# Theory of Complex Systems

Tutorials

Lecturer: Clélia de Mulatier

May 27, 2024

This document contains the collective contributions of:

Wout Merbis, Ebo Peerbooms, Janusz Meylahn
Niki Stratikopoulou, Gal Ross
Leander Post, Maxim Zewe

Thank you very much for your help in making the course and exercises better!

# Contents

**Symbol "(⋆)":** Questions and exercises indicated with a (⋆) are optional. No worries if you don't have time to try to solve them, or if you don't manage to solve them on your own.

**The symbol "</>":** indicates optional questions with numerical simulation.

# Chapter 0

# Markov Processes and Simple Models of Complex Systems

**Goal of the tutorial** is for students to get familiar with Poisson processes, and more generally, with Markovian processes. After this tutorial, students should:

- know what a homogeneous Poisson process is, know of examples and where they can be used.
- know that the distribution of number events of a Poisson process in a given time interval is the Poisson distribution.
- know that the distribution of waiting time in a Poisson process is the exponential distribution, and how to recover such distribution.
- know how to generate events following a Poisson process with rate $\lambda$.
- know how to check if a distribution is normalized.
- know how to compute averages for discrete probability distributions and continuous probability distributions.
- write down the equation of evolution for quantities following a continuous time Markov process.
- understand what *stationary* means for a stochastic process.

## 0.1 Homogeneous Poisson process

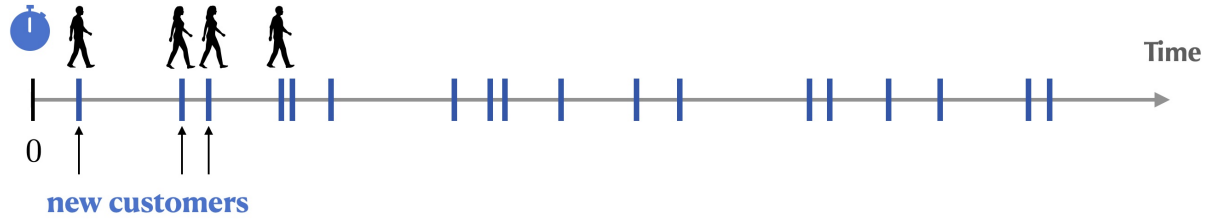
The **Poisson process** is one of the most widely-used counting processes. It is a model for a series of discrete events where the average time between events is known, but the exact time at which events occur is random. In particular, the occurrence of a new event is independent of the previous events: we say that the process is memoryless, or *Markovian*.

**(Homogeneous) Poisson process:** Collection of **random events** that:

- are **independent from each other**;
- occur with a **constant rate**.

For example, consider a supermarket in which a new customer arrives on average every 2 minutes. The arrival time of a new customer is random, independent of the previous customers, and will not influence the arrival time of the next customers. The Poisson process is a good model for this problem. In practice, Poisson processes have been used to model many other phenomena, such as timing of calls at a help center, of visitors to a website, of patients arriving in emergency room, of meteors hitting Earth, radioactive decay in atoms, occurrences of earthquakes, spiking time of neurons, or movements in a stock price.

Poisson processes are generally associated with time, but they can also be associated with space. For instance, consider trees randomly placed in a forest and the problem of counting the number of trees in an acre (events per area).




### 0.1.1 Simple considerations

Let us take the example of modeling the times at which customers arrive at a supermarket. At time $t = 0$, we start counting the number of customers that arrive. We denote by $N(t)$ the total number of customers that have entered the shop until a time $t$ , with $N(0) = 0$.

**Q1.** We know that on average there is one customer entering the shop every 2 minutes. What is the rate $\lambda$ at which customers enter the shop? What is the average number of customers $\overline{N(t)}$ that one can expect to see enter the shop within $t = 10$ minutes?

### 0.1.2 Definitions of a Poisson process and Poisson distribution.

**Definition 1.** The counting process $N(t)$ is a **Poisson process** with rate $\lambda > 0$ , if:

a. $N(0) = 0$;

b. $N(t)$ has independent increments;

c. the number of arrivals in any interval of length $t > 0$ follows a **Poisson distribution** with parameter $\mu = \lambda\, t$, i.e. the probability to observe $k$ arrivals during the time $t$ is given by:

$$\mathbb{P}[N(t) = k] = \frac{\mu^k \exp(-\mu)}{k!}\,. \tag{1}$$

**Q2.** Can you check that the probability distribution defined in Eq. (1) is well normalized, and show that the mean of the distribution is $\mu$? Compare with the expression for the average number of customers $\overline{N(t)}$ that you used in question Q1.

**Q3.** For the supermarket problem described above, assuming that the number of customers arriving at the shop follows a Poisson process, can you compute the probability that only 2 customers have arrived within a 10 minutes time window? What about 5 customers? 15 customers? Can you plot the distribution of $N(t)$ for $t = 10$ minutes, for $N(t) = 0$ to $N(t) = 20$ (included)?

**Q4.** Consider a very short time interval $dt \ll 1$. Using a Taylor expansion for small $dt$, can you show that the probability that no customer arrives during $dt$ is close to 1 and takes the form:

$$\mathbb{P}[N(dt) = 0] = 1 - \lambda\, dt + o(dt), \tag{2}$$

where $o(dt)$ indicate terms that are negligable (very small) compared to $dt$. Similarly, can you show that the probability that the probability that 1 customer (respectively or 2 or more customers) arrives during dt takes the form:

$$\mathbb{P}[N(dt) = 1] = \lambda\, dt + o(dt) \tag{3}$$

$$\mathbb{P}[N(dt) \geq 2] = o(dt) \tag{4}$$

**Definition 2.** The counting process $N(t)$ is a **Poisson process** with rate $\lambda > 0$, if:

a. $N(0) = 0$;

b. $N(t)$ has independent and stationary increments;

c. during a very short time interval $dt$:
   - the probability that 1 customer enter the shop is: $\lambda\, dt + o(dt)$
   - the probability that 2 or more customers enter the shop is: $o(dt)$
   - and, therefore, the probability that no customer enter the shop is: $1 - \lambda dt + o(dt)$

   which can be written formally as:

$$\mathbb{P}[N(dt) = 0] = 1 - \lambda\, dt + o(dt) \tag{5}$$
$$\mathbb{P}[N(dt) = 1] = \lambda\, dt + o(dt) \tag{6}$$
$$\mathbb{P}[N(dt) \geq 2] = o(dt) \tag{7}$$

**Q5.** How can you simulate events that follow such rules? Write a small program that simulate the supermarket problem described above, by generating random times (in minutes) at which customers arrive at the supermarket. Simulate data for three days at the supermarket: a day starts at 8 am and finishes at 10 pm. What is the order of magnitude of the number of time steps that the simulation will do if you take $dt = 0.1$ for instance? Discuss the choice of $dt$.

### 0.1.3 Waiting time (or interarrival times)

**The general question of this section is:** How long do we have to wait until the next event occurs? (i.e. until the next customer arrives)

Let's assume that $N(t)$ follows a Poisson process. We would like to study the distribution of waiting times $\tau$ between two customers. Let us denote $P(\tau)$ this distribution.

We recall that by definition of a probability distribution $P(\tau)$, one has the following:
Assume that a customer just entered, the probability that the next customer enters within the time interval $[\tau;\, \tau + d\tau)$ later is given by $P(\tau)\, d\tau$, where $d\tau$ is a very short time interval ("infinitesimal time interval") chosen such that at maximum only 1 customer arrives during $d\tau$.

**Q6.** Let's introduce the **cumulative function** $U(T)$ as the probability that the next customer arrives at any time $\tau$ between 0 and $T$ . Can you express $U(T)$ as a function of $P(\tau)$? What is the value of $U(T)$ for $T = 0$? What is the limit value of $U(T)$ for infinitely large time $T$?

**Q7.** Consider an infinitesimal time interval $dT$. The quantity $U(T+dT)$ is then the probability that the next customer arrives before the time $T + dT$. Assuming that $N(t)$ is a Poisson process, can you express $U(T + dT)$ as a function of $U(T)$, $dT$ and $\lambda$? Deduce that $U(T)$ verifies the following equation of evolution:

$$\frac{d\,U(T)}{dT} = \lambda\,(1 - U(T))\,. \tag{8}$$

**Q8.** Solve this equation to find $U(T)$. Which probability distribution do you obtain for $P(\tau)$? Check that you recover an average waiting time of $1/\lambda$, as original defined in the exercise.

**Poisson process: Distribution of waiting times.**

In this question, you have recovered that the **distribution of the waiting times** $\tau$ between two consecutive events of a Poisson process with rate $\lambda$ is the **exponential distribution**:

$$P(\tau) = \lambda\, \exp(-\lambda\, \tau)\,. \tag{9}$$

**Q9.** In the data generated in question Q2, can you check that the distribution of time intervals between two successive customers corresponds to the distribution $P(\tau)$ that you just found? You can also test this on the data provided with this tutorial. What is the average waiting time in the data? Compare with $1/\lambda$.

**Q10.** If we arrive at a random time in the shop, how long can we expect to wait to see the next customer arriving? What is the probability that the next customers arrives in less than 1 minute? What is the probability that the next customer arrives in more than 5 minutes?

**Q11. Sampling from an exponential distribution.** Using the expression of the cumulative distribution $U(T)$ found in question Q5, how can you directly sample the next time at which a customer will arrive? Can you use this result to generate random data for the supermarket problem more efficiently?

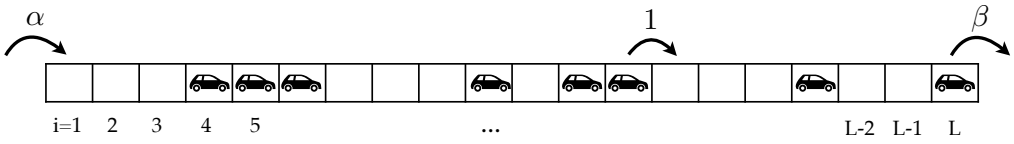
### 0.1.4 Summary

Can you summarize what you have seen in this section?

## 0.2 TASEP: a simple Traffic Model

We consider a one-dimensional lattice with $L$ sites opened on both sides (which represents a portion of road). Cars randomly hop into the lattice from the left with rate $\alpha$, jump along the lattice to the right with rate 1, and exit the lattice from the right with rate $\beta$ (see figure). Each site can only be occupied by one car at a time, and a car can hop to the next site only if it is empty (exclusion interaction). This model for instance the fact that, on a one-lane road, a car moves forward provided that there is no vehicle in front of it.

This simple traffic model is called Totally Asymmetric Simple Exclusion Process (TASEP): "exclusion" because there can only be one car per site, and "Totally Asymmetric" because cars can only move towards the right. It was originally introduced to model the translation of mRNA in protein synthesis, and is useful in modeling a wide range of transport phenomena, including road traffic. The model is known for exhibiting phase transitions even in one dimension, between low-density, high-density and maximum-current phases.




**Q1.** Consider a very short time interval $dt$, what are the possible events that can happen during $dt$ and what are the probability that each of these events happen?

**Q2.** How can you simulate the dynamics of this model? Give the main lines of your algorithm.

**Q3.** We denote by $n_i(t)$ the number of cars at the site $i$ at time $t$, which can only take two values: 0 if the site $i$ is empty, and 1 if the site $i$ is occupied. We denote by $\rho_i(t)$ the probability that there is a car at the site $i$ at time $t$: $\rho_i(t) = P[n_i(t) = 1]$. The probability that there is no car at site $i$ is then: $1 - \rho_i$ (as a site can either be occupied or not).
Can you show that the average number of cars in $i$ at time $t$ is $\overline{n_i(t)} = \rho_i(t)$?

**Q4.** We define the current $J_i(t)$ of cars that exit site $i$ to the right at time $t$ by:

$$J_i(t) = P[n_i(t) = 1; n_{i+1}(t) = 0]\,, \tag{10}$$

which corresponds to the joined probability that site $i$ is occupied and that site $i+1$ is empty at time $t$. What is the probability that a car moves from site $i$ to site $i+1$ during the very short time interval $[t, t+dt)$?

**Q5.** Show that the local density of cars follows the equation of evolution:

$$\frac{d\rho_i}{dt}(t) = J_{i-1}(t) - J_i(t)\,, \tag{11}$$

for all the sites $i$ between $i = 2$ and $i = L - 1$.

**Q6.** What are the equations of evolution for $\rho_1$ and $\rho_L$? Introduce a current $J_0(t)$ that enters the site 1 from the left, and a current $J_L(t)$ that leaves site $L$ to the right.

**Q7.** In the stationary state, the local densities and currents become time-independent. Deduce from the previous equations that the current is uniform in the stationary state, i.e. for all $i$, $J_i = J$ is a constant. Show that we have the boundary equations:

$$\rho_1 = 1 - \frac{J}{\alpha} \qquad \text{and} \qquad \rho_L = \frac{J}{\beta}\,. \tag{12}$$

**Q8.** Play with the simulation provided in the mathematica notebook. The graph in the bottom left display the evolution of the current that goes through the last site $i = L$ as a function of time. Observe that the average current initially increases to then reach a stationary value (the value of $J$ fluctuates around a fixed value). Play with the parameters $\alpha$ and $\beta$ and observe the different stationary states of the systems. Can you comment on the behavior of the system for $(\alpha,\ \beta) = (1.0, 0.1)$, $(0.1, 1.0)$, $(0.2, 0.2)$ and $(0.6, 0.6)$?

**For Thursday:** Using the simulation provided in the mathematica notebook, fill in 4 or 5 values of $J$ in the table here. You can accelerate the simulation by selecting '20' for the simulation speed (that will change the refreshing rate of the graphs). Fixing the values of the parameters, the system will slowly evolve towards stationarity: the value of $J$ will fluctuate around a fixed value. Report an estimate of this value in the table of the google doc.

**Bonus question:** If you have some spare time, feel free to try to implement yourself a numerical simulation of the TASEP model!

## 0.1 Solutions

### 0.1.1 Homogeneous Poisson process

**Useful definitions**

- **Discrete probability distribution:** The Poisson distribution in Eq. (1) is a discrete probability distribution, for which $P(k)$ is the probability that $k$ event happens.
- **Continuous probability distribution**, also called **probability density function (PDF):** The Exponential distribution $P(\tau)$ in Eq. (9) is a probability density function (i.e. $P(\tau)$ is not a probability per se, but a probability density): $P(\tau)\, d\tau$ is the probability that the next event happens within the small time interval $[\tau, \tau + d\tau]$.
- **Normalization and Ensemble Average:**

| | Discrete distribution | Continuous distribution |
|---|---|---|
| Normalization | $\sum_k P(k) = 1$ | $\int \mathrm{d}x\, P(x) = 1$ |
| Average | $\langle A \rangle = \sum_k A(k)\, P(k)$ | $\langle A \rangle = \int A(x)\, P(x)\, dx$ |

  where the summation (resp. the integration) is over the entire domain where the discrete distribution $P(k)$ (resp. the continuous distribution $P(x)$) is defined.
- **Cumulative Distribution Function (CDF):**

$$U(T) = \int_0^T P(\tau) \mathrm{d}\,\tau.$$

- Series expansion of the exponential function:

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \dots$$

  Note: you don't need to remember this expansion for the exam. If it is needed for the exam, the expansion will be written in the first page of the exam sheet.

**Notations.** In the literature (and in this course), you can find the (ensemble) average denoted with different notations, the most common being:

$$\langle A \rangle \qquad \text{or} \qquad \overline{A} \qquad \text{or} \qquad E[A] \tag{13}$$

**A1.**

$$\lambda = \frac{1}{2\ [\mathrm{min}]} = \frac{1}{2 \times 60\ [\mathrm{s}]} = 8.3 \cdot 10^{-3}\ [\mathrm{s}^{-1}] = 0.5\ [\mathrm{min}^{-1}] \tag{14}$$

$$\overline{N(t = 10\,\mathrm{min})} = \lambda \times t = 5 \text{ customers} \tag{15}$$

**A2.** To check that the probability distribution is normalized, we must check that $\sum_{k\geq 0} \mathbb{P}[k] = 1$.

$$\sum_{k\geq 0} \mathbb{P}[k] = \sum_{k=0}^{\infty} \frac{\mu^k\ \exp(-\mu)}{k!} = \exp(-\mu) \sum_{k=0}^{\infty} \frac{\mu^k}{k!} = \exp(-\mu)\ \exp(\mu) = 1\,, \tag{16}$$

where we recognized the power series of the exponential function: $\exp(\mu) = \sum_{k=0}^{\infty} \frac{\mu^k}{k!}$.

The mean $\overline{N(t)}$ of the distribution at a given time $t$ is defined as $\overline{N(t)} = \sum_{k\geq 0} k\,\mathbb{P}[k]$, in which $k = N(t)$.

$$\overline{N(t)} = \sum_{k\geq 0} k\,\mathbb{P}[k], \qquad \text{in this sum, the term for k=0 is null} \tag{17}$$

$$= \exp(-\mu) \sum_{k=1}^{\infty} k\,\frac{\mu^k}{k!} = \exp(-\mu) \sum_{k=1}^{\infty} \frac{\mu^{k-1}\,\mu}{(k-1)!}, \tag{18}$$

$$= \mu\,\exp(-\mu) \sum_{k'=0}^{\infty} \frac{\mu^{k'}}{k'!}, \qquad \text{where we took } k' = k-1 \tag{19}$$

$$= \mu\,\exp(-\mu)\,\exp(\mu) = \mu. \tag{20}$$

**A3.** We want to compute the probability that $N(t = 10\,\text{min}) = 2$. Assuming that $N(t)$ follows a Poisson process, we have that the probability that $N(t) = k$ is:

$$\mathbb{P}[N(t) = k] = \frac{\mu^k\,\exp(-\mu)}{k!}, \tag{21}$$

with $\mu = \lambda\,t$, where $\lambda = 1/2\,[\text{min}^{-1}]$ and $t = 10\,\text{min}$, i.e. $\mu = 10/2 = 5$. Therefore:

$$\mathbb{P}[N(t) = 2] = \frac{5^2\,\exp(-5)}{2!} \simeq 0.08$$

$$\mathbb{P}[N(t) = 5] = \frac{5^5\,\exp(-5)}{5!} \simeq 0.18$$

$$\mathbb{P}[N(t) = 15] = \frac{5^{15}\,\exp(-5)}{15!} \simeq 1.6 \cdot 10^{-4}$$

For other numerical values, see Fig. 1 Left a few pages below.

**A4.** The probability that no customer arrives during $dt$ is given by the Poisson distribution at $k = 0$ with $\mu = \lambda\,dt$:

$$\mathbb{P}[N(dt) = 0] = \exp(-\lambda\,dt) \tag{22}$$

$$= 1 - \lambda\,dt + \frac{(\lambda\,dt)^2}{2} - \cdots \qquad \text{Taylor series} \tag{23}$$

If $dt$ is small, the terms that include second or higher powers of $dt$ are negligible compared to $dt$ (i.e. when $dt$ goes to 0, these terms are going much faster to 0 than $dt$ itself). We write this as

$$\mathbb{P}[N(dt) = 0] = 1 - \lambda\,dt + o(dt)\,. \tag{24}$$

> **Little-"o" notation:**
>
> If $dt$ is small, the terms that include second or higher powers of $dt$ are negligible compared to $dt$ (i.e. when $dt$ goes to 0, these terms are going much faster to 0 than $dt$ itself). In the equations, we denote all terms that are negligible compared to $dt$ (i.e. the terms that are of order $dt^2$, $dt^3$, or higher) by $o(dt)$. For instance, Eq. (23) becomes:
>
> $$\mathbb{P}[N(dt) = 0] = 1 - \lambda\,dt + o(dt)\,. \tag{25}$$

The probability that one customer arrives during $dt$ is given by:

$$\mathbb{P}[N(dt) = 1] = \lambda\,dt\,\exp(-\lambda\,dt) \tag{26}$$

$$= \lambda\,dt - (\lambda\,dt)^2 + \frac{(\lambda\,dt)^3}{2} - \cdots \qquad \text{Taylor series} \tag{27}$$

$$= \lambda\,dt + o(dt)\,. \tag{28}$$

The probability that two or more customers arrives during $dt$ can be obtained using the normalisation:

$$\mathbb{P}[N(dt) = 0] + \mathbb{P}[N(dt) = 1] + \mathbb{P}[N(dt) \geq 2] = 1\,, \tag{29}$$

which leads to:

$$\mathbb{P}[N(dt) \geq 2] = 1 - \mathbb{P}[N(dt) = 0] + \mathbb{P}[N(dt) = 1]\,, \tag{30}$$

$$= o(dt)\,. \tag{31}$$

**A5.** Take a small time interval $dt$ such that the probability that 1 event occurs during $dt$ is reasonable, but the probability that 2 events occur during $dt$ is very small. For instance, if we consider a minute as a unit time, choosing $dt = 0.1$ minute gives that the probability that 1 event occurs during $dt$ is $\lambda\, dt = 0.05$, the probability that two events occur during $dt$ is $(\lambda\, dt)^2 = 0.0025$ (note that the probability that two events or more occur during $dt$ is $(\lambda\, dt)^2 + o(dt^2)$, which is of the order of $(\lambda\, dt)^2 = 0.0025$). Note that we don't want to take $dt$ too small, as otherwise we would need to take too many time steps. With $dt = 0.1$ minute, one needs 10 steps to simulate a minute, 600 steps for an hour, $14 * 600 = 8400$ steps for a whole day, and finally 25200 steps for three days.

**A6.** By definition:

$$U(T) = \int_0^T P(\tau) d\tau \qquad \text{and} \qquad \lim_{T\to\infty} U(T) = 1\,. \tag{32}$$

The first equation comes from the definition of $U(T)$, the second from the normalization of $P(\tau)$ (as $P(\tau)$ is defined for $\tau \in [0; +\infty)$).

**A7.**

$$\underbrace{U(T + dT)}_{(1)} = \underbrace{U(T)}_{(2)} + \underbrace{(1 - U(T))}_{(3)} \times \underbrace{\lambda\, dT}_{(4)}\,. \tag{33}$$

where:

(1) is the probability that the next customer enters between time 0 and time $T + dT$;

(2) is the probability that the next customer enters between time 0 and time $T$;

(3) is the probability that the next customer has not yet entered between time 0 and time $T$ (which is also equal to the probability that the next customer arrives at a time later than $T$);

(4) is the probability that a customer enters during the short time $dT$. Note that, as the process is memoryless process, this probability is completely independent of what could have happened before.

This equation can be re-written as:

$$\frac{U(T + dT) - U(T)}{dT} = \lambda\,(1 - U(T))\,. \tag{34}$$

which finally leads to the equation of evolution in the text, in which we use the first order approximation of the derivative:

$$U'(T) = \frac{dU(T)}{dT} = \frac{U(T + dT) - U(T)}{dT}\,. \tag{35}$$

**Approximation for the derivative of a function:**

We will often use the following approximation for the derivative of a function $f(x)$:

$$f'(x) = \frac{f(x + dx) - f(x)}{dx}\,, \qquad \text{for very small } dx \tag{36}$$

Note: this is equivalent to cutting the Taylor expansion of $f(x+dx)$ to first order:

$$f(x+dx) = f(x) + dx\, f'(x) + o(dx^2)\,. \tag{37}$$

Note for computer scientists: this is similar to using the finite difference as an approximation of the derivative:

$$f'(T) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \tag{38}$$

**A8.** We can re-write equation (8) as

$$\frac{d\,V(T)}{dT} = -\lambda\, V(T) \qquad \text{where } V(T) = 1 - U(T) = \int_T^{+\infty} P(\tau)\, d\tau\,. \tag{39}$$

The solution of this equation is $V(T) = A\, \exp(-\lambda\, T)$, where $A$ is a constant. Using that $V(0) = 1$ (because $P(\tau)$ is normalized), we obtained that $A = 1$. Hence $U(T) = 1 - \exp(-\lambda\, T)$. Finally, $P(\tau)$ can be obtained by deriving $U(T)$:

$$P(T) = U'(T) = \lambda \exp(-\lambda T)\,. \tag{40}$$

We just proved that the time intervals between successive events of a Poisson process follow an Exponential distribution (see Fig. 1 Right in the next page).

The average waiting time is calculated using:

$$\begin{aligned}
\langle\tau\rangle &= \int_0^{+\infty} \tau\, P(\tau)\, d\tau = \int_0^{+\infty} \tau\, \lambda\, \exp(-\lambda\tau)\, d\tau \\
&= [-\tau\, \exp(-\lambda\tau)]_0^{+\infty} + \int_0^{+\infty} \exp(-\lambda\tau)\, d\tau \qquad \text{(integration by parts)} \\
&= 0 + [-\frac{1}{\lambda}\, \exp(-\lambda\tau)]_0^{+\infty} \\
\langle\tau\rangle &= \frac{1}{\lambda}
\end{aligned}$$

**A9.** Note that, if we were to sample many dataset, for each of them, we will find a different value of the average waiting time. However, if one were to run a large number of experiments, one would find that the average waiting time (computed for each experiment) follows a Gaussian distribution (normal distribution) centered around the average value $1/\lambda$, and with a standard deviation that is proportional to $1/\sqrt{N}$, where $N$ is the number of experiments performed.

**A10.** We can expect to wait an average time of $1/\lambda = 2$ minutes. The probability that the next customer arrives in less than 1 minute is given by $U(T = 1\,\text{min}) = 1 - \exp(-\lambda T) = 1 - \exp(-1[\text{min}]/2[\text{min}]) = 0.39$. The probability that the next customer arrives in more than 5 minutes is $1 - U(T = 5\,\text{min}) = \exp(-\lambda T) = \exp(-5[\text{min}]/2[\text{min}]) = 0.08$.

**A11.** To sample a random time $T$ from an exponential distribution, one can sample uniformly a variable $\epsilon \in [0, 1]$, then find the value of $T$ such that:

$$\int_0^T P(\tau)\, d\tau = \epsilon\,. \tag{41}$$

i.e. such that $U(T) = \epsilon$. As $U(T) = 1 - \exp(-\lambda\, T)$ is a strictly growing function of $T$, this equation has only a unique solution, which can be easily computed:

$$1 - \exp(-\lambda\, T) = \epsilon \qquad \Longleftrightarrow \qquad T = -\frac{1}{\lambda} \log(1 - \epsilon)\,. \tag{42}$$

Note that if $\epsilon$ is uniformly distributed over $[0, 1]$, then $\eta = 1 - \epsilon$ is also uniformly distributed over $[0, 1]$. Therefore, to sample a random time $T$ from an exponential distribution with rate $\lambda$, one can sample uniformly a variable $\eta \in [0, 1]$ and return $T = -\frac{1}{\lambda} \log(\eta)$.

To generate data for the supermarket problem more efficiently, one can directly sample from the exponential distribution $P(\tau)$ the waiting time until the next customer arrives using the method just described.

**Summary 1. Poisson process.**

If $N(t)$ follows a Poisson process with parameter $\lambda > 0$, then:

- the **probability to observe $k$ events** in a given time interval $t$ is given by the **Poisson distribution**:

$$\mathbb{P}[N(t) = k] = \frac{\mu^k \exp(-\mu)}{k!}, \qquad \text{where } \mu = \lambda t\,. \tag{43}$$

- the **distribution of the waiting time** $\tau$ between two events is the **exponential distribution**:

$$P(\tau) = \lambda \exp(-\lambda \tau)\,. \tag{44}$$

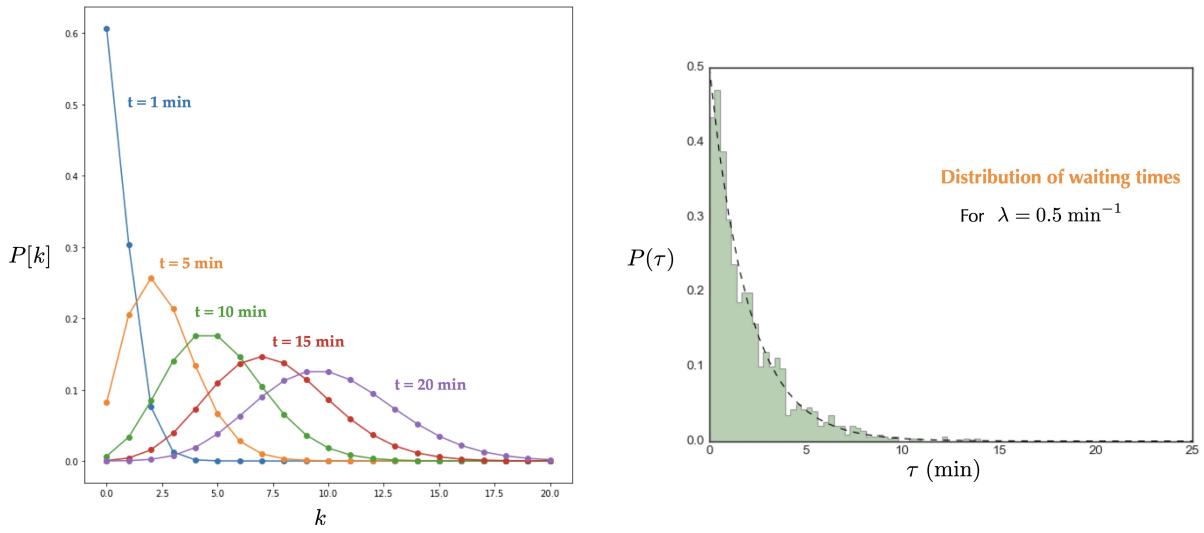



Figure 1: **Left: Poisson distribution.** The probability to observe $k$ events in a given time interval $t$ is given by the Poisson distribution (which is a discrete probability distribution). The figure shows the probability that $k$ customers arrives in a given time interval $t$ for the process in the exercise above with rate $\lambda = 0.5$ customers per minute. **Right: Distribution of waiting times.** The duration of the time-intervals between two consecutive events (also called *waiting times*) in a Poisson process follows an exponential distribution. This is a continuous probability distribution. The figure shows the distribution of waiting times $\tau$ for the process in the exercise above with rate $\lambda = 0.5$ customers per minute.

**Summary 2. Sampling from an Exponential Distribution.**

If $\eta$ is uniformly distributed over $[0, 1]$, then $T = -\frac{1}{\lambda}\log(\eta)$ follows an exponential distribution with rate $\lambda$.

### 0.1.2 TASEP: a simple Traffic Model

**A1.** During $dt$, the following events can happen:

- if the site $i = 1$ is empty, then a car can enter from the left with probability $\alpha\, dt$;
- if the site $i = L$ is occupied, then the car in $i = L$ can exit the lattice from the right with probability $\beta\, dt$;
- each car inside the lattice can hop to the right with probability $dt$ if the site to their right is empty.

**A2.** A first possible algorithm consists in simulating the evolution of the system by small time steps of length $dt$. During each time step, one must check if any of the events listed in the answer of question Q1 happens using their respective probability to occur. One must take $dt$ sufficiently small so that the probability that the same car moves twice (or more) during

$dt$ is negligeable, but large enough for the simulation to be executed in a reasonable time.

A second version of the algorithm consists in directly sampling the next time at which one event will happen, and thus to jump forward in time directly to that time. This approach is based on the fact that, for each of the possible action (car enters from the left, car moves forward, car exits from the right), the time interval between two actions are exponentially distributed (with respective rates $\alpha$, 1, $\beta$). This version of the algorithm will be much faster than the first version. In this version, we iterate in the following way: (1) at the current time $t$, we sample:

- if the first site $i = 1$ is empty, we sample the time $T_{in}$ at which the next car with enter the lattice from the left; this time is sampled from an exponential distribution with rate $\alpha$;
- if the last site $i = L$ is occupied, we sample the time $T_{out}$ at which the car will leave the lattice from the right; this time is sampled from an exponential distribution with rate $\beta$;
- the time $T_{hop}$ at which the next car on the lattice will move to the right; this time is sampled from an exponential distribution with rate $N \times 1$, where $N$ is the number of cars on the lattice that have an empty spot on their right (i.e. that can move to the right).

We then take the shorter of the sampled times ($T = \min(T_{in}, T_{out}, T_{hop})$), move forward in time to that time ($t = t + T$), and perform the corresponding action. Note that if $T_{hop}$ is the shortest time, we then uniformly sample one of the $N$ cars on the lattice to hop to the right.

**A3.** By definition, the average number of cars in $i$ at time $t$ is given by: $\overline{n_i(t)} = \sum_{n_i} n_i P[n_i(t)]$, where the sum is over the values that can be taken by $n_i$. As $n_i$ can only take the two values 0 and 1, we get:

$$\overline{n_i(t)} = 1 \times P[n_i(t) = 1] + 0 \times P[n_i(t) = 0] \tag{45}$$
$$= \rho_i(t) \tag{46}$$

**A4.** The probability that a car hops from node $i$ to node $i + 1$ during the time interval $[t, t + dt)$ is:

$$J_i(t)\, dt = \underbrace{P[n_i(t) = 1; n_{i+1}(t) = 0]}_{(1)} \underbrace{dt}_{(2)} \tag{47}$$

where:

(1) is the probability that there is a car in $i$ and no car in $i + 1$ at time $t$;

(2) is the probability that the car moves to the right.

