
**Università degli Studi di Padova**

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"


Corso di Laurea Triennale in Matematica

# Phase transition in the 2D Ising model: the theory behind and simulations via the Metropolis-Hastings algorithm


Relatore: Prof. Marco Formentin

Laureando: Lorenzo Luzzi

Matricola: 1173312



Anno Accademico 2021/2022

23/09/2022

# Contents

**Introduction** **iv**

**1 Basics of Homogeneous Markov Chain and the Ergodic Theorem** **1**

1.1 Homogeneous Markov Chains ..... 1

1.2 Stationarity ..... 2

1.3 Strong Markov Property and Regeneration ..... 3

1.4 Recurrence ..... 4

1.5 The Ergodic Theorem ..... 9

**2 A MCMC Method: The Metropolis-Hasting Algorithm** **12**

2.1 General Principle and MCMC methods ..... 12

2.2 Metropolis-Hasting Algorithm ..... 13

**3 Markov Random Fields and Gibbs Fields** **16**

3.1 Neighborhoods and Local Specifications ..... 16

3.2 Cliques, Potential, and Gibbs Distributions ..... 18

3.3 Gibbs-Markov Equivalence ..... 19

**4 2D Ising Model** **21**

4.1 Ising's finite model ..... 21

4.2 Phase Transitions ..... 22

4.2.1 Peierls's Argument ..... 25

4.3 Simulations via the Metropolis-Hastings Algorithm ..... 29

4.3.1 Motivations ..... 29

4.3.2 Simulations ..... 30

**Conclusions** **33**

**Matlab code** **35**

**Bibliography** **38**

## Introduction

The Ising model was born in the field of statistical mechanics, to describe the ferromagnetism of materials. This model dates back to 1920 [1], when the physicist Wilhelm Lenz proposed it to his student Ernst Ising, as a problem to investigate the phase transition, from paramagnetic state to ferromagnetic state of some materials. Ising solved it in 1-dimension (1924), publishing in his thesis (1925) that no phase transition occurs [2].
For what concerns the 2-dimensional (or more) square-lattice Ising model, much harder of the 1-dimensional one, we have to wait until 1936 when Rudolph Peierls proved for the firs time the possibility of a spontaneous magnetization, using what is now called a Peierls argument [3]. Only in 1944, Lars Onsager proposed an analytical solution to the problem of phase transition in the Ising model, with no external magnetic field [4].
Even if the model was introduced in the field of physics, it's now widely accepted in the mathematical literature, thanks to the many tools it uses, such as graph theory, combinatorics and certainly the theory of probability, to name a few ones. In its simplest formulation, the Ising model is based on discrete variables, which we call spins. As happens in reality, such spins are characterized by having interactions that strongly depend on the proximity they have with each other. For this purpose, a topology that makes this idea precise and formal is introduced. As we will see, this topology does not have particularly stringent rules. For this reason, this model represents the paradigm of collective phenomena, that are certainly presented not only in many other realities of physics, but also in many other areas such as opinion theory, genetic statistics, neuroscience, economics and environmental sciences.
In this regard, simulating the model computationally is very important and of practical interest. As it often happens, evaluating numerically this model in a direct way is difficult, due to the large amount of data that must be processed from a number of inputs not high at all. This is why the model is simulated via Monte Carlo methods. Since it has a characteristic stochastic aspect, the Markov chain Monte Carlo methods are even more useful in this case. Indeed, this class of computational algorithms uses randomness to obtain estimates of quantities that converge in probability to those desired. In particular do it by constructing a Markov chain that has a specific stationary distribution .

In this thesis, we present the 2-dimension Ising model and its occurrence of phase transition from a purely probabilistic point of view, without going into the more physical as well as analytic aspects of the issue.
An extensive initial part is dedicated to introducing the fundamental and basic concepts of the theory behind the model. For this purpose we introduce the homogeneous Markov

chains to get to the Markov random fields and Gibbs fields. We will also focus on some introductory aspects of the theory of the stability of Markov chains to prove the very powerful Ergodic theorem and to complete this theoretical framework. The theorem is the operating principle of the Metropolis-Hastings algorithm (one of the first and fundamental among the MCMC methods), whose usefulness in this thesis can be appreciated by showing the phase transition from a practical point of view throughout the simulations. In addition the Peierls argument is also considered for its historical importance.

# Chapter 1

# Basics of Homogeneous Markov Chain and the Ergodic Theorem

## 1.1 Homogeneous Markov Chains

Let us start introducing the key object that will make the whole work possible.
In order to represent a process that evolves in a discrete-time, from a probabilistic point of view, the starting point is to define what is a Stochastic Process.

**Definition 1.1.1** (Stochastic Process)**.** A sequence $\{X_n\}_{n\in\mathbb{N}}$, of random variables, with with values in a set $E$, called the state space, is called a *discrete-time Stochastic Process*.

In this context the time is indexed by $n \in \mathbb{N}$ and we assume that $E$ is countable. Moreover we will say that if $X_n = i$, the process is in state i at time $n$.
Among all these possible processes, in particular we are interested to Markov chains and more specifically in which ones are homogeneous over the time.

