For example, the notes and videos for week 1 (8–15 February) will be available on or before Sunday, 31 January. In case subsequent corrections or additions to the notes are needed, the date of last modification will be included at the bottom of each page.

### Version control and co-creation

An easy way to check any corrections or additions that may need to be made to these notes is provided by the public GitHub repository where their LaTeX source is maintained under version control:

github.com/daschaich/MATH327_2021

In addition, you are also welcome to use GitHub to report issues with the notes[1] and create pull requests to address them. Such co-creation is entirely optional. If you are interested in learning more about version control with `git`, the Software Carpentry project provides free resources at

swcarpentry.github.io/git-novice/

### Expected background

You may have seen quantum statistics and quantum gases listed in the module schedule on the first page, so I will emphasize that no prior knowledge of quantum mechanics is assumed. Any necessary information on this topic will be provided. It is also not necessary to be familiar with statistical analysis or combinatorics. In those contexts I just anticipate that you have previously seen the standard deviation and the binomial coefficient

$$\binom{N}{k} = \frac{N!}{k!\,(N-k)!} = \binom{N}{N-k}$$

that counts the number of possible ways to choose $k$ objects out of a set of $N \geq k$ total objects. I also anticipate that you have previously seen gaussian integrals,

$$\int_{-\infty}^{\infty} e^{-a(x+b)^2}\,dx = \sqrt{\frac{\pi}{a}}.$$

[1] Questions about the mathematical content should go to the Canvas discussions.

## Assessment and academic integrity

The assessment workload has also been kept the same as last year, though weights have been distributed more evenly across the term to accommodate current circumstances and avoid a high-stakes final assessment. Deadlines for in-term assignments have been coordinated within the Department to minimize pile-up. MATH327 has been allocated the following deadlines at 23:59 on **Tuesdays**:

**20%** A computer-based project divided into two equally weighted parts, the first due **Tuesday, 2 March** and the second due **Tuesday, 13 April**

**15%** A homework assignment available by Monday, 22 February and due **Tuesday, 16 March**

**15%** A homework assignment available by Monday, 19 April and due **Tuesday, 27 April**

**50%** A final assessment to be centrally scheduled within the period 17 May through 4 June

Since the second part of the computer project is due shortly after the term break, feedback for the first part will be provided before the break, and office hours will continue during the break in case any assistance is needed with this project.

Following the University's Code of Practice on Assessment, late submissions completed within five days (120 hours) after the submission deadline will have 5% of the total marks deducted for each 24-hour post-deadline period. Submissions more than five days late will be awarded zero marks, though I will still endeavour to provide feedback on them. I will aim to return feedback and solutions 7–10 days after the deadline.

Because the computer-based project will be done remotely, rather than in a computer lab on campus, you are free to complete it using the programming language of your choice. Instruction on the necessary programming concepts will be provided using Python, the free programming language recommended for the project. If you have difficulty setting up Python on your device, you can run it for free online at repl.it. (Make sure to keep a copy of the code to submit for assessment!) Alternative languages could include C, R, or even MATLAB (through the University's site license). Maple may struggle to handle parts of the project.

I encourage you to discuss the in-term assignments with each other, since discussing and debating concepts and procedures is a very effective way to learn. The examination must be done on your own, and your submissions for all assignments must be your own work representing your own understanding. Clear and neat presentations of your workings and the logic behind them will contribute to your mark. It is unacceptable to copy solutions in part or in whole from other students (current or prior) or from other sources (commercial or otherwise). Should you make use of resources beyond the module materials, these must be explicitly referenced in your work.

By now you should have successfully passed the Academic Integrity Tutorial and Quiz to affirm that you have read and understood the Academic Integrity Policy detailed in Appendix L of the Code of Practice on Assessment. If you have any questions about what is or is not acceptable under this policy, please ask me or Departmental Assessment Officer Kamila Zychaluk. In all cases, the spirit of the Academic Integrity Policy should take precedence over legalistic convolutions of the text.

## Additional resources

The lists of further reading below provide some optional additional resources that may be helpful. Hyperlinks lead either to the resource itself or to the corresponding record page in our library.

**Books and lecture notes at roughly the level of this module:**

1. David Tong, *Lectures on Statistical Physics* (2012), www.damtp.cam.ac.uk/user/tong/statphys.html
2. Daniel V. Schroeder, *An Introduction to Thermal Physics* (first edition, 2000)
3. C. Kittel and H. Kroemer, *Thermal Physics* (second edition, 1980)
4. F. Reif, *Fundamentals of Statistical and Thermal Physics* (first edition, 1965)

**More advanced and more specialized books**, which may be useful to consult concerning specific questions or topics:

5. R. K. Pathria, *Statistical Mechanics* (second edition, 1996)
6. Sidney Redner, *A Guide to First-Passage Processes* (first edition, 2007)
7. Pavel L. Krapivsky, Sidney Redner and Eli Ben-Naim, *A Kinetic View of Statistical Physics* (first edition, 2010)
8. Kerson Huang, *Statistical Mechanics* (second edition, 1987)
9. Weinan E, Tiejun Li and Eric Vanden-Eijnden, *Applied Stochastic Analysis* (first edition, 2019)
10. Michael Plischke and Birger Bergersen, *Equilibrium Statistical Physics* (third edition, 2005)
11. L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part 1* (third edition, 1980)

**A general book about learning**, which (among other strategies) emphasizes the value of retrieval practice compared to re-reading lecture notes or re-watching videos:

12. Peter C. Brown, Henry L. Roediger III and Mark A. McDaniel, *Make it Stick: The Science of Successful Learning* (first edition, 2014)
    A short summary video is also available

**Maple and MATLAB resources:**

13. Ian Thompson, *Understanding Maple* (ebook edition, 2017)
    Related videos: *Running Maple for the first time* (stream.liv.ac.uk/4k67bdzt)
    *Configuring Maple* (stream.liv.ac.uk/7pjge23a)

14. Stormy Attaway, *MATLAB: A Practical Introduction to Programming and Problem Solving* (third edition, 2013)

15. B. Barnes and G. R. Fulford, *Mathematical Modelling with Case Studies: Using Maple and MATLAB* (third edition, 2014)

Finally, there is a vast constellation of purely online resources, such as Wikipedia. These are often fine places to *start* learning about a subject, but may be terrible places to *stop*.

# Week 1: Central limit theorem and diffusion

## Introductory remarks: What is Statistical Physics?

Mathematical sciences such as physics aim to determine the laws of nature and understand how these govern experimental observations—both in everyday circumstances and under extreme conditions. This mathematical understanding is typically guided by reproducing a set of observations, with the resulting framework then used to make predictions for other "observables".

Over the past few centuries this process has been tremendously successful, with theoretical physics accurately predicting experimental and observational results from sub-atomic through to extra-galactic scales. Modern physics labs can create a vacuum better than in outer space and the coldest temperatures in the known universe, as well as going to the other extreme to reach temperatures of millions of degrees and pressures millions of times atmospheric pressure at sea level. Amazingly, many aspects of these realms of physics can be theoretically described by mathematics developed centuries ago.[2]

The domain of **statistical physics** is one in which simple mathematical principles enable amazing predictive capabilities. Initially developed in the nineteenth century, statistical physics remains a core component of modern physics, and will retain this position in years to come. The foundations of statistical physics lie in the use of probability theory to mathematically describe experimental observations and corresponding laws of nature that involve stochastic randomness rather than being perfectly predictable.

The lack of perfect predictability in statistical physics is a matter of practicality rather than one of principle. It results from working with a large number of degrees of freedom (i.e., a large number of independent objects such as particles or balls). For example, Avogadro's number $N_A \approx 6.022 \times 10^{23}$ is the large number of molecules in everyday amounts of familiar substances—about 18 grams of water or about 22 litres of air at sea-level atmospheric pressure ($\approx 101$ kPa). Specifying the positions and velocities of $\sim 10^{23}$ objects would require far more information than could be stored even in the memory of the biggest existing supercomputers. Statistical physics instead produces simple mathematical descriptions of large-scale properties such as temperature, pressure and diffusion, which are generally of such outstanding quality that the underlying 'randomness' is effectively undetectable.

Historically, the difficulty detecting the stochastic processes underlying such *thermodynamic* properties made it challenging to convince skeptics that atoms and molecules really exist. Ludwig Boltzmann, a prominent early developer of statistical physics, endured a constant struggle to defend his ideas, which likely contributed to his deteriorating mental health and eventual suicide in 1906. A crucial advance to convincingly establish the existence of atoms was Albert Ein-

[2] Eugene Wigner's famous article, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" (1960), and subsequent work in the philosophy of physics, elaborates on why this may be considered 'amazing'. These lecture notes will not comment extensively on philosophy.

stein's use of statistical physics to explain the observed "Brownian motion" of particles suspended in fluids—this work was part of Einstein's "miracle year" in 1905, along with special relativity and early contributions to quantum physics. More modern applications of statistical physics include explaining why stars don't collapse under the 'weight' of their own gravity, and identifying effects of dark matter in temperature fluctuations observable in the *cosmic microwave background* lingering from the early years of the universe.

For this week we will focus on some of the foundational mathematics that will underlie our later development and application of statistical ensembles. Looking back to Boltzmann's times, we can consider the following question one of his opponents might have asked: *If the pressure of a gas in a container results from molecules stochastically colliding with the walls of that container and pushing them out, then how can the pressure be so stable and reproducible, rather than itself fluctuating stochastically?* The mathematical answer lies in the **law of large numbers** and the **central limit theorem**, which this week we will learn and apply to the physics of diffusion in one dimension.

## 1.1 Probability foundations

Let's begin by placing some familiar everyday concepts into a more formal mathematical framework through the following definitions:

- A (random) **experiment** $\mathcal{E}$ involves setting up, manipulating and/or observing some (physical or hypothetical) system with some element of randomness. Flipping a coin is a simple random experiment. In the context of the statistical ensembles that will be the focus of this module, the typical experiment will be simply allowing a collection of particles to evolve in time, subject to certain constraints.

- Each time an experiment is performed, the world comes out in some **state** $\omega$. The definition of the experiment and the state must include all objects of interest, and may include more besides. When flipping a coin, for example, the full state could contain information not only about the final orientation of the coin, but also about its position—did it land on the floor or on a cat?

- The **set of all states** $\Omega$ collects all possible states $\omega$ that the given experiment $\mathcal{E}$ can produce, and is therefore intricately tied to $\mathcal{E}$ itself.

- We are generally not interested in all aspects of the full state $\omega$. For example, we won't care where a flipped coin lands. Instead we're typically only interested in whether it lands heads up or tails up—and we may want to set aside any state that doesn't cleanly map on to those options. The **measurement** $X(\omega)$ extracts and quantifies this information, acting as a function that maps the state $\omega$ to a number that we can mathematically manipulate. If we repeat a fixed experiment $\mathcal{E}$ many times and carry out the measurement $X$ on each resulting state $\omega$, we will obtain a sequence of numbers $X(\omega)$ that behave as a *random variable*.

- Acting with the measurement $X$ on all of the possible states in the set $\Omega$ defines the **set of all outcomes** (or **outcome space**) $A$:

$$X : \Omega \to A.$$

  That is, $A$ collects all possible measurement results that the given experiment $\mathcal{E}$ and measurement $X$ can produce. $A$ can be finite, countably infinite, or uncountably infinite (i.e., continuous).

- Finally, defining an **event** to be any subset of the set of all outcomes $A$, we further group these subsets together to define a **set of events** $\mathcal{F}$.

Let's consider some **examples** to clarify these definitions. With an experiment of rolling a six-sided die and measuring the number ($1$–$6$) that comes out on top, what is the set of all outcomes $A$, and what additional information could be present in the set of all states $\Omega$?

What is the outcome space $A$ if we toss a coin four times and measure whether it lands heads up ($H$) or tails up ($T$)?

What information would characterize a state $\omega$ for a gas of $6\times10^{23}$ argon atoms in a container?

We can generalize the concept of measurement by introducing a unique number as a *label* to characterize each state $\omega$ in the set $\Omega$. This would provide a label function $L(\omega)$ as a random variable. Our condition of uniqueness makes $L(\omega)$ isomorphic, so that the label can be used interchangeably with the full state,

$$\omega \longleftrightarrow L(\omega).$$

While the measurements $X(\omega)$ we consider will generally not produce a unique number for each $\omega$, we will design them (as best we can) precisely to remove irrelevant information that doesn't interest us. Ignoring that irrelevant information leaves us free to interchange the set of outcomes $A$ for the set of states $\Omega$, which we will do from now on. (Some textbooks may never distinguish between $A$ vs $\Omega$ in the first place, though this can be a source of confusion.)

We are now prepared for the final foundational definition in this section, the **probability** $P$ that an event in the set $\mathcal{F}$ has of occurring. Mathematically, $P$ is a *measure function*,

$$P : \mathcal{F} \to [0, 1],$$

which must satisfy the following two requirements:

1. The probability of a countable union of mutually exclusive events must equal the countable sum of the probabilities of each of these events.
2. The probability of the outcome space ($\mathcal{F} = A$) must equal $1$ (even if $A$ is uncountable). This simply means that the experiment $\mathcal{E}$ must produce an outcome. If no outcome were produced, it would not make sense to say that the experiment had occurred.

Combining the outcome space, event space and probability measure gives us a *probability space* $(A, \mathcal{F}, P)$.

For **example**, consider an experiment that can only produce $N$ possible states, so that

$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_N\}.$$

If two states are identical, $\omega_i = \omega_j$, they must produce the same measurement outcomes $X(\omega_i) = X(\omega_j)$, which implies the contra-positive

$$X(\omega_i) \neq X(\omega_j) \qquad \Longrightarrow \qquad \omega_i \neq \omega_j.$$

On the other hand, as described above, it is possible to have $X(\omega_i) = X(\omega_j)$ even when $\omega_i \neq \omega_j$. This means that the size $n$ of the outcome space $A$ may be smaller than the size of $\Omega$, $n \leq N$. We can write

$$A = \{X_1, X_2, \cdots, X_n\},$$

where each $X_\alpha$ is distinct and its index does *not* necessarily correspond to that on $\omega_i$. We can take the individual $X_\alpha$ themselves to be the events we're interested in, with an overall event space

$$\mathcal{F} = \{X_1, X_2, \cdots, X_n\} = A.$$

These events are all mutually exclusive by construction, so if we assign them probabilities

$$P(X_\alpha) \equiv p_\alpha \qquad \text{for } \alpha = 1, \cdots, n,$$

then the above requirements on probabilities demand that for any $\alpha \neq \beta$ we have

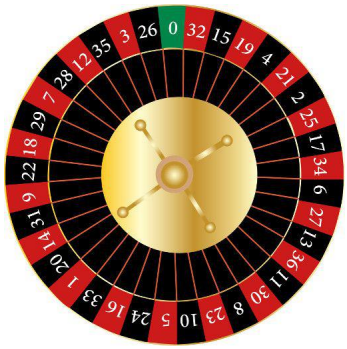
$$P(X_\alpha \text{ or } X_\beta) = p_\alpha + p_\beta$$
$$P(X_1 \text{ or } X_2 \text{ or } \cdots \text{ or } X_n) = \sum_{\alpha=1}^{n} p_\alpha = 1.$$

Similarly choosing an event space $\mathcal{F} = A$ for the six-sided die considered in an earlier gap, what are the probabilities $p_1$ through $p_6$ that result from assuming the die is *fair* (meaning that each outcome is equally probable)?

Again taking $\mathcal{F} = A$ for the case of tossing a coin four times, what are the probabilities $p_\alpha$ that result from assuming the coin is fair? If we instead consider the event space

$$\mathcal{F} = \{\text{equal number of } H \text{ and } T,\ \text{different numbers of } H \text{ and } T\}\,,$$

what are the probabilities $p_{\text{equal}}$ and $p_{\text{diff}}$ for the two events in this $\mathcal{F}$?




The standard European roulette wheel shown to the left (source) has 37 pockets labelled "0" through "36". 18 of these pockets are coloured red, 18 are coloured black and 1 (pocket "0") is coloured green.

What is the outcome space $A$ for a spin of the roulette wheel? With $\mathcal{F} = A$, what are the probabilities $p_\alpha$ for a fair wheel? With

$$\mathcal{F} = \{\text{ball in a red pocket, ball in a black pocket, ball in the green pocket}\}\,,$$

what are the corresponding probabilities $p_{\text{red}}$, $p_{\text{black}}$ and $p_{\text{green}}$?

We conclude this section with two **comments** on the process of assigning probabilities to events (which is called *modelling*):

- We saw above that *symmetries* are a powerful way to constrain probabilities. The symmetry between the six sides of a fair die, the two sides of a fair coin, and the 37 pockets of a fair roulette wheel each sufficed to completely fix the corresponding probabilities $p_\alpha$.

- Modelling can also be guided by empirical information obtained by repeating an experiment many times. For example, if we don't know whether a set of dice are fair, we will be able to infer their probabilities $p_\alpha$ (with a certain confidence level) by rolling them enough times. The need to repeat the experiment many times comes from the law of large numbers, to which we now turn.

## 1.2 Law of large numbers

Let's continue considering the setup introduced above, with

$$\mathcal{F} = A = \{X_1, X_2, \cdots, X_n\} \tag{1}$$

for finite $n$, and probabilities $p_\alpha = P(X_\alpha)$ that obey

$$p_\alpha \in [0,1] \qquad\qquad \sum_{\alpha=1}^{n} p_\alpha = 1.$$

We can generalize this notation by writing instead

$$\sum_{X \in A} P(X) = 1,$$

which provides simple expressions for the **mean** $\mu$ and **variance** $\sigma^2$ of the probability space,

$$\mu = \langle X \rangle = \sum_{X \in A} X P(X) \tag{2}$$

$$\sigma^2 = \left\langle (X-\mu)^2 \right\rangle = \sum_{X \in A} (X-\mu)^2 P(X). \tag{3}$$

The angle bracket notation indicates the **expected** (or **expectation**) **value** with general definition

$$\langle f(X) \rangle = \sum_{X \in A} f(X) P(X), \tag{4}$$

which is a linear operation,

$$\langle c \cdot f(X) + g(X) \rangle = c \langle f(X) \rangle + \langle g(X) \rangle .$$

The square root of the variance, $\sqrt{\sigma^2} = \sigma$, is the **standard deviation**. What is $\sigma$ expressed in terms of $\langle X^2 \rangle$ and $\langle X \rangle^2$?

We now define a new experiment that consists of *repeating* the original experiment $R$ times, with each repetition independent of all the others. Using the same measurement as before for each repetition, we obtain a new outcome space that we can call $B$. For $R = 4$, what are some representative outcomes in the set $B$? What is the total size of $B$?

Each outcome in $B$ contains $R$ different $X^{(r)} \in A$ (with $\left\langle X^{(r)} \right\rangle = \mu$ and $\left\langle (X^{(r)} - \mu)^2 \right\rangle = \sigma^2$), one for each repetition $r = 1, \cdots, R$. Considering the case $R = 4$ for simplicity, any element of $B$ can be written as $X_i^{(1)} X_j^{(2)} X_k^{(3)} X_l^{(4)} \in B$ with corresponding probability

$$P_B \left( X_i^{(1)} X_j^{(2)} X_k^{(3)} X_l^{(4)} \right) = P_A \left( X_i^{(1)} \right) P_A \left( X_j^{(2)} \right) P_A \left( X_k^{(3)} \right) P_A \left( X_l^{(4)} \right),$$

using subscripts to distinguish between the single-experiment ($A$) and repeated-experiment ($B$) probability spaces.

Averaging over all $R$ repetitions defines the *arithmetic mean*

$$\overline{X}_R = \frac{1}{R} \sum_{r=1}^{R} X^{(r)}. \tag{5}$$

Unlike the true mean $\mu$, the arithmetic mean $\overline{X}_R$ is a random variable—a number that may be different for each element of $B$. That said, $\overline{X}_R$ and $\mu$ are certainly related, and so long as the standard deviation exists (i.e., $\sigma^2$ is finite), this relation can be proved rigorously in the limit $R \to \infty$.[3]

[3] In the computer project we will numerically investigate a case with divergent $\sigma^2$.

Here we will not be fully rigorous, and take it as given that

$$\left\langle \left(X^{(i)} - \mu\right)\left(X^{(j)} - \mu\right)\right\rangle = \sigma^2 \delta_{ij} = \begin{cases} \sigma^2 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases},$$

where the *Kronecker delta* $\delta_{ij} = 1$ for $i = j$ and vanishes for $i \neq j$. This is a consequence of the assumed independence of the different repetitions. Using this result and the relation $\left(\sum_i a_i\right)\left(\sum_j b_j\right) = \sum_{i,j}(a_i b_j)$, express the following quantity in terms of $\sigma$ and $R$:

$$\left\langle \left( \frac{1}{R} \sum_{r=1}^{R} X^{(r)} - \mu \right)^2 \right\rangle =$$

You should find that your result vanishes in the limit $R \to \infty$, so long as $\sigma^2$ is finite. Since the square makes this expectation value a sum of non-negative terms, it can vanish only if every one of those terms is individually zero.

This establishes the **law of large numbers**:

$$\lim_{R\to\infty} \frac{1}{R} \sum_{r=1}^{R} X^{(r)} = \mu, \tag{6}$$

where we have assumed $\left\langle X^{(r)} \right\rangle = \mu$ and $\left\langle (X^{(r)} - \mu)^2 \right\rangle = \sigma^2$ are finite.

## 1.3 Probability distributions

It is not necessary to make the assumption (Eq. 1) that our outcome space contains only a countable number of possible outcomes. The considerations above continue to hold even if the random variable $X$ is a continuous real number. In this case, however, the identification of probabilities with outcomes is slightly more complicated, which will be relevant when we consider the central limit theorem in the next section.

When the outcome can be any number on the real line, the fundamental object is a **probability distribution** (or **density function**) $p(x)$ defined for all $x \in \mathbb{R}$. Starting from this density, a probability is determined by integrating over a given interval. Calling this interval $[a, b]$, the integration produces the probability that the outcome $X$ lies within the interval,

$$P\left(a \leq X \leq b\right) = \int_a^b p(x)\, dx.$$

We similarly generalize the definition of an expectation value (Eq. 4) to an integral over the entire domain of the probability distribution,

$$\langle f(x) \rangle = \int f(x)\, p(x)\, dx.$$

We will omit the limits on integrals over the entire domain, so for $x \in \mathbb{R}$ we implicitly have $\int dx = \int_{-\infty}^{\infty} dx$. An important set of expectation values is

$$\left\langle x^{\ell} \right\rangle = \int x^{\ell}\, p(x)\, dx, \tag{7}$$

which provides the mean and variance of the probability distribution $p(x)$, through generalizations of Eqs. 2–3:

$$\mu = \langle x \rangle = \int x\, p(x)\, dx \qquad\qquad \sigma^2 = \left\langle x^2 \right\rangle - \langle x \rangle^2 . \tag{8}$$

The expression for the variance should be familiar from your determination of the standard deviation in an earlier gap. Unless stated otherwise, we will assume the mean and variance are both finite for the probability distributions we consider.

## 1.4 Central limit theorem

The central limit theorem is a major result of probability theory. Over the years it has been expressed in several equivalent ways, and there are also many distinct variants of the theorem accommodating different conditions and assumptions. In this module we are interested in applying rather than proving the central limit theorem; the curious can find proofs in many textbooks.

The version of the theorem we use in this module assumes we have $N$ independent random variables $x_1, \cdots, x_N$, each of which has the same (finite) mean $\mu$ and variance $\sigma^2$. (Such random variables are said to be *identically distributed*, and a common way to obtain them is to repeat an experiment $N$ times, as we considered in Section 1.2.) Just as in Eq. 5, the sum

$$s = \sum_{i=1}^{N} x_i \tag{9}$$

is itself a random variable.

The **central limit theorem** states that for large $N \gg 1$ the probability distribution for $s$ is

$$p(s) \approx \frac{1}{\sqrt{2\pi N \sigma^2}} \exp\left[ -\frac{(s - N\mu)^2}{2N\sigma^2} \right], \tag{10}$$

with the approximation becoming exact in the $N \to \infty$ limit.

In addition to asserting that the collective behaviour of many independent and identically distributed random variables $x_i$ is governed by a **normal** (or **gaussian**) **distribution**, the central limit theorem further specifies the precise form of this distribution in terms of the mean and variance of *each individual* $x_i$.

As an **example** to illustrate the applicability of the central limit theorem even for a modest $N = 5$, consider the roulette wheel discussed in Section 1.1. A simple game of roulette would let us place bets on whether or not the ball will end up in a red- or black-coloured pocket: If we bet correctly we get back twice the money we put in; otherwise we lose our money. Define our (potentially negative) *gain* to be the amount we receive minus the amount we spend on bets.

Suppose we place £5 bets on 'black' for each of $N$ spins of the roulette wheel. What are the probabilities and gains of winning and of losing for any single one of those spins? Letting $W = 0, \cdots, N$ be the number of spins where we win, what is our total gain $G_W$ as a function of $(N, W)$?

Recall that the number of different ways we could win $W$ times out of $N$ attempts is given by the binomial coefficient

$$\binom{N}{W} = \frac{N!}{W!\,(N-W)!},$$

with $0! = 1$. Setting $N = 5$, what are the six probabilities $p_0$ through $p_5$ that we win $W = 0, \cdots, 5$ times? What is the general expression for $p_W$ for any $(N, W)$?

Now let's apply the central limit theorem to this setup. What are the mean gain and its variance for a single spin of the wheel? What is the resulting probability distribution $p(G)$ given by the central limit theorem for the gain after $N$ spins?

In order to compare this approximation with the exact $p_i$ computed above, we need to extract probabilities by integrating over appropriate intervals as discussed in Section 1.3. Natural intervals to consider are

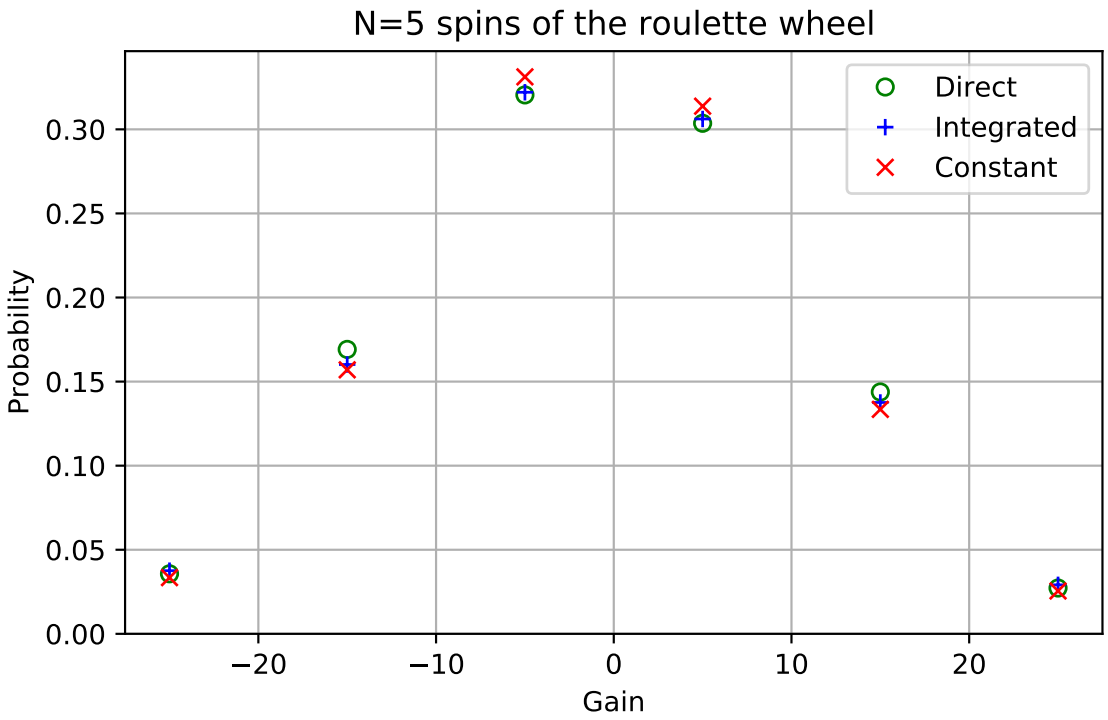
$$P_{\text{integ}}(G_W) \equiv \int_{G_W-\Delta G/2}^{G_W+\Delta G/2} p(g)\, dg,$$

where $\Delta G = G_{W+1} - G_W$ is a constant you can read off from your work above. These numerical integrations are not convenient to do by hand, but can easily be performed by Maple, Python, MATLAB, Mathematica, etc. Alternately, we can simplify further by approximating $p(G)$ as a constant within each interval, which would give us

$$P_{\text{const}}(G_W) \equiv p(G_W)\ \Delta G.$$

What are the six $P_{\text{integ}}$ and $P_{\text{const}}$?

You can check your results against those plotted below, which come from this python code that is available for you to peruse, run for different $N$ or modify.




## 1.5 Diffusion and the central limit theorem

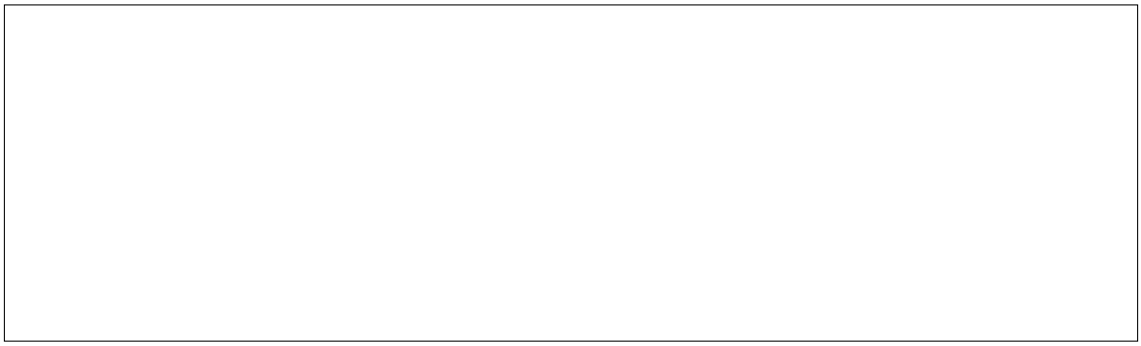
### 1.5.1 Random walk on a line

As a more general application and illustration of the central limit theorem, let's consider the behaviour of a randomly moving object. Such **random walks** appear frequently in mathematical modelling of stochastic phenomena (including Brownian motion), and can be applied to movement through either physical space or more abstract vector spaces. They are examples of Markov processes, in which the state of the system (in this case the position of the 'walker') at any time probabilistically depends only on the system's prior state at the previous point in time—there is no 'memory' of any earlier states. The resulting sequence of system states is known as a *Markov chain*, since each state is produced from the one before, like links in a chain.

To start with a simple case, let's consider a random walker that moves only in a single spatial dimension—to the left or to the right on a line—and can only take 'steps' of a fixed length, which we can set to $\ell = 1$ without loss of generality. At each point in time, the walker takes either a step to the right ($R$) with probability $p$ or a step to the left ($L$) with probability $q = 1 - p$. We will further assume that each step takes a constant amount of time $\Delta t$, so a walk of $N$ steps will last for total time

$$t = N\Delta t. \tag{11}$$

As an example, for $N = 6$ a representative walk can be written as $LRLRRR$, which leaves the walker $x = 2$ steps to the right of its starting point ($x = 0$). The opposite walk $RLRLLL$ would leave the walker at $x = -2$, with negative numbers indicating positions to the left of the starting point. How many possible walks are there for $N = 6$, and what is the probability (in terms of $p$ and $q$) for these particular walks $LRLRRR$ and $RLRLLL$ to occur? How many possible walks are there for general $N$, and what is the probability for any particular walk involving $r$ steps to the right to occur?



We will be interested in the walker's final position $x$ at time $t$ after it has taken $N$ steps. Just as for the possible gains after $N$ spins of the roulette wheel considered in Section 1.4, there are a range of possible final positions $x$, each of which has some probability $P(x)$ of being realized. The key pieces of information we want to determine are the expectation value $\langle x \rangle$ and the variance $\langle x^2 \rangle - \langle x \rangle^2$ that indicates the scale of fluctuations we can expect around $\langle x \rangle$ as the $N$-step walk is repeated many times from the same starting point.[4]

Suppose the $N$ total steps involve $r$ steps to the right. What is the final position $x$ of the walker in terms of $N$ and $r$? Check your general answer for the cases $N = 6$ and $r = 4$, 2 considered above.

This relation makes it equivalent to consider either the probability $P_r$ of taking $r$ steps to the right, or the probability $P(x)$ of ending up at final position $x$. This equivalence will not hold for more general random walks in which the step length is no longer fixed and $\ell_i$ can vary from one step to the next.

[4] We will reserve the variables $\mu$ and $\sigma^2$ for the mean and variance (respectively) of the single-step process, which we will connect to the central limit theorem in Section 1.5.3.

Because the order in which steps are taken does not affect the final position $x$, to determine the probability $P_x$ we have to count all possible ways of walking to $x$. For $N = 6$, what are all the possible walks that produce $x = 4$, and what is the corresponding probability $P_4$?

Your answer should have a factor of $6$ that corresponds to the binomial coefficient $\binom{N}{r} = \binom{6}{5} = 6$. In terms of this binomial coefficient, what is the general probability $P_r$ that an $N$-step walk will include $r$ steps to the right in any order?

Given this probability $P_r$, we can apply Eqs. 2–3 to find the expectation value $\langle x \rangle$ and the variance $\langle x^2 \rangle - \langle x \rangle^2$. We will need to compute

$$\begin{aligned} \langle x \rangle &= \sum_{r=0}^{N} (2r - N) P_r = 2 \langle r \rangle - N \\ \langle x^2 \rangle &= \sum_{r=0}^{N} (2r - N)^2 P_r = 4 \langle r^2 \rangle - 4N \langle r \rangle + N^2, \end{aligned} \tag{12}$$

where, as usual, $\langle r^n \rangle = \sum_{r=0}^{N} r^n P_r$. An excellent trick to calculate these expectation values $\langle r^n \rangle$ is to define the **generating function**

$$T(\theta) = \sum_{r=0}^{N} e^{r\theta} P_r. \tag{13}$$

This approach introduces a parameter $\theta$ that we subsequently remove by setting $\theta = 0$. For example, $T(0) = \sum_{r=0}^{N} P_r = 1$. What do you obtain upon taking derivatives of the generating function and then setting $\theta = 0$?

$$\left.\frac{d}{d\theta}T(\theta)\right|_{\theta=0} =$$

$$\left.\frac{d^n}{d\theta^n}T(\theta)\right|_{\theta=0} =$$

For the current case of a fixed-step-length random walk in one dimension, the probabilities $P_r$ provide a simple closed-form expression for the generating functional:

$$\left(e^{\theta}p + q\right)^N = \sum_{r=0}^{N} e^{r\theta}\binom{N}{r}p^r q^{N-r} = \sum_{r=0}^{N} e^{r\theta}P_r = T(\theta), \tag{14}$$

making use of the binomial formula $(a+b)^N = \sum_{i=0}^{N}\binom{N}{i}a^i b^{N-i}$. It's straightforward to take the necessary derivatives of Eq. 14:

$$\left.\frac{d}{d\theta}\left(e^{\theta}p + q\right)^N\right|_{\theta=0} =$$

$$\left.\frac{d^2}{d\theta^2}\left(e^{\theta}p + q\right)^N\right|_{\theta=0} =$$

Since $p + q = 1$, the simplification $\left(e^{\theta}p + q\right)^n\big|_{\theta=0} = 1$ provides

$$\langle r\rangle = Np \qquad\qquad \left\langle r^2\right\rangle = Np(Np+q).$$

Plugging these results into Eq. 12, we have:

$\langle x \rangle =$

$\langle x^2 \rangle =$

In the end, you should obtain

$$\langle x \rangle = N(2p-1) \qquad\qquad \langle x^2 \rangle - \langle x \rangle^2 = 4Npq. \tag{15}$$

We can check that this $\langle x \rangle$ produces the expected results in the special cases $p = 0$, $1/2$ and $1$, while the variance also behaves appropriately by vanishing for both $p = 0$ and $1$.

### 1.5.2 Law of diffusion

It's possible to gain a more intuitive interpretation of the results in Eq. 15 by expressing them in terms of the total time $t$ taken by the random walk (Eq. 11). Inserting $N = t/\Delta t$ into Eq. 15,

$$\langle x \rangle = \frac{t}{\Delta t}(2p-1) = \frac{2p-1}{\Delta t} t \equiv v_{\mathrm{dr}} t,$$

we see that the *expected* final position of the walker depends linearly on time, with **drift velocity**

$$v_{\mathrm{dr}} = \frac{2p-1}{\Delta t} = \frac{N(2p-1)}{t} = \frac{\langle x \rangle}{t}. \tag{16}$$

The sign of the drift velocity indicates whether the walker is drifting to the right ($p > \frac{1}{2}$) or to the left ($p < \frac{1}{2}$). The standard deviation of the final position of the walker provides a measure of the scale of fluctuations (or 'uncertainty') around the expectation value that we should anticipate:

$$\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = 2\sqrt{Npq} = 2\sqrt{\frac{pq}{\Delta t}}\sqrt{t},$$

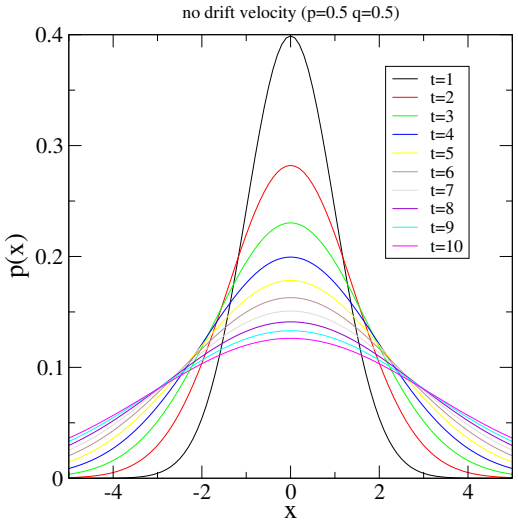
which increases proportionally to $\sqrt{t}$. This is a particular realization of a very general result.

The **law of diffusion** states that

$$\Delta x = D\sqrt{t}, \tag{17}$$

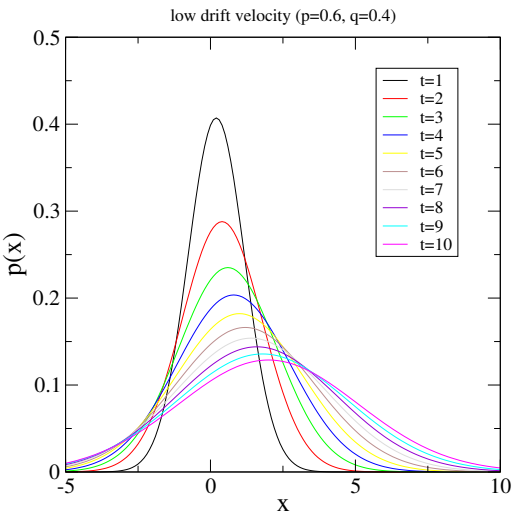
where $D$ is the **diffusion constant** and the uncertainty $\Delta x$ is sometimes called the *diffusion length*.

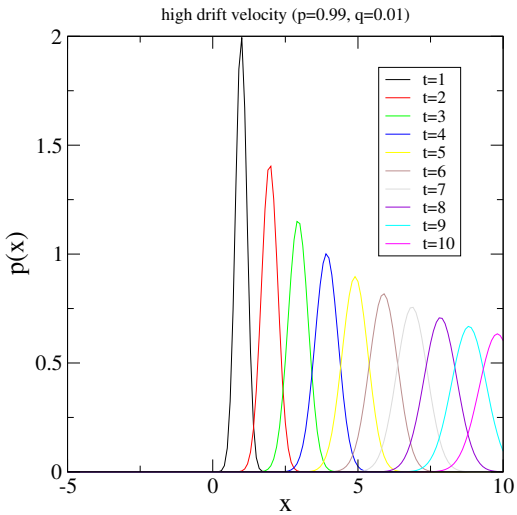
The diffusion constant $D = 2\sqrt{\frac{pq}{\Delta t}}$ that we computed above is specific to the current case of a fixed-step-length random walk in one dimension. The figure below illustrates why this behavior is called diffusion.




This figure shows probability distributions $p(x)$ (computed below) for the special case $p = q = \frac{1}{2}$ where $v_{\rm dr} = 0$, so the expectation value is always $\langle x \rangle = 0$ for any walk time $t$. However, as time goes on, there is a steady decrease in the probability that any given walker will end up at its starting point $x = 0$.[5] Instead, the probability distribution spreads (or *diffuses*) outward, leading to a constant ('one-sigma' or 68%) probability that the walker remains within the steadily increasing interval $-D\sqrt{t} \leq x \leq D\sqrt{t}$.

Except in the trivial cases $p = 0$ or $q = 0$, this same diffusive phenomenon occurs when the drift velocity is non-zero. This is shown in the two figures below, which compare the case of a low but non-zero drift velocity on the left with a high drift velocity on the right.




[5] As in Section 1.4, we can extract this probability by integrating the distribution $p(x)$ over the interval $-0.5 \leq x \leq 0.5$.

In the figure on the left, each individual probability distribution looks similar to the corresponding one for $v_{\mathrm{dr}} = 0$, but now their central peaks (and expectation values $\langle x \rangle$) drift steadily to the right. The distributions in the figure on the right look a bit different, but still diffuse to exhibit shorter and broader peaks as time goes on.

When $p \neq \frac{1}{2}$ so that $\langle x \rangle \neq 0$, it is interesting to compare how 'fast' the expectation value drifts compared to how fast its uncertainty grows due to diffusion. We can do this by considering the following *relative* uncertainty:

$\frac{\Delta x}{\langle x \rangle} =$

You should find that at large times this ratio vanishes proportionally to $1/\sqrt{t} \propto 1/\sqrt{N}$. Although the absolute uncertainty grows by diffusion, $\Delta x = D\sqrt{t}$, for $v_{\mathrm{dr}} \neq 0$ the linear drift in the expectation value becomes increasingly dominant as time goes on.

### 1.5.3 Applying the central limit theorem

Based on our work in Section 1.4, we can see how to apply the central limit theorem to analyze this fixed-step-length random walk in one dimension, for large numbers of steps $N$ or equivalently large times $t = N\Delta t$. Each step in the random walk can be considered an independent and identically distributed random variable $x_i$, analogous to each spin of the roulette wheel. The corresponding probability space involves only two possible outcomes: a step of length $\ell = 1$ to the right or to the left with probability $p$ or $q$, respectively. From this we can easily compute the mean and variance *of the single-step process*:

$\mu = \langle x_i \rangle =$

$\langle x_i^2 \rangle =$

$\sigma^2 = \langle x_i^2 \rangle - \langle x_i \rangle^2 =$

The final position $x$ of the walker after $N$ steps is exactly the sum over these $x_i$ given in Eq. 9. Its probability distribution $p(x)$ from the central limit theorem is therefore obtained simply by plugging in these single-step $\mu$ and $\sigma^2$, which we can then rewrite in terms of the drift velocity and diffusion constant:

$$\begin{aligned} p(x) &= \frac{1}{\sqrt{2\pi(4Npq)}} \exp\left[-\frac{(x-N(2p-1))^2}{8Npq}\right] \\ &= \frac{1}{\sqrt{2\pi D^2 t}} \exp\left[-\frac{(x-v_{\mathrm{dr}}t)^2}{2D^2 t}\right], \end{aligned} \tag{18}$$

which was used to produce the three figures above. We could have jumped straight to the final line by considering Eq. 10 and noting

$$v_{\mathrm{dr}}t = N(2p-1) = N\mu \qquad\qquad D^2 t = 4pq\frac{t}{\Delta t} = N\sigma^2. \tag{19}$$

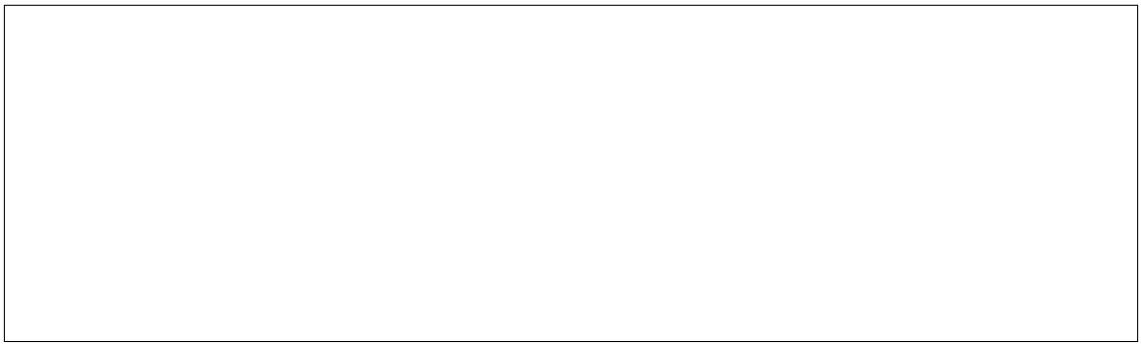
While the functions of $p$ and $q$ that we computed are specific to the particular type of random walk we're currently considering, Eq. 19 relating $\{v_{\mathrm{dr}}, D^2\}$ with the single-step $\{\mu, \sigma^2\}$ turn out to be generic. This is remarkable, because it means that so long as the single-step mean and variance are finite, we end up with Eq. 18 as the large-$t$ probability distribution for any random walk in a single variable $x$.