**Comment:** there is a common mistake done by students here, which is to write the probability of $n_i(t) = 1$ and $n_{i+1} = 0$ as a product of the two probabilities, $P[n_i(t) = 1] \times P[n_{i+1}(t) = 0]$, instead of as the joint probability $P[n_i(t) = 1; n_{i+1}(t) = 0]$. This is not correct. $P[n_i(t); n_{i+1}(t)]$ is a joint probability distribution over 4 possibles states $((n_i(t), n_{i+1}(t)) \in \{(0, 0), (0, 1), (1, 0), (1, 1)\})$. Writing down that:

$$P[n_i(t); n_{i+1}(t)] = P[n_i(t)] \times P[n_{i+1}(t)] \qquad \text{(not true in general)} \tag{48}$$

assumes that the states taken by the variables $n_i$ are independent of the states taken by $n_{i+1}$ (and reciprocally that the states of $n_{i+1}$ doesn't depends on the state of $n_i$). For instance, Eq. (48) assumes that the probability that $n_{i+1} = 1$ doesn't depends on the state of $n_i$. A priori, there is no reason to assume independence between $n_i$ and $n_{i+1}$ without any specific knowledge about the system, so in general Eq. (48) is not valid. Note that in the tutorial on mean-field theory, you will see that this independence is assumed in the mean-field approximation.

**A5.** To compute the equation of evolution of $\rho_i(t)$, one must consider the evolution of the system during a very short time interval $[t, t + dt)$. For that, the idea is to compute the probability $\rho_i(t + dt)$ that the site $i$ is occupied by a car at time $t + dt$ by relating it to the states of the system at time $t$. There are two possibilities for the site $i$ to be occupied at time $t + dt$:

1. at time $t$, the site $i$ was empty and the site $(i-1)$ was occupied, AND the car moved from $(i-1)$ to $i$ during $dt$; The probability of this event corresponds to item (1) in the equation below.
2. at time $t$, the site $i$ was occupied by a car that didn't move out to the right during $dt$. This corresponds to item (2) in the equation below and can be decomposed into two cases:
   - 2a. at time $t$, the site $i$ and $i+1$ were both occupied (and therefore the car in $i$ can't move to the right during $dt$). The probability of this event corresponds to item (2a) in the equation below.
   - 2b. at time $t$, the site $i$ was occupied and the site $i+1$ was empty, but the car didn't move to the right during $dt$. The probability of this event corresponds to item (2b) in the equation below.

These considerations can be translated into the following equation:

$$P[n_i(t+dt)=1] = \underbrace{P[n_{i-1}(t)=1; n_i(t)=0] \times \underbrace{dt}_{(\star)}}_{(1)} + \underbrace{\underbrace{P[n_i(t)=1; n_{i+1}(t)=1]}_{(2a)} + \underbrace{P[n_i(t)=1; n_{i+1}(t)=0]\,(1-dt)}_{(2b)}}_{(2)}, \tag{49}$$

where

(1) is the probability that a car moves from $(i-1)$ to $i$ between the times $t$ and $t+dt$;

($\star$) is the probability that a chosen car moves to the right during $dt$;

(2) is the probability that the site $i$ was already occupied at time $t$ and stays occupied during $dt$;

Using that $P[n_i(t)=1; n_{i+1}(t)=1] + P[n_i(t)=1; n_{i+1}(t)=0] = P[n_i(t)=1]$, we can re-write the term (2) as:

$$(2) = P[n_i(t)=1] - P[n_i(t)=1; n_{i+1}(t)=0]\,dt\,, \tag{50}$$

which could have also been directly obtained as the probability that $i$ was occupied at time $t$ minus the probability that $i$ becomes empty during $dt$. Finally, using the definition of $\rho_i(t)$ and $J_i(t)$, we get:

$$\frac{\rho_i(t+dt)-\rho_i(t)}{dt} = J_{i-1}(t) - J_i(t)\,, \qquad \text{for all } i \in \{2, \cdots, (L-1)\}\,, \tag{51}$$

which leads to the equation of evolution for $\rho_i(t)$ given in Eq. (3.9).

**Comment:** Note that Eq. (3.9) can also be interpreted in the following way: the variation of the density of cars (where here the density of cars is the average number of cars) in cell i during dt is equal to the average number of cars that come in the cell during dt (i.e., $J_{i-1}dt$) minus the average number of cars that come out during dt (i.e., $J_i dt$).

**A6.** Similarly to question Q5, we can write the probability that the site $i=1$ is occupied at time $t+dt$:

$$P[n_1(t+dt)=1] = P[n_1(t)=0]\alpha\,dt + P[n_1(t)=1] - P[n_1(t)=1; n_2(t)=0]\,dt\,, \tag{52}$$

which gives:

$$\frac{d\rho_1}{dt}(t) = J_0(t) - J_1(t)\,, \qquad \text{where} \qquad J_0(t) = \alpha\,(1-\rho_1(t))\,. \tag{53}$$

Similarly, one gets for the current on the right end of the lattice:

$$\frac{d\rho_L}{dt}(t) = J_{L-1}(t) - J_L(t)\,, \qquad \text{where} \qquad J_L(t) = \beta\,\rho_L(t)\,. \tag{54}$$

**A7.** At stationarity (i.e. in the stationary state), the local densities are independent from time, which means that $\frac{d\rho_i}{dt}(t) = 0$. Replacing this value in Eq. (3.9), we obtain that $J_{i-1} = J_i$ for all $i$ from 2 to $(L-1)$, i.e. the current is constant:

$$J_i = J = \text{constant}\,, \qquad \text{for all } i \in \{1, \cdots, (L-1)\}\,. \tag{55}$$

| beta \ alpha | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 | 0,6 | 0,7 | 0,8 | 0,9 | 1,0 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1,0 | 0,09 | 0,16 | 0,21 | 0,24 | 0,25 | 0,26 | 0,27 | 0,27 | 0,27 | 0,26 |
| 0,9 | 0,09 | 0,16 | 0,21 | 0,24 | 0,25 | 0,26 | 0,26 | 0,27 | 0,27 | 0,27 |
| 0,8 | 0,09 | 0,16 | 0,21 | 0,24 | 0,24 | 0 | 0 | 0,27 | 0 | 0,27 |
| 0,7 | 0,09 | 0,16 | 0,21 | 0,24 | 0 | 0 | 0,26 | 0 | 0,25 | 0,27 |
| 0,6 | 0,09 | 0,17 | 0,21 | 0,24 | 0 | 0,26 | 0 | 0,22 | 0 | 0,26 |
| 0,5 | 0,09 | 0,15 | 0,21 | 0,24 | 0,25 | 0,25 | 0,25 | 0 | 0,26 | 0,25 |
| 0,4 | 0,09 | 0,16 | 0,2 | 0,23 | 0,24 | 0 | 0 | 0 | 0,23 | 0,24 |
| 0,3 | 0,09 | 0,16 | 0,2 | 0,2 | 0,21 | 0,21 | 0,21 | 0,21 | 0,21 | 0,21 |
| 0,2 | 0,09 | 0,15 | 0,15 | 0,15 | 0,16 | 0,16 | 0,16 | 0,16 | 0,15 | 0,16 |
| 0,1 | 0,08 | 0,09 | 0,09 | 0,09 | 0,09 | 0,09 | 0,09 | 0,08 | 0,09 | 0,09 |

Figure 2: **TASEP phase diagram:** Numerical values of the stationary current $J$ for different values of $\alpha$ and $\beta$, filled in by the students of year 2023, based on the mathematica program available in canvas.

Similarly, at stationarity, $\frac{d\rho_1}{dt}(t) = 0$ and $\frac{d\rho_L}{dt}(t) = 0$. Replacing this respectively in Eq. (53) and in Eq. (54), we get that:

$$J_0 = J = J_L \,, \tag{56}$$

where $J_0 = \alpha\,(1 - \rho_1)$ and $J_L = \beta\,\rho_L$.

**A8.** The answers to this question are discussed at the beginning of the lecture L1 (and see Fig. 2 below).

# Chapter 1

# Markov Chains
# Understanding the Metropolis-Hastings algorithm for simulating the Ising model

**Goal of this tutorial** is for students to get familiar with the basis of the theory behind the Metropolis-Hastings algorithm, and with its application to the numerical simulation of the Ising model. After this tutorial, students should:

- be more familiar with the notion of microscopic states VS macroscopic properties of a complex systems;
- become more confident on how to write a continuous time Markov chain;
- know what is the condition of *detailed balance*.
- become more familiar with the Boltzmann distribution and with the Ising model.
- understand where the Metropolis-Hastings algorithm comes from and know how to apply it for the Ising model.
- know how to compute ensemble averages from a Metropolis simulation of the system.

The Metropolis-Hastings algorithm is a well-known Markov Chain Monte Carlo (MCMC) algorithm that is commonly used to sample from probability distribution for which direct sampling is difficult (such as multi-variate probability distributions, like the Boltzmann distribution in Eq. (1.2)).

**Important comment about the notations:** Note that to match the notation used in the lectures and to avoid confusion with the external magnetic field (which is often denoted $h$ or $H$), we denote the energy as $E(\boldsymbol{s})$ instead of $H(\boldsymbol{s})$. In the literature, you may conventionally find the energy expressed as the Hamiltonian of the system and denoted $H(\boldsymbol{s})$.

## 1.1 Markov chain for sampling from an equilibrium system

One of the simplest way to study the behavior of a stochastic system is to resort to numerical simulation. To study the possible phases of a system at equilibrium, we must look at its macroscopic properties, which are obtained by computing the ensemble averages (or "thermal average") of different quantities characterizing the system. By definition, the ensemble average of a quantity $A(\boldsymbol{s})$ is given by the average of $A(\boldsymbol{s})$ over all the states $\boldsymbol{s}$ accessible to the system at equilibrium:

$$\langle A(\boldsymbol{s}) \rangle = \sum_{\boldsymbol{s}} A(\boldsymbol{s}) P(\boldsymbol{s}) \,. \tag{1.1}$$

In this equation, the sum is over all the states $\boldsymbol{s}$ of the system, and $P(\boldsymbol{s})$ is the probability to observe the system in the state $\boldsymbol{s}$ at equilibrium, which is often written under the form:

$$P(\boldsymbol{s}) = \frac{\exp(-\beta E(\boldsymbol{s}))}{\mathcal{Z}}, \tag{1.2}$$

where $E(\boldsymbol{s})$ is the energy[a] of the system when it is in the state $\boldsymbol{s}$, the parameter $\beta = 1/T$ is the inverse temperature, and $\mathcal{Z}$ is the partition function, which is obtained from the normalization of $P(\boldsymbol{s})$. The parameter $\beta$ (or $T$) can also be seen as controlling the noise level of the system: the larger is $\beta$, the smaller is the noise. In practice, it is not computationally efficient (and, most of the time, impossible) to compute all the terms of the sum in Eq. (1.1). Just the computation of the partition function itself would be already computationally demanding.

[a] In physics, $E(\boldsymbol{s})$ is also known as the Hamiltonian of the system and is often denoted $H(\boldsymbol{s})$ (instead of $E(\boldsymbol{s})$).

**Q1.** For instance, in the context of the Ising model, the simplest quantity to look at is the local average magnetisation $m_i = \langle s_i \rangle$. What are the possible microstates $\boldsymbol{s}$ of a system of $n$ spins in the Ising model? Using the definition of the ensemble average in Eq. (1.1), what is the expression of $m_i$? What is $A(\boldsymbol{s})$ in this case? what is $E(\boldsymbol{s})$? and how is $Z$ computed?

**Q2.** Consider an Ising model on a 1d-lattice with 20 sites: how many terms are there in the sum of Eq. (1.1) in that case? How many terms are there in the sum to compute $Z$?

**What can we do?** If one was able to generate many random states of the system with their respective probabilities[1] $P(\boldsymbol{s})$, then an approximation of the ensemble average would be given by:

$$\langle A(\boldsymbol{s}) \rangle \simeq \frac{1}{N} \sum_{r=1}^{N} A(\boldsymbol{s}^{(r)}), \tag{1.3}$$

where $N$ is the total number of random states that were sampled, and $\boldsymbol{s}^{(r)}$ denotes the $r$-th sampled states. $A(\boldsymbol{s}^{(r)})$ is simply the value of the function $A(\boldsymbol{s})$ when the system is in state $\boldsymbol{s}^{(r)}$. The number of terms needed to compute $\langle A \rangle$ can now be way smaller. However, this is not sufficient, as in order to sample states from $P(\boldsymbol{s})$ one would still need to compute $Z$ (which still a sum over a very large number of terms).

**Q3.** Let's try to clarify Eq. (1.3). Consider a dataset composed of $N$ random states of the system sampled from $P(\boldsymbol{s})$. The right-hand side of Eq. (1.3) corresponds to the average value of $A(\boldsymbol{s})$ over the dataset, let us denote it $\langle A \rangle_{data}$:

$$\langle A \rangle_{data} = \frac{1}{N} \sum_{r=1}^{N} A(\boldsymbol{s}^{(r)}), \tag{1.4}$$

whereas the left-hand side of Eq. (1.3) is the thermal average of $A(\boldsymbol{s})$ defined in Eq. (1.1). Let's try to clarify Eq. (1.3) with a few additional steps. How can one define the empirical probability $P_{data}(\boldsymbol{s})$ to observe a given state $\boldsymbol{s}$ in the data? Can you re-write $\langle A \rangle_{data}$ using $P_{data}(\boldsymbol{s})$? What happens to $\langle A \rangle_{data}$ when the number of datapoints become very large?

### 1.1.1 Defining a Stochastic Markovian process

The trick proposed by Metropolis, Rosenbluth and Teller in 1953 consists of introducing a stochastic Markovian process between successive configurations, $\boldsymbol{s}^{(r)} \to \boldsymbol{s}^{(r+1)}$, that converges towards the desired equilibrium distribution. One can then obtain many samples of the desired distribution by recording states from the random process at equilibrium. The more steps are included, the more closely the distribution of the sample matches the actual desired distribution.

[1] i.e., randomly generate more often states that are more likely to occur (with a larger $P(\boldsymbol{s})$), and less often states that are less likely to occur (with a smaller $P(\boldsymbol{s})$).

**Introducing a Markovian process.** To do so, we must first introduce a numerical time $t$, which starts at $t = 0$ at the beginning of the simulation. We then denote by $p(\boldsymbol{s}, t)$ the probability that the system is in the configuration $\boldsymbol{s}$ at time $t$ during the simulation. To introduce a stochastic process, we must define the probabilities with which the system jumps from state to state during a very small time interval $[t, t + dt)$ (for instance, take $dt = 1/n$, where $n$ is the number of spins in the system). Besides, defining a *Markovian process* means that the probability of jumping from a state $\boldsymbol{s}$ to a state $\boldsymbol{s}'$ during $dt$ only depends on the state $\boldsymbol{s}$ where the system was at time $t$, and not on the previous states of the system (the process is memory-less). We thus define the probability that the system jumps from $\boldsymbol{s}$ to $\boldsymbol{s}'$ during $dt$ as: $w(\boldsymbol{s} \to \boldsymbol{s}')\, dt$.

**Q4.** Can you show that the probability to find the system in state $\boldsymbol{s}$ at time $t + dt$ is given by:

$$p(\boldsymbol{s}, t + dt) = p(\boldsymbol{s}, t) + \sum_{\boldsymbol{s}' \neq \boldsymbol{s}} p(\boldsymbol{s}', t)\, w(\boldsymbol{s}' \to \boldsymbol{s})\, dt - \sum_{\boldsymbol{s}' \neq \boldsymbol{s}} p(\boldsymbol{s}, t)\, w(\boldsymbol{s} \to \boldsymbol{s}')\, dt\,, \tag{1.5}$$

where the sums are over all states $\boldsymbol{s}'$ of the system that are different from $\boldsymbol{s}$? Deduce the equation for the time evolution of $p(\boldsymbol{s}, t)$.

### 1.1.2 Convergence to equilibrium and Detailed Balance equation

**Driving the system towards the desired equilibrium.** The stationary solution of the previous equation is:

$$\sum_{\boldsymbol{s}'} p_{st}(\boldsymbol{s}')\, w(\boldsymbol{s}' \to \boldsymbol{s}) = \sum_{\boldsymbol{s}'} p_{st}(\boldsymbol{s})\, w(\boldsymbol{s} \to \boldsymbol{s}')\,, \tag{1.6}$$

where $p_{st}(\boldsymbol{s})$ are the stationary probabilities of the states of the system for the simulated process. As we would like this process to converge towards the desired probability distribution $p_{st}(\boldsymbol{s}) = P(\boldsymbol{s})$ given by Eq. (1.2), this latter equation therefore gives us a set of constraints that must be satisfied by $w(\boldsymbol{s} \to \boldsymbol{s}')$.
A simple solution is given by:

$$P(\boldsymbol{s}')\, w(\boldsymbol{s}' \to \boldsymbol{s}) = P(\boldsymbol{s})\, w(\boldsymbol{s} \to \boldsymbol{s}')\,, \tag{1.7}$$

which is known as the condition of **detailed balance**. It expresses that at stationarity, there is the same probability to observe the system transition from a state $\boldsymbol{s}$ to a state $\boldsymbol{s}'$ then to observe the reverse transition. Note that this not necessarily the unique solution, but in general, it is difficult to prove that a system that doesn't satisfy detailed balance will converge to equilibrium, which is why many Monte Carlo algorithms use detailed balance.

**Q5.** Can you show that Eq. (1.7) implies that $w(\boldsymbol{s} \to \boldsymbol{s}')$ does not depend on the partition function $Z$, but only on the Boltzmann factor $\exp(-\beta\, E(\boldsymbol{s}))$? We recall that the desired equilibrium distribution $P(\boldsymbol{s})$ is given by Eq. (1.2).

This is great, because it means that if we find a solution $w(\boldsymbol{s} \to \boldsymbol{s}')$ that satisfies Eq. (1.7), we will be able to design a Markovian stochastic process that converges towards the desired equilibrium $P(\boldsymbol{s})$ without the need to sample the partition function! Once the process is at equilibrium, it will stay there (by definition), and we will be able to get many samples of the system at equilibrium and therefore compute the thermal average of any quantity $A(\boldsymbol{s})$ using Eq. (1.3).

## 1.2 The Metropolis–Hastings algorithm

A Monte Carlo algorithm works as follows. We start from a randomly sampled initial state (out-of-equilibrium) at time $t = 0$. We then consider very short time increments $dt$ (typically, for an Ising model it is common to take $dt = 1/n$, where $n$ is the total number sites), and during each time step $dt$ we perform the two operations:

- from a current state $\boldsymbol{s}$ of the system, sample a random new state $\boldsymbol{s}'$;
- the new state is accepted with probability $\Pi(\boldsymbol{s} \to \boldsymbol{s}')$.

The choice for $\Pi(\boldsymbol{s} \to \boldsymbol{s}')$ introduced by Metropolis et al. is:

$$\Pi(\boldsymbol{s} \to \boldsymbol{s}') = \begin{cases} 1\,, & \text{if } E(\boldsymbol{s}') \leq E(\boldsymbol{s})\,, \\ \exp[-\beta\,(E(\boldsymbol{s}') - E(\boldsymbol{s}))]\,, & \text{otherwise}\,. \end{cases} \tag{1.8}$$

Note that each step of the numerical simulation corresponds to a very short time increment $dt$, which means that $\Pi(\boldsymbol{s} \to \boldsymbol{s}')$ is the probability that the systems goes from $\boldsymbol{s}$ to $\boldsymbol{s}'$ during $dt$, and therefore $\Pi(\boldsymbol{s} \to \boldsymbol{s}') = w(\boldsymbol{s} \to \boldsymbol{s}')\,dt$.

**Q6.** Can you show that this choice of $\Pi(\boldsymbol{s} \to \boldsymbol{s}')$ satisfies the detailed balance equation (1.7)?

**Q7.** Can you give a general interpretation of the stochastic process defined by Eq. (1.8)? We recall that:

$$\exp[-\beta\,(E(\boldsymbol{s}') - E(\boldsymbol{s}))] = \frac{P(\boldsymbol{s}')}{P(\boldsymbol{s})}\,. \tag{1.9}$$

## 1.3 Application to the Ising model

For a trial state $\boldsymbol{s}'$ to be accepted, this state needs to be "close" to the previous state $\boldsymbol{s}$ in energy, because the acceptance probability is proportional to the exponential of the energy difference between the two states. If this difference is large and positive, the probability of acceptance becomes very small and the system can stay "trapped" for a long time in a local minimum. For the Metropolis algorithm, the single spin flip is generally used which leads to a dynamics where energy changes are reasonable. Because the dynamics must remain stochastic, a spin must be chosen randomly for each trial configuration, and not according to a regular sequence.

**Q8.** What is the energy difference, $\Delta E = E(\boldsymbol{s}') - E(\boldsymbol{s})$, between a state $\boldsymbol{s}$ and a new state $\boldsymbol{s}'$ in which the spin $s_k$ at site $k$ has been flipped? Show that one can re-write the ratio of the state probabilities as:

$$\frac{P(\boldsymbol{s}')}{P(\boldsymbol{s})} = \exp\Big(-\beta\, h_k^{loc} s_k\Big)\,, \tag{1.10}$$

where $s_k$ is the value of the spin at site $k$ in the current state $\boldsymbol{s}$, and where $h_k^{loc}$ is a local field to be specified.

**Q9.** If a spin is flipped, what is the energy of a new configuration $E(\boldsymbol{s}')$ as a function of the old one? What is the value of the magnetization of the whole system $M(\boldsymbol{s}')$ in the new state as a function of the old one?

Note that in the case where the old configuration is kept, instantaneous quantities are unchanged, but time must be incremented.

**Q10.** Using these recursive relations provides a computationally efficient way to keep track of the energy $E(\boldsymbol{s})$ and total magnetization $M(\boldsymbol{s})$ of the system during the simulation. For the Ising model, all the thermodynamic quantities of interest can be expressed as statistical moments of $E(\boldsymbol{s})$ and $M(\boldsymbol{s})$ (i.e. as the average of powers of $E$ and $M$), which means that the computation of thermodynamic quantities can be easily performed (not computationally demanding). For instance, can you recall the expression of the specific heat and that of the susceptibility as functions of $E(\boldsymbol{s})$ and $M(\boldsymbol{s})$?

**Q11.** The simulation starts from a random initial state (out-of-equilibrium), we then impose a fixed temperature $T$ (i.e. a fixed value of $\beta$), and let the system evolves towards equilibrium using a Metropolis algorithm. We would like to record the average value of the magnetization $\langle M \rangle = \sum_{\boldsymbol{s}} M(\boldsymbol{s})$ in the stationary state. Can you describe the different steps of the program?

**Q12. 2D Ising model.** in 1944, Onsager obtained an exact solution for the 2D Ising model with external magnetic field. He showed that there exists a ferromagnetic transition at a critical temperature:

$$T_c = \frac{2\,J}{\log\left(1+\sqrt{2}\right)} \simeq 2.2691\,J\,. \tag{1.11}$$

You can find here an online implementation of the Metropolis algorithm for a 2D Ising model. For the Ising model, it is known that the phase transition happens at $T_c \simeq 2.27\ldots$. Play with the parameters of the algorithm. How does the systems look like when you take $T > T_c$? How does it looks like when you take $T < T_c$? Can you observe a phase transition? Can you see appearing magnetic domains for $T < T_c$? How does it looks like when you take $T = 0$? Can you explain why?

**Bonus question:** If you have some spare time, feel free to implement your own numerical simulation of the Ising model!

## 1.4 Solutions

**Q1:** We consider a system of Ising spins. We denote by $s_i$ is the value $\pm 1$ of the spin located at the site $i$, and by $\boldsymbol{s}$ the state of all the spins of the system. The local magnetization $m_i$ is the ensemble average of $A(\boldsymbol{s}) = s_i$, which can be written as:

$$m_i = \langle s_i \rangle = \sum_{\boldsymbol{s}} s_i \, P(\boldsymbol{s}) \, . \tag{1.12}$$

For the Ising model, the *state probability* (i.e. the probability that the system is in a given state $\boldsymbol{s}$) is defined by:

$$P(\boldsymbol{s}) = \frac{\exp(-\beta \, E(\boldsymbol{s}))}{Z} \, , \qquad \text{where} \;\; E(\boldsymbol{s}) = -h \sum_i s_i - J \sum_{<i,j>} s_i s_j \, . \tag{1.13}$$

In this equation, $h$ is a parameter that corresponds to an external magnetic field applied to all the spins of the system, the sum over $< i, j >$ denotes a sum over neighboring pairs of spins $s_i$ and $s_j$, and $J$ is a coupling parameter that corresponds to the strength of the coupling between neighboring spins $s_i$ and $s_j$. The partition function $Z$ is then obtained from the normalisation of $P(\boldsymbol{s})$ (i.e. $\sum_{\boldsymbol{s}} P(\boldsymbol{s}) = 1$), which gives that:

$$Z = \sum_{\boldsymbol{s}} \exp(-\beta \, E(\boldsymbol{s})) \, . \tag{1.14}$$

**Q2:** A system with 20 spins can be in $2^{20}$ different states, which means that there are $2^{20} \simeq 10^6$ terms in the sum in Eq. (1.1). Idem for the sum to compute $Z$. These are a lot of terms to sum over to compute $Z$, or to compute any ensemble average (and 20 spins is only a small system!). This also means that the distribution $P(\boldsymbol{s})$ is a distribution over $\sim 10^6$ elements, so sampling from that distribution is not very practical, and many of the states will actually have a very small probability. Ideally, we would like to avoid having to compute these huge sums to compute observable of the systems. Ideally, it would also be good to be able to take into account mostly the contributions of the most likely states to the sum.

**Q3.** The empirical probability $P_{data}(\boldsymbol{s})$ to observe the state $\boldsymbol{s}$ in the data is given by the ratio of the number of times $K(\boldsymbol{s})$ that the state $\boldsymbol{s}$ appears in the data to the total number of datapoints:

$$P_{data}(\boldsymbol{s}) = \frac{K(\boldsymbol{s})}{N} \, . \tag{1.15}$$

We can re-write $\langle A \rangle_{data}$ as:

$$\langle A \rangle_{data} = \frac{1}{N} \sum_{r=1}^{N} A(\boldsymbol{s}^{(r)}) = \frac{1}{N} \sum_{\boldsymbol{s}} K(\boldsymbol{s}) \, A(\boldsymbol{s}) = \sum_{\boldsymbol{s}} \frac{K(\boldsymbol{s})}{N} \, A(\boldsymbol{s}) \tag{1.16}$$

$$= \sum_{\boldsymbol{s}} P_{data}(\boldsymbol{s}) \, A(\boldsymbol{s}) \, , \tag{1.17}$$

where we replaced the sum over all the datapoints by a sum over all the states weighted by the number of times each state occurs. Note that most of the times, many states are never observed in the dataset, and for these states $K(\boldsymbol{s}) = 0$. As the number of datapoints $N$ increases, the empirical distribution $P_{data}(\boldsymbol{s})$ converges to the true distribution $P(\boldsymbol{s})$, and we have that:

$$\langle A \rangle_{data} \underset{N \to \infty}{\longrightarrow} \langle A \rangle \, , \tag{1.18}$$

in which we recall that $\langle A \rangle = \sum_{\boldsymbol{s}} P(\boldsymbol{s}) \, A(\boldsymbol{s})$ and $\langle A \rangle_{data} = \sum_{\boldsymbol{s}} P_{data}(\boldsymbol{s}) \, A(\boldsymbol{s})$.

**Q4:** The probability to find the system in state $\boldsymbol{s}$ at time $t + dt$ is given by:

$$p(\boldsymbol{s}, t + dt) = \underbrace{\sum_{\boldsymbol{s}' \neq \boldsymbol{s}} p(\boldsymbol{s}', t) \, w(\boldsymbol{s}' \to \boldsymbol{s}) \, dt}_{(1)} + \underbrace{p(\boldsymbol{s}, t) \left( 1 - \sum_{\boldsymbol{s}' \neq \boldsymbol{s}} w(\boldsymbol{s} \to \boldsymbol{s}') \, dt \right)}_{(2)} , \tag{1.19}$$

where

(1) is the probability that the system was in state $\boldsymbol{s}'$ at time $t$ and transition from $\boldsymbol{s}'$ to $\boldsymbol{s}$ during $dt$.

(2) is the probability that the system was in state $\boldsymbol{s}$ at time $t$ and stayed in $\boldsymbol{s}$ during $dt$ (i.e. did not transition to another state during $dt$).

Re-organizing the terms gives Eq. (1.19).

**Q5:** Replacing the expression of $P(\boldsymbol{s})$ in Eq. (1.7) gives:

$$\frac{w(\boldsymbol{s} \to \boldsymbol{s}')}{w(\boldsymbol{s}' \to \boldsymbol{s})} = \frac{P(\boldsymbol{s}')}{P(\boldsymbol{s})} = \mathrm{e}^{-\beta(E(\boldsymbol{s}') - E(\boldsymbol{s}))} \,. \tag{1.20}$$

which doesn't depends on $Z$ anymore. This implies that $w(\boldsymbol{s} \to \boldsymbol{s}')$ does not depend on the partition function $Z$.

**Q6.** We can decompose the problem in two cases:

- if $E(\boldsymbol{s}') \leq E(\boldsymbol{s})$, then:

$$w(\boldsymbol{s} \to \boldsymbol{s}') = 1 \qquad \text{and} \qquad w(\boldsymbol{s}' \to \boldsymbol{s}) = \mathrm{e}^{-\beta(E(\boldsymbol{s}) - E(\boldsymbol{s}'))} = \frac{P(\boldsymbol{s})}{P(\boldsymbol{s}')} \,, \tag{1.21}$$

- or $E(\boldsymbol{s}') > E(\boldsymbol{s})$, then:

$$w(\boldsymbol{s} \to \boldsymbol{s}') = \mathrm{e}^{-\beta(E(\boldsymbol{s}') - E(\boldsymbol{s}))} = \frac{P(\boldsymbol{s}')}{P(\boldsymbol{s})} \qquad \text{and} \qquad w(\boldsymbol{s}' \to \boldsymbol{s}) = 1 \,. \tag{1.22}$$

In both case, we recover the detailed balance Eq. (1.7).

**Q7.** According to Eq. (1.8), if the system is more likely to be in the new state $\boldsymbol{s}'$ than in the current state $\boldsymbol{s}$ at equilibrium, then the algorithm will move the system to the new state with probability 1. If instead the new state is less likely than the current state at equilibrium, then the algorithm will still allow a transition of to the new state, but with a probability that is proportional to the ratio of the probabilities of the two states at equilibrium. Thus, if $\boldsymbol{s}'$ is very unlikely compared to $\boldsymbol{s}$, then the ratio will be close to zero and a transition to $\boldsymbol{s}'$ will be unlikely, and if $\boldsymbol{s}'$ is almost as probable as $\boldsymbol{s}$, then the ratio will be close to 1 and a transition to $\boldsymbol{s}'$ will still have a high chance to occur.

Another way to see this, is that: The stochastic process is driving the system towards lower energy states, while still allowing small moves towards higher energy states. This in fact prevents the system from getting stuck in local minima.

**Q8.** One can write the two states $\boldsymbol{s} = (s_1, \cdots, s_k, \cdots, s_n)$ and $\boldsymbol{s}' = (s_1, \cdots, -s_k, \cdots, s_n)$, which differ only by a flip of $s_k$. Let us separate the contributions of the spin $s_k$ from the rest in $E(\boldsymbol{s})$:

$$E(\boldsymbol{s}) = -\left( h \sum_{i \neq k} s_i + J \sum_{<i,j>, i,j \neq k} s_i s_j \right) - \left( h\, s_k + J \sum_{i:<i,k>} s_i s_k \right) . \tag{1.23}$$

Idem for $E(\boldsymbol{s}')$:

$$E(\boldsymbol{s}) = -\left( h \sum_{i \neq k} s_i + J \sum_{<i,j>, i,j \neq k} s_i s_j \right) - \left( h\,(-s_k) + J \sum_{i:<i,k>} s_i (-s_k) \right) , \tag{1.24}$$

$$= -\left( h \sum_{i \neq k} s_i + J \sum_{<i,j>, i,j \neq k} s_i s_j \right) + \left( h\, s_k + J \sum_{i:<i,k>} s_i s_k \right) . \tag{1.25}$$

The energy difference $\Delta E = E(\boldsymbol{s}') - E(\boldsymbol{s})$ is then:

$$\Delta E = +2 \left( h\, s_k + J \sum_{i:<i,k>} s_i s_k \right) , \tag{1.26}$$

$$= h_k^{loc}\, s_k \,, \qquad \text{where} \quad h_k^{loc} = 2\,h + 2\,J \sum_{i:<i,k>} s_i \,. \tag{1.27}$$

The quantity $h_i^{loc}$ is a local field created by the 4 neighbors of the spin $s_k$ (the notation "$i :< i, k >$" indicates a summation over the sites $i$, such that $i$ is a neighbor of $k$). As $P(\boldsymbol{s}')/P(\boldsymbol{s}) = \exp(-\beta\,(E(\boldsymbol{s}') - E(\boldsymbol{s}))) = \exp(-\beta\,\Delta E)$, one directly recovers Eq. (1.10).