**Definition 1.1.2** (Markov Chains and Homogeneous Markov Chains)**.** A stochastic process $\{X_n\}_{n\geqslant 0}$ with state space $E$, is called a *Markov Chain* if it satisfies the following property, called *Markov property*

$$P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, .., X_0 = i_0) = P(X_{n+1} = j|X_n = i) \tag{1.1}$$

$\forall\ n \in \mathbb{N}; \forall\ j, i, i_{n-1}, .., i_0 \in E$.
In particular, if the right-hand side of (1.1) is independent of $n$ it is called an *Homogeneous Markov Chains* [HMC], i.e. it holds

$$P(X_{n+1} = j|X_n = i) = m_{ij}. \tag{1.2}$$

To paraphrase the above definition, in a Markov chain the future $X_n = j$ and the past $X_{n-1} = i_{n-1}$ are independent, given the present $X_n = i_n$.

Focusing our attention on the HMCs, it is natural to define the main object that characterized any chain. From the property (1.2) it is possible to define the matrix $M = (m_{ij})_{i,j\in E}$, called the *Transition Matrix* of the HMC $\{X_n\}_{n\geqslant 0}$.

**Remark 1.** Since the entries of the transition matrix $M = (m_{ij})_{i,j\in E}$ are probabilities and the chain in state $i$ must move to some state, it holds true that $M$ is a *Stochastic Matrix*, that is

$$m_{ij} \geqslant 0, \quad \sum_{k\in E} m_{ik} = 1 \qquad \forall i, j \in E. \tag{1.3}$$

## 1.2 Stationarity

Let us introduce the notion of stationarity, an important concept in the stability theory of HMCs.

**Definition 1.2.1** (Invariant Measure/Distribution)**.** Let $X = \{X_n\}_{n\geqslant 0}$ be an HMC with values in the state space $E$ and transition matrix $M = (m_{ij})_{i,j\in E}$. The measure, or the probability distribution, $\pi = (\pi(i))_{i\in E}$ is said *stationary* or *invariant* for the chian $X$ if

$$\pi M = \pi \tag{1.4}$$

or equivalently,

$$\sum_{i\in E} \pi(i) m_{ij} = \pi(j) \quad \forall j \in E \tag{1.5}$$

Iterating (1.4) it follows that $\pi M^n = \pi M M^{n-1} = \ldots = \pi$. Therefore, if the initial state of the chain is chosen according to the stationary distribution, i.e. $P(X_0 = i) = \pi(i)$, we obtain that

$$P(X_n = i) = (\pi M^n)_i = \pi(i)$$

and also

$$\begin{aligned} P(X_n = i, X_{n+1} = i_{n+1}, \ldots, X_{n+k} = i_{n+k}) &= P(X_n = i) m_{i(i+1)} \cdots m_{(n+k-1)(n+k)} \\ &= \pi(i) m_{i(i+1)} \cdots m_{(n+k-1)(n+k)}. \end{aligned}$$

In this sense we say that the chain is *stationary*, because if the chain starts according to the stationary distribution, the probability $P(X_n = i)$ depends only on $\pi$ and not on the time $n$, in other words a chain starting in equilibrium remains in equilibrium.

Let us give a sufficient condition for the stationarity of a distribution.

**Theorem 1.2.1** (Detailed Balance Test [DBT]). Let $X = \{X_n\}_{n\geqslant 0}$ be an HMC with values in the state space $E$ and transition matrix $M = (m_{ij})_{i,j\in E}$.
If the probabity distribution $\pi$ on E satisfies the *detailed balance equation* (DBE):

$$\pi(i)m_{ij} = \pi(j)m_{ji} \quad \forall i,j \in E, \tag{1.6}$$

then $\pi$ is stationary for $X$.

*Proof.* Summing on the all possible state and recalling that $M$ is a stochastic matrix, it follows that $\pi(i) = \sum_{j\in E}\pi(i)m_{ij} = \sum_{j\in E}\pi(j)m_{ji}$.

□

## 1.3 Strong Markov Property and Regeneration

In order to give a generalization of the Markov property let us introduce the notion of stopping time.

**Definition 1.3.1** (Stopping Time). Let $X = \{X_n\}_{n\geqslant 0}$ be a stochastic process.
A random variable $\tau$ with values in $\mathbb{N}\cup\{+\infty\}$ is called *stopping time*, with respect to $X$, if the event $\{\tau = m\}$ does depend only on $X_0, X_1, .., X_m$, i.e. $\{\tau = m\}$ depends on the trajectories of the process $X$ up to $m$.

**Example 1** (Return times). Let $X = \{X_n\}_{n\geqslant 0}$ be an HMC with values in $E$ and define
$T_i = inf\{n \geqslant 1 : X_n = i\}$ the *return time* to state $i$.
$T_i$ is a stopping time since we can write $\{T_i = m\} = \{\sum_{k=1}^{m}\mathbb{1}_{\{X_k=i\}} = 1, X_m = i\}$.

**Example 2** (Successive return times). In the same setting as Example 1, define $\{\tau_j\}_{j\geqslant 1}$ by

$$\tau_1 = T_i = inf\{n \geqslant 1 : X_n = i\}, \quad \tau_j = inf\{n > \tau_{j-1} : X_n = i\} \qquad \forall j \geqslant 2.$$

$\{\tau_j\}_{j\geqslant 1}$ is a sequence of stopping times, since we can write
$\{\tau_j = m\} = \{\sum_{k=1}^{m}\mathbb{1}_{\{X_k=i\}} = j, X_m = i\}$.