This result is related to the generality of the law of diffusion, which we can recognize in the structure of Eq. 18. For any time $t$, the factor $\delta x^2 \equiv (x - v_{\mathrm{dr}}t)^2$ simply measures the distance from the drifting expectation value where the gaussian $p(x)$ is peaked. As $t$ increases, so does the factor $2D^2t$ dividing this $\delta x^2$, meaning that a larger distance from the peak is needed for the overall argument of the exponential to reach a given value—in other words, the peak becomes broader. This in turn requires a shorter peak, reflected in the $\frac{1}{\sqrt{t}}$ in the overall coefficient, which is set by requiring $\int p(x)\, dx = 1$. Therefore we can conclude that the law of diffusion holds whenever the mean and variance of the single-step process are finite. In the computer project we will numerically investigate the *anomalous diffusion* that occurs when this condition is not satisfied.

# Week 2: Micro-canonical ensemble

## 2.1 Statistical ensembles and thermodynamic equilibrium

We begin this week by developing the concept of *statistical ensembles* (introduced by J. Willard Gibbs in the early 1900s), building on the probability foundations we laid last week. As forecast last week, we will be interested in 'experiments' that simply allow a collection of degrees of freedom to evolve in time, subject to certain constraints. At a given time $t_1$, the disposition of these degrees of freedom defines the state $\omega_1$ of the system. To consider a couple of examples, what would be a representative state for a system of $8$ *spins* (arrows that can point either 'up' or 'down') arranged in a line?[6] What information would characterize the state of $N$ hydrogen ($H_2$) molecules in a container?



At a different time $t_2$, the system's state $\omega_2$ is generally different from $\omega_1$. However, there are some measurements we can perform on these states that always produce the same outcome as the system evolves in time. These measurements define *conserved quantities*, an important example of which is the total energy $E$ *inside* an isolated (or 'closed') system,

$$E(\omega_1) = E(\omega_2).$$

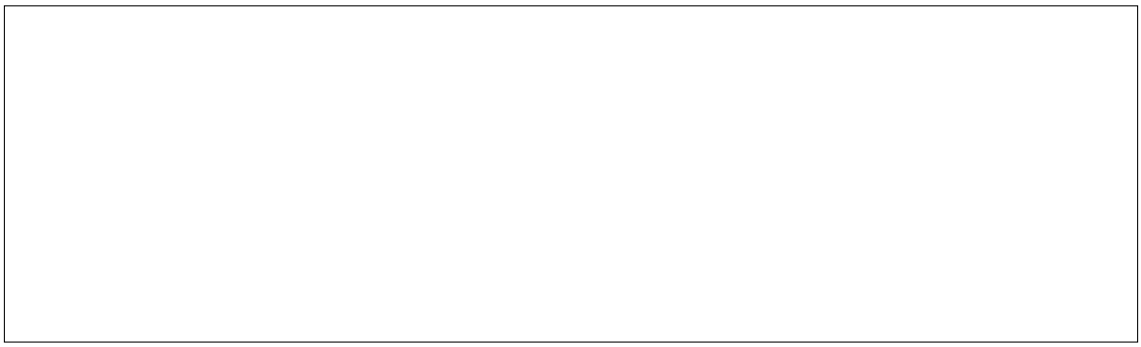
The conservation of energy is presumably a familiar concept, and you may also know that it can be rigorously proven through Emmy Noether's theorem.[7] Because statistical physics was first developed when conservation of energy was primarily an empirical observation rather than a proven result, it was given a more grandiose name: the **first law of thermodynamics**. Another way of stating the first law is that any change in the internal energy of one particular system $\Omega$ must be matched by an equal and opposite change in the energy of some other system(s) with which $\Omega$ is in contact. We will return to this formulation of the first law in future weeks.

For now, let's return to the **examples** above, and suppose that the spin system is placed in an external magnetic field. If a spin is parallel to the field,

[6] We will consider spin systems extensively in this module. In addition to obeying simple mathematics analogous to flipping coins, spins also serve as good models of physical systems such as magnetic molecules.

[7] This proof holds only in 'flat' space-time as opposed to the curved space-time manifolds than arise in general relativity—which is far beyond the scope of this module.

it contributes energy $-H$ to the total energy $E$ of the system. If a spin is anti-parallel to the field, it instead contributes energy $H \geq 0$. What is the total energy $E$ of a system of $N$ spins in a state with $n_+$ spins anti-parallel to the field and $n_- = N - n_+$ parallel to it? What is $E$ for the representative $N = 8$-spin state you wrote down above? How many of the $2^8 = 256$ states of the spin system have this energy?



For the $N$ hydrogen molecules in a container, we can write a simple expression for the energy $E$ by treating each molecule as a *point-like particle*, with no size or structure. In this case each molecule contributes only its kinetic energy to $E$,

$$E = \frac{m}{2}\sum_{i=1}^{N} \vec{v}_i^{\,2} = \frac{1}{2m}\sum_{i=1}^{N} \vec{p}_i^{\,2},$$

where $\vec{v}_i$ is the velocity of the $i$th molecule, $\vec{p}_i = m\vec{v}_i$ is its momentum, and all molecules have exactly the same mass $m$.

As emphasized in last week's introduction, we treat the time evolution of the system as a stochastic process in which the system probabilistically adopts a sequence of states $\omega_i \in \Omega$:

$$\omega_1 \longrightarrow \omega_2 \longrightarrow \omega_3 \longrightarrow \omega_4 \longrightarrow \cdots$$

This attitude is adopted as a matter of practicality rather than one of principle. In principle, Newton's laws would allow us to exactly predict the time evolution of (say) $\sim 10^{23}$ hydrogen molecules, but only by specifying $\sim 10^{23}$ initial conditions and solving $\sim 10^{23}$ differential equations. Since we cannot hope to write down so much information or carry out so many computations, we instead apply probability theory in order to analyze these systems.

This leads us to the following core definition: A **statistical ensemble** is the set of all states $\Omega = \{\omega_1, \omega_2, \cdots\}$ that a system can possibly adopt through its time evolution. Each state $\omega_i$ has some probability $p_i$ of being adopted by the system, so we can recognize a statistical ensemble as a probability space.

Because these states $\omega_i$ depend on the 'microscopic' degrees of freedom that compose the overall system, we will refer to them as **micro-states** from now

on. From last week's definition of probability, we have the requirement $\sum_i p_i = 1$, which simply means that the system must be in *some* micro-state at any point in time. The fact that time evolution cannot change any conserved quantities, as discussed above, means that such conserved quantities characterize statistical ensembles. We will define different types of statistical ensemble that depend on the specific set of conserved quantities.

This week we define a **micro-canonical ensemble** to be a statistical ensemble characterized by conserved internal energy $E$ and conserved number of degrees of freedom $N$ (which we will call **particle number** for short).

According to the discussion above, this means that a system governed by the micro-canonical ensemble is *isolated* in the sense that it cannot exchange energy or particles with any other system.

Now that the micro-canonical ensemble is defined, we can connect it to our intuition from everyday physical systems. Let's consider a collection of particles moving around and bouncing (or '*scattering*') off each other in a sealed container. To a first approximation, this should describe the behaviour of air in a room, which our lived experience indicates is spread quite uniformly throughout the room in a way that is stable as time passes. We do not expect all the air in a room to be concentrated in any one corner, nor do we expect strong collective gusts of wind without some clear external influence.

These qualitative expectations illustrate the idea of **thermodynamic equilibrium**, an axiomatic concept in statistical physics.[8] We can mathematically define thermodynamic equilibrium through the probabilities $p_i$ that appear in the micro-canonical ensemble.

A micro-canonical system $\Omega$ with $M$ micro-states $\omega_i$ is in thermodynamic equilibrium if and only if all probabilities $p_i$ are equal. If $M$ is finite, the requirement $\sum_i p_i = 1$ implies

$$p_i = \frac{1}{M}. \tag{20}$$

The full meaning and significance of this definition are not immediately obvious, and we will continue exploring them through consideration of derived quantities such as entropy and temperature. First, we emphasize that this equilibrium is indeed *dynamic*: There is not a single 'equilibrium state' that the system approaches; instead, the system continues probabilistically adopting different states as it evolves in time.

[8] Our expectation that physical systems generically evolve towards thermodynamic equilibrium as time passes is more formally expressed as the ergodic hypothesis.

## 2.2 Entropy and its properties

### 2.2.1 Definition of entropy

We can gain further insight into thermodynamic equilibrium by considering a famous derived quantity.

The **entropy** of a statistical ensemble $\Omega$ with a countable number of micro-states $M$ is defined to be

$$S = -\sum_{i=1}^{M} p_i \log p_i, \tag{21}$$

where $p_i$ is the probability for micro-state $\omega_i$ to occur. Unless otherwise specified, "$\log$" indicates the natural logarithm of base $e$.

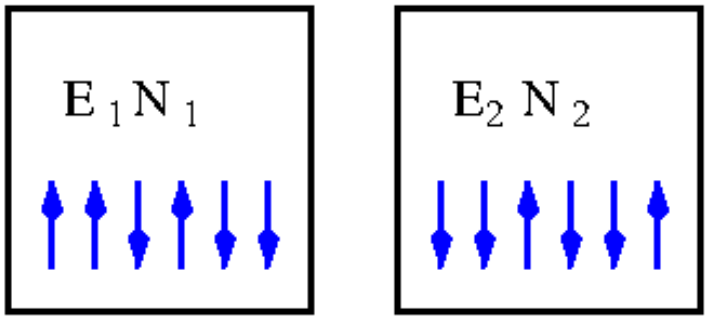
When the system under consideration is in thermodynamic equilibrium, we expect derived quantities such as the entropy to be stable over time, even as different micro-states are probabilistically adopted. This implies that such derived quantities are functions of the conserved quantities that are the same for all micro-states. Therefore, for the micro-canonical ensemble, the equilibrium entropy $S(E, N)$ would be a function of the conserved energy and particle number.

By inserting Eq. 20 into Eq. 21 you can quickly compute a simple expression for the entropy of a micro-canonical ensemble in thermodynamic equilibrium:

Your result should depend only on the number of micro-states $M$, and diverge as $M \to \infty$. While the energy $E$ and particle number $N$ are not explicit in this expression, $\{E, N, M\}$ are inter-related and might be expressed in terms of each other depending on the particular situation under consideration. For example, what is the equilibrium entropy of the system of $N$ spins considered above, if the external magnetic field is turned off (so $H = 0$ implying $E = 0$)?

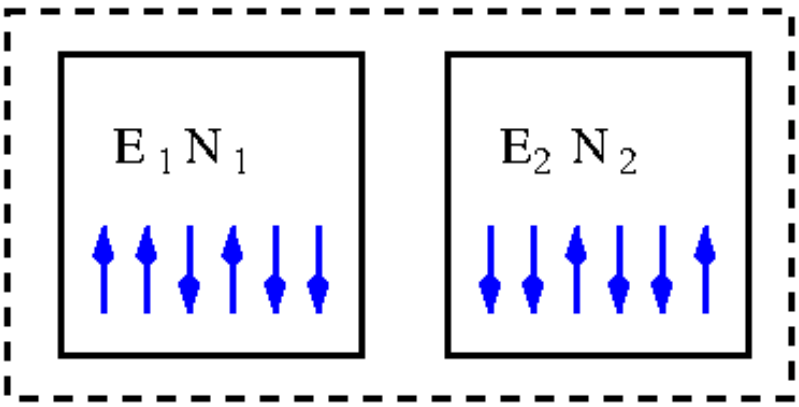
### 2.2.2 Extensivity

The increase in entropy for an increasing number of micro-states $M$ is a reflection of entropy being an *extensive* quantity. Extensive quantities are formally defined by considering how they behave if two isolated systems are *analyzed* as a single system—while still remaining isolated from each other, exchanging neither energy nor particles. This is clearest to consider through the specific example shown below of two isolated spin systems, $\Omega_1$ & $\Omega_2$, respectively characterized by the corresponding energies $E_1$ & $E_2$ and particle numbers $N_1$ & $N_2$. To simplify the subsequent analysis, we can assume that both systems are placed in external magnetic fields with the same $H$, so that $E_S = H\left(n_+^{(S)} - n_-^{(S)}\right)$ for $S \in \{1, 2\}$.




In the figure above, we can take system $\Omega_1$ to have $M_1$ micro-states with probabilities $p_i$ while system $\Omega_2$ has $M_2$ micro-states with probabilities $q_k$. (As discussed above, $M_S$ is a function of $E_S$ and $N_S$ for $S = 1, 2$.) Then, even without assuming thermodynamic equilibrium, the entropies of the two systems are

$$S_1 = -\sum_{i=1}^{M_1} p_i \log p_i \qquad\qquad S_2 = -\sum_{k=1}^{M_2} q_k \log q_k.$$




Now we keep these two (sub)systems isolated from each other, but consider them as a combined system $\Omega_{1+2}$, as illustrated above. In order to compute

the entropy $S_{1+2}$, we first need to figure out how many micro-states the combined system could possibly adopt ($M_{1+2}$), and then determine the corresponding probability for each micro-state. Both steps are simplified by the systems being isolated from each other, so that they are statistically independent. Specifically, with subsystem $\Omega_1$ in a fixed micro-state $\omega_i^{(1)}$, subsystem $\Omega_2$ could independently inhabit any of its $M_2$ micro-states. What is the resulting $M_{1+2}$ in terms of $M_1$ and $M_2$?

$M_{1+2} =$

Similarly, statistical independence means that the combined probability of subsystem $\Omega_1$ adopting micro-state $\omega_i^{(1)}$ while subsystem $\Omega_2$ adopts $\omega_k^{(2)}$ is the product of the individual probabilities, $p_i q_k$. We can check that this is a well-defined probability, with

$$\sum_{M_{1+2}} p_i q_k = \sum_{i=1}^{M_1} \sum_{k=1}^{M_2} p_i q_k = \left[\sum_{i=1}^{M_1} p_i\right] \cdot \left[\sum_{k=1}^{M_2} q_k\right] = 1 \cdot 1 = 1.$$

Inserting the probability $p_i q_k$ into Eq. 21, and recalling $\log(a \cdot b) = \log a + \log b$, what is the combined entropy $S_{1+2}$ of these two independent subsystems?

$S_{1+2} =$

You should find that the entropies of the two isolated subsystems add up to form the total, which is also the case for the energies and particle numbers,

$$E_{1+2} = E_1 + E_2 \qquad N_{1+2} = N_1 + N_2 \qquad S_{1+2} = S_1 + S_2$$

**Extensive** quantities are defined to be those that add up across independent subsystems. Examples include the energy, particle number and entropy as shown above. This can be contrasted with **intensive** quantities, which are defined to be independent of the extent of the system (and hence the same for subsystems as for the combined system). We will see an example of this below. It is possible for quantities to be neither extensive nor intensive.

Finally, if we were to assume that each subsystem is (independently) in thermodynamic equilibrium, with finite $M_1$ and $M_2$,

$$p_i = \frac{1}{M_1} \qquad\qquad q_k = \frac{1}{M_2}$$
$$S_1 = \log M_1 \qquad\qquad S_2 = \log M_2.$$

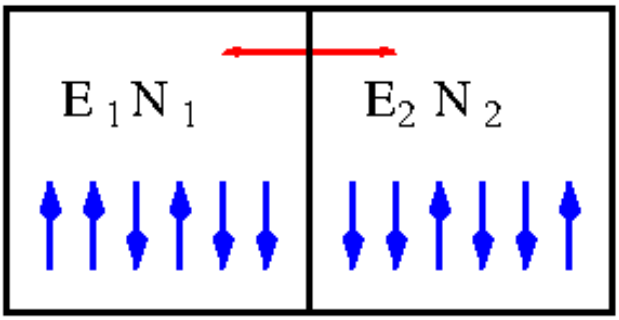
then we would find as a consequence that their combination is also in thermodynamic equilibrium, since

$$p_i q_k = \frac{1}{M_1 M_2} = \frac{1}{M_{1+2}}$$

is the same for every combined micro-state.

### 2.2.3 Second law of thermodynamics

Let's continue considering the two spin (sub)systems discussed above, with one significant change: We suppose the two subsystems are now able to exchange energy (but not particles) with each other. We'll say they are in *thermal contact* with each other, rather than being fully isolated. This is illustrated by the figure below:




The total energy $E = E_1 + E_2$ remains conserved, so the overall system $\Omega$ is still governed by the micro-canonical ensemble. However, the individual energies $E_1$ and $E_2$ can now change as time passes, meaning that *each subsystem is no longer micro-canonical*.

The overall $\Omega$ is *not* the same as the combined $\Omega_{1+2}$ considered above. We need to reconsider the total number of micro-states $M$ that $\Omega$ could adopt, which is much more difficult than before because we can no longer apply statistical independence. Our main remaining tool is the conservation of the total energy $E$.

Considering a micro-state in which the $N_1$ spins contribute energy $e_1$ to the total, we know that the $N_2$ spins must contribute the remaining $e_2 = E - e_1$. Our work above implies there are $M_{e_1} = M^{(1)}_{e_1} M^{(2)}_{E-e_1}$ micro-states providing this particular distribution of energies, where $M^{(1)}_{e_1}$ is the number of micro-states of the formerly isolated subsystem $\Omega_1$ with energy $e_1$, and $M^{(2)}_{E-e_1}$ similarly corresponds

to $\Omega_2$ with energy $E - e_1$. We also know that it's possible to have $e_1 = E_1$, since that's the initial energy of $\Omega_1$ before it was brought into thermal contact with $\Omega_2$. When $e_1 = E_1$, we have $M_{e_1} = M_1 M_2$, covering all the micro-states of the combined system when the two subsystems were isolated. *In addition*, we also have to count any other micro-states for which $e_1 \neq E_1$:

$$M = \sum_{e_1} M^{(1)}_{e_1} M^{(2)}_{E-e_1} = M_1 M_2 + \sum_{e_1 \neq E_1} M^{(1)}_{e_1} M^{(2)}_{E-e_1} \geq M_1 M_2. \tag{22}$$

Equality holds when $e_1 = E_1$ is the only possibility, which is an extremely special case. This is all we can say without specifying more details of a particular example, but it allows us to obtain a famous result for the total entropy $S$ of $\Omega$ *in thermodynamic equilibrium*:

$S = \log M \geq$

You have now derived a form of the **second law of thermodynamics**,

$$S \geq S_{1+2} = S_1 + S_2.$$

In words, whenever isolated (sub)systems in thermodynamic equilibrium are brought into thermal contact with each other and allowed to exchange energy, the total entropy of the overall system can never decrease. Indeed, it generically increases except in extremely special cases.

Though we won't go through the more general derivation here, it turns out that the total entropy never decreases (and generically increases) as time passes, under *any* circumstances. This has many far-reaching consequences, the first of which is a more general definition of thermodynamic equilibrium that (unlike Eq. 20) will also apply when we consider statistical ensembles other than the micro-canonical ensemble. For simplicity we assume that any system under consideration has a finite number of micro-states, which means that its entropy is bounded from above. To motivate the definition below, note that the overall system $\Omega$ may have undergone an equilibration process to reach its thermodynamic equilibrium after its two (independently equilibrated) subsystems were brought into thermal contact—and in this process the entropy was non-decreasing.

A system is defined to be in **thermodynamic equilibrium** if its entropy is maximal.

We can *derive* Eq. 20 from this definition. All we need to do is maximize the entropy $S = -\sum_i p_i \log p_i$ subject (for the micro-canonical ensemble) to the three constraints of conserved energy, conserved particle number, and well-defined probabilities $\sum_i p_i = 1$. Only the last will turn out to matter, and can be incorporated into the maximization through the method of **Lagrange multipliers**. In case you are not familiar with this method, it involves maximizing the modified entropy

$$\overline{S} = S + \lambda \left( \sum_{i=1}^{M} p_i - 1 \right) = -\sum_{i=1}^{M} p_i \log p_i + \lambda \left( \sum_{i=1}^{M} p_i - 1 \right),$$

where the parameter $\lambda$ is called the 'multiplier'. In short, this procedure is valid because $\dfrac{\partial \overline{S}}{\partial \lambda} = 0$, so that any extremum of $\overline{S}$ corresponds to an extremum of $S$ when the multiplier is set to $\lambda = 0$. Recalling $\dfrac{\partial}{\partial x_k} \sum_i f(x_i) = \dfrac{\partial f(x_k)}{\partial x_k}$, what is the probability $p_k$ that maximizes $\overline{S}$?

$$0 = \frac{\partial \overline{S}}{\partial p_k} =$$

You should find that $p_k$ is some constant that depends on $\lambda$. We don't care about $\lambda$; so long as we know $p_k$ is constant, then $p_k = \frac{1}{M}$ to satisfy $\sum_k p_k = 1$. As advertised, we recover Eq. 20 from our new definition of thermodynamic equilibrium based on the second law.

## 2.3 Temperature

In the micro-canonical ensemble, the conserved internal energy and particle number are fundamental, while the temperature (like the entropy) is a derived quantity. As discussed below Eq. 21, in thermodynamic equilibrium such derived quantities are functions of the conserved $\{E, N\}$. In this section we will state the definition of temperature for the micro-canonical ensemble and apply this to a spin system. In the next section we will check that our definition reproduces our expectations from everyday experiences.

In thermodynamic equilibrium, the **temperature** in the micro-canonical ensemble is defined by

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_N . \tag{23}$$

In words, the (inverse) temperature is set by the dependence of the entropy on the internal energy for a fixed number of degrees of freedom.

Since this definition is not terribly intuitive, we will again gain insight by working through the **example** of $N$ spins in a line, in an external magnetic field of strength $H$. We saw above that $E = H(n_+ - n_-)$ for $n_+$ and $n_- = N - n_+$ spins respectively aligned anti-parallel and parallel to the magnetic field. Each (conserved) value of $E$ defines a *different* micro-canonical system, which we can expect to have a different number of micro-states $M(E)$, different entropy $S(E)$ and different temperature $T(E)$. We will compute the functional forms of each of these three quantities, starting with $M(E)$.

Even though the total energy $E$ remains fixed as time passes, individual spins can 'flip' between pointing parallel or anti-parallel to the magnetic field. Such spin flips simply have to come in pairs so that the overall $n_\pm$ both remain the same. As illustration, what are the spin configurations that produce the minimal energy $E_{\text{min}} \equiv E_0$ and the next-to-minimal $E_1$? What are $E_0$ and $E_1$ in terms of $\{N, H\}$, and how many distinct micro-states are there for each of $E_0$ and $E_1$?

Your results should generalize to

$$M(E_{n_+}) = \binom{N}{n_+} = \frac{N!}{n_+!\,(N-n_+)!}. \tag{24}$$

To take the derivative in Eq. 23, we need to express $n_+$ in terms of $\{E, N\}$. It will also be convenient to avoid the factorial operation, which is inconvenient to differentiate. For $N \gg 1$, we can accomplish both these goals by interpreting the spin system as a random walk (in the abstract 'line' of its possible energies $E$) and applying the central limit theorem:

- Each spin adds to $x \equiv \frac{E}{H} = 2n_+ - N$ a 'step' of fixed 'length' $\pm 1$. Our task therefore coincides with the special case we considered in Section 1.5.
- We don't impose any preference for positive vs. negative energies, meaning $p = q = \frac{1}{2}$ in the language of Section 1.5.
- With $p = q = \frac{1}{2}$, every one of the $2^N$ possible configurations of $N$ spins is equally probable. Therefore the probability $P_{n_+}$ that our overall 'walk' ends

up producing a configuration with $n_+ = \frac{1}{2}(x+N)$ is simply the fraction of those $2^N$ configurations with this $n_+$, in which we can recognize Eq. 24:

$$P_{n_+} = \frac{1}{2^N}\binom{N}{n_+} = \frac{M(E_{n_+})}{2^N} \qquad \Longrightarrow \qquad M(E_{n_+}) = 2^N P_{n_+}.$$

- To estimate $P_{n_+}$ for $N \gg 1$, we apply the central limit theorem just as in Section 1.5.3. In particular, we can re-use our computation that $\mu = 2p-1 = 0$ and $\sigma^2 = 4pq = 1$, finding

$$p(x) = \frac{1}{\sqrt{2\pi N}} \exp\left[-\frac{x^2}{2N}\right] = \frac{1}{\sqrt{2\pi N}} \exp\left[-\frac{E^2}{2NH^2}\right]$$

  for the probability *distribution* from which we want to extract $P_{n_+}$.

- We saw in Section 1.4 that $P_{\text{const}}(n_+) = p(2n_+ - N)\Delta n_+$ is a good approximation, and $\Delta n_+ = 1$. Therefore we find

$$M(E) \approx 2^N p(2n_+ - N) \approx \frac{2^N}{\sqrt{2\pi N}} \exp\left[-\frac{E^2}{2NH^2}\right]. \tag{25}$$

What is the derivative of the log of Eq. 25 with $N$ fixed?

$$\frac{\partial}{\partial E}\log M\bigg|_N =$$

You should find the temperature

$$T \approx -\frac{NH^2}{E} \qquad\qquad N \gg 1, \tag{26}$$

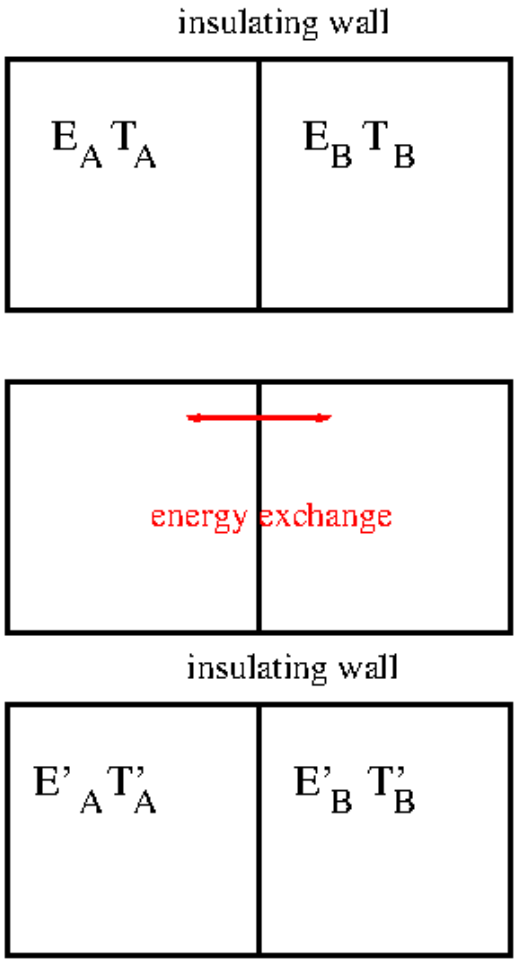
which in several ways does *not* seem to match our expectations from everyday experiences: This $T$ diverges as $E \to 0$ for $n_+ \approx n_-$, and it is negative whenever $E > 0$ (corresponding to $n_+ > n_-$).[9] When $H = 0$, we also have $E = 0$ and $T$ is ill-defined. Restricting our attention to $H > 0$, with $n_+ < n_-$ producing a positive temperature, we also see that this temperature cannot vanish. It is minimized by the most-negative energy you found above, $T_{\text{min}} = H > 0$ for $E_{\text{min}} = -NH$. The non-zero minimum temperature is specific to spin systems, while some of the other oddities result from the micro-canonical approach more generally. This will motivate turning to the canonical ensemble next week, but first we can check that some aspects of the temperature defined in Eq. 23 do match our everyday expectations. These include the temperature being *intensive* (independent of $N$) since $\langle E\rangle \propto N$ from Eq. 15, as well as its implications for heat exchange.

[9] You can check that $T < 0$ whenever the number of micro-states decreases for larger internal energies, $\frac{\partial M}{\partial E} < 0$. In so-called 'natural' systems, larger energies 'open up' more possible micro-states, producing $\frac{\partial M}{\partial E} > 0$ and a positive temperature.

## 2.4 Heat exchange

In the 'natural' positive-temperature regime of the spin system considered above, we see that larger (i.e., less-negative) energies produce higher temperatures. This direct relation between energy and temperature is very generic, and we will study it in more detail when considering thermodynamic cycles in a few weeks. For now, considering unspecified systems that exhibit this behaviour, let's ask what would happen if we take two initially isolated micro-canonical systems ($\Omega_A$ and $\Omega_B$ with temperatures $T_A$ and $T_B$ in thermodynamic equilibrium) and bring them into thermal contact.

In micro-canonical terms, the temperatures $T_A$ and $T_B$ are derived from the corresponding energies $E_A$ and $E_B$, while thermal contact allows the two systems to exchange energy (but not particles) as non-isolated subsystems of a combined micro-canonical system $\Omega_C$. Once the two subsystems have been in thermal contact long enough for the combined system to have reached thermodynamic equilibrium, it will have temperature $T_C$. We can then re-isolate the two subsystems, which will now have energies $E'_A$ and $E'_B$ with (in thermodynamic equilibrium) temperatures $T'_A$ and $T'_B$. This three-step procedure is illustrated below.




From everyday experience, we expect that this energy exchange will result in a net flow of energy from the hotter system to the colder system, cooling the

former by heating the latter. We will now check that the micro-canonical definition of temperature in Eq. 23 predicts this expected behaviour. We define

$$E_i' = E_i + \Delta E_i \qquad \left|\frac{\Delta E_i}{E_i}\right| \ll 1 \qquad i \in \{A, B\}\,,$$

for simplicity considering the case where the change in energy is relatively small (suggesting that $E_A$ and $E_B$ are not extremely different). We also know $\Delta E_B = -\Delta E_A$ thanks to conservation of energy.

Equation 23 tells us that we need to consider the entropies as functions of $E_i$ and $E_i'$ in order to connect the temperatures to any flow of energy. (Because we don't change the number of particles in each system, we only need to consider the energy dependence of the entropy.) We assume that we can expand each of the final entropies $S(E_i')$ in a Taylor series,

$$S(E_i') = S(E_i + \Delta E_i) \approx S(E_i) + \left.\frac{\partial S}{\partial E}\right|_{E_i} \Delta E_i,$$

neglecting all $\mathcal{O}\left(\Delta E_i^2\right)$ terms because we consider relatively small changes in energy.[10] What is the linearized Taylor series above in terms of the initial temperatures $T_i$?

From the second law of thermodynamics, we know that the total entropy of these systems can never decrease as time passes:

$$S(E_A') + S(E_B') = S(E_A + E_B) \geq S(E_A) + S(E_B). \tag{27}$$

The fact that re-isolating the two subsystems doesn't change the entropy may be surprising—it's tempting to guess that this procedure might reduce the entropy, by reversing the argument used when the two subsystems were brought into thermal contact in Section 2.2.3. The key is that we don't fix $E_A'$, which could have any value from zero to $E_A + E_B$ at the moment when the subsystems are re-isolated. Computing the final entropy $S(E_A') + S(E_B')$ therefore requires summing over all possible values of $E_A'$, producing exactly the sum in Eq. 22 for the overall system. We will see something similar when we consider 'the Gibbs paradox' in a couple of weeks.

[10] The assumption here is that $S(E)$ is continuous and infinitely differentiable, which is not the case at a *first-order phase transition*, where we would need to be more careful. We will learn about phase transitions towards the end of the term.

What do you find when you insert your linearized Taylor series into Eq. 27?

Applying conservation of energy should produce

$$\left(\frac{1}{T_A} - \frac{1}{T_B}\right) \Delta E_A \geq 0.$$

Recalling from Section 2.2.3 that equality holds only in extremely special cases, we can identify three possibilities consistent with this result. If $T_A > T_B$, then $\left(\frac{1}{T_A} - \frac{1}{T_B}\right)$ is negative and we will generically have $\Delta E_A < 0$, so that energy flows out of the hotter system $\Omega_A$ and into the colder one. Similarly, if $T_A < T_B$, we will generically have $\Delta E_A > 0$, and energy still flows from the hotter system $\Omega_B$ into the colder one. We can finally conclude that $T_A = T_B$ is the very special case where there is no energy flow, $\Delta E_i = 0$. All of this is exactly what we expected based on everyday experience.

# Week 3: Canonical ensemble

## 3.1 The thermal reservoir

### 3.1.1 Replicas and occupation numbers

While it is relatively easy to prevent particle exchange, for example by sealing gases inside airtight containers, it is not practical to forbid energy exchange as would be needed to fully isolate statistical systems. Any thermal insulator is imperfect, and even in the deepest reaches of space we would still be bombarded by cosmic microwave radiation. In practice it is more convenient to work with physical systems that are characterized by their (intensive) temperatures rather than their (extensive) internal energies.

This leads us to define a **canonical ensemble** to be a statistical ensemble characterized by its fixed temperature $T$ and conserved particle number $N$, with the temperature held fixed through contact with a **thermal reservoir**.

The second part of this definition connects the fixed temperature to the fundamental fact of energy conservation (the first law of thermodynamics). This is done by proposing that our system of interest $\Omega$ is in thermal contact with a much larger external system $\Omega_{\text{res}}$—the thermal reservoir, sometimes called a "heat bath". The overall combined system $\Omega_{\text{tot}} = \Omega_{\text{res}} \otimes \Omega$ is governed by the micro-canonical ensemble, with conserved total energy $E_{\text{tot}} = E_{\text{res}} + E \approx E_{\text{res}}$, while the energy $E$ of $\Omega$ is allowed to fluctuate. The key qualitative idea is that, in thermodynamic equilibrium, $\Omega$ has a negligible effect on the overall system. In particular, the temperature of that overall system—and therefore the temperature of $\Omega$, by intensivity—is set by the reservoir and remains fixed even as $E$ fluctuates. This effectively generalizes the setup we used to analyze heat exchange last week, where we saw that thermal contact causes a net flow of energy from hotter systems to colder systems, cooling the former by heating the latter.

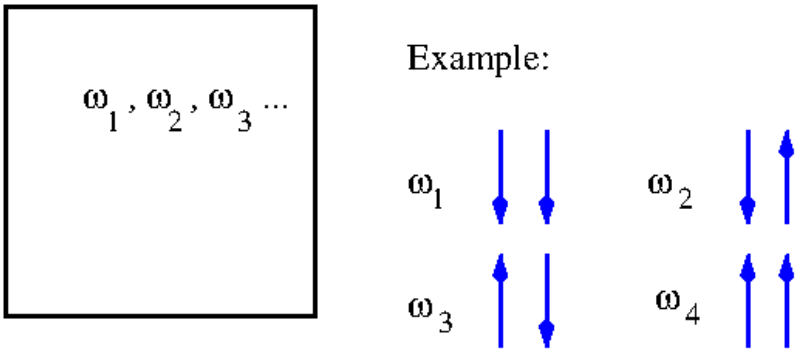
The mathematical realization of this argument, as developed by J. Willard Gibbs, proceeds by considering a well-motivated ansatz for the form of the thermal reservoir $\Omega_{\text{res}}$. The goal, which will be useful to keep in mind as we go through the lengthy analysis, is to show that the specific form of $\Omega_{\text{res}}$ is ultimately irrelevant. This will allow us to work directly with the system of interest, $\Omega$, independent of the details of the thermal reservoir that fixes its temperature.

Without further ado, we take $\Omega_{\text{tot}}$ to consist of many ($R \gg 1$) identical **replicas** of the system $\Omega$ that we're interested in. All of these replicas are in thermal contact with each other, and in thermodynamic equilibrium.[11] Choosing one of

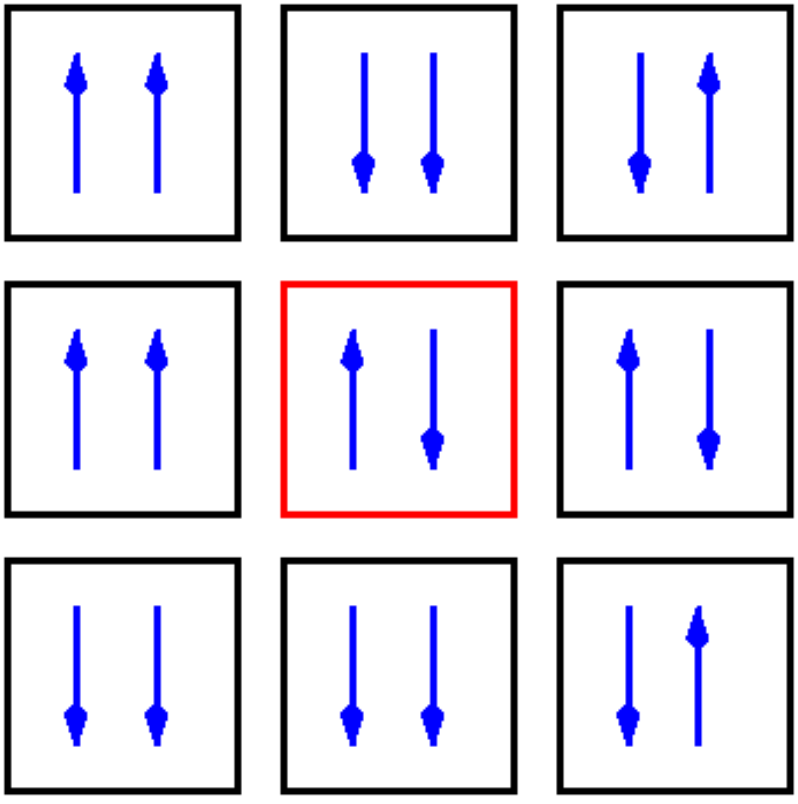
[11] The thermal contact between any two replicas can be indirect, mediated by a sequence of intermediate replicas. This transitivity of thermodynamic equilibrium is sometimes called the zeroth law of thermodynamics. It declares that if systems $\Omega_A$ & $\Omega_B$ are in thermodynamic equilibrium while systems $\Omega_B$ & $\Omega_C$ are in thermodynamic equilibrium, then $\Omega_A$ & $\Omega_C$ must also be in thermodynamic equilibrium.

the replicas to be the system of interest, $\Omega$, the other $R-1 \gg 1$ replicas provide the thermal reservoir $\Omega_{\text{res}}$. Assuming we want to study reasonable systems $\Omega$, this ansatz ensures that $\Omega_{\text{res}}$ is also reasonable, simply much larger.

An extremely small example of this setup is illustrated by the figures below, where the system of interest consists of only $N = 2$ spins. For now we assume the spins are *distinguishable*, so that $\downarrow\uparrow$ and $\uparrow\downarrow$ are both distinct micro-states. This means that each individual replica has only the $M = 4$ micro-states $\omega_i$ defined below.




To form the overall system $\Omega_{\text{tot}}$ we now bring together the $R = 9$ replicas shown below. We draw boxes around each replica to remind us that they are allowed to exchange only energy with each other, while the $N = 2$ spins are conserved in each replica. We pick out one of these replicas (coloured red) to serve as the system $\Omega$ we will consider. The other $8$ are the thermal reservoir $\Omega_{\text{res}}$ that fixes the temperature of $\Omega$.



A convenient way to analyze the overall system of $R$ replicas, $\Omega_{\text{tot}}$, is to define the **occupation number** $n_i$ to be the number of replicas that adopt the micro-state $\omega_i \in \Omega$ in any given micro-state of $\Omega_{\text{tot}}$. The index $i \in \{1, 2, \cdots, M\}$ runs over all $M$ micro-states of $\Omega$. In the example above, three of the replicas in the second figure have the micro-state $\omega_1 = \downarrow\downarrow$, meaning $n_1 = 3$. What are the occupation numbers $\{n_2, n_3, n_4\}$ for the other three $\omega_i$ in the figures above? Are all replicas are accounted for, $\sum_i n_i = R$?

Normalizing the occupation number by $R$ gives us a well-defined *occupation probability*, $p_i = n_i/R$ with $\sum_i p_i = 1$. This $p_i$ is the probability that if we choose a replica at random it will be in micro-state $\omega_i$.

Now let us consider conservation of energy, which continues to apply to the total energy $E_{\text{tot}}$ of the overall system $\Omega_{\text{tot}}$. We assume that each replica's energy $E_r$ is independent of all the other replicas. This is guaranteed for the non-interacting systems we will focus on until week $10$, and also holds when interactions are allowed within each replica but not between different replicas. The thermal contact between replicas allows $E_r$ to fluctuate (subject to conservation of $E_{\text{tot}}$), but there are only $M$ possible values $E_i$ it can have, corresponding to the $M$ micro-states $\omega_i \in \Omega$. This allows us to rearrange a sum over replicas into a sum over the micro-states of $\Omega$:

$$E_{\text{tot}} = \sum_{r=1}^{R} E_r = \sum_{i=1}^{M} n_i E_i, \tag{28}$$

with the occupation number $n_i$ counting how many times micro-state $\omega_i$ appears among the $R$ replicas. We can assume that $R$ and $M$ are both finite, so we don't need to worry about rearranging infinite sums.

### 3.1.2 Partition function

Following Gibbs, we have already taken the thermal reservoir $\Omega_{\text{res}}$ to consist of $R - 1$ replicas of the system of interest, $\Omega$. The next step is to further simplify the mathematics by assuming that the overall $R$-replica system $\Omega_{\text{tot}}$ is fully specified by a fixed set of $M$ occupation numbers $\{n_i\}$. From Eq. 28, we see that this ensures conservation of the total energy $E_{\text{tot}}$, and we can apply the microcanonical tools we developed last week. Recall our ultimate goal of showing that such details of the thermal reservoir are irrelevant to the system $\Omega$.

Based on the conservation of $E_{\text{tot}}$, we want to determine the (intensive) temperature of $\Omega_{\text{tot}}$, which fixes the temperature of the system of interest, $\Omega$. According to our work last week, to do this we first need to compute the overall number

of micro-states $M_{\text{tot}}$ as a function of $E_{\text{tot}}$, from which we can derive the entropy and temperature since the system is in thermodynamic equilibrium. From the fixed occupation numbers $n_i$, we already know how many times each micro-state $\omega_i$ appears among the $R$ replicas. To determine $M_{\text{tot}}$ we just need to count how many possible ways there are of distributing the $\{n_i\}$ micro-states among the $R$ replicas.

If we consider first the micro-state $\omega_1$, the number of possible ways of distributing $n_1$ copies of this micro-states among the $R$ replicas is just the binomial coefficient

$$\binom{R}{n_1} = \frac{R!}{n_1!\,(R-n_1)!}.$$

Moving on to $\omega_2$, we need to keep in mind that $n_1$ replicas have already been assigned micro-state $\omega_1$, so there are only $R - n_1$ replicas left to choose from. What is the resulting number of possible ways of distributing these $n_2$ micro-states?

Repeating this process for all micro-states $\{\omega_1, \omega_2, \cdots, \omega_M\}$, and recalling that $(R - \sum_i n_i)! = 0! = 1$, you should obtain a product that 'telescopes' to

$$M_{\text{tot}} = \frac{R!}{n_1!\; n_2!\; \cdots\; n_M!}. \tag{29}$$

From this we can see that the order in which we assign micro-states to replicas is irrelevant, since integer multiplication is commutative.