**Q9.** Since $\Delta E = E(\boldsymbol{s}') - E(\boldsymbol{s})$, we have that the new energy is expressed as a function of the old energy by:

$$E(\boldsymbol{s}') = E(\boldsymbol{s}) + \Delta E = E(\boldsymbol{s}) + h_k^{loc} s_k \tag{1.28}$$

where $s_k$ is spin that is flipped from the old to the new configuration. To compute the new magnetization $M(\boldsymbol{s}')$, we recall that the total average magnetization is given by:

$$M(\boldsymbol{s}) = \sum_{i=1}^{N} s_i \tag{1.29}$$

Flipping the spin $s_k$ changes the magnetization as follows

$$M(\boldsymbol{s}') = -s_k + \sum_{i\neq k} s_i = -2s_k + \sum_i s_i = M(\boldsymbol{s}) - 2s_k\,, \tag{1.30}$$

where $M(\boldsymbol{s})$ is the total average magnetization of the original state.

**Q10.** **This question is in fact discussed in the next tutorial (TCS T2 – exercise 1, Q5).** The specific heat is given by the derivative of the average energy $\langle E\rangle$ with respect to the temperature $T$ at fixed external field $h$:

$$c = \frac{1}{N}\left(\frac{\partial\langle E\rangle}{\partial T}\right)_h\,. \tag{1.31}$$

The fluctuation-dissipation theorem connects $c$ with the variance of the total energy $E$ of the system (see proof in tutorial 2, exercise 1, Q5):

$$\frac{\langle E^2\rangle - \langle E\rangle^2}{N} = k_B T^2\, c\,. \tag{1.32}$$

Similarly, the susceptibility is given by the derivative of the average total magnetization with respect to the strength of the external magnetic field at fixed value of the temperature $T$:

$$\chi = \frac{1}{N}\left(\frac{\partial\langle M\rangle}{\partial h}\right)_T\,. \tag{1.33}$$

The fluctuation-dissipation theorem connects $\chi$ with the variance of $M$ (see proof in tutorial 2, exercise 1, Q5):

$$\frac{\langle M^2\rangle - \langle M\rangle^2}{N} = k_B T\,\chi\,. \tag{1.34}$$

**Q11.** We consider a system with $n$ spins. In the algorithm, the system initially starts from a random initial state $\boldsymbol{s}$, out-of-equilibrium. During each small time step $dt = 1/n$, we perform the following operations:

- choose a random spin $s_k$ (we will consider the candidate state $\boldsymbol{s}'$ corresponding to flipping the spin $s_k$ in the state $\boldsymbol{s}$);
- using Eq. (1.27), compute the difference in energy $\Delta E$ between the two states $\boldsymbol{s}$ and $\boldsymbol{s}'$;
- if $\Delta E \leq 0$ the new state $\boldsymbol{s}'$ is accepted: we flip $s_k$ and compute the new magnetisation $M(\boldsymbol{s}')$ using Eq. (1.30);
- else (i.e., $\Delta E > 0$) the new state $\boldsymbol{s}'$ is accepted with probability $\Pi(\boldsymbol{s} \to \boldsymbol{s}') = \exp(-\beta\Delta E)$.

It is typical to consider a unit time to be of the order of $n\,dt = 1$, i.e. after we have performed $n$ times the Monte Carlo steps described above. Before we start recording the average magnetization $\langle M\rangle$, we must first wait that the system has reached equilibrium. For instance, we may wait of the order of 5000 unit time before we start recording the magnetization to compute the averages (see additional comment below). We can then start recording the successive values of $M(\boldsymbol{s})$ to compute $\langle M\rangle$ using Eq. (1.3). To avoid too much correlation between the successive values of $M$ that are recorded, one can also record values every 10 unit time.

**Important comments:** Computation of a thermal average only starts when the system reaches equilibrium, namely when $p(s,t) \simeq p_{st}(s)$. Therefore, in a Monte Carlo run, there are generally two parts: The first starting from an initial configuration, where a run is performed in order to lead the system close to equilibrium; the second where the system evolves in the vicinity of equilibrium with the computation of thermodynamic quantities. To estimate the relaxation time towards equilibrium, a naive approach consists of following the evolution of the instantaneous energy of the system and in considering that equilibrium is reached when the energy is stabilized around a quasi-stationary value.

# Chapter 2

# Examples of Equilibrium Critical Phenomena

Questions and exercises indicated with a (⋆) are optional. No worries if you don't have time to try to solve them, or if you don't manage to solve them on your own.

Most of the exercises below are mostly based on the lectures and are here to help you re-derive some of the results we have seen in class. For equilibrium systems, all the information about the properties of the systems are contained in the partition function (denoted $Z$ below). The exercises on the Ising and Potts models thus focus on deriving the partition function.

## 2.1 Ising Model

### 2.1.1 Macroscopique quantities of interest ("Observables")

Consider a system of $N$ spin interacting placed in a uniform external field $H$. The total energy of the system is:

$$E(\vec{s}) = -J \sum_{\langle i,j \rangle} s_i s_j - H \sum_{i=1}^{N} s_i \,, \tag{2.1}$$

where $J$ is a coupling constant and $H$ is the external field. The notation $\langle i, j \rangle$ means that the summation is over pairs of spins $i$ and $j$ that are nearest neighbors. Note that which pairs of spins are considered "neighbors" will depend on the underlying lattice structure of the system, which we haven't yet specified for the moment

**Q1.** The temperature $T$ of the system is set by a thermostat, such that the probability distribution of the microstates of the system is given by the Boltzmann distribution. Can you recall the form of this distribution? What is the partition function $Z(T, H)$?

**Q2.** The free energy $F$ is a good potential thermodynamic for this system. Can you recall it's thermodynamic definition? How can one compute the free energy of this system from its microscopic description?
Because the free energy is an extensive quantities (i.e. it scales linearly with the system size), we introduce the free energy per spin $f = F/N$ (which is an intensive quantity, i.e. it doesn't scale with $N$ anymore).

**Q3.** Interesting quantities for the system are the average total magnetization $\langle M \rangle$ and the average total energy $\langle E \rangle$ of the system. Can you recall their definitions? Observe that the average magnetization and the average energy are both extensive quantities (i.e. they are proportional to the number of spin $N$). We introduce the average total magnetization per spin $m = \langle M \rangle / N$ and the average total energy per spin $\epsilon = \langle E \rangle / N$.

**Q4.** The average magnetization and the average energy can be obtained from the free energy (or the log-partition function)

by derivation:

$$m(T,H) = -\left(\frac{\partial f}{\partial H}\right)_T = \frac{1}{N}\frac{1}{\beta}\left(\frac{\partial \log Z}{\partial H}\right)_T \qquad \text{and} \qquad \epsilon(T,H) = -\frac{1}{N}\left(\frac{\partial \log Z}{\partial \beta}\right)_H . \tag{2.2}$$

Can you prove these relations?

We recall that $f$ is the free energy per spin, and that $\beta = \frac{1}{k_B T}$ is the inverse temperature.

> **Partial derivative:** for a multivariate function $f(x,y)$, one can derive $f$ with respect to any of the two variables $x$ or $y$. To indicate precisely according to which variable we derive $f$, one introduces partial derivatives.
> We denote respectively:
>
> $$\left(\frac{\partial f}{\partial x}\right)_y (x,y) \qquad \text{and} \qquad \left(\frac{\partial f}{\partial y}\right)_x (x,y)\,, \tag{2.3}$$
>
> the partial derivative of $f$ with respect to the variable $x$ and the partial derivative of $f$ with respect to the variable $y$. The first partial derivative indicates that we derive $f(x,y)$ with respect to $x$ while keeping the variable $y$ constant; while the second indicates that we derive $f(x,y)$ with respect to $y$ while keeping the variable $x$ constant.

**Q5.** (⋆)

> **Fluctuation-dissipation theorem.** The susceptibility per spin $\chi(T,H)$ characterizes how sensitive is the magnetisation of the system to small modifications of the external field $H$ at any given temperature. Thus, for any given temperature, the susceptibility corresponds to the slope of the magnetisation $m(T,H)$ as a function of $H$:
>
> $$\chi(T,H) = \left(\frac{\partial m}{\partial H}\right)_T . \tag{2.4}$$
>
> Similarly, the specific heat $c(T,H)$ characterizes how much the average energy per spin $\epsilon(T,H)$ changes under small modification of the temperature $T$. For any given external field, the susceptibility corresponds to the slope of the energy $\epsilon(T,H)$ as a function of $T$:
>
> $$c(T,H) = \left(\frac{\partial \epsilon}{\partial T}\right)_H . \tag{2.5}$$
>
> According to the fluctuation-dissipation theorem, the variance of the total energy is related to the specific heat, and the variance of the magnetisation is related to the susceptibility per spin through:
>
> $$\frac{\langle E^2\rangle - \langle E\rangle^2}{N} = k_B T^2\, c \qquad \text{and} \qquad \frac{\langle M^2\rangle - \langle M\rangle^2}{N} = k_B T\, \chi\,. \tag{2.6}$$

Can you prove these relations?

### 2.1.2 System of non-interacting Ising spins

> Consider a system of $N$ spin non-interacting ($J = 0$) placed in a uniformed external field $H$. Since the spin are non-interacting, the problem is independent of the underlying lattice and of the dimensionality. The result can serve as a cross reference for the behaviour of the Ising model in the weak-coupling limit $\beta J \to 0$ (i.e. $T \to +\infty$). The total energy of the system is:
>
> $$E(\vec{s}) = -H \sum_{i=1}^{N} s_i\,. \tag{2.7}$$

**Q1.** The behavior of the system is determined by the relative strength between the external field $H$ and the thermal energy $k_BT$. Can you comment on the behavior that you expect for the system in the limit $H/k_BT \to 0$ and in the limit $H/k_BT \to \pm\infty$? Which values do you expect for the average magnetisation $\langle M\rangle$? For the average energy $\langle E\rangle$?
Observe that the average magnetisation and the average energy are both extensive quantities (i.e. they are proportional to the number of spin $N$). We introduce the average total magnetisation per spin $m = \langle M\rangle/N$ and the average total energy per spin $\epsilon = \langle E\rangle/N$.

**Q2.** Can you compute the partition function $Z(T,H)$ for this system?

**Q3.** From your result for the partition function, can you compute the free energy for this system, using the statistical formulation of the free energy? Deduce that the expression for the free energy per spin is: $f(T,H) = F(T,H)/N = -k_BT\log(2\cosh(\beta H))$.

**Q4.** Using the relation in Eq. (2.2), can you compute $m(T,H)$ and $e(T,H)$ for this system? Observe that when there is no external field ($H = 0$), there is no spontaneous magnetization of the system ($m = 0$) at any temperature.

**Q5.** Using the relation in Eq. (2.6), can you show that, in presence of an external field ($H \neq 0$), the relative fluctuations of the energy and the relative fluctuations of the magnetization both behave as:

$$\frac{\sqrt{\langle E^2\rangle - \langle E\rangle^2}}{\langle E\rangle} \propto \frac{1}{\sqrt{N}} \qquad \text{and} \qquad \frac{\sqrt{\langle M^2\rangle - \langle M\rangle^2}}{\langle M\rangle} \propto \frac{1}{\sqrt{N}} \tag{2.8}$$

As the system size is increased (larger and larger $N$), the distribution (over many realization of the system) of a quantity of interests (energy and magnetization, here) becomes more and more sharply peaked around its average value. This is a standard behavior for a system with non-interacting components.

### 2.1.3 (⋆) 1-dimensional system with no external field

> Consider a system of $N$ interacting spin placed on a 1-dimensional line. The total energy of the system can be written as:
>
> $$E(\vec{s}) = -J\sum_{i=1}^{N-1} s_i s_{i+1}\,. \tag{2.9}$$
>
> where $J$ is a coupling constant.

**Q1.** Let $\phi$ be is a binary variable that can take values $\pm 1$ and $x$ a real number. Can you show that:

$$\exp(x\phi) = \cosh(x\phi)(1 + \phi\,\tanh x) \tag{2.10}$$

**Q2.** Using this relation, can you show that the partition function can be re-written under the form:

$$Z = (\cosh\beta J)^{N-1}\sum_{\vec{s}}\prod_{i=1}^{N-1}\left[1 + s_i s_{i+1}\tanh(\beta J)\right]? \tag{2.11}$$

**Q3.** Can you show that for any spin $s_i$, its (non-weighted) average value over all the states of the system is zero:

$$\sum_{\vec{s}} s_i = 0\,? \tag{2.12}$$

Can you show that this is also true for any sub-product of spins? i.e., let $\phi(\vec{s})$ be a product of a subset of spins:

$$\sum_{\vec{s}} \phi(\vec{s}) = 0\,? \tag{2.13}$$

**Q4.** Expand the product over $i$ in Eq. (2.11). Using the results of question Q3, can you show that the only term in this expansion that doesn't cancel after summing over the states $\vec{s}$ is the term "1"?

**Q5.** Finally, deduce that, for a one-dimensional Ising system with no external field:

$$Z = 2^N \left(\cosh \beta J\right)^{N-1} . \tag{2.14}$$

**Q6.** How is this result modified if we were to consider periodic boundary conditions? i.e.

$$E(\vec{s}) = -J \sum_{i=1}^{N} s_i s_{i+1} \qquad \text{with } s_{N+1} = s_1 . \tag{2.15}$$

Can you show that the two results are identical in the thermodynamic limit?

## 2.2 (⋆) Potts model

Consider a Potts model with $q$ states ($s_i \in \{1, 2, \cdots, q\}$) on a 1-d lattice with periodic boundary conditions:

$$E(\vec{s}) = -J \sum_{i=1}^{N-1} \delta(s_i, s_{i+1}) - J\delta(s_1, s_N) \tag{2.16}$$

where $\delta(x, y) = 1$ if $x = y$ and $\delta(x, y) = 0$ if $x \neq y$.

**Q1.** Compute the partition function $Z(T)$ of this system using the transfer matrix method.

**Q2.** We place $N$ dots on a circle and we have $q$ colors. From the result of the last question, can you obtain the number of ways to color the $N$ dots such that there is no two consecutive dots with the same color.

## 2.3 Percolation

Try to re-derive the results seen during the lectures.

## 2.4 Short reminder on conditional probabilities

**Independent events.** Two events $A$ and $B$ are independent if the probability of their simultaneous occurrence is the product of the probabilities that each of them occurs: $P(A \cap B) = P(A)\, P(B)$.

**Conditional probability.** The probability of an event $A$ conditional on the occurrence of an event $B$, is defined

$$P(A \,|\, B) \doteq \frac{P(A \cap B)}{P(B)} .$$

This probability can also be read as the probability that an event $A$ occurs, given that another event $B$ has already occurred.
Equivalently, $P(A \cap B) = P(A \,|\, B)\, P(B)$, i.e. that the probability that both $A$ and $B$ occur (*joint probability of A and B*) is the probability that $B$ occurs, times the probability that $A$ occurs given $B$. If $A$ and $B$ are independent then $P(A \,|\, B) = P(A)$.

**Q1.** Consider families with two children $\Omega = \{bb, bg, gb, gg\}$ where the first (second) character stands for the sex (boy or girl) for the elder (younger) child. Imagine that all four possibilities occur with the same probability $P(\omega) = 1/4$.

(a) Given that a family has a boy, what is the probability that the other child is also a boy?

(b) Given that the older child of a family is a boy, what is the probability that the younger is also a boy?

## 2.5 Solutions

### 2.5.1 Ising model: Macroscopic quantities of interest

**Q1.** The probability to find the system in a particular state $\mathbf{s}$ is weighted by the energy of that stated in the for of a Boltzmann factor $\left(e^{-\beta E(\mathbf{s})}\right)$. The corresponding Boltzmann distribution is given by

$$P(\mathbf{s}) = \frac{1}{Z} \exp\left(-\beta E(\mathbf{s})\right) , \tag{2.17}$$

where $\beta = 1/(k_B T)$. The partition function ensures that this distribution is normalized and is consequently given by the sum of the Boltzmann factors over all the system configurations.

$$Z = \sum_{\mathbf{s}} \exp\left(-\beta E(\mathbf{s})\right) . \tag{2.18}$$

**Q2.** The free energy is defined as

$$F = E - TS , \tag{2.19}$$

where $E$ is the internal energy of the system and $S$ is the entropy. An equivalent expression is given by

$$F = -k_B T \log Z = -\frac{1}{\beta} \log Z . \tag{2.20}$$

Since $Z$ contains the sum over all specific configurations, this is how to compute the free energy given the knowledge of its microscopic description in terms of the partition function.

**Q3.** We can find averages of any relevant observable by computing its thermal average

$$\langle A \rangle = \sum_{\mathbf{s}} A(\mathbf{s}) P(\mathbf{s}) = \frac{1}{Z} \sum_{\mathbf{s}} A(\mathbf{s}) \exp\left(-\beta E(\mathbf{s})\right) . \tag{2.21}$$

The total magnetization is given by the sum over all the values of the spins, i.e.

$$M = \sum_{i=1}^{N} s_i . \tag{2.22}$$

Hence,

$$\langle M \rangle = \sum_{\mathbf{s}} M(\mathbf{s}) \exp\left(-\beta E(\mathbf{s})\right) = \sum_{\mathbf{s}} \sum_{i=1}^{N} s_i \exp\left(-\beta E(\mathbf{s})\right)$$
$$\langle E \rangle = \sum_{\mathbf{s}} E(\mathbf{s}) \exp\left(-\beta E(\mathbf{s})\right) .$$

**Q4.** We can find many of the thermal averages by taking appropriate derivatives of the partition function (2.18). For instance, to find the average total energy we take a derivative with respect to $\beta$,

$$\frac{\partial Z}{\partial \beta} = \sum_{\mathbf{s}} (-E(\mathbf{s})) \exp\left(-\beta E(\mathbf{s})\right) = -Z \langle E \rangle . \tag{2.23}$$

This is almost in the form of a thermal average. We can complete the expression by writing

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \equiv -\frac{\partial \log Z}{\partial \beta} \tag{2.24}$$

For the magnetization we can do something similar. Note that the expression for the magnetization occurs in the expression for the system energy (2.1) with the prefactor $H$. Therefore, if we take a derivative of the partition function with respect to $H$ we find

$$\frac{\partial Z}{\partial H} = \sum_{\mathbf{s}} \beta \left[ \sum_{i=1}^{N} s_i \right] \exp\left(-\beta E(\mathbf{s})\right) = \beta Z \langle M \rangle \tag{2.25}$$

and thus we can also write that the average total magnetization is given by

$$\langle M\rangle = \frac{1}{\beta Z}\frac{\partial Z}{\partial H} = \frac{1}{Z}\frac{\partial}{\partial H}\frac{Z}{\beta} = \frac{\partial}{\partial H}\left(\frac{\log Z}{\beta}\right) \equiv -\frac{\partial}{\partial H}F \tag{2.26}$$

Having defined the magnetization and energy per spin,

$$m \equiv \langle M\rangle/N\,, \quad \epsilon \equiv \langle E\rangle/N\,, \tag{2.27}$$

We can divide equations (2.24), (2.5.1) with N and derive the desired expressions:

$$m = -\frac{\partial f}{\partial H}\,, \quad \epsilon = -\frac{1}{N}\left(\frac{\partial \log Z}{\partial \beta}\right) \tag{2.28}$$

**Q5.**

- For the specific heat: We showed that

$$\epsilon = -\frac{1}{N}\left(\frac{\partial \log Z}{\partial \beta}\right), \tag{2.29}$$

but we now want to link it with $c$ which involves a derivative with respect to T. It is useful to use the chain rule to switch to a derivative with respect to $\beta$

$$\beta \equiv \frac{1}{k_B T} \Longrightarrow \mathrm{d}\beta = -\frac{1}{k_B T^2}\,\mathrm{d}T \Longrightarrow \frac{\partial}{\partial \beta} = -k_B T^2\frac{\partial}{\partial T} \Longrightarrow \frac{\partial}{\partial T} = -k_B\beta^2\frac{\partial}{\partial \beta} \tag{2.30}$$

Using this relation we find that

$$c \equiv \frac{\partial \epsilon}{\partial T} = \frac{k_B\beta^2}{N}\frac{\partial^2 \log Z}{\partial \beta^2} = \frac{k_B\beta^2}{N}\frac{\partial}{\partial \beta}\left[\frac{1}{Z}\frac{\partial Z}{\partial \beta}\right] = \frac{k_B\beta^2}{N}\left[\frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} - \left(\frac{1}{Z}\frac{\partial Z}{\partial \beta}\right)^2\right] \tag{2.31}$$

The last term is equal to the square of $\langle E\rangle$ (see eq. (2.24)). The second derivative of the partition function with respect to $\beta$ is given by

$$\frac{\partial^2 Z}{\partial \beta^2} = \frac{\partial}{\partial \beta}\left[\sum_{\mathbf{s}}\left(-E(\mathbf{s})\right)\exp\left(-\beta E(\mathbf{s})\right)\right] = \sum_{\mathbf{s}}\left[-E(\mathbf{s})\right]^2\exp\left(-\beta E(\mathbf{s})\right) = Z\langle E^2\rangle \tag{2.32}$$

Upon substitution of both these relations we obtain the relation between the specific heat and the variance of the energy

$$c = k_B\beta^2\left[\frac{\langle E^2\rangle - \langle E\rangle^2}{N}\right] = \frac{1}{k_B T^2}\left[\frac{\langle E^2\rangle - \langle E\rangle^2}{N}\right] \tag{2.33}$$

- For the magnetization we do the same trick

$$\chi = \frac{\partial m}{\partial H} = -\frac{\partial^2 f}{\partial H^2} = \frac{1}{\beta N}\frac{\partial^2}{\partial H^2}\log Z\,, \tag{2.34}$$

and the proof follows in the same way as for the specific heat.

### 2.5.2 System of non-interacting Ising spins

**Q1.** The probability of finding the system in a state **s** is given by

$$P(\mathbf{s}) = \frac{1}{Z}\exp\left(\beta H\sum_{i=1}^{N} s_i\right) = \frac{1}{Z}\exp(\beta H M)\,.$$

- In the limit $H/k_BT = \beta H \to 0$ , the thermal energy "wins" over the external field. Then we have

$$P(\mathbf{s}) = \frac{1}{Z}\,. \tag{2.35}$$

So each state becomes equally likely, which means that the most probable microstates consist of spins randomly pointing up and down. Therefore we expect a vanishing magnetization $\langle M\rangle = 0$ and zero energy $\langle E\rangle = 0$.

- In the limit $\beta H \to \pm\infty$, the external field "wins" over the thermal energy: The most probable microstates consists of spins that are aligned in the same direction as the external magnetic field $H$, corresponding to a minimum disorder. The behaviour depends on the sign of the external field. If $H > 0$, as $\beta H$ goes to (plus) infinity, the state with all spins pointing up becomes the lowest energy state, to such an extent as to completely dominate the probability distribution. We thus have

$$\langle M\rangle = N, \quad \langle E\rangle = -HN, \quad \beta H \to \infty \tag{2.36}$$

**NOTE:** The way to derive $\langle E\rangle$ is to notice that in eq. (2.1), when $|H| \to \infty$, the second term dominates, so we can safely neglect the first one.

The minus infinity (i.e. $H < 0$) case is the same but with all the spins pointing down. Thus

$$\langle M\rangle = -N, \quad \langle E\rangle = -HN, \quad \beta H \to -\infty\,. \tag{2.37}$$

**Q2.** The partition function is sum of the Boltzmann factors over all configurations of the system. So explicitly we have

$$Z = \sum_{\mathbf{s}} \exp\left(\beta H \sum_{i=1}^{N} s_i\right) = \sum_{\mathbf{s}} \prod_{i=1}^{N} \exp(\beta H s_i) = \underbrace{\sum_{s_1} \exp(\beta H s_1)}_{Z_1} \sum_{s_2} \exp(\beta H s_2) \cdots \sum_{s_N} \exp(\beta H s_N) = (Z_1)^N\,. \tag{2.38}$$

Note that because the spins are all independent, the partition function factorizes into contributions from each individual spin. Each spin $s_i$ can take the value $\pm 1$, so each factor is just

$$Z_1 = \sum_{s_i=\pm 1} \exp(\beta H s_i) = 2\cosh(\beta H)\,. \qquad \text{Reminder: } \cosh x = \frac{e^x + e^{-x}}{2} \tag{2.39}$$

So, for a system with N spins, the partition function is

$$Z = (Z_1)^N = 2^N[\cosh(\beta H)]^N\,. \tag{2.40}$$

**Q3.** Since we know that $F = -k_BT \log Z$ we find that

$$F = -k_BT \log Z = -Nk_BT \log[2\cosh(\beta H)]\,, \tag{2.41}$$

and therefore

$$f = \frac{F}{N} = -k_BT \log[2\cosh(\beta H)]\,. \tag{2.42}$$

**Q4.** The average magnetization per site follows directly from the free energy expression

$$m = -\frac{df}{dH} = k_BT \frac{1}{2\cosh(\beta H)} 2\beta \sinh(\beta H) = \tanh(\beta H)\,. \tag{2.43}$$

Likewise, the average energy is given by

$$\varepsilon = -\frac{1}{N}\left(\frac{\partial \log Z}{\partial \beta}\right) = -\frac{\partial}{\partial \beta} \log[2\cosh(\beta H)] = -\frac{1}{2\cosh(\beta H)} 2H \sinh(\beta H) = -H\tanh(\beta H)\,. \tag{2.44}$$

**Q5.** For the energy, we use the relation

$$\langle E^2\rangle - \langle E\rangle^2 = Nk_BT^2c \sim N, \qquad \langle E\rangle = \varepsilon N = -HN\tanh(\beta H) \sim N\ . \tag{2.45}$$

Since $c$ is related to $\varepsilon$ $\left(c \equiv \frac{\partial\varepsilon}{\partial T}\right)$, and $\varepsilon$ does not depend on $N$, then $c$ does not depend on N too. Thus,

$$\frac{\sqrt{\langle E^2\rangle - \langle E\rangle^2}}{\langle E\rangle} \sim \frac{\sqrt{N}}{N} \sim \frac{1}{\sqrt{N}}. \tag{2.46}$$

The same argument holds for the fluctuations of the magnetization.

### 2.5.3 One-dimensional Ising system with no external field

**Q1.** Note first that

$$\cosh(x) = \frac{e^x + e^{-x}}{2}, \quad \sinh(x) = \frac{e^x - e^{-x}}{2}. \tag{2.47}$$

From this follows that

$$\exp(x\phi) = \cosh(x\phi) + \sinh(x\phi). \tag{2.48}$$

Now remember that $\cosh(x)$ is an even function and $\sinh(x)$ is an odd function of $x$. This we know that

$$\cosh(x) = \cosh(-x), \quad \sinh(x) = -\sinh(-x). \tag{2.49}$$

Since $\phi$ can only be $\pm1$ we can use the above relations to write

$$\cosh(\phi x) = \cosh(x), \quad \sinh(\phi x) = \phi\sinh(x). \tag{2.50}$$

With these relations we can write that

$$\exp(x\phi) = \cosh(x) + \phi\sinh(x) = \cosh(x)\left(1 + \phi\frac{\sinh(x)}{\cosh(x)}\right) = \cosh(x)(1 + \phi\tanh(x)). \tag{2.51}$$

**Q2.** Starting from the formal definition of the partition function we have

$$Z = \sum_{\mathbf{s}} \exp\left(\beta J\sum_{i=1}^{N-1} s_is_{i+1}\right). \tag{2.52}$$

We identify $\phi_i = s_is_{i+1} = \pm1$ and write

$$Z = \sum_{\mathbf{s}}\prod_{i=1}^{N-1}\exp(\beta J\phi_i). \tag{2.53}$$

Since $\phi_i$ is a variable that takes binary values $\pm1$, we can use the identity from before and write

$$Z = \sum_{\mathbf{s}}\prod_{i=1}^{N-1}\cosh(\beta J)\left[1 + \phi_i\tanh(\beta J)\right]. \tag{2.54}$$

Since $\cosh(\beta J)$ does not depend on $s_i$ we can take it out of the product and obtain

$$Z = \cosh(\beta J)^{N-1}\sum_{\mathbf{s}}\prod_{i=1}^{N-1}\left[1 + \phi_i\tanh(\beta J)\right]. \tag{2.55}$$

**Q3.** Consider a state $\mathbf{s} = (s_1, s_2, \ldots, s_i, \ldots, s_N)$.
For all such states, there exists another state of the form $\mathbf{s}' = (s_1, s_2, \ldots, -s_i, \ldots, s_N)$ where $s_i$ is flipped. Since there are $2^N$ states, there are $2^{N-1}$ states with $s_i = 1$ and $2^{N-1}$ states with $s_i = -1$. Thus we can write

$$\sum_{\mathbf{s}} s_i = 2^{N-1} - 2^{N-1} = 0. \tag{2.56}$$

A product oif spins can always be written as $\phi(\mathbf{s}) = s_i \times \phi'(\mathbf{s})$ where $\phi'(\mathbf{s})$ is the remainder of the product without $s_i$, and so

$$\sum_{\mathbf{s}} \phi(\mathbf{s}) = \sum_{\mathbf{s}} s_i \times \phi'(\mathbf{s}) = \sum_{s_i=\pm 1} s_i \sum_{\mathbf{s}'} \phi'(\mathbf{s}) = \sum_{\mathbf{s}'} \phi'(\mathbf{s}) - \sum_{\mathbf{s}'} \phi'(\mathbf{s}) = 0. \tag{2.57}$$

**Q4.** The product has the form

$$\begin{aligned}\sum_{\mathbf{s}} (1+\phi_1)(1+\phi_2)\ldots(1+\phi_N) &= \sum_{\mathbf{s}} (1 + \phi_1 + \phi_2 + \cdots + \phi_1\phi_2 + \phi_1\phi_3 + \cdots + \phi_1\phi_2\ldots\phi_N) \\ &= \sum_{\mathbf{s}} 1 + \sum_{\mathbf{s}} \phi_1 + \cdots + \sum_{\mathbf{s}} \phi_1\phi_2\ldots\phi_N = \sum_{\mathbf{s}} 1 = 2^N\end{aligned}$$

since we just showed that single-spin and product-operators vanish when summed over the possible configurations of the system. Since there are $2^N$ possible configurations for $N$ binary spins, the remaining sum is just the number of configurations.