With these elements we can give the following theorem that allows us to generalize the Markov property, that is very useful in applications.

**Theorem 1.3.1** (Strong Markov Property). Let $X = \{X_n\}_{n\geqslant 0}$ be an HMC with values in the state space $E$ and transition matrix $M = (m_{ij})_{i,j\in E}$.
Let $\tau$ be a *finite* stopping time, with respect to $X$.
Then:

(i) $\forall\, i_0, i_1, .., i, j_1, j_2, .., j_k \in E$,

$$\begin{aligned} &P(X_{\tau+1} = j_1, X_{\tau+2} = j_2, .., X_{\tau+k} = j_k | X_\tau = i, X_{\tau-1} = i_{\tau-1}, .., X_0 = i_0) \\ =&P(X_{\tau+1} = j_1, X_{\tau+2} = j_2, .., X_{\tau+k} = j_k | X_\tau = i) \\ =&m_{ij_1} m_{j_1 j_2} \cdots m_{j_{k-1} j_k}; \end{aligned}$$

(ii) the process $Y = \{Y_n\}_{n\geqslant 0} = \{X_{\tau+n}\}_{n\geqslant 0}$ is an HMC with the same transition matrix $M = (m_{ij})_{i,j\in E}$ of $X$.

For the proof see [5].

Now, let us denote by

$$N_i = \sum_{n\geqslant 1} \mathbb{1}_{\{X_n = i\}} \tag{1.7}$$

the number of visits to state i, in particular after the initial state at time 0.

**Remark 2.** Consider an HMC starting from state 0 and such that $P_0(N_0 = +\infty) = 1$ and let the sequence $\tau_0 = 0, \tau_1 = T_0, \tau_2, ..$ be that one of the successive return times to 0.
By the Theorem (1.3.1), for any $k \geqslant 1$, the process after $\tau_k$ is independent of the process before $\tau_k$. In particular the process after $\tau_k$ is a Markov chain with the same transition matrix of the process that generated it.

Therefore, as a consequence of the Strong Markov Property we have the following

**Corollary 1.3.1.1** (Regenerative Cycles)**.** Let $\{X_n\}_{n\geqslant 0}$ be an HMC, with initial state 0 and such that $P_0(N_0 = +\infty) = 1$.
Let $(\tau_n)_{n\geqslant 0}$, $\tau_0 = 0$, be the successive return times to 0.
Then, the pieces of trajectories

$$\{X_{\tau_k}, X_{\tau_k+1}, .., X_{\tau_{k+1}-1}\}, \quad k \geqslant 0$$

called *regenerative clicles*, are independent and identically distributed.
In particular, the sequence $\{\tau_k - \tau_{k-1}\}_{k\geqslant 1}$ is i.i.d. .

## 1.4 Recurrence

Let us focus on the number of visits to a state.
Denoting by $f_{ji} = P_j(T_i < +\infty)$ the probability of the return time to state $i$

starting from $j$,then it follows

$$P_j(N_i = r) = \begin{cases} f_{ji} f_{ii}^{r-1}(1 - f_{ii}), & r \geqslant 1 \\ 1 - f_{ji}, & r = 0 \end{cases} \tag{1.8}$$

**Remark 3.** Observing that $P_i(N_i = r) = f_{ii}^r(1 - f_{ii})$, $r \geqslant 0$, it follows

$$P_i(T_i < +\infty) = 1 \quad \Longleftrightarrow \quad P_i(N_i = +\infty) = 1 \tag{1.9}$$

Furthermore,

$$\begin{aligned} E_i(N_i) &= \sum_{r \geqslant 1} r f_{ii}^r (1 - f_{ii}) \\ &= f_{ii} \sum_{r \geqslant 1} r f_{ii}^{r-1} (1 - f_{ii}) \\ &= f_{ii} E(Geo(1 - fii)) \\ &= \frac{f_{ii}}{1 - f_{ii}}, \end{aligned}$$

Therefore, we obtain the following equivalence

$$E_i(N_i) < +\infty \quad \Longleftrightarrow \quad P_i(T_i < +\infty) < 1. \tag{1.10}$$

It is possible to classify a state, with respect to the chain under consideration, depending on its number of visits. To this aim we give the following

**Definition 1.4.1** (Positive/Null-recurrent and transient state)**.** Consider a stochastic process with values in $E$ and let $T_i$ be first return time to state $i \in E$.
Then the state $i$ is called

1. *recurrent* if $P_i(T_i < +\infty) = 1$, in particular is said

   - *positive recurrent* if it also holds $E_i(T_i) < +\infty$

   - *null recurrent* if it also holds $E_i(T_i) = +\infty$

2. *transient* if $P_i(T_i < +\infty) < 1$.