Thanks to thermodynamic equilibrium, the entropy of the micro-canonical overall system $\Omega_{\text{tot}}$ is

$$S(E_{\text{tot}}) = \log M_{\text{tot}} = \log(R!) - \sum_{i=1}^{M} \log(n_i!),$$

where the dependence on $E_{\text{tot}}$ enters through the occupation numbers via Eq. 28. With $R \gg 1$ and $n_i \gg 1$ for all $i = 1, \cdots, M$, we can approximate each of these logarithms using the first two terms in Stirling's formula,

$$\log(N!) = N \log N - N + \mathcal{O}(\log N) \approx N \log N - N \qquad \text{for } N \gg 1.$$

Back in Eq. 25 we used the central limit theorem to derive a form of this approximation that included a leading term from the $\mathcal{O}(\log N)$ we neglect here. In order for *every* occupation number to be large, $n_i \gg 1$, the number of replicas must be much larger than the number of micro-states of $\Omega$, so that $R \gg M$. As we have discussed before, the number of micro-states $M$ is typically a very large number, so we are formally considering truly enormous thermal reservoirs! This enormity helps ensure that the detailed form of the reservoir will be irrelevant.

Applying the approximation above, what do you find for $S(E_{\text{tot}})$ in terms of $R$ and $n_i$? What is the entropy in terms of the occupation probabilities $p_i = n_i/R$?

$$S(E_{\text{tot}}) = \log(R!) - \sum_{i=1}^{M} \log(n_i!) \approx$$

In your result, the dependence on $E_{\text{tot}}$ now enters through the occupation probabilities $p_i$. In order to determine the temperature, we have to express $S(E_{\text{tot}})$ directly in terms of $E_{\text{tot}}$. We do this by applying our knowledge from last week that thermodynamic equilibrium implies maximal entropy.

Following the same steps as last week, we maximize the entropy, including two Lagrange multipliers to account for the two constraints on the occupation probabilities:

$$\sum_{i=1}^{M} p_i = 1 \qquad\qquad \sum_{i=1}^{M} n_i E_i = R \sum_{i=1}^{M} p_i E_i = E_{\text{tot}}.$$

Writing everything in terms of occupation probabilities we therefore need to maximize the modified entropy

$$\overline{S} = -R \sum_{i=1}^{M} p_i \log p_i + \alpha \left( \sum_{i=1}^{M} p_i - 1 \right) - \beta \left( R \sum_{i=1}^{M} p_i E_i - E_{\text{tot}} \right).$$

(The sign of $\beta$ is irrelevant, and chosen for later convenience.) What is the occupation probability $p_k$ that maximizes $\overline{S}$?

$$0 = \frac{\partial \overline{S}}{\partial p_k} =$$

You should find a probability

$$p_k = \frac{1}{Z} e^{-\beta E_k}, \tag{30}$$

where we define $Z = \exp\left[1 - \frac{\alpha}{R}\right]$ to put the result in its traditional form. In place of $\{\alpha, \beta\}$, our free parameters are now $\{Z, \beta\}$. Still following last week's procedure, we need to fix these free parameters by demanding that the two constraints above are satisfied. The first of these constraints is straightforward and produces an important result:

$$1 = \sum_{i=1}^{M} p_i = \frac{1}{Z} \sum_{i=1}^{M} e^{-\beta E_i} \qquad \Longrightarrow \qquad Z(\beta) = \sum_{i=1}^{M} e^{-\beta E_i}. \tag{31}$$

Equation 31 defines the canonical **partition function** $Z(\beta)$, a fundamental quantity in the canonical ensemble, from which many other derived quantities can be obtained.

$Z(\beta)$ still depends on the other as-yet-unknown free parameter $\beta(E_{\text{tot}})$. If we apply our second constraint, Eq. 28, we can relate $\beta$ to $E_{\text{tot}}$:

$$E_{\text{tot}} = R \sum_{i=1}^{M} p_i E_i = \frac{R}{Z(\beta)} \sum_{i=1}^{M} E_i \, e^{-\beta E_i} = R \frac{\sum_{i=1}^{M} E_i \, e^{-\beta E_i}}{\sum_{i=1}^{M} e^{-\beta E_i}}. \tag{32}$$

This is a bit opaque, but will suffice for our goal of expressing the entropy in terms of $E_{\text{tot}}$. Inserting Eq. 30 for $p_i$ into your earlier result for the entropy, what do you obtain upon applying Eqs. 31 and 32?

$$S(E_{\text{tot}}) = -R \sum_{i=1}^{M} p_i \log p_i =$$

There is a pleasant simplification when we take the derivative to determine the temperature. Defining $\beta' = \frac{\partial}{\partial E_{\text{tot}}} \beta(E_{\text{tot}})$, we have

$$\frac{1}{T} = \frac{\partial}{\partial E_{\text{tot}}} S(E_{\text{tot}}) = \frac{\partial}{\partial E_{\text{tot}}} \left[ E_{\text{tot}} \beta + R \log Z(\beta) \right] = \beta + E_{\text{tot}} \beta' + R \frac{1}{Z} \frac{\partial Z(\beta)}{\partial \beta} \beta'.$$

Using Eq. 32 we can recognize

$$\frac{1}{Z} \frac{\partial Z(\beta)}{\partial \beta} = \frac{1}{Z} \frac{\partial}{\partial \beta} \sum_{i=1}^{M} e^{-\beta E_i} = -\frac{1}{Z} \sum_{i=1}^{M} E_i \, e^{-\beta E_i} = -\frac{E_{\text{tot}}}{R},$$

so that we don't need to figure out the explicit form of $\beta'$:

$$\frac{1}{T} = \beta + E_{\text{tot}}\beta' - E_{\text{tot}}\beta' = \beta. \tag{33}$$

What's truly remarkable about Eq. 33 is that all the details of the thermal reservoir have vanished—there is no reference to the $R$ replicas or any extensive quantity such as $E_{\text{tot}}$. This is the goal we have been pursuing since the start of the notes for this week! The large thermal reservoir is still present to fix the temperature $T$ characterizing the canonical system $\Omega$, but beyond that nothing about it is relevant—or even knowable in the canonical approach. Every aspect of $\Omega$ can now be specified in terms of its fixed temperature $T$ and conserved particle number $N$, starting with the parameter $\beta = 1/T$.

In particular, the partition function from Eq. 31 is simply

$$Z(T) = \sum_{i=1}^{M} e^{-E_i/T}. \tag{34}$$

and together with $\beta$ specifies the probabilities

$$p_i = \frac{1}{Z} e^{-E_i/T} \tag{35}$$

from Eq. 30. This $p_i$ is now the probability—in thermodynamic equilibrium—that $\Omega$ adopts micro-state $\omega_i$ with (non-conserved) internal energy $E_i$. This probability distribution is called either the **Boltzmann distribution** or the **Gibbs distribution**, while $e^{-E_i/T}$ itself is known as a **Boltzmann factor**. All micro-states with the same energy have the same probability in thermodynamic equilibrium, which is consistent with the micro-canonical behaviour we saw last week.

## 3.2 Internal energy, heat capacity, and entropy

In addition to fixing the temperature of the system $\Omega$, the thermal reservoir also allows the internal energy of $\Omega$ to fluctuate, by exchanging energy with this reservoir. Although the internal energy fluctuates, its expectation value $\langle E \rangle$ is an important derived quantity in thermodynamic equilibrium. Applying the general definition from Eq. 4 to the probability space of the canonical ensemble,

$$\langle E \rangle(T) = \sum_{i=1}^{M} E_i \, p_i = \frac{1}{Z} \sum_{i=1}^{M} E_i \, e^{-\beta E_i}.$$

Here we highlight the dependence of $\langle E \rangle$ on the temperature, and also freely interchange $\beta = 1/T$, in part because the last expression may look familiar from our work above:

$$\frac{\partial}{\partial \beta} \log Z =$$

In this case it is easier to take the derivative with respect to $\beta$ as opposed to

$$\frac{\partial}{\partial \beta} = \frac{\partial T}{\partial \beta} \frac{\partial}{\partial T} = -\frac{1}{\beta^2} \frac{\partial}{\partial T} = -T^2 \frac{\partial}{\partial T}. \tag{36}$$

Last week, we saw that 'natural' micro-canonical systems exhibit higher (derived) temperatures for larger (conserved) internal energies. Now, in the canonical approach, the average internal energy $\langle E \rangle$ is the derived quantity while the temperature is fixed. From our everyday experience, we expect a similar direct relation between temperature and energy, which the following result confirms.

The **heat capacity** is defined to be

$$c_v = \frac{\partial}{\partial T} \langle E \rangle , \tag{37}$$

and is always non-negative, $c_v \geq 0$.

In a homework assignment you will show that $c_v \geq 0$,[12] by deriving a so-called **fluctuation–dissipation** (or **fluctuation–response**) **relation**. That relation will be a special case of a more general theorem, and will connect the fluctuations of the internal energy around its expectation value, $\sum_i (E_i - \langle E \rangle)^2$, to the energy's *response* to a change in temperature, $\frac{\partial}{\partial T} \langle E \rangle$. Equality will hold only in extremely special cases, meaning that the heat capacity is generically positive, which confirms our expectation from everyday experiences that higher temperatures produce larger (average) internal energies.

[12] The subscript indicates that the volume of the system is kept fixed; details will wait until we properly introduce the volume over the next couple of weeks.

We finally need to compute the entropy of $\Omega$ itself, still in thermodynamic equilibrium but with no reference to the thermal reservoir apart from its role fixing $T$. Since the definition of the entropy in Eq. 21 continues to hold for the canonical ensemble, we just need to insert the probabilities $p_i$ from Eq. 35:

$$S(T) = -\sum_{i=1}^{M} p_i \log p_i =$$

You should find

$$S(T) = \frac{\langle E \rangle}{T} + \log Z. \tag{38}$$

## 3.3 Helmholtz free energy

Last week we saw that the micro-canonical entropy in thermodynamic equilibrium is the logarithm of the number of micro-states. Now Eq. 38 features a similar logarithm of the partition function (Eq. 34), which is a sum over all micro-states that we identified as a fundamental quantity in the canonical ensemble. This motivates the following definition of a quantity with the dimensions of energy that is related to $\log Z$, which provides simpler and more elegant expressions for the derived quantities we considered above.

The **Helmholtz free energy** of a canonical ensemble is defined to be

$$F(T) = -T \log Z(T) \qquad\qquad F(\beta) = -\frac{\log Z(\beta)}{\beta}, \tag{39}$$

where $Z$ is the partition function of the ensemble. In terms of this free energy, Eqs. 34 and 35 are

$$Z = e^{-F/T} \qquad\qquad p_i = e^{(F-E_i)/T}.$$

The Helmholtz free energy is named after Hermann von Helmholtz and reveals its usefulness when we take its derivative. The derivative involves $\frac{\partial}{\partial T}\log Z$, which is worth collecting in advance based on Eq. 36:

$$-\frac{\partial}{\partial T}\left(\frac{F(T)}{T}\right) = \frac{\partial}{\partial T}\log Z(T) =$$

$$\frac{\partial}{\partial T}F(T) =$$

From these results and Eq. 36 we can read off the more elegant expressions promised above:

$$S(T) = -\frac{\partial}{\partial T}F(T) \tag{40}$$

$$\langle E\rangle(T) = -T^2\frac{\partial}{\partial T}\left(\frac{F(T)}{T}\right) = \frac{\partial}{\partial \beta}\left[\beta F(\beta)\right] = TS(T) + F(T). \tag{41}$$

## 3.4 The physics of information

As a first application of the canonical ensemble, we will explore physically observable effects that depend on the pure information content of a statistical system. A famous topic related to these effects is the black hole information loss paradox, but this example is well beyond the scope of this module since it involves quantum mechanics and general relativity in addition to statistical physics. Here we will consider simple spin systems as introduced last week, contrasting the behaviour of their average internal energy $\langle E\rangle$ and entropy $S$ depending on whether or not the spins can (in principle) be distinguished from each other. It's important to appreciate that the "information" discussed here is an intrinsic property of the system—what is *knowable* about it in principle. It does not matter whether or not any observer actually knows this information; so long as it can possibly be known it will have an effect.

### 3.4.1 Distinguishable spins in a solid

We begin with the setup introduced last week: A system of $N$ spins arranged in a line, placed in an external magnetic field of strength $H$, and in thermodynamic equilibrium. We further specify that the spins are embedded in a solid material that fixes their positions and prevents them from moving. This allows them to be distinguished from one another: An observer can target an appropriate position

in the solid to measure the corresponding spin. The spins distinguished in this way will be aligned either anti-parallel or parallel to the magnetic field. The canonical system therefore has $M = 2^N$ distinct micro-states $\omega_i$ with energies $E_i$ and probabilities $p_i = \frac{1}{Z} e^{-E_i/T}$, each defined by the orientations of all $N$ spins.

Building on our work last week, we can represent the orientation of the $n$th spin as $s_n \in \{1, -1\}$, where $s_n = 1$ indicates alignment anti-parallel to the field and $s_n = -1$ indicates alignment parallel to the field. Since the spins don't interact with each other, the internal energy of the system in micro-state $\omega_i$ specified by the $N$ spins $\{s_n\}$ is therefore

$$E_i = H \sum_{n=1}^{N} s_n. \tag{42}$$

To compute the partition function $Z_D$ (with the subscript reminding us about the spins' distinguishability), we have to sum over all possible spin configurations $\{s_n\}$. In this process we can save some space by defining the dimensionless variable $x = \beta H = \frac{H}{T}$:

$$\begin{aligned} Z_D &= \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} e^{-\beta E_i} = \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} \exp\left[-x \sum_{n=1}^{N} s_n\right] \\ &= \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} e^{-x s_1} \cdots e^{-x s_N} = \left(\sum_{s_1=\pm 1} e^{-x s_1}\right) \cdots \left(\sum_{s_N=\pm 1} e^{-x s_N}\right) \\ &= \left(\sum_{s=\pm 1} e^{-xs}\right)^N = \left(e^x + e^{-x}\right)^N = \left[2\cosh(\beta H)\right]^N, \end{aligned} \tag{43}$$

distributing the summations since all the spins are independent of each other.

The corresponding Helmholtz free energy

$$F_D(\beta) = -\frac{N \log\left[2\cosh(\beta H)\right]}{\beta} \tag{44}$$

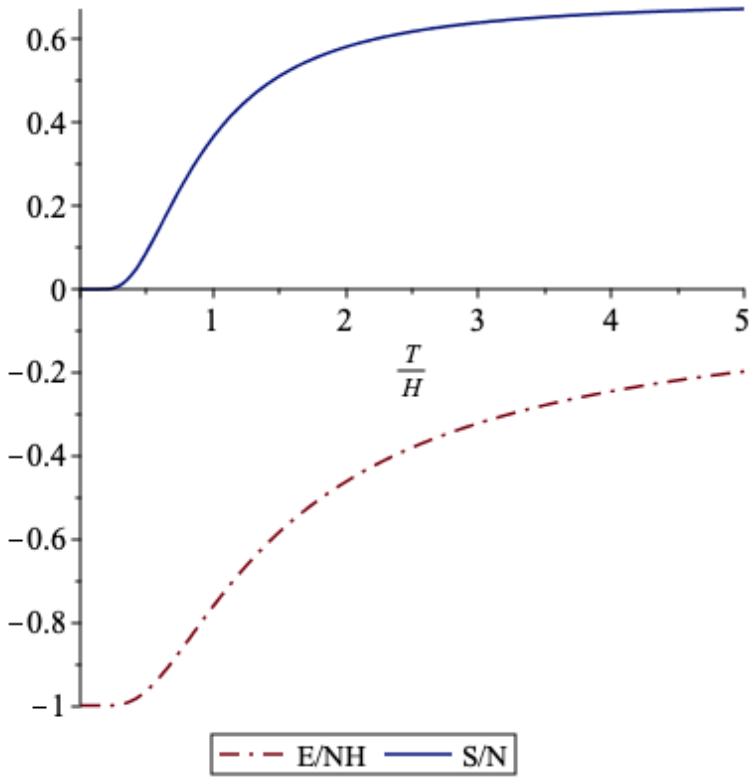
is all we need to compute the average internal energy:

$$\langle E \rangle_D = \frac{\partial}{\partial \beta}\left[\beta F_D(\beta)\right] =$$

From this we immediately obtain the entropy

$$S_D = \beta\left(\langle E \rangle_D - F_D\right) = -N\beta H \tanh(\beta H) + N \log\left[2\cosh(\beta H)\right]. \tag{45}$$

These results for $\langle E \rangle_D$ and $S_D$ are plotted below as functions of $\frac{T}{H} = \frac{1}{\beta H}$. Since both these quantities are extensive, we show them in natural units, $\frac{\langle E \rangle_D}{NH}$ and $\frac{S_D}{N}$.




Let's check the asymptotic behaviour of these functions, starting with **low temperatures**. In the canonical ensemble, there is no issue with taking the independent variable $T \to 0$. This corresponds to $\beta H \to \infty$ and $\tanh(\beta H) \to 1$, approaching the "ground-state" energy $E_{\text{min}} = E_0 = -NH$ you computed last week, which is realized only by the single micro-state in which all the spins are aligned with the magnetic field (every $s_n = -1$). At the same time, $\log[2\cosh(\beta H)] \to \log e^{\beta H} = \beta H$ and the two terms in Eq. 45 cancel out, so that $S_D \to 0$. This vanishing entropy is a generic consequence of temperatures approaching *absolute zero*.

Even at low temperatures, $\langle E \rangle_D$ and $S_D$ will be affected by the non-zero probability for the system to adopt higher-energy "excited states". Last week you also computed the energy $E_1 = -(N-2)H$ of the first excited state, which is realized by $N$ distinct micro-states. The *energy gap* between the ground state and the first excited state is $\Delta E = E_1 - E_0 = 2H$. (For this system all energy levels are separated by the same $E_{n_+ +1} - E_{n_+} = 2H$, which we will use below.)

We can see the effects of the higher-energy states at low temperatures $\beta H \gg 1$ by expanding $\langle E \rangle_D$ in powers of $e^{-\beta H} \ll 1$. What is the first $\frac{T}{H}$-dependent term in this expansion?

$$\frac{\langle E \rangle_D}{NH} =$$

You should find that the excited-state effects are *exponentially* suppressed by the energy gap $\Delta E$ at low temperatures,

$$\frac{\langle E \rangle_D}{NH} = -1 + 2e^{-\beta \Delta E} + \mathcal{O}\left(e^{-2\beta \Delta E}\right).$$

This is a generic feature of systems with a non-zero energy gap, and is due to the exponentially small probability $\propto e^{-\beta \Delta E}$ of the system adopting any of the microstates with the higher energy.

The low-temperature expansion of Eq. 45 for the entropy $S_D$ in powers of $e^{-\beta H} \ll 1$ is similar:

$$\frac{S_D}{N} =$$

Here the exponential suppression in the leading term is slightly counteracted by a linear factor of $\beta \Delta E$,

$$\frac{S_D}{N} = \beta \Delta E e^{-\beta \Delta E} + e^{-\beta \Delta E} + \mathcal{O}\left(e^{-2\beta \Delta E}\right).$$

In the limit of **high temperatures** the expansion of $\langle E \rangle_D$ and $S_D$ in powers of $\beta H \ll 1$ is more straightforward. The $\tanh$ in $\langle E \rangle_D$ can be expanded as usual,

$$\frac{\langle E \rangle_D}{NH} = -\tanh(\beta H) = -\beta H + \frac{(\beta H)^3}{3} + \mathcal{O}\left([\beta H]^5\right),$$

and vanishes $\sim\frac{1}{T}$ as $T \to \infty$. This matches the micro-canonical behaviour we saw for this system last week, where the derived temperature diverged as the conserved energy approached zero.

For the entropy, there is a similar connection to micro-canonical behaviour at high temperatures:

$\frac{S_D}{N} =$

As $\frac{T}{H} \to \infty$, the result

$$\frac{S_D}{N} = \log 2 - \frac{(\beta H)^2}{2} + \mathcal{O}\left([\beta H]^4\right)$$

approaches the asymptotic value $S_D \to N \log 2 = \log M$ for the $M = 2^N$ micro-states (with different energies). Qualitatively, in this limit the energy of each spin is negligible compared to the temperature, and the system approximately behaves as though the energy were zero for all micro-states (and hence conserved).

### 3.4.2 Indistinguishable spins in a gas

We now consider nearly the same setup, with $N$ spins in thermodynamic equilibrium, in an external magnetic field of strength $H$. The only difference is that now the spins are allowed to move, like particles in a one-dimensional gas. We demand that they move slowly so that we can neglect their kinetic energy and the total energy of the system continues to be given by Eq. 42. Since the spins don't interact with each other, they can freely move past each other, and even occupy the same space, making it impossible for them to be distinguished from one another in any way.

To define the fundamental canonical partition function (Eq. 34), we have to sum over the micro-states of the system. These micro-states are no longer in one-to-one correspondence with the full configurations $\{s_n\}$ of the $N$ spins. Because the spins are now indistinguishable, certain spin configurations also cannot be distinguished from each other. The simplest example comes from the two-spin system considered in Section 3.1.1, where the configurations $\downarrow\uparrow$ and $\uparrow\downarrow$ now both map onto a single micro-state. In this micro-state, we know only that one spin is $s_i = 1$ while the other is $s_k = -1$; it's not possible to distinguish which is which.

What remains observable is the energy of the micro-state. Because permutations of the spin configurations don't change the internal energy (the sum in

Eq. 42 is commutative), we can conclude that all such permutations correspond to a single distinct micro-state. For this particular system, we saw above that the energies are best described as *energy levels* separated by a constant gap $\Delta E = 2H$. As a quick example, enumerate the energy levels when $N = 4$ and list the spin configurations associated with the corresponding micro-states. How many micro-states are there for $N$ spins?

A convenient way to label these micro-states and energy levels is to define

$$E_k = -NH + 2Hk$$

for micro-state $\omega_k$ with $k = 0, \cdots N$. (Previously we used the label $n_+$ instead of $k$.) To compute the partition function $Z_I$ (with the subscript reminding us about the spins' indistinguishability), we simply sum

$$Z_I = \sum_{k=0}^{N} e^{-\beta E_k} = \sum_{k=0}^{N} e^{\beta H(N-2k)} = e^{N\beta H} \sum_{k=0}^{N} \left(e^{-2\beta H}\right)^k = e^{N\beta H} \frac{1 - e^{-2(N+1)\beta H}}{1 - e^{-2\beta H}}. \quad (46)$$

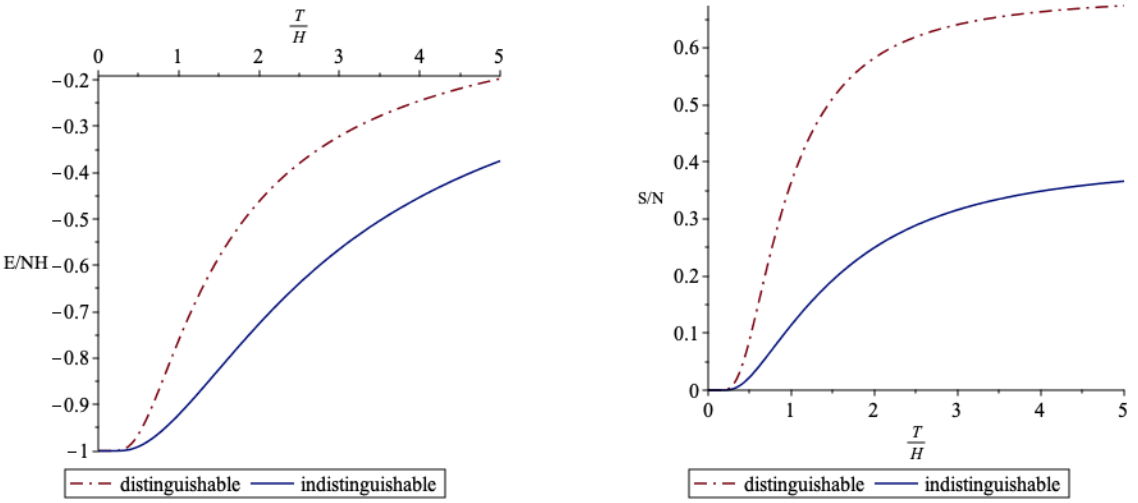
The geometric series in the last step can be reconstructed by considering

$$\sum_{k=0}^{N} x^k = \sum_{k=0}^{\infty} x^k - \sum_{k=N+1}^{\infty} x^k = \frac{1}{1-x} - x^{N+1} \sum_{\ell=0}^{\infty} x^\ell = \frac{1}{1-x} - \frac{x^{N+1}}{1-x}.$$

The corresponding Helmholtz free energy is

$$F_I(\beta) = -NH - \frac{\log\left[1 - e^{-2(N+1)\beta H}\right]}{\beta} + \frac{\log\left[1 - e^{-2\beta H}\right]}{\beta}. \quad (47)$$

In contrast to Eq. 44, $F_i(\beta)$ is no longer proportional to $N$. In a homework assignment you will use $F_I$ to determine the average internal energy $\langle E \rangle_I$ and entropy $S_I$ shown in the figures below, and also check the low- and high-temperature expansions like we did for the distinguishable case above. Unlike our results for the distinguishable case, you will find that $\frac{\langle E \rangle_I}{NH}$ and $\frac{S_I}{N}$ depend on $N$, which requires us to fix $N = 4$ in the plots below.




The dash-dotted lines in these figures are exactly the distinguishable-spin results shown above. The solid lines are the new results for indistinguishable spins. We see that the same $T \to 0$ limits are approached in both cases: $E \to -NH$ and $S \to 0$. At low temperatures, the indistinguishable results approach these limits more quickly—they still feature exponential suppression of excited-state effects by the energy gap, $e^{-\beta\Delta E}$, but this now comes with additional factors of $N$.

At high temperatures there is an even more striking difference. While the average internal energy $\langle E \rangle_I$ continues to vanish $\sim \frac{1}{T}$ as $T \to \infty$ (with different $N$ dependence), the entropy approaches the asymptotic value $S_I \to \log(N+1) = \log M$ for the $M = N+1$ micro-states. This logarithmic dependence on $N$ is very different from the $S_D \to N \log 2$ limit we found for distinguishable spins, and reflects the exponentially smaller number of micro-states that exist for indistinguishable spins, $N+1$ vs. $2^N$.

Finally, away from those low- and high-temperature limits, the results above show a significant difference in the internal energy of the spin systems, depending only on whether or not the spins can be distinguished from each other in principle. This is a physically measurable effect caused by the intrinsic information content of a statistical system, and a simple illustration of phenomena that remain at the leading edge of ongoing research. The conclusion was pithily stated by Rolf Landauer in 1991: “Information is physical.”

# Week 4: Ideal gases

## 4.1 Volume, energy levels, and partition function

This week we apply the canonical ensemble to investigate non-relativistic, classical, ideal gases. Using statistical physics we will explore how the large-scale behaviours of such gases emerge from the properties of the particles that compose them. The three adjectives above specify these properties in the case we consider this week:

- **Classical** systems are those for which we can ignore the effects of quantum mechanics. At a technical level, this means we assume we can simultaneously define both the position $(x, y, z)$ and the momentum $\vec{p} = (p_x, p_y, p_z)$ of each particle with arbitrary precision.
- **Non-relativistic** particles move with speeds small compared to the speed of light, which allows us to ignore small effects due to special relativity. The particles are therefore governed by the laws Isaac Newton published all the way back in 1687. In particular, the energy of each particle of mass $m$ is

$$E_n = \frac{1}{2m} p_n^2,$$

  where $p_n^2 = \vec{p}_n \cdot \vec{p}_n = (p_x)_n^2 + (p_y)_n^2 + (p_z)_n^2$ is the inner (or 'dot') product of the momentum vector for the $n$th particle.
- **Ideal** gases are those whose constituent particles don't interact with each other. As a result, the total energy of the gas is simply the sum of the energies of the $N$ individual particles,

$$E = \frac{1}{2m} \sum_{i=n}^{N} p_n^2. \tag{48}$$

As usual for the canonical ensemble, we consider the gas to be in thermodynamic equilibrium, and in thermal contact with a large external thermal reservoir with which it can exchange energy but not particles. To prevent particle exchange, we can specify that the gas is enclosed in a box with volume $V = L^3$. The thermal reservoir fixes the temperature $T$ of the gas.

The starting point for our analysis is to compute the partition function

$$Z = \sum_i e^{-E_i/T}.$$

Here the sum is over all possible micro-states $\omega_i$ of the $N$-particle system, which is problematic. These micro-states depend on the momenta $\vec{p}_n$ for all $N$ particles, and it's intuitive to suppose that each component of $(p_x, p_y, p_z)_n$ is a continuously varying real number that can (in principle) be distinguished with arbitrary precision. This implies an uncountably infinite set of distinct momenta and hence an uncountably infinite set of micro-states, making the summation above ill-defined.

To proceed, we need to *regulate* the system so that there are a countable number of micro-states and we can define the partition function. We do this by positing that the particles' momentum components can take only discrete (or 'quantized') values that depend on the volume of the box. Specifically, we declare that the possible momenta are

$$\vec{p} = (p_x, p_y, p_z) = \hbar \frac{\pi}{L}(k_x, k_y, k_z) \qquad\qquad k_{x,y,z} = 0, 1, 2, \cdots . \tag{49}$$

Each component of the non-negative-integer vector $\vec{k} = (k_x, k_y, k_z)$ is independent. The constant factor $\hbar$ ("h-bar"), known as the (reduced) Planck constant (named after Max Planck), simply converts units from inverse-length ($\frac{1}{L}$) to momentum ($p$). Very similar discrete momenta turn out to be realized in nature, thanks to quantum mechanics—if you have previously studied quantum physics, you may recognize the momenta for a particle in a box, but for the purposes of this module we can just adopt this result as an ansatz.

What are the energies that correspond to these discretized momenta?

You should find energies that fall into discrete *energy levels*, somewhat similar to the spin system considered last week.

Even though there are still an infinite number of possible momenta and energy levels for each particle in the gas, these are now countable, making our partition function well-defined. Let's start by considering the partition function $Z_1$ for a single particle in the box. The micro-states for this single-particle system are completely specified by that particle's momentum $\vec{p}$,

$$Z_1 = \sum_i \exp\left[-\frac{E_i}{T}\right] = \sum_{\vec{p}} \exp\left[-\frac{p^2}{2mT}\right] = \sum_{k_{x,y,z}=0}^{\infty} \exp\left[-\frac{\hbar^2\pi^2}{2mTL^2}\left(k_x^2 + k_y^2 + k_z^2\right)\right].$$

Because $(k_x, k_y, k_z)$ are all independent, we can sum over each of them independently,

$$Z_1 = \sum_{k_x=0}^{\infty} \exp\left[-\frac{\hbar^2\pi^2}{2mTL^2}k_x^2\right] \sum_{k_y=0}^{\infty} \exp\left[-\frac{\hbar^2\pi^2}{2mTL^2}k_y^2\right] \sum_{k_z=0}^{\infty} \exp\left[-\frac{\hbar^2\pi^2}{2mTL^2}k_z^2\right].$$

For technical reasons,[13] the three properties listed above allow us to assume that Planck's constant $\hbar$ is extremely small compared to $L\sqrt{mT}$, so that

$$\frac{\hbar^2\pi^2}{2mTL^2} \ll 1.$$

[13] If $m$ is too small, effects due to special relativity become non-negligible. If $T$ is too small, effects due to quantum physics become non-negligible. If $L$ is too small, we don't have a sufficiently large-scale (*macroscopic*) system to justify analysis via statistical ensembles in the first place.

This means that the energy levels are spaced very close to each other, and the summations over discrete integer $k$ can be well approximated by integrals over continuous real $\widehat{k}$, so that

$$\sum_{k_x=0}^{\infty} \exp\left[-\frac{\hbar^2\pi^2 k_x^2}{2mTL^2}\right] \approx \int_0^{\infty} d\widehat{k}_x \exp\left[-\frac{\hbar^2\pi^2 \widehat{k}_x^2}{2mTL^2}\right] = \frac{1}{2}\int_{-\infty}^{\infty} d\widehat{k}_x \exp\left[-\frac{\hbar^2\pi^2 \widehat{k}_x^2}{2mTL^2}\right].$$

The final equality simply notes that the integrand is an even function of $\widehat{k}_x$ (depending only on $\widehat{k}_x^2$).

Since we're back to working with continuous real momenta, we may as well use Eq. 49 to return to the original $dp_x = \hbar\frac{\pi}{L}d\widehat{k}_x$,

$$\sum_{k_x=0}^{\infty} \exp\left[-\frac{\hbar^2\pi^2}{2mTL^2}k_x^2\right] = \frac{1}{2}\int \frac{L}{\pi\hbar}dp_x \exp\left[-\frac{p_x^2}{2mT}\right].$$

Combining the three summations that appear in $Z_1$ above produces

$$Z_1 = \left(\frac{L}{2\pi\hbar}\right)^3 \int d^3p \exp\left[-\frac{p^2}{2mT}\right],$$

again with $p^2 = p_x^2 + p_y^2 + p_z^2$. We can now account for all $N$ particles in the ideal gas, which are completely independent and don't interact with each other. Assuming we can distinguish these particles from each other, then each of them simply contributes an independent factor of $Z_1$ to the overall partition function

$$Z_D = \left(\frac{L}{2\pi\hbar}\right)^{3N} \int d^{3N}p \exp\left[-\sum_{n=1}^{\infty}\frac{p_n^2}{2mT}\right], \tag{50}$$

where the subscript reminds us about the particles' distinguishability, and we will consider the indistinguishable case below.

First, we can recognize that each of the $3N$ independent integrations in Eq. 50 is a gaussian integral,

$$\frac{L}{2\pi\hbar}\int dp_x \exp\left[-\frac{p_x^2}{2mT}\right] = \frac{L}{2\pi\hbar}\sqrt{2\pi mT} = \sqrt{\frac{mTL^2}{2\pi\hbar^2}} \equiv \frac{L}{\lambda_{\text{th}}(T)}.$$

In the last step we have made the notation more compact by defining the *thermal de Broglie wavelength* (named after Louis de Broglie),

$$\lambda_{\text{th}}(T) = \sqrt{\frac{2\pi\hbar^2}{mT}}. \tag{51}$$

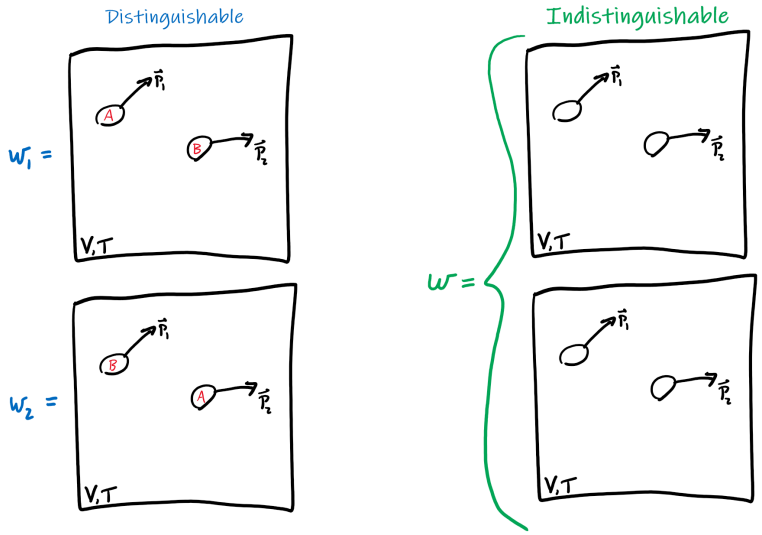
Performing all $3N$ gaussian integrals,

$$Z_D = \left(\frac{mTL^2}{2\pi\hbar^2}\right)^{3N/2} = \left(\frac{L}{\lambda_{\text{th}}}\right)^{3N} = \left(\frac{V}{\lambda_{\text{th}}^3}\right)^N, \tag{52}$$

since the volume of the box is $V = L^3$. It is worth emphasizing here that the partition function **depends on the volume of the gas**, $V = L^3$, in addition to

the fixed temperature $T$ and conserved particle number $N$. This dependence may persist in other quantities derived from the partition function, which we will consider in the next section.

Now let's consider what we would have with indistinguishable particles. In the ideal gas context, distinguishability means that we can label the particles and use those labels to tell them apart as they bounce around inside the box. In the simple two-particle example illustrated below, these labels mean we have a different micro-state $\omega_1$ when particle $A$ has momentum $\vec{p}_1$ while particle $B$ has momentum $\vec{p}_2$, compared to micro-state $\omega_2$ in which particle $A$ has momentum $\vec{p}_2$ while particle $B$ has momentum $\vec{p}_1$.




If the particles are indistinguishable, no such labeling is possible, and there is only one micro-state for these $\{\vec{p}_1, \vec{p}_2\}$, rather than two. This factor of $2$ is not accidental, as you can explore by counting how many micro-states there are for three distinguishable particles with momenta $\{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$, compared to the single micro-state for the indistinguishable case:

Generalizing to $N$ particles, ideal gases with distinguishable particles have $N!$ times more micro-states compared to otherwise-identical ideal gases with indistinguishable particles: There are $N$ possible ways to label the particle with

momentum $\vec{p}_1$, then $N-1$ possible labels for $\vec{p}_2$, and so on.[14] The partition function sums over these micro-states, but depends only on their energies, which are independent of any labeling. Therefore this factor of $N!$ is the only difference between Eq. 52 and the partition function $Z_I$ for indistinguishable particles,

$$Z_I = \frac{1}{N!}\left(\frac{mTL^2}{2\pi\hbar^2}\right)^{3N/2} = \frac{1}{N!}\left(\frac{L}{\lambda_{\text{th}}}\right)^{3N} = \frac{1}{N!}\left(\frac{V}{\lambda_{\text{th}}^3}\right)^N . \tag{53}$$

## 4.2 Internal energy, and entropy

Now that we have the canonical partition function, let's apply our work from last week to predict the large-scale behaviour of the ideal gas it describes. Our first targets are the average internal energy $\langle E\rangle$ and entropy $S$ for the gas, as functions of its fixed temperature $T$, conserved particle number $N$, and the volume $V = L^3$ in which it is placed. Let's begin with the slightly more complicated case of indistinguishable particles, Eq. 53. Recalling the derivatives in Eqs. 40–41, we should keep the temperature dependence explicit in our workings, rather than hidden inside the thermal de Broglie wavelength $\lambda_{\text{th}}(T)$.

Starting by writing down the Helmholtz free energy,

$$F_I = -T\log Z_I = -\frac{3NT}{2}\log\left(\frac{mTL^2}{2\pi\hbar^2}\right) + T\log\left(N!\right),$$

we can quickly extract the internal energy,

$$\langle E\rangle_I = -T^2\frac{\partial}{\partial T}\left(\frac{F_I}{T}\right) = -T^2\frac{\partial}{\partial T}\left(-\frac{3N}{2}\log T + T\text{-independent}\right) = \frac{3}{2}NT.$$

This in turn provides the entropy

$$S_I = \frac{\langle E\rangle_I - F_I}{T} = \frac{3}{2}N + \frac{3N}{2}\log\left(\frac{mTL^2}{2\pi\hbar^2}\right) - \log\left(N!\right).$$

We can clean this up by reintroducing the thermal de Broglie wavelength,

$$\frac{3N}{2}\log\left(\frac{mTL^2}{2\pi\hbar^2}\right) = \frac{3N}{2}\log\left(\frac{L^2}{\lambda_{\text{th}}^2}\right) = N\log\left(\frac{V}{\lambda_{\text{th}}^3}\right),$$

and by applying Stirling's formula as we did last week, again neglecting $\mathcal{O}(\log N)$ terms,

$$\log\left(N!\right) = \log\left(N^N e^{-N}\right) + \mathcal{O}(\log N) \approx -N\log\left(\frac{e}{N}\right).$$

Putting these pieces together,

$$S_I = \frac{3}{2}N + N\log\left(\frac{eV}{N\lambda_{\text{th}}^3}\right).$$

[14] We here assume that the momenta themselves are distinguishable: $\vec{p}_i \neq \vec{p}_k$ for any $i \neq k$. This is a reliable assumption for classical gases, but will need to be revisited when we introduce quantum statistics.

What are the corresponding results for the case of distinguishable particles, where the partition function $Z_D$ is given by Eq. 52?

$F_D =$

$\langle E \rangle_D =$

$S_D =$

You should find that the energy is insensitive to whether or not we can label the particles:

$$\langle E \rangle_D = \langle E \rangle_I = \frac{3}{2} NT. \tag{54}$$

This is in contrast to the spin system we considered last week.[15] The entropy, however, does reflect the extra information that distinguishability provides:

$$S_D = \frac{3}{2} N + N \log \left( \frac{V}{\lambda_{\text{th}}^3} \right) \qquad S_I = \frac{3}{2} N + N \log \left( \frac{eV}{N \lambda_{\text{th}}^3} \right). \tag{55}$$

As the temperature approaches absolute zero, $T \to 0$, it appears as though these entropies become infinitely negative—this is a warning sign that our classical assumptions are breaking down and quantum effects would need to be taken into account.

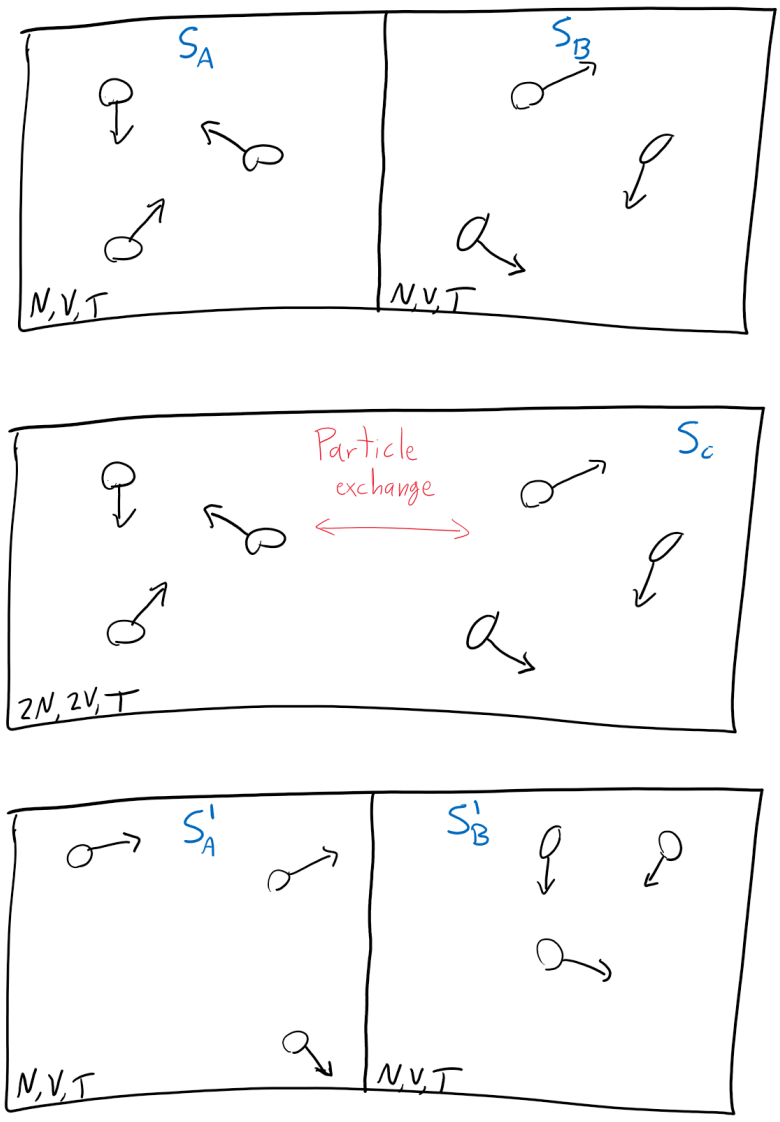
## 4.3 The mixing entropy and the 'Gibbs paradox'

Back in Section 2.4 we considered what would happen if we allowed two micro-canonical ensembles to exchange energy, and then re-isolated them. We saw that this procedure obeyed the second law of thermodynamics: the entropy never decreased, though we had to be careful to account for all of the entropy after re-isolating the two subsystems.