**Q5.** We have all the ingedients necessary to find that

$$Z = 2^N \cosh(\beta J)^{N-1}. \tag{2.58}$$

**Q6.** A first guess for the partition function would be $Z_{WRONG} = 2^N \cosh(\beta J)^N$, but this is just the partition function for a line of strings of length $N+1$ instead of $N$ and doesn't take the periodic boundary conditions into account. To include these, consider again the product

$$\sum_{\mathbf{s}} \prod_{i=1}^{N} [1 + s_i s_{i+1} \tanh(\beta J)]. \tag{2.59}$$

Since each of the spins $s_i$ can only take values $\pm 1$, a square $(s_i)^2 = 1$ necessarily. Now note that because of the periodic boundary conditions, there is a term in the product expansion of the form

$$\sum_{\mathbf{s}} \prod_{i=1}^{N} [1 + s_i s_{i+1} \tanh(\beta J)] = 1 + \cdots + (s_1 s_2 \times s_2 s_3 \times \cdots \times s_{N-1} s_N \times s_N s_1) \tanh(\beta J)^N.$$

In the final term each spin $s_i$ occurs as a square and so the product evaluates to 1 and therefore

$$\sum_{\mathbf{s}} \prod_{i=1}^{N} [1 + s_i s_{i+1} \tanh(\beta J)] = (1 + \tanh(\beta J)^N) 2^N, \tag{2.60}$$

and so the partition function for the periodic system is given by

$$Z_{\text{periodic}} = 2^N \cosh(\beta J)^N \left(1 + \tanh(\beta J)^N\right) = 2^N \left(\cosh(\beta J)^N + \sinh(\beta J)^N\right). \tag{2.61}$$

**Q7.** We can use the same expansion to derive the average magnetization $m = \langle s_i \rangle$ and the spin-spin correlation function $\langle s_i s_j \rangle$. We start with the average magnatization which we write formally as

$$\langle s_j \rangle = \frac{1}{Z} \sum_{\mathbf{s}} s_j \exp\left(\beta J \sum_{i=1}^{N-1} s_i s_{i+1}\right). \tag{2.62}$$

Note that we use the label $j$ to distinguish the expectation of the spin from the spin's index in the sum. We use the same tricks as above to write this as

$$\langle s_j \rangle = \frac{1}{Z} \sum_{\mathbf{s}} s_j \prod_{i=1}^{N-1} \cosh(\beta J) \left[1 + s_i s_{i+1} \tanh(\beta J)\right]. \tag{2.63}$$

Since the hyperbolic cosine function does not depend on the state of the system, we can take it out of both the sum and the product and write

$$\langle s_j \rangle = \frac{\cosh(\beta J)^{N-1}}{Z} \sum_{\mathbf{s}} s_j \prod_{i=1}^{N-1} \left[1 + s_i s_{i+1} \tanh(\beta J)\right]. \tag{2.64}$$

Let $\phi_i \equiv s_i s_{i+1}$ and define $t = \tanh(\beta J)$. Then we can write the product as

$$s_j \prod_{i=1}^{N-1} [1 + \phi_i t] = s_j \left[1 + \phi_1 t + \phi_2 t + \cdots + \phi_1 \phi_2 t^2 + \cdots + \phi_1 \cdots \phi_{N-1} t^{N-1}\right]. \tag{2.65}$$

Note that, due to the nearest-neighbour lattice, no spin can occur more than twice in any of the products of $\phi$ operators (e.g. $\phi_1 \phi_2 = s_1 s_2 s_2 s_3 = s_1 s_3$). Thus, all of these products consist of a multiple of two spins, all of which are distinct. Therefore, multiplying by any $s_j$ will result in terms that involve $2n+1$ different spins for $n = 0, 1, 2, \ldots$. If we sum these over all states we find

$$\langle s_j \rangle = \frac{\cosh(\beta J)^{N-1}}{Z} \sum_{\mathbf{s}} \left[s_j + s_j \phi_1 t + \cdots + s_j \phi_1 \cdots \phi_{N-1} t^{N-1}\right] = 0. \tag{2.66}$$

For the spin-spin correlation function we have to evaluate the product

$$s_j s_k \prod_{i=1}^{N-1} [1 + \phi_i t]. \tag{2.67}$$

Note that we can express $s_j s_k$ as the following product of the $\phi$ ( assuming $j < k$ without loss of generality)

$$\phi_j \phi_{j+1} \ldots \phi_{k-2} \phi_{k-1} = s_j (s_{j+1})^2 (s_{j+2})^2 \ldots (s_{k-2})^2 (s_{k-1})^2 s_k = s_j s_k. \tag{2.68}$$

The number of operators involved in constructing $s_j s_k$ like this is directly related to the distance between the spins $|k - j|$. If we call this distance $r$, then the term in the product expansion comes with a factor of $\tanh(\beta J)^r$, i.e. there is a single term in the product expansion of the form

$$s_j s_{j+r} \prod_{i=1}^{N-1} [1 + s_i s_{i+1} \tanh(\beta J)] = s_j s_{j+r} + \cdots + \tanh(\beta J)^r + \ldots \tag{2.69}$$

Summing over all the spins, the hyperbolic tangent term is the only surviving term and we obtain

$$\langle s_j s_{j+1}\rangle = \frac{\cosh(\beta J)^{N-1}}{Z}\sum_{\mathbf{s}}\left[s_j s_{j+r} + \cdots + \tanh(\beta J)^r + \ldots\right] = \frac{2^N\cosh(\beta J)^{N-1}}{2^N\cosh(\beta J)^{N-1}}\tanh(\beta J)^r = \tanh(\beta J)^r. \tag{2.70}$$

We now define the correlation length $\xi$ and write

$$e^{-r/\xi} = \tanh(\beta J)^r \Rightarrow \xi = -\frac{1}{\log(\tanh(\beta J))}. \tag{2.71}$$

### 2.5.4 Potts model

In the transfer matrix method we define the transfer matrix as a matrix with the appropriate Boltzmann factor in each of its entries. In this case, the diagonal and off-diagonal elements are given by

$$(T_{s_i,s_{i+1}})_{ii} = \exp(\beta J)\,, \quad (T_{s_i,s_{i+1}})_{ij} = 1\,, j \neq i\,. \tag{2.72}$$

Consequently, the transfer matrix is a $q\times q$ matrix of the form

$$T_{s_i,s_{i+1}} = \exp(\beta J)\,\mathbb{I}_q + \left(\mathbf{1}_{q\times q} - \mathbb{I}_q\right) = \left[\exp(\beta J) - 1\right]\mathbb{I}_q + \mathbf{1}_{q\times q}\,, \tag{2.73}$$

where $\mathbb{I}_q$ is the identity matrix and $\mathbf{1}_{q\times q}$ is a matrix filled with ones.

Let $a \equiv \exp(\beta J)$. To solve for the eigenvalues we write

$$\left[(a-1)\mathbb{I}_q + \mathbf{1}_{q\times q}\right]\mathbf{v} = \lambda\mathbf{v} \tag{2.74}$$

Note that since $\mathbf{1}_{q\times q}$ applied to any vector, results to a vector with the sum of its components as entries, a vector with an equal number of $+1$ and $-1$ returns the zero vector. Denoting these vectors as $\mathbf{u}_\pm$ we find that

$$\left[(a-1)\mathbb{I} + \mathbf{1}_{q\times q}\right]\mathbf{u}_\pm = (a-1)\mathbf{u}_\pm \tag{2.75}$$

And so $a-1$ is an eigenvalue. The eigenvectors of this type are generated by a basis of $q-1$ eigenvectors which all have the same eigenvalue (check!).

Also note that if we have a vector with all ones, which we denote $\mathbf{u}_1 = (1,1,\ldots,1)$ then $\mathbf{1}_{q\times q}\mathbf{u}_1 = q\mathbf{u}_1$. So in that case

$$\left[(a-1)\mathbb{I} + \mathbf{1}_{q\times q}\right]\mathbf{u}_1 = (a-1+q)\mathbf{u}_1\,, \tag{2.76}$$

thus $a-1+q$ are also eigenvalues ($q$ of those). For $q=0$ we recover the previous eigenvalue. We have thus found all the eigenvalues of the $q\times q$ transfer matrix.

The eigenvectors of the transfer matrix are the same as those of $\mathbf{1}_{q\times q}$. The partition function is given by the trace of the transfer matrix to the $N-th$ power and therefore

$$Z = \mathrm{Tr}\left[T^N\right] = \sum_{i=1}^{q}(\lambda_i)^N = \left(e^{\beta J} - 1 + q\right)^N + (q-1)\left(e^{\beta J} - 1\right)^N\,. \tag{2.77}$$

**Q2.** This translates to taking the limit $J\to-\infty$. Then, a state with two consecutive dots with the same color, would cost an infinite amount of energy to construct. In that case, only states with no consecutive dots contribute to the partition function (with contribution $e^0 = 1$) and so the number of ways, $n$, of constructing such a state is equal to the partition evaluated at $J\to-\infty$

$$n = \lim_{J\to-\infty} Z = (q-1)^N + (-1)^N(q-1)\,. \tag{2.78}$$

# Chapter 3

# Mean-field approximation

**Symbol "(⋆)":** Questions and exercises indicated with a (⋆) are optional. No worries if you don't have time to try to solve them, or if you don't manage to solve them on your own.

**The symbol "</>":** indicates optional questions with numerical simulation.

> **Toolbox.** We recall the following Taylor expansions:
>
> $$\tanh(x) = x - \frac{x^3}{3} + \frac{2x^5}{15} + o(x^5)\,, \qquad \text{for } x \text{ close to } 0 \tag{3.1}$$
>
> $$\cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + o(x^4)\,, \qquad \text{for } x \text{ close to } 0 \tag{3.2}$$
>
> $$\frac{1}{1-x^2} = 1 + x^2 + x^4 + o(x^4)\,, \qquad \text{for } x \text{ close to } 0 \tag{3.3}$$
>
> We recall the derivative of the hyperbolic tangent:
>
> $$[\tanh(x)]' = \frac{1}{\cosh^2(x)} = 1 - \tanh^2(x)\,. \tag{3.4}$$

## 3.1 Critical exponents for the Mean-Field Ising model

In the mean-field approximation, the avegare magnetisation per site is given by the self-consistency equation:

$$m = \tanh(\beta(H + Jmq)) \tag{3.5}$$

where $H$ is the external magnetic field, $J$ is the coupling parameter, and $q$ is the number of nearest neighbors. In absence of any external field, $H = 0$, finding the solution of this equation graphically gives the critical temperature $T_c = qJ/k_B$ (see lecture). For clarity, we will denote $m_0$ the magnetization per spin when $H = 0$, i.e. $m_0(T) = m(H = 0, T)$.

**Q1.** In absence of any external field ($H = 0$), study the behavior of $m_0$ close to the critical point ($T \to T_c$) and show that:

$$\text{for } T < T_c\,, \qquad m_0 \sim a\,(T_c - T)^{\beta}\,, \text{ when } T \to T_c^-\,, \tag{3.6}$$

where $\beta$ is an exponent to specify and $a$ is a constant coefficient. What is the critical exponent $\beta$? What is the multiplicative coefficient $a$?

**Q2.** Similarly, can you show that, near the critical point, the magnetic susceptibility behaves as:

$$\chi \sim |T - T_c|^{-\gamma}, \tag{3.7}$$

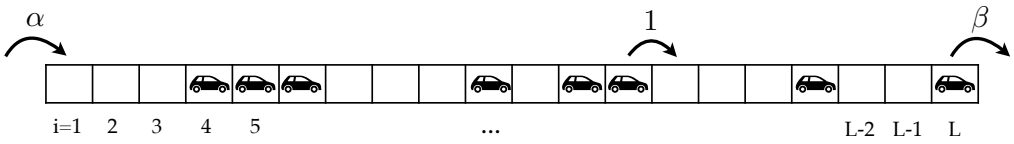
for $T < T_c$, as well as for $T > T_c$.

**Q3. (⋆)** Similarly, can you find out what are the critical exponents for: a) the heat capacity close to the critical point $T_c$ for the phase transition at $H = 0$; and b) the behavior of the magnetization at $T = T_c$ for small values of external field $H$:

$$c \sim |T - T_c|^{-\alpha}, \qquad \text{and} \qquad m \sim H^{1/\delta}. \tag{3.8}$$

## 3.2 Back to TASEP!

We recall the TASEP model for traffic that we have discussed in the first tutorial T0.




Cars evolve on a one-dimensional lattice with $L$ sites opened on both sides (which represents a portion of road). Cars randomly hop into the lattice from the left with rate $\alpha$, jump along the lattice to the right with rate 1, and exit the lattice from the right with rate $\beta$ (see figure). Each site can only be occupied by one car at a time, and a car can hop to the next site only if it is empty (exclusion interaction).

We denote by $\rho_i(t)$ the probability that there is a car at the site $i$ at time $t$: $\rho_i(t) = P[n_i(t) = 1]$. The probability that there is no car at site $i$ is then $(1 - \rho_i)$ (as a site can either be occupied or not). We define $J_i(t) = P[n_i(t) = 1; n_{i+1}(t) = 0]$ as the current of cars that exit site $i$ to the right at time $t$.

We recall that the local density of cars $\rho_i(t)$ follows the equation of evolution:

$$\frac{d\rho_i}{dt}(t) = J_{i-1}(t) - J_i(t), \text{ for all the sites } i \in \{1, \cdots, L\}, \tag{3.9}$$

with the current at the left and right boundaries respectively given by:

$$\begin{cases} J_0(t) &= \alpha\,(1 - \rho_1(t)), \\ J_L(t) &= \beta\,\rho_L(t). \end{cases} \tag{3.10}$$

**Q1.** In the stationary state, the local densities and currents become time-independent. Deduce from the previous equations that the current is uniform in the stationary state, i.e. for all $i$, $J_i = J$ is a constant. Can you give the relation between $\rho_1$ and $J$, and between $\rho_L$ and $J$?

**Q2.** We are interested in studying the stationary state and understand if the system can be in different phases. Using a mean-field approximation, can you obtain a recurrence relation between $\rho_i$ and $\rho_{i+1}$ in the stationary state?

**Q3.** Can you find what are the fixed points of the recurrence relation?

**Q4.** Studying the stability of the fixed points: Depending on the values of $\alpha$ and $\beta$: which fixed points are stable?

**Q5.** From this analysis, can you obtain the phase diagram in $(\alpha, \beta)$ for the TASEP model?

## 3.3 Ising model with long-range interactions; Mean-field Ising and Landau theory

In this exercise, we are interested in showing that the mean-field Ising model becomes exact, in the case of an Ising model in which all the spins interact identically (not just the nearest neighbors).

Let us consider $N$ spins ($N \gg 1$). There are $N(N-1)/2$ pairs of spins. We take the coupling parameter to be inversely proportional to $N$, so that the coupling energy remains proportional to $N$. The energy of the system take the form:

$$E(\vec{s}) = -\frac{J}{N} \sum_{pair(i,j)} s_i s_j \,, \tag{3.11}$$

in which the sum is over all possible pairs of spins (and not just the nearest neighbors), and $J$ is a coupling constant.

**Q1.** Can you show that the energy can be re-written under the form:

$$E(\vec{s}) = -\frac{J}{2N} \left[ \left( \sum_{i=1}^{N} s_i \right)^2 - N \right] \; ? \tag{3.12}$$

**Q2.** We denote $M = \sum_{i=1}^{N} s_i$ the total magnetization. We recall that $-N \leq M \leq N$. Can you show that the number of states with the same magnetization $M$ is given by:

$$\Omega(M) = \frac{N!}{\left(\frac{N+M}{2}\right)! \, \left(\frac{N-M}{2}\right)!} \tag{3.13}$$

**Q3.** Can you show that the partition function $Z$ can be re-written under the form:

$$Z = \sum_{M} \Omega(M) \exp\left(-\beta \, E(M)\right) \,, \tag{3.14}$$

where $E(M)$ is the energy of a state with magnetization $M$. We introduce $Z(M,T) = \Omega(M) \exp\left(-\beta \, E(M)\right)$ such that:

$$Z = \sum_{M} Z(M,T) \,, \tag{3.15}$$

**Q4.** Using the Stirling formula[1] for $N!$ for large $N$, can you show that the function $F(M,T) = -k_B T \ln Z(M)$ can be written under the form:

$$F(M,T) = -\frac{JM^2}{2N} + \frac{J}{2} + k_B T \left( N_+ \ln N_+ + N_- \ln N_- - N \ln N \right) \,, \tag{3.16}$$

where $N_+ = (N+M)/2$ and $N_- = (N-M)/2$. In Eq. (3.15), $Z$ can be approximated by the largest term in the sum, which is the term $Z(M_0, T)$ for which $M_0$ minimizes $F(M,T)$.

**Q5.** Let us define $m = M/N$ the magnetization per spin, and $f(m,T) = F(m,T)/N$. Using Eq. (3.16), can you give the expression of $f(m,T)$ as a function of $m$? Can you show that the values of $m$ that minimize $f(m,T)$ satisfy a self-consistency equation that is similar to the characteristic equation of the mean-field Ising model? Using that $\tanh^{-1}(x) = \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right)$ will help recover the equation of the mean-field Ising model. Can you show that the system has a phase transition and compute

[1] Stirling formula: $\ln(N!) \simeq N \ln N - N$ as $N$ goes to $+\infty$.

the critical temperature $T_c$?

**Q6.** Show that for $m$ close to 0, $f(m,T)$ has the form expected by the Landau theory, i.e.:

$$f(m,T) = f_0 + \frac{1}{2}a_2(T)m^2 + \frac{1}{4}a_4(T)m^4\,, \tag{3.17}$$

where $a_2(T_c) = 0$. Can you give the expressions of $a_2(T)$ and $a_4(T)$?

## 3.1 Solutions

### 3.1.1 Critical exponents for the Mean-Field Ising model

**A1.** To find out the behavior of $m$ close to the critical point, we must solve the self-consistency equation Eq. (3.5) with $H = 0$ for $T$ close to $T_c$. We know already that for all temperature $T > T_c$, there is only one solution that is $m_0 = 0$. We are therefore interested in what happens in the region $T < T_c$. In general, it is not easy to solve Eq. (3.5), even with $H = 0$. However, for temperature smaller than the critical point but close to the critical point, we expect the magnetization per spin $m_0$ to be very small, i.e. to be close to zero. This is because the value of $m$ evolves continuously through the phase transition at $T_c$ (continuous phase transition) and we know that for $T > T_c$ the magnetization per spin is zero $m_0 = 0$. Thanks to this, we can make an expansion of the non-linear function on the right-hand-side of Eq. (3.5) for small values of $m_0$.

We start from Eq. (3.5), in which we take $H = 0$ and replace the expression of $T_c = qJ/k_B$:

$$m_0 = \tanh\left(\frac{T_c}{T}\, m_0\right) \tag{3.18}$$

We then expend the hyperbolic tangent around $m_0 = 0$ (see Taylor expansion in Eq. (3.1)), which gives:

$$m_0 = \frac{T_c}{T}\, m_0 - \left(\frac{T_c}{T}\right)^3 \frac{m_0^3}{3} + o(m_0^3)\,. \tag{3.19}$$

Note that an expansion to the 3 order is sufficient to extract a value for $m$. An expansion to the first order would give us that "$T = T_c$" to first order, which is not very useful. This then yields a quadratic equation for $m_0$ of the form

$$\left(\frac{T_c - T}{T}\right) = \left(\frac{T_c}{T}\right)^3 \frac{m_0^2}{3} + o(m_0^2) \tag{3.20}$$

We can check that the terms on each part of the equality are both positive (as $T < T_c$). We can then take the square root on both sides and get two solutions, for positive and negative $m_0$ close to $T_c$:

$$m_0 \sim \pm\sqrt{3}\left(\frac{T}{T_c}\right)^{3/2}\left(\frac{T_c - T}{T}\right)^{1/2}, \qquad \text{for } T < T_c\,. \tag{3.21}$$

In the limit $T \to T_c$, the fraction $\frac{T_c}{T} \to 1$, so we obtain:

$$m_0 \sim \pm\sqrt{3}\left(\frac{T_c - T}{T_c}\right)^{1/2}, \quad \text{for } T \to T_c^-\,. \tag{3.22}$$

From the above, we read off that $\beta = 1/2$, and $a = \sqrt{3/T_c}$.

To summarize, close to the critical point, $T \to T_c$, we have that:

$$m_0 \sim \begin{cases} \pm\sqrt{\dfrac{3}{T_c}}\,(T_c - T)^{1/2}\,, & \text{for } T < T_c \\ 0\,, & \text{for } T > T_c \end{cases} \tag{3.23}$$

**A2.** The magnetic susceptibility is given by $\chi = \lim_{H\to 0}\left(\frac{\partial m}{\partial H}\right)_T$.

We start from Eq. (3.5) in which we replace the expression of $T_c = qJ/k_B$:

$$m = \tanh\left(\frac{T_c}{T}\, m + \frac{H}{k_B T}\right) \tag{3.24}$$

**Version 1:** Deriving both side of this equation according to $H$, we obtain:

$$\left(\frac{\partial m}{\partial H}\right)_T = \frac{1}{\cosh^2\left(\frac{T_c}{T}\, m + \frac{H}{k_B T}\right)}\left(\frac{T_c}{T}\left(\frac{\partial m}{\partial H}\right)_T + \frac{1}{k_B T}\right) \tag{3.25}$$

Here we used the first expression in Eq. (3.4) for the derivative of $\tanh(x)$. Note that we also had to keep the derivative of $m$ with respect to $H$ for the derivation of the term inside the hyperbolic tangent. Taking the limit $H \longrightarrow 0$ leads to

$$\chi = \frac{1}{\cosh^2\left(\frac{T_c}{T}\, m_0\right)}\left(\frac{T_c}{T}\,\chi + \frac{1}{k_B T}\right), \tag{3.26}$$

where we used the definition of $\chi$ and the relabeling $m_0(T) = m(H = 0, T)$. Re-organising the terms to extract $\chi$ from this equation gives:

$$\chi = \frac{1}{k_B T}\,\frac{1}{\cosh^2\left(\frac{T_c}{T}\, m_0\right) - \frac{T_c}{T}} \tag{3.27}$$

• **For** $T > T_c$, we have $m_0 = 0$, which gives (we recall that $\cosh(0) = 1$):

$$\chi = \frac{1}{k_B T}\,\frac{1}{1 - \frac{T_c}{T}} = \frac{1}{k_B}(T - T_c)^{-1}\,. \tag{3.28}$$

Therefore, we get the critical exponent $\gamma = 1$ for $T > T_c$.

• **For** $T < T_c$ **and** $T$ **close to** $T_c$, the value of $m_0$ is close to 0, and its behavior is given by Eq. (3.22). We can expand $\cosh(T_c m_0/T)$ for small values of $m_0$ in Eq. (3.27):

$$\cosh^2\left(\frac{T_c}{T}\, m_0\right) = \left[1 + \frac{1}{2}\left(\frac{T_c}{T}\, m_0\right)^2 + o(m_0^2)\right]^2 \tag{3.29}$$

$$= \left[1 + 2\,\frac{1}{2}\left(\frac{T_c}{T}\, m_0\right)^2 + o(m_0^2)\right] = \left[1 + \left(\frac{T_c}{T}\, m_0\right)^2 + o(m_0^2)\right] \tag{3.30}$$

in which we neglected all the terms of order higher than $m_0^2$ in $m_0$. Replacing this expansion in Eq. (3.27), we get:

$$\chi = \frac{1}{k_B T}\,\frac{1}{1 + \left(\frac{T_c}{T}\right)^2 m_0^2 - \frac{T_c}{T}} \tag{3.31}$$

In which we can replace $m_0(T)$ by its behavior close to the critical temperature Eq. (3.22), $m_0^2 \sim 3\,(T_c - T)/T_c$:

$$\chi = \frac{1}{k_B T}\,\frac{1}{1 + \left(\frac{T_c}{T}\right)^2 3\left(\frac{T_c - T}{T_c}\right) - \frac{T_c}{T}} = \frac{1}{k_B}\,\frac{1}{T + \left(\frac{T_c}{T}\right) 3\,(T_c - T) - T_c} \tag{3.32}$$

$$\chi = \frac{1}{k_B}\,\frac{1}{\left(\frac{T_c}{T}\right) 3\,(T_c - T) - (T_c - T)} \tag{3.33}$$

In the limit $T \longrightarrow T_c^-$, one finally gets:

$$\chi = \frac{1}{k_B}\,\frac{1}{2\,(T_c - T)} = \frac{1}{2k_B}(T_c - T)^{-1} \tag{3.34}$$

Therefore, we also get the critical exponent $\gamma = 1$ for $T < T_c$. This is the same exponent than for $T > T_c$, only the coefficient differ: $1/(2k_B)$ instead of $1/k_B$.

To summarize, close to the critical point, $T \to T_c$, we have that:

$$\chi \sim \begin{cases} \dfrac{1}{2k_B}\,(T_c - T)^{-1}, & \text{for } T < T_c \\ \dfrac{1}{k_B}\,(T - T_c)^{-1}, & \text{for } T > T_c \end{cases} \tag{3.35}$$

**Version 2:** Note, one can obtain these results "faster" by using another expression for the derivative of $\tanh(x)$ (second expression in Eq. (3.4)). This is much simpler, as it doesn't require to take any additional expansion in small $m$, because the non-linear part, $\tanh(\cdots)$, is directly equal to $m$, thanks to the self-consistency equation (3.5). Restarting from Eq. (3.24), taking the derivative according to $H$:

$$\left(\frac{\partial m}{\partial H}\right)_T = \left[1 - \tanh^2\left(\frac{T_c}{T}\,m + \frac{H}{k_B T}\right)\right]\left(\frac{T_c}{T}\left(\frac{\partial m}{\partial H}\right)_T + \frac{1}{k_B T}\right) \tag{3.36}$$

In this equation, we can directly replace the tanh-term by $m$ using Eq. (3.24) we just started from:

$$\left(\frac{\partial m}{\partial H}\right)_T = \left[1 - m^2\right]\left(\frac{T_c}{T}\left(\frac{\partial m}{\partial H}\right)_T + \frac{1}{k_B T}\right) \tag{3.37}$$

Taking the limit $H \to 0$ on both sides, we then get:

$$\chi = \left[1 - m_0^2\right]\left(\frac{T_c}{T}\,\chi + \frac{1}{k_B T}\right) \tag{3.38}$$

from which we can extract $\chi$:

$$\chi = \frac{1}{k_B T}\,\frac{1}{\left(\frac{1}{\left[1-m_0^2\right]} - \frac{T_c}{T}\right)} = \frac{(1 - m_0^2)}{k_B T}\,\frac{1}{1 - (1 - m_0^2)\frac{T_c}{T}} \tag{3.39}$$

- For $T > T_c$, we know that $m_0 = 0$, which gives back Eq. (3.28).
- For $T < T_c$, we replace in Eq. (3.39) the expression of Eq. (3.22) for the behavior of $m_0$ close to the critical point, $m_0^2 \sim 3\,(T_c - T)/T_c$, which gives:

$$(1 - m_0^2) = \frac{T_c - 3\,(T_c - T)}{T_c} = \frac{T + 2(T - T_c)}{T_c} \underset{T \to T_c}{\longrightarrow} 1 + 2\frac{(T - T_c)}{T_c} \tag{3.40}$$

and therefore gives:

$$\chi = \frac{1}{k_B T}\left[\frac{T + 2(T - T_c)}{T_c}\right]\frac{1}{1 - \left[\frac{T+2(T-T_c)}{T_c}\right]\frac{T_c}{T}} \qquad \text{for } T < T_c \tag{3.41}$$

$$= \frac{1}{k_B T}\left[\frac{T + 2(T - T_c)}{T_c}\right]\frac{1}{1 - 1 - \left[\frac{2(T-T_c)}{T}\right]} = \frac{1}{k_B T}\left[\frac{T + 2(T - T_c)}{T_c}\right]\frac{1}{\left[\frac{2(T_c-T)}{T}\right]} \tag{3.42}$$

$$\underset{T \to T_c}{\longrightarrow} \frac{1}{k_B}\,[1]\,\frac{1}{[2(T_c - T)]}\,. \tag{3.43}$$

We recover equation (3.34) in the limit $T \to T_c^-$ (coming from below the critical point):

$$\chi = \frac{1}{2k_B}\,(T_c - T)^{-1} \tag{3.44}$$

From Eq. (3.39), one can also simply expand $1/(1-m_0^2)$ in second order in $m_0$, which gives:

$$\frac{1}{\chi} = k_B T \left(\frac{1}{1-m_0^2} - \frac{T_c}{T}\right) = k_B T \left(1 + m_0^2 + o(m_0^2) - \frac{T_c}{T}\right) \tag{3.45}$$

$$\sim k_B T \left(1 + \frac{3\,(T_c - T)}{T_c} - \frac{T_c}{T}\right) \sim k_B \left(T + \frac{3\,T}{T_c}\,(T_c - T) - T_c\right) \tag{3.46}$$

$$\sim k_B \left(\frac{3\,T}{T_c} - 1\right)\,(T_c - T) \tag{3.47}$$

$$\underset{T \to T_c^-}{\sim} 2\,k_B\,(T_c - T) \tag{3.48}$$

which gives back eq. (3.44).

**A3. (⋆)** See lecture slides on Landau theory: In the slides the free energy is expended in small values of $m$ and the steps on how to find out the critical behavior of $c$ and $m$ are then detailed.

### 3.1.2 Back to TASEP!

**A1.** In the stationary state, $d\rho_i/dt = 0$ and so

$$0 = J_{i-1}(t) - J_i(t) \implies J_i(t) = J_{i-1}(t) \equiv J \tag{3.49}$$

Setting $J_0(t) = J_L(t) = J$ allows us to find a relation between $\rho_1$ and $\rho_L$ and $J$.

$$J = \alpha(1-\rho_1), \quad J = \beta\rho_L \implies \rho_1 = 1 - \frac{J}{\alpha}, \quad \rho_L = \frac{J}{\beta} \tag{3.50}$$

**A2.** There are two different ways to approach the mean-field approximation.

**Version 1:** The first approach consists in applying directly the approximation to the joint probability distribution $J_i = P[n_i = 1;\ n_{i+1} = 0]$, by replacing the joint distribution by a product of independent distributions:

$$J_i = P[n_i = 1;\ n_{i+1} = 0] \simeq P[n_i = 1]\ P[n_{i+1} = 0]\,. \tag{3.51}$$

In this approach to the mean-field approximation, we assume that the probability that a given site is occupied is independent of the state of the other sites; each site only "feels" the mean influence of the other sites. Replacing this in the stationary equation (3.49), one gets:

$$J = P[n_i = 1]\ P[n_{i+1} = 0] \tag{3.52}$$

$$= \rho_i(1 - \rho_{i+1})\,, \tag{3.53}$$

where we used that $\rho_i = P[n_i = 1]$ and $P[n_{i+1} = 0] = 1 - \rho_{i+1}$. This finally yields the recurrence relation:

$$\rho_i = \frac{J}{1-\rho_{i+1}} \qquad \text{or equivalently} \qquad \rho_{i+1} = 1 - \frac{J}{\rho_i}\,. \tag{3.54}$$

**Version 2:** In the second approach to the mean-field approximation, we write the number of cars at site $i$ is equal to the mean number of cars $\langle n_i(t)\rangle$, plus some fluctuations $\delta n_i(t)$ around the mean value:

$$n_i(t) = \langle n_i(t)\rangle + \delta n_i(t)\,. \tag{3.55}$$

We then assume that these fluctuations are negligeable compared to $\langle n_i\rangle$. This version of the mean-field approximation is useful only if we have written down an equation for the evolution of the number of cars in the system. We observe that, as

there can only be one car at a time at each site, the number $n_i(t)$ of cars at site $i$ can only be 0 or 1 at any time. This allows to obtain the following expression for the mean number of cars in site $i$ at time $t$:

$$\langle n_i(t)\rangle = 1 \times \rho_i(t) + 0 \times [1 - \rho_i(t)] \tag{3.56}$$
$$= \rho_i(t)\,, \tag{3.57}$$

as $\rho_i(t) = P[n_i(t) = 1]$ is the probability that site $i$ is occupied at time $t$. Similarly, we observe that the current $J_i(t)$ in $i$ at time $t$ is equal to the ensemble average $\langle n_i(t)(1 - n_{i+1}(t))\rangle$. Indeed, $n_i(t)(1 - n_{i+1}(t)$ also only takes two values: value 1 if $i$ is occupied AND $(i + 1)$ is empty, and value 0 in all other cases, and therefore:

$$\langle\, n_i(t)\,(1 - n_{i+1}(t))\,\rangle = 1 \times P[n_i = 1;\ n_{i+1} = 0] + 0 \times [1 - P[n_i = 1;\ n_{i+1} = 0]] \tag{3.58}$$
$$= J_i(t)\,, \tag{3.59}$$

The equation (3.9) could then be re-written under the form:

$$\frac{d\,\langle n_i\rangle}{dt}(t) = J_{i-1}(t) - J_i(t)\,,\ \text{where } J_i(t) = \langle n_i(t)(1 - n_{i+1}(t))\rangle\ , \tag{3.60}$$

and the stationary equation take the form: $J_i = \langle n_i(1 - n_{i+1})\rangle = J$ is a constant. Using the mean-field approximation Eq. (3.55) in the stationary equation we get:

$$J = \langle(\langle n_i\rangle + \delta n_i)(1 - \langle n_{i+1}\rangle - \delta n_{i+1})\rangle \tag{3.61}$$
$$\simeq \langle n_i\rangle\ (1 - \langle n_{i+1}\rangle) \tag{3.62}$$
$$\simeq \rho_i\,(1 - \rho_{i+1})\,, \tag{3.63}$$

where we have neglected all the terms of order $\delta n_i$, $\delta n_{i+1}$, or higher. This equation finally yields the recurrence relations Eq. (3.54). Currently, this relation only holds for sites $i \in [2, L - 1]$. By defining the auxiliary site

$$\rho_0 \equiv \alpha, \quad \rho_{L+1} \equiv 1 - \beta \tag{3.64}$$

we make sure that the recurrence relations holds for all sites on the chain.