Now we give a criterion for recurrence which is called the Potenial Matrix Criterion. Let $\{X_n\}_{n \geqslant 0}$ be an HMC with values in $E$ and with transition matrix $M = (m_{ij})$, we will

use the following notation

$$m_{ij}(n) = (M^n)_{ij} = P_i(X_n = j).$$

**Theorem 1.4.1** (Potential Matrix Criterion [PMC])**.** State $i \in E$ is recurrent if, and only if,

$$\sum_{n=0}^{+\infty} m_{ii}(n) = +\infty \tag{1.11}$$

*Proof.* Observing first that

$$P_i(T_i < +\infty) = 1 \quad \overset{remark3}{\Longleftrightarrow} \quad E_i(N_i) = +\infty$$

the thesis follows rewriting the condition (1.11), indeed

$$\begin{aligned}
\sum_{n=0}^{+\infty} m_{ii}(n) &= \sum_{n\geqslant 0} P_i(X_n = i) \\
&= \sum_{n\geqslant 0} E_i(\mathbb{1}_{\{X_n=i\}}) \\
&= E_i(\sum_{n\geqslant 0} \mathbb{1}_{\{X_n=i\}}) \\
&= E_i(N_i).
\end{aligned}$$

□

Continuing the study of the recurrence properties, we shall now prove that the recurrence is a (communication) class property.
Let us start with a definition.

**Definition 1.4.2** (Communicating states)**.** Given an HMC $X = \{X_n\}_{n\geqslant 0}$ with values in E, we say that the state $i$ and $j$ *communicate* if there exist a trajectory of $X$ from $i$ to $j$,
i.e. $\exists\, n$ s.t. $P_i(X_n = j) > 0$.

As a direct theoretical application of the PMC we have the following

**Theorem 1.4.2** (Recurrence is a class property)**.** If $i$ and $j$ communicate, they are either both recurrent or both transient.

Since in an irreducible HMC all the states communicate, it holds the following

**Corollary 1.4.2.1.** An *irreducible* HMC has all its states of the same nature: recurrent or transient.

In this sense we will say that a chain is recurrent or transient.

A slight generalization of the Potential Matrix Criterion is possible when we are treating an *irreducible* HMC, let us call $X$ such a chain.
Indeed, with the same notations of the Theorem 1.4.1, we have that

1. $X$ is *recurrent* if, and olnly if,

$$E_j(N_i) = \sum_{n \geqslant 1} m_{ji}(n) = +\infty \qquad \forall i, j \in E; \tag{1.12}$$

2. $X$ is *transient* if, and olnly if,

$$E_j(N_i) = \sum_{n \geqslant 1} m_{ji}(n) < +\infty \qquad \forall i, j \in E. \tag{1.13}$$

These observations are useful to obtain the following

**Theorem 1.4.3.** A *finite* (i.e. $|E| < +\infty$) and *irreducible* HMC is recurrent.

*Proof.* Assume by contradiction that the chain $X$ is transient.
By (1.13) it follows

$$\sum_{n \geqslant 1} m_{ji}(n) < +\infty, \forall i, j \in E.$$

Therefore, since $|E| < +\infty$, it holds

$$\sum_{i \in E} \sum_{n \geqslant 1} m_{ji}(n) < +\infty,$$

then, recalling that $M$ is a stochastic matrix, it is also true

$$\sum_{i \in E} \sum_{n \geqslant 1} m_{ji}(n) = \sum_{n \geqslant 1} \sum_{i \in E} m_{ij}(n) = +\infty.$$

□

The continuation of the theory that follows will be aimed at proving that a finite state irreducible HMC is positive recurrent.
This result is clear that it is quite understandable because, if it were not so, the states would be visited in a finite number of times and therefore after a random time no state would be visited anymore. Let us we now proceed formally.

**Theorem 1.4.4.** Let $X = \{X_n\}_{n\geqslant 0}$ be an *irreducible* and *recurrent* HMC with values in $E = \{0, 1, ..\}$ and transition matrix $M = (m_{ij})_{i,j\in E}$.
For $k \in E$, let $T_k$ be the return time to state $k$.
Set for any $i \in E$, $\bar{x}_i^k = E_k(\sum_{n\geqslant 1} \mathbb{1}_{\{X_n=i\}} \mathbb{1}_{\{n\leqslant T_k\}})$.
Then:

1. $\bar{x}_i^k > 0 \quad \forall i \in E$
2. $\bar{x}^k = (\bar{x}_i^k)_{i\in E}$ is a stationary measure for $M$.

Moreover, the invariant measure of such a chain is unique up to a multiplicative factor.

For the proof see [5].

**Remark 4.** Observe that

$$\begin{aligned}\sum_{i\in E}\sum_{n\geqslant 1} \mathbb{1}_{\{X_n=i\}} \mathbb{1}_{\{n\leqslant T_k\}} &= \sum_{n\geqslant 1}\left(\sum_{i\in E} \mathbb{1}_{\{X_n=i\}}\right) \mathbb{1}_{\{n\leqslant T_k\}} \\ &= \sum_{n\geqslant 1} \mathbb{1}_{\{n\leqslant T_k\}} \\ &= T_k.\end{aligned}$$

Therefore,

$$\sum_{i\in E} \bar{x}_i^k = E_k(T_k). \tag{1.14}$$

The following useful corollary is easily obtained from the Theorem 1.4.4 and the Remark 4.

**Corollary 1.4.4.1.** An *irreducible, recurrent* HMC is *positive recurrent* if, and only if,

$$\sum_{i\in E} y_i < +\infty$$

where $(y_i)_{i\in E}$ is an invariant measure for the chain.

**Remark 5.** Thanks to this last result, it follows by the normalization of the invariant measure, that an irreducible and positive recurrent HMC admits an invariant distribution, moreover which is unique.