We can now carry out a similar thought experiment by allowing two canonical ensembles to exchange particles, and then re-separating them. We demand that both canonical ensembles are in thermodynamic equilibrium with each other,

[15] The origin of this contrast is that each ideal gas micro-state $\omega_i$ in the indistinguishable case corresponds to $N!$ micro-states in the distinguishable case, independent of the energy $E_i$. For the spin system this factor was $\binom{N}{n_+}$ and varied with the energy $E_{n_+}$.

for instance by sharing the same thermal reservoir with temperature $T$. This procedure is illustrated below, where we simplify the setup by taking the two initial systems to have equal volumes, $V_A = V_B = V$, and numbers of particles, $N_A = N_B = N$.




We can represent the process of combining and then re-separating these systems by writing

$$\Omega_A + \Omega_B \longrightarrow \Omega_C \longrightarrow \Omega'_A + \Omega'_B.$$

What is the entropy for each of these three stages? Since the entropies depend on whether or not the particles in the gas are distinguishable from each other, let's first consider the case of *indistinguishable* particles.

The initial entropy is the sum of the contributions from the two canonical systems, $S_A + S_B$, both of which are the same thanks to our simplification above, and are given by Eq. 55:

$S_A + S_B =$

To find the entropy $S_C$ of the combined system, we just need to consider what happens when we double the volume and also double the number of particles:

$S_C =$

You should find $S_C = S_A + S_B$, which is consistent with the second law of thermodynamics.

Things are more complicated when we re-separate the systems. Analogously to our considerations in Section 2.4, we need to sum over all the possible ways of dividing the $2N$ particles between the two re-separated boxes. In particular, we need to carry out this sum to compute the partition function $Z'$ for $\Omega'_A + \Omega'_B$, since this is the fundamental quantity from which the entropy is then derived as $S = \frac{\partial}{\partial T}\left(T \log Z'\right)$ (Eq. 40). If $\nu$ particles end up in the left box (system $\Omega'_A$), then the other system $\Omega'_B$ must contain the remaining $2N - \nu$ particles, giving us

$$Z_\nu = \frac{1}{\nu!}\left(\frac{V}{\lambda_{\text{th}}^3}\right)^\nu \times \frac{1}{(2N-\nu)!}\left(\frac{V}{\lambda_{\text{th}}^3}\right)^{2N-\nu} = \frac{1}{\nu!\,(2N-\nu)!}\left(\frac{V}{\lambda_{\text{th}}^3}\right)^{2N}$$

$$Z' = \sum_{\nu=0}^{2N} Z_\nu = \left(\frac{V}{\lambda_{\text{th}}^3}\right)^{2N} \sum_{\nu=0}^{2N} \frac{1}{\nu!\,(2N-\nu)!} = \left(\frac{V}{\lambda_{\text{th}}^3}\right)^{2N} \frac{1}{(2N)!} \sum_{\nu=0}^{2N} \binom{2N}{\nu}$$

$$S'_A + S'_B = 2N \frac{\partial}{\partial T}\left(T \log\left[\frac{V}{\lambda_{\text{th}}^3}\right]\right) - \log[(2N)!] + \log\left[\sum_{\nu=0}^{2N} \binom{2N}{\nu}\right]. \tag{56}$$

Since this is hard to compare with $S_C = S_A + S_B = 3N + 2N \log\left(\frac{eV}{N\lambda_{\text{th}}^3}\right)$, back in the 1870s J. Willard Gibbs suggested considering only the case $N'_A = N'_B = N$. This is justified for large numbers of particles, $N \gg 1$, because there are *very* many ways of choosing $N$ out of the total $2N$ particles to end up in $\Omega'_A$,

$$\binom{2N}{N} = \frac{(2N)!}{N!\,N!} \gg 1, \tag{57}$$

compared to (for example) the single way of choosing $N'_A = 0$. This is the case drawn in the illustration above, and with this approximation we have $\Omega'_A = \Omega_A$ and

$\Omega'_B = \Omega_B$. Therefore the final entropy is simply $S'_A + S'_B = S_A + S_B$ and the second law of thermodynamics is satisfied:

$$S'_A + S'_B = S_C = S_A + S_B.$$

This is what we would expect from everyday experience: opening a door between two identical rooms doesn't lead to any observable effects, nor does reversing that process by closing the door.

Something interesting happens when we repeat the same procedure for the case of *distinguishable* particles, using our result for $S_D(N, V)$ in Eq. 55. If we consider the difference between the combined entropy $S_C$ and the initial entropy $S_A + S_B$,

$$\begin{aligned}\Delta S_{\text{mix}} &= S_C - (S_A + S_B) = S_D(2N, 2V) - 2S_D(N, V) \\ &= 3N + 2N \log\left(\frac{2V}{\lambda_{\text{th}}^3}\right) - \left[3N + 2N \log\left(\frac{V}{\lambda_{\text{th}}^3}\right)\right] = 2N \log 2 > 0, \quad (58)\end{aligned}$$

we find that the entropy increases upon combining the two initial systems. This $\Delta S_{\text{mix}} > 0$ is known as the **mixing entropy**.

This result $S_C > S_A + S_B$ is what we would expect from the second law of thermodynamics. However, repeating the argument above that we can approximate $\Omega'_A = \Omega_A$ and $\Omega'_B = \Omega_B$ would then imply $S'_A + S'_B < S_C$, indicating a *decrease* in the entropy by $\Delta S_{\text{mix}}$ and an apparent violation of the second law of thermodynamics. This is known as the 'Gibbs paradox', though Gibbs himself explained how the paradox is avoided.

The explanation is that because the particles are now distinguishable, $N'_A = N_A$ is no longer sufficient to establish $\Omega'_A = \Omega_A$. We need to specify not only that the total number of particles is the same, but also that the same distinguishable particles initially in $\Omega_A$ also end up in the final $\Omega'_A$. For example, if we start with particles $\{A, B, C\}$ in system $\Omega_A$ with $\{D, E, F\}$ in $\Omega_B$, then there is only a single way of ending up with $\{A, B, C\}$ in $\Omega'_A$ and $\{D, E, F\}$ in $\Omega'_B$, not the factorially large number (Eq. 57) that let us justify Gibbs's approximation. Therefore we have to sum over all the possible ways of distributing all $2N$ distinct particles into the re-separated systems, analogously to Eq. 56. It's not illuminating to grind through that calculation; we can be content to observe that the 'Gibbs paradox' does not arise and the second law of thermodynamics remains valid.

We can also observe another example of physical effects that differ depending only on the intrinsic information content of the system—whether or not the particles in an ideal gas can be distinguished from each other in principle. While mixing gases of distinguishable particles introduces a positive mixing entropy, Eq. 58, for gases of indistinguishable particles there is no change in entropy when we let two subsystems mix, or when we reverse that process and re-separate them.

## 4.4 Pressure, ideal gas law, and equations of state

Below Eq. 52 we emphasized that the ideal gas partition function depends on the volume of the gas, $V$, in addition to the fixed temperature $T$ and conserved

particle number $N$ that always characterize systems governed by the canonical ensemble. Parameters like $V$ that appear in the partition function are called **control parameters**, with the idea that they can (in principle) be controlled in experiments. Control parameters often enter the partition function through the definition of the energies $E_i$ for the micro-states $\omega_i$. In the spin systems we considered in previous weeks, the strength $H$ of the external magnetic field is another example of a control parameter.

Focusing on ideal gases for now, we see that all dependence on $V$ drops out of our result for the average internal energy, Eq. 54. On the other hand, the volume does appear in Eq. 55 for the entropy $S$. For both cases of distinguishable and indistinguishable particles, the entropy depends on the same combination of volume and temperature: $V\lambda_{\text{th}}^{-3} \propto VT^{3/2}$. If we keep $N$ fixed and consider using our experimental control to change the volume and the temperature of the system, the entropy will typically change as a consequence, unless the following relation is satisfied:

$$VT^{3/2} = \text{constant} \qquad \Longrightarrow \qquad S = \text{constant}.$$

Such constant-entropy processes are important enough to merit special terminology.

We define an **adiabatic** process to be a change in the control parameters of a system that does not change the system's entropy.

We have previously seen from the micro-canonical temperature (Eq. 23) and the canonical heat capacity (Eq. 37) that it can be interesting to consider how a system's derived quantities depend on its control parameters. If we consider trying to squeeze an inflated balloon into a small box, we can motivate the following connection between pressure and changes in a system's volume.

The **pressure** is defined to be

$$P = -\left.\frac{\partial}{\partial V}\langle E\rangle\right|_S, \tag{59}$$

with constant entropy $S$. In words, the pressure is the adiabatic response of the system's internal energy to a change in the volume.

Next week we will look in detail at processes that change some or all of the pressure, volume, temperature, or internal energy of a gas (with $N$ fixed). Although changing the temperature departs from the assumptions of the canonical ensemble, we will be able to understand these processes as changing the system from one canonical ensemble (in thermodynamic equilibrium with a thermal reservoir that fixes the initial temperature $T_0$) into another (in thermodynamic equilibrium with a different thermal reservoir that fixes the final temperature $T_f$).

If we demand an adiabatic process,

$$VT^{3/2} = c' \qquad \longrightarrow \qquad T = cV^{-2/3},$$

with $c$ a constant, then starting from Eq. 54 we can relate the average internal energy to the volume (with $N$ fixed),

$$\langle E \rangle = \frac{3}{2} NT = \frac{3c}{2} NV^{-2/3} \qquad \text{for constant entropy.}$$

This in turn allows us to determine the pressure for the ideal gas,

$$P = -\left. \frac{\partial}{\partial V} \langle E \rangle \right|_S =$$

You should find the **ideal gas law**,

$$PV = NT, \tag{60}$$

which is an example of an **equation of state**.

The "state" being referred to by this terminology is different from the micro-states that we have discussed up until now. Whereas each micro-state is defined by detailed information about the microscopic degrees of freedom constituting the system, this **thermodynamic state** or **macro-state** concerns only the large-scale properties of the system: its pressure, volume, temperature, or internal energy.

Historically, equations of state were empirically observed and experimentally studied well before the mathematical development of statistical physics. In the 1660s, for instance, Robert Boyle experimented with changing the pressure of a gas while holding its temperature fixed, finding a special case of the ideal gas law,

$$PV = \text{constant} \qquad \text{for constant } N \text{ and } T,$$

which became known as Boyle's law. Other aspects of the ideal gas law were uncovered during the Industrial Revolution:

- $\frac{V}{T} =$ constant for constant $N$ and $P$ (1787)
- $\frac{P}{T} =$ constant for constant $N$ and $V$ (1802)

- $\frac{V}{N} =$ constant for constant $P$ and $T$ (1812)

In the 1830s these empirical results were combined into the ideal gas law itself, which was subsequently derived on the basis of statistical physics in the 1850s. These historical data are useful to illustrate how progress in scientific and mathematical understanding went hand-in-hand with industrial developments, including the design of engines and related machines, which are connected to our next topic of thermodynamic cycles.

# Week 5: Thermodynamic cycles

## 5.1 Work, pressure and force

Last week we defined the pressure in the canonical ensemble as the thermodynamic response of the internal energy to an adiabatic change in the volume (Eq. 59). At the same time, we motivated this definition by thinking about 'squeezing' the system—exerting a force on it—which suggests a connection between pressure and force. Here we make this connection explicit by considering how the energy of an object changes when a force acts on it.

Consider an object at position $\vec{r} = (x, y, z)$, and suppose it is displaced by a vector $d\vec{r}$ due to a force $\vec{F}(\vec{r})$. The **work** done by this force is defined to be the resulting change in the energy of the object. Infinitesimally, $W = dE = \vec{F} \cdot d\vec{r}$, which generalizes to the line integral $W = \Delta E = \int \vec{F}(r) \cdot d\vec{r}$.

A famous example is an object falling due to the force of the Earth's gravity. That force is $\vec{F} = (0, 0, -mg)$, where $m$ is the mass of the object, $g \approx 9.8\ \mathrm{m/s^2}$ (metres per second per second) is the strength of gravity near the surface of the Earth, and the negative sign indicates that the gravitational force is directed downward. The object starts from rest, with initial (kinetic) energy $E_0 = 0$, and falls downward (parallel to $\vec{F}$) from a height $h$. Its final energy $E_f$ upon hitting the ground comes from the work done by the Earth's gravity:

$$W = \int \vec{F}(r) \cdot d\vec{r} = -mg \int_h^0 dz = mgh > 0$$

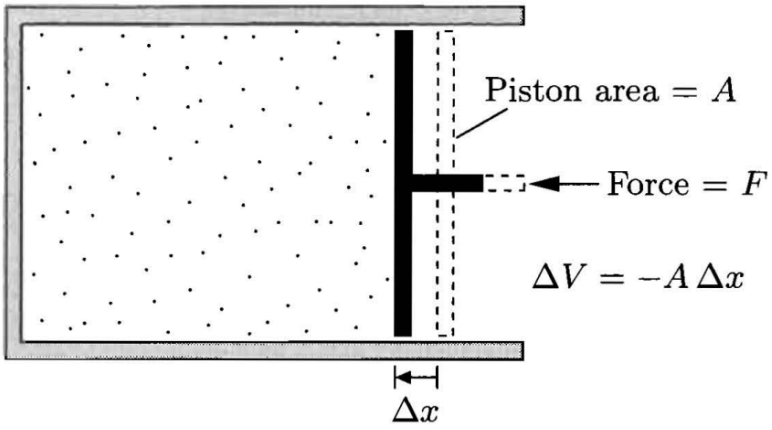
$$E_f = E_0 + \Delta E = 0 + W = mgh = \frac{p_z^2}{2m} \qquad \longrightarrow \qquad p_z = m\sqrt{2gh},$$

where $\vec{p} = (p_x, p_y, p_z)$ is the momentum we considered last week (Eq. 48).

To connect the concept of work to the pressure of a statistical system described by the canonical ensemble, let's consider the setup shown below (copied from Schroeder's *Introduction to Thermal Physics*). Here we have an gas in a container of volume $V$, with one wall of that container being a piston that we can displace by applying a force $F$. Let's demand that this process is adiabatic—it does not change the entropy of the gas. The displacement $\Delta x$ shown in the figure reduces the volume of the gas, by $\Delta V = -A\Delta x < 0$ where $A$ is the area of the piston. Since the force $F$ is parallel to the piston's displacement $\Delta x$, it does positive work $W = F\Delta x > 0$. Therefore the energy of the gas increases, at the same time as its volume decreases adiabatically, so from Eq. 59 we have

$$P = -\frac{\partial}{\partial V} \langle E \rangle \bigg|_S = -\frac{W}{\Delta V} = \frac{F\Delta x}{A\Delta x} = \frac{F}{A} \tag{61}$$

This identifies the pressure of a gas in a container as the force per unit area that the gas exerts on the container wall, reassuringly consistent with our everyday experiences.




Rearranging the expressions above, we can obtain an expression for the work *put into* the gas by its surroundings (that is, by the external force applied to move the piston). Still assuming an adiabatic (constant-entropy) process, this input work must match the increase in the gas's average internal energy,

$$W = \Delta\langle E\rangle = -P\Delta V \qquad \text{for constant entropy.}$$

If the entropy is allowed to change, this relation between pressure and work will still hold—as we will see in the next section, we will simply no longer be able to identify the work as the total change in the average internal energy:

$$W = -P\Delta V \qquad \text{more generally.} \tag{62}$$

Later we will be interested in using the gas as a thermodynamic engine that *does work on* its surroundings. This removes energy from the gas, corresponding to a negative $W < 0$, and we will need to be careful to keep track of the negative signs and their meaning.

Of course, as we change the volume of the gas, the pressure itself will change as described by the gas's equation of state—such as the ideal gas law, Eq. 60. With the equation of state providing an expression $P(V)$ for the pressure as a function of the volume, Eq. 62 generalizes to

$$W = -\int_{V_0}^{V_f} P(V)\, dV. \tag{63}$$

## 5.2 Heat and entropy

Now let's switch things up by changing the temperature $T$ of an ideal gas while keeping its volume constant. (As always for the canonical ensemble, the number of particles $N$ is also constant.) Since the volume is constant, Eq. 63 indicates that no work is done, $W = 0$. Even so, from Eq. 54 we have $\langle E\rangle = \frac{3}{2}NT$ and can see that the average internal energy still changes,

$$d\langle E\rangle = \frac{3}{2}N dT. \tag{64}$$

In order to remain consistent with our discussion in the previous section, we should expect a change in the entropy to accompany this change in the internal energy that occurs with no work done. Indeed, for both cases of distinguishable and indistinguishable particles, the temperature dependence of the entropy in Eq. 55 is the same:

$$S = N \log\left(\lambda_{\text{th}}^{-3}\right) + T\text{-independent} = N \log\left(T^{3/2}\right) + T\text{-independent}.$$

What is the change in the entropy that results from changing the temperature by $dT$?

$dS =$

Looking back to Eq. 64, you should find $d\langle E\rangle = TdS$, which leads us to another important definition.

The **heat** added to or removed from a statistical system is defined to be

$$Q = TdS, \tag{65}$$

and corresponds to the change in the average internal energy of the system when the volume and particle number are kept constant.

Just as for the work $W$ considered in previous section, the heat $Q$ is positive when energy is added to the system to increase $\langle E\rangle$, and negative when energy is removed. We are used to seeing the entropy as a derived function of the temperature, $S(T)$. We can generally invert this relation to integrate over the infinitesimal[16] definition in Eq. 65,

$$Q = \int_{S_0}^{S_f} T(S)\, dS, \tag{66}$$

with $Q = \Delta\langle E\rangle$ when the volume is constant.

This relation between heat exchange and changes in the entropy provides useful insight into what it means, physically, for a process to be adiabatic. In order to keep the entropy constant, an adiabatic process has to occur with no heat moving into or out of the system. In other words, **adiabatic processes are fast** enough that the system does not have time to exchange heat with its

[16] Sometimes infinitesimal heat and work are written $dQ$ and $dW$, but this invites misinterpretation as a 'change' in heat or work, while the heat and work themselves are already changes in the internal energy.

surroundings The opposite extreme would be a process slow enough that any and all possible heat exchange can be completed while it is underway. Based on our work in Section 2.4, we can see that such heat exchange will keep the system's temperature equal to the temperature of its surroundings. Taking that surrounding temperature to be constant, we reach the conclusion that **constant-temperature (or isothermal) processes are slow**. Most real processes exist in between these two extremes, usually closer to the adiabatic limit.

## 5.3 Thermodynamic cycles

Now we can generalize our work in the previous two sections to consider simultaneous changes in the temperature $T$ and the volume $V$ of the gas (still with fixed particle number $N$). We are used to working with the internal energy $\langle E\rangle(T,V)$ and entropy $S(T,V)$ as functions of the temperature and volume. Inverting these relations allows us to instead express the temperature $T(S,V)$ and therefore the internal energy as functions of the entropy and volume,

$$\langle E\rangle(T,V) \to \langle E\rangle(S,V).$$

Expanding the internal energy to first order in a multi-variable Taylor expansion, we have

$$\langle E\rangle(S,V) \approx \langle E\rangle(S_0,V_0) + (S-S_0)\left.\frac{\partial\langle E\rangle}{\partial S}\right|_V + (V-V_0)\left.\frac{\partial\langle E\rangle}{\partial V}\right|_S.$$

This approximation becomes exact in the limit of infinitesimal changes, $(S-S_0)\to dS$, $(V-V_0)\to dV$ and $\langle E\rangle(S,V)-\langle E\rangle(S_0,V_0)\to d\langle E\rangle$. At the same time, we can recognize the temperature from Eq. 23 and the (negative) pressure from Eq. 59, to obtain

$$d\langle E\rangle = TdS - PdV = Q + W. \tag{67}$$

This is the general form of the **first law of thermodynamics**: Any change in the internal energy of a statistical system must be matched by (either or both) heat exchange with its surroundings or work done by or on those surroundings.

We now have all the concepts and **key equations** needed to consider a variety of ways to manipulate an ideal gas in a container:

- Eq. 54 for the internal energy $\langle E\rangle = \frac{3}{2}NT$
- Eq. 55 for the condition of constant entropy, $VT^{3/2} =$ constant
- Eq. 60 for the equation of state (ideal gas law): $PV = NT$
- Eq. 67 for the first law of thermodynamics, $d\langle E\rangle = TdS - PdV = Q + W$

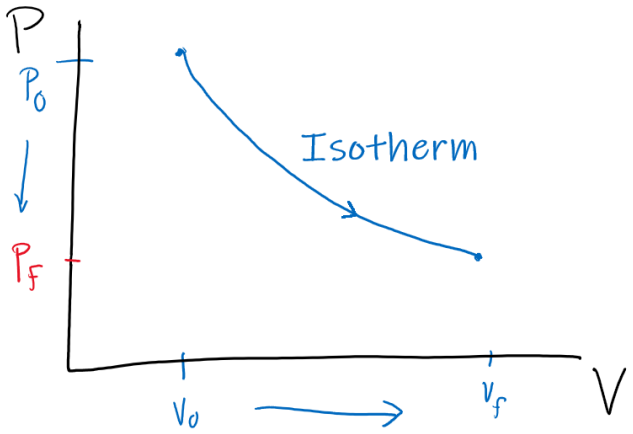
As examples of manipulations we could carry out just by changing the system's control parameters, the piston shown in the figure above allows us to compress or expand the gas. This change in volume can be either fast enough to

keep the entropy constant (adiabatic) or slow enough to keep the temperature constant (isothermal). Alternately, we can clamp the piston in place to keep the volume constant, and add heat to the gas to increase its temperature—which will also increase its pressure according to the ideal gas law. Or we can add heat while keeping the pressure constant by applying a constant force to the piston. The ideal gas law then implies the volume will increase, pushing out the piston and potentially doing work on the external world.

It's possible to carry out a sequence of such manipulations that cause the system to end up in the same thermodynamic (macro-)state in which it started, with the same pressure, volume, temperature and internal energy. That sequence can then be repeated over and over again, always ending up back at its starting point. Such a repeatable process is known as a **thermodynamic cycle**. As we will see in the next section, such cycles can make use of heat to have the system do work on its surroundings (providing an *engine*), or make use of work to remove heat from the system (providing a *refrigerator*), among other applications.

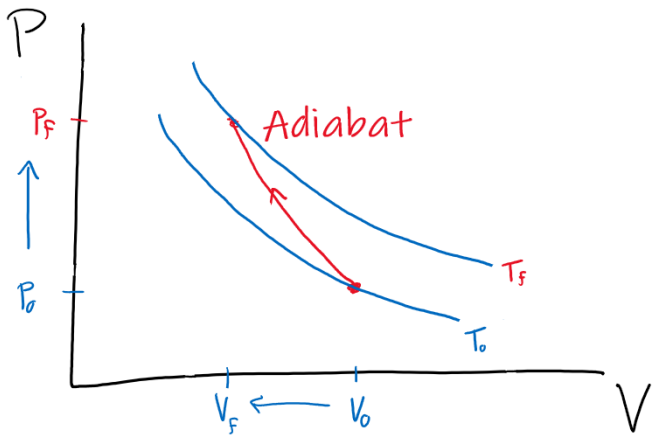
Thanks to the key equations above, we can see that the full ideal gas macro-state can be specified solely in terms of the pressure $P$ and the volume $V$. With fixed $N$, the ideal gas law fixes the temperature $T = \frac{PV}{N}$, which then provides the internal energy $\langle E \rangle \propto NT$. It is therefore convenient to represent the system's macro-state in the form of a **pressure–volume** (or $\mathbf{PV}$) **diagram**—a plot with the volume on the horizontal axis and the pressure on the vertical axis. The manipulations discussed above correspond to lines in $PV$ diagrams. In the case of a thermodynamic cycle, the lines must meet up to form a closed path for the system to go around as the cycle is repeated.

As a first example, the figure below shows the $PV$ diagram for a (slow) isothermal expansion of the gas.



The line in a $PV$ diagram for an isothermal process is known as an *isotherm*. As the volume expands from $V_0$ to $V_f$, the temperature (and therefore $PV$) is constant. What is the change in pressure $\Delta P = P_f - P_0$ in terms of $P_0$, $V_0$ and $V_f$? Can isotherms ever cross?

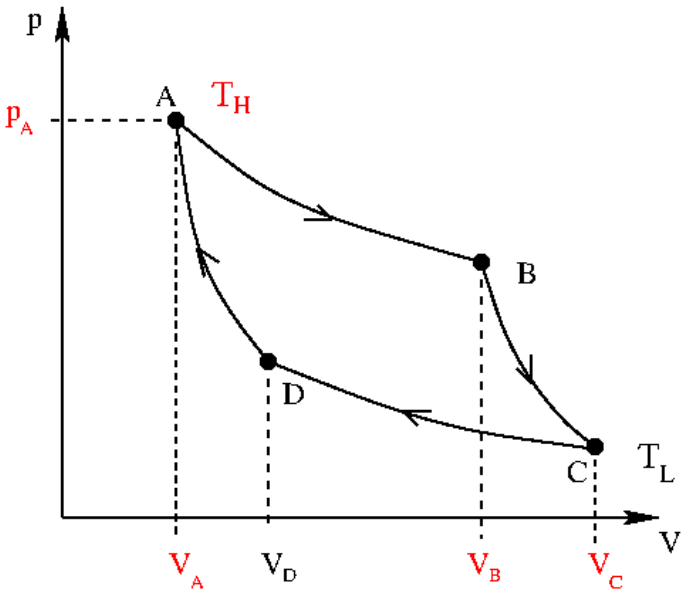
Similarly, we can consider the $PV$ diagram below for a (fast) adiabatic compression of the gas.

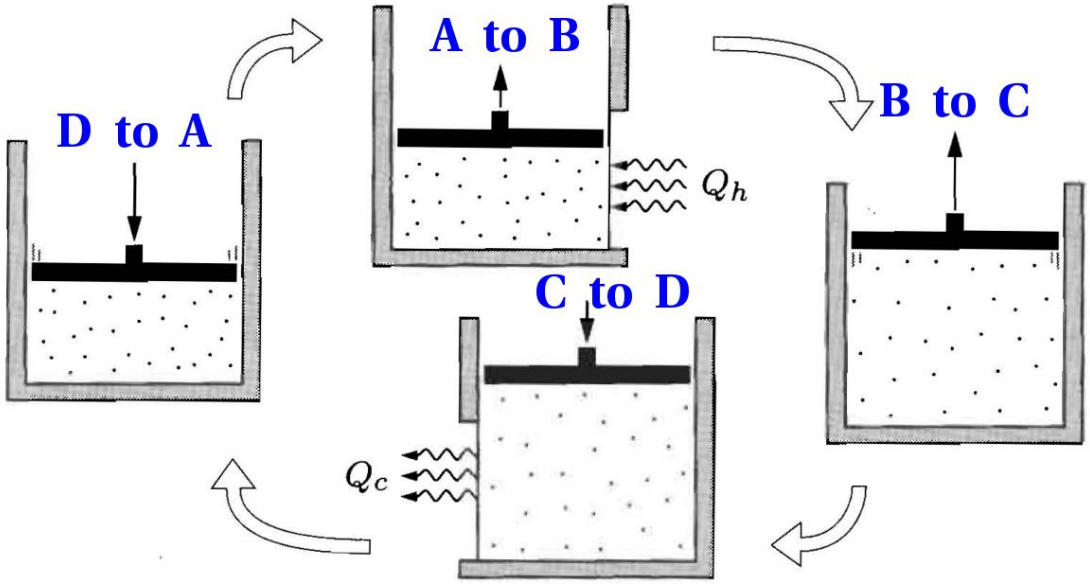



In this case both the pressure and temperature change, while the entropy (and therefore $VT^{3/2}$) is constant. What are $\Delta P$ and the change in the temperature $\Delta T = T_f - T_0$ in terms of $P_0$, $V_0$, $V_f$ and the fixed number of particles $N$?

## 5.4 The Carnot cycle

A famous thermodynamic cycle was proposed by Sadi Carnot in 1824, and laid the groundwork for subsequent development of engines and refrigerators later in the nineteenth century. The key idea is to propose that the ideal gas's container can exchange energy with either of two different thermal reservoirs: a 'hot' reservoir with temperature $T_H$ and a 'cold' reservoir with temperature $T_L$. The Carnot cycle consists of four stages, which are first shown below in the form of a $PV$ diagram, then illustrated in a sketch (adapted from Schroeder's *Introduction to Thermal Physics*) that provides a more concrete picture of the physical processes, and finally summarized in words.



The illustration above supposes that the hot reservoir is located to the right of the system, while the cold reservoir is located to its left. In words, the four stages are the following:

- From point $A$ to point $B$ the system undergoes slow isothermal expansion, bringing in heat $Q_{\text{in}}$ from the hot reservoir in order to keep its temperature fixed at $T_H$.
- From point $B$ to point $C$ the system undergoes fast adiabatic expansion, with no heat exchange, until its temperature falls from $T_H$ down to $T_L$.
- From point $C$ to point $D$ the system undergoes slow isothermal compression, expelling heat $Q_{\text{out}}$ into the cold reservoir in order to keep its temperature fixed at $T_L$.
- From point $D$ to point $A$ the system undergoes fast adiabatic compression, with no heat exchange, until its temperature rises from $T_L$ back up to $T_H$.

We need to make sure that these four processes really do produce a closed, self-consistent cycle that our system could repeatedly follow. In a real experiment, we would have full control over the four variables $\{P_A, V_A, V_B, V_C\}$ coloured red in the $PV$ diagram above. Specifically, we can prepare our $N$-particle system in initial macro-state $A$ with our choice of pressure $P_A$ and volume $V_A$, which together fix the temperature $T_H = \frac{P_A V_A}{N}$ through the ideal gas law. We can then choose how much isothermal expansion occurs to reach volume $V_B$, and finally choose how much adiabatic expansion occurs to reach volume $V_C$. The question then becomes: Is it possible to return to our starting macro-state $A$ by choosing the appropriate volume $V_D$ at which to switch from isothermal compression to adiabatic compression?

To answer this question, we should try to express all the remaining variables $\{P_B, P_C, T_L, P_D, V_D\}$ in terms of the four (red) inputs described above, along with the fixed number of particles $N$. At point $B$, we know the system's temperature remains $T_H = P_A V_A / N$. What is the pressure $P_B$ in terms of $\{P_A, V_A, V_B, V_C, N\}$?

At point $C$, we know the system's entropy is the same as at point $B$. What are the temperature $T_L$ and pressure $P_C$ in terms of $\{P_A, V_A, V_B, V_C, N\}$?

At point $D$, we know the system's temperature remains $T_L$, and demand that its entropy is the same as at point $A$. What are the pressure $P_D$ and volume $V_D$ in terms of $\{P_A, V_A, V_B, V_C, N\}$?

You should find that all of $\{P_B, P_C, T_L, P_D, V_D\}$ can be consistently specified by the (red) inputs under our control, which establishes that the Carnot cycle is a valid thermodynamic cycle. Now we can consider the more interesting question of how much work (if any) this cycle can do on its surroundings, compared to the amount of heat it would need to operate. It will simplify this calculation to use the following positive quantities, with subscripts (rather than negative signs) indicating whether energy is flowing into or out of the gas:

- When work is done on the system by its surroundings, $W_{\text{in}} = W > 0$ from Eq. 63
- When work is done by the system on its surroundings, $W_{\text{out}} = -W > 0$
- When heat enters the system, $Q_{\text{in}} = Q > 0$ from Eq. 65
- When heat leaves the system, $Q_{\text{out}} = -Q > 0$

We can now define a convenient combination of heat and work to consider.

The **efficiency** $\eta$ of a thermodynamic engine is defined to be

$$\eta = \frac{W_{\text{done}}}{Q_{\text{in}}} = \frac{W_{\text{out}} - W_{\text{in}}}{Q_{\text{in}}}, \tag{68}$$

where $W_{\text{done}} = W_{\text{out}} - W_{\text{in}}$ is the *net* amount of work done by each repetition of the cycle, while $Q_{\text{in}}$ is the *total* amount of heat that enters the system in each repetition.

By specifying a thermodynamic *engine*, we assume $W_{\text{out}} > W_{\text{in}}$, so that the overall cycle does more work on its surroundings than it requires as input to operate. This corresponds to $\eta > 0$, and we can also put an upper bound on the efficiency, due to the first law of thermodynamics, Eq. 67. Because the system returns to its initial macro-state after each repetition of the cycle, we have

$$\begin{aligned} \Delta\langle E\rangle = 0 &= Q_{\text{in}} - Q_{\text{out}} + W_{\text{in}} - W_{\text{out}} \\ &\implies W_{\text{out}} - W_{\text{in}} = Q_{\text{in}} - Q_{\text{out}} \leq Q_{\text{in}}, \end{aligned} \tag{69}$$

or $\eta \leq 1$, with equality occurring when no waste heat is expelled throughout the entire cycle, $Q_{\text{out}} = 0$. All together, $0 < \eta \leq 1$ lets us interpret the efficiency as the fraction of the input heat that the engine is able to use to do work on its surroundings.

Let's illustrate these ideas by computing the efficiency of the Carnot cycle. We can divide this task into smaller pieces by considering the contributions to $W_{\text{done}}$ and $Q_{\text{in}}$ from each of the cycle's four stages. First, in the isothermal expansion from point $A$ to point $B$, the ideal gas law provides $P(V)$ to insert into Eq. 63:

$$W_{AB} = -\int_{V_A}^{V_B} P(V)\, dV =$$

You should find $W_{AB} < 0$, meaning the system does work on its surroundings during this stage. At the same time, the constant temperature means $\Delta\langle E\rangle \propto \Delta T = 0$ from Eq. 54, so that $Q_{AB} = -W_{AB} > 0$ and heat flows into the system.

Next, in the adiabatic expansion from point $B$ to point $C$, we know $Q_{BC} = 0$ and therefore

$$W_{BC} = \Delta\langle E\rangle = \frac{3}{2}N\Delta T =$$

You should find that the system continues doing work on its surroundings during this stage, $W_{BC} < 0$.

Finally, the computations for the two compression stages are directly analogous to those above. For the isothermal compression from point $C$ to point $D$, we have

$$W_{CD} = -\int_{V_C}^{V_D} P(V)\, dV =$$

Now you should find $W_{CD} > 0$, meaning this compressions requires work to be done on the system by its surroundings, while $Q_{CD} = -W_{CD} < 0$ means heat flows out of the system. For the adiabatic compression from point $D$ to point $A$, we know $Q_{DA} = 0$ while the change in temperature is exactly opposite the $\Delta T$ of the $B \to C$ adiabatic expansion. Therefore $W_{DA} = -W_{BC} > 0$ and more work has to be done on the system to complete the cycle.

Putting everything together,

$$\begin{aligned} W_{\text{out}} &= -W_{AB} - W_{BC} \\ W_{\text{in}} &= W_{CD} + W_{DA} = W_{CD} - W_{BC} \\ Q_{\text{in}} &= Q_{AB} = -W_{AB} \\ \eta &= \frac{-W_{AB} - W_{BC} - W_{CD} + W_{BC}}{-W_{AB}} = 1 + \frac{W_{CD}}{W_{AB}} = 1 - \frac{T_L}{T_H}. \end{aligned} \tag{70}$$

We can check that our result $\eta = 1 - \frac{T_L}{T_H}$ for the efficiency of the Carnot cycle makes sense. Since $T_L < T_H$, we have $\eta > 0$. If the temperatures of the hot and cold reservoirs approached each other, $\frac{T_L}{T_H} \to 1$, then there would no longer be any heat exchange, and the engine would cease to function, with its efficiency $\eta \to 0$. In the opposite limit of a large difference in the temperatures $T_L \ll T_H$, the efficiency would improve, with $\eta \to 1$ as $\frac{T_L}{T_H} \to 0$.

It turns out to be generic for heat engines to operate more efficiently as the temperature difference between their hot and cold reservoirs increases, and

they always cease working as $\frac{T_L}{T_H} \to 1$. The Carnot cycle is special because its efficiency $\eta = 1 - \frac{T_L}{T_H}$ is the theoretical maximum allowed by the second law of thermodynamics. We can show this by using Eq. 69 to rewrite

$$\eta = \frac{Q_{\text{in}} - Q_{\text{out}}}{Q_{\text{in}}} = 1 - \frac{Q_{\text{out}}}{Q_{\text{in}}} = 1 - \frac{T_L \Delta S_{\text{out}}}{T_H \Delta S_{\text{in}}},$$

where the last equality uses Eq. 66 and the generic fact that the input heat $Q_{\text{in}} = T_H \Delta S_{\text{in}}$ enters the engine from the hot reservoir while the waste heat $Q_{\text{out}} = T_L \Delta S_{\text{out}}$ is expelled to the cold reservoir. After each repetition of the cycle, the gas returns to its original macro-state, with its original entropy, after absorbing entropy $\Delta S_{\text{in}}$ from its surroundings and expelling $\Delta S_{\text{out}}$ back out again. The second law therefore demands $\Delta S_{\text{out}} \geq \Delta S_{\text{in}}$, so that

$$\eta = 1 - \frac{T_L}{T_H} \frac{\Delta S_{\text{out}}}{\Delta S_{\text{in}}} \leq 1 - \frac{T_L}{T_H}$$

in principle, for any thermodynamic engine.

Finally, if we were to operate the Carnot cycle in reverse, with isothermal expansion at temperature $T_L$ and compression at $T_H$, we would do work on the system in order to bring heat in from the cold reservoir and expel it to the hot reservoir. In other words, we would have a refrigerator rather than an engine. The 'efficiency' of a refrigerator is called its *coefficient of performance*, and defined as

$$\text{COP} = \frac{Q_{\text{in}}}{W_{\text{in}} - W_{\text{out}}} = \frac{Q_{\text{in}}}{Q_{\text{out}} - Q_{\text{in}}} = \frac{1}{Q_{\text{out}}/Q_{\text{in}} - 1} \leq \frac{1}{T_H/T_L - 1} = \frac{T_L}{T_H - T_L},$$

which can be greater than one. The reversed Carnot cycle provides the best possible $\text{COP}$ for a refrigerator. Despite its efficiency, the Carnot cycle does not provide a practical engine or refrigerator, simply because its slow isothermal stages take too long! Real engines and refrigerators sacrifice efficiency for functionality.

# Week 6: Grand-canonical ensemble

## 6.1 The particle reservoir and chemical potential

This week we define and develop the third and final statistical ensemble to be studied in this module, which is known as the grand-canonical ensemble. Our approach will follow the pattern set by our previous development of the canonical ensemble. Recall that statistical ensembles are probability spaces describing the micro-states that a system can adopt as it evolves in time, subject to certain constraints. We began in week 2 with the micro-canonical ensemble, in which these constraints were conservation of the internal energy $E$ and particle number $N$. We then introduced the canonical ensemble in week 3 by allowing the system's internal energy to fluctuate, while keeping its temperature $T$ fixed through thermal contact with a large external thermal reservoir.

The next step is to allow *both* the system's energy and its particle number to fluctuate. Just as we saw for the canonical ensemble in week 3, these fluctuations occur through contact between the system and a large external reservoir. This is now a **particle reservoir**, with which the system can exchange both energy and particles.

In the same way that energy exchange leads to a fixed temperature, we expect there to be some quantity that will be fixed due to particle exchange. Recall that we initially defined the temperature in the context of the micro-canonical ensemble in thermodynamic equilibrium (Eq. 23), as the dependence of the entropy on the internal energy for a fixed number of degrees of freedom:

$$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_N .$$

The quantity we are now interested in comes from the complementary analysis interchanging the roles of $E$ and $N$.

In thermodynamic equilibrium, the **chemical potential** in the micro-canonical ensemble is defined by

$$\mu = -T \left.\frac{\partial S}{\partial N}\right|_E . \tag{71}$$

This definition is not terribly intuitive, nor is the chemical potential a familiar concept from everyday experiences. To gain some insight into the chemical potential, we can note (from either Eq. 23 or Eq. 67) that $\mu$ has dimensions of energy. We can also expect the partial derivative $\frac{\partial S}{\partial N}$ to be positive in general, since systems with more degrees of freedom generically have more entropy, reflecting the greater amount of information they can contain. This can be checked explicitly for the spin system (Eq. 45) and ideal gas (Eq. 55) we previously analyzed.

The choice of sign in Eq. 71 therefore means that we should expect the chemical potential to be negative, in 'natural' systems with positive temperatures.

The motivation for this negative sign comes from considering a net flow of particles between two systems $\Omega_A$ and $\Omega_B$ with the same temperature $T$ but different

$$\left(\frac{\partial S}{\partial N}\right)_A > \left(\frac{\partial S}{\partial N}\right)_B \qquad \Longrightarrow \qquad \mu_A < \mu_B.$$

Due to the negative sign in Eq. 71, the system with the larger partial derivative has the smaller (more-negative) chemical potential. According to the second law of thermodynamics, particles will flow from $\Omega_B$ to $\Omega_A$, because

$$\Delta S_A = \left(\frac{\partial S}{\partial N}\right)_A \Delta N > \left(\frac{\partial S}{\partial N}\right)_B \Delta N = \Delta S_B,$$

meaning that more entropy is gained by adding particles to system $\Omega_A$ than is lost by removing particles from system $\Omega_B$. This ensures that the process increases the total entropy of the universe, $\Delta S = \Delta S_A - \Delta S_B > 0$. In other words, we expect particles to flow *from* systems with larger chemical potentials *to* systems with smaller chemical potential. This provides a useful similarity to heat flowing from hotter systems with larger temperatures to colder systems with smaller temperatures, allowing us to borrow our intuition from temperature to apply to the less-familiar chemical potential.[17] Like the temperature, the chemical potential is an intensive quantity.

We are now able to define a **grand-canonical ensemble** to be a statistical ensemble characterized by its fixed temperature $T$ and fixed chemical potential $\mu$, with the temperature and chemical potential held fixed through contact with a particle reservoir.

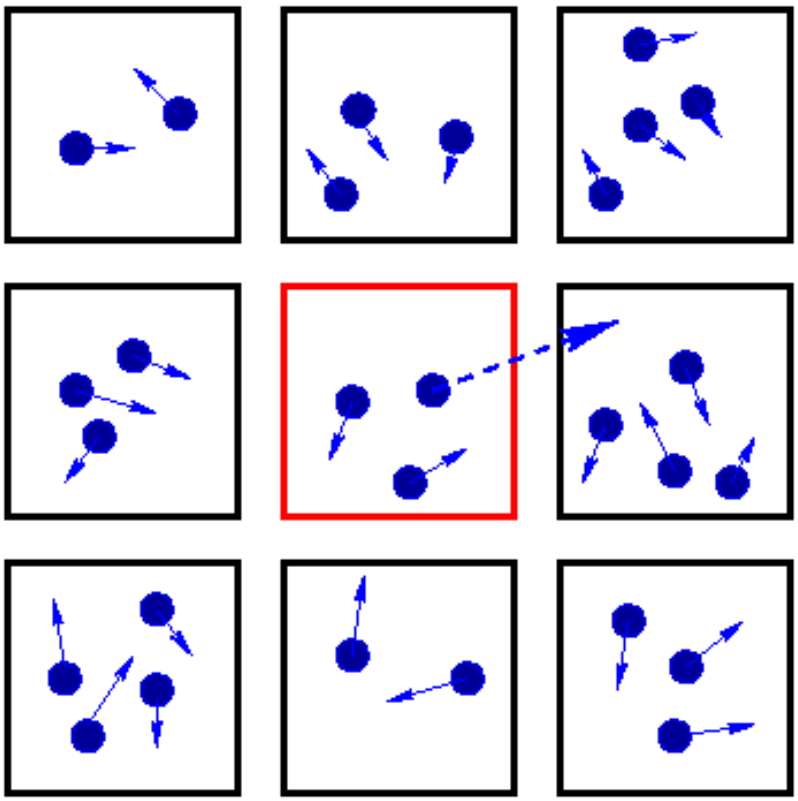
## 6.2 The grand-canonical partition function

We now place the grand-canonical ensemble on a more concrete mathematical foundation, following the same path as we did when developing the canonical ensemble. That is, we introduce a well-motivated ansatz for the form of the particle reservoir $\Omega_{\text{res}}$, then show that this specific form of $\Omega_{\text{res}}$ is ultimately irrelevant. This will allow us to work directly with the system of interest, $\Omega$, independent of the details of the particle reservoir that fixes its temperature and chemical potential.