**A3.** For a recurrence relation of the form $a_{i+1} = f(a_i)$, the fixed points are given by the condition $f(a_i) = a_i$. So for our particular recurrence relation we have

$$\rho_i = 1 - \frac{J}{\rho_i} = \frac{\rho_i - J}{\rho_i} \implies (\rho_i)^2 - \rho_i + J = 0\,. \tag{3.65}$$

This quadratic equation has solutions only if its discriminant $\Delta = 1 - 4J$ is positive. Therefore, we have three cases:

1) $\Delta < 0 \iff J > 1/4$: there are no solution such that $J > 1/4$: the current of cars in the system cannot be larger than $1/4$.
2) $\Delta = 0 \iff J = 1/4$: this is the maximal current reachable, and in that case the fixed point is $\rho^\infty = 1/2$.
3) $\Delta > 0 \iff J < 1/4$: in this case there are two possible fixed points:

$$\rho^\infty_\pm = \frac{1 \pm \sqrt{1 - 4J}}{2}\,. \tag{3.66}$$

**A4.** For $J > 1/4$ there are no steady state solutions.

For $J = 1/4$ we have the solution $\rho_\infty = \frac{1}{2}$. Starting from $\rho_0 = \alpha$, we find that if

$$\rho_1 = 1 - \frac{1}{4\alpha} < \rho_\infty \tag{3.67}$$

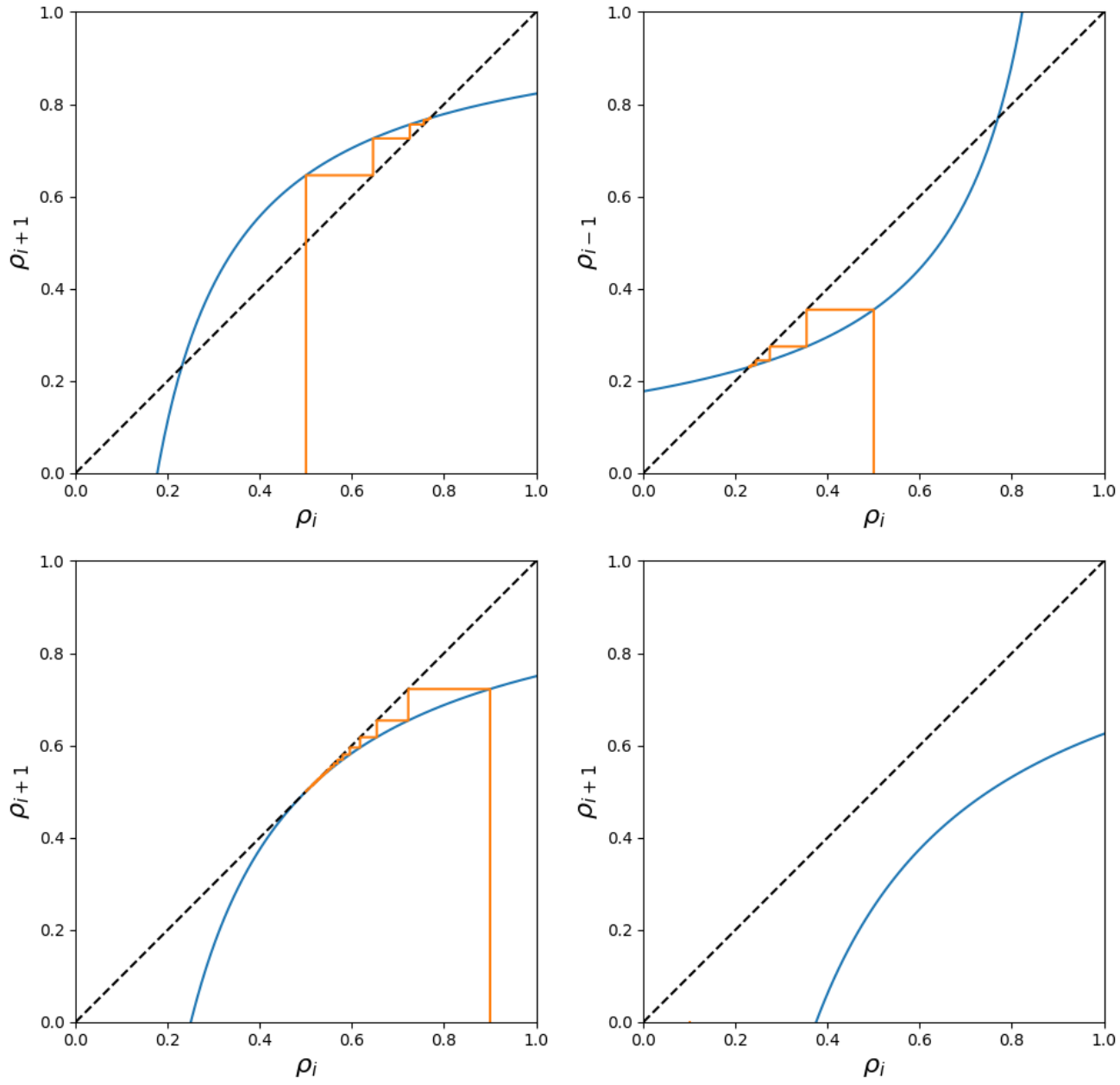



Figure 3.1: Graphical representation of the different steady-solutions. Top left: $J < 1/4$. The iteration converges to $\rho_+$ as long as $\alpha > \rho_-$. Top right: $J < 1/4$ but starting from the right boundary with initial condition $\rho_{L+1} = 1 - \beta$. In this case the iteration only converges to the steady-state solution if $\beta > \alpha$. Bottom left: $J = 1/4$. The iteration converges as long as $\alpha \geq \frac{1}{2}$. Bottom right: $J > 1/4$. There are no steady-state solutions.

we will move away from the fixed point. Solving for $\alpha$ gives that for $\alpha < \frac{1}{2}$ we will not converge to the steady-state density. So for $J = 1/4$ the steady-state is stable for $\alpha \geq \frac{1}{2}$. Similarly, requiring that

$$\rho_L = \frac{1}{4\beta} \geq \rho_\infty \tag{3.68}$$

gives a stability condition in terms of $\beta$. Solving for $\beta$, we find $\beta \geq \frac{1}{2}$. So the fixed point $\rho_\infty = \frac{1}{2}$ corresponding to $J = 1/4$ is stable if $\alpha \geq \frac{1}{2}$ and $\beta \geq \frac{1}{2}$.

If $J < 1/4$ we have two steady-state solutions given by

$$\rho_\pm = \frac{1 \pm \sqrt{1-4J}}{2} \tag{3.69}$$

and we can distinguish between the regimes $\alpha \leq \frac{1}{2}$ and $\beta \leq \frac{1}{2}$. For $\alpha \leq \frac{1}{2}$, we find that the solution converges to a steady state as long as (see Fig. 3.1 - top left)

$$\rho_0 = \alpha \geq \rho_- \tag{3.70}$$

Picking $\alpha = \rho_-$ we find that

$$\alpha = \frac{1-\sqrt{1-4J}}{2} \implies J = \alpha(1-\alpha) \tag{3.71}$$

At the other end of the chain, we then find

$$\rho_{i-1} = \frac{\alpha(1-\alpha)}{1-\rho_i} \tag{3.72}$$

We find that for initial condition $\rho_{L+1} = 1-\beta$, the solution only converges to the stable steady-state solution if $1-\beta < 1-\alpha$, or equivalently $\beta > \alpha$ (see Fig. 3.1 - top right). For $\beta \leq \frac{1}{2}$, the argument is essentially the same. In that case we get $J = \beta(1-\beta)$ and $\alpha > \beta$.

**A5.** From the above analysis, we have found three different phases. The **maximum current phase** is given by $\alpha \geq \frac{1}{2}, \beta \geq \frac{1}{2}$ and $J = 1/4$ with bulk density $\rho_\infty = \frac{1}{2}$. The **low density phase** is characterized by $\alpha \leq \frac{1}{2}, \beta > \alpha$ and $J = \alpha(1-\alpha)$, where $\rho_0 = \rho_\infty = \alpha$ and $\rho_L = \frac{\alpha(1-\alpha)}{\beta}$. The **high density phase** is characterized by $\beta \leq \frac{1}{2}, \alpha > \beta$ and $J = \beta(1-\beta)$. In this case $\rho_{L+1} = \rho_\infty = 1-\beta$ and $\rho_1 = 1 - \frac{\beta(1-\beta)}{\alpha}$.

### 3.1.3 Ising model with long-range interactions; Mean-field Ising and Landau theory

**A1.** Expanding out the square of the sum yields

$$\left(\sum_{i=1}^N s_i\right)^2 = \sum_{i=1}^N \sum_{j=1}^N s_i s_j = \sum_{i=1}^N \sum_{j\neq i}^N s_i s_j + \sum_{i=j}^N s_i s_j \tag{3.73}$$

Note that when $s_i = s_j$, we obtain $s_i s_i = (s_i)^2 = 1$ and therefore we find that

$$\left(\sum_{i=1}^N s_i\right)^2 = 2 \sum_{pair(i,j)} s_i s_j + N \tag{3.74}$$

where the factor 2 is due to the fact that we count both the pairs $(i, j)$ and the pairs $(j, i)$. Rearranging the equation above yields the re-written form of the energy.

**A2.** Since the magnetization is a sum over the spins, the configurations with equal magnetization are those for which we have the same number of spins in the +1 configuration. If $N_+$ is the number of spins in this configuration, then we can formally write that the number of states is

$$\Omega(M) = \binom{N}{N_+} = \frac{N!}{N_+!(N-N_+)!} \tag{3.75}$$

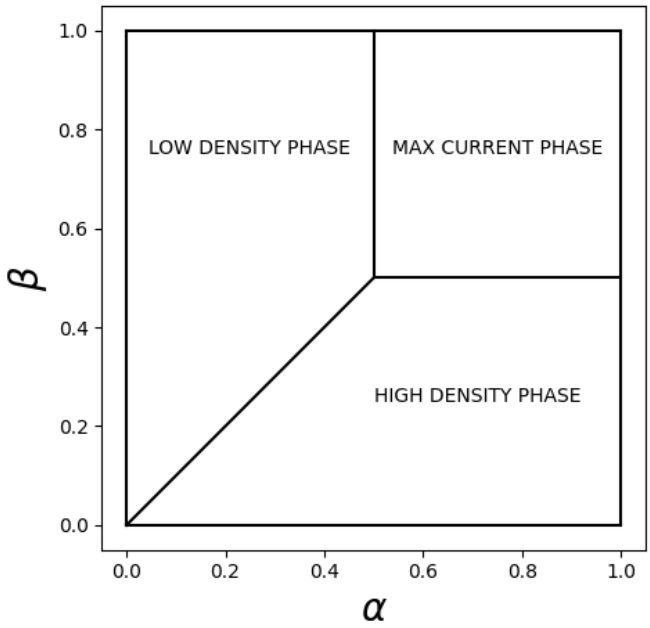



Figure 3.2: The TASEP phase diagram as function of $\alpha$ and $\beta$.

To rewrite this in terms of $M$, we note that

$$M = N_+(+1) + N_-(-1) = N_+ - N_- = N_+ - (N - N_+) = 2N_+ - N \tag{3.76}$$

Using this relation, we can express $N_+$ in terms of $M$ and find

$$N_+ = \frac{M+N}{2} \tag{3.77}$$

Plugging this relation into our expression for the number of states with the same magnetization yields the requested formula.

**A3.** Since the partition function is a sum over all the possible states, and because we can write the energy as

$$E(\vec{s}) = -\frac{J}{2N}\left[\left(\sum_{i=1}^{N} s_i\right)^2 - N\right] = -\frac{J}{2N}\left[M^2 - N\right], \tag{3.78}$$

clearly each state with the magnetization has the same energy. We can therefore sum over the states with equal magnetization and account for their multiplicity with the factor $\Omega(N)$ that we have derived.

**A4.** We take the partition function $Z$ and only keep the term with magnetization $M^*$, which minimizes the free energy. In this approximation, the partition function is just

$$Z^* = \Omega(M^*)\exp(-\beta E(M^*)) \tag{3.79}$$

We write just $M$ instead of $M^*$ and the logarithm to find

$$-k_B T \ln Z = -k_B T[\ln \Omega(M) - \beta E(M)] \tag{3.80}$$

We already expressed the energy in terms of the magnetization $M$ in our answer to Q3, so we can write

$$-k_B T \ln Z = E(M) - k_B T \ln \Omega(M) = -\frac{J(M^2 - N)}{2N} - k_B T \ln \Omega(M) \tag{3.81}$$

The Stirling formula allows us to rewrite factorials for large values as a more tractable expression. It states

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \tag{3.82}$$

or we can immediately use its logarithmic form, i.e.

$$\ln(n!) \approx n\ln n - n \tag{3.83}$$

Taking the logarithm of the multiplicity, we find

$$\ln \Omega(M) = \ln(N!) - \ln(N_+!) - \ln(N_-!) \approx N\ln N - N - N_+\ln N_+ + N_+ - N_-\ln N_- + N_- \tag{3.84}$$

Since $N_+ + N_- = N$, the non-logarithmic terms cancel and we have

$$F(M,T) = -\frac{JM^2}{2N} + \frac{J}{2} + k_BT(N_+\ln N_+ + N_-\ln N_- - N\ln N) \tag{3.85}$$

**A5.** Taking the derivative of $F(M,T)$ with respect to $M$ (and remembering that $N_+$ and $N_-$ depend on $M$) we get

$$\frac{\partial F(M,T)}{\partial M} = -\frac{JM}{N} + \frac{k_BT}{2}\left(\log\frac{N+M}{N-M}\right) = -\frac{JM}{N} + k_BT\tanh^{-1}\left(\frac{M}{N}\right) = 0 \tag{3.86}$$

Recognizing that $m = M/N$, we find the following transcendental equation for $m$

$$m = \tanh(\beta Jm) \tag{3.87}$$

If we plot both $m$ and $\tanh(\beta Jm)$ as a function of $m$, one notices that for $\beta J < 1$, the graphs only intersect at the origin. For $\beta J > 1$, the graph of $\tanh(\beta J)$ is intersected twice at the corresponding values of the magnetization $\pm m$. Therefore, the critical temperature occurs when $\beta J = 1$ and so

$$T_c = \frac{J}{k_B} \tag{3.88}$$

Note that because we scaled our coupling parameter with $N$, we got rid of the dependence on $q$ in the expression for the critical temperature. (Strictly speaking, we cancelled a factor of $N/N$.)

**A6.** We first write the free energy in terms of the per-spin-magnetization $m$ as follows:

$$F(m,T) = -\frac{JNm^2}{2} + \frac{J}{2} + k_BT\left[\frac{N(1+m)}{2}\log\frac{N(1+m)}{2} + \frac{N(1-m)}{2}\log\frac{N(1-m)}{2} - N\log N\right] \tag{3.89}$$

Dividing by $N$ we have

$$f(m,T) = -\frac{Jm^2}{2} + \frac{J}{2N} + k_BT\left[\frac{1+m}{2}\log\frac{1+m}{2} + \frac{1-m}{2}\log\frac{1-m}{2}\right] \tag{3.90}$$

Expanding around $m = 0$ and using the fact that $\log(1+x) \approx x - \frac{x^2}{2} + \frac{x^3}{3} + O(x^4)$ we find

$$f(m,T) = -\frac{Jm^2}{2} + \frac{J}{2N} + k_BT\left[\frac{m^2}{2} + \frac{m^4}{12} - \log 2\right] \tag{3.91}$$

Collecting like powers of $m$ we have

$$f(m,T) = \frac{J}{2N} - k_BT\log 2 + m^2\left(\frac{k_BT - J}{2}\right) + \frac{k_BT}{12}m^4 \tag{3.92}$$

Since we had $T_c = k_BJ$ we can write

$$f(m,T) = f_0 + \frac{1}{2}m^2k_B(T - T_c) + \frac{1}{4}\frac{k_BT}{3}m^4 \tag{3.93}$$

and so

$$f_0 = \frac{J}{2N} - k_BT\log 2, \quad a_2(T) = k_B(T - T_c), \quad a_4(T) = \frac{k_BT}{3} \tag{3.94}$$

# Chapter 4

# Scale-invariance and Universality: Random Walks, Renormalization

Niki Stratikopoulou, Clélia de Mulatier

**Symbol "(⋆)":** Questions and exercises indicated with a (⋆) are optional. No worries if you don't have time to try to solve them, or if you don't manage to solve them on your own.

**The symbol "</>":** indicates optional questions with numerical simulation.

## 4.1 Discrete Space Random walks

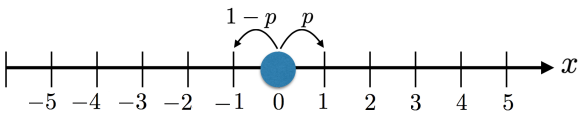



**Fig 1:** 1D random walk on a lattice.

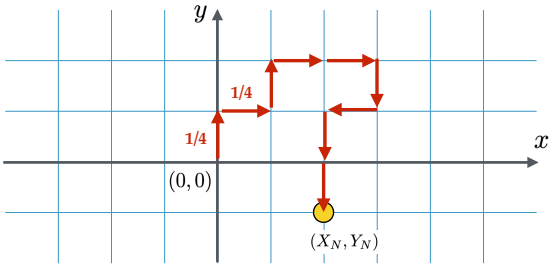



**Fig 2:** 2D random walk on a lattice.

### 4.1.1 1D random walker

Take a walker hopping on the sites of a 1D lattice (see Fig. 1). The walker starts from the position $X_0 = 0$. It then randomly hops:

- to the next site on its right with probability $p$;
- or, to the next site on its left with probability $(1 - p)$.

We denote by $X_N$ the position of the walker after $N$ steps. This position is a random variable, as its value can change for each new realization of the walk. We would like to characterize better this random position.

For example, imagine playing a "Heads or Tails?" game with a (unfair) coin that gives "Heads" with probability $p$. You have chosen "Heads", and will win \$1 each times a toss returns "Heads" and lose \$1 each time its returns "Tails". The evolution of the amount of the money that you earn or lose during the course of the game is represented by the simple 1d random walk

model described above. The position $X_i$ of the walker after $i$ steps models the amount of money you have earned ($X_i > 0$) or lost ($X_i < 0$) after $i$ tosses of the coin.

> **Symmetric walk: $p = 0.5$.**
> Let us start with the simple case $p = (1 - p) = 0.5$: the walker jumps to right or to the left with equal probability.

We denote $r_i$ the jumps performed by the walker at the $i$-th step (or toss): $r_i = +1$ if the walker moves to the next site on its right and $r_i = -1$ if it moves to the left, such that the position after $i$ steps is given by: $X_i = X_{i-1} + r_i$.

**Q1.** Can you show that $X_N = \sum_{i=1}^{N} r_i$? Can you deduce that for $p = 0.5$, $\langle X_N \rangle = 0$?

**Q2.** Can you show that $X_N^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} r_i r_j = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} r_i r_j + \sum_{i=1}^{N} r_i^2$? Can you deduce the value of $\langle X_N^2 \rangle$?

> **Asymmetric walk: $p \neq 0.5$.**
> The walker now jumps to the right with probability $p$ and to the left with probability $(1 - p)$.

**Q3.** Using a numerical simulation, sample $10^4$ realizations of the walk until 100 steps. What seems to be the distribution of final positions $X_{100}$ of the walker?

**Q4.** Using the same reasoning as in Q1, can you deduce the value of $\langle X_N \rangle$ for any value of $p$? Check that you recover $\langle X_N \rangle = 0$ for $p = 0.5$.

**Q5.** Can you show that the variance of $X_N$ is equal to $N$ times the variance of $r_i$:

$$\sigma^2_{X_N} = \langle X_N^2 \rangle - \langle X_N \rangle^2 = N(\langle r_i^2 \rangle - \langle r_i \rangle^2) = N\, \sigma^2_{r_i} \,. \tag{4.1}$$

Can you deduce the value of $\sigma^2_{X_N}$ for any value of $p$? Check that you recover the value obtained in Q2 for $p = 0.5$.

**Q6.** Using a numerical simulation, sample $10^4$ realizations of the walk with $p = 2/3$ until 100 steps. What seems to be the distribution of final positions $X_{100}$ of the walker?

**Q7.** Let denote $X_N^+$, resp. $X_N^-$, the total number of jumps to the right, resp. to the left, performed by the walker until the $N$-th steps. Both $X_N^+$ and $X_N^-$ are random variables. What is the distribution of $X_N^+$? What is its mean value? its standard deviation? For large values of $N$ to which distribution is the distribution of $X_N^+$ converging to (see de Moivre–Laplace theorem)?

**Q8.** Can you express $X_N$ as a function of $X_N^+$? Using this result, can you check that you re-obtain the values of the $\langle X_N \rangle$ and $\sigma^2_{X_N}$ derived in the previous question? For large values of $N$, which distribution is then expected for $X_N$? Can you compare this distribution to the distribution of the $10^4$ final positions obtained in Q6?

### 4.1.2 Random walk on a 2D lattice

> **2D random walk on a lattice.** We consider the following 2D generalization of the previous walker (see Fig. 2). The walker is now randomly evolving on a 2D square lattice: it jumps in each direction (up, down, left or right) with the same probability 1/4. The position of the walker after $N$ steps on the 2D lattice is indicated by its coordinate $(X_N, Y_N)$. The walker initially starts from the origin $(X_0, Y_0) = (0, 0)$.

**Q9.** Can you generate and plot one realization of such 2D random walk with 100 000 steps? How does the plot of this trajectory looks like if you cut it to its first 10 000 steps? Observe the scale invariance structure of the random trajectory

(looks the same at all scale when $N$ is sufficiently large).

**Q10.** We are interested in characterizing better the statistics of the position of the walker after $N$ steps. Can you generate $10^4$ trajectories of 100 steps and plot their final positions? What do you expect regarding the distribution of the walker positions? Can you plot this distribution?

**Q11.** We denote by $\vec{R}_N = (X_N, Y_N)$ the vector indicating the position of the walker after $N$ steps in the 2D plane. Can you show that the average final position of the walker is $\langle \vec{R}_N \rangle = (0, 0)$?

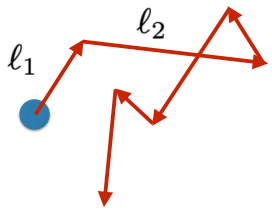
**Q12.** We observe that $\vec{R}_N = \sum_{i=1}^{N} \vec{r}_i$, where $\vec{r}_i = (x_i, y_i)$ is the vector of coordinates $(x_i, y_i)$ corresponding to the $i$-th jump. Can you show that $\langle \vec{R}_N^2 \rangle = N\ell^2$, where $\ell = ||r_i|| = \sqrt{2}$?

### 4.1.3 Random walk on a lattice of any dimension $d$.

**Q13.** Can you show that the result of Q12 holds for any symmetric random walks on a square lattice in any dimension?

**Q14.** Can you show that the fractal dimension associated with the trajectory of the random walker is $d_f = 2$?

## 4.2 Continuous Space Random walks




### 4.2.1 Universality

Let's go hunt for mushroom in the forest... We are looking for a specific type of mushrooms that grows on trees. Our strategy is the following: we start from a tree, we then take a random direction and go straight in that direction until we reach another tree, we check if there are mushroom on that tree, and re-iterate.

**Q15.** Let's assume that trees are uniformly distributed in the forest. Starting from a random tree and a random direction, the average distance to the next tree is of the order of 5 meters. What do you expect for the distribution of the distance to travel before reaching the next tree?
A similar model (extended to 3 dimensions) is used to simulate the diffusion of light in a diffusive material or the diffusive transport of neutrons in a nuclear reactor.

**Q16.** Can you generate a trajectory of 10 000 steps of this walker? Can you compare the trajectory to the one sampled in question Q9?

**Q17.** Can you generate $10^4$ realizations of 100 steps of this walkers with an average distance between trees of $\ell = \sqrt{2}$ and plot the distribution of their final positions? Compare to the distribution found in Q10.

### 4.2.2 Diffusion equation – Fokker-Planck equation

**Q18.** Going back to the discrete 1D case described in Sec. 4.1.1, let $P(x, N)$ be the probability to find the walker in position $x$ after $N$ jumps. Can you give the master equation describing the discrete evolution of the walker on the lattice? The equation gives the probability to find the walker in $x$ after $N+1$ steps, $P(x, N+1)$, as a function of the probabilities to find the walker on other places on the lattice after $N$ steps.

**Q19.** Let's assume that the walker starts from $x_0$ at time $t = 0$ and that jumps occur at regular time intervals $dt$. After $N$ steps, the corresponding time is $t = N\,dt$, and $P(x, N) = P(x, t)$. We also take the lattice step size to be equal to a small value $\ell$. Assuming that $dt$ is very small, can you write a Taylor expansion of $P(x, t + dt)$ to first order in $dt$? Assuming that $\ell$ is very small, can you write a Taylor expansion of $p(x + \ell, N)$ in $\ell$ to order 2?

**Q20.** Let us take $p = 1/2$. Combining the Taylor expansions in the master equation, can you recover the diffusion equation? Take the limit of very small jump length $\ell$ and very small duration $dt$, such that the ratio $\ell^2/dt$ goes to a constant.

**Q21.** Let us take $p \neq 1/2$. How is this equation modified? Do you recognize this equation?

## 4.3 (⋆) Real Space renormalization

### 4.3.1 Renormalization in $d = 1$ Ising model

We consider the usual Ising Hamiltonian (Energy) with nearest-neighbor interactions

$$E(S_i; J) = -J \sum_i S_i S_{i+1}\,, \quad \text{where} \quad S_i = \pm 1\,. \tag{4.2}$$

The Boltzmann weight for each pair of spins is

$$W(S_i, S_{i+1}; v) = e^{K S_i \cdot S_{i+1}}\,, \qquad \text{where } K = \frac{J}{k_B T}\,. \tag{4.3}$$

**Q1.** Write the Boltzmann weight in the form

$$e^{K S_i S_{i+1}} = \cosh(K)\,(1 + \tanh(K)\, S_i S_{i+1})\,. \tag{4.4}$$

**Q2. Warm up**: Consider three consecutive spins $S_i$, $S_{i+1}$ and $S_{i+2}$, with $S_i$ and $S_{i+2}$ being fixed. The corresponding Boltzmann factor where they appear is

$$\exp\left(K\, S_i S_{i+1} + K\, S_{i+1} S_{i+2}\right)\,. \tag{4.5}$$

Using eq. (4.4) we can "integrate out" the middle spin $S_{i+1}$, by taking the sum over all its possible values. Can you prove that

$$\sum_{S_{i+1}=\pm 1} e^{K S_i S_{i+1} + S_{i+1} S_{i+2}} = 2\cosh^2(K)(1 + t^2\, S_i S_{i+2})\,, \tag{4.6}$$

where for convenience $t \equiv \tanh(K)$.

Note: this procedure of "integrating out" a random variable from a probability distribution is often called "*marginalization*": i.e., by computing the probability distribution:

$$P(s_1, \cdots, s_{i-1}, s_{i+1}, \cdots, s_n) = \sum_{s_i} P(s_1, \cdots, s_{i-1}, s_i, s_{i+1}, \cdots, s_n)\,, \tag{4.7}$$

we are *marginalizing out* the variable $s_i$.

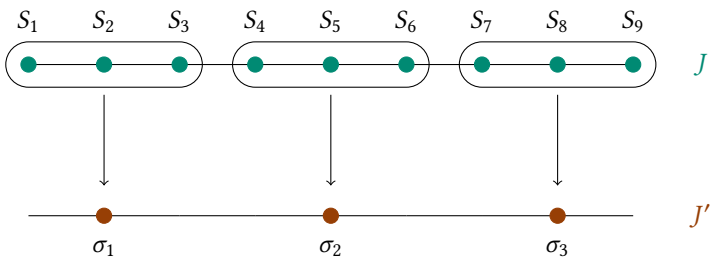



Figure 4.1: RG transformation: block spins and decimate.

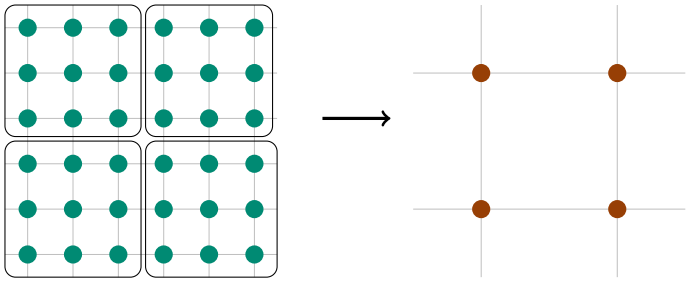


Figure 4.2: One possible decimation procedure.

**RG transformation.** One possible RG transformation is to divide the system into blocks, each including three spins, and apply the *decimation* rule: for each block we choose as a spin of the new system, the one that is at the center (see examples in Fig. 4.1 and 4.2). Applying this procedure to the probability original probability distribution would mean marginalizing over all the spins that are not kept at the next step.

**Q3.** For the first RG step, we can consider having two neighboring blocks and summing over the intermediate spins $(S_3, S_4)$, while keeping fixed the values of the spins at the center of the blocks $\sigma_1 \equiv S_2$ and $\sigma_2 \equiv S_5$. Write the new (*effective*) Boltzmann factor and show that the new couplings $K'$ can be rewritten from $K$, through a recursion relation.

**Q4.** Plot the recursion relation, and graphically show that there are two fixed points. Remembering that we have absorbed a factor $1/k_BT$ in $K$, which fixed points correspond to the high and low temperature phases of the system? In which phase does the 1d Ising model live? *Hint:* Use Cobweb plots.

**Q5** Now we would like to generalize eq. (4.6), by *marginalizing out* ($\Leftrightarrow$ summing over) $(k-1)$ spins consecutive spins. Given eq. (4.5) can you induce the equivalent, generalized, form of the summing procedure and the recursion relation for the coupling constants?

### 4.3.2 Renormalization in $d = 2$ Ising model

Moving to the 2d lattice, soon we figure out that the decimation procedure we applied in $1d$ cannot be done exactly in $2d$, as depicted in Fig. 4.2, since new spin interaction terms are generated at the end of each step. Nevertheless, it is possible to construct another, more sophisticated, decimation scheme that allows us to have the same recursion relation that we had in the $1d$ case. Specifically the scheme relies on *moving* the bonds that are not connected, to the spins that are kept. These spins are then connected by bonds of strength $2J$ instead of $J$.

**Q 2.1** Argue that the new recursion relation after decimating 2 spins is

$$t' \equiv \tanh(K') = \tanh^2(2K) \tag{4.8}$$

**Q 2.2** Using the recursion relation (4.8), find the fixed points $J_c$ of this decimation scheme. Give a qualitative analysis of the fixed points.
*Hint:* Use the identity $\tanh(2x) = \frac{2\tanh(x)}{1+\tanh^2(x)}$ and set $t = \tanh(K)$ for convenience.

## 4.4 Solutions

### 4.4.1 Discrete Space Random walks

#### 1D random walker

**A1.** The walker initially starts from the position $X_0 = 0$. Then, at each step (indexed by $i$), it performs a random jump $r_i \in \{-1, +1\}$. Therefore its position is:

$$\text{after } N = 1 \text{ step}: \quad X_1 = X_0 + r_1 = r_1\,, \tag{4.9}$$
$$\text{after } N = 2 \text{ steps}: \quad X_2 = X_1 + r_2 = r_1 + r_2\,, \tag{4.10}$$
$$\text{after } N = 3 \text{ steps}: \quad X_3 = X_2 + r_3 = r_1 + r_2 + r_3\,, \tag{4.11}$$
$$\text{after } N \text{ steps}: \quad X_N = X_{N-1} + r_N = (r_1 + \cdots + r_{i-1}) + r_N = \sum_{i=1}^{N} r_i\,. \tag{4.12}$$

Taking the ensemble average on both sides gives:

$$\langle X_N \rangle = \left\langle \sum_{i=1}^{N} r_i \right\rangle = \sum_{i=1}^{N} \langle r_i \rangle \tag{4.13}$$

as the average of a sum of terms is equal to the sum of the averages of each terms. Here we observe that the random variables $r_i$ are independent random variables, all sampled from the same distribution, which is $P(r_i) = [\delta(r_i + 1) + \delta(r_i - 1)]/2$, i.e. that $r_i$ is equal to $+1$ with probability $1/2$ and is equal to $-1$ with probability $1/2$. Therefore all the variables $r_i$ have the same mean which is given by:

$$\langle r_i \rangle = \frac{1}{2} \times (+1) + \frac{1}{2} \times (-1) = 0\,. \tag{4.14}$$

Replacing this result in Eq. (4.13) finally gives that $\langle X_N \rangle = N \times \langle r_i \rangle = 0$.

**A2. Erratum:** In the text of question Q2, the correct expression is $X_N^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} r_i r_j$ (the expression previously appearing in the tutorial text had an extra term that was not correct). This has been adjusted in the present text. The last result was correct though, i.e. $X_N^2 = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} r_i r_j + \sum_{i=1}^{N} r_i^2$ is correct.

Starting from $X_N = \sum_{i=1}^{N} r_i$, we take the square:

$$X_N^2 = \left( \sum_{i=1}^{N} r_i \right) \times \left( \sum_{i=1}^{N} r_i \right), \tag{4.15}$$
$$= (r_1 + r_2 + \cdots + r_N) \times (r_1 + r_2 + \cdots + r_N)\,, \tag{4.16}$$
$$= (r_1 \times r_1 + r_1 \times r_2 + \cdots + r_1 \times r_N) + (r_2 \times r_1 + r_2 \times r_2 + \cdots + r_2 \times r_N) + \cdots \tag{4.17}$$
$$= \sum_{i=1}^{N} \sum_{j=1}^{N} r_i r_j \tag{4.18}$$

which gives all possible products of the type $r_i r_j$ for all indexes $i$ and $j$ (up to $N$). Separating the terms for which $i = j$ from the terms with $i \neq j$, one finally gets:

$$X_N^2 = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} r_i r_j + \sum_{i=1}^{N} r_i^2\,. \tag{4.19}$$

Note: you certainly recognize here the identity $(r_1 + r_2)^2 = r_1^2 + 2\,r_1 r_2 + r_2^2$, which can be re-written under the form: $(r_1 + r_2)^2 = r_1^2 + r_1 r_2 + r_2 r_1 + r_2^2$. The result above is just a generalization of this identity when there are more terms in the sum.