There is also the converse part of the Remark 5, that we give it in the following

**Theorem 1.4.5.** An irreducible HMC which admits a stationary distribution is positive recurrent.

*Proof.* Let $M$ be the transition matrix of the chain and $\pi$ be the invariant distribution. Therefore,

$$\pi(i) = \sum_{j \in E} \pi(j) m_{ji}(n) \qquad \forall i \in E, \forall n \geqslant 1 \tag{1.15}$$

Assume that the chain is transient, then by the observation in (1.13) it follows that

$$\lim_{n \to +\infty} m_{ji}(n) = 0.$$

Taking the limit in (1.15), for any $i \in E$ we have

$$\begin{aligned} \pi(i) &= \lim_{n \to +\infty} \sum_{j \in E} \pi(j) m_{ji}(n) \\ &\overset{*}{=} \sum_{j \in E} \pi(j) \lim_{n \to +\infty} m_{ji}(n) = 0 \end{aligned}$$

in contradiction with the normalisation constraint.
Therefore, the chain must be recurrent and the thesis follows thank to the Corollary 1.4.4.1.
(The equality * follows from the Dominated Convergence theorem for series)

□

Finally, we have all the tools to demonstrated the following

**Theorem 1.4.6.** A *finite* state space *irreducible* HMC is positive recurrent.

*Proof.* A finite space HMC, which is irreducible, is recurrent.
Therefore, by Theorem 1.4.4 it admits an invariant measure that can be normalized since the space state is finite. The thesis follows by Theorem 1.4.5.

□

# 1.5 The Ergodic Theorem

We now want to present an important theorem, very powerful in applications, which is the Markov chain equivalent version of the Strong Law of Large Numbers (SLLN) recalled below, the Ergodic Theorem.

Let us start first with a definition.

**Definition 1.5.1** (Almost-surely convergence)**.** A sequence $(Z_n)_{n \geqslant 1}$ of real random variables *converges almost surely* to a real number $z$, with respect to probability measure $P$,

if it holds

$$P\Big(\lim_{n\to+\infty} Z_n = z\Big) = 1.$$

**Theorem 1.5.1** (SLLN)**.** Let $\{X_n\}_{n\geqslant 1}$ be an i.i.d. (independent and identically distributed) sequence of random variables, such that $E(|X_1|) < +\infty$.
Consider the empirical mean $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, then

$$\lim_{n\to+\infty} S_n \overset{P-a.s.}{=} E(X_1) \tag{1.16}$$

**Theorem 1.5.2** (Ergodic Theorem)**.** Let $\{X_n\}_{n\geqslant 0}$ be an *irreducible*, *positive recurrent* and *homogeneous* Markov chain with values in the state space $E$.
Let $\pi$ be the *invariant distribution* of the chain and $f : E \to \mathbb{R}$ be a function such that

$$E_\pi(|f|) = \sum_{i\in E} |f(i)|\pi(i) < +\infty.$$

Then, for any initial distribution $\mu$, it hold true the following:

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) \overset{P_\mu-a.s.}{=} E_\pi(f) = \sum_{i\in E} f(i)\pi(i), \tag{1.17}$$

where $P_\mu$ is the law of chain when the initial distribution is $\mu$.

*Proof.* Without loss of generality, take $f : E \to \mathbb{R}$, $f \geqslant 0$. The general case follows considering the positive and negative part of $f$, $f = f^+ - f^-$.
Let $\tau_0 \equiv 0, \tau_1 \equiv T_0, \tau_2, \ldots$ be the sequence of the successive return times to $0 \in E$ and define

$$\omega_p = \sum_{\tau_p+1}^{\tau_{p+1}} f(X_i), \quad p \geqslant 0.$$

We know that $\{\omega_p\}_{p\geqslant 0}$ is an i.i.d. sequence of random variables, which have finite mean

$$\begin{aligned}
E(\omega_0) &= E_0\left(\sum_{i=1}^{T_0} f(X_i)\right) \\
&= E_0\left(\sum_{i=1}^{T_0}\sum_{j\in E} f(X_j)\mathbb{1}_{\{X_i=j\}}\right) \\
&= \sum_{j\in E} f(X_j) E_0\left(\sum_{i=1}^{T_0} \mathbb{1}_{\{X_i=j\}}\right) \\
&= \sum_{j\in E} f(X_j)\bar{x}_j^0,
\end{aligned}$$

where $(\bar{x_j}^0)_{j\in E}$ is the invariant measure defined in the Theorem 1.4.4.
Now we can apply the SLLN to say

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{\tau_{n+1}} f(X_i) = \lim_{n\to+\infty}\frac{1}{n}\sum_{p=0}^{n}\omega_p \overset{P_\mu-a.s.}{=} \sum_{j\in E} f(X_j)\bar{x_j}^0. \tag{1.18}$$

Observing that

$$\tau_{g(n)} \quad \leqslant \quad n \quad \leqslant \quad \tau_{g(n)+1} \tag{1.19}$$

where $\tau_{g(n)} = \left(\sum_{k=1}^{n} \mathbb{1}_{\{X_k=0\}}\right)$, we can write

$$\frac{\sum_{k=1}^{\tau_{g(n)}} f(X_k)}{g(n)} \quad \leqslant \quad \frac{\sum_{k=1}^{\tau_n} f(X_k)}{g(n)} \quad \leqslant \quad \frac{\sum_{k=1}^{\tau_{g(n)+1}} f(X_k)}{g(n)}.$$