As before, our ansatz is to take $\Omega_{\text{tot}} = \Omega_{\text{res}} \otimes \Omega$ to consist of many ($R \gg 1$) identical replicas of the system $\Omega$ that we're interested in. All of these replicas are in thermodynamic equilibrium, and can exchange both energy and particles with each other. The overall system $\Omega_{\text{tot}}$ is governed by the micro-canonical ensemble, with conserved total energy $E_{\text{tot}}$ and conserved total particle number $N_{\text{tot}}$. An extremely small example of this setup is illustrated by the figure below, where the

[17] Of course, the mathematics would work just as well with the opposite sign convention—in that case we would just need to develop the opposite intuition for chemical potential vs. temperature.

system of interest is an ideal gas in a volume $V$. This week we will consider only indistinguishable particles, so that we don't need to keep track of which particular particles are exchanged between the replicas, only the overall number.



Although we draw a box around each replica (and colour one red to pick out the system $\Omega$ we will consider), these boxes are now merely mental constructions, and don't interfere with particles moving from one replica to another. For example, we could take our system to be a cubic centimetre of air in a room, with the rest of the room forming its reservoir. As in Section 3.1.1, we assume that this system $\Omega = \{\omega_1, \omega_2, \cdots, \omega_M\}$ has a finite number of $M$ possible micro-states, where now different micro-states may involve different numbers of particles.

This again allows us to analyze the overall system of $R$ replicas in terms of occupation numbers $n_i$ and the corresponding occupation probabilities $p_i$. Recall that $n_i$ is the number of replicas that adopt the micro-state $\omega_i \in \Omega$ in any given micro-state of the overall system $\Omega_{\text{tot}}$, so that $\sum_i n_i = R$. Similarly, $p_i = n_i/R$ is the probability that a randomly chosen replica will be in micro-state $\omega_i$, with $\sum_i p_i = 1$ as usual. In terms of $n_i$ and $p_i$, the total number of micro-states of $\Omega_{\text{tot}}$, and the corresponding entropy, are the same as we derived in Section 3.1.2,

$$M_{\text{tot}} = \frac{R!}{n_1!\, n_2!\, \cdots\, n_M!} \qquad \longrightarrow \qquad S(E_{\text{tot}}, N_{\text{tot}}) = -R \sum_{i=1}^{M} p_i \log p_i,$$

assuming $R \gg 1$ and $n_i \gg 1$ for all $i = 1, \cdots, M$. In this expression, the dependence on both $E_{\text{tot}}$ and $N_{\text{tot}}$ now enters through the occupation probabilities $p_i$,

since the micro-states $\omega_i$ may involve different numbers of particles in addition to different energies.

Continuing as before, we want to determine the (intensive) temperature and chemical potential of $\Omega_{\text{tot}}$ through Eqs. 23 and 71, which requires expressing $S(E_{\text{tot}}, N_{\text{tot}})$ directly in terms of $E_{\text{tot}}$ and $N_{\text{tot}}$. We again do this by maximizing the entropy subject to the constraints on the conserved quantities of the micro-canonical overall system $\Omega_{\text{tot}}$. Labelling the energy and particle number of each replica $E_r$ and $N_r$, respectively, as in Eq. 28 we can again rearrange sums over replicas into sums over the micro-states of $\Omega$:

$$1 = \sum_{i=1}^{M} p_i \qquad E_{\text{tot}} = \sum_{r=1}^{R} E_r = \sum_{i=1}^{M} n_i E_i = R \sum_{i=1}^{M} p_i E_i$$
$$N_{\text{tot}} = \sum_{r=1}^{R} N_r = \sum_{i=1}^{M} n_i N_i = R \sum_{i=1}^{M} p_i N_i, \tag{72}$$

where $E_i$ and $N_i$ are the energies and particle numbers of the $M$ micro-states $\omega_i \in \Omega$. The first two constraints, on the occupation probabilities and the total energy, are the same as we had in Section 3.1.2. The third constraint, on the total particle number, is the new ingredient for us to incorporate.

Writing everything in terms of occupation probabilities (choosing signs and normalization factors for later convenience), we see that we need to maximize the modified entropy

$$\overline{S} = -R \sum_{i=1}^{M} p_i \log p_i + \alpha \left( \sum_{i=1}^{M} p_i - 1 \right) - \beta \left( R \sum_{i=1}^{M} p_i E_i - E_{\text{tot}} \right) + \gamma \left( R \sum_{i=1}^{M} p_i N_i - N_{\text{tot}} \right),$$

with the three Lagrange multipliers $\alpha$, $\beta$ and $\gamma$. What is the occupation probability $p_k$ that maximizes $\overline{S}$?

$$0 = \frac{\partial \overline{S}}{\partial p_k} =$$

You should find a probability of the form

$$p_k = \frac{1}{Z_g} e^{-\beta E_k + \gamma N_k}, \tag{73}$$

defining $Z_g = \exp\left[1 - \frac{\alpha}{R}\right]$ to work in terms of the free parameters $\{Z_g, \beta, \gamma\}$. As usual, we fix these three free parameters by demanding that the three constraints above are satisfied. Using the first constraint, what is $Z_g$ in terms of $\beta$ and $\gamma$?

$$1 = \sum_{i=1}^{M} p_i =$$

Guided by our work in Section 3.1.2, we can expect to need the partial derivatives

$$\frac{1}{Z_g}\frac{\partial}{\partial \beta} Z_g(\beta, \gamma) = \frac{1}{Z_g}\frac{\partial}{\partial \beta} \sum_{i=1}^{M} e^{-\beta E_i + \gamma N_i} = -\frac{1}{Z_g} \sum_{i=1}^{M} E_i \, e^{-\beta E_i + \gamma N_i} = -\frac{E_{\text{tot}}}{R} \tag{74}$$

$$\frac{1}{Z_g}\frac{\partial}{\partial \gamma} Z_g(\beta, \gamma) = \frac{1}{Z_g}\frac{\partial}{\partial \gamma} \sum_{i=1}^{M} e^{-\beta E_i + \gamma N_i} = \frac{1}{Z_g} \sum_{i=1}^{M} N_i \, e^{-\beta E_i + \gamma N_i} = \frac{N_{\text{tot}}}{R}, \tag{75}$$

where we have used $E_{\text{tot}} = R \sum_i p_i E_i$ and $N_{\text{tot}} = R \sum_i p_i N_i$ from Eq. 72. These partial derivatives appear when we express the entropy in terms of $E_{\text{tot}}$, $N_{\text{tot}}$ and the free parameters

$$\{Z_g(\beta, \gamma), \quad \beta(E_{\text{tot}}, N_{\text{tot}}), \quad \gamma(E_{\text{tot}}, N_{\text{tot}})\},$$

then take the partial derivatives that define the temperature and chemical potential. What do you obtain upon inserting Eq. 73 for $p_i$ into the formula for the entropy?

$$S(E_{\text{tot}}, N_{\text{tot}}) = -R \sum_{i=1}^{M} p_i \log p_i =$$

Taking the derivative of the result with respect to $E_{\text{tot}}$, keeping $N_{\text{tot}}$ fixed, gives us the temperature (Eq. 23):

$$\begin{aligned}\frac{1}{T} &= \frac{\partial}{\partial E_{\text{tot}}}\left[R\log Z_g + \beta E_{\text{tot}} - \gamma N_{\text{tot}}\right]_{N_{\text{tot}}} \\ &= \frac{R}{Z_g}\left(\frac{\partial Z_g}{\partial \beta}\frac{\partial \beta}{\partial E_{\text{tot}}} + \frac{\partial Z_g}{\partial \gamma}\frac{\partial \gamma}{\partial E_{\text{tot}}}\right) + E_{\text{tot}}\frac{\partial \beta}{\partial E_{\text{tot}}} + \beta - N_{\text{tot}}\frac{\partial \gamma}{\partial E_{\text{tot}}} \\ &= -E_{\text{tot}}\frac{\partial \beta}{\partial E_{\text{tot}}} + N_{\text{tot}}\frac{\partial \gamma}{\partial E_{\text{tot}}} + E_{\text{tot}}\frac{\partial \beta}{\partial E_{\text{tot}}} + \beta - N_{\text{tot}}\frac{\partial \gamma}{\partial E_{\text{tot}}} = \beta,\end{aligned} \tag{76}$$

where we insert Eqs. 74 and 75 in the last line. In the same way, the derivative with respect to $N_{\text{tot}}$, keeping $E_{\text{tot}}$ fixed, gives us the chemical potential:

$$\begin{aligned}\mu &= -T\frac{\partial}{\partial N_{\text{tot}}}\left[R\log Z_g + \beta E_{\text{tot}} - \gamma N_{\text{tot}}\right]_{N_{\text{tot}}} \\ &= -T\left[\frac{R}{Z_g}\left(\frac{\partial Z_g}{\partial \beta}\frac{\partial \beta}{\partial N_{\text{tot}}} + \frac{\partial Z_g}{\partial \gamma}\frac{\partial \gamma}{\partial N_{\text{tot}}}\right) + E_{\text{tot}}\frac{\partial \beta}{\partial N_{\text{tot}}} - N_{\text{tot}}\frac{\partial \gamma}{\partial N_{\text{tot}}} - \gamma\right] \\ &= -T\left[-E_{\text{tot}}\frac{\partial \beta}{\partial N_{\text{tot}}} + N_{\text{tot}}\frac{\partial \gamma}{\partial N_{\text{tot}}} + E_{\text{tot}}\frac{\partial \beta}{\partial N_{\text{tot}}} - N_{\text{tot}}\frac{\partial \gamma}{\partial N_{\text{tot}}} - \gamma\right] = T\gamma.\end{aligned} \tag{77}$$

Putting everything together, we have

$$\beta = \frac{1}{T} \qquad\qquad \gamma = \beta\mu = \frac{\mu}{T} \tag{78}$$

and the desired result that all the details of the particle reservoir have vanished, with no remaining reference to $R$, $E_{\text{tot}}$ or $N_{\text{tot}}$. The large particle reservoir is still present to fix the temperature $T$ and chemical potential $\mu$ that characterize the grand-canonical system $\Omega$, but beyond that nothing about it is relevant—or even knowable in the grand-canonical approach.

Every aspect of $\Omega$ can now be specified in terms of its fixed temperature $T$ and chemical potential $\mu$, starting with the parameters $\beta = 1/T$ and $\gamma = \mu/T$. In particular, the probability—in thermodynamic equilibrium—that $\Omega$ adopts microstate $\omega_i$ with (non-conserved) internal energy $E_i$ and particle number $N_i$ is

$$p_i = \frac{1}{Z_g}e^{-\beta(E_i - \mu N_i)} = \frac{1}{Z_g}e^{-(E_i-\mu N_i)/T}. \tag{79}$$

These probabilities depend on the **grand-canonical partition function**

$$Z_g(T,\mu) = \sum_{i=1}^{M} e^{-\beta(E_i-\mu N_i)} = \sum_{i=1}^{M} e^{-(E_i-\mu N_i)/T}. \tag{80}$$

Analogously to the canonical partition function, this $Z_g$ is a fundamental quantity in the grand-canonical ensemble, from which many other derived quantities can be obtained.

Since the particle number $N_i$ is dimensionless, the combination $E_i - \mu N_i$ that appears in Eqs. 79 and 80 is consistent with our observation below Eq. 71 that the chemical potential $\mu$ has dimensions of energy.

## 6.3 The grand-canonical potential, internal energy, entropy, and particle number

The development of the grand-canonical ensemble we have seen so far closely resembles our earlier work setting up the canonical ensemble. We have generalized the thermal reservoir to a particle reservoir that allows both the internal energy and particle number of the system $\Omega$ to vary, while keeping its temperature $T$ and chemical potential $\mu$ fixed. By adapting the replica ansatz to this setup, we determined the grand-canonical partition function $Z_g$, and found it to be independent of the details of the particle reservoir.

We now continue by considering a similar set of derived quantities for the grand-canonical ensemble in thermodynamic equilibrium. In addition to the expectation value of the internal energy introduced in Section 3.2, the fluctuations of the particle number mean that we also need to consider its expectation value,

$$\langle E\rangle(T,\mu) = \sum_{i=1}^{M} E_i\, p_i = \frac{1}{Z_g}\sum_{i=1}^{M} E_i\, e^{-\beta(E_i-\mu N_i)}$$
$$\langle N\rangle(T,\mu) = \sum_{i=1}^{M} N_i\, p_i = \frac{1}{Z_g}\sum_{i=1}^{M} N_i\, e^{-\beta(E_i-\mu N_i)}.$$

We can anticipate that both of these derived quantities will be related to the logarithm of the grand-canonical partition function, in analogy to the Helmholtz free energy for the canonical ensemble considered in Section 3.3.

This leads us to define **grand-canonical potential** of a grand-canonical ensemble to be

$$\Phi(T,\mu) = -T\log Z_g(T,\mu) = -\frac{\log Z_g(\beta,\mu)}{\beta}, \tag{81}$$

where $Z_g$ is the grand-canonical partition function of the ensemble. In terms of this free energy, Eqs. 79 and 80 are

$$Z_g = e^{-\Phi/T} \qquad\qquad p_i = e^{(\Phi-E_i+\mu N_i)/T}.$$

The grand-canonical potential is sometimes called the *Landau free energy*, named after Lev Landau, to highlight its similarity with the Helmholtz free energy.

An aspect of this similarity is the importance of derivatives of the grand-canonical potential. The simplest derivative to consider first is with respect to the chemical potential,

$$\frac{\partial}{\partial \mu}\Phi(\beta,\mu) =$$

The derivative with respect to the temperature is a little messier. As in Section 3.3, it involves $\frac{\partial}{\partial T}\log Z_g$, which is again worth collecting in advance, recalling $\frac{\partial}{\partial T} = -\beta^2\frac{\partial}{\partial \beta}$ from Eq. 36.

$$-\frac{\partial}{\partial T}\left(\frac{\Phi(T,\mu)}{T}\right) = \frac{\partial}{\partial T}\log Z_g(T,\mu) =$$

$$\frac{\partial}{\partial T}\Phi(T,\mu) =$$

You should find

$$\frac{\partial \Phi}{\partial T} = \frac{\Phi - \langle E \rangle + \mu \langle N \rangle}{T} = -\log Z_g - \beta \langle E \rangle + \beta\mu \langle N \rangle ,$$

which we can connect to the entropy by inserting the probabilities $p_i$ from Eq. 79 into the definition of the entropy from Eq. 21:

$$S(T,\mu) = -\sum_{i=1}^{M} p_i \log p_i =$$

From this work we can read off the following relations involving the grand-canonical potential $\Phi(T,\mu)$:

$$\langle N \rangle(T,\mu) = -\frac{\partial}{\partial \mu}\Phi(T,\mu) \tag{82}$$

$$S(T,\mu) = -\frac{\partial}{\partial T}\Phi(T,\mu) \tag{83}$$

$$\langle E \rangle(T,\mu) = -T^2 \frac{\partial}{\partial T}\left[\frac{\Phi(T,\mu)}{T}\right] + \mu \langle N \rangle(T,\mu) \tag{84}$$

$$\Phi(T,\mu) = -TS(T,\mu) + \langle E \rangle(T,\mu) - \mu \langle N \rangle(T,\mu) \tag{85}$$

Finally, the connections between the energy, entropy and particle number provided by these relations motivate a further extension of the general first law of thermodynamics we derived last week (Eq. 67). Now allowing the particle number $N$ of our thermodynamic system to change, we can express its entropy as a function of the internal energy, volume and particle number, $S(E,V,N)$, and consider the change in the entropy due to changes in each of these three parameters,[18]

$$dS = \left.\frac{\partial S}{\partial E}\right|_{V,N} dE + \left.\frac{\partial S}{\partial V}\right|_{E,N} dV + \left.\frac{\partial S}{\partial N}\right|_{V,E} dN = \frac{1}{T} dE + \left.\frac{\partial S}{\partial V}\right|_{E,N} dV - \frac{\mu}{T} dN.$$

We can interpret the remaining partial derivative by considering the first law, Eq. 67, in the case of fixed internal energy $E$,

$$dE = 0 = T\, dS - P\, dV \qquad \Longrightarrow \qquad \left.\frac{\partial S}{\partial V}\right|_{E,N} = \frac{P}{T}.$$

[18] To make the notation less cumbersome here, we write $\langle E \rangle$ and $\langle N \rangle$ as $E$ and $N$, respectively, keeping in mind that these are properties of the system's thermodynamic macro-state rather than its fluctuating micro-state.

The fixed particle number $N$ is already incorporated into Eq. 67, since that expression was derived in the framework of the canonical ensemble.

Putting things together, we obtain the generalized thermodynamic identity

$$dE = T\ dS - P\ dV + \mu\ dN. \tag{86}$$

Due to this result, the term $\mu\ dN$ is sometimes referred to as "chemical work", in analogy to the mechanical work $W = -P\ dV$ done on the system by changing its volume. This thermodynamic identity provides a convenient way to remember (or derive) relations between the internal energy, entropy, volume and particle number in thermodynamic equilibrium, by considering processes in which any two of these are fixed. For example, fixing $N$ and $V$ gets us back to Eq. 23 for the temperature,

$$dE = T\ dS \qquad \Longrightarrow \qquad \frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{N,V},$$

while fixing $N$ and $S$ gives Eq. 59 for the pressure,

$$dE = -P\ dV \qquad \Longrightarrow \qquad P = -\left.\frac{\partial E}{\partial V}\right|_{N,S}.$$

If we fix the entropy $S$ and volume $V$, we end up with another way of understanding the chemical potential,

$$dE = \mu\ dN \qquad \Longrightarrow \qquad \mu = \left.\frac{\partial E}{\partial N}\right|_{S,V}. \tag{87}$$

That is, the chemical potential is the change in the internal energy when we adiabatically add a particle to the system (in a constant volume). We argued below Eq. 71 that systems with more degrees of freedom generically have more entropy, and can recall that 'natural' systems have positive temperatures that correspond to the entropy increasing as the internal energy increases. In order to keep the entropy *fixed* as we increase $N$, we would therefore have to reduce $E$, and vice-versa, so that Eq. 87 confirms our earlier finding that the chemical potential is negative in general.

# Week 7: Quantum statistics

## 7.1 Quantized energy levels and their micro-states

This week we begin applying the grand-canonical ensemble to investigate quantum statistical systems. The first step is to introduce quantum statistics itself, building on the initial glimpse that we got in Section 4.1. It is worth reiterating that no prior knowledge of quantum physics is assumed, nor will this module attempt to teach quantum mechanics. We will simply consider quantum behaviour as an ansatz (that turns out to be realized in nature), and analyze the resulting systems by making use of the statistical physics tools we have developed.

Looking back to our derivation of the canonical partition function for a classical (that is, non-quantum) ideal gas in Section 4.1, we can recall that we engaged in slightly circular argumentation. First, because the partition function is defined as a sum over micro-states $\omega_i$,

$$Z = \sum_i e^{-E(\vec{p}_i)/T},$$

we had to conjecture that the gas particles' momenta $\vec{p}_i$ are *quantized* and can take only particular discrete values, rather than varying continuously. These quantized momenta produce a countable number of discrete *energy levels*, leading to a countable number of micro-states and hence a well-defined partition function that takes the form of a sum over all possible discrete momenta for each particle. Second, we then assumed that the energy levels are spaced very close to each other, allowing us to approximate that sum as a multi-dimensional gaussian integral. That is, we went right back to working with continuously varying momenta, despite the formal need to regulate the system by quantization.

For the next few weeks, we will work in the quantum regime where all energy levels of a system remain discrete. In addition, a more subtle change of approach is required by the fundamental indistinguishability of particles governed by quantum mechanics. This quantum indistinguishability is a fact about nature that we will take as given.

To appreciate the consequences of quantum indistinguishability, let's first apply our usual (classical) approach to compute the grand-canonical partition function for a system with discrete energy levels. Despite the quantized energy levels, this calculation will still produce a non-quantum result known as **Maxwell–Boltzmann** (MB) statistics (named after James Clerk Maxwell and Ludwig Boltzmann). We will be able to see when this result is a good approximation and when it breaks down.

To define some notation, let's label the discrete energy levels $\mathcal{E}_\ell$ for $\ell = 0$, $1$, $\cdots$, $L$, where $L$ can be taken to infinity while retaining a countable number of micro-states and hence well-defined partition functions. The energy level $\mathcal{E}_\ell$ may be characterized by extra information in addition to the actual value of its energy, $E_\ell$. As we saw in Section 4.1, it is therefore possible for distinct energy levels $\{\mathcal{E}_m, \mathcal{E}_n\}$ to have the same energy $E_m = E_n$ for $m \neq n$. Such energy

levels with the same value of the energy are said to be *degenerate*. We will label energy levels so that $E_m \leq E_n$ for $m < n$. Without loss of generality, we can take $E_\ell \geq E_0 \geq 0$.

Now starting from the general expression for the grand-canonical partition function, Eq. 80,

$$Z_g(\beta, \mu) = \sum_i e^{-\beta(E_i - \mu N_i)},$$

we just need to define the micro-states $\omega_i$ with energy $E_i$ and particle number $N_i$. In the classical Maxwell–Boltzmann approach, we first sum over all possible particle numbers,

$$Z_g^{\mathsf{MB}}(\beta, \mu) = \sum_{i, N_i = 0} e^{-\beta E_i} + \sum_{j, N_j = 1} e^{-\beta(E_j - \mu)} + \sum_{k, N_k = 2} e^{-\beta(E_k - 2\mu)} + \cdots,$$

where the micro-states labelled $\{\omega_i, \omega_j, \omega_k, \cdots\}$ are those that have $N = 0$, $1$, $2$, $\cdots$ particles, respectively. We can recognize $N$-particle canonical partition functions $Z_N(\beta)$ in the expression above,

$$Z_g^{\mathsf{MB}}(\beta, \mu) = Z_0(\beta) + e^{\beta\mu} Z_1(\beta) + e^{2\beta\mu} Z_2(\beta) + \cdots = \sum_{N=0}^{\infty} \left[e^{\beta\mu}\right]^N Z_N(\beta), \tag{88}$$

allowing us to benefit from our experience with the canonical ensemble. (This is a general result known as the *fugacity expansion*, where $e^{\beta\mu}$ is called the fugacity.)

In particular, because we continue to consider only 'ideal' systems in which the particles don't interact with each other, each $Z_N(\beta)$ is simply the product of the single-particle partition functions $Z_1(\beta)$ for all $N$ independent particles,

$$Z_N(\beta) = \frac{1}{N!} \left[Z_1(\beta)\right]^N,$$

with the factor of $N!$ included to correct for over-counting indistinguishable particles. This is exactly the derivation we performed in Section 4.1, to obtain Eq. 53 for the classical ideal gas. Inserting this into Eq. 88, we have

$$Z_g^{\mathsf{MB}}(\beta, \mu) = \sum_{N=0}^{\infty} \frac{1}{N!} \left[e^{\beta\mu}\right]^N \left[Z_1(\beta)\right]^N = \sum_{N=0}^{\infty} \frac{1}{N!} \left[e^{\beta\mu} Z_1(\beta)\right]^N = \exp\left[e^{\beta\mu} Z_1(\beta)\right].$$

In the case of a system with discrete energy levels $E_\ell$, the single-particle partition function is simply

$$Z_1(\beta) = \sum_{\ell=0}^{L} e^{-\beta E_\ell},$$

leading to the Maxwell–Boltzmann grand-canonical partition function

$$Z_g^{\mathsf{MB}}(\beta, \mu) = \exp\left[e^{\beta\mu} \sum_{\ell=0}^{L} e^{-\beta E_\ell}\right] = \exp\left[\sum_{\ell=0}^{L} e^{-\beta(E_\ell - \mu)}\right]. \tag{89}$$

Unfortunately, as mentioned in a footnote accompanying Eq. 53, this derivation relies on the assumption that every particle occupies a different energy level.

While this would be effectively guaranteed when the particles' energies vary continuously, and can be an excellent approximation when there are many more energy levels than there are particles to occupy them, the assumption breaks down if there is a non-negligible chance of two particles occupying the same energy level.

We can illustrate this with a simple exercise of considering a system with $N = 2$ particles that can occupy any of five energy levels. For a further simplification, let's suppose that all five energy levels are degenerate, with $E_\ell = 0$ for $\ell = 0$, $\cdots$, $5$. This means the canonical partition function simply counts the (integer) number of micro-states, for example

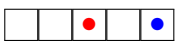
$$Z_1 = \sum_{\ell=0}^{4} e^{-\beta E_\ell} = \sum_{\ell=0}^{4} 1 = 5$$

for all $\beta = 1/T$. The mathematics is the same as counting the number of ways two balls can be placed in five boxes, with possible micro-states that can be represented as  and  . What is the two-particle partition function if the balls are distinguishable?

$Z_D =$

For indistinguishable particles, our derivation above would predict the partition function $Z_I = \frac{1}{2} Z_D$, which is not an integer and therefore cannot be correct.

We can spot the error by explicitly writing down all micro-states in both cases of distinguishable and indistinguishable particles. In the distinguishable case, we can suppose that the balls are red (●) and blue (●), and compactly label micro-states by recording whether each box is empty ("0"), contains the red ball ("$R$"), the blue ball ("$B$") or both balls ("2"):

 $= 00R0B$  $= 00200.$

The full catalog of micro-states is then

| | | | | |
|---|---|---|---|---|
| $RB000$ | $0R0B0$ | $BR000$ | $0B0R0$ | $20000$ |
| $R0B00$ | $0R00B$ | $B0R00$ | $0B00R$ | $02000$ |
| $R00B0$ | $00RB0$ | $B00R0$ | $00BR0$ | $00200$ |
| $R000B$ | $00R0B$ | $B000R$ | $00B0R$ | $00020$ |
| $0RB00$ | $000RB$ | $0BR00$ | $000BR$ | $00002$ |

If we now consider indistinguishable particles where we can only know the number $R = B = 1$, we see that the third and fourth columns above duplicate the first two columns. This is exactly the over-counting that the usual factor of $\frac{1}{N!} = \frac{1}{2}$

corrects, which leaves us with the micro-states

| | | |
|---|---|---|
| 11000 | 01010 | 20000 |
| 10100 | 01001 | 02000 |
| 10010 | 00110 | 00200 |
| 10001 | 00101 | 00020 |
| 01100 | 00011 | 00002 |

But we see that the micro-states in the final column, with both particles in the same energy level, were not over-counted, and must not be divided by $N!$.

In order to generalize this simple exercise, we note that the micro-states for indistinguishable particles can be systematically labelled by *occupation numbers* $n_\ell$, similar to those that we encountered when using replicas to derive the canonical partition function in Section 3.1 and the grand-canonical partition function in Section 6.2. Here the occupation number $n_\ell$ is simply the number of particles in energy level $\mathcal{E}_\ell$. This change of perspective is all we need to define quantum statistics as opposed to classical statistics.

In **quantum statistics**, the micro-states are defined by considering each energy level $\mathcal{E}_\ell$ in turn, and summing over the possible occupation numbers $n_\ell$ that it could have. This contrasts with the classical approach in which we define the micro-states by considering each particle in turn, and summing over the possible energies $E_\ell$ that it could have.

Recalling the fundamental indistinguishability of particles governed by quantum mechanics, we have seen that the classical approach over-counts micro-states, but this over-counting depends on how likely it is for multiple particles to occupy the same energy level. The quantum approach of summing over the occupation numbers of the quantized energy levels avoids this issue, and requires no additional factors to correct over-counting.

## 7.2 Bosons and fermions

In the two sections below we will carry out explicit computations to clarify what the above definition of quantum statistics means in practice. First, there is one more fact about nature that we need to mention. This concerns the occupation numbers $n_\ell$ that are possible for each energy level $\mathcal{E}_\ell$.

A theorem, based on quantum mechanics and special relativity, states that all particles are either *bosons* (named after Satyendra Nath Bose) or *fermions* (named after Enrico Fermi).[19] These two classes of particles obey different rules for their possible occupation numbers, and therefore give rise to distinct quantum statistics.

[19] The proof assumes three-dimensional physical space, and more exotic behaviour is possible for particles confined to two-dimensional surfaces.

Any non-negative number of identical bosons can simultaneously occupy the same energy level, corresponding to occupation numbers $n_\ell = 0, 1, 2, \cdots$. Physical examples of bosons include photons (particles of light), pions, helium-4 atoms and the famous Higgs particle.

On the other hand, it is impossible for multiple identical fermions to occupy the same energy level, meaning that their only possible occupation numbers are $n_\ell = 0$ and $1$. This behaviour is known as the *Pauli exclusion principle* (named after Wolfgang Pauli) and has extremely important consequences, including the existence of chemistry and life. Physical examples of fermions include electrons, protons, neutrons, neutrinos and helium-3 atoms.

The reason multiple identical fermions cannot occupy the same energy level is due to a feature of quantum mechanics, and not because they physically repel each other. This paragraph will imprecisely describe that aspect of quantum physics for the curious, and can be skipped without any problem. Consider a system of identical quantum particles occupying various energy levels. Loosely speaking, if we interchange any pair of those particles, we end up with the same system, up to a factor of $\pm 1$. Bosons correspond to the intuitive $+1$ case, where interchanging indistinguishable particles has no effect. Fermions correspond to the unintuitive $-1$ case (which is related to Grassmann numbers named after Hermann Grassmann). In this case, if the particles we interchange are occupying the same energy level, the resulting system is exactly the same as the starting point — but it also has to differ by this factor of $-1$. Since no non-zero system of particles can equal its negative, no systems with multiple identical fermions in the same energy level can possibly exist.

Looking back at the example system of $N = 2$ particles with five energy levels in the previous section, all $15$ micro-states we wrote down are possible if the particles are bosons. Which of those micro-states are allowed if the particles are fermions?

This difference in the possible micro-states ensures that bosons and fermions exhibit different quantum statistics. We will now consider each case in turn.

## 7.3 Bose–Einstein statistics

The quantum statistics of bosons is known as **Bose–Einstein** (BE) statistics, named after Satyendra Nath Bose and Albert Einstein. As described above, to carry out the sum over all micro-states in the grand-canonical partition function

$$Z_g(\beta, \mu) = \sum_i e^{-\beta(E_i - \mu N_i)},$$

we first sum over all energy levels $\mathcal{E}_\ell$, and then over all possible occupation numbers $n_\ell \in \mathbb{N}_0$ for each energy level.

Consider first the simple case of a system that only has a single energy level $\mathcal{E}$, with energy $E$. In this case, each micro-state $\omega_i$ is uniquely described by its particle number $N_i$, which is just the occupation number of $\mathcal{E}$. What is the energy $E_i$ of micro-state $\omega_i$ with occupation number $n = N_i$?

$E_i =$

Plugging this into the sum over all possible particle numbers, the Bose–Einstein grand-canonical partition function for this single-level system is

$$Z_g^{\mathsf{BE}}(\beta,\mu) = \sum_{n=0}^{\infty} e^{-\beta(E-\mu)n} = \sum_{n=0}^{\infty} \left[e^{-\beta(E-\mu)}\right]^n = \frac{1}{1-e^{-\beta(E-\mu)}}. \tag{90}$$

In the last step we recognized the geometric series

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots,$$

which only converges for $x = e^{-\beta(E-\mu)} < 1$. For natural systems with $\beta = 1/T > 0$, this condition requires $E - \mu > 0$ or equivalently $\mu < E$. Since we can also take $E_\ell \geq 0$ without loss of generality, this constraint is consistent with the negative chemical potential $\mu < 0$ that we discussed last week.

At this point, it is straightforward to generalize to multiple energy levels $\mathcal{E}_\ell$ with $\ell = 0, 1, \cdots, L$. Because we consider only ideal systems with non-interacting particles, the micro-state $\omega_i$ defined by the set of occupation numbers $\{n_\ell\}$ has total energy and particle number

$$E_i = \sum_{\ell=0}^{L} E_\ell n_\ell \qquad\qquad N_i = \sum_{\ell=0}^{L} n_\ell. \tag{91}$$

The general Bose–Einstein grand-canonical partition function is therefore

$$Z_g^{\mathsf{BE}}(\beta,\mu) = \sum_{n_0=0}^{\infty}\sum_{n_1=0}^{\infty}\cdots\sum_{n_L=0}^{\infty} \exp\left[-\beta\sum_{\ell=0}^{L}(E_\ell-\mu)\,n_\ell\right].$$

We can apply the identity $e^{\sum_i x_i} = \prod_i e^{x_i}$ to rewrite

$$\exp\left[-\beta\sum_{\ell=0}^{L}(E_\ell-\mu)\,n_\ell\right] = \left(e^{-\beta(E_0-\mu)n_0}\right)\left(e^{-\beta(E_1-\mu)n_1}\right)\cdots\left(e^{-\beta(E_L-\mu)n_L}\right).$$

Recalling $\mu < E_\ell$ for all $\ell$, by rearranging the terms we find

$$\begin{aligned} Z_g^{\mathsf{BE}}(\beta,\mu) &= \left(\sum_{n_0=0}^{\infty} e^{-\beta(E_0-\mu)n_0}\right)\left(\sum_{n_1=0}^{\infty} e^{-\beta(E_1-\mu)n_1}\right)\cdots\left(\sum_{n_L=0}^{\infty} e^{-\beta(E_L-\mu)n_L}\right) \\ &= \prod_{\ell=0}^{L}\frac{1}{1-e^{-\beta(E_\ell-\mu)}}. \end{aligned} \tag{92}$$

This grand-canonical partition function defines the quantum Bose–Einstein statistics of bosons. Its structure as the product of an independent contribution for each energy level is reminiscent of the result $Z_N \propto Z_1^N$ for the classical $N$-particle canonical partition function discussed in Section 7.1. In both cases we say that the computation *factorizes* into a product of many similar pieces, which can enable drastic simplifications. Looking back to Eq. 92, we can also observe factorization in the classical Maxwell–Boltzmann grand-canonical partition function,

$$Z_g^{\mathsf{MB}}(\beta,\mu) = \exp\left[\sum_{\ell=0}^{L} e^{-\beta(E_\ell-\mu)}\right] = \prod_{\ell=0}^{L} \exp\left[e^{-\beta(E_\ell-\mu)}\right]. \tag{93}$$

In all of these cases, factorization is possible because the particles are non-interacting. In a few weeks we will consider non-ideal systems in which the particles can interact with each other, where the absence of factorization will make analyses much more difficult.

## 7.4 Fermi–Dirac statistics

The quantum statistics of fermions is known as **Fermi–Dirac** (FD) statistics, named after Enrico Fermi and Paul Dirac. The derivation of the Fermi–Dirac grand-canonical partition function is very similar to the Bose–Einstein case considered in the previous section. We again proceed by summing over all energy levels $\mathcal{E}_\ell$, and just have to account for the more limited possible occupation numbers $n_\ell \in \{0, 1\}$ for each energy level.

Taking the same approach of first considering a simple system with only a single energy level, Eq. 90 would just change to

$$Z_g^{\mathsf{FD}}(\beta,\mu) = \sum_{n=0}^{1} e^{-\beta(E-\mu)n} = 1 + e^{-\beta(E-\mu)}.$$

Generalizing to multiple energy levels $\mathcal{E}_\ell$ with $\ell = 0, 1, \cdots, L$, the energy and particle number remain the same as in Eq. 91, and the computation again factorizes,

$$\begin{aligned} Z_g^{\mathsf{FD}}(\beta,\mu) &= \sum_{n_0=0}^{1}\sum_{n_1=0}^{1}\cdots\sum_{n_L=0}^{1} \exp\left[-\beta\sum_{\ell=0}^{L}(E_\ell-\mu)\,n_\ell\right] \\ &= \left(\sum_{n_0=0}^{1} e^{-\beta(E_0-\mu)n_0}\right)\left(\sum_{n_1=0}^{1} e^{-\beta(E_1-\mu)n_1}\right)\cdots\left(\sum_{n_L=0}^{1} e^{-\beta(E_L-\mu)n_L}\right) \\ &= \prod_{\ell=0}^{L}\left[1 + e^{-\beta(E_\ell-\mu)}\right]. \end{aligned} \tag{94}$$

This grand-canonical partition function defines the quantum Fermi–Dirac statistics of fermions. In this case it is possible for the chemical potential $\mu$ to be either positive or negative.

Over the next two weeks we will take $Z_g^{\mathsf{BE}}$ and $Z_g^{\mathsf{FD}}$ as starting points to analyze quantum gases of bosons and fermions, respectively. Before beginning

those more detailed analyses, let's quickly compare the three types of statistics that we have derived this week, while they are all close to hand.

## 7.5 The classical limit

In Section 7.1 we claimed that the classical Maxwell–Boltzmann statistics leading to Eq. 93 can be an excellent approximation to quantum statistics (both bosonic and fermionic), if the probability of multiple particles occupying the same energy level is negligible. We will wrap up this week by demonstrating this fact and clarifying the conditions that correspond to this 'classical limit' of quantum statistics.

It is useful to start by considering the contrapositive of the statement in the previous paragraph: If Maxwell–Boltzmann statistics predicts a non-negligible probability for multiple particles to occupy the same energy level, then it will disagree with the true quantum statistics. This is a useful way to state the problem, because it reveals that we have already found the answer, back in Section 3.4 (and the first homework assignment). There we considered (in the canonical ensemble) classical spin systems with discrete energy levels, finding that at low temperatures the systems are dominated by their lowest-energy micro-states, with only exponentially suppressed corrections coming from any higher-energy micro-states. In the present context, this corresponds to an exponentially small probability for particles to occupy any energy levels with $E_\ell > E_0$, effectively guaranteeing that the lowest energy level $\mathcal{E}_0$ will be occupied by multiple particles and classical statistics will break down.

In short, the low-temperature regime is where quantum and classical statistics disagree, while **high temperatures correspond to the classical limit** of quantum statistics.[20] However, it can be subtle to work with the grand-canonical ensemble at high temperatures, due to the dependence of the average number of particles on the temperature. To demonstrate this subtlety, let's compute the average particle number $\langle N\rangle(T,\mu)$ starting from the grand-canonical partition function, for both classical and quantum statistics.

For convenience, let's collect our earlier results for the grand-canonical partition functions corresponding to classical Maxwell–Boltzmann statistics (Eq. 93), the quantum Bose–Einstein statistics of bosons (Eq. 92) and the quantum Fermi–Dirac statistics of fermions (Eq. 94):

$$Z_g^{\mathsf{MB}} = \prod_{\ell=0}^{L} \exp\left[e^{-\beta(E_\ell-\mu)}\right]$$

$$Z_g^{\mathsf{BE}} = \prod_{\ell=0}^{L} \frac{1}{1-e^{-\beta(E_\ell-\mu)}} \qquad\qquad Z_g^{\mathsf{FD}} = \prod_{\ell=0}^{L} \left[1+e^{-\beta(E_\ell-\mu)}\right].$$

[20] If you are not convinced by the argument above, you can find a more detailed derivation based on the equation of state and thermal de Broglie wavelength in Section 3.5 of David Tong's *Lectures on Statistical Physics* (reference 1 in the list of further reading on page 6).

Recalling $\log\left[\prod_i x_i\right] = \sum_i \log x_i$, the corresponding grand-canonical potentials $\Phi = -T \log Z_g$ for these three cases are

$$\Phi_{\mathsf{MB}} = -T \sum_{\ell=0}^{L} e^{-\beta(E_\ell - \mu)}$$

$$\Phi_{\mathsf{BE}} = T \sum_{\ell=0}^{L} \log\left[1 - e^{-\beta(E_\ell-\mu)}\right] \qquad \Phi_{\mathsf{FD}} = -T \sum_{\ell=0}^{L} \log\left[1 + e^{-\beta(E_\ell-\mu)}\right].$$

We are now ready to compute the average particle numbers. Using the result we derived last week,

$$\langle N \rangle = -\frac{\partial \Phi}{\partial \mu},$$

what are the average particle numbers resulting from the three grand-canonical potentials above?

$$\langle N \rangle_{\mathsf{MB}} = T \sum_{\ell=0}^{L} \frac{\partial}{\partial \mu} e^{-\beta(E_\ell-\mu)} =$$

$$\langle N \rangle_{\mathsf{BE}} = -T \sum_{\ell=0}^{L} \frac{\partial}{\partial \mu} \log\left[1 - e^{-\beta(E_\ell-\mu)}\right] =$$

$$\langle N \rangle_{\mathsf{FD}} = T \sum_{\ell=0}^{L} \frac{\partial}{\partial \mu} \log\left[1 + e^{-\beta(E_\ell-\mu)}\right] =$$

You should find that the average particle number in all three cases can be expressed as a sum over the average occupation numbers,

$$\langle N \rangle = \sum_{\ell=0}^{L} \langle n_\ell \rangle \,,$$

where the average occupation numbers for Maxwell–Boltzmann statistics, Bose–Einstein statistics and Fermi–Dirac statistics are

$$\langle n_\ell \rangle_{\mathsf{MB}} = \frac{1}{e^{\beta(E_\ell - \mu)}}$$

$$\langle n_\ell \rangle_{\mathsf{BE}} = \frac{1}{e^{\beta(E_\ell - \mu)} - 1} \qquad\qquad \langle n_\ell \rangle_{\mathsf{FD}} = \frac{1}{e^{\beta(E_\ell - \mu)} + 1}.$$

Note that $0 \leq \langle n_\ell \rangle_{\mathsf{FD}} \leq 1$, as required for fermions. From these results it is easy to see that the classical limit $\langle n_\ell \rangle_{\mathsf{BE}} \approx \langle n_\ell \rangle_{\mathsf{FD}} \approx \langle n_\ell \rangle_{\mathsf{MB}}$ corresponds to

$$e^{\beta(E_\ell - \mu)} \pm 1 \approx e^{\beta(E_\ell - \mu)} \qquad \Longrightarrow \qquad e^{\beta(E_\ell - \mu)} \gg 1.$$

We can also confirm that in this limit $\langle n_\ell \rangle_{\mathsf{MB}} \ll 1$ for all energy levels $\mathcal{E}_\ell$, indicating very small probabilities for multiple particles to occupy the same energy level.