Taking the ensemble average on both sides of Eq. (4.19):

$$\langle X_N^2 \rangle = \left\langle \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} r_i r_j + \sum_{i=1}^{N} r_i^2 \right\rangle , \tag{4.20}$$

$$= \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \langle r_i r_j \rangle + \sum_{i=1}^{N} \langle r_i^2 \rangle , \tag{4.21}$$

where we used again that the average of a sum of terms is the sum of the averages of the terms. Besides, for $i \neq j$, the two random variables $r_i$ and $r_j$ are independently sampled. This implies that they are not correlated, i.e. that[1] $\langle r_i r_j \rangle = \langle r_i \rangle \langle r_j \rangle$. Given that, $\langle r_i \rangle = 0$ at any step $i$, we get that:

$$\langle r_i r_j \rangle = \langle r_i \rangle \langle r_j \rangle = 0 , \qquad \text{for all } i \neq j . \tag{4.22}$$

Finally, here again we observe that the random variables $r_i$ are independent random variables, all sampled from the same distribution, $P(r_i) = [\delta(r_i + 1) + \delta(r_i - 1)]/2$, i.e. that $r_i$ is equal to $+1$ with probability $1/2$ and is equal to $-1$ with probability $1/2$. Therefore all the variables $r_i$ have the same second order moment, which is given by:

$$\langle r_i^2 \rangle = \frac{1}{2} \times (+1)^2 + \frac{1}{2} \times (-1)^2 = 1 . \tag{4.23}$$

Replacing the results (4.22) and (4.23) in Eq. (4.21) finally gives that $\langle X_N^2 \rangle = N \times \langle r_i^2 \rangle = N$.

We observe that the variance of the position of the walker after $N$ steps grows linearly with $N$ ($\sigma_N^2 = \langle X_N^2 \rangle - \langle X_N \rangle^2 = N$), and therefore the standard deviation of the position of the walker grows as $\sqrt{N}$.

**A3.** We sample $10^4$ realizations of the 1D random walk described in this section (see Fig. 1). Each realization of the walk is stopped at $N = 100$ steps. The distribution of the final positions of the $10^4$ walkers is expected to be a Gaussian distribution with mean $\langle X_{100} \rangle = 0$ and variance $\sigma_{100}^2 = \langle X_{100}^2 \rangle - \langle X_{100} \rangle^2 = N = 100$ (i.e. a standard deviation of $\sigma_{100} = \sqrt{100} = 10$).

**A4.** Starting from Eq. (4.13), and using that the $r_i$ are all sampled from the same distribution, we obtain:

$$\langle X_N \rangle = \sum_{i=1}^{N} \langle r_i \rangle = N \langle r_i \rangle . \tag{4.24}$$

The random variables $r_i$ can take two values: either $r_i = +1$ with the probability $p$, or $r_i = -1$ with the probability $(1 - p)$. Therefore the mean of $r_i$ is simply given by:

$$\langle r_i \rangle = (+1) \times p + (-1) \times (1 - p) = 2p - 1 , \tag{4.25}$$

which finally gives:

$$\langle X_N \rangle = N (2p - 1) . \tag{4.26}$$

Taking $p = 0.5$ gives back $\langle X_N \rangle = 0$ for the symmetric random walk. Besides, if $p > 0.5$ (resp. $p < 0.5$), we have that $\langle X_N \rangle > 0$ (resp. $\langle X_N \rangle < 0$): the walker is drifting towards the right (resp. left) side of the $x$-axis.

**A5.** Starting from Eq. (4.21), and using that the $r_i$ are all independently sampled from the same distribution, we obtain:

$$\langle X_N^2 \rangle = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \langle r_i r_j \rangle + \sum_{i=1}^{N} \langle r_i^2 \rangle , \tag{4.27}$$

$$= \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \langle r_i \rangle \langle r_j \rangle + \sum_{i=1}^{N} \langle r_i^2 \rangle , \qquad \text{using the independence between } r_i \text{ and } r_j \tag{4.28}$$

$$\langle X_N^2 \rangle = N(N-1) \langle r_i \rangle^2 + N \langle r_i^2 \rangle , \qquad \text{using that all the } r_i \text{ are sampled from the same distribution.} \tag{4.29}$$

[1] Note that this is not true in general: the average of a product of terms is not equal to the product of the averages of the terms. This is only correct when the random terms in the product are statistically independent. For instance, this doesn't work for the product $r_i \times r_i$: $\langle r_i^2 \rangle \neq \langle r_i \rangle^2$.

There are $N(N-1)$ terms in the first sum: $\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} 1 = N(N-1)$ (think for instance of the number of elements of an $N \times N$-matrix, excluding the diagonal elements).

The variance of $X_N$ is defined as $\sigma^2_{X_N} = \langle X_N^2\rangle - \langle X_N\rangle^2$. Replacing that $\langle X_N\rangle = N\langle r_i\rangle$ (see Eq. (4.24)) and the previous results Eq. (4.29) gives:

$$\sigma^2_{X_N} = \left(N(N-1)\langle r_i\rangle^2 + N\langle r_i^2\rangle\right) - \left(N\langle r_i\rangle\right)^2, \tag{4.30}$$

$$= N^2\langle r_i\rangle^2 - N\langle r_i\rangle^2 + N\langle r_i^2\rangle - \left(N\langle r_i\rangle\right)^2, \tag{4.31}$$

$$= N\left(\langle r_i^2\rangle - \langle r_i\rangle^2\right) \tag{4.32}$$

$$\sigma^2_{X_N} = N\,\sigma^2_{r_i}\,. \tag{4.33}$$

The variance of $X_N$ is equal to $N$ times the variance of $r_i$. This is true for any variable $X_N$ defines as the sum of $N$ independently and identically distributed (i.i.d.) random variables $r_i$ for which $\sigma_{r_i}$ is finite.

**Independent and Identically Distributed random variables (or, i.i.d. random variables):**
For any variable $X_N$ defines as the sum of $N$ independently and identically distributed (i.i.d.) random variables $r_i$ for which $\sigma_{r_i}$ is finite, we have that:

- the mean of $X_N$ is equal to $N$ times the mean of $r_i$:

$$\langle X_N\rangle = N\langle r_i\rangle\,; \tag{4.34}$$

- the variance of $X_N$ is equal to $N$ times the variance of $r_i$:

$$\sigma^2_{X_N} = N\,\sigma^2_{r_i} = N\left(\langle r_i^2\rangle - \langle r_i\rangle^2\right). \tag{4.35}$$

As a consequence, the standard deviation of the final positions of a 1-dimensional random walk defined by $X_N$ always grows as $\sqrt{N}$, as long as $\sigma_{r_i}$ is finite (no matter the detailed shape of the distribution of $r_i$).

The random variables $r_i$ can take two values: either $r_i = +1$ with the probability $p$, or $r_i = -1$ with the probability $(1-p)$. Therefore the second moment of $r_i$ is given by:

$$\langle r_i^2\rangle = (+1)^2 \times p + (-1)^2 \times (1-p) = 1\,. \tag{4.36}$$

The mean of $r_i$ was already computed in Eq. (4.25), which gives the variance:

$$\sigma^2_{r_i} = \langle r_i^2\rangle - \langle r_i\rangle^2 = 1 - (2p-1)^2 \tag{4.37}$$

$$= 4\,p\,(1-p)\,. \tag{4.38}$$

Finally, the variance of the position of the walker after $N$ steps, computed using $\sigma^2_{X_N} = N\,\sigma^2_{r_i}$, is:

$$\sigma^2_{X_N} = 4\,p\,(1-p)\,N\,. \tag{4.39}$$

For a symmetric random walk, $p = 1/2$, we recover that $\sigma^2_{X_N} = N$.

For $p = 0$ and for $p = 1$, $\sigma^2_{X_N} = 0$. This is normal, as for both values the walk becomes deterministic (there is no randomness/stochasticity anymore), and therefore we can expect that there will be zero variance in the final position of the walker: for $p = 1$ (resp. $p = 0$) the position of the walker after $N$ steps is always exactly $X_N = N$ (resp. $X_N = -N$), i.e. the walker has done $N$ steps to the right (resp. to the left).

**A6.** Taking $p = 2/3$, we expect the final positions of the walker to be Gaussian distributed with a mean $X_{100} = N\,(2p-1) \simeq 33.33$ and a variance of $\sigma^2_{100} = 4\,p\,(1-p)\,N \simeq 88.88$.

## Random walk on a 2D lattice

## Random walk on a lattice of any dimension $d$

### 4.4.2 Continuous Space Random walks

## Universality

## Diffusion equation – Fokker-Planck equation

**A18.** Because the walker can only jump to one of the two nearest site at each step, for the walker to be in position $x$ after $(N+1)$ steps, it must have been in position $(x+1)$ or in position $(x-1)$ just before, i.e. at the $N$-th step. Therefore, the probability to find the walker in position $x$ after $(N+1)$ steps is given by:

$$P(x,\, N+1) = \underbrace{P(x-1,\, N) \times p}_{(1)} + \underbrace{P(x+1,\, N) \times (1-p)}_{(2)}\,, \tag{4.40}$$

where

(1) is the probability that the walker was at the position $(x-1)$ at the previous step $N$ and then jumped to the right from the position $(x-1)$ to the position $x$,

(2) is the probability that the walker was at the position $(x+1)$ at the previous step $N$ and then jumped from to the left from the position $(x+1)$ to the position $x$.

**A19.** Note that taking the lattice step size equal to $\ell$ instead of just "1", and replacing jump numbers $N$ by the time $t = N\,dt$, the previous equation is modified as:

$$P(x,\, t+dt) = \underbrace{P(x-\ell,\, t) \times p}_{(1)} + \underbrace{P(x+\ell,\, t) \times (1-p)}_{(2)}\,. \tag{4.41}$$

> We recall that the Taylor expansion of a function $f(x)$ (of a single variable $x$) that is infinitely differentiable at a point $x_0$ is given by the infinite sum:
>
> $$f(x_0+h) = f(x_0) + h\,f'(x_0) + \frac{h^2}{2} f''(x_0) + \frac{h^3}{3!} f'''(x_0) + \cdots\,. \tag{4.42}$$
>
> For very small values of $h$, one may then truncate this sum at a chosen order.

Assuming $dt$ small, a Taylor expansion of $P(x,\, t+dt)$ to first order in $dt$ gives:

$$P(x,\, t+dt) = P(x,t) + dt\,\frac{\partial P}{\partial t}(x,t) + o(dt)\,. \tag{4.43}$$

Note that this expansion is done by considering $P(x,t)$ as a function of the variable $t$, while $x$ is kept constant.

Assuming $\ell$ small, a Taylor expansion of $P(x+\ell,\, t)$ and of $P(x-\ell,\, t)$ to second order in $\ell$ gives:

$$P(x+\ell,\, t) = P(x,t) + \ell\,\frac{\partial P}{\partial x}(x,t) + \frac{\ell^2}{2}\,\frac{\partial^2 P}{\partial x^2}(x,t) + o(\ell^2)\,. \tag{4.44}$$

and

$$P(x-\ell,\, t) = P(x,t) - \ell\,\frac{\partial P}{\partial x}(x,t) + \frac{\ell^2}{2}\,\frac{\partial^2 P}{\partial x^2}(x,t) + o(\ell^2)\,. \tag{4.45}$$

Note that these two expansions are done by considering $P(x,t)$ as a function of the variable $x$, while $t$ is kept constant. Also note that the sign in front of the first order term is the only thing that changes between Eq. (4.44) and Eq. (4.45).

**A20.** Taking $p = 1/2$, and replacing the three Taylor expansions above into the master equation (4.41) gives:

$$P(x,t) + \frac{\partial P}{\partial t}(x,t)\,dt + o(dt) = \left[P(x,t) - \ell\,\frac{\partial P}{\partial x}(x,t) + \frac{\ell^2}{2}\,\frac{\partial^2 P}{\partial x^2}(x,t) + o(\ell^2)\right] \times p + \left[P(x,t) + \ell\,\frac{\partial P}{\partial x}(x,t) + \frac{\ell^2}{2}\,\frac{\partial^2 P}{\partial x^2}(x,t) + o(\ell^2)\right] \times (1-p)\,. \tag{4.46}$$

After replacing $p = 1/2$, simplifying and reorganizing of the terms in Eq. (4.46), we get:

$$\frac{\partial P}{\partial t}(x,t)\,dt + o(dt) = \frac{\ell^2}{2}\,\frac{\partial^2 P}{\partial x^2}(x,t) + o(\ell^2)\,. \tag{4.47}$$

Note that because the first order terms in $\ell$ cancel each other, one needs to go to the next order in $\ell$ (second order) to get a non-zero term in the previous equation. Besides, the text of the question assumes that $\ell^2/dt$ goes to a constant, which implies that $dt$ will be of the same order as $\ell^2$. It then makes sense to take the expansion in Eq. (4.46) simultaneously to second order in $\ell$ and to first order in $dt$.

**Diffusion equation:** Taking the limit $dt$ and $\ell$ very small as $\ell^2/dt$ goes to a constant finally gives the one-dimensional diffusion equation:

$$\frac{\partial P}{\partial t}(x,t) = D\,\frac{\partial^2 P}{\partial x^2}(x,t)\,. \tag{4.48}$$

in which $D = \ell^2/(2\,dt)$ can be identified as the diffusion constant.

**A21.** More generally, for any values of $p$, we can restart from Eq. (4.46), and after simplying and reorganizing the terms, we get:

$$\frac{\partial P}{\partial t}(x,t)\,dt + o(dt) = (1-2p)\,\ell\,\frac{\partial P}{\partial x}(x,t) + \frac{\ell^2}{2}\,\frac{\partial^2 P}{\partial x^2}(x,t) + o(\ell^2)\,. \tag{4.49}$$

**Fokker-Planck equation:** For small values of $\ell$ and $dt$, we therefore obtain the equation:

$$\frac{\partial P}{\partial t}(x,t) = -v\,\frac{\partial P}{\partial x}(x,t) + D\,\frac{\partial^2 P}{\partial x^2}(x,t)\,, \tag{4.50}$$

which is a one-dimensional Fokker-Planck equation, with a constant diffusion coefficient $D = \ell^2/(2\,dt)$ and a constant drift $v = (2p-1)\,\ell/dt$.

**Comments:**

- The diffusion constant $D$ is the same as in the diffusion equation (4.48). In fact if we take $p = 1/2$ we obtain $v = 0$ and recover Eq. (4.48).
- For $p > 1/2$, the drift is positive $v > 0$, which means that the walker would tend to drift towards larger values of $x$ (towards the right of the $x$-axis). This is compatible with a probability to jump to the right that is larger than one half, $p > 1/2$. Similarly, if $p < 1/2$, then the walker would tend to drift towards the left.

### 4.4.3 (⋆) Real space renormalization

## Renormalization in d=1 Ising model

**A1.**

$$\begin{aligned} e^{KS_iS_{i+1}} &= \frac{1}{2}\left(e^K + e^{-K} + (e^K - e^{-K})S_iS_{i+1}\right) \quad \text{since} \quad S_iS_{i+1} = \pm 1 \\ &= \cosh(K) + \sinh(K)\,S_iS_{i+1} \\ &= \cosh(K)\,(1 + \tanh(K)\,S_iS_{i+1})\,. \end{aligned} \tag{4.51}$$

Another way is to directly use the relation $\exp(x) = \cosh(x) + \sinh(x)$ and the facts that:
– $\cosh(s_i\, x) = \cosh(x)$, because $s_i \in \{-1, +1\}$ and cosh is an even function;
– $\sinh(s_i\, x) = s_i\, \sinh(x)$, because $s_i \in \{-1, +1\}$ and sinh is an odd function.

**A2.**

$$\begin{aligned}\sum_{S_{i+1}=\pm 1} e^{KS_iS_{i+1}+S_{i+1}S_{i+2}} &= \cosh^2(K)(1+tS_iS_{i+1})(1+tS_{i+1}S_{i+2}) \\ &= \sum_{S_{i+1}=\pm 1} \cosh^2(K)\left(1+tS_iS_{i+1}+tS_{i+1}S_{i+2}+t^2S_iS_{i+2}\right) \\ &= 2\cosh^2(K)(1+t^2S_iS_{i+2}).\end{aligned}$$

**A3.**

$$\begin{aligned}\sum_{S_3,S_4} e^{K\sigma_1S_3}e^{KS_3S_4}e^{KS_4\sigma_2} &= (\cosh K)^3(1+t\sigma_1S_3)(1+tS_3S_4)(1+tS_4\sigma_2) \\ &= 2^2(\cosh K)^3(1+t^3\sigma_1\sigma_2). \qquad (4.52)\end{aligned}$$

Besides the multiplicative normalization constant (independent of the spins), (4.52) is the same as (4.51) and so we can define a new Boltzmann weight $W(\sigma_1, \sigma_2; K')$ of the block spins $\sigma_1$ and $\sigma_2$ with

$$t' \equiv \tanh K' = t^3 \equiv (\tanh K)^3 \qquad (4.53)$$

$$\Leftrightarrow K' = \tanh^{-1}[(\tanh K)^3] \qquad (4.54)$$

such that the new Hamiltonian of the system is given by

$$H(\sigma_i; K') = A(K) - K' \sum_i \sigma_i\sigma_{i+1} \qquad (4.55)$$

where $A(K)$ comes from the multiplicative factors in (4.52) and contributes only to the Free energy of the system.

**A4.** At the fixed points, the coupling constants will no longer change, which translates to

$$t' = t \qquad (4.56)$$

The intersections of the above line with the recursion relation (4.53), will give the location of these fixed points, and using the Cobweb plots we can deduce their stability, as shown in Fig. 4.3.

**Note**: We have defined $K = \frac{J}{k_B T}$. This means approaching $t = 0$, or equivalently $K = 0$, corresponds to approaching the high temperature phase, while approaching $t \to 1$ (i.e., $K \to \infty$), corresponds to $T \to 0$.

Since the stable fixed point is at $t \equiv \tanh(K) = 0 \Leftrightarrow K = 0$, when starting from any point $t_0 \neq 1$, we conclude that the **1d Ising model is always in its disordered phase**, with non interacting spins ($J = 0$). A system with non-interacting spins is self-similar with correlation length $\xi = 0$ and qualitatively it looks the same on all length scales. Quantitatively, the probability distribution of microstates is invariant under the renormalization transformation.

Overall, the renormalization transformation correctly predicts that there is no phase transition in the $d = 1$ Ising model.

**A5.** The middle terms, containing the spins to be summed over, will vanish like in the case of one, or two intermediate spins, and the only term that will survive, besides the identity, will be the last one containing the 'boundary spins', with a coefficient $t^{\text{len( summed spins )}+1}$. Hence, we get the following recursion relation:

$$\sum_{S_1,\ldots,S_{k-1}} e^{K(S_0S_1+\cdots+S_{k-1}S_k)} = 2^{(k-1)}\cosh^k(K)\left(1+t^k\, S_0S_k\right). \qquad (4.57)$$

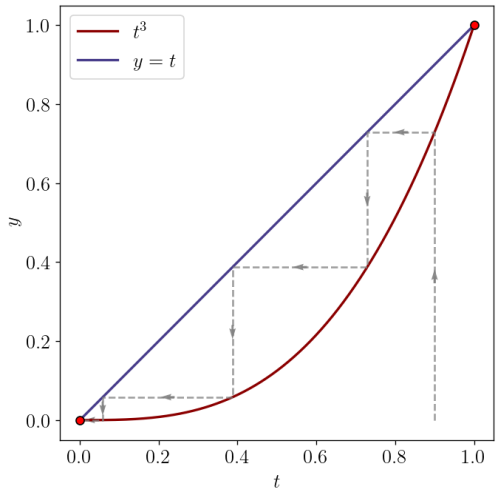



Figure 4.3: Cobweb plot, showing the fixed point $(0, 0)$ of the recursion relation when starting from $t_0 = 0.9$. $(1, 1)$ is an unstable fixed point.

Again, Besides the multiplicative normalization constant (independent of the spins), (4.57) is the same as (4.51) and so we can define a new Boltzmann weight $W(S_0, S_k; J')$ with

$$t' \equiv \tanh(K') = t^k = \tanh^k(K)\,. \tag{4.58}$$

We see that the mapping has two fixed points at $t_1^* = 0$ and $t_2^* = 1$. The first is an attractive fixed point and the second repulsive; Unless $t$ is exaclty $t = 1$, each iteration moves the values of $t$ towards the origin.

## Renormalization in $d = 2$ Ising model

**A 2.1.** Set $K \to 2K$ and $k = 2$ in eq. (4.58)

**A 2.2.** At the fixed points, the couplings will no longer change, so $t' = t$. We should therefore look for the intersections of the curves

$$y = \frac{4t^2}{(1 + t^2)} \quad \text{and} \quad y = t\,. \tag{4.59}$$

This time there are three intersections $\Rightarrow$ three fixed points. If we use the diagram as a Cobweb plot, we see that the points $t = 0, 1$ are stable fixed points, that correspond to $K = 0$ (high T) and $K \to \infty$ (low T) respectively. This implies that there must be also a critical point at a non-trivial $t_c$ ($K_c$), for which the 2d Ising model undergoes a **phase transition**.

With the current choice of decimation scheme, $t_c \simeq 0.2956$, which corresponds to $K_c = \frac{J}{k_B T_C} \simeq 0.30469*$. The exact theoretical value obtained by Onsager (1944) is $K_c \simeq 0.44$ (Onsager: $k_B\, T_c \simeq 2.26\, J$).

**Note:** Different decimation schemes will lead to slightly different values for $K_c$. However, the qualitative behavior of the system predicted by any RG transformation, will remain the same.

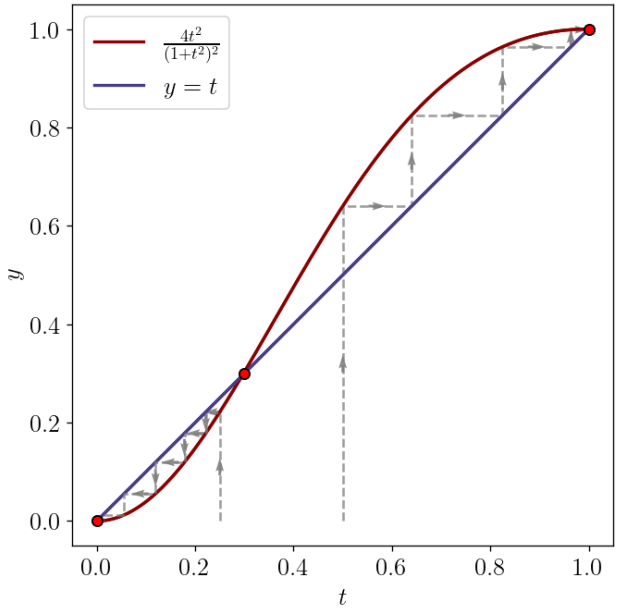



Figure 4.4: Graphical solution of the recursion relation and Cobweb plots showing the stability of the three fixed points.

# Chapter 5

# Examples of non-equilibrium critical phenomena

Wout Merbis, Ebo Peerbooms, Clélia de Mulatier

In the first part of the course, phase transitions were studied in the equilibrium framework, and the goal was to characterize the order arising below or above a critical value of the control parameters. Here we would like to investigate how this order is formed dynamically, i.e. when the system is evolving out-of-equilibrium.

**Symbol "(⋆)":** Questions and exercises indicated with a (⋆) are optional. No worries if you don't have time to try to solve them, or if you don't manage to solve them on your own.

**The symbol "</>":** Indicates optional questions with numerical simulation.

## 5.1 Epidemic spreading

In the compartmental SIS model, the fraction of infected in a population is changing in time as:

$$\frac{d\,\mathrm{i}(t)}{dt} = (\beta - \mu)\,\mathrm{i}(t) - \beta\,\mathrm{i}(t)^2\,, \tag{5.1}$$

where $\mu$ is the recovery rate and $\beta$ is the transmission rate.

**Q1.** What are the steady solutions to (5.1)? Analyze the stability of the steady state solutions by linearizing the equation (5.1) around the steady states $i_0$ as $i(t) = i_0 + \epsilon(t)$. Depending on the value of $R_0 = \beta/\mu$, which steady states are stable under which circumstances?

> The **SIR model** has individuals recover into a recovered compartment. The rate equations are:
>
> $$\frac{ds(t)}{dt} = -\beta\,\mathrm{i}(t)s(t)\,, \qquad \frac{\mathrm{d}\,\mathrm{i}(t)}{\mathrm{d}t} = \beta\,\mathrm{i}(t)\,\mathrm{s}(t) - \mu\,\mathrm{i}(t)\,, \qquad \frac{\mathrm{d}\,\mathrm{r}(t)}{\mathrm{d}t} = \mu\,\mathrm{i}(t)\,, \tag{5.2}$$
>
> where $s, i, r$ are the fractions of susceptible, infected and recovered respectively, such that $1 = \mathrm{s}(t) + \mathrm{i}(t) + \mathrm{r}(t)$ at all times. At late times ($t \to \infty$), the total fraction of recovered individuals $r_\infty$ is a control parameter for the epidemic outbreak. Here we will derive a self-consistency relation for $r_\infty$ using the above equations.

**Q2.** Show that:

$$C(t) = \mathrm{s}(t)\exp\left(\frac{\beta}{\mu}\,\mathrm{r}(t)\right), \tag{5.3}$$

is a **constant of motion**, i.e. show that $\frac{\mathrm{d}}{\mathrm{d}t}C(t) = 0$ by using (5.2).

**Q3.** Using (5.3), derive an expression for $r_\infty$ by equating $C(t=0) = C(t=\infty)$. Choose as initial conditions a completely susceptible population and use the fact that at late times $i_\infty = 0$.

In the SIS model on a network, before taking a mean-field approximation, the probability of any node $i$ being infected ($X_i(t) = 1$) depends on the probability of it forming an SI-pair with a neighboring node in the network

$$\frac{d\mathbb{P}[X_i(t)=1]}{dt} = -\mu\,\mathbb{P}[X_i(t)=1] + \beta \sum_j A^{ij}\mathbb{P}[X_i(t)=0; X_j(t)=1]\,. \tag{5.4}$$

**Q4.** Derive the dynamical equation for the joint probability $\mathbb{P}[X_i(t)=0; X_j(t)=1]$ of node $i$ being susceptible and node $j$ being infected. To do so, think about all the different ways that $\mathbb{P}[X_i(t)=0; X_j(t)=1]$ can change, when the dynamical rules of the system are given by recovery of single nodes from infected to susceptible with rate $\mu$ and infection by neighboring nodes with rate $\beta$.

In the heterogeneous mean-field models on networks, we bin nodes of the same degree $k$ (i.e. nodes with the same number of neighbors). The fraction of infected nodes of degree $k$ is $i_k(t)$ and follows (for the SIS model):

$$\frac{di_k(t)}{dt} = -\mu i_k(t) + \beta k\,(1-i_k(t))\Theta_k(t) \tag{5.5}$$

where $\Theta_k(t)$ is the fraction of infected neighbors of nodes of degree $k$. Assuming the absence of degree correlations, this quantity is independent of $k$ and given by:

$$\Theta_k(t) = \Theta(t) = \sum_{k'} \frac{k'-1}{\langle k\rangle} P(k') i_{k'}(t) \tag{5.6}$$

where $\langle k\rangle = \sum_k kP(k)$ is the mean degree and $P(k)$ the degree distribution.

**Q5.** Assuming that initially $i_k(t) \ll 1$ for all $k$, neglect quadratic terms in $i(t)$ and derive a linearized equation from (5.5). Use this to derive the rate equation for $\Theta(t)$. How will the number of infected grow with time? What does this mean for the epidemic threshold?

## 5.2 The (linear) voter model

This exercise is inspired from the Chapter on "Spin dynamics" (Chapter 8) of the book "A Kinetic View of Statistical Physics" by P. L. Krapivsky, S. Redner, and E. Ben Naim [1].

> The **linear voter model** is the simplest kinetic spin system. It has the advantage that, on a lattice structure, it is exactly soluble in any dimensions. The model describes how consensus emerges in a population of individuals that have no firmly fixed opinion. Individuals randomly take the opinion of one of their neighbors. A finite population of such voters eventually achieves consensus in a time that depends on the system size, the spatial dimension, and the number of opinions. Note that a consensus (i.e. everyone has the same opinion) will eventually be reached as it is the only absorbing state of the system.
>
> In the voter model, social interactions between voters are represented by a graph, in which each node represents an individual and links between two nodes indicate which individuals interact. This graph could be a regular lattice in $d$ dimension, or it could be any other type of graph (an Erdös-Rényi random graph, or a graph with broad degree distribution, for example). Let us label the nodes of the graph with an index $i \in \{1, 2, \ldots, N\}$. The opinion of each individual $i$ is indicated by its state $s_i$ , which can take discrete values. For simplicity, we restricted the dynamics to just two opinions (*for* or *against* some issue), such that the $s_i$ are binary variables: $s_i = \pm 1$. For instance, individuals in the state $s_i = +1$ could be "Democrats", and individuals in the state $s_i = -1$ could be "Republicans" (such as in the homework exercise of H1). A state of the system of the $N$ voters is denoted by $\boldsymbol{s} = (s_1, s_2, \ldots, s_N)$.

The dynamics of the voter model is very simple: each voter has no confidence and looks randomly to one of its neighbors to decide what to do. At each time step:

1. Pick a random voter $i$;
2. Select a random neighbor $j$ of the voter $i$, and align the opinion of $i$ to the opinion of $j$, i.e set $s_i = s_j$;
3. Repeat steps 1 & 2 *ad infinitum*, or until a consensus is reached.

Note that in step 2, the voter $i$ changes opinion only when its neighbor $j$ has opposite opinion. Figure 5.2 shows the evolution of the voter model on a square lattice with two different types of initial conditions. In the bottom row, starting from a random initial condition with equal densities of up and down states, one can observe how the system tends to organize in domains with single opinion as time increases.

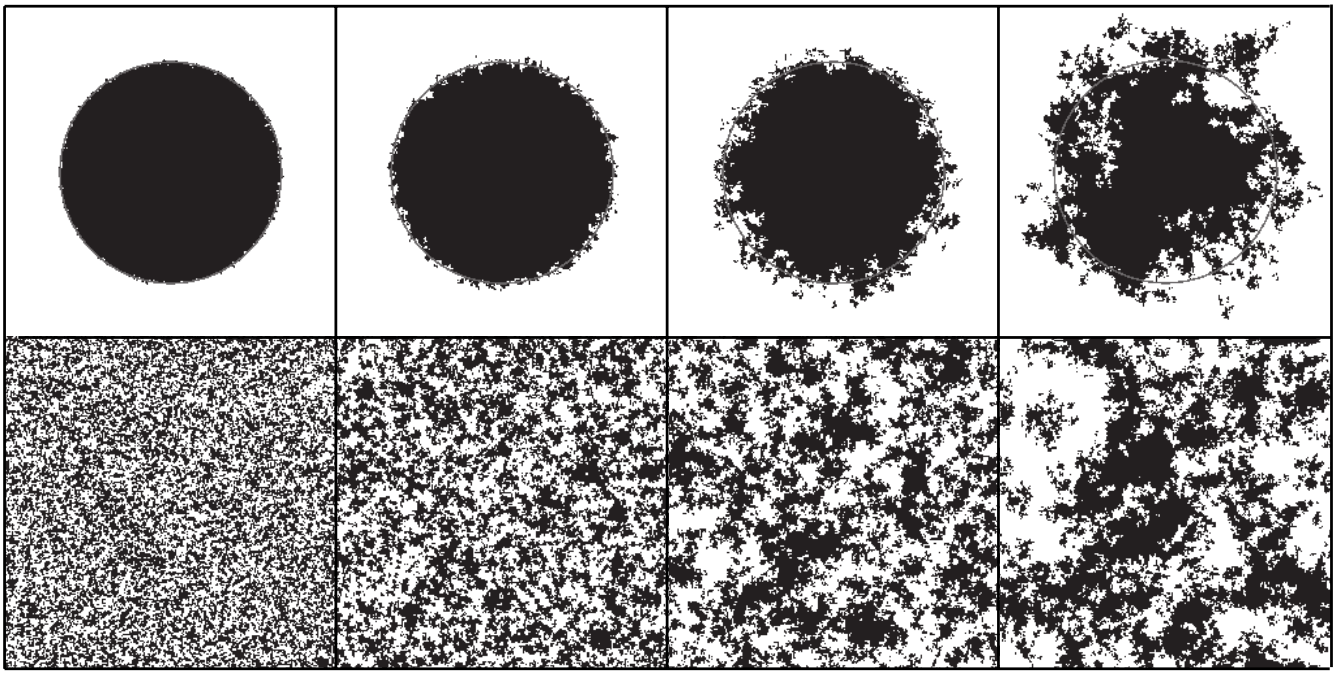


Figure 5.1: Figure is from Dornic et al. [2]: Evolution of the voter model on a 256 × 256 square lattice. Black and white pixels indicate the different opinion states. The top row shows snapshots at times $t = 4$, 16, 64, and 256 starting with an initial circular bubble of one opinion of radius 180 in a sea of the opposite opinion. The bottom row shows the same evolution starting with a random initial condition with equal densities of the two opinion states.

**Q0. (bonus) (</>)** Starting from one of the two initial conditions of Fig. 5.2, can you perform a Monte Carlo simulation of this simple voter model on a 2D lattice? Do you recover a behavior similar to that of Fig. 5.2?