Therefore,
taking the limits for $n \to +\infty$ and observing that $g(n) \overset{n\to+\infty}{\longrightarrow} +\infty$ since the chain is positive recurrent, it follows from (1.18)

$$\lim_{n\to+\infty}\frac{1}{g(n)}\sum_{k=1}^{n} f(X_k) \overset{P_\mu-a.s.}{=} \sum_{j\in E} f(j)\bar{x_j}^0 \tag{1.20}$$

Now, taking $f \equiv 1$, the latter equality implies that

$$\lim_{n\to+\infty}\frac{n}{g(n)} \overset{P_\mu-a.s.}{=} \sum_{j\in E}\bar{x_j}^0 \tag{1.21}$$

Therefore

$$\begin{aligned}
\lim_{n\to+\infty}\sum_{k=1}^{n} f(X_k) &= \lim_{n\to+\infty}\frac{g(n)}{n}\frac{1}{g(n)}\sum_{k=1}^{n} f(X_k) \\
&\overset{\mathrm{P}_\mu-\mathrm{a.s.}}{=} \sum_{j\in E} f(j)\frac{\bar{x_j}^0}{\sum_{j\in E}\bar{x_j}^0} \\
&= \sum_{j\in E} f(j)\pi(j) \\
&= E_\pi(f)
\end{aligned}$$

□

# Chapter 2

# A MCMC Method: The Metropolis-Hasting Algorithm

## 2.1 General Principle and MCMC methods

Let us start from the following problem.
Consider the configuration space $E_N = \Lambda^{S_N}$, where $\Lambda = \{-1, +1\}$ and the set of the sites $S_N = \{1, .., N\}$ is finite. Take a distribution $\pi = (\pi(i))_{i=1,..,|E_N|}$ on $E_N$ and a function $f : E_N \to \mathbb{R}$.
Suppose we want to compute numerically the average

$$E_\pi(f) = \sum_{i \in E_N} f(i)\pi(i) \tag{2.1}$$

In the vast cases of applications, it is not taken granted the previous computation is effectively numerically possible. For instance:
- the cardinality $|E_N| = 2^N$ may be very large even $N$ is not,
assuming $N = 100 \Rightarrow |E_N| \gtrsim 10^{30}$;
- the distribution $\pi$ is known only up to a normalizing factor,
we will see in the chapter 3 that this is the case of a Gibbs distribution , for which the partition function is usually uncomputable in closed form.
In order to overcome such a difficulty we can think of giving up the exact result of the average and proceed through ergotic estimates. To this aim, a powerful tools are the *Makov chain Monte Carlo* (MCMC) *methods*.

MCMC methods are algorithms to build an homogeneous Markov chain (HMC) $\{X_n\}_{n \geqslant 0}$ with a given stationary measure/distribution $\pi$.

Therefore if we are able to construct an irreducible HMC, which is an ergodic chain by the finiteness of the state space $E_N$, thanks to the Ergodic Theorem we can estimate

the average through

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{n=1}^{N} f(X_n) \overset{P_\mu a-s}{=} E_\pi(f) \tag{2.2}$$

for any initial distribution $\mu$.

In the sequel we will see the Metropolis-Hastings algorithm, a particular form of a MCMC methods.

## 2.2 Metropolis-Hasting Algorithm

*Idea*: starting from *some* homogeneous chain, i.e. an *arbitrary* transition matrix, on a finite state space $E$ and a given distribution $\pi$ we want modify the chain so that the new chain obtained has as stationary distribution $\pi$.

Let us take an arbitrary *irreducible* transition matrix $Q = (q_{ij})_{i,j\in E}$,
called the $candidate-generating$ matrix, and a collection of probabilities $(\alpha_{ij})_{i,j\in E}$.
When the present state is $i$, the MCMC algorithm has the following steps:

*Step* 1. (Proposal) The next $tentative$ state $j$ is chosen with probability $q_{ij}$.

*Step* 2. (Acceptance-Rejection) If $j \neq i$, the new state $j$ is *accepted* with probability $\alpha_{ij}$ and *rejected*, i.e. the chain does not move from i, with probability $1-\alpha_{ij}$.

According to these two steps we can construct the transition matrix $\mathbf{P} = (p_{ij})_{i,j\in E}$ defined by:

$$p_{ij} = \begin{cases} q_{ij}\alpha_{ij} & \text{if } j \neq i, \\ 1-\sum_{z\neq i} q_{iz}\alpha_{iz} & \text{if } j = i. \end{cases} \tag{2.3}$$

Indeed, for $j = i$, we have

$$\begin{aligned} p_{ii} &= q_{ii} + \sum_{z\neq i} q_{iz}(1-\alpha_{iz}) \\ &= \sum_{z\in E} q_{iz} - \sum_{z\neq i} q_{iz}\alpha_{iz} \\ &= 1 - \sum_{z\neq i} q_{iz}\alpha_{iz}. \end{aligned}$$