Now we can appreciate the subtlety promised above, because

$$\beta(E_\ell - \mu) = \frac{E_\ell - \mu}{T} \gg 1. \tag{95}$$

does not look like a high-temperature limit. Indeed, if we consider the naive high-temperature limit $\beta = 1/T \to 0$ with fixed $(E_\ell - \mu)$, we would find large average occupation numbers,

$$\langle n_\ell \rangle_{\mathsf{MB}} \approx 1 \qquad\qquad \langle n_\ell \rangle_{\mathsf{BE}} \to \infty \qquad\qquad \langle n_\ell \rangle_{\mathsf{FD}} \approx \frac{1}{2}.$$

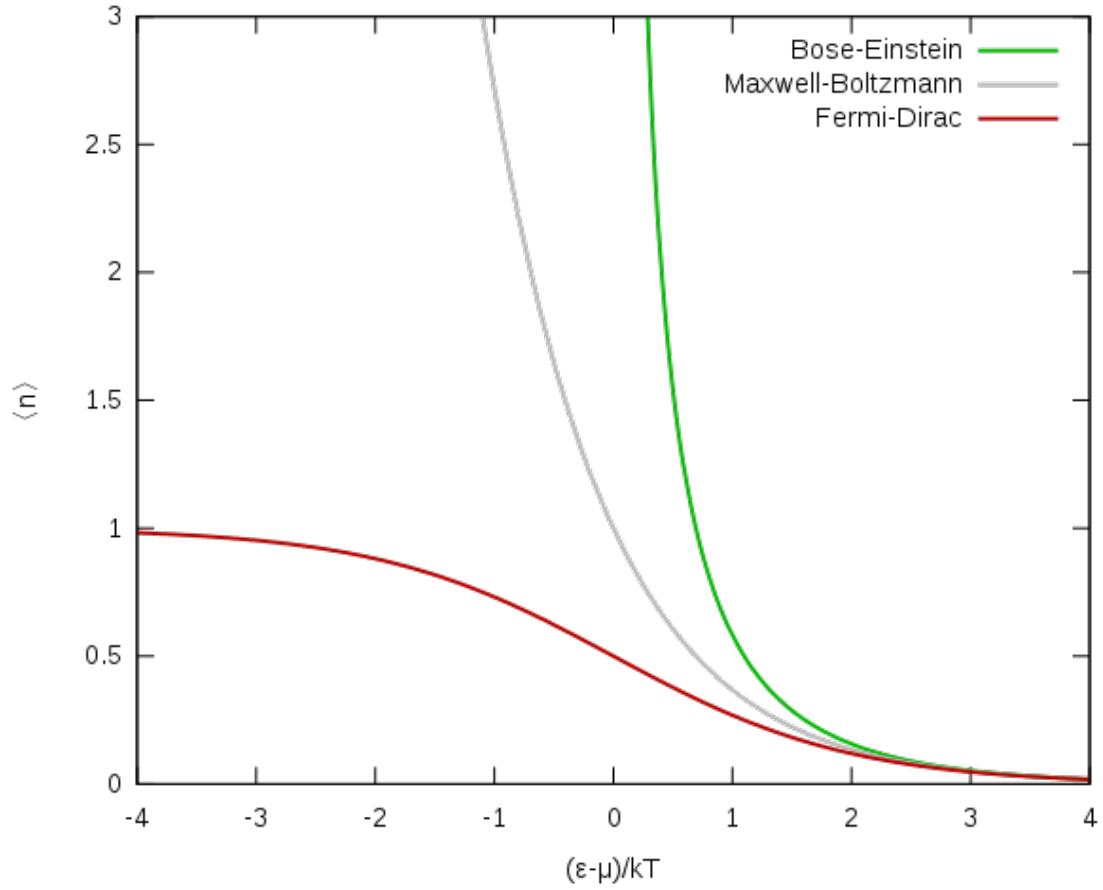
In addition to implying non-negligible classical probabilities for multiple particles to occupy the same energy level, this result indicates that higher temperatures in the grand-canonical ensemble lead to larger particle numbers in total.

In order to balance this effect, we need to adjust the other free parameter offered by the grand-canonical ensemble, the chemical potential $\mu$. Specifically, in order to satisfy Eq. 95 in the high-temperature limit, we need $E_\ell - \mu \gg T$, requiring that $\mu \to -\infty$ as $T \to \infty$. In order to keep the average particle number finite, the true high-temperature limit in which quantum statistics becomes classical turns out to be

$$-\mu \gg T \gg E_\ell \qquad \Longrightarrow \qquad \frac{E_\ell - \mu}{T} \gg 1. \tag{96}$$

This corresponds to the right portion of the plot below,[21] where we can confirm excellent agreement between all three predictions for the average occupation number $\langle n_\ell \rangle$

[21] The plot comes from Wikimedia Commons. The "$k$" in its x-axis label is a unit conversion factor that we set to $k = 1$.




In the low-temperature regime $\frac{E_\ell - \mu}{T} \ll 1$ corresponding to the left portion of the plot, we see dramatically different behaviour for the three cases. The classical Maxwell–Boltzmann prediction for the average occupation number grows exponentially, while the quantum Bose–Einstein prediction diverges as $E_\ell \to \mu$ and the Fermi–Dirac prediction slowly approaches its maximum possible value $\langle n_\ell \rangle_{\text{FD}} \to 1$. Over the next two weeks we will study in more detail the quantum gases of bosons and fermions that produce these results.

# Week 8: Quantum gases of bosons

## 8.1 The photon gas

Last week we derived the grand-canonical partition function (Eq. 92) that defines quantum Bose–Einstein statistics for systems of non-interacting bosons,

$$Z_g^{\mathsf{BE}}(\beta,\mu) = \prod_{\ell=0}^{L} \frac{1}{1-e^{-\beta(E_\ell-\mu)}}.$$

This expression results from summing over the possible occupation numbers $n_\ell \in \mathbb{N}_0$ for each energy level $\mathcal{E}_\ell$ with energy $E_\ell$. The corresponding grand-canonical potential is

$$\Phi_{\mathsf{BE}} = -T\log Z_g = T\sum_{\ell=0}^{L} \log\left[1-e^{-\beta(E_\ell-\mu)}\right],$$

from which we can determine the large-scale properties of the system, including its average internal energy $\langle E\rangle$, average particle number $\langle N\rangle$, entropy $S$, and pressure $P$.

To do so, we have to specify the energy levels of the particles that compose the system of interest, and the degeneracies of those energy levels. One example of this that we have already seen is the analysis of non-relativistic ideal gas particles in Section 4.1. For a single particle with mass $m$ in a volume $V = L^3$, we determined the quantized energies

$$E(k_x,k_y,k_z) = \frac{\hbar^2\pi^2}{2mL^2}\left(k_x^2+k_y^2+k_z^2\right), \tag{97}$$

where the integers $k_{x,y,z}$ specify the possible momenta of the particle,

$$\vec{p} = (p_x,p_y,p_z) = \hbar\frac{\pi}{L}(k_x,k_y,k_z) \qquad\qquad k_{x,y,z} = 1,2,\cdots.$$

(For technical reasons, quantum mechanics requires $k_{x,y,z} \geq 1$, leading us to adjust our ansatz compared to Eq. 49.) This system has a unique ground state $\mathcal{E}_0$ with $\vec{k} = (1,1,1)$ and energy $E_0 = \frac{3}{2}\frac{\hbar^2\pi^2}{mL^2}$. The next three energy levels are degenerate, with energy $3\frac{\hbar^2\pi^2}{mL^2}$ corresponding to $\vec{k} = (2,1,1)$ and permutations, followed by another three degenerate energy levels with energy $\frac{9}{2}\frac{\hbar^2\pi^2}{mL^2}$ corresponding to $\vec{k} = (2,2,1)$ and permutations.

This week we will build on that experience to consider a gas of *photons*, massless bosonic quantum particles of light. For our purposes, with no prior knowledge of particle physics, we can define photons simply by specifying two relevant details of their energy levels. First, a photon's energy is proportional to the magnitude of its momentum,

$$E_{\mathsf{ph}}(p) = c\sqrt{p_x^2+p_y^2+p_z^2} \equiv cp.$$

Here the speed of light $c$ is really just a unit conversion factor that we could set to $c = 1$ by using appropriate units. Second, for each momentum $\vec{p}$, a photon has two degenerate energy levels with the same energy $E(p)$.

In a volume $V = L^3$, only the same discrete momenta as above are allowed,

$$p = \hbar\frac{\pi}{L}\sqrt{k_x^2 + k_y^2 + k_z^2} \equiv \hbar\frac{\pi}{L}k \qquad\qquad k_{x,y,z} = 1, 2, \cdots,$$

so that the quantized photon energies are

$$E_{\text{ph}}(k) = \hbar c\frac{\pi}{L}k. \tag{98}$$

It is conventional to use the speed of light to work with photons in terms of their wavelength $\lambda$ and angular frequency $\omega = 2\pi f$ (not to be confused with generic micro-states $\omega_i$), given the relation

$$c = \frac{\lambda\omega}{2\pi}.$$

Just like the momenta, the wavelengths $\lambda$ are also quantized in volume $V = L^3$,

$$\lambda = \frac{2L}{k} \qquad \Longrightarrow \qquad c = \frac{\omega}{\frac{\pi}{L}k},$$

and we can rewrite Eq. 98 as

$$E_{\text{ph}}(\omega) = \hbar\omega. \tag{99}$$

Low (*infrared*) frequencies correspond to small energies and long wavelengths, while high (*ultraviolet*) frequencies correspond to large energies and short wavelengths.

We are now ready to write down the grand-canonical potential for a photon gas:

$$\Phi_{\text{ph}} = T\sum_{\ell=0}^{L}\log\left[1 - e^{-\beta(E_\ell - \mu)}\right] = 2T\sum_{\vec{k}}\log\left[1 - e^{-\beta(E_{\text{ph}}(k) - \mu)}\right],$$

where the factor of $2$ in the final expression accounts for the doubly degenerate energy levels. We can simplify this expression by appreciating that photons are easy to create and destroy. Every time a light is switched on, it begins emitting a constant flood of photons (with wavelengths of several hundred nanometres). Food in a microwave oven gets hot by absorbing many lower-energy photons (with longer wavelengths around $12$ centimetres). In both cases an enormous number of photons is required to make even a small change in energy, so that Eq. 87 implies the chemical potential of a photon gas must be negligible,

$$\mu = \left.\frac{\partial E}{\partial N}\right|_S \approx 0 \qquad \Longrightarrow \qquad \Phi_{\text{ph}} \approx 2T\sum_{\vec{k}}\log\left[1 - e^{-\beta E_{\text{ph}}(k)}\right].$$

Another simplification comes from considering the photon gas in a large volume, so that we can approximate the sum over discrete integer $k_{x,y,z}$ by integrals over continuous real $\widehat{k}_{x,y,z}$,

$$\Phi_{\text{ph}} \approx 2T\int d\widehat{k}_x d\widehat{k}_y d\widehat{k}_z \log\left[1 - e^{-\beta E_{\text{ph}}(\widehat{k})}\right].$$

Since the energy $E_{\text{ph}}(\widehat{k})$ depends only on the magnitude $\widehat{k}$, we can profit from converting to spherical coordinates. When we do so, we have to keep in mind that $k_{x,y,z} > 0$ corresponds only to the positive octant of the sphere,

$$\int_0^\infty d\widehat{k}_x \int_0^\infty d\widehat{k}_y \int_0^\infty d\widehat{k}_z = \int_0^\infty d\widehat{k}\ \widehat{k}^2 \int_0^{\pi/2} d\theta\ \sin\theta \int_0^{\pi/2} d\phi = \frac{\pi}{2} \int_0^\infty d\widehat{k}\ \widehat{k}^2,$$

so that

$$\Phi_{\text{ph}} \approx \pi T \int_0^\infty d\widehat{k}\ \widehat{k}^2 \log\left[1 - e^{-\beta E_{\text{ph}}(\widehat{k})}\right].$$

We can finally change variables to integrate over the photon angular frequency $\omega = c\frac{\pi}{L}k$, with $E_{\text{ph}} = \hbar\omega$, to find

$$\begin{aligned}\Phi_{\text{ph}} &\approx \pi T \left(\frac{L}{\pi c}\right)^3 \int_0^\infty d\omega\ \omega^2 \log\left[1 - e^{-\beta\hbar\omega}\right] \\ &= \frac{VT}{c^3\pi^2} \int_0^\infty d\omega\ \omega^2 \log\left[1 - e^{-\beta\hbar\omega}\right], \end{aligned} \tag{100}$$

recognizing $L^3 = V$. With this grand-canonical potential derived, we just need to take the appropriate derivatives to determine the thermodynamics and equation of state for the photon gas.

## 8.2 The sun and the void

We are now ready to analyze the average internal energy from the grand-canonical potential for a photon gas, Eq. 100. With $\mu = 0$, Eq. 84 from week 6 becomes

$$\langle E\rangle_{\text{ph}} = -T^2 \frac{\partial}{\partial T}\left[\frac{\Phi_{\text{ph}}}{T}\right] = \frac{\partial}{\partial\beta}\left[\beta\Phi_{\text{ph}}\right].$$

To begin, we will consider the energy density expressed as an integral over photon frequencies,

$$\frac{\langle E\rangle_{\text{ph}}}{V} = \int_0^\infty P(\omega)\ d\omega,$$

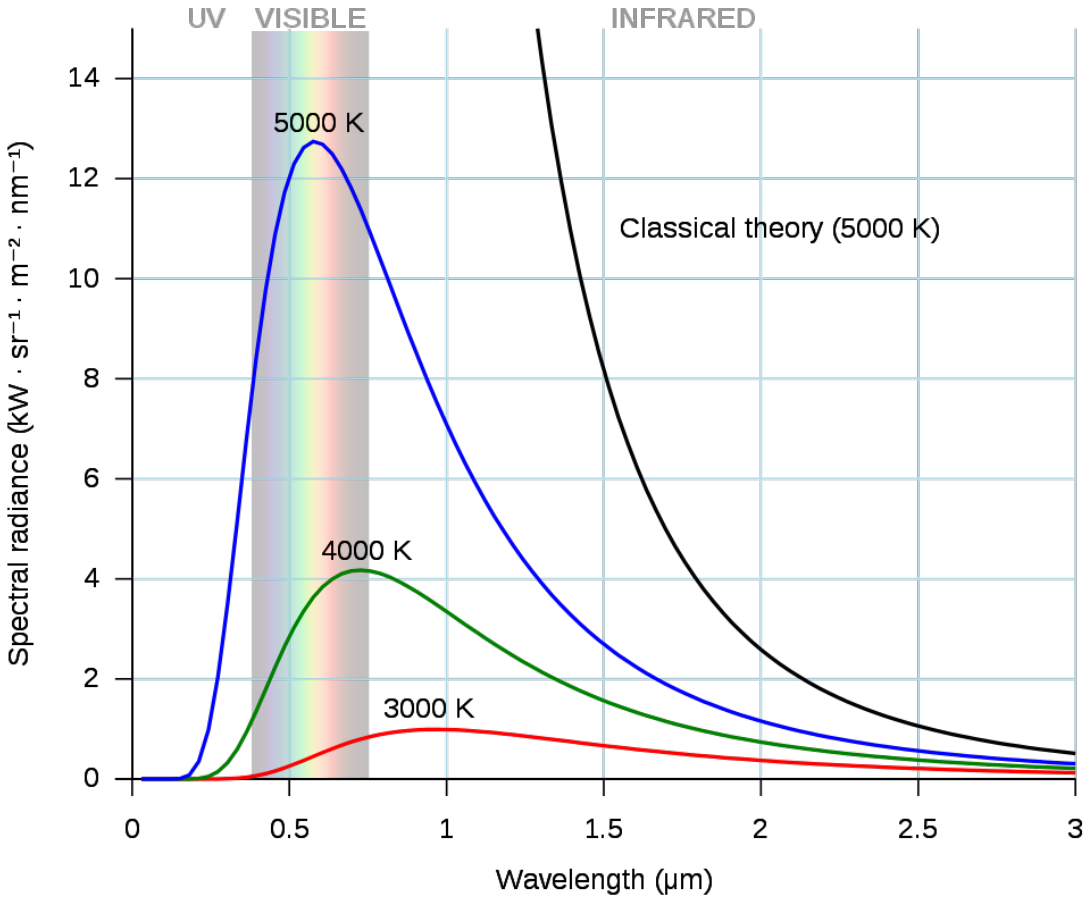
where the function $P(\omega)$ is known as the *spectral density*, or simply the *spectrum*. What is the spectrum for a photon gas?

$$\frac{\langle E\rangle_{\text{ph}}}{V} = \frac{1}{c^3\pi^2} \int_0^\infty d\omega\ \omega^2 \frac{\partial}{\partial\beta} \log\left[1 - e^{-\beta\hbar\omega}\right] =$$

You should find

$$P(\omega) = \left(\frac{\hbar}{c^3\pi^2}\right)\frac{\omega^3}{e^{\beta\hbar\omega}-1}, \tag{101}$$

which is known as the Planck spectrum, named after Max Planck. The Planck spectrum is plotted in the figure below, which comes from Wikimedia Commons.




In this plot the horizontal axis uses the wavelength $\lambda = 2\pi c/\omega$. Changing variables in your work above, what is Planck spectrum $P(\lambda)$ as a function of wavelength?

$$\frac{\langle E\rangle_{\text{ph}}}{V} = \frac{\hbar}{c^3\pi^2}\int_0^\infty \frac{\omega^3}{e^{\beta\hbar\omega}-1}\,d\omega =$$

You should find

$$P(\lambda) = \left(\frac{16\pi^2\hbar c}{\lambda^5}\right)\frac{1}{e^{2\pi\beta\hbar c/\lambda}-1}, \tag{102}$$

which is plotted[22] for three temperatures $T = 1/\beta$ in the figure above. Considering first the high-energy ultraviolet (UV) limit of small wavelengths $\lambda$, we can see from Eq. 102 that $P(\lambda)$ is exponentially suppressed, which overwhelms the diverging factor $\propto 1/\lambda^5$ in parentheses.

In the low-energy infrared limit, the large $\lambda$ has the same effect that a large temperature ($\beta \ll 1$) would have: $e^{2\pi\beta\hbar c/\lambda} - 1 \approx 2\pi\beta\hbar c/\lambda$ and

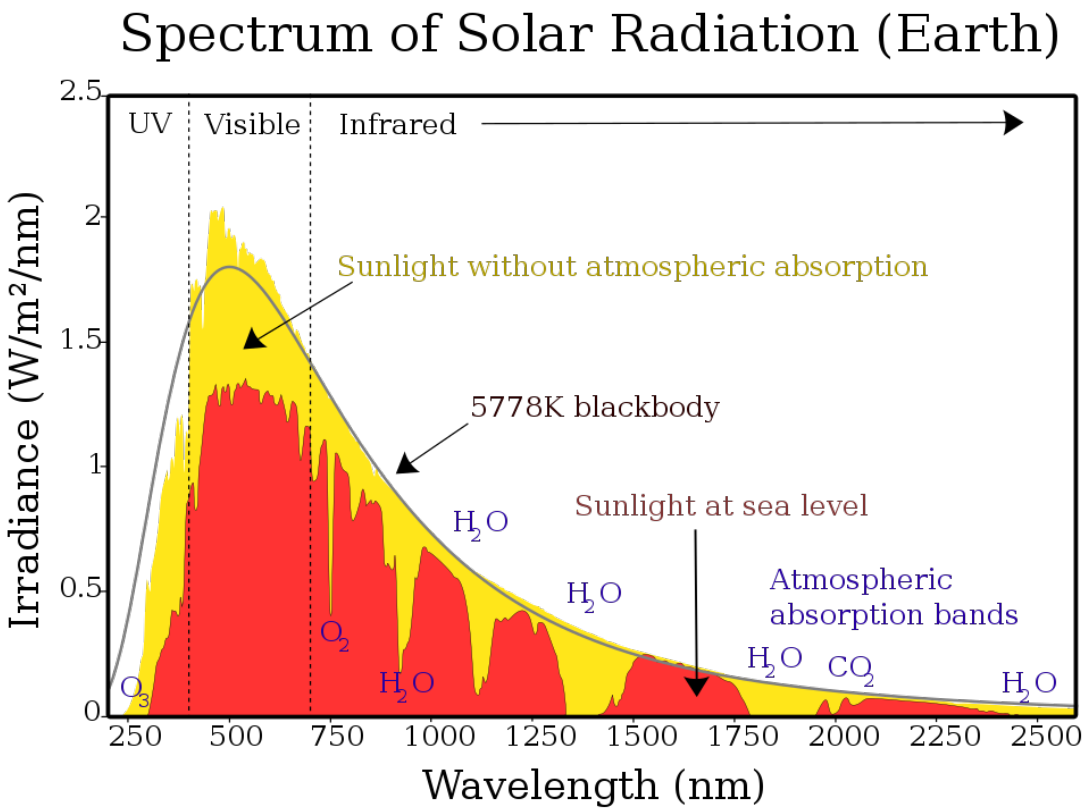
$$P(\lambda) \approx \left(\frac{16\pi^2\hbar c}{\lambda^5}\right)\frac{\lambda}{2\pi\beta\hbar c} = \frac{8\pi T}{\lambda^4}.$$

The connection to large temperatures indicates that this is what classical statistics would predict for the energy spectrum of light. It is known as the Rayleigh–Jeans spectrum, named after the third Baron Rayleigh and James Jeans. Recall that the classical approach sums over all possible energies for each degree of freedom, corresponding to a light-emitting object (historically known as a *black body*) emitting light of all wavelengths $\lambda$. According to the classical Rayleigh–Jeans spectrum, in the limit $\lambda \to 0$ this light would carry an infinite amount of energy, an obvious error that became known as the *ultraviolet catastrophe*. Planck's (heuristic) solution to this conundrum in 1900 was one of the first steps towards the quantum physics that introduces the exponential suppression we computed above.

One final observation we will make about the Planck spectrum shown above is that as the temperature increases, the maximum of the Planck spectrum moves to shorter wavelengths and correspondingly larger energies. The fact that the peak of the spectrum for $T \approx 5000$ K falls within the wavelengths of visible light (roughly $400$–$700$ nm) is not a coincidence. As shown in the figure below (from Wikimedia Commons), the amount of sunlight that reaches the surface of the earth is also maximized around visible wavelengths, which are visible to us because we have evolved to make the most efficient use of this sunlight.

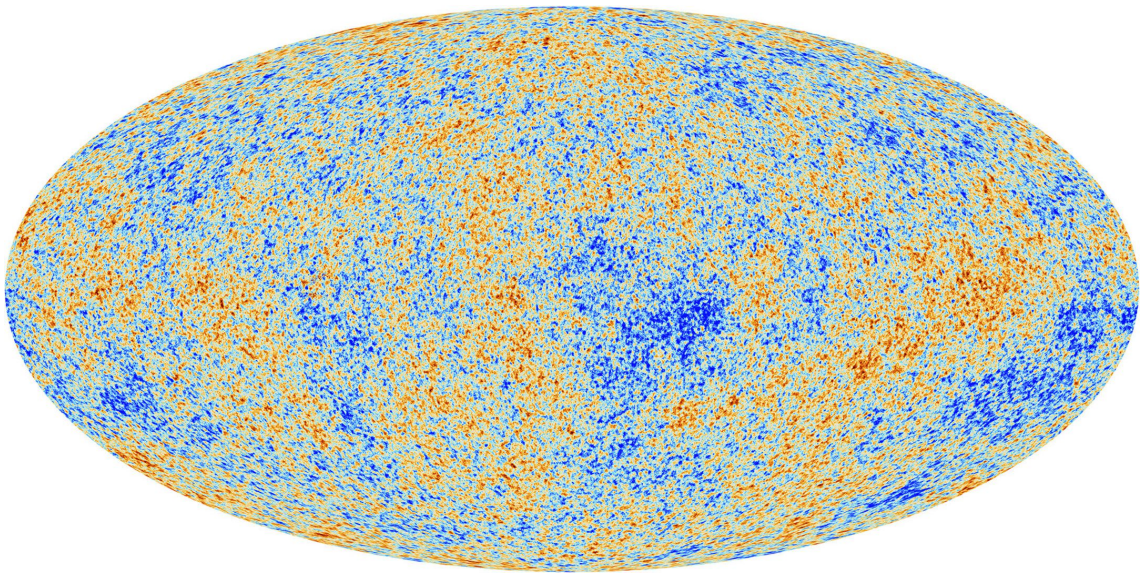
Taking into account the absorption of some sunlight by molecules in the atmosphere, we can see from the figure below that the energy spectrum of the sunlight reaching the top of the atmosphere is quite close to a Planck (or 'blackbody') spectrum with temperature $T \approx 5778$ K. The agreement isn't perfect, which is to be expected since the Planck spectrum relies on the non-trivial assumption of an ideal gas of non-interacting particles. Even with that caveat, numerically fitting the measured sunlight to the Planck spectrum is how the surface temperature of the sun is determined to be $T \approx 5778$ K. This same fitting procedure can even be done for distant stars, with red stars corresponding to relatively low temperatures $T \lesssim 3500$ K and blue stars corresponding to relatively high temperatures $T \gtrsim 10{,}000$ K.

[22] The plot divides our $P(\lambda)$ by $4\pi$ steradian and multiplies it by $c$ to convert from an energy density to a spectral power per unit area per unit of solid angle. For our purposes only the functional form is significant.

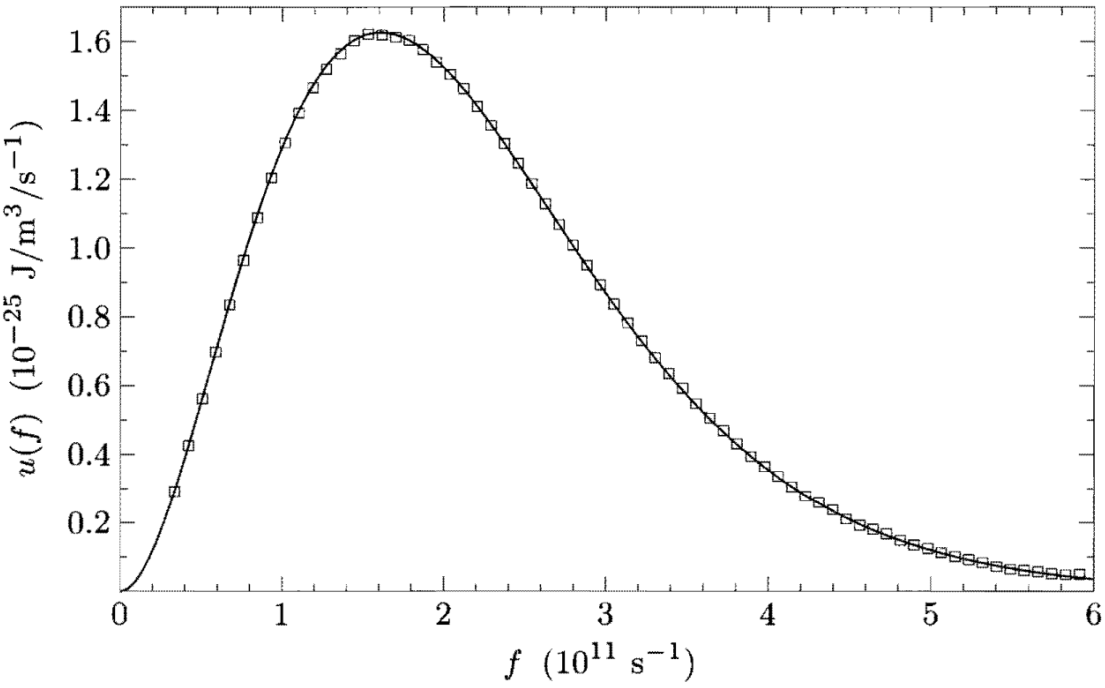



In fact, we can even use the Planck spectrum to determine the temperature of inter-galactic space. Rather than being empty, these voids are actually permeated by a very low-temperature photon gas left over from the Big Bang roughly $14$ billion years ago. This photon gas is known as the cosmic microwave background (CMB), and carries information about the early evolution of the universe, including some of the strongest evidence for the existence of dark matter.

The picture below is a famous visualization of the CMB, provided by the European Space Agency and produced from measurements taken by their 'Planck' satellite. To produce this image, for each point in the sky the satellite measures the photon spectrum reaching it from that direction. The contributions coming from stars and galaxies are subtracted, and the remaining data is fit to the Planck spectrum to find the temperature of the intergalactic CMB photon gas at that point. From point to point, there are only small temperature fluctuations around the average $T_{\text{CMB}} \approx 2.725$ K. That average temperature is subtracted and the fluctuations themselves are shown below, with warmer red-coloured regions only $\Delta T \approx 0.0002$ K hotter than the cooler blue-coloured regions.



The final figure below illustrates such a fit of CMB data to the Planck spectrum, using measurements taken by the Cosmic Background Explorer (COBE) satellite and published in 1990. (This version of the figure is adapted from that publication, and copied from Schroeder's *Introduction to Thermal Physics*.) The squares are the measured data, and their size represents a cautious estimate of uncertainties. They are plotted with the frequency $f = \omega/(2\pi)$ on the horizontal axis, with $f \approx 3 \cdot 10^{-11}$ s$^{-1}$ corresponding to a low-energy wavelength $\lambda = c/f \approx 1$ mm, roughly $1000$ times longer than the wavelengths of visible light. The solid line is a fit to the data that produces $T_{\text{CMB}} = 2.735 \pm 0.060$ K. While more recent satellites have increased the precision with which we know $T_{\text{CMB}}$, this first result was awarded the 2006 Nobel Prize in Physics.



Even though we derived the Planck spectrum by assuming an ideal gas of non-interacting photons, we see that it provides an excellent mathematical model for real physical systems, stretching from the hottest to the coldest places in the universe.

## 8.3 Equation of state for the photon gas

Having looked in some detail at the integrand for the photon gas energy density, Eq. 101, let's complete the integration, which involves a famous integral related to the Riemann zeta function:

$$I_4 = \int_0^\infty \frac{x^3}{e^x - 1} dx = \Gamma(4)\zeta(4) = \frac{\pi^4}{15}.$$

Using this result, what is the ideal photon gas energy density?

$$\frac{\langle E \rangle_{\text{ph}}}{V} = \frac{\hbar}{c^3\pi^2} \int_0^\infty \frac{\omega^3}{e^{\beta\hbar\omega} - 1}\, d\omega =$$

You should find a result proportional to $T^4$. It would be ideal (sorry) if we could put this in a form as simple as Eq. 54 for the energy of an $N$-particle non-relativistic ideal gas in the canonical ensemble. Now that we are working in the grand-canonical ensemble, this requires computing the average photon number from Eq. 82,

$$\langle N \rangle_{\text{ph}} = -\frac{\partial}{\partial\mu}\Phi_{\text{ph}}\bigg|_{\mu=0} = -\frac{VT}{c^3\pi^2}\int_0^\infty d\omega\, \omega^2 \frac{\partial}{\partial\mu} \log\left[1 - e^{-\beta\hbar\omega}e^{\beta\mu}\right]\bigg|_{\mu=0}.$$

recalling $\mu = 0$ for photon gases. The calculation is quite similar to that for the average internal energy density, now involving the integral

$$I_3 = \int_0^\infty \frac{x^2}{e^x - 1} dx = \Gamma(3)\zeta(3) = 2\zeta(3).$$

Using this result, what is the ideal photon gas number density?

$$\frac{\langle N\rangle_{\text{ph}}}{V} = \frac{1}{c^3\pi^2}\int_0^\infty \frac{\omega^2}{e^{\beta\hbar\omega}-1}\,d\omega =$$

You should find a result proportional to $T^3 \propto \langle E\rangle_{\text{ph}}/T$, so that

$$\langle E\rangle_{\text{ph}} = \frac{\pi^2}{15\hbar^3c^3}VT^4 = \frac{\pi^4}{30\zeta(3)}\langle N\rangle_{\text{ph}}\,T. \tag{103}$$

The functional form is the same as Eq. 54, with a larger numerical factor

$$\frac{\pi^4}{30\zeta(3)} = \frac{\Gamma(4)\zeta(4)}{\Gamma(3)\zeta(3)} \approx 2.7$$

compared to the $\frac{3}{2}$ for the classical non-relativistic case.

To get the rest of the way to the equation of state for the photon gas, we need to compute the *radiation pressure*

$$P_{\text{ph}} = -\frac{\partial}{\partial V}\langle E\rangle_{\text{ph}}\bigg|_{S_{\text{ph}}},$$

which requires first figuring out the condition of constant entropy $S_{\text{ph}}$ for a photon gas. From Eq. 85 with $\mu = 0$, we have

$$S_{\text{ph}} = \frac{\langle E\rangle_{\text{ph}} - \Phi_{\text{ph}}}{T}.$$

Looking back to Eq. 100 for the grand-canonical potential, we see

$$\frac{\Phi_{\text{ph}}}{T} = \frac{V}{c^3\pi^2}\int_0^\infty d\omega\,\omega^2\log\left[1-e^{-\beta\hbar\omega}\right] = \frac{VT^3}{\hbar^3c^3\pi^2}\int_0^\infty dx\,x^2\log\left[1-e^{-x}\right],$$

changing variables to $x = \beta\hbar\omega = \hbar\omega/T$. The final factor in this expression is yet another delightful integral,

$$\int_0^\infty dx\,x^2\log\left[1-e^{-x}\right] = -2\zeta(4) = -\frac{\pi^4}{45}.$$

Since this gives us $S \propto VT^3$, we can conclude that the condition of constant entropy for a photon gas is $VT^3 =$ constant, in contrast to the $VT^{3/2}$ dependence of Eq. 55 for classical non-relativistic particles.

At this point we just need to insert the constant-entropy condition $T = bV^{-1/3}$ (with $b$ a constant) into the average internal energy and take the derivative:

$$P = -\left.\frac{\partial}{\partial V}\left\langle E\right\rangle_{\mathsf{ph}}\right|_{S_{\mathsf{ph}}} = -\frac{\partial}{\partial V}\frac{\pi^2}{15\hbar^3c^3}b^4V^{-1/3} =$$

You should find the following equation of state for the photon gas,

$$P_{\mathsf{ph}}V = \frac{1}{3}\left\langle E\right\rangle_{\mathsf{ph}} = \frac{\pi^4}{90\zeta(3)}\left\langle N\right\rangle_{\mathsf{ph}}T. \tag{104}$$

The functional form is the same as the (classical, non-relativistic) ideal gas law, with just an additional numerical factor

$$\frac{\pi^4}{90\zeta(3)} = \frac{\zeta(4)}{\zeta(3)}.$$

# Week 9: Quantum gases of fermions

## 9.1 Non-relativistic ideal fermion gas

This week we wrap up our applications of the grand-canonical ensemble to investigate ideal gases of non-interacting particles. We again take the quantum statistical approach of defining micro-states by summing over the possible occupation numbers $n_\ell$ for each energy level $\mathcal{E}_\ell$ with energy $E_\ell$. In contrast to the bosonic case considered last week, we now focus on quantum gases of fermions, where the only possible occupation numbers are $n_\ell = 0$ and $1$, since the Pauli exclusion principle prevents multiple identical fermions from occupying the same energy level.

In Section 7.4 we derived the grand-canonical partition function (Eq. 94) that defines quantum Fermi–Dirac statistics for such systems of non-interacting fermions,

$$Z_g^{\mathsf{FD}}(\beta,\mu) = \prod_{\ell=0}^{L} \left[1 + e^{-\beta(E_\ell - \mu)}\right],$$

for inverse temperature $\beta = 1/T$ and chemical potential $\mu$. Recall that it is possible for systems of fermions to have any value for the chemical potential (either positive or negative), in contrast to the systems of bosons we considered last week. From the corresponding grand-canonical potential,

$$\Phi_{\mathsf{FD}} = -T \sum_{\ell=0}^{L} \log\left[1 + e^{-\beta(E_\ell - \mu)}\right]$$

we can determine the large-scale properties of the system, including its average internal energy $\langle E\rangle$, average particle number $\langle N\rangle$, entropy $S$, and pressure $P$, along with the equation of state relating these quantities.

A concrete calculation requires specifying the energy levels of the particles that compose the gas, and the degeneracies of those energy levels. Let's begin this week by considering non-relativistic particles of the sort we previously analyzed in Section 4.1. In a volume $V = L^3$, the energy levels are defined by the non-zero quantized energies

$$E(k) = \frac{p^2}{2m} = \frac{\hbar^2\pi^2}{2mL^2}\left(k_x^2 + k_y^2 + k_z^2\right) \qquad\qquad k_{x,y,z} = 1, 2, \cdots.$$

In addition to the usual degeneracies coming from permutations of $(k_x, k_y, k_z)$ that we discussed in previous weeks, for each distinct $\vec{k}$ typical fermions such as electrons have two degenerate energy levels with the same energy. This factor of $2$ has a different origin compared to the double degeneracy discussed for photons last week. Rather than worry about the physical origins of this behaviour, in both cases we simply incorporate this given information into our ansatz.

The grand-canonical potential for an ideal gas of non-relativistic fermions is therefore

$$\Phi_{\mathsf{f}} = T\sum_{\ell=0}^{L} \log\left[1 + e^{-\beta(E_\ell - \mu)}\right] = 2T\sum_{\vec{k}} \log\left[1 + \exp\left(-\frac{\hbar^2\pi^2k^2}{2mL^2T} + \frac{\mu}{T}\right)\right].$$

We can again proceed by considering the gas in a large volume and approximating the sum over discrete integer $k_{x,y,z}$ by integrals over continuous real $\widehat{k}_{x,y,z}$:

$$\Phi_{\mathsf{f}} \approx 2T \int d\widehat{k}_x d\widehat{k}_y d\widehat{k}_z \log\left[1 + \exp\left(-\frac{\hbar^2\pi^2\widehat{k}^2}{2mL^2T} + \frac{\mu}{T}\right)\right].$$

Converting to spherical coordinates and carrying out the angular integrations over the $\frac{\pi}{2}$ solid angle of the octant of the sphere with $k_{x,y,z} > 0$, we have

$$\Phi_{\mathsf{f}} \approx \pi T \int_0^\infty d\widehat{k}\, \widehat{k}^2 \log\left[1 + \exp\left(-\frac{\hbar^2\pi^2\widehat{k}^2}{2mL^2T} + \frac{\mu}{T}\right)\right].$$

In the same spirit as the change of variables we carried out last week, to integrate over photon frequencies $\omega = E_{\mathsf{ph}}/\hbar$, we will now change variables to integrate over the fermion energy:

$$E = \frac{\hbar^2\pi^2}{2mL^2}\widehat{k}^2 \quad \longrightarrow \quad \widehat{k} = \frac{L\sqrt{m}}{\pi\hbar}\sqrt{2E}$$
$$d\widehat{k} = \frac{L\sqrt{m}}{\pi\hbar}\frac{dE}{\sqrt{2E}}.$$

Plugging this in produces

$$\begin{aligned}\Phi_{\mathsf{f}} &\approx \pi T\left(\frac{L^3m^{3/2}}{\pi^3\hbar^3}\right)\int_0^\infty dE \frac{2E}{\sqrt{2E}} \log\left[1 + e^{-\beta(E-\mu)}\right] \\ &= \frac{\sqrt{2m^3}VT}{\pi^2\hbar^3}\int_0^\infty dE\sqrt{E}\log\left[1 + e^{-\beta(E-\mu)}\right]. \qquad (105)\end{aligned}$$

recognizing $L^3 = V$.

With this grand-canonical potential derived, we just need to take the appropriate derivatives to determine the thermodynamics and equation of state for non-relativistic fermions. When doing so, we'll focus on the low-temperature regime where we expect quantum Fermi–Dirac statistics to differ significantly from the classical case we considered back in Section 4.1. As we saw in Section 7.5, at high temperatures (with large negative chemical potential) the classical approach provides a good approximation to the true quantum physics.

## 9.2 Low-temperature equation of state

Rather than the average internal energy $\langle E\rangle_{\mathsf{f}}$, it will prove profitable to first analyze the average particle number

$$\langle N\rangle_{\mathsf{f}} = -\frac{\partial}{\partial\mu}\Phi_{\mathsf{f}}$$

coming from the grand-canonical potential for non-relativistic fermions, Eq. 105. In analogy to the Planck spectrum we derived for the photon gas last week, we first express the particle number density as an integral over energies,

$$\frac{\langle N\rangle_{\mathsf{f}}}{V} = \frac{\sqrt{2m^3}}{\pi^2\hbar^3}\int_0^\infty F(E)\sqrt{E}\, dE,$$

where the function $F(E)$ is known as the Fermi function. In contrast to the Planck spectrum, all the constant factors are kept separate from $F(E)$:

$$\frac{\langle N\rangle_{\mathrm{f}}}{V} = \frac{\sqrt{2m^3T}}{\pi^2\hbar^3}\int_0^\infty dE\sqrt{E}\frac{\partial}{\partial\mu}\log\left[1+e^{-\beta(E-\mu)}\right]$$
$$= \frac{\sqrt{2m^3}}{\pi^2\hbar^3}\int_0^\infty \frac{1}{e^{\beta(E-\mu)}+1}\sqrt{E}\,dE = \frac{\sqrt{2m^3}}{\pi^2\hbar^3}\int_0^\infty F(E)\sqrt{E}\,dE. \quad (106)$$

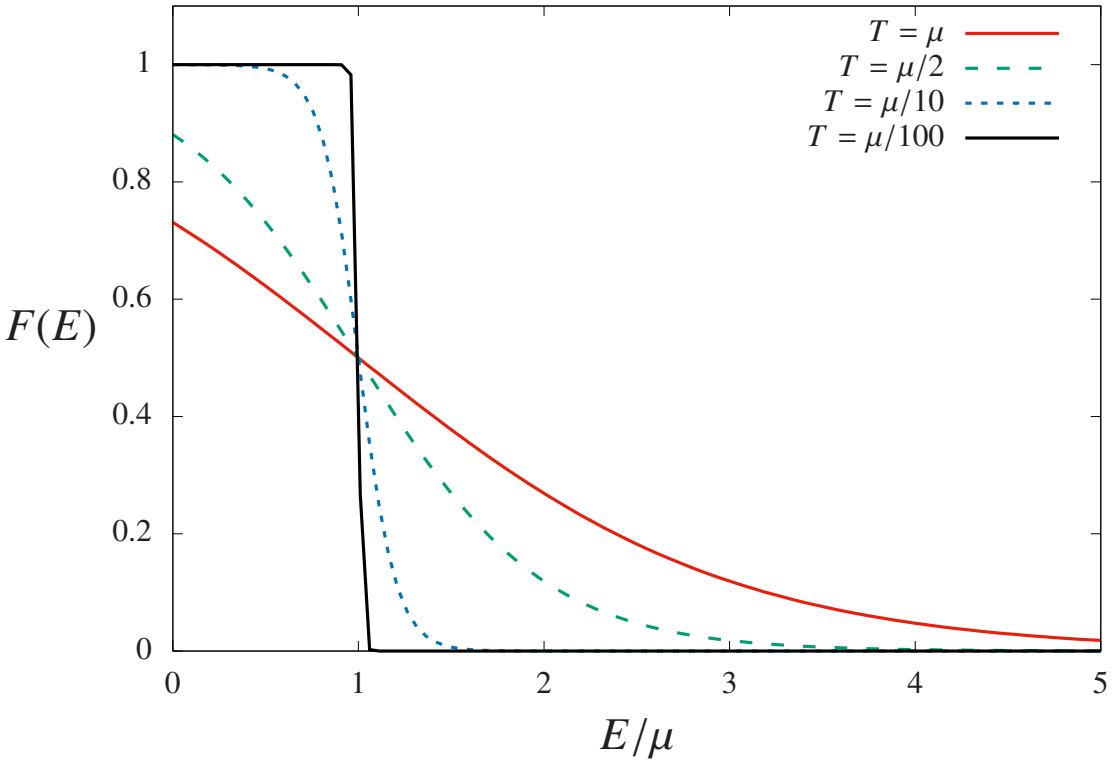
This allows the Fermi function to closely resemble the average occupation numbers $\langle n_\ell\rangle$ we computed in Section 7.5:

$$F(E) = \frac{1}{e^{\beta(E-\mu)}+1}. \quad (107)$$

As usual in the grand-canonical approach, the average particle number and Fermi function depend on both the inverse temperature $\beta$ and the chemical potential $\mu$. Expressing $F(E)$ in terms of the dimensionless ratios $E/\mu$ and $T/\mu$,

$$F(E) = \frac{1}{\exp\left[\frac{E-\mu}{T}\right]+1} = \frac{1}{\exp\left[\frac{\mu}{T}\left(\frac{E}{\mu}-1\right)\right]+1} = \frac{1}{\left(\exp\left[\frac{E}{\mu}-1\right]\right)^{\mu/T}+1},$$

we can highlight the two main features of the figure below, which plots the Fermi function against $E/\mu$ for various temperatures $T/\mu$.




First, we can see that the point $E=\mu$, where $F(E)=1/2$ for any non-zero temperature, is a threshold at which the behaviour of the Fermi function changes.