**Q1.** Can you show that, at any time step $n$ of the simulation described above, the probability that the voter $i$ changes its opinion from $s_i$ to $-s_i$ is equal to:

$$\mathbb{P}[(s_i, n) \to (-s_i, n+1)] = \frac{1}{N} \frac{q_i[\boldsymbol{s}(n)]}{k_i}, \tag{5.7}$$

where $N$ is the total number of nodes on the graph, $k_i$ is the total number of neighbors of node $i$ (i.e. the multiplicity of node $i$ in the graph), and $q_i[\boldsymbol{s}(n)]$ is the number of these neighbors that disagree with voter $i$ (i.e. that hold the opinion opposite to $s_i$) at the $n$-th step of the simulation. Note that $q_i$ depends on the state $\boldsymbol{s}(n)$ of the system at the $n$-th step.

**Q2.** Let us define a small time interval $dt = 1/N$. The simulation starts at time $t = 0$, and each step of the simulation last a small time interval $dt$. After $n$ steps, the time is $t = n\,dt$. Can you re-write Eq. (5.7) as a function of $t$ instead of $n$? Deduce that the rate at which a voter $i$ switches opinion is equal to the fraction of its neighbors that hold the opposite opinion, $w_i[\boldsymbol{s}] = q_i[\boldsymbol{s}]/k_i$, and changes with time. Observe that the rate $w_i(\boldsymbol{s})$ depends on the state of the system, which evolves in time. Each voter switch its opinion following an **inhomogeneous Poisson process**.

**Q3.** What is the value of the product $s_i s_j$ if the two voters $i$ and $j$ have the same opinion? if they disagree? Using this observation, can you show that $w_i(\boldsymbol{s})$ can be re-written as:

$$w_i(\boldsymbol{s}) = \frac{1}{2}\left[1 - \frac{s_i}{k_i}\sum_{j\in\langle i\rangle} s_j\right], \tag{5.8}$$

where $j \in \langle i\rangle$ denotes a sum over all the neighbors $j$ of $i$? Tips: you can start by computing the number $q_i[\boldsymbol{s}]$ of neighbors of $i$ that disagree with $i$ when the system is in a state $\boldsymbol{s}$.

**Q4. Master equation.** In principle, to solve the voter model and describe the dynamical evolution of the system, we would need to compute the probability $P(\boldsymbol{s}, t)$ of the system to be in the state $\boldsymbol{s}$ at any time $t$. Can you show that $P(\boldsymbol{s}, t)$ satisfies the master equation:

$$\frac{dP(\boldsymbol{s},t)}{dt} = \sum_{i=1}^{N} P(\boldsymbol{s}_{-i},t)\, w_i(\boldsymbol{s}_{-i}) - P(\boldsymbol{s},t)\sum_{i=1}^{N} w_i(\boldsymbol{s}), \tag{5.9}$$

in which $\boldsymbol{s}_{-i}$ denotes the state of the voter identical to the state $\boldsymbol{s}$ but in which spin $i$ has been flipped?

**Q5.** In principle, the master equation (5.9) can be used to derive equations for all the moments of the probability distribution, namely all multi-spin correlation functions of the form $\langle s_i s_j \cdots s_k\rangle$ where the brackets denote the ensemble average $\langle A(\boldsymbol{s})\rangle = \sum_{\boldsymbol{s}} A(\boldsymbol{s})P(\boldsymbol{s})$. Let us consider the simplest of such moments: the mean $\langle s_i\rangle$. The equation for $\langle s_i\rangle$ could be obtained from the master equation, however it is quite cumbersome. Let us take a more direct approach. Assuming that the system is in a state $\boldsymbol{s}$ at time $t$, depending on the value of $s_i(t)$, which values can be taken by $s_i(t+dt)$ at time $t+dt$, and with which probability? Can you then show that the mean opinion of $s_i$ evolves as:

$$\frac{d\langle s_i\rangle}{dt} = -2\langle s_i\, w_i(\boldsymbol{s})\rangle \ ? \tag{5.10}$$

**Q6.** Replacing $w_i(\boldsymbol{s})$ by its value, can you show that:

$$\frac{d\langle s_i\rangle}{dt} = -\langle s_i\rangle + \frac{1}{k_i}\sum_{j\in<i>}\langle s_j\rangle \ ? \tag{5.11}$$

**Q7.** As for the Ising model, we can define the total magnetization (per spin) of the system as: $\sum_{i=1}^{N} s_i/N$. We denote by $m$ the mean magnetization: $m = \sum_{i=1}^{N}\langle s_i\rangle/N$. This represents the average orientation of the system. For instance, if the state $s_i = +1$ represents a "Democrat", and the state $s_i = -1$ represents a "Republican", then the sign of $m$ would represent the average political orientation of the whole population.

For simplicity, let us consider graphs in which all the nodes have the same multiplicity $k$, i.e. for all node $i$, $k_i = k$. Using Eq. (5.11), can you show that the mean magnetization is conserved over time, i.e. that:

$$\frac{dm}{dt} = 0 \ ? \tag{5.12}$$

Notice that while the magnetization of a specific system does change in a single update event, the average over all sites and over all trajectories of the dynamics is conserved. The consequences of this conserved mean magnetization are profound.

**Q8.** Consider a finite system with an **initial** fraction $\rho_0^+$ of voters in the +1 state and a fraction $1-\rho_0^+$ in the −1 state, so that the initial magnetization is $m_0 = 2\rho_0^+ - 1$. After a very long time, the system will reach a consensus, i.e. the final total magnetization will be +1 (all voters will be +1), or the final total magnetization will be −1 (all voters will be −1). Let us define $E(\rho)$ as the probability that the system reach the "+1" consensus. Using the conservation of the mean magnetization, can you compute the value of $E(\rho)$ as a function of the original fraction $\rho_0^+$ of +1 voters in the initial state of the system?

**Q9 (bonus) (</>).** Write an algorithm that implements the dynamics of the voter model and verify the properties you just derived using many realizations of the dynamics.

## References

[1] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A kinetic view of statistical physics*. Cambridge University Press, 2010.

[2] I. Dornic, H. Chaté, J. Chave, and H. Hinrichsen, "Critical coarsening without surface tension: The universality class of the voter model," *Physical Review Letters*, vol. 87, no. 4, p. 045 701, 2001.

## 5.3 Stability analysis

Bifurcation theory is a mathematical framework closely related to the study of phase transitions. The difference is that in bifurcation theory the object of study are dynamical systems , while for phase transitions we look at the behaviour of the free energy. For one-dimensional systems, it is possible to define a bifurcation potential $U(x)$ such that

$$\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{\mathrm{d}U(x)}{\mathrm{d}x} \,. \tag{5.13}$$

The local minima of such a potential are equivalent to the steady states of the corresponding dynamical system. As an example, consider

$$\dot{x} = \mu x^2 + x \,. \tag{5.14}$$

In bifurcation analysis, $\mu$ is considered to be parameter we can change . As a function of $\mu$ interesting things can happen to the nature of the solutions. The steady state solution is given by setting $\dot{x} = 0$, which gives $\mu x^2 + x = 0$. The steady-state solutions are therefore $x_0 = 0$ and $x_0 = -\frac{1}{\mu}$. To find the stability of this solution, one typically expands around the steady-state solution by writing $x(t) = x_0 + \varepsilon(t)$ and plugging this back into the equation. For this system we find

$$\frac{d}{dt}(x+\varepsilon) = \dot{x} + \dot{\varepsilon} = \mu(x_0+\varepsilon)^2 + (x_0+\varepsilon) = \mu x_0^2 + x_0 + \varepsilon(2\mu x_0 + 1) + \mu\varepsilon^2 \,. \tag{5.15}$$

Since $\dot{x} = \mu x_0^2 + x_0$ is a solution to $\dot{x} = 0$, the only remaining terms are those involving $\varepsilon$. Since $\varepsilon$ is assumed to be small, we only keep the terms proportional to $\varepsilon$. This is called linearizing the equation . We now have an equation for the evolution of small perturbations to the steady-state solution in the form of

$$\dot{\varepsilon} = (2\mu x_0 + 1)\varepsilon \,. \tag{5.16}$$

Solving for $\varepsilon(t)$ we find

$$\varepsilon(t) = \varepsilon(0)\exp([2\mu x_0 + 1]t) \,. \tag{5.17}$$

Depending on the sign of $2\mu x_0 + 1$, the perturbation either decays to zero or blows up to infinity. Plugging in $x_0 = 0$, we have

$$\varepsilon(t) = \varepsilon(0)e^t \,. \tag{5.18}$$

Thus regardless of the value of $\mu$ the origin is an unstable solution. For $x_0 = -\frac{1}{\mu}$ we get

$$\varepsilon(t) = \varepsilon(0)e^{-t} \,. \tag{5.19}$$

Again, regardless of the value of $\mu$, the solution $x_0 = -\frac{1}{\mu}$ is a stable steady-state solution. This example illustrates the procedure of stability analysis. Using it, we found that no bifurcation occurs. For the following systems we will see that more interesting behaviour is possible.

**Q1.** For each of the following dynamical systems

$$a)\ \ \dot{x} = \mu - x^2 \qquad b)\ \ \dot{x} = \mu x - x^2 \qquad c)\ \ \dot{x} = \mu x - x^3 \tag{5.20}$$

**a.** Find the steady-state solutions.

**b.** Find the value of $\mu$ at which a bifurcation occurs.

**c.** Make a sketch or plot of the steady-state solution branches as a function of $\mu$.

**d.** Write down the stability of all of the solution branches on either side of the bifurcation point.

**e.** Find a suitable bifurcation potential $U(x)$ that reproduces the dynamical system. Is the potential a suitable candidate for a physical free energy?

**Q2.** Now consider the dynamical system

$$\dot{x} = \mu x^2 - x^4 \tag{5.21}$$

Clearly, $x_0 = 0$ is a fixed point of this system.

**a.** Do a linear stability analysis of the fixed point $x_0 = 0$ and comment on the result. How is this possible?

**b.** Think of a way to solve the problem you found in the previous question and solve the differential equation that you obtain.

**c.** Find the stability of the resulting solution. In particular, what happens at $\mu = 0$?

**d.** For several positive and negative values of $\mu$, sketch or plot $\dot{x}$ as a function of $x$. What can this plot tell you about the stability of the steady-state solutions?

## 5.4 Solutions

### 5.4.1 Epidemic spreading

**A1.** The steady states are solutions to $\frac{d}{dt}i(t) = 0$. There are two:

$$i_0 = 0\,, \qquad i_1 = 1 - \frac{1}{R_0}\,, \qquad \text{with: } R_0 = \frac{\beta}{\mu}\,. \tag{5.22}$$

- If we expand the equation around $i_0$ as $i(t) = i_0 + \epsilon(t)$, we see that, for small positive $\epsilon(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{i}(t) = \frac{\mathrm{d}\epsilon(t)}{\mathrm{d}t} \sim (\beta - \mu)\epsilon(t)\,. \tag{5.23}$$

  This means that the fluctuations $\epsilon(t)$ will grow in size if $(\beta - \mu) > 0$ and decay in size when $\beta - \mu < 0$, or phrased in terms of $R_0$: when $R_0 < 1$ the steady state $i_0 = 0$ is stable, and when $R_0 > 1$ it is unstable.

- Likewise, we can expand the equation around the endemic steady state $i_1 = 1 - \frac{1}{R_0}$ as

$$\frac{\mathrm{d}\epsilon(t)}{\mathrm{d}t} = \mu(R_0 - 1)\left(1 - \frac{1}{R_0} + \epsilon(t)\right) - \beta\left(1 - \frac{1}{R_0} + \epsilon(t)\right)^2 \sim -\beta\left(1 - \frac{1}{R_0}\right)\epsilon(t)\,, \tag{5.24}$$

  where we use that $\frac{\mathrm{d}i_1}{\mathrm{d}t} = 0$.

  When $R_0 > 1$, we have $(1 - \frac{1}{R_0}) > 0$ and fluctuation around the endemic steady state will decay in size. When $R_0 < 1$ the fluctuation around the endemic steady state will grow in size and the system will flow to the other stable steady state at $i = i_0 = 0$.

**A2.** The solution is straight forward by application of the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}t}C(t) = \frac{ds(t)}{dt}\exp\left(\frac{\beta}{\mu}\,\mathrm{r}(t)\right) + \mathrm{s}(t)\frac{\beta}{\mu}\frac{\mathrm{d}\,\mathrm{r}(t)}{\mathrm{d}t}\exp\left(\frac{\beta}{\mu}\,\mathrm{r}(t)\right) \tag{5.25}$$

$$= -\beta\,\mathrm{i}(t)\,\mathrm{s}(t)\exp\left(\frac{\beta}{\mu}\,\mathrm{r}(t)\right) + \beta\,\mathrm{i}(t)\,\mathrm{s}(t)\exp\left(\frac{\beta}{\mu}\,\mathrm{r}(t)\right) = 0\,. \tag{5.26}$$

**A3.** Because $C(t)$ is actually constant in time, we can equate $C(t = 0) = C(t = \infty)$ to obtain

$$s(0)\exp\left(\frac{\beta}{\mu}r(0)\right) = s_\infty \exp\left(\frac{\beta}{\mu}r_\infty\right). \tag{5.27}$$

At $t = 0$ we choose a completely susceptible population $s(0) = 1$, such that $r(0) = 0$. At late times, the infection has died out, so $i_\infty = 0$ and by conservation of total probabilities $s_\infty = 1 - i_\infty - r_\infty = 1 - r_\infty$. We hence obtain an expression solely in terms of $r_\infty$ and $R_0$

$$1 = (1 - r_\infty)e^{R_0 r_\infty}\,, \qquad \text{or: } e^{-R_0 r_\infty} = 1 - r_\infty\,. \tag{5.28}$$

**A4.** We have to analyze all possible ways in which SI pairs of nodes $i$ and $j$ can be created or destroyed. There are five possibilities, two contribute positively and three contribute negatively:

**Positive contributions:**

- The node pair $i$ and $j$ were both infected and node $i$ recovers to form the SI pair. This increases the probability of being $i$ and $j$ forming an SI pair and happens with probability $\mu\mathbb{P}[X_i(t) = 1; X_j(t) = 1]$
- The node pair $i$ and $j$ were both susceptible, then node $j$ became infected by one of its neighbours. This also contributes positively and happens with probability $\beta\sum_k A^{jk}\mathbb{P}[X_i(t) = 1; X_j(t) = 0; X_k(t) = 1]$.

**Negative contributions:**

- The node pair $i$ and $j$ are already in the SI configuration, and node $j$ infects node $i$. This happens with probability $\beta A^{ij}\mathbb{P}[X_i(t) = 0; X_j(t) = 1]$.

– The node pair $i$ and $j$ are already in the SI configuration, and node $j$ recovers. This happens with probability $\mu\mathbb{P}[X_i(t) = 0; X_j(t) = 1]$.

– The node pair $i$ and $j$ are already in the SI configuration, and node $i$ is infected by another neighbor $k$. This happens with probability $\beta \sum_{k\neq j} A^{ik}\mathbb{P}[X_i(t) = 0; X_j(t) = 1; X_k(t) = 1]$.

Putting it all together we get the equation

$$\begin{aligned}\frac{d}{dt}\mathbb{P}[X_i(t) = 0; X_j(t) = 1] &= \mu\mathbb{P}[X_i(t) = 1; X_j(t) = 1] - \beta A^{ij}\mathbb{P}[X_i(t) = 0; X_j(t) = 1] \\ &\quad - \mu\mathbb{P}[X_i(t) = 0; X_j(t) = 1] + \beta \sum_k A^{jk}\mathbb{P}[X_i(t) = 1; X_j(t) = 0; X_k(t) = 1] \\ &- \beta \sum_{k\neq j} A^{ik}\mathbb{P}[X_i(t) = 0; X_j(t) = 1; X_k(t) = 1]\,.\end{aligned} \tag{5.29}$$

**A5.** The linear approximation for small fractions of infected gives

$$\frac{d}{dt} i_k(t) = -\mu i_k(t) + \beta k \Theta(t)\,, \tag{5.30}$$

with

$$\Theta(t) = \sum_{k'} \frac{k' - 1}{\langle k\rangle} P(k') i_{k'}(t)\,. \tag{5.31}$$

Taking the time derivative of $\Theta(t)$ and using (5.30) gives:

$$\frac{d}{dt}\Theta(t) = \frac{1}{\langle k\rangle} \sum_{k'} (k' - 1)P(k')\left(-\mu i_{k'}(t) + \beta k' \Theta(t)\right) \tag{5.32}$$

$$= -\mu\Theta(t) + \frac{\beta}{\langle k\rangle} \sum_{k'} (k'^2 - k')P(k')\Theta(t) \tag{5.33}$$

$$= -\mu\Theta(t) + \frac{\beta}{\langle k\rangle} \left(\langle k^2\rangle - \langle k'\rangle\right)\Theta(t)\,. \tag{5.34}$$

$$\tag{5.35}$$

This implies that

$$\Theta(t) \sim e^{t/\tau}\,, \qquad \text{with: } \tau^{-1} = \mu\left(\frac{\beta}{\mu}\frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle} - 1\right)\,. \tag{5.36}$$

And so the basic reproductive number $R_0$ is defined in terms of transmission and recovery rates and the network degree fluctuations as

$$R_0 = \frac{\beta}{\mu}\frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle}\,. \tag{5.37}$$

For large $\langle k^2\rangle$, the reproduction number is increased, meaning that the epidemic threshold at $R_0 = 1$ is lowered!

### 5.4.2 Voter model

**A1.** At each time step:

– the probability that the voter $i$ is chosen is $1/N$, where $N$ is the number of voters (= the number of nodes);

– if selected, the probability that the voter $i$ changes its opinion is then equal to the probability that the randomly chosen neighbor $j$ has the opposite opinion; this probability is equal to the fraction of neighbors of $i$ that have opposite opinion, which is the ratio $q_i(n)/k_i$ of the number of neighbors with opposite opinion $q_i(n)$ and the total number of neighbors $k_i$.

This finally gives that the probability that the voter $i$ changes its opinion between the time steps $n$ and $(n+1)$ in Eq. (5.7).

**A2.** Re-writing Eq. (5.7) as a function of the time $t$, we get:

$$\mathbb{P}[s_i(t) \to -s_i(t+dt)] = \frac{q_i[\boldsymbol{s}(t)]}{k_i}\, dt = w_i[\boldsymbol{s}(t)]\, dt\,. \tag{5.38}$$

The voter $i$ therefore changes its opinion with a rate $w_i[\boldsymbol{s}(t)] = \frac{q_i[\boldsymbol{s}(t)]}{k_i}$, which may evolve in time, while the opinions of the neighbors of $i$ are still changing.

> Observe that the rate $w_i(\boldsymbol{s})$ depends on the state of the system, which evolves in time. Each voter switch its opinion following an **inhomogeneous Poisson process**.

**A3.** $s_i s_j = -1$ if the two voters $i$ and $j$ have opposite opinions, and $s_i s_j = +1$ if the two voters have the same opinion. Using this, we can define the following marker function :

$$\delta(s_i, s_j) = \frac{1}{2}(1 - s_i s_j)\,, \tag{5.39}$$

which is equal to 1 if $i$ and $j$ have opposite opinions, and 0 if they have the same opinion. The number of neighbors of $i$ that have an opinion opposite to $s_i$ is therefore given by:

$$q_i(\boldsymbol{s}) = \sum_{j \in <i>} \delta(s_i, s_j) \tag{5.40}$$

$$= \frac{1}{2} \sum_{j \in <i>} (1 - s_i s_j) = \frac{1}{2}\left[k_i - s_i \sum_{j \in <i>} s_j\right], \tag{5.41}$$

where we used that $\sum_{j \in <i>} 1 = k_i$ is equal to the total number of numbers of $i$. Using that $w_i(\boldsymbol{s}) = q_i(\boldsymbol{s})/k_i$ finally gives Eq. (5.8).

**A4.** In the master equation, the loss term accounts for all possible transitions out of the state $\boldsymbol{s}$, while the gain term accounts for transitions to the state $\boldsymbol{s}$ from states in which a single spin differ. In more details: the probability that the system is in state $\boldsymbol{s}$ at time $t + dt$ can be connected to the states of the system at time $t$ in the following way:

$$P(\boldsymbol{s}, t+dt) = \underbrace{\sum_{i=1}^{N} P(\boldsymbol{s}_{-i}, t)\; w_i(\boldsymbol{s}_{-i})\, dt}_{(1)} + \underbrace{P(\boldsymbol{s}, t)\left[1 - \sum_{i=1}^{N} w_i(\boldsymbol{s})\, dt\right]}_{(2)}, \tag{5.42}$$

where

(1) is the probability that, at time $t$, the system was in a state $\boldsymbol{s}_{-i}$ that differs from the state $\boldsymbol{s}$ by only a single spin $s_i$, and that this spin flipped from $-s_i$ to $s_i$ during $dt$;

(2) is the probability that the system was already in state $\boldsymbol{s}$ at time $t$ and that no spin were flipped during $dt$ (i.e., no voter changed its opinion during $dt$).

Re-organizing the terms in Eq. (5.42), we get:

$$\frac{P(\boldsymbol{s}, t+dt) - P(\boldsymbol{s}, t)}{dt} = \underbrace{\sum_{i=1}^{N} P(\boldsymbol{s}_{-i}, t)\; w_i(\boldsymbol{s}_{-i})}_{(1)} + \underbrace{P(\boldsymbol{s}, t)\left[\sum_{i=1}^{N} w_i(\boldsymbol{s})\right]}_{(2)}, \tag{5.43}$$

which finally leads to the master equation (5.9).

**A5.** Assuming that the system is in a state $\boldsymbol{s}$ at time $t$, and knowing $s_i(t)$, the value of $s_i(t+dt)$ is given by:

$$s_i(t+dt) = \begin{cases} s_i(t) & \text{with probability } \; 1 - w_i(\boldsymbol{s},t)\,dt \\ -s_i(t) & \text{with probability } \; w_i(\boldsymbol{s},t)\,dt \end{cases} \tag{5.44}$$

i.e., the spin $i$ is flipped with probability $w_i(\boldsymbol{s},t)\,dt$ and remains unchanged with probability $(1 - w_i(\boldsymbol{s},t)\,dt)$. Note that in the case where the spin $i$ remains unchanged, other spins could have been flipped. Using this equation, the average value of $s_i(t+dt)$ at time $t+dt$ can be computed by summing over all possible state $\boldsymbol{s}$ in which the system could have been at time $t$, which gives:

$$\langle s_i(t+dt)\rangle = \sum_{\boldsymbol{s}} P(\boldsymbol{s},t)\, s_i(t)[1 - w_i(\boldsymbol{s},t)\,dt] + \sum_{\boldsymbol{s}} P(\boldsymbol{s},t)\,(-s_i(t))\, w_i(\boldsymbol{s},t)\,dt\,. \tag{5.45}$$

Cleaning up a bit:

$$\langle s_i(t+dt)\rangle = \sum_{\boldsymbol{s}} P(\boldsymbol{s},t)\, s_i(t) - 2\sum_{\boldsymbol{s}} P(\boldsymbol{s},t)\, s_i(t)\, w_i(\boldsymbol{s},t)\,dt\,, \tag{5.46}$$

$$= \langle s_i(t)\rangle - 2\,\langle s_i(t)\, w_i(\boldsymbol{s},t)\rangle\,dt \tag{5.47}$$

Re-organizing the terms finally leads to:

$$\frac{\langle s_i(t+dt)\rangle - \langle s_i(t)\rangle}{dt} = -2\,\langle s_i(t)\, w_i(\boldsymbol{s},t)\rangle \tag{5.48}$$

which corresponds to equation (5.10).

**A6.** Replacing $w_i(\boldsymbol{s})$ by its value in Eq. (5.10), one gets:

$$\frac{d\,\langle s_i\rangle}{dt} = -\left\langle s_i\left[1 - \frac{s_i}{k_i}\sum_{j\in<i>} s_j\right]\right\rangle \tag{5.49}$$

$$= -\langle s_i\rangle + \frac{1}{k_i}\left\langle s_i^2 \sum_{j\in<i>} s_j\right\rangle \tag{5.50}$$

$$= -\langle s_i\rangle + \frac{1}{k_i}\sum_{j\in<i>}\langle s_i\rangle\;, \tag{5.51}$$

where we used that $s_i^2 = 1$.

**A7.** To obtain the evolution of the average total magnetization in time, we must sum Eq. (5.11) over all the nodes:

$$\frac{dm(t)}{dt} = \sum_{i=1}^{N}\frac{d\,\langle s_i(t)\rangle}{dt} \tag{5.52}$$

$$= -\sum_{i=1}^{N}\langle s_i(t)\rangle + \frac{1}{k}\sum_{i=1}^{N}\sum_{j\in<i>}\langle s_i(t)\rangle \tag{5.53}$$

$$= -m(t) + \frac{1}{k}\sum_{i=1}^{N}\sum_{j\in<i>}\langle s_i(t)\rangle\;. \tag{5.54}$$

As each node has exactly $k$ neighbors, each of the term $\langle s_i\rangle$ is counted $k$ times in the double sum:

$$\sum_{i=1}^{N}\sum_{j\in<i>}\langle s_i(t)\rangle = k\sum_{i=1}^{N}\langle s_i(t)\rangle = k\,m(t)\,. \tag{5.55}$$

Replacing this result in Eq. (5.54), we finally get:

$$\frac{dm(t)}{dt} = -m(t) + \frac{1}{k}\, k\, m(t) = 0\,. \tag{5.56}$$

**A8.** As a consequence of Eq. (5.12), we have that $m(t)$ is a constant. The average magnetization over all sites and over all trajectories of the dynamics is constant throughout the evolution of the system, and therefore at all time $m(t) = m_0$. In particular, the final magnetization is equal to the initial magnetization: $m_\infty = m_0$.

The initial magnetization is given by:

$$m_0 = (+1) \times \rho_0^+ + (-1) \times (1 - \rho_0^+) = 2\rho_0^+ - 1\,. \tag{5.57}$$

At infinite time, the system will always reach a consensus, which is either everyone votes +1, or everyone votes −1. These two states are the absorbing states of the system, so as soon as the system is in one of them, it will stay there. Averaging over all possible realizations of the dynamics, the final magnetization can be computed as:

$$m_\infty = (+1) \times E(\rho) + (-1) \times (1 - E(\rho)) = 2\,E(\rho) - 1\,, \tag{5.58}$$

where $E(\rho)$ is the probability that the system reaches the "+1" consensus. Using that $m_\infty = m_0$, we finally obtain that the probability that the system reach the consensus "+1" is equal to the original fraction of +1 voters in the initial state, i.e.: $E(\rho) = \rho_0^+$.

### 5.4.3 Stability analysis

**A1.** **a. Saddle-node bifurcation.** For the first system $\dot{x} = \mu - x^2$, the steady-state solutions are given by

$$\dot{x} = \mu - x^2 = 0 \implies x_0 = \pm\sqrt{\mu}\,. \tag{5.59}$$

To find the bifurcation point, we expand around the steady-state solution. We let $x(t) = x_0 + \varepsilon(t)$ and plug this into the governing equation

$$\dot{x} = \cancel{\dot{x}_0} + \dot{\varepsilon} = \dot{\varepsilon} = \mu - (x_0 + \varepsilon)^2 = \mu - x_0^2 - 2x_0\varepsilon - \varepsilon^2 \approx -2x_0\varepsilon\,. \tag{5.60}$$

Note that we neglect terms quadratic in $\varepsilon$. The above computation results in an evolution equation for the perturbation $\varepsilon$,

$$\dot{\varepsilon} = -2x_0\varepsilon\,, \tag{5.61}$$

which is solved by

$$\varepsilon(t) = \varepsilon(0)\exp(-2x_0 t)\,. \tag{5.62}$$

It is clear that the behavior of the perturbation depends on the sign of $x_0$. If $x_0 < 0$, the perturbation grows exponentially and if $x_0 > 0$, the perturbation decays exponentially.
To find the point where the bifurcation occurs, we look at $x_0 = \pm\sqrt{\mu}$. We see that the steady-state solutions are created at $\mu = 0$. For $\mu < 0$ there are no (real) steady-state solutions.
The stability of the solutions is found by plugging the different steady-state solutions into the governing equation for the perturbation $\varepsilon$. For $x_0 = +\sqrt{\mu}$, perturbations decay to zero and so this solution is stable. For $x_0 = -\sqrt{\mu}$, perturbations grow rapidly and therefore this solution is unstable.

**b. Transcritical bifurcation.** For the second system $\dot{x} = \mu x - x^2$, the steady-state solutions are given by

$$\dot{x} = \mu x - x^2 = x(\mu - x) \implies x_0 = 0\,,,\mu\,. \tag{5.63}$$

We again expand around the steady-solution to find their stability and the bifurcation point.

$$\dot{\varepsilon} = \mu(x_0 + \varepsilon) - (x_0 + \varepsilon)^2 = \mu x_0 + \mu\varepsilon - x_0^2 - 2x_0\varepsilon - \varepsilon^2 \approx (\mu - 2x_0)\varepsilon\,. \tag{5.64}$$

In this case, the equation governing the perturbations is

$$\dot{\varepsilon} = (\mu - 2x_0)\varepsilon \tag{5.65}$$

and is solved by

$$\varepsilon(t) = \varepsilon(0)\exp([\mu - 2x_0]t)\,. \tag{5.66}$$

Note that now the stability of the solution is determined by the sign of the term $\mu - 2x_0$. For the solution $x_0 = 0$, the stability depends on the sign of $\mu$. If $\mu < 0$, this solution is stable. For the solution $x_0 = \mu$, we have $\mu - 2x_0 = -\mu$ and so the stability depends on the sign of $\mu$ but opposite to the $x_0 = 0$ solution. In this case, if $\mu < 0$, this solution is unstable. Clearly, the stability of the solutions switches at $\mu = 0$, which is the location of the bifurcation point.

**c. Pitchfork bifurcation.** For the third system $\dot{x} = \mu x - x^2$, the steady-state solutions are given by

$$\dot{x} = \mu x - x^3 = x(\mu - x^2) \implies x_0 = 0, \pm\sqrt{\mu}\,. \tag{5.67}$$

As usual, we expand around the steady-state solution.

$$\dot{\varepsilon} = \mu(x_0 + \varepsilon) - (x_0 + \varepsilon)^3 = \mu x_0 + \mu\varepsilon - (x_0^3 + 3x_0^2\varepsilon + 3x_0\varepsilon^2 + \varepsilon^3) \approx (\mu - 3x_0^2)\varepsilon \tag{5.68}$$

and so the perturbations are governed by

$$\varepsilon(t) = \varepsilon(0)\exp\big([\mu - 3x_0^2]t\big) \tag{5.69}$$

The stability now depends on the sign of $\mu - 3x_0^2$. We plug in the three different steady-state solutions

$$x_0 = 0, \quad \mu - 3x_0^2 = \mu, \quad x_0 = \pm\sqrt{\mu}, \quad \mu - 3x_0^2 = -2\mu\,. \tag{5.70}$$

Thus, the solutions $x_0 = \pm\sqrt{\mu}$ are stable if $\mu > 0$ and the solution $x_0 = 0$ is stable for $\mu < 0$. As with the other two systems, the bifurcation point is $\mu = 0$.

**A2.** **a.** Doing a linear stability analysis around $x_0 = 0$ we find

$$\dot{\varepsilon} = \mu x_0^2 - x_0^4 + \varepsilon\big(2x_0\big(\mu - 2x_0^2\big)\big) + \varepsilon^2\big(\mu - 6x_0^2\big) + O\big(\varepsilon^3\big) \approx 2x_0\varepsilon\big(\mu - 2x_0^2\big)\,. \tag{5.71}$$

If we solve this for $\varepsilon$ we find

$$\varepsilon(t) = \varepsilon(0)\exp\left(2x_0\big(\mu - 2x_0^2\big)\right)\,. \tag{5.72}$$

Plugging in $x_0 = 0$ gives

$$\varepsilon(t) = \varepsilon(0)\,, \tag{5.73}$$

implying that the perturbation is a constant and neither grows or decays. This is due to the fact that the term in front of $\varepsilon$ in our expansion is zero at $x_0 = 0$. In order to find the behavior of small perturbations, we therefore have include terms of order $\varepsilon^2$.

**b. c.** Including terms of order $\varepsilon^2$ and setting $x_0 = 0$ gives us the equation

$$\dot{\varepsilon} = \mu\varepsilon^2\,. \tag{5.74}$$

Solving this equation gives

$$\varepsilon(t) = \frac{\varepsilon(0)}{1 - \mu\varepsilon(0)t}\,. \tag{5.75}$$

If $\mu$ and $\varepsilon(0)$ have the same sign, perturbations grow. If $\mu$ and $\varepsilon(0)$ have different signs, then perturbations decay. This means that the solution $x_0 = 0$ is semi-stable. On one side of $x_0 = 0$ perturbations decay, on the other side perturbations grow, depending on the sign of $\mu$. If $\mu = 0$, the original equation becomes $\dot{x} = -x^4$ and therefore terms of order $\varepsilon^2$ give no information about the stability of solutions.