In order to have $\pi$ as stationary distribution of $\mathbf{P}$ we impose the DBE with the quite general

form, due to Hastings [6], of the acceptance probabilities

$$\alpha_{ij} = \frac{s_{ij}}{1+t_{ij}} \tag{2.4}$$

where $S = (s_{ij})_{i,j\in E}$ is a *symmetric* matrix:

$$\pi(i)p_{ij} = \pi(j)p_{ji} \tag{2.5}$$

$$\pi(i)q_{ij}\frac{s_{ij}}{1+t_{ij}} = \pi(j)q_{ji}\frac{s_{ji}}{1+t_{ji}}$$

$$(1+t_{ji})\pi(i)q_{ij} = (1+t_{ji})\pi(j)q_{ji}$$

In conclusion we can take

$$t_{ij} = \frac{\pi(i)q_{ij}}{\pi(j)q_{ji}} \tag{2.6}$$

and therefore the DBE (2.5) is satisfied.
To conclude, it remains to impose the constraints $\alpha_{ij} \in [0,1]$, so one must have

$$s_{ij} \leqslant 1 + (t_{ij} \wedge t_{ji})$$

since the symmetry of $S$.
Taking equality, which corresponds to the Metropolis algorithm [7], we obtain

$$\alpha_{ij} = 1 \wedge \frac{\pi(j)q_{ji}}{\pi(i)q_{ij}}. \tag{2.7}$$

**Remark 6.** It is important to observe that in the construction of **P** through the Hastings algorithm, the dependence on $\pi$ is only on the ratio $\pi(j)/\pi(i)$ and therefore to apply this method it is sufficient to know $\pi$ only up to a multiplicative constant.

**Remark 7.** If we take an *irreducible* candidate-generating matrix $Q$, we will obtain that **P** is *irreducible*.

**Metropolis-Hastings algorithm**

**Input:** - finite state space $E$
- target distribution $\pi$ on $E$
- candidate-generating irreducible transition matrix $Q = (q_{ij})_{i,j\in E}$

**Output:** - transition matrix $\mathbf{P} = (p_{ij})_{i,j\in E}$, defined by

$$p_{ij} = \begin{cases} q_{ij}\big(1 \wedge \frac{\pi(j)q_{ji}}{\pi(i)q_{ij}}\big) & \text{if } j \neq i \\ 1 - \sum_{z\neq i} q_{iz}\big(1 \wedge \frac{\pi(z)q_{zi}}{\pi(i)q_{iz}}\big) & \text{if } j = i \end{cases}$$

# Chapter 3

# Markov Random Fields and Gibbs Fields

## 3.1 Neighborhoods and Local Specifications

We now want to extend the concept of Markov chains on the size of the working space, that will allow us to model more complex situations in which the 1-dimensionality of chains may not be enough. To this aim we present the random fields, as we did for a generic stochastic process, and we will then give the markovianity conditions on them.

**Definition 3.1.1** (Random Field)**.** Let $S$ be a $finite$ set, with elements denoted by $s$ and called *sites*, and let $\Lambda$ be a finite set called the *phase space*. A *random field* on $S$ with *phase space* in $\Lambda$ is a collection $X = \{X(s)\}_{s\in S}$ of random variables $X(s)$ with values in $\Lambda$.

A random field can be regarded as a random variable taking its values in the $configuration$ *space* $\Lambda^S$.

A natural generalization of the Markov property comes out observing that the Markov property of a stochastic sequence $\{X_n\}_{n\geqslant 0}$ implies that for all $n \geqslant 1$, $X_n$ is independent of $(X_k, k \notin \{n-1, n, n+1\})$ given $(X_{n-1}, X_{n+1})$. In order to make rigorous this property for Random fields let's introduce a $topology$ on the sites.

**Definition 3.1.2** (Neighborhoods)**.** A *Neighborhood system* on $S$ is a family $N = \{\mathcal{N}_s\}_{s\in S}$ of subsets of $S$ such that for all $s \in S$,

(i) $s \notin \mathcal{N}_s$

(ii) $t \in \mathcal{N}_s \Rightarrow s \in \mathcal{N}_t$

The subset $\mathcal{N}_s$ is called the *neighborhood* of sites $s$,
the couple $(S, N)$ is called *topology*,
the *boundary* of $A \subset S$ is, by definition, the set $\partial A = (\bigcup_{s\in A} \mathcal{N}_s) \backslash A$.

**Definition 3.1.3** (Markov Random Filed)**.** The random field $X$ is called a *Markov random filed* (MRF), with respect to the neighborhood system $N$, if for all sites $s \in S$ the random variables $X(s)$ and $X(S\backslash\tilde{\mathscr{N}_s})$ are independent given $X(\mathscr{N}_s)$,
that is

$$P(X(s) = x(s)|X(S\backslash s) = x(S\backslash s)) = P(X(s) = x(s)|X(\mathscr{N}_s) = x(\mathscr{N}_s)) \tag{3.1}$$

$\forall\ s \in S, x \in \Lambda^S$,
where $\tilde{\mathscr{N}_s} := \mathscr{N}_s \cup \{s\}$.

Markov fields are characterized by local interactions.

**Definition 3.1.4** (Local Specification)**.** The *local characteristic* of the MRF at sites $s$ is the function $\pi^s : \Lambda^S \to [0,1]$ defined by

$$\pi^s(x) = \pi^s(x(s)|x(\mathscr{N}_s)) = P(X(s) = x(s)|X(\mathscr{N}_s) = x(\mathscr{N}_s)) \tag{3.2}$$

The family $\{\pi^s\}_{s\in S}$ is called the *local specification* of the MRF.

Let us see with the following condition that the distribution of a MRF is univocally determined by the local specification.

**Definition 3.1.5** (Positivity Condition)**.** The probability distribution $\pi$ on the finite configuration space $\Lambda^S$, where $S = \{1,2,..,K\}$, is said to satisfy the *positivity condition*
if $\forall\ j \in S, x_j \in \Lambda$,

$$(\pi_j(x_j) = 0) \Rightarrow (\pi(y_1,..,y_{j-1},x_j,y_{j+1},..,y_K) = 0) \tag{3.3}$$

$\forall\ y_1,..,y_K \in \Lambda$, where $\pi_j$ is the marginal distribution of site $j$.

**Theorem 3.1.1** (At Most One Distribution for a Local Specification)**.** Two distributions of a MRF with a *finite* configuration space $\Lambda^S$, where $S = \{1,2,..,K\}$, that satisfy the positivity condition and have the same local specification are identical.

*Proof.* Let $x, y$ be two configurations in $\Lambda^S$ with non-null probability and $\pi$ the distribution of the field.
The positivity condition and $\pi(x), \pi(y) > 0$ imply that: $\forall j \in \{1,..,K\}$,
$\pi(x_1,..,x_j,y_{j+1},..,y_K) > 0$. Therefore, using Bayes's formula, we have:

$$\pi(x) = \frac{\pi(x_k|x_1,..,x_{k-1})}{\pi(y_k|x_1,..,x_{k-1})}\pi(x_1,..,x_{k-1},y_k)$$

First iterating the same argue, thanks to the finiteness of $S$, and then using the Markovianity of the field we obtain:

$$\pi(x) = \prod_{i=1}^{K} \frac{\pi(x_i|x_1,..,x_{i-1},y_{i+1},..,y_K)}{\pi(y_i|x_1,..,x_{i-1},y_{i+1},..,y_K)}\pi(y)$$

$$= \prod_{i=1}^{K} \frac{\pi^i(x_1,..,x_i,y_{i+1},..,y_K)}{\pi^i(x_1,..,x_{i-1},y_i,..,y_K)}\pi(y)$$

Therefore, if the local specification is specified and $y$ is fixed, then $\pi$ is determined up to the factor $\pi(y)$, which is determined by normalization.

□

Note that Theorem 3.1.1 is not true in general for an infinite number of sites, and this is why the basic assumption on the finiteness of $S$ was recalled. We will see, specified to the Ising model dealing the *phase transition* that where $S$ is infinite there may be several distributions corresponding to a given local specification.

## 3.2 Cliques, Potential, and Gibbs Distributions

Consider the probability distribution on the configuration space $\Lambda^S$

$$\pi_T(x) = \frac{1}{Z_T} e^{-\frac{1}{T}\varepsilon(x)} \tag{3.4}$$

where $T > 0$ is the *temperature*, $\varepsilon(x)$ is the *Energy function* of a configuration $x$ and $Z_T$ is the normalizing constant called *Partition function*.

The definition of such a measure, called under some conditions Gibbs Measure, is particularly useful when we want to describe, from a mathematical point of view, an equilibrium state of a physical system which consists of a very large number of interacting components.

To manage the interactions of neighbors let us define, in this environment, what a clique is.

**Definition 3.2.1** (Clique)**.** Any singleton $\{s\}$ is a *clique*. A suset $C \subset S$, $|C| \neq 1$, is a *clique* of the topology $(S,N)$ if $\forall\ s,t \in S$, $s \neq t$, $s \in \mathcal{N}_t$ and $t \in \mathcal{N}_s$.

**Definition 3.2.2** (Gibbs Potential, Gibbs Distribution and Gibbs field)**.** A *Gibbs Potential* on $\Lambda^S$ relative to the neighborhood system $N$ is a collection $\{V_C\}_{C\subset S}$ of functions $V_C : \Lambda^S \to \mathbb{R}\cup\{+\infty\}$ such that

(i) $V_C \equiv 0$ if $C$ is not a clique,

(ii) $\forall\, x,\bar{x}\in\Lambda^S$, $\forall\, C\subset S$,

$$x(C)=\bar{x}(C) \quad \Rightarrow \quad V_C(x)=V_C(\bar{x}) \tag{3.5}$$

The energy function $\varepsilon : \Lambda^S \to \mathbb{R}\cup\{+\infty\}$ is said to derive from the potential $\{V_C\}_{C\subset S}$, if

$$\varepsilon(x)=\sum_C V_C(x) \tag{3.6}$$

In this context, the distribution in (3.4), characterized by an energy which derives from a Gibbs potential is called a *Gibbs distribution*. Similarly a MRF with a distribution equal to a Gibbs distribution is called *Gibbs field*.

## 3.3 Gibbs-Markov Equivalence

Gibbs distributions, characterized by Gibbs potentials defined on a topology, are distributions of Markov field with respect to the same topology.

**Theorem 3.3.1** (Gibbs field are Markov field)**.** If $X$ is a random field with a Gibbs distribution $\pi$, relative to