For larger energies $E > \mu$, the exponential factor $\exp\left[\frac{E}{\mu} - 1\right] > 1$ and drives $F(E) \to 0$ as the energy increases. For smaller energies $E < \mu$, the exponential factor $\exp\left[\frac{E}{\mu} - 1\right] < 1$ and vanishes as the energy decreases, leaving $F(E) \to 1$. These two asymptotic limits reflect the possible energy level occupation numbers for fermions, $n_\ell = 0$ and $1$. Second, smaller temperatures cause much more rapid approach to these two limits, with the exponential factor either enhanced (if $E > \mu$) or suppressed (if $E < \mu$) by a power $\mu/T \gg 1$. Therefore, for small temperatures $T \ll \mu$, we can simplify our calculations by approximating the Fermi function as a step function,

$$F(E) \approx \begin{cases} 1 & \text{for } E < \mu \\ 0 & \text{otherwise} \end{cases} . \tag{108}$$

Using this approximation, what is the resulting particle number density?

$$\frac{\langle N \rangle_\mathrm{f}}{V} = \frac{\sqrt{2m^3}}{\pi^2\hbar^3} \int_0^\infty F(E)\sqrt{E}\, dE =$$

You should find a result proportional to $\mu^{3/2}$ but independent of $T$. The temperature independence can be understood by viewing this as the leading-order term in an expansion in the small temperature (known as a *Sommerfeld expansion*, named after Arnold Sommerfeld). The $\mu^{3/2}$ dependence on the chemical potential is also what we would predict even without doing the detailed calculation. The step function in Eq. 108 corresponds to a single fermion occupying each and every energy level with $E_\ell < \mu$, while all energy levels with $E_\ell > \mu$ are unoccupied. Since $E(k) \propto k^2$, summing over all $k_{x,y,z}$ corresponds to computing (a portion of) the volume of a sphere of radius $r = \sqrt{\mu}$, which is proportional to $r^3 = \mu^{3/2}$ as we found directly above. If we were to invert this relation, we would obtain the so-called **Fermi energy** as a function of the particle number density,

$$E_F = \mu = \frac{\hbar^2}{2m}\left(\frac{3\pi^2 \langle N \rangle_\mathrm{f}}{V}\right)^{2/3} . \tag{109}$$

Now we can consider the average energy density of the non-relativistic fermion gas at low temperatures. Rather than taking another derivative of the grand-canonical potential, we can note from Eq. 91 and from our work on the photon gas last week that

$$\frac{\langle E \rangle_\mathrm{f}}{V} = \frac{\sqrt{2m^3}}{\pi^2\hbar^3} \int_0^\infty E\, F(E)\sqrt{E}\, dE. \tag{110}$$

That is, instead of simply counting the number of fermions in the system, we need to add up their energies, introducing an extra factor of $E$ compared to Eq. 106. Still using the low-temperature step-function approximation for the Fermi function in Eq. 108, what is the average energy density?

$$\frac{\langle E\rangle_{\mathsf{f}}}{V} = \frac{\sqrt{2m^3}}{\pi^2\hbar^3}\int_0^\infty F(E)E^{3/2}\,dE =$$

You should find

$$\langle E\rangle_{\mathsf{f}} = \frac{3}{5}\mu\,\langle N\rangle_{\mathsf{f}}\,, \tag{111}$$

which means that the average energy of the fermions in a low-temperature ideal gas is $3/5$ the Fermi energy $E_F = \mu$. In particular, because this result is also independent of the temperature, we find that non-interacting quantum fermions retain a positive energy even as the temperature approaches absolute zero, $T \to 0$. This can be understood by recalling that the lowest-energy pair of degenerate energy levels can each only hold a single fermion, forcing all additional fermions to 'fill' energy levels with larger energies $E_\ell > 0$, up to the Fermi energy set by the chemical potential. This is a stark contrast to the classical ($\langle E\rangle \propto T$) and bosonic ($\langle E\rangle \propto T^4$) cases we considered earlier, where the average energy vanishes in the zero-temperature limit. As discussed in Sections 3.4 and 7.5, in those cases all the particles in the system are able to fall into the lowest energy level, with only an exponentially small probability for particles to occupy any energy levels with $E_\ell > E_0$.

To get the rest of the way to the equation of state for the ideal gas of non-relativistic fermions, we need to compute the pressure

$$P_{\mathsf{f}} = -\left.\frac{\partial}{\partial V}\langle E\rangle_{\mathsf{f}}\right|_{N,S_{\mathsf{f}}}.$$

In the low-temperature limit, the condition of constant entropy $S_{\mathsf{f}} = -\sum_{i=1}^{M} p_i \log p_i$ is automatically satisfied, since the step function in Eq. 108 restricts the system to a single micro-state, resulting in $S_{\mathsf{f}} = 0$. This micro-state is the one in which each and every energy level with $E_\ell < \mu$ is occupied, while all energy levels with $E_\ell > \mu$ are unoccupied.

Inserting Eq. 109 into Eq. 111, we have

$$\langle E\rangle_{\mathsf{f}} = \frac{3}{5}\mu\,\langle N\rangle_{\mathsf{f}} = \frac{3}{5}\left(\frac{\hbar^2}{2m}\right)\left(\frac{3\pi^2}{V}\right)^{2/3}\langle N\rangle_{\mathsf{f}}^{5/3}.$$

The pressure is therefore

$$P_{\mathrm{f}} = -\frac{\partial}{\partial V}\left[\frac{3}{5}\left(\frac{\hbar^2}{2m}\right)\left(\frac{3\pi^2}{V}\right)^{2/3}\langle N\rangle_{\mathrm{f}}^{5/3}\right] = \frac{2}{3V}\left[\frac{3}{5}\left(\frac{\hbar^2}{2m}\right)\left(\frac{3\pi^2}{V}\right)^{2/3}\langle N\rangle_{\mathrm{f}}^{5/3}\right]$$
$$= \left(3\pi^2\right)^{2/3}\frac{\hbar^2}{5m}\left(\frac{\langle N\rangle_{\mathrm{f}}}{V}\right)^{5/3} = \frac{2}{5}\mu\frac{\langle N\rangle_{\mathrm{f}}}{V} = \frac{2}{3}\frac{\langle E\rangle_{\mathrm{f}}}{V}. \tag{112}$$

The three expressions in the second line above present several relations between the pressure, the energy density, the Fermi energy $E_F = \mu$ and the particle number density. In particular, we can see that the pressure (like the energy) remains positive even as the temperature approaches absolute zero, with

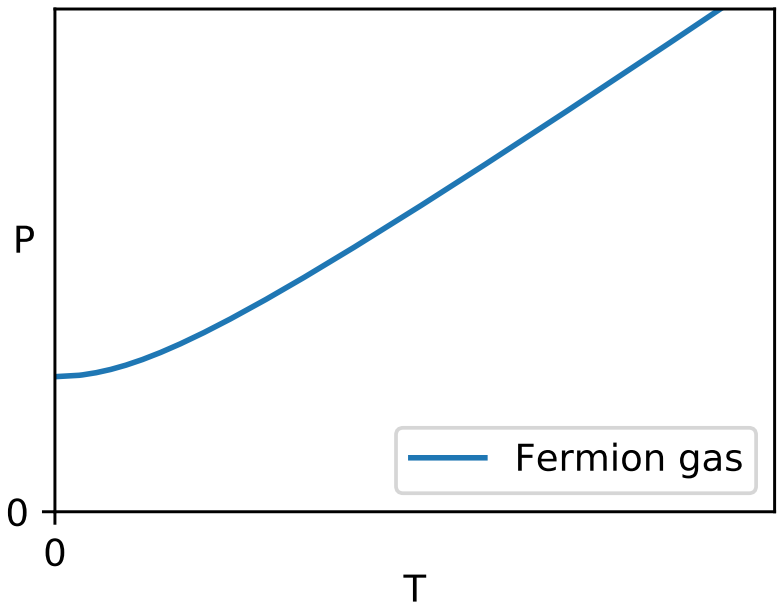
$$P_{\mathrm{f}} = \left(3\pi^2\right)^{2/3}\frac{\hbar^2}{5m}\rho_{\mathrm{f}}^{5/3}, \tag{113}$$

where we define the density $\rho_{\mathrm{f}} = \langle N\rangle_{\mathrm{f}}/V$. This positive pressure in the zero-temperature limit is not due to any direct force between the fermions, which remain non-interacting in this ideal gas. Instead, it is a purely quantum effect resulting from the Pauli exclusion principle.

As we saw earlier in this section, the temperature independence of the pressure $P_{\mathrm{f}}$ is due to approximating the low-temperature Fermi function as a step function in Eq. 108, and systematic corrections to this approximation can be computed through a Sommerfeld expansion. Even without getting into such detailed calculations, we know that at high temperatures the quantum ideal gas of massive fermions will be well approximated by the classical ideal gas we considered in Section 4.4, with equation of state

$$PV = NT \qquad \Longrightarrow \qquad P = \frac{N}{V}T = \rho T. \tag{114}$$

In words, at high temperatures the pressure depends linearly on the temperature, with the slope corresponding to the density $\rho$. The plot below (produced by this Python code) shows how the pressure changes from a positive constant as $T \to 0$ to this linear behaviour at higher temperatures.



## 9.3 Type-Ia supernovas

The positive pressure that remains for a fermion gas even at zero temperature, Eq. 113, is known as the *degeneracy pressure*. (This use of the word 'degenerate' is completely unrelated to its other use describing multiple energy levels with the same value of the energy.) The degeneracy pressure plays a key role in a famous cosmic phenomenon—a certain class of supernova explosions of stars.

As a opening remark to this topic, note that the temperature doesn't need to be exactly zero in order for the degeneracy pressure to be significant. The temperature just needs to be small compared to the Fermi energy, $T \ll E_F$. From Eq. 109 we can see that $E_F \propto \rho_{\mathrm{f}}^{2/3}$ increases for larger densities $\rho_{\mathrm{f}} = \langle N \rangle_{\mathrm{f}} / V$. Not surprisingly, the densities of stars can be very large indeed, due to the enormous amount of matter that is being squeezed together by gravitational attraction. Everyday matter has densities around $10^{28}$–$10^{29}$ atoms per cubic metre (roughly corresponding to Avogadro's number per cubic centimetre), which results in Fermi energies $E_F \sim 10^4$ K. Fermi energies for particularly dense stars known as *white dwarfs* are a hundred thousand times larger, $E_F \sim 10^9$ K, corresponding to densities of roughly one tonne per cubic centimetre. This is around a million times more dense than our sun—while white dwarf stars have a mass similar to our sun's $M_\odot$, their radius is a hundred times smaller, comparable to the radius of the earth.

White dwarf stars are so dense because they have exhausted the hydrogen and helium 'fuel' for the nuclear fusion that generates heat and light—and therefore radiation pressure—in stars such as our sun. For actively 'burning' stars, this radiation pressure counteracts the gravitational attraction of the star's enormous mass. Without nuclear fusion, white dwarfs end up gravitationally compressed into much denser and more compact objects. The degeneracy pressure, Eq. 113, coming from the (fermionic) electrons in the white dwarf is what stabilizes these stars and prevents them from collapsing further into even denser objects such as black holes.

It is remarkable that even under these extreme conditions the electrons in white dwarf stars are well described by the non-interacting ideal fermion gas we have analyzed above. In particular, it is crucial that white dwarfs' Fermi energies are so large, $E_F \sim 10^9$ K. Even though white dwarfs have burned all their nuclear fuel, their interiors remain quite hot by everyday standards, roughly ten million kelvin ($T \sim 10^7$ K). It is only due to their large densities and Fermi energies that $T \ll E_F$ and white dwarfs can be treated as zero-temperature objects to a good approximation.

So far we've seen no sign of supernovas. In isolation, white dwarfs will happily cool for trillions of years, supported by their degeneracy pressure, until they reach thermal equilibrium with the $T_{\mathrm{CMB}} \approx 2.725$ K cosmic microwave background radiation we considered last week. Things become more interesting for a white dwarf that forms a binary system with a companion star. If this companion star that is still burning hydrogen or helium through nuclear fusion, it will emit matter

that is then captured by the white dwarf, slowly increasing the white dwarf's mass. Such a binary system is pictured below, in an artist's illustration provided by the European Space Agency.

As the white dwarf accumulates the matter emitted by its companion, its mass and its density will steadily increase. As the mass of the white dwarf approaches a value about 40% larger than the mass of our sun (known as the Chandrasekhar limit, named after Subrahmanyan Chandrasekhar), its density becomes large enough for new types of nuclear fusion reactions to occur. Instead of hydrogen or helium, which the white dwarf has already burned, these new fusion reactions involve carbon and oxygen, which the white dwarf has in plenty. In the space of just a few seconds, these fusion reactions run away, increase the temperature of the white dwarf to billions of kelvin, and blast it apart in a supernova explosion about five billion times brighter than the sun.

For obscure historical reasons, these particular stellar explosions are known as *type-Ia* ("one-A") *supernovas*. They rely on the degeneracy pressure (Eq. 113) of a low-temperature gas of non-interacting fermions, which allows a specific amount of mass to build up before the explosion is triggered. The specificity of the process results in a great deal of regularity among type-Ia supernovas, which is very useful to astronomers. In particular, these type Ia supernovas play a key role in demonstrating that the expansion of the universe is accelerating (a phenomenon popularly called 'dark energy'), which was awarded the 2011 Nobel Prize in Physics.

## 9.4 Relativistic ideal fermion gas

While our main focus this week is on non-relativistic gases with $E \propto p^2$, gases of relativistic fermions also appear in nature. In fact, by changing units we can see that the white dwarf Fermi energy discussed above, $E_F \sim 10^9$ K $\sim 0.3$ MeV is comparable to the $0.511$ MeV rest-energy of electrons, suggesting that

relativistic effects may be non-negligible in white dwarfs. Such relativistic effects were indeed taken into account by Chandrasekhar and others investigating these compact stars in the twentieth century. While the full calculations required to analyze massive relativistic particles are beyond the scope of this module, we can benefit from last week's consideration of gases of massless photons to briefly consider similar gases of massless fermions. *Neutrinos* (denoted '$\nu$') are physical examples of particles whose masses are so small that they can be very well approximated as massless fermions.[23]

In the same way as photons, such massless fermions travel at the speed of light, $c$, and have energies $E = cp$ that depend on their angular frequencies,

$$E_\nu = \hbar\omega. \tag{115}$$

In a volume $V = L^3$, these energies are quantized as usual,

$$\omega = \frac{2\pi c}{\lambda} = c\frac{\pi}{L}k,$$

where $k^2 = k_x^2 + k_y^2 + k_z^2$ and $k_{x,y,z} > 0$ are positive integers. Just as for the massive fermions considered in Section 9.1, for each distinct set of integers $\vec{k} = (k_x, k_y, k_z)$, typical massless fermions such as neutrinos have two degenerate energy levels with the same energy.

The detailed analysis of a gas of massless fermions is nearly the same as the work we did last week for photon gases. In particular, massless fermions are also well described by a vanishing chemical potential, $\mu \approx 0$. Again approximating sums over discrete integer $k_{x,y,z}$ by integrals over continuous real $\widehat{k}_{x,y,z}$, and changing variables to integrate over the angular frequency, we end up with the grand-canonical potential

$$\Phi_\nu = -\frac{VT}{c^3\pi^2}\int_0^\infty d\omega\, \omega^2 \log\left[1 + e^{-\beta\hbar\omega}\right]. \tag{116}$$

The only change in $\Phi_\nu$ compared to Eq. 100 for photons are a couple of negative signs, precisely as we would expect from comparing the Bose–Einstein and Fermi–Dirac grand-canonical potentials in Section 7.5.

Carrying out the derivatives to obtain the average particle number density and the average internal energy density produces

$$\frac{\langle E\rangle_\nu}{V} = \frac{1}{c^3\pi^2}\int_0^\infty \frac{\hbar\omega^3}{e^{\beta\hbar\omega}+1}d\omega \qquad \frac{\langle N\rangle_\nu}{V} = \frac{1}{c^3\pi^2}\int_0^\infty \frac{\omega^2}{e^{\beta\hbar\omega}+1}d\omega, \tag{117}$$

now differing only by negative signs in their denominators compared to the photon gas results we obtained last week. The condition of constant entropy is still $VT^3 =$ constant, and the resulting pressure leads to an equation of state in the usual form,

$$P_\nu V \propto \langle N\rangle_\nu T,$$

with the constant of proportionality another $\mathcal{O}(1)$ number involving the Riemann zeta function.

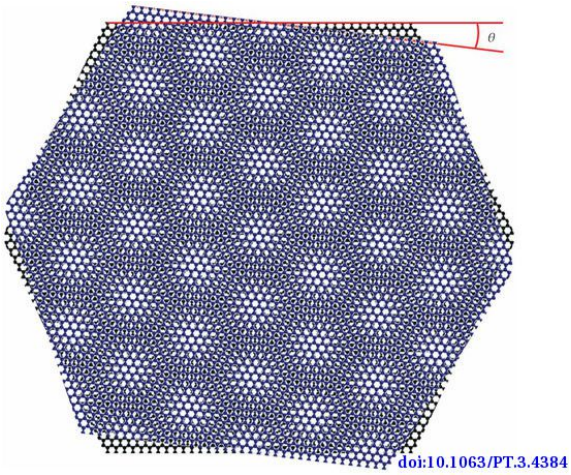
[23] Neutrinos' masses are so small that it was extremely difficult to determine that they are not exactly massless. The 2015 Nobel Prize in Physics was awarded for this discovery.

# Week 10: Interacting systems

## 10.1 From non-interacting spins to the Ising model

So far in this module we have considered 'ideal' systems whose constituent objects do not interact with each other. While we have seen that excellent mathematical models for real physical systems (such as stars and the cosmic microwave background) can be obtained despite this approximation of non-interacting particles, there are important statistical physics phenomena that cannot be captured by this approach.

An important class of examples, which we will investigate this week, are **phase transitions**, where interactions allow the same particles to produce extremely different large-scale behaviours, depending on control parameters such as the temperature or pressure. An everyday example is the transition of $H_2O$ molecules from liquid water to solid ice as the temperature decreases. As the temperature of the universe itself decreased during the first few micro-seconds following the big bang, elementary particles transitioned from a so-called quark–gluon plasma to the protons and neutrons we are made out of today. An intermediate example illustrated in the figure below (source) involves two layers of graphene at a low temperature $T \approx 1.7$ K. If these two layers are rotated with respect to each other by a small "magic angle" $\theta \approx 1.1°$, the system transitions from being an electrical insulator to being a superconductor.

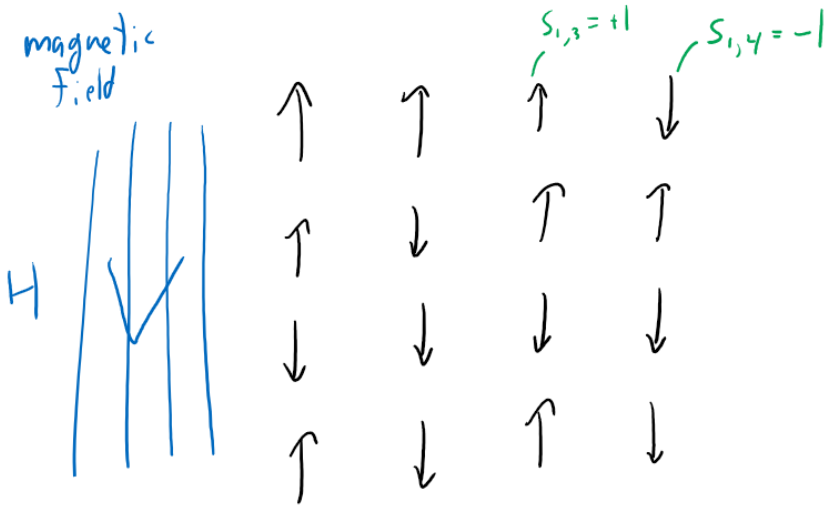



We will introduce interactions and explore their effects using simple spin systems of the sort we previously analyzed in some depth during weeks $2$ and $3$. In the non-interacting case we previously considered, the internal energy of the system (Eq. 42) is

$$E = H \sum_{n=1}^{N} s_n \qquad \text{(non-interacting)},$$

where $H > 0$ is the constant strength of an external magnetic field and the orientation of the $n$th spin, $s_n$, takes one of only two possible values: $s_n = 1$ if the spin is aligned anti-parallel to the field and $s_n = -1$ if the spin is aligned parallel

to the field. The ground state of the system features all $N$ spins aligned parallel to the magnetic field, with minimal energy $E_0 = -NH$. This week we will only consider systems of distinguishable spins, which can be labeled by their fixed position in a $d$-dimensional simple cubic **lattice** like that shown below for $d = 2$ dimensions. The $d = 1$ case of a one-dimensional lattice is precisely the system of spins arranged in a line that we analyzed in Section 3.4.1.




We can see that the total internal energy of the non-interacting system can easily be written as a sum over energies for each individual spin,

$$E_n = Hs_n \qquad\qquad E = \sum_{n=1}^{N} E_n \qquad \text{(non-interacting)}.$$

This is a generic feature of non-interacting systems, and an aspect of the **factorization** that enormously simplifies calculations by causing the $N$-particle partition function (Eq. 43) to take the form of a product of $N$ identical terms, $Z = [2\cosh(\beta H)]^N = Z_1^N$. However, a stronger condition needs to be satisfied in order for such factorization to be guaranteed, which rigorously defines what it means for a system to be non-interacting.

Let $\Delta E_i$ be the change in the system's internal energy caused by changing its $i$th degree of freedom. Then the system is defined to be **non-interacting** if and only if $\Delta E_i$ is independent of all other degrees of freedom $k \neq i$.

For our system of $N$ distinguishable spins, the only possible change we can make to a degree of freedom is to negate it, $s_i \to -s_i$, which corresponds to flipping its alignment relative to the external magnetic field. This spin flip causes the total energy to change,

$$E = H\sum_{n=1}^{N} s_n = H\left(s_i + \sum_{k\neq i} s_k\right) \quad \longrightarrow \quad H\left(-s_i + \sum_{k\neq i} s_k\right),$$

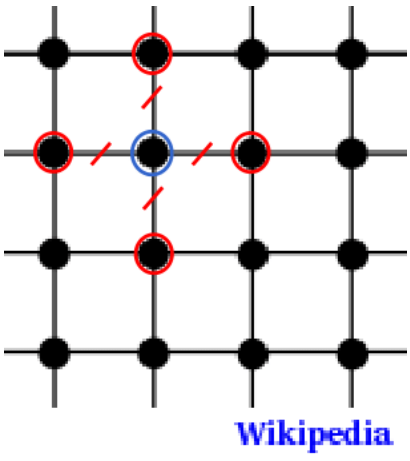
corresponding to $\Delta E_i = -2Hs_i$, which is indeed independent of all spins $s_k$ with $k \neq i$. This simple check confirms that our definition works for the non-interacting spin system under consideration.

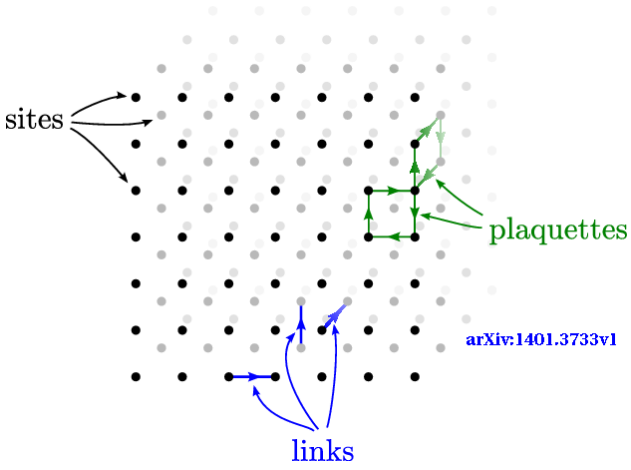
Now let's convert this setup into a system of interacting spins by adding the simplest possible two-spin contribution to its energy:

$$E = -\sum_{(ij)} s_i s_j + H \sum_{n=1}^{N} s_n. \tag{118}$$

The first sum runs over all pairs of nearest-neighbour spins in the lattice, denoted $(ij)$. What is the change in energy $\Delta E_i$ from Eq. 118 upon negating $s_i \to -s_i$? Does this indicate an interacting or non-interacting system?

The pictures below illustrate nearest-neighbour pairs for simple cubic lattices in $d = 2$ and $3$ dimensions, while also introducing some additional lattice terminology.

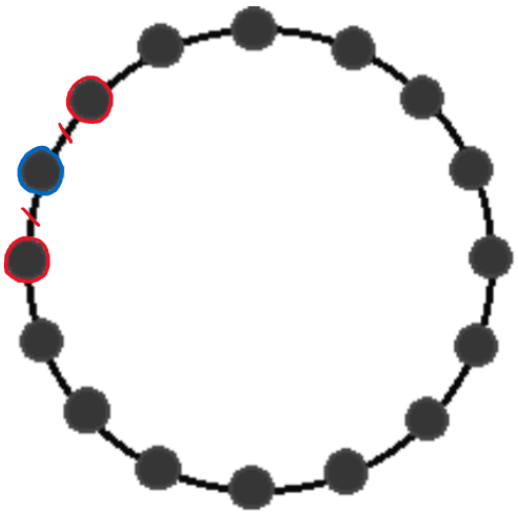



Instead of drawing up- and down-pointing arrows, these pictures identify the spins with *sites* in the lattice represented as points (or larger filled circles). In simple cubic lattices, all sites are positioned in a regular grid, separated by a constant distance along each basis vector. In between nearest-neighbour sites, we can draw *links* as solid lines. The picture of a two-dimensional lattice on the left highlights the four links (with red hatch marks) that correspond to the four

nearest neighbours (circled in red) of a particular site (circled in blue). While we can only have physical lattices with $d = 1$, $2$ or $3$ in nature, the mathematical construction works just as well for any integer $d \geq 1$. For $d \geq 2$, an elementary unit of surface area is called a *plaquettes*, while for $d \geq 3$ the elementary unit of volume is called a *cube*.

Computing the energy in Eq. 118 requires determining all of the nearest-neighbour pairs to be summed in the first term, which is equivalent to all of the links in the lattice, $\ell = (ij)$. The only potential complication to this task is the need to consider what happens at the edges of the (finite) lattice. We can avoid this complication by imposing **periodic boundary conditions**, which add an extra link between each site on the left edge of the lattice and a corresponding site on the right edge (and similarly in all other dimensions). This is illustrated below for the simple one-dimensional lattice, which has been drawn as a circle to emphasize that all $N$ sites remain separated by a constant distance. In higher dimensions, periodic boundary conditions produce flat (zero-curvature) $d$-dimensional tori that preserve the simple cubic lattice structure.



With periodic boundary conditions, we can easily see that the $N$-site one-dimensional lattice drawn above has $N$ links. Each site has two links connecting it to its two nearest neighbours, and each of those links is shared between two sites, so that $\#\ell = 2N/2 = N$. Looking back to the two-dimensional lattice drawn farther above, the four links per site produce $\#\ell = 4N/2 = 2N$. How many terms are there in the sum $\sum_{(ij)}$ in Eq. 118 for $N$-site lattices with periodic boundary conditions in $d$ dimensions?

The energy in Eq. 118, with nearest-neighbour spins specified by the underlying simple cubic lattice structure, defines a famous system known as the

$d$-dimensional **Ising model**. Since the 1960s, the Ising model has been the basis of thousands of scientific studies analyzing everything from ferromagnetism to neural networks to urban segregation.[24] The model was proposed in 1920 by Wilhelm Lenz, whose PhD student Ernst Ising solved the one-dimensional system as a research project in 1924. Exactly solving the two-dimensional case (with $H = 0$) took another twenty years, culminating in renowned work by Lars Onsager in 1944. The three-dimensional Ising model remains an open mathematical question, with no known exact solution.

In this context, 'solving' the Ising model means deriving a closed-form expression for its canonical partition function,

$$Z(\beta, N, H) = \sum_{\{s_n\}} \exp\left[-\beta E(s_n)\right] = \sum_{\{s_n\}} \exp\left[\beta \sum_{(ij)} s_i s_j - \beta H \sum_n s_n\right].$$

As in Section 3.4, the partition function sums over all possible spin configurations $\{s_n\}$, which amounts to a sum of $2^N$ exponential factors for $N$ spins, with $\mathcal{O}(N)$ terms within each exponential. Now that the system is interacting, the partition function can no longer be factorized into $N$ identical two-term factors, making it extremely difficult to evaluate. This is why there is no known exact solution to the three-dimensional Ising model, and it also makes 'brute-force' numerical computations impractical. Even for a system of $N = 1023$ spins (tiny compared to Avogadro's number $\sim 10^{23}$) there would be roughly $2^{1023} \sim 10^{310}$ terms in the partition function, far beyond the capabilities of existing or foreseeable supercomputers.

## 10.2 Ising model phases and phase transition

Similarly to how we analyzed non-interacting spin systems in Section 3.4, we can simplify the Ising model by considering its behaviour in the limits of high and low temperatures. For another simplification, we will set $H = 0$ in this section, and consider

$$E = -\sum_{(ij)} s_i s_j \qquad Z(\beta, N) = \sum_{\{s_n\}} \exp\left[\beta \sum_{(ij)} s_i s_j\right]. \tag{119}$$

We will see that the large-scale behaviour of the Ising model is qualitatively different at high temperatures compared to low temperatures. In other words, the system exhibits two distinct phases for different temperatures. This is a necessary but not sufficient condition for there to be a true phase transition—a priori, it is possible for there to be a gradual *crossover* between these two phases, as opposed to a rapid transition. We will use the Ising model to more rigorously define what exactly constitutes a phase transition, and how this can be distinguished from a crossover.

[24] For a brief summary, see Charlie Wood, "The Cartoon Picture of Magnets That Has Transformed Science", *Quanta Magazine*, 24 July 2020.

The Ising model partition function becomes extremely simple in the **high-temperature** limit $\beta \to 0$. What is its asymptotic limit?

$Z(\beta = 0, N) =$

You should find a result identical to that for a micro-canonical system with energy $E = 0$, which is clearly non-interacting since $\Delta E_i = 0$. Every spin configuration is a different micro-state of the system, all with the same probability $p_i = 1/2^N$, as in Eq. 20.

Effectively, we are considering temperatures so high that the energy from Eq. 119 is negligible for any spin configuration. Although the energy no longer distinguishes between different micro-states, we can define a quantity that continues to be sensitive to the details of the spin configuration. This is the **magnetization** $M = n_+ - n_-$, where we redefine $n_\pm$ to be the number of spins with value $\pm 1$, so that $N = n_+ + n_-$.[25] It is convenient to normalize the total magnetization $M$ by the number of spins,

$$|m| \equiv \frac{|M|}{N} = \frac{|n_+ - n_-|}{n_+ + n_-} \tag{120}$$

so that $0 \leq |m| \leq 1$ for any value of $N$.

Our task is now to determine the expectation value of the magnetization at high temperatures. Above we found that all spin configurations are equally probable in this regime, so $\langle |m| \rangle$ will be determined by how likely it is for these micro-states to have a particular magnetization. For example, there are only two micro-states with $|m| = 1$, corresponding to $(n_+, n_-) = (N, 0)$ and $(0, N)$. In general, just as we saw in Eq. 24, there are

$$\binom{N}{n_+} = \binom{N}{n_-} = \frac{N!}{n_+!\, n_-!}$$

equally probable micro-states with a given $n_+ = N - n_-$. For large $N \gg 1$ this binomial coefficient is factorially peaked around

$$n_+ = n_- = \frac{1}{2}N \qquad \longrightarrow \qquad |m| = 0,$$

which defines a **disordered phase** with similar numbers of up- and down-pointing spins producing a small magnetization. In the so-called *thermodynamic limit* $N \to \infty$, the expectation value of the magnetization in the disordered phase vanishes exactly, $\langle |m| \rangle \to 0$.

[25] In week 2 we defined $n_\pm$ based on spins' alignment with or against the external magnetic field, which no longer applies now that we have set $H = 0$.

We now need to determine $\langle|m|\rangle$ in the **low-temperature** limit $\beta \to \infty$. In this regime, as we saw in Section 3.4.1, the Boltzmann factor $\exp\left[\beta \sum_{(ij)} s_i s_j\right]$ makes it exponentially more likely for the system to adopt micro-states with lower energies. In particular, we can expect the ground state to dominate the expectation value of the magnetization, $\langle|m|\rangle$, up to exponentially suppressed corrections from higher-energy excited states. With $H = 0$, the Ising model has two degenerate ground states corresponding to the two ways all the spins can be aligned with each other: $(n_+, n_-) = (N, 0)$ and $(0, N)$. What is the ground-state energy of the $N$-site Ising model in $d$ dimensions?

$$E_0 = -\sum_{(ij)} s_i s_j =$$

As mentioned above, both of these degenerate ground states have the maximal magnetization $|m| = 1$. Let's check what effect the first excited state would have on the overall magnetization of the system. The first excited state involves negating (or 'flipping') a single spin, corresponding to $(n_+, n_-) = (N-1, 1)$ and $(1, N-1)$. Because any one of the $N$ spins in the lattice could be flipped, the degeneracy of the first excited state grows with $N$:

$$\binom{N}{1} + \binom{N}{N-1} = 2N.$$

At the same time, as $N$ increases the magnetization of each of these micro-states gets closer to that of the ground state,

$$|m| = \frac{N-1}{N} = 1 - \frac{1}{N}.$$

The key factor is the probability for the system to be in one of these micro-states, which depends on the energy of the first excited state, $E_1$. What is the first-excited-state energy of the $N$-site Ising model in $d$ dimensions?

$$E_1 =$$

Let's put things together by computing the relative probability for the $d$-dimensional Ising model to be in its ground state with $|m| = 1$ compared to its

first excited state with $|m| = 1 - \frac{1}{N}$, accounting for the different number of microstates in each case:

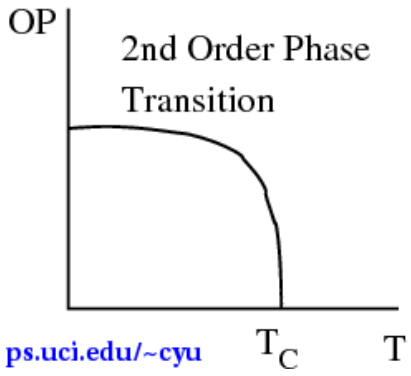
$$\frac{p(E_0)}{p(E_1)} = \frac{2 \cdot \exp\left[\beta d \cdot N\right]}{2N \cdot \exp\left[\beta\left(d \cdot N - 4d\right)\right]} = \frac{\exp\left[4\beta d\right]}{N}.$$

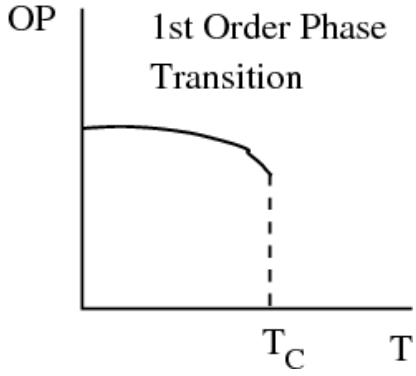
For any fixed $N$, a sufficiently low temperature will cause the ground state to dominate. This defines an **ordered phase** in which essentially all spins are aligned in the same direction, producing a large expectation value for the magnetization, $\langle |m| \rangle = 1$.

We have seen that the magnetization $\langle |m| \rangle$ distinguishes between the high- and low-temperature behaviour of the $d$-dimensional Ising model. In the high-temperature disordered phase, the magnetization is small and $\langle |m| \rangle \to 0$ in the thermodynamic limit $N \to \infty$. In the low-temperature ordered phase, the magnetization is large and $\langle |m| \rangle \to 1$ as $T \to 0$. This is typical behaviour for interacting statistical systems, where the quantity distinguishing between these two phases (here the magnetization) is known as the **order parameter**. The behaviour of the order parameter is what distinguishes gradual crossovers from rapid phase transitions.

A phase transition is characterized by a discontinuity in the order parameter or its derivative(s), in the $N \to \infty$ thermodynamic limit. The value(s) of the control parameter(s) at which the discontinuity occurs define the *critical point* corresponding to the transition.

For the zero-field ($H = 0$) Ising model, the control parameter is the temperature $T$, and any phase transition would occur at a **critical temperature** $T_C$. The sketches below illustrate the most common types of phase transitions. When the order parameter (OP) itself is discontinuous (shown by a dashed line), the transition is said to be a *first-order* phase transition. When the order parameter is continuous at $T_C$ but its first derivative is discontinuous, the transition is said to be a *second-order* phase transition. This naming scheme generalizes to higher-order phase transitions, the most remarkable of which is the infinite-order BKT phase transition (named after Vadim Berezinskii, J. Michael Kosterlitz and David Thouless), which was awarded the 2016 Nobel Prize in Physics.



Because true discontinuities are only possible with an infinite number of degrees of freedom, it is really the way in which the system approaches the $N \to \infty$ thermodynamic limit that distinguishes crossovers from true phase transitions. We will conclude this week's consideration of interacting systems by developing a useful approximation for the Ising model, which will turn out to be more reliable as the dimensionality $d$ increases.

## 10.3 The mean-field approximation

Looking back to Eq. 120, we can identify the magnetization (with no absolute value) as the average spin,

$$m = \frac{M}{N} = \frac{1}{N}\left(n_+ - n_-\right) = \frac{1}{N}\sum_{n=1}^{N} s_n.$$

The 'mean' in the mean-field approximation refers to the expectation value of the average spin,

$$\langle m \rangle = \frac{1}{Z}\sum_{\{s_n\}} m\, e^{-\beta E(s_n)} = \frac{1}{N}\sum_{n=1}^{N} \langle s_n \rangle\,,$$

which is independent of the spin configuration $\{s_n\}$ and is simply a function of the inverse temperature $\beta$ and magnetic field strength $H$.[26] By adding and subtracting factors of $\langle m \rangle$, we can exactly rewrite each nearest-neighbour term in the Ising model energy, Eq. 118, as

$$\begin{aligned} s_i s_j &= \left[\left(s_i - \langle m \rangle\right) + \langle m \rangle\right] \times \left[\left(s_j - \langle m \rangle\right) + \langle m \rangle\right] \\ &= \left(s_i - \langle m \rangle\right)\left(s_j - \langle m \rangle\right) + \left(s_i + s_j\right)\langle m \rangle - \langle m \rangle^2 . \end{aligned} \tag{121}$$

The factors of $\left(s_i - \langle m \rangle\right)$ correspond to the spins' fluctuations around their mean value $\langle m \rangle$. By conjecturing that these fluctuations are small *on average*, we can approximate the Ising model energy by neglecting the first term in Eq. 121 when summing over all links:

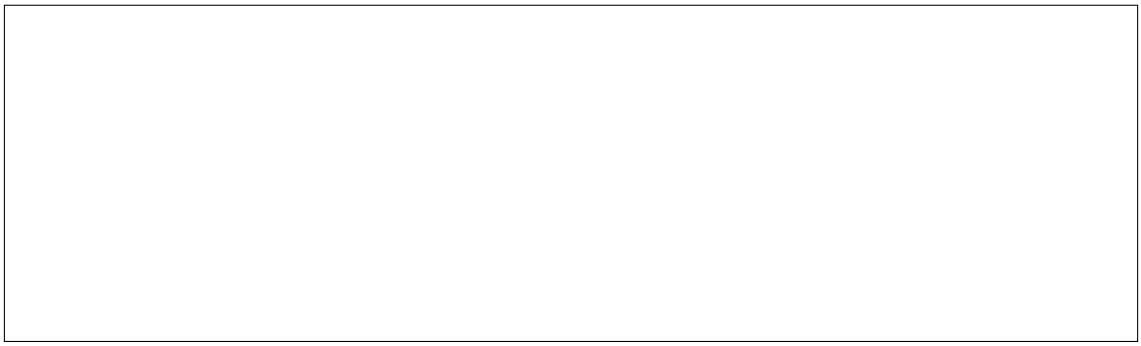
$$E = -\sum_{(ij)} s_i s_j - H\sum_{n=1}^{N} s_n \quad \longrightarrow \quad E_{\mathsf{MF}} = -\sum_{(ij)}\left[\left(s_i + s_j\right)\langle m \rangle - \langle m \rangle^2\right] - H\sum_{n=1}^{N} s_n.$$

The sum over the links $\ell = (ij)$ in $d$ dimensions simply counts $d \cdot N$ factors of the constant $\langle m \rangle^2$. Similarly, since the first term includes both spins $(s_i + s_j)$ on each end of the link, every individual spin appears $2d$ times in the sum over links, which we can combine with the sum over sites:

$$E_{\mathsf{MF}} = d \cdot N \langle m \rangle^2 - \left(2d\langle m \rangle + H\right)\sum_{n=1}^{N} s_n \equiv d \cdot N \langle m \rangle^2 - H_{\mathsf{eff}}\sum_{n=1}^{N} s_n, \tag{122}$$

[26] To be consistent with the redefinition of $n_\pm$ above Eq. 120, we reverse our sign convention for the magnetic field, $H \to -H$.

defining an effective magnetic field $H_{\text{eff}}$ that depends on the mean spin. What is the change in energy $\Delta E_i$ from Eq. 122 upon negating $s_i \to -s_i$? Does this indicate an interacting or non-interacting system?



The mean-field approximation producing Eq. 122 makes it very easy to compute the corresponding canonical partition function

$$\begin{aligned} Z_{\text{MF}} &= \sum_{\{s_n\}} \exp\left[-\beta E(s_n)\right] = \exp\left[-\beta d \cdot N \left\langle m \right\rangle^2\right] \sum_{s_1=\pm 1} \cdots \sum_{s_N=\pm 1} \exp\left[-x \sum_{n=1}^{N} s_n\right] \\ &= \exp\left[-\beta d \cdot N \left\langle m \right\rangle^2\right] \left(2\cosh\left[\beta H_{\text{eff}}\right]\right)^N \\ &= \exp\left[-\beta d \cdot N \left\langle m \right\rangle^2\right] \left(2\cosh\left[\beta\left(2d\left\langle m \right\rangle + H\right)\right]\right)^N, \end{aligned} \tag{123}$$

where we defined $x = -\beta H_{\text{eff}}$ to put the sums into exactly the same form as Eq. 43. We see that the mean-field partition function exhibits complicated dependence on $\langle m \rangle$. In order to get a handle on this dependence, we need to determine another relation between $\langle m \rangle$ and $Z_{\text{MF}}$.

We can do this by recalling the connection between the magnetization and the average spin discussed at the start of this section. Because the total $N$-spin magnetization

$$M = n_+ - n_- = \sum_{n=1}^{N} s_n,$$

we can write the full Ising model energy as

$$E = -\sum_{(ij)} s_i s_j - H \sum_{n=1}^{N} s_n = -\sum_{(ij)} s_i s_j - HM,$$

with corresponding canonical partition function

$$Z = \sum_{\{s_i\}} \exp\left[\beta \sum_{(ij)} s_i s_j + \beta H M\right].$$

From our earlier experience with the canonical ensemble, it now comes as no surprise that $\langle M \rangle = N \langle m \rangle$ can be related to a derivative of the Helmholtz free energy $F = -T \log Z$. What is this relation?

$$\frac{\partial}{\partial H} F =$$

Returning to the mean-field approximation,

$$\log Z_{\mathsf{MF}} = N \log \cosh \left[ \beta \left( 2d \langle m \rangle + H \right) \right] + \{ H\text{-independent terms} \} ,$$

we can use this relation between $\langle m \rangle$ and $Z_{\mathsf{MF}}$ to find

$$\langle m \rangle = \frac{1}{N\beta} \frac{\partial}{\partial H} \log Z_{\mathsf{MF}} = \frac{1}{\beta} \frac{1}{\cosh \left[ \beta \left( 2d \langle m \rangle + H \right) \right]} \frac{\partial}{\partial H} \cosh \left[ \beta \left( 2d \langle m \rangle + H \right) \right] .$$

Simplifying, we obtain a **self-consistency condition** for the magnetization in the mean-field approximation:

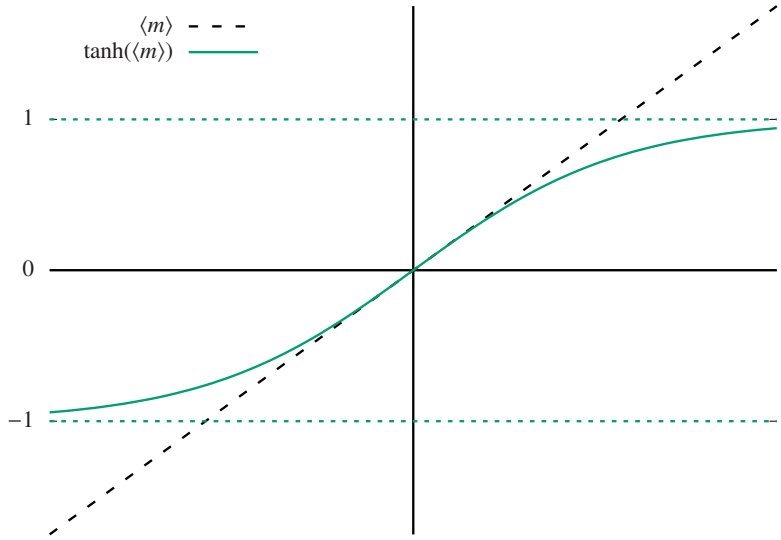
$$\langle m \rangle = \tanh \left[ \beta \left( 2d \langle m \rangle + H \right) \right] . \tag{124}$$

Solving this equation for $\langle m \rangle$ is equivalent to finding the roots of the equation $\tanh \left[ \beta (2d \cdot x + H) \right] - x = 0$.

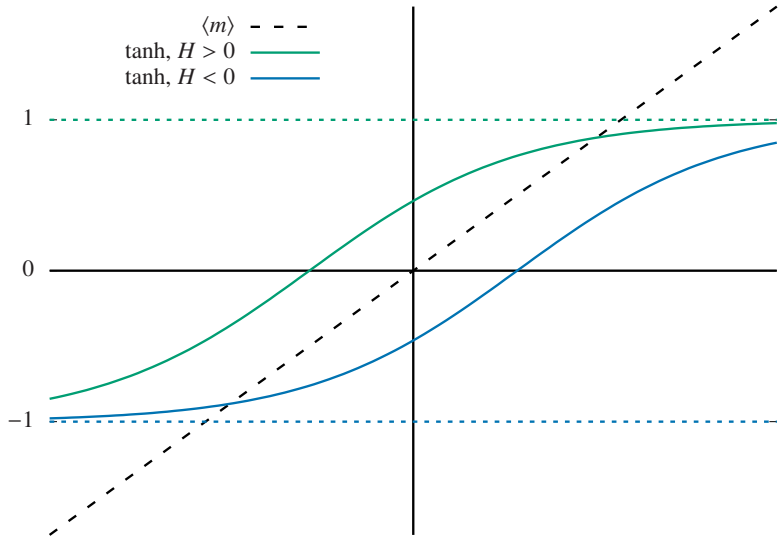
A straightforward way to inspect such solutions is by plotting both

$$f(\langle m \rangle) = \langle m \rangle \qquad\qquad g(\langle m \rangle) = \tanh \left[ \beta (2d \langle m \rangle + H) \right]$$

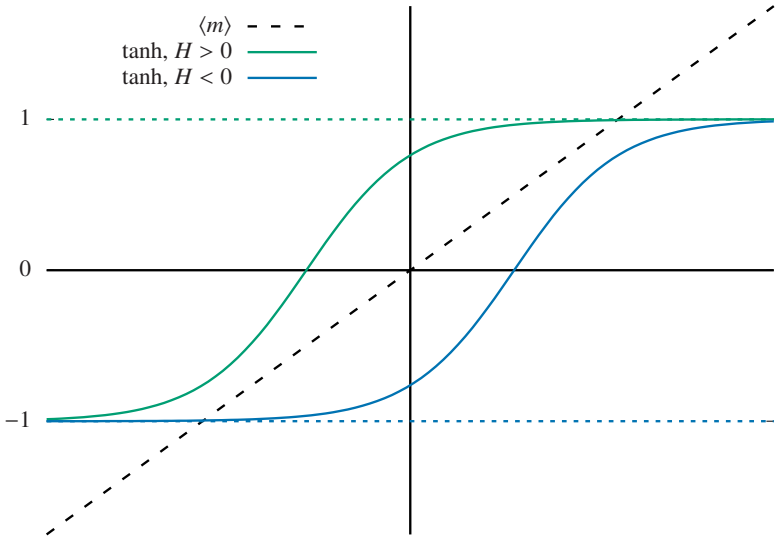
and monitoring the intersections of these two functions. Fixing $d = 2$ dimensions, the plot below considers the simplest case $\beta = \frac{1}{4}$ and $H = 0$ for which $g(\langle m \rangle) = \tanh \left[ \langle m \rangle \right]$ (the solid line). There is only a single intersection between this function and $f(\langle m \rangle)$ (the dashed line), at $\langle m \rangle = 0$.



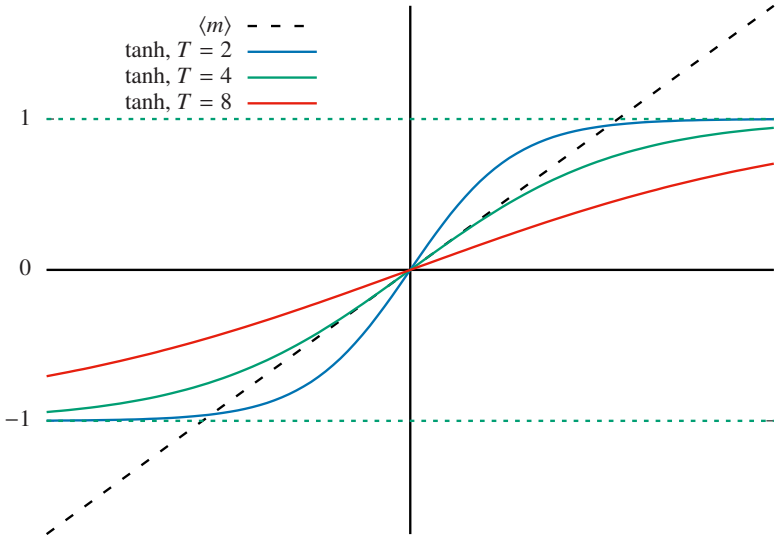

Before interpreting this result, let's check how the intersections depend on $\beta$ and $H$. In the next plot below we keep the same temperature $T = 1/\beta = 4$ while turning the external magnetic field back on. A positive $H > 0$ simply shifts $g(\langle m \rangle)$ to the left (the green line), while a negative $H < 0$ shifts it to the right (the blue line). For $H = \pm 2$, there is still only a single intersection, at $\langle m \rangle \approx \pm 0.88$. We can see that because $-1 \leq \tanh x \leq 1$, the mean-field self-consistency condition can only ever predict $-1 \leq \langle m \rangle \leq 1$, in accord with the definition of the magnetization.




Next, if we decrease the temperature, $g(\langle m \rangle)$ becomes steeper, more rapidly interpolating between those limiting values $-1 \leq \tanh x \leq 1$. The plot below illustrates this for $T = 1/\beta = 2$, so that $\beta = \frac{1}{2}$ is doubled. Already for this temperature and magnetic field $H = \pm 2$, the intersection is $\langle m \rangle \approx \pm 1$ to a very good approximation. We can interpret this result as an indication that the system is in an ordered phase where all spins are aligned with the external field.



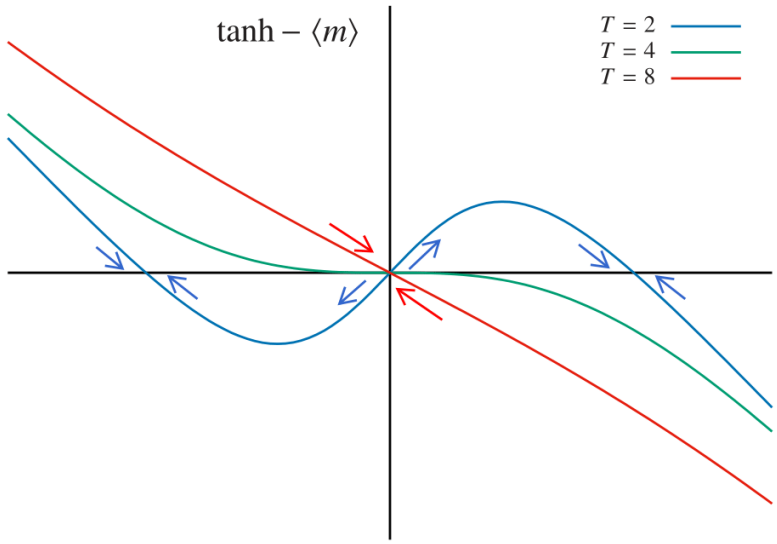

To search for a phase transition we need to turn off that external field by setting $H = 0$, and consider how the solutions of the self-consistency condition depend on the temperature. We do this in the next plot below, considering a low temperature $T = 2$ with $\beta = \frac{1}{2}$ (the red line), the same green curve for $T = 4$ shown in the first plot above, and a high temperature $T = 8$ with $\beta = \frac{1}{8}$ (the blue line). While the $\langle m \rangle = 0$ expected in the disordered phase is always a possible solution, something interesting happens at lower temperatures, where the steeper $\tanh$ function introduces two additional solutions at $\langle m \rangle = \pm m_0$. As $T \to 0$, this $m_0 \to 1$, approaching the magnetization of the ordered phase.




When there are three solutions $\langle m \rangle = \{-m_0, 0, m_0\}$ at low temperatures, we can determine that the $\langle m \rangle = 0$ solution is actually *unstable*. In this case, the slope of the $\tanh$ at $\langle m \rangle = 0$ must be greater than $1$. Any positive value $\langle m \rangle = \varepsilon > 0$ would therefore produce $\tanh\left[2\beta d \langle m \rangle\right] > \langle m \rangle$, violating the self-consistency condition in Eq. 124. Since the violation comes from $\langle m \rangle$ being too

small compared to the $\tanh$, $\langle m \rangle$ would be driven to increase further, until it eventually reached the non-zero solution $\langle m \rangle = m_0$. The analogous argument implies any negative value $\langle m \rangle = -\varepsilon < 0$ would drive $\langle m \rangle$ away from zero and to the $\langle m \rangle = -m_0$ solution.

This argument can be visualized by plotting $\tanh\left[2\beta d \langle m \rangle\right] - \langle m \rangle$ vs. $\langle m \rangle$ as shown in the final plot below. Whenever this difference is negative, it implies $\langle m \rangle$ is larger than the self-consistency condition allows, and therefore consistency requires reducing $\langle m \rangle$, shown by arrows pointing to the left. Conversely, whenever the difference is positive, $\langle m \rangle$ would have to increase to recover consistency, shown by the arrows pointing to the right. For the low temperature $T = 2$, we see that the arrows move the system away from the unstable solution $\langle m \rangle = 0$ and to the stable solutions $\langle m \rangle = \pm m_0$.




Reassuringly, the mean-field approximation with $H = 0$ therefore reproduces the behaviour we derived in the previous section. For high temperatures we have $\langle m \rangle = 0$ consistent with the disordered phase, while for low temperatures we have $|\langle m \rangle| = m_0 \to 1$ consistent with the ordered phase. We can also determine the temperature at which the $\langle m \rangle = \pm m_0$ solutions appear and the $\langle m \rangle = 0$ solution becomes unstable. As described above, this occurs whenever the slope of the $\tanh$ function at $\langle m \rangle = 0$ is greater than $1$. Using $\tanh(x) = x + \mathcal{O}\left(x^3\right)$ for $x \approx 0$, what is the slope of this $\tanh$?

$$\frac{d}{d\langle m \rangle} \tanh\left[2\beta d \langle m \rangle\right]\Big|_{\langle m \rangle = 0} =$$

You should find that the change from the high-temperature disordered phase to the low-temperature ordered phase occurs at $T_c = 2d$ in $d$ dimensions, or equiv-

alently $\beta_c = \frac{1}{2d}$. Despite the subscript, we don't yet know if this $T_c$ is a true critical temperature. In order to check whether the mean-field approximation of the $H = 0$ Ising model predicts a crossover or a true phase transition, we need to check whether or not the order parameter $\langle m \rangle$ or its derivatives are discontinuous at $T_c$. We can do this by considering the self-consistency condition for a temperature $T = 1/\beta$ lower than but close to $T_c = 2d$, which would produce $0 < |\langle m \rangle| \ll 1$ and allow us to expand

$$\langle m \rangle = \tanh\left[\frac{2d}{T}\langle m \rangle\right] \approx \frac{T_c}{T}\langle m \rangle - \frac{1}{3}\left(\frac{T_c}{T}\langle m \rangle\right)^3 .$$

What is the resulting prediction for $\langle m \rangle$?

Approximating $\frac{T}{T_c} \approx 1$, your result should resemble

$$\langle m \rangle = \pm\sqrt{3}\left(\frac{T_c - T}{T_c}\right)^{1/2} \qquad \text{for } T \lesssim T_c.$$

From this, we can see that the order parameter $\langle m \rangle$ is continuous at $T_c$:

$$\langle m \rangle \propto \begin{cases} (T_c - T)^{1/2} & \text{for } T \lesssim T_c \\ 0 & \text{for } T \gtrsim T_c \end{cases} . \tag{125}$$

However, its first derivative

$$\frac{d\langle m \rangle}{dT} \propto \frac{1}{(T_c - T)^{1/2}}$$

diverges as $T \to T_c$ from below, predicting a second-order phase transition with critical temperature $T_c = 2d$ in $d$ dimensions. The power-law dependence $\langle m \rangle \propto (T_c - T)^b$ is a generic feature of second-order phase transitions, where the power $b$ is known as a **critical exponent**, in this case $b = 1/2$.

At this point we have invested some effort to find that the mean-field approximation of the $d$-dimensional Ising model, with $H = 0$, predicts a second-order phase transition at $T_c = 2d$ with critical exponent $1/2$. Let's wrap up by quickly checking the reliability of the mean-field approximation and the accuracy of these results it has given us.

The accuracy of the mean-field results turns out to depend on the number of dimensions. For the one-dimensional ($d = 1$) Ising model that Ising himself

solved, there is no phase transition at all.[27] The mean-field approximation fails badly in this case.

The situation improves for the two-dimensional Ising model. Onsager's exact $H = 0$ solution reveals a second-order phase transition, with inverse critical temperature $\beta_c = \frac{1}{2}\log\left(1+\sqrt{2}\right) \approx 0.44$ and $\langle m \rangle \propto (T_c - T)^{1/8}$ for $T \lesssim T_c$ corresponding to a critical exponent $1/8$. While the mean-field approach now provides the correct qualitative behaviour, its prediction $\beta_c = \frac{1}{2d} = 0.25$ is off by almost a factor of $2$, while the mean-field critical exponent $b = 1/2$ is four times larger than the true $b = 1/8$.

For higher dimensions $d \geq 3$ there is no known exact solution for the Ising model, but the existence of a second-order phase transition can be established, while the corresponding critical temperature and critical exponents can be computed numerically. In three dimensions the mean-field $T_c = 2d = 6$ and $b = 1/2$ are still only rough first approximations to the true $T_c \approx 4.5$ and $b \approx 0.32$. The mean-field prediction for the critical exponent $b = 1/2$ turns out to be correct for $d \geq 4$, while the critical temperature $T_c = 2d$ gradually approaches the true value as the number of dimensions increases. (Numerical computations find $T_c \approx 6.7$, $8.8$, $10.8$ and $12.9$ for $d = 4$, $5$, $6$ and $7$, respectively.) Formally, the mean-field approximation exactly reproduces the Ising model in the unphysical limit of infinite dimensions, $d \to \infty$. Roughly speaking, the greater reliability of the mean-field approach in higher dimensions is due to the larger number of nearest neighbours for each site, $2d$, which allow that site to interact with a better approximation to the mean spin.

[27] You can find the exact solution of the one-dimensional Ising model in Section 5.3.1 of David Tong's *Lectures on Statistical Physics* (reference 1 in the list of further reading on page 6).

# Week 11: Synthesis and broader applications

## 11.1 Monte Carlo importance sampling

Last week we wrapped up our discussion of the mean-field approximation to the Ising model by comparing some of its predictions against results from numerical computations for systems in $d \geq 3$ dimensions where no exact solution is known. These numerical results may have come as a surprise given the statement at the end of Section 10.1 that numerically evaluating the Ising model partition function even for tiny systems with $N \sim 1000$ spins is far beyond the capabilities of existing or foreseeable supercomputers. To quantify 'tiny', consider that $N \sim 1000$ would correspond to a $10 \times 10 \times 10$ lattice in three dimensions or a $6 \times 6 \times 6 \times 6$ lattice in four dimensions, both very far from the thermodynamic limit of interest for phase transitions.

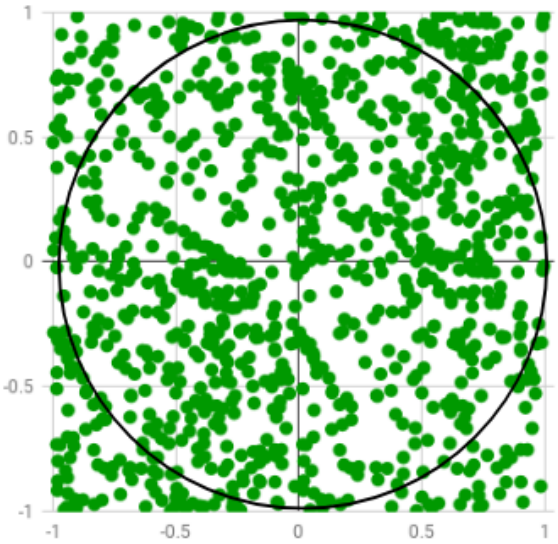
The key is that practical numerical computations do not perform a 'brute-force' evaluation of every single micro-state that is summed over in the formal definition of the partition function (and hence in the expectation values that we have derived from the partition function). Instead, they proceed by (pseudo-)randomly selecting (or **sampling**) a small subset of those micro-states, and using this subset to compute results for the average energy, magnetization, and other thermodynamic quantities. The law of large numbers allows us to treat these averages as controlled approximations to the true ensemble expectation values.

As we saw in the computer project, such numerical calculations employ pseudo-random numbers rather than complete randomness, which allows them to be reproducible up to very high precision by different people using different computers. Due to the role of randomness, these numerical approaches have become known as **Monte Carlo** methods, based on a whimsical reference to the famous gambling centre in Monaco. Monte Carlo methods are crucial in statistical physics because they are applicable even to the interacting systems introduced last week, where there are no longer dramatic simplifications from factorization.

An intuitive illustration of how Monte Carlo methods work is provided by using them to numerically evaluate integrals. The idea is that the integral can be numerically approximated by evaluating its integrand at randomly sampled points in the integration domain, and normalizing by the number of samples. A simple example is to compute

$$\int_{-1}^{1} dx \int_{-1}^{1} dy \; H\big(1 - \{x^2 + y^2\}\big) = \pi \qquad\qquad H(r) = \begin{cases} 1 & \text{for } r \geq 0 \\ 0 & \text{for } r < 0 \end{cases} ,$$

which is just the area of a disk with radius $R = 1$, in a square integration domain with area $4$, as shown below. Since the integrand is either $0$ or $1$ for each randomly sampled point in that domain, simply counting the fraction of the $S$ samples that lie in the disk provides a numerical determination of $\pi$, with a *statistical uncertainty* that vanishes $\sim 1/\sqrt{S}$. In just a few minutes, this Python code predicts $\pi = 3.14155 \pm 0.00024$ purely from the statistics of pseudo-random numbers.




Of course, Monte Carlo integration is an extremely inefficient way to determine $\pi$. These numerical methods are most useful when analytic solutions are not available, and especially for very high-dimensional integrals such as partition functions of statistical systems. State-of-the-art research in theoretical physics routinely uses Monte Carlo methods to numerically evaluate billion-dimensional integrals.

Switching back to the language of statistical ensembles, there are an enormous number of possible micro-states for any interacting systems of interest, only a vanishingly small fraction of which can be sampled in a reasonable amount of time. Even if we put in the time to sample a trillion ($10^{12}$) micro-states of the tiny $N \sim 1000$ Ising systems considered above, this would account for only about one part in $10^{288}$ of the total $2^N \sim 10^{300}$ micro-states. Even worse, as $N$ increases the number of possible Ising model micro-states grows exponentially quickly, $2^N$, and each micro-state takes more work to sample. For a concrete example, this research publication from 2015 includes a calculation with $N \sim 10^9$ that was only able to sample $\sim 10^4$ out of roughly $2^{10^9} \sim 10^{300,000,000}$ micro-states.

For sufficiently high temperatures, our consideration of the disordered phase of the Ising model in Section 10.2 suggests that randomly sampling even such a tiny fraction of the micro-states would still suffice. All of those micro-states are equally probably in the infinite-temperature limit, where we would just need to consider enough samples $S$ to produce reasonably small statistical uncertainties $\sim 1/\sqrt{S}$. However, the low-temperature ordered phase is more challenging. We saw that the large-scale behaviour of the system in this phase is dominated by a very small number of micro-states. For sufficiently low temperatures, just the two degenerate ground states effectively determine the magnetization, with only exponentially suppressed corrections from higher-energy excited states. As there is essentially no chance of randomly sampling either of those two ground states, the approach described above seems doomed to fail.

The cure is to carefully guide the procedure so that the probability of sampling a particular micro-state $\omega_i$ is proportional to its probability $p_i \propto e^{-\beta E_i}$, without introducing any bias. This is known as **importance sampling**, since it preferentially samples the important micro-states that make the most significant contributions to the partition function and derived quantities. As $\beta \to \infty$ in the low-temperature phase, the probability of sampling the ground states would be exponentially enhanced, as desired. As $\beta \to 0$ in the high-temperature phase, there would be little change compared to the more straightforward approach described above, since all micro-states would become equally probable.

A challenge facing importance sampling is that the energies $E_i$ are not known in advance. They are generally only computed for those few micro-states that are sampled. An ingenious way to get around this challenge[28] is by incorporating the concept of Markov chains that we encountered all the way back in Section 1.5. Recall that a Markov chain is a process in which the next micro-state to be sampled is chosen based on the micro-state currently under consideration, with no 'memory' of any other micro-states that may already have been sampled.

Applying this to the Ising model, we can begin with any random spin configuration. Randomly selecting one spin, $s_i$, we compute $\Delta E_i$, the change in the system's energy that would be caused by flipping $s_i \to -s_i$. We then update the spin configuration by 'accepting' this spin flip with probability

$$P_{\text{accept}} = \min\left\{1, e^{-\beta \Delta E}\right\}, \tag{126}$$

which defines the next micro-state in the Markov chain. Importantly, this new micro-state may be identical to the previous micro-state with probability $P_{\text{reject}} = 1 - P_{\text{accept}}$. This is exactly what we would want in the zero-temperature limit where the ground state should dominate. We then repeat this single-spin update procedure as many times as our computers can handle.

Digging into Eq. 126, we can see that a spin flip that lowers the energy will always be accepted, since $\Delta E < 0 \implies e^{-\beta \Delta E} > 1$. The algorithm is therefore free to approach the minimum-energy ground state of the system. If it is in the ground state, then any spin flip will increase the energy, $\Delta E > 0$, and will only be accepted with an exponentially suppressed probability $e^{-\beta \Delta E} \to 0$ as $T = 1/\beta \to 0$, as desired. More generally, if we consider two micro-states $\omega_A$ and $\omega_B$ with $E_A \leq E_B$, then the relative probabilities of moving between these two micro-states are

$$\frac{P(A \to B)}{P(B \to A)} = \frac{\min\left\{1, e^{-\beta(E_B - E_A)}\right\}}{\min\left\{1, e^{-\beta(E_A - E_B)}\right\}} = \frac{e^{-\beta(E_B - E_A)}}{1} = \frac{e^{-\beta E_B}}{e^{-\beta E_A}},$$

confirming that the micro-states $\omega_i$ will indeed be sampled with probabilities proportional to the Boltzmann factors $e^{-\beta E_i}$ that quantify their importance.

[28] This approach was developed in 1953 by Nick Metropolis, Arianna Rosenbluth, Marshall Rosenbluth, Mici Teller and Edward Teller. In an infamous misfiring of alphabetical ordering, it remains widely known as the "Metropolis algorithm" even though Metropolis's role was providing specialized computing equipment rather than creating the algorithm itself. In addition, the key contributions of Arianna Rosenbluth and Mici Teller were under-appreciated for many years.

You can find more information about this *Metropolis–Rosenbluth–Teller* algorithm in Section 8.2 of Schroeder's *Introduction to Thermal Physics* (reference 2 in the list of further reading on page 6), including a single-page annotated code applying it to the two-dimensional Ising model. While this is the most famous and most widely used Monte Carlo algorithm in existence, it is far from the only one. There is an enormous amount of ongoing research developing, optimizing and applying more elaborate Monte Carlo methods to investigate topics throughout the mathematical sciences and beyond. In Section 11.3 we will briefly look at some of these broader applications. First, there is another important concept to introduce, called universality, which helps to explain why interacting statistical systems are so useful to apply to such a diverse range of scientific investigations.

## 11.2 Universality

In Section 10.3 we defined the critical exponent $b$ as the power governing the behaviour of the order parameter $\langle m \rangle \propto (T_c - T)^b$ as the temperature $T$ approaches the critical temperature $T_c$ of the second-order phase transition. In addition to being an important feature characterizing any given phase transition, it was discovered during the twentieth century that precisely the same critical exponents turn out to govern the behaviour of phase transitions that we initially would not have expected to have any connection. For example, the critical exponent $b \approx 0.32$ for the three-dimensional Ising model magnetization mentioned last week also appears at the phase transition where an interacting (non-ideal) gas transforms into a liquid,

$$\frac{1}{\rho_{\text{gas}}} - \frac{1}{\rho_{\text{liquid}}} \propto (T_c - T)^b \qquad\qquad b \approx 0.32,$$

where $\rho$ is the density of the atoms, which are equal for the two phases at the transition.[29]

This is not simply a numerical coincidence, but an example of an amazing phenomenon is known as **universality**. In essence, universality states that the specific details of interacting statistical systems become irrelevant close to critical points at which phase transitions occur. It doesn't matter whether we are considering a three-dimensional lattice of Ising spins or a liquid such as water—the behaviour of both systems is governed by the same set of critical exponents (known as their *universality class*).

The detailed mathematics underlying universality is well beyond the scope of this module. Important contributions to its development were made by Leo Kadanoff and Ken Wilson, among many others. The purpose of this section is just to quickly introduce the concept and emphasize that universality causes the same large-scale behaviour to appear in the vicinity of critical points even for systems that initially may seem completely different. This independence of the

[29] We're not covering the liquid–gas transition in this module. If you are curious about it, you can find a discussion in Section 5.1 of David Tong's *Lectures on Statistical Physics* (reference 1 in the list of further reading on page 6).

details of the system helps to explain the power of even simple interacting statistical systems such as the Ising model, and their applicability in so many different domains.

## 11.3 Broader applications

In this section we will continue to provide quick introductions rather than detailed derivations, considering a few ways in which the concepts and tools of statistical physics are applied beyond the mathematical sciences. The cases considered here are far from exhaustive, and many more examples may become apparent now that we have gained familiarity with the underlying statistical methods and perspectives.

### Voter models

Sociology is one domain where the applicability of statistical physics may be unexpected. Despite (possible) expectations, there is a diverse and active field of "sociophysics", which uses the statistical physics concepts and tools that we have been learning about to describe various aspects of social and political behavior.[30]

A particular branch of sociology where connections to statistical physics may be more apparent is the field of opinion formation, where so-called voter models have been widely used since the 1970s, and have proved capable of capturing outcomes of elections. Voter models are interacting statistical systems not too different from the Ising model, where spin degrees of freedom that we have been working with are interpreted as voters' opinions on a certain topic. For example, support for a proposal or candidate can be represented as $s_i = +1$, while opposition would be $s_i = -1$. Just like the interactions in the Ising model encourage spins to align with each other, voter models build in a tendency for voters to align (i.e., agree) with the majority of other voters they interact with.

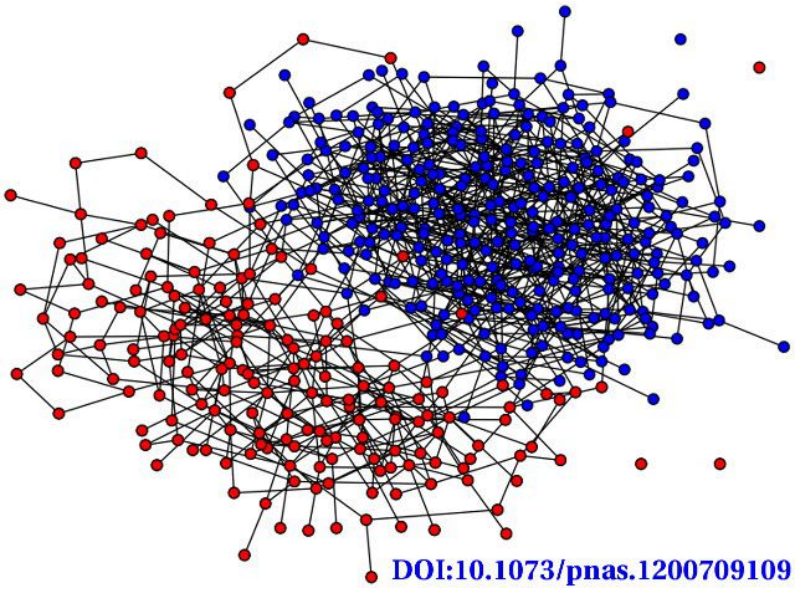
There are many generalizations that can then be incorporated to better describe the outcome of polls and elections. A simple extension would be to allow voters to be neutral or indifferent to the question, represented as $s_i = 0$. Similarly, the strength of a voter's commitment to their opinion can be modelled by extending the range of possible spin values,

$$s_i \in \{\cdots, -2, -1, 0, +1, +2, \cdots\}\,.$$

Another way to make voter models more realistic is to define them on more flexible *graphs* rather than regular lattices. The figure below (source) illustrates how such a graph can look for a two-state system with $s_i \in \{-1, +1\}$ coloured red and blue. As in everyday experience (and inspired by social networks), different voters can have connections with different numbers of other voters, who do not

[30] For a review, see Serge Galam, *Sociophysics: A Physicist's Modeling of Psycho-political Phenomena* (2012).

need to be nearest neighbours. In this particular investigation, network connections between disagreeing voters are probabilistically severed, and transitions are observed between a phase in which one of the opinions becomes the consensus, or all connections are severed between the two groups with differing opinions.



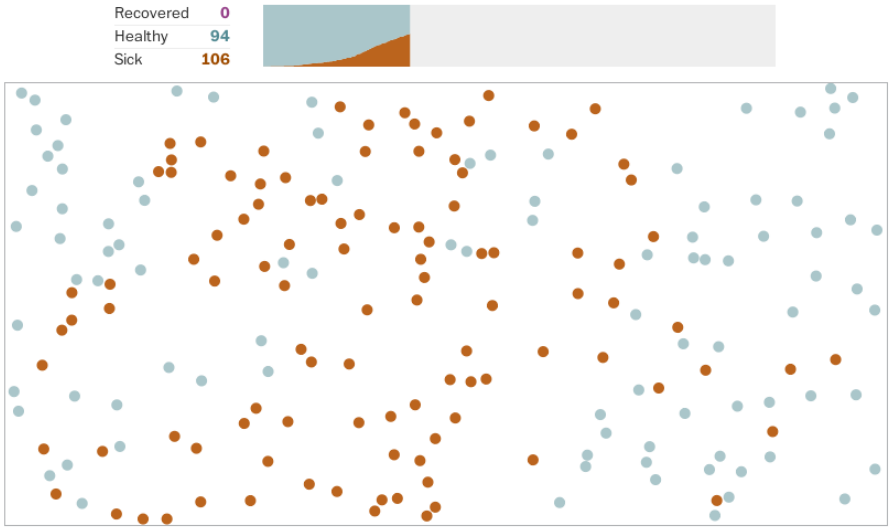

Simpler studies that don't involve severing connections have observed a similar consensus phase occurring once an opinion reaches a critical concentration. Remarkably, this was found to be described by a second-order phase transition in the universality class of the two-dimensional Ising model. Another outcome possible in certain voter models is a "stable non-consensus" phase, in which the two opinions persist indefinitely, each one held by a 'cluster' of aligned voters who interact mainly with each other rather than with voters holding the opposite opinion. These investigations are described by Alexander Balankin et al., "Ising percolation in a three-state majority vote model" (2016).

## Epidemiology

A particularly topical application of interacting statistical systems is to model the spread of diseases in populations, something most of us may have seen in the news over the past fifteen months. As in the case of voter models, the degrees of freedom under consideration are again individuals, whose interactions with each other allow infection to spread to those who are susceptible. A Monte Carlo calculation of the sort described in Section 11.1 can then be used to model how many people may be infected as time passes, guided by data on typical movement and contacts.

The figure below comes from an over-simplified simulation provided by the *Washington Post* to illustrate these concepts. Here individuals are modeled simply as a gas of interacting particles in two dimensions, and various ways of restricting their motion are used to explore the likely effects of measures such as

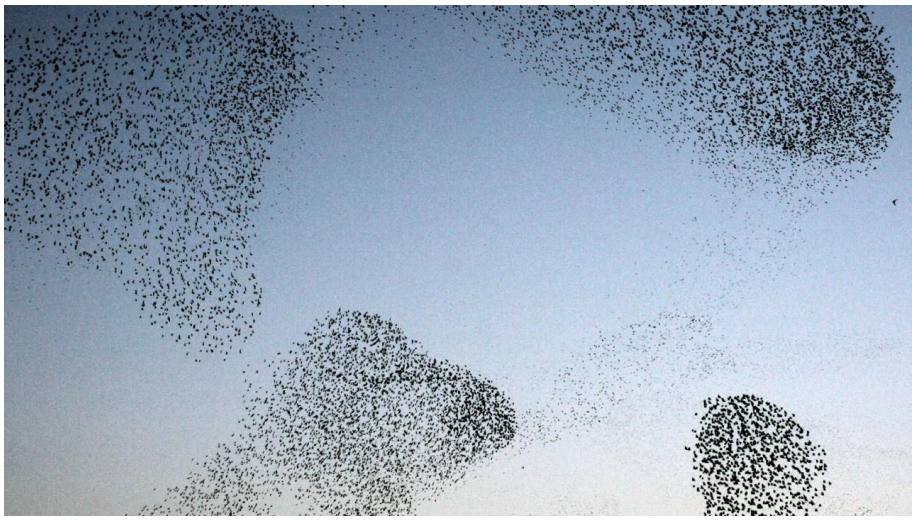
quarantines and social distancing. By coincidence, the scale of this simulation, which considers a population of just $200$ people, is comparable to the first importance sampling Monte Carlo calculation carried out in 1953 as a first application of the Metropolis–Rosenbluth–Teller algorithm. That historic investigation computed the pressure (i.e., equation of state) for $224$ interacting particles in a two-dimensional 'volume'. At the time, this required several days of computing time on a state-of-the-art machine; now these sorts of calculations are easily done on a smartphone.




Larger-scale and more realistic versions of these epidemiological simulations provide important input into government deliberations regarding what restrictions (such as lockdowns) would be most beneficial to reduce the spread of disease, and how long such restrictions may need to be maintained. Rather than investigating a phase transition, the goal is to quantify likely effects of restrictions (or of the absence of restrictions). As described in Section 1.2, the numerical experiments are therefore repeated many times with different sequences of pseudo-random numbers. This produces an ensemble of possibilities from which the likely outcomes of various interventions can be inferred.

### Flocking

Let's conclude our quick glimpses of some broader applications of statistical physics by considering another biological example, where interacting statistical systems are used to model the large-scale collective motion of certain groups of animals. The image below (from Marcus Woo, "How bird flocks are like liquid helium", *Science*, 27 July 2014) illustrates the so-called flocking behaviour of large groups of starlings, which fly through the sky in surprisingly tight coordination. This same sort of behaviour is also seen in schools of fish, swarms of insects, and even crowds of humans.



Many models based on interacting statistical systems have been (and continue to be) developed to describe this emergent collective behaviour. A particularly simple and widely studied example was introduced in 1995. It proposes that each particle (i.e., bird) will interact with others near to it, and this interaction will encourage it to move in the same direction as its neighbours, in much the same way that the nearest-neighbour interactions in the Ising model encourage its spins to align.

This model exhibits a transition between two distinct phases. When there is a low density of particles, there are relatively few interactions and the particles' motion is disordered, with no formation of flocks or swarms. At high densities, in contrast, large-scale collective motion appears, based solely on the interactions between the individual particles in the system. These two regimes turn out to be separated by a first-order phase transition, with a discontinuity (in the thermodynamic limit) in an order parameter related to the average angle of flight. In addition, as the critical density is approached, flights in the disordered phase exhibit anomalous super-diffusion like that investigated in the computer project.

## 11.4 Wrap-up recap and synthesis

We have covered a lot of ground in this module, building on foundations from probability theory to develop and apply the core concepts of statistical physics. In week 1 we defined probability spaces, expectation values and variances, and used these to establish the law of large numbers through which stable large-scale behaviour occurs for stochastic systems involving many degrees of freedom $N \gg 1$. We also saw how the central limit theorem relates large-$N$ probability distributions to the underlying mean and variance of the elementary degree of freedom, and practiced extracting probabilities from such distributions. The law of diffusion results from the central limit theorem, and the computer project applied inverse transform sampling to study how anomalous diffusion can arise when the central limit theorem's assumption of finite elementary mean and variance breaks down.

Starting in week 2 we specialized to particular probability spaces known as statistical ensembles, which consist of the micro-states and associated probabilities that a system can possibly adopt through its time evolution. The laws of physics, such as conservation of energy (the first law of thermodynamics), impose constraints on statistical ensembles. The micro-canonical ensemble directly implements such constraints, by requiring the system's internal energy and particle number be constant, which implies that the system is completely isolated from the rest of the universe. This makes both the entropy (Eq. 21) and temperature (Eq. 23) derived quantities, which we explored via non-interacting spin systems. In particular, we derived a form of the second law of thermodynamics, which states that the total entropy never decreases as time passes and therefore indicates that maximal entropy corresponds to thermodynamic equilibrium.

Motivated by the impracticality of demanding that a system be completely isolated, in week 3 we turned to the canonical ensemble, which allows systems to exchange energy with a large thermal reservoir that imposes a constant temperature. The particle number is still fixed, while the entropy, internal energy and heat capacity are now important derived quantities, with the last two related by a fluctuation–dissipation relation. Maximizing the entropy to consider systems in thermodynamic equilibrium defines the partition function (Eq. 34) and Boltzmann distribution (Eq. 35). Derived quantities are determined from the partition function, or equivalently the Helmholtz free energy (Eq. 39). By analyzing both distinguishable and indistinguishable non-interacting spins, we demonstrated that the intrinsic information content of statistical systems has physically measurable effects.

In week 4 we developed another application of the canonical ensemble by analyzing non-relativistic, classical, ideal (non-interacting) gases. At the start of this analysis we had to take care to ensure that the partition function was well-defined, by assuming that only discrete momenta are possible in a finite volume $V$, to ensure that the number of micro-states is countable. Again determining the internal energy (Eq. 54) and entropy (Eq. 55) from the resulting partition functions for both distinguishable and indistinguishable gas particles, we computed the mixing entropy that is created by combining gases of distinguishable particles. Finally we defined the pressure as the adiabatic change in the system's energy upon changing its volume (Eq. 59), and derived the ideal gas law (Eq. 60) as a famous equation of state.

Building on our analyses of ideal gases in the canonical ensemble, in week 5 we considered thermodynamic cycles as systems that perform a repeatable sequence of expansions, compressions and heat transfers in order to act as heat engines or refrigerators. This required first considering the work done on the system through these processes (Eq. 63), the heat added to or removed from it (Eq. 65), and the first law of thermodynamics expressed in terms of these quantities (Eq. 67). After introducing $PV$ diagrams as a convenient way to visualize thermodynamic cycles and the individual processes that comprise them, we analyzed the Carnot heat engine and computed how much work it can do by transferring heat from a hot thermal reservoir to a cold reservoir. This combination of work and heat defines the efficiency (Eq. 68), and we computed that the Carnot

heat engine achieves the maximum possible efficiency allowed by the second law of thermodynamics.

The logical next step, after allowing the energy to fluctuate with a constant temperature, is also allowing the system to exchange particles with a large particle reservoir. This leads to the grand-canonical ensemble that we developed in week 6, beginning by defining the chemical potential (Eq. 71) as the new quantity (in addition to the temperature) that the particle reservoir keeps constant. Carrying out another round of entropy maximization defined the grand-canonical partition function (Eq. 80) and the corresponding grand-canonical potential (Eq. 81) that determine the derived quantities that now include the particle number in addition to the energy and entropy. We were also able to derive a generalized thermodynamic identity (Eq. 86) that relates the chemical potential to the change in energy upon adiabatically adding a particle to the system.

Our main applications of the grand-canonical ensemble were to analyze gases based on the quantum statistics that we introduced (as an ansatz) in week 7. After demonstrating how the classical (non-quantum) approach we took when first considering ideal gases breaks down when there is a non-negligible probability for multiple identical particles to occupy the same energy level, we got around this problem by organizing the micro-states in terms of the possible occupation numbers of the energy levels. These possible occupation numbers distinguish between the two types of particles that appear in nature: bosons that can have any non-negative occupation numbers $n_\ell \in \mathbb{N}_0$, and fermions that obey the Pauli exclusion principle and can have only $n_\ell \in \{0, 1\}$. We derived the respective Bose–Einstein and Fermi–Dirac statistics for these two types of quantum particles, and checked that both approach classical (Maxwell–Boltzmann) statistics in the limit of high temperature with large negative chemical potential $-\mu \gg T$.

We carried out a more thorough application of the grand-canonical ensemble to quantum gases of bosons in week 8, focusing on ideal (non-interacting) gases of ultra-relativistic photons with energies defined in terms of frequencies (Eq. 99). Based on the grand-canonical potential, we derived the Planck spectrum (Eq. 101) governing the frequency dependence of the energy density for photon gases, and found how it solves the ultraviolet catastrophe of the classical Rayleigh–Jeans spectrum. We also saw how the Planck spectrum provides an excellent mathematical model for both stars (some of the hottest places in the universe) as well as the cosmic microwave background that fills frigid inter-galactic space and provides strong evidence for the existence of dark matter. We finally derived the radiation pressure of photon gases, and the corresponding equation of state, which has the same form as the ideal gas law, just with a different numerical factor.

In week 9 we carried out a similar application of the grand-canonical ensemble to quantum gases of non-interacting fermions, focusing mainly on non-relativistic particles and considering the low-temperature regime where quantum Fermi–Dirac statistics differs the most from the classical case. Again based on the grand-canonical potential, we derived the Fermi function (Eq. 107) and saw how it becomes a step function at low temperatures. This corresponds to all the

low-lying energy levels being filled by the single fermion that can occupy each of them by the Pauli exclusion principle, up to the Fermi energy that (at zero temperature) is simply the chemical potential (Eq. 109). The resulting internal energy (Eq. 111) and pressure (Eq. 113) remain positive even as the temperature approaches absolute zero. This non-zero degeneracy pressure helps to explain the regularity of type-Ia supernovas, which play a key role in establishing the existence of so-called dark energy.

The last major topic of the module was to explore effects of interactions in weeks 10 and 11. Statistical systems in which the constituent degrees of freedom can interact with each other exhibit a broader array of phenomena, such as phase transitions. At the same time, they become enormously more difficult to analyze, because their partition functions and derived quantities no longer factorize into independent single-particle contributions. We focused on the famous Ising model, which is simple to write down as a system of spins interacting with their nearest neighbours on a $d$-dimensional simple cubic lattice, but extremely difficult to solve exactly in two or more dimensions. We discussed how to analyze the Ising model through the mean-field approximation (last week) and by using numerical Monte Carlo importance sampling algorithms (Section 11.1). For $d \geq 2$, the Ising model exhibits a second-order phase transition between its high-temperature disordered phase and its low-temperature ordered phase. The critical exponents characterizing these transitions also arise in many other contexts, a phenomenon known as universality, which helps to explain the broad applicability and far-reaching power of statistical physics, including the examples briefly considered in Section 11.3.

In summary, we have learned foundations including statistical ensembles, entropy, and the laws of thermodynamics. We have studied applications including diffusion, ideal gases, thermodynamic cycles, and phase transitions. And we have previewed advanced topics including numerical methods and universality. All together, our new knowledge of statistical physics enables us to observe and appreciate many further applications of these concepts and tools across the mathematical sciences and beyond.