# Chapter 6

# Collective Behavior

Janusz Meylahn, Ebo Peerbooms, Clélia de Mulatier

## 6.1 Noisy Kuramoto Model

For a popular science introduction to the Kuramoto model you can watch the following video. Here a couple of applications of the Kuramoto model are mentioned, such as for the synchronization of clocks, humans walking, fireflies and planets or moons. Another application is in the study of synchrony between neurons in the brain. The suprachiasmatic nucleus, also known as the body-clock, is a cluster of neurons in the brain responsible for dictating the rhythm of bodily functions. For a short introduction into the application of the Kuramoto model see the following article. The article also contains an interactive animation that you can use to develop intuition regarding this problem set.

The noisy Kuramoto model is defined by the following system of coupled differential equations

$$\frac{\mathrm{d}\theta_i(t)}{\mathrm{d}t} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin\left[\theta_j(t) - \theta_i(t)\right] + \xi_i(t), \tag{6.1}$$

with $\xi_i(t)$ independent noise with expected values $\langle\xi_i(t)\rangle = 0$ and $\langle\xi_i(t)\xi_j(t')\rangle = 2D\delta(t-t')\delta_{i,j}$. Here we have

- $N$ – number of oscillators
- $\omega_i \sim \mu(\omega)$ – natural frequency of $i^{\text{th}}$ oscillator drawn from distribution $\mu(\cdot)$
- $\theta_i(t)$ – phase of $i^{th}$ oscillator
- $K \in (0, \infty)$ – interaction strength
- $D \in (0, \infty)$ – noise strength (contained in $\xi_i(t)$).

**Q1.** Consider a system of two oscillators. What is the effect of the interaction term $\frac{K}{N}\sum_{j=1}^{N}\sin\left[\theta_j(t) - \theta_i(t)\right]$? Note that $K$ is a positive constant. Try this situation in the simulation here and check that your intuition is correct.

We define the order parameter as

$$r_N(t)\mathrm{e}^{i\psi_N(t)} = \frac{1}{N}\sum_{j=1}^{N}\mathrm{e}^{i\theta_j(t)}. \tag{6.2}$$

**Q2.** Show that (6.1) can be written as

$$\frac{\mathrm{d}\theta_i(t)}{\mathrm{d}t} = \omega_i + Kr_N(t)\sin\left[\psi_N(t) - \theta_i(t)\right] + \xi_i(t). \tag{6.3}$$

**Q3.** What is the advantage of (6.3)?

Let $p_t(\theta, \omega)$ represent the density of phase oscillator with angle $\theta$, natural frequency $\omega$ at time $t$. In the continuum limit ($N \to \infty$) this density evolves according to:

$$\frac{\partial p_t(\theta,\omega)}{\partial t} = -\frac{\partial}{\partial \theta}\Big[p_t(\theta,\omega)v_t(\theta,\omega)\Big] + D\frac{\partial^2 p_t(\theta,\omega)}{\partial \theta^2} \tag{6.4}$$

with

$$v_t(\theta,\omega) = \omega + Kr(t)\sin(\psi(t) - \theta) \tag{6.5}$$

and satisfies

$$p_t(\theta + 2\pi, \omega) = p_t(\theta,\omega) \quad \text{and} \quad \int_0^{2\pi} p_t(\theta,\omega)\mathrm{d}\theta = 1. \tag{6.6}$$

**Q4.** Solve for the stationary density of (6.4) in the case that $\mu(\omega) = \delta_0$.

**Q5.** In the large $N$ limit the expression for the order parameter becomes:

$$r(t)\mathrm{e}^{i\psi(t)} = \int_0^{2\pi}\int_{\mathbb{R}} \mathrm{e}^{i\theta} p_t(\theta,\omega)\mathrm{d}\omega\mathrm{d}\theta. \tag{6.7}$$

1. Use the solution you calculated in **Q4.** to find an implicit equation for the stationary synchronization level $r$ in the case that $\mu(\omega) = \delta_0$.
2. Taylor expand (around $r = 0$) of the equation you found in a) to determine the critical threshold for $K$ above which nonzero solutions for $r$ exists. *Hint: Define the right hand side of the equation from a) to be a function $V(2Kr/D)$. Use that $V(x)$ is monotonically increasing, concave and that* $\lim_{x\to\infty} V(x) = 1$.
3. Reflect on whether or not the critical threshold makes sense.

## 6.2 Swarm behavior and Vicsek Model

### 6.2.1 Bird flocks

Watch this 20min talk (starting at 1h 11 min of the video), which discusses this paper on bird flocks [1].

**Q1.** In the Vicsek model, how is the evolution of the positions of the birds described? Can you recall the expression of the position and the direction of the birds at time $t+1$ as a function of its position and direction at time $t$? Can you comment on the meaning of each term?

**Q2.** How would you generalize this equation of evolution in the case of birds that move at a constant speed $v_0$?

**Q3.** In the Vicsek model, what is the meaning of the metric range? The authors of Ref. [1] argue that animal collective behavior depends on topological distance rather than metric distance. What is the difference between a metric range and a topological range? How do the authors of Ref. [1] suggest to modify the Vicsek model to capture better the behavior of bird flocks?

**Q4.** The author of the paper give the following relation between the distance $r_1$ to the first neighbor and the distance $r_N$ to the $N$-th neighbor:

$$r_N \sim r_1 N^{1/3}, \tag{6.8}$$

where the $\sim$ indicates the scaling of $r_N$ with $N$ to some constant factor (to be more specific, this equation means that we expect $r_N$ to behave like: $r_N = \alpha\, r_1 N^{1/3} + \beta$, where $\alpha$ and $\beta$ don't depend on $N$). Assuming that a flock has a constant density of birds, can you re-derive this result?

**Q5.** Considering the following two hypotheses:

(a) the birds interact with a metric interaction: they align to their neighbors that are within a fixed distance $r_c$ from them;

(b) the birds interact with a topological interaction: they align with their $N_c$ first neighbors, where $N_c$ is fixed.

Assume that you have access to the recordings of many bird flocks, how would you use (6.8) to check which of these two hypotheses is the good one? Comment on Fig. 3.c and d of Ref. [1].

We are interested in understanding how the collective behavior (flocking) emerges in the Vicsek model as a dynamical phase transition. In section 2.2, we remove the movement of the "animals" by fixing them on a lattice, but we still allow them to have different directions. In this case, the system can only exhibit a collective behavior (long-range order in the alignment of the orientation of the "animals") if the dimension $d$ is strictly larger than 2. In section 2.3, we see that the Vicsek model can already exhibit such collective behavior in dimension $d = 2$ thanks to the spatial movements of the "animals" (which allows information to propagate faster and further in the system). **Note: the following exercises are optional.**

### 6.2.2 (Bonus) Scaling arguments for the XY models – "Birds on a lattice"

Consider "bird-spins" on a $d$-dimensional (rigid) lattice that have the ability to align themselves to their neighbouring spins. Instead of taking binary values, these spins can take on continuous values in the range $\theta_i \in (-\pi, \pi)$. Suppose we prepare the lattice in a configuration where all spins point in the same direction, $\langle \boldsymbol{\theta} \rangle = \theta_0$. We now introduce a single mis-aligned spin that has alignment $\theta_i = \theta_0 + \delta\theta_0$. Due to the interaction between adjacent lattice sites, the error will propagate throughout the lattice.

**Q1.** Assuming that the propagation on the lattice follows a diffusion law (i.e. behaves like a random walker), derive (using scaling arguments) that the spin error decays as

$$\delta\theta \sim \delta\theta_0 \tau^{-d/2} \tag{6.9}$$

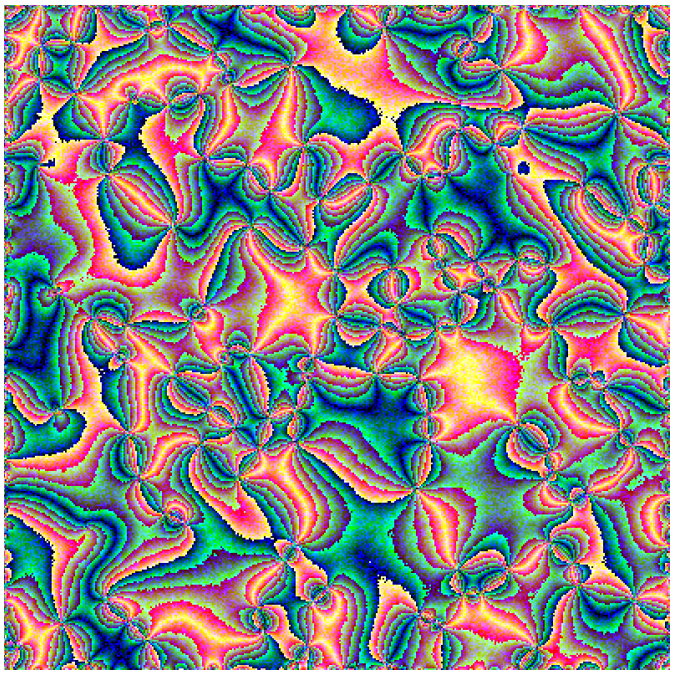


Figure 6.1: Realization of the "birds on a lattice" model (also known as the $xy$-model) on a $400 \times 400$ grid. Colors indicate the angle of each of the spins.

where $\tau$ is the time elapsed since introducing the mis-aligned spin and $d$ is the lattice dimension.

**Q2.** Having seen what happens to a single defect, we now wish to find the influence of noise on all the spins on the lattice. Assuming that noise fluctuations produce a certain number of errors within some volume and that this number scales with the time $\tau$, derive that the number of errors is proportional to

$$n_e \sim \tau^{1+d/2} \tag{6.10}$$

**Q3.** Show that the total expected error $\Omega_e$ in the system scales with $n_e$ as

$$\Omega_e \sim \sqrt{n_e} \tag{6.11}$$

Hint: since the errors can point in any direction, how would you measure the total error?

**Q4.** **a.** Combining these results, show that the error amplitude per spin is given by

$$\Delta\theta \sim r_e^{1-d/2} \tag{6.12}$$

where $r_e$ is the radius of the volume within which the errors have propagated.

**b.** Comment on the behaviour of the systems in different dimensions. In which dimensions is it possible to have long-range order?

**c.** For the physicists; do you know of a theorem that could have predicted the absence of long-range order in some dimensions? How does it apply to this system?

### 6.2.3 (Bonus) Birds not on a lattice: the Vicsek model

In the Vicsek model, the fluctuations are coupled to the motion of the "bird-spins". Assume that the mean motion of the spins is in the $\theta_0$ direction. We align our frame of reference to this direction, such that $\theta_0 = 0$ in this frame.

**Q1.** Argue or draw a diagram to show that we can write the seperations induced by angular fluctuations around the mean direction as

$$\delta x_\perp \sim v_0 \tau \sin(\delta\theta), \quad \delta x_\parallel \sim v_0 \tau (1 - \cos(\delta\theta)) \tag{6.13}$$

**Q2.** Use the fact that $\delta\theta$ is small to derive the leading order terms in $\delta\theta$ for the expressions of $\delta x_\perp$ and $\delta x_\parallel$. What does this tell you about the nature of the transport of orientation information?

**Q3.** Decomposing the error propagation volume as $w_\perp^{d-1} w_\parallel$ and introducing two diffusion constants $D_\perp$ and $D_\parallel$, which encode the loss of orientation information due to diffusion. The growth of the error propagation volume is given by

$$w_\perp \sim \delta x_\perp + D_\perp \tau^{1/2}, \quad w_\parallel \sim \delta x_\parallel + D_\parallel \tau^{1/2} \tag{6.14}$$

Using the same reasoning as you did in the question on the lattice spins, show that

$$\Delta\theta \sim \frac{\tau^{1/2}}{\sqrt{w_\perp^{d-1} w_\parallel}} \tag{6.15}$$

**Q4.** We introduce three timescales which govern the propagation of orientation errors:

$$w_\perp \sim \tau^\alpha, \quad w_\parallel \sim \tau^\beta, \quad \Delta\theta \sim \tau^\gamma \tag{6.16}$$

Show that for large $\tau$ these yield a system of three equations given by

$$2\gamma = 1 - \beta - (d-1)\alpha, \quad \alpha = \max\left(1 + \gamma, \frac{1}{2}\right), \quad \beta = \max\left(1 + 2\gamma, \frac{1}{2}\right) \tag{6.17}$$

**Q5. a.** Solve (by hand or by using Mathematica) this system of equations. Determine the range of validity of each of the solutions you find by taking the constraints on $\alpha$ and $\beta$ into account. You can use that

$$\max(x, y) = \frac{1}{2}(x + y + |x - y|) \tag{6.18}$$

Comment of the nature of the orientation information diffusion in the different regimes. You should find three distinct regimes.

**b.** What is the minimal dimension for which long-range order is possible?

**c.** For the physicists; why doesn't this contradict the theorem of question **2.1,Q4c**?

# References

[1] M. Ballerini, N. Cabibbo, R. Candelier, *et al.*, "Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study," *Proceedings of the national academy of sciences*, vol. 105, no. 4, pp. 1232–1237, 2008.

## 6.3 Solutions

### 6.3.1 Kuramoto model

**A1.** The interaction is attractive.

**A2.** Multiplying both sides of (6.2) by $e^{-i\theta_i}$ gives

$$r_N(t)e^{i(\psi_N(t)-\theta_i(t))} = \frac{1}{N}\sum_{j=1}^{N} e^{i(\theta_j(t)-\theta_i(t))}. \tag{6.19}$$

Now expressing both complex exponentials using Euler we have

$$\begin{aligned} r_N(t)\cos(\psi_N(t)-\theta_i(t)) + ir_N(t)\sin(\psi_N(t)-\theta_i(t)) \\ = \frac{1}{N}\sum_{j=1}^{N}\Big[\cos\big(\theta_j(t)-\theta_i(t)\big) + i\sin\big(\theta_j(t)-\theta_i(t)\big)\Big]. \end{aligned} \tag{6.20}$$

Here we just equate the terms without an $i$ on the left with the terms without an $i$ on the right and do the same for terms with an $i$ so that

$$r_N(t)\cos(\psi_N(t)-\theta_i(t)) = \frac{1}{N}\sum_{j=1}^{N}\cos\big(\theta_j(t)-\theta_i(t)\big)$$

$$r_N(t)\sin(\psi_N(t)-\theta_i(t)) = \frac{1}{N}\sum_{j=1}^{N}\sin\big(\theta_j(t)-\theta_i(t)\big).$$

The right hand side of the second equation matches the sum that appears in our interaction term so we can replace it with the left hand side, giving us:

$$Kr_N(t)\sin(\psi_N(t)-\theta_i(t)).$$

**A3.** The equation basically says that you only interact with the average angle of the oscillators and that your interaction is modulated by the amount of synchronization there is. This makes it much easier to analyze and simulate the model.

**A4.** Since $\omega = 0$ for all oscillators and we are considering the stationary density (i.e. time independent density) we can write $p(\theta)$ instead of $p_t(\theta,\omega)$ as well as $r$ and $\psi$ instead of $r(t)$ and $\psi(t)$. This simplifies (6.4) to

$$0 = -\frac{\partial}{\partial\theta}\Big[p(\theta)Kr\sin(\psi-\theta)\Big] + D\frac{\partial^2 p(\theta)}{\partial\theta^2}. \tag{6.21}$$

We can rewrite this in a more suggestive form:

$$\frac{\partial^2 p(\theta)}{\partial\theta^2} = \frac{\partial}{\partial\theta}\Big[p(\theta)\frac{Kr}{D}\sin(\psi-\theta)\Big]. \tag{6.22}$$

This implies that

$$\frac{\partial p(\theta)}{\partial\theta} = p(\theta)\frac{Kr}{D}\sin(\psi-\theta) \tag{6.23}$$

The solution has the form

$$p(\theta) = \frac{1}{Z}\exp\Big[\frac{Kr}{D}\cos(\psi-\theta)\Big] + C. \tag{6.24}$$

Using the boundary conditions we have

$$p(\theta) = \frac{1}{Z}\exp\big[\frac{Kr}{D}\cos(\psi-\theta)\big] + C = \frac{1}{Z}\exp\big[\frac{Kr}{D}\cos(\psi-\theta+2\pi)\big] + C = p(\theta+2\pi), \tag{6.25}$$

which implies that $C = 0$. For normalization

$$\int_0^{2\pi} p(\theta)\mathrm{d}\theta = 1 \tag{6.26}$$

so that

$$Z = \int_0^{2\pi} \mathrm{e}^{\frac{Kr}{D}\cos(\psi-\theta)}\mathrm{d}\theta, \tag{6.27}$$

which is a special function called a *modified Bessel function of the first kind.*

**A5.**

**a.** Plugging the solution of the previous question into (6.7) we have

$$r\mathrm{e}^{i\psi} = \int_0^{2\pi} \mathrm{e}^{i\theta}\frac{1}{Z}\exp\left[\frac{Kr}{D}\cos(\psi-\theta)\right]\mathrm{d}\theta. \tag{6.28}$$

Multiplying both sides by $\mathrm{e}^{-i\psi}$ and collecting real and imaginary parts gives

$$r = \frac{\int_0^{2\pi}\cos(\psi-\theta)\exp\left[\frac{Kr}{D}\cos(\psi-\theta)\right]\mathrm{d}\theta}{\int_0^{2\pi}\mathrm{e}^{\frac{Kr}{D}\cos(\psi-\theta)}\mathrm{d}\theta}. \tag{6.29}$$

**b.** We can consider the equation for the synchronization level to be of the form $r = V(\frac{Kr}{D})$ where.

$$V(x) := \frac{\int_0^{2\pi}\cos(\psi-\theta)\exp\left[x\cos(\psi-\theta)\right]\mathrm{d}\theta}{\int_0^{2\pi}\mathrm{e}^{x\cos(\psi-\theta)}\mathrm{d}\theta}. \tag{6.30}$$

From this expression we can see that $V(0) = 0$. The derivative of $V(x)$ is

$$V'(x) = \frac{\int_0^{2\pi}\cos^2(\psi-\theta)\exp\left[x\cos(\psi-\theta)\right]\mathrm{d}\theta\int_0^{2\pi}\mathrm{e}^{x\cos(\psi-\theta)}\mathrm{d}\theta}{\left(\int_0^{2\pi}\mathrm{e}^{x\cos(\psi-\theta)}\mathrm{d}\theta\right)^2} \tag{6.31}$$

$$- \frac{\left(\int_0^{2\pi}\cos(\psi-\theta)\exp\left[x\cos(\psi-\theta)\right]\mathrm{d}\theta\right)^2}{\left(\int_0^{2\pi}\mathrm{e}^{x\cos(\psi-\theta)}\mathrm{d}\theta\right)^2} \tag{6.32}$$

$$= \frac{\int_0^{2\pi}\cos^2(\psi-\theta)\exp\left[x\cos(\psi-\theta)\right]\mathrm{d}\theta\int_0^{2\pi}\mathrm{e}^{x\cos(\psi-\theta)}\mathrm{d}\theta}{\left(\int_0^{2\pi}\mathrm{e}^{x\cos(\psi-\theta)}\mathrm{d}\theta\right)^2} - V(x). \tag{6.33}$$

From this we see that

$$V'(0) = \frac{1}{2\pi}\int_0^{2\pi}\cos^2(\psi-\theta)\mathrm{d}\theta = \frac{1}{2}. \tag{6.34}$$

the equation for $r$ around small $r$ is thus

$$r = \frac{Kr}{2D}. \tag{6.35}$$

By plotting both left and right hand side of the full equation as functions of $r$, we see that there is a non-zero solution for $r$ when $K > K_c$ with

$$K_c := 2D. \tag{6.36}$$

**c.** For synchronization to occur the interaction strength must overcome the tendency of the noise term to distribute the oscillators evenly on the circle. It thus makes sense that the critical value increases as the noise strength increases.

### 6.3.2 Swarm behavior and Vicsek Model

**A1.** At each time step, a randomly chosen bird moves forward by a small step of length 1 in the direction given by $\theta_t$ (equation given in the case of a 2-dimensional model):

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t + \cos\theta_t \boldsymbol{e}_x + \sin\theta_t \boldsymbol{e}_y \,. \tag{6.37}$$

It then updates its direction $\theta_{t+1}$ to be equal to the average value of the directions of the other birds in its neighbourhood plus some random noise:

$$\theta_{t+1} = \langle\theta_t\rangle_{r_c} + \eta_t \,, \tag{6.38}$$

The term $\eta_t$ represents a Gaussian noise, i.e. $\langle\eta_t\rangle = 0$ and $\langle\eta_t\,\eta_{t'}\rangle = \alpha\,\delta_{t,t'}$, where $\alpha$ is a constant controlling the amplitude of the noise. The constant $r_c$ is the radius of the disk around the bird that defines its neighbourhood. The considered bird interacts with other birds in its neighbourhood by updating its new direction to be the average direction of these birds.

**A2.** If the birds move at a constant speed $v_0$, then, at each time steps of the simulation, the time is increased by a small quantity $dt$ and the positions of the birds are updated using:

$$\boldsymbol{x}_{t+dt} = \boldsymbol{x}_t + v_0\,dt\,(\cos\theta_t \boldsymbol{e}_x + \sin\theta_t \boldsymbol{e}_y) \,. \tag{6.39}$$

The update of the directions of the birds remain given by Eq. (6.38).

**A3. Metric range:** In the Vicsek model, the metric range is the constant denoted $r_c$ in the previous equations. It represents the radius of the area around each bird that is considered to be the neighbourhood of the bird. Birds aligned their direction with other birds in that neighbourhood.

**Topological range:** The authors of Ref. [1] suggest that birds don't use a metric range to align their direction, as described above, but instead they align their direction with their $N_c$-th nearest neighbours. The constant $N_c$, which represents the number of birds to which they align to, is the topological range.

**A4.** The quantity $r_1$ is the average distance of the closest neighbor of any bird in the flock. The quantity $r_N$ is the average distance of the $N$-th closest neighbor of any bird in the flock. Let us denote $\rho$ the density of birds in the flock; we assume that this density is constant over the flock. We then obtained that:

$$\rho \sim \frac{1}{r_1^3} \qquad \text{and} \qquad \rho \sim \frac{N}{r_N^3} \,, \tag{6.40}$$

as we expect to find on average 1 bird in a ball of radius $r_1$ and $N$ birds in a ball of radius $r_N$. Assuming that the density is constant in the flock, this leads to:

$$\frac{1}{r_1^3} \sim \frac{N}{r_N^3} \,, \tag{6.41}$$

which thus gives Eq. (6.8): $r_N \sim r_1\,N^{1/3}$. This relation applies to $N = n_c$, for which $r_N = r_c$; one gets that:

$$r_c \sim r_1\,n_c^{1/3} \,, \tag{6.42}$$

**A5.** Considering the following two hypotheses:

(a) **metric scenario:** birds align to their neighbors that are within a fixed distance $r_c$ from them, where the distance $r_c$ is a constant over all flocks: $r_c$ = constant. Equation (6.8) then gives a dependence of $n_c$ in $r_1$:

$$r_c = \text{constant} \qquad \text{and} \qquad n_c^{1/3} \sim r_1^{-1} \,. \tag{6.43}$$

(b) **topological scenario:** the birds align with their $n_c$ first neighbors, where the number $n_c$ is fixed and constant over all flocks (of these type of birds): $n_c$ = constant. From Eq. (6.8) we then obtain a dependence of $r_c$ in $r_1$:

$$n_c = \text{constant} \qquad \text{and} \qquad r_c \sim r_1 \,. \tag{6.44}$$

The authors of Ref. [1] recorded multiple flocks for the same species of birds and computed the values of $n_c$ (using a measure of anisotropy in the positions $N$-th birds), $r_1$, and finally $r_c$ (using the relation (6.42)). Fig. 3.c and d shows the behavior of $n_c$ and $r_c$ as a function of $r_1$ for different flocks. We observe that the value of $n_c$ doesn't seem to depend on the flock, while $r_c$ depends linearly in $r_1$. The topological distance seems to be the appropriate hypothesis for this problem.

### Birds on a lattice

**A1.** Since the propagation of errors was assumed to be diffusive, we know that in a time $\tau$ the error will spread out over a distance $r \sim \sqrt{\tau}$. In $d$ dimensions, the volume over which the error propagates is given by $V_e \sim r^d \sim \tau^{d/2}$. Because the error is conserved, the spin error decays as

$$\delta\theta \sim \delta\theta_0 / V_e = \delta\theta_0 \tau^{-d/2} \tag{6.45}$$

**A2.** If we assume that errors are produced within a volume $V_e$ at a rate proportional to $\tau$, then the total number of errors should scale with the volume. We also know that errors propagate over the volume $V_e$ as $V_e \sim \tau^{d/2}$. Combining both assumptions we have

$$n_e \sim \tau V_e \sim \tau^{1+d/2} \tag{6.46}$$

**A3.** The total error is given by

$$\Omega_e = \sqrt{\sum_{i=1}^{n_e} (\delta\theta_i)^2} \approx \sqrt{n_e \langle (\delta\theta)^2 \rangle} \sim \sqrt{n_e} \tag{6.47}$$

**A4. a.** The error amplitude per spin is the total expected error divided by the volume over which the errors have propagated. Therefore

$$\Delta\theta \sim \Omega_e / V_e \sim \frac{\sqrt{n_e}}{V_e} \sim \frac{\sqrt{\tau V_e}}{V_e} \sim \sqrt{\frac{\tau}{V_e}} \sim \sqrt{\frac{r^2}{r^d}} \sim r^{1-d/2} \tag{6.48}$$

**b.** If $d < 2$, then $1 - d/2 > 0$ and therefore errors propagate throughout the entire system and long-range order is not possible. If $d > 2$, then $1 - d/2 < 0$ and errors will decay. This means that order is resistant to fluctuations, enabling the possibility of long-range order. In $d = 2$, fluctuations can still propagate throughout the entire system, but order will be disturbed on a very slow timescale.

**c.** The absence of long-range order is predicted by the Mermin-Wagner theorem, which states that continuous symmetries cannot be spontaneously broken in systems with sufficiently short-range interactions in $d \leq 2$. For this model, that means that in $d \leq 2$, there is no spontaneous magnetization, i.e. $\langle s_i \rangle = 0$.

### Birds not on a lattice: the Vicsek model

**A1.** If two birds start from the same location $\boldsymbol{x}_0(\tau = 0) = (0, 0)$ with angles $\theta_1 = 0$ and $\theta_2 = \delta\theta$, then after a time $\tau$ we have

$$\boldsymbol{x}_1(\tau) = (v_0\tau, 0), \quad \boldsymbol{x}_2(\tau) = (v_0\tau \cos\delta\theta, v_0\tau \sin\delta\theta) \tag{6.49}$$

The difference between these locations is

$$\delta\boldsymbol{x} = \boldsymbol{x}_2 - \boldsymbol{x}_1 = (\delta x_\parallel, \delta x_\perp) = (v_0\tau(1 - \cos(\delta\theta)), v_0\tau \sin(\delta\theta_0)) \tag{6.50}$$

**A2.** Expanding the sine and cosine we get

$$\delta x_\perp = v_0\tau \sin(\delta\theta) \approx v_0\tau\delta\theta \sim \tau\delta\theta \tag{6.51}$$

$$\delta x_\parallel = v_\tau(1 - \cos(\delta\theta)) \approx v_0\tau\left(1 - 1 + \frac{1}{2}(\delta\theta)^2\right) \sim \tau(\delta\theta)^2 \tag{6.52}$$

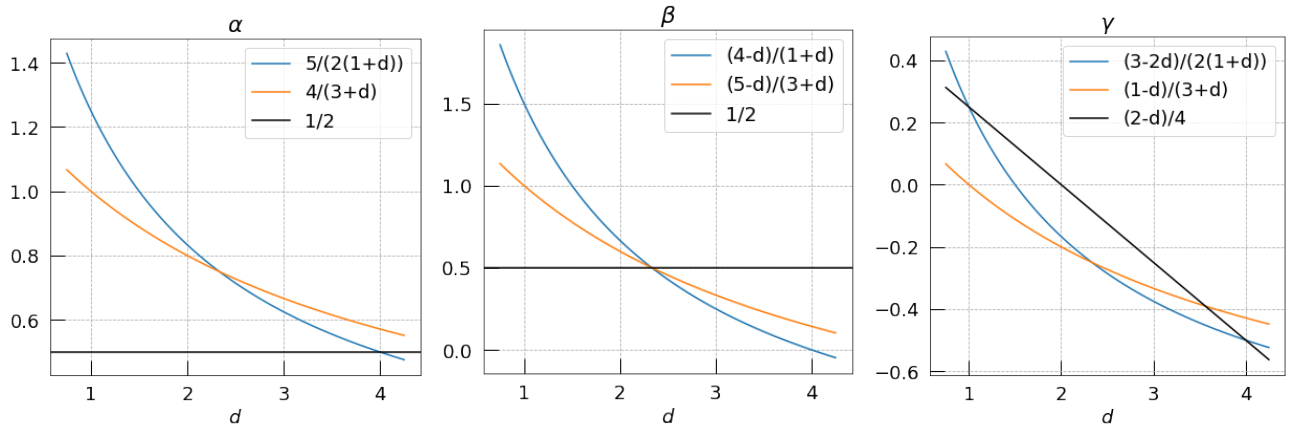



Figure 6.2: Solution regimes of the exponents $\alpha, \beta$ and $\gamma$.

**A3.** Using the same reasoning as for the lattice spins, we have

$$\delta\theta \sim \frac{\Omega_e}{V_e} \sim \sqrt{\frac{\tau}{V_e}} \sim \frac{\tau^{1/2}}{\sqrt{w_\perp^{d-1} w_\parallel}} \tag{6.53}$$

**A4.** We start with the equation for $\delta\theta$ and write

$$\delta\theta \sim \frac{\tau^{1/2}}{\sqrt{\tau^{\alpha(d-1)}\tau^{\beta}}} = \tau^{\gamma} \tag{6.54}$$

Squaring both sides we have

$$\frac{\tau}{\tau^{\alpha(d-1)}\tau^{\beta}} = \tau^{2\gamma} \tag{6.55}$$

and therefore

$$\tau = \tau^{2\gamma+\alpha(d-1)+\beta} \implies 1 = 2\gamma + \alpha(d-1) + \beta \tag{6.56}$$

Secondly, we have

$$\tau^{\alpha} = \tau(\delta\theta) + D_\perp \tau^{1/2} = \tau^{\gamma+1} + D_\perp \tau^{1/2} \tag{6.57}$$

and

$$\tau^{\beta} = \tau(\delta\theta)^2 + D_\parallel \tau^{1/2} = \tau^{2\gamma+1} + D_\parallel \tau^{1/2} \tag{6.58}$$

At large values of $\tau$, the behavior of the exponent will tend to the largest exponent in the sum terms, thus

$$\alpha = \max\left(1+\gamma, \frac{1}{2}\right), \quad \beta = \max\left(1+2\gamma, \frac{1}{2}\right) \tag{6.59}$$

**A5. a.** Plugging the system of equations into Mathematica and using the analytic expression for the maximum of two numbers, we obtain four solutions. The trick here is to look at the value of $\gamma$ as a function of $d$ and check the consistency relations $\alpha = \max\left(1+\gamma, \frac{1}{2}\right)$ and $\beta = \max\left(1+2\gamma, \frac{1}{2}\right)$. An easier way to find the different regimes is to plot the different solution curves and find the only consistent way to match the three regimes. Looking at Fig. 6.2, we see that in order to continuously move between the different solution regimes, we should move from the black curve to the blue curve at $d = 4$ and move from the blue curve to the orange curve at $d = 7/3$.
For $d \geq 4$ the solution is consistent if

$$\alpha = \beta = \frac{1}{2}, \quad 2\gamma - 1 + \frac{1}{2}(d-1) + \frac{1}{2} = 0 \implies \gamma = \frac{1}{2} - \frac{d}{4} \tag{6.60}$$

In this case, we have a diffusive system (characterized by the exponent $1/2$) and $\gamma < 0$ indicates the existence of long-range order. For $7/3 \leq d < 4$, the solution is given by

$$\alpha = \frac{5}{2(d+1)}, \quad \beta = \frac{1}{2}, \quad \gamma = \frac{3-2d}{2(d+1)} \tag{6.61}$$

For $d$ in this range, $\gamma < 0$. Note that in this case we have super-diffusive propagation (due to $\alpha > \frac{1}{2}$) in the direction transverse to the motion. For $d < 7/3$, the valid solution is given by

$$\alpha = \frac{4}{d+3}, \quad \beta = \frac{5-d}{d+3}, \quad \gamma = \frac{1-d}{d+3} \tag{6.62}$$

giving $\gamma < 0$ for any $d > 1$ and super-diffusive propagation in both the longitudinal and transverse direction.

**b.** From the above, we see that the minimal dimension in which long-range order is possible is $d > 1$.

**c.** The Mermin-Wagner theorem holds for equilibrium systems. The Vicsek model - due to the active nature of the particles - is out of equilibrium. The ability to move allows for faster propagation of orientation information through the flock, thereby allowing for long-range order in dimensions lower than 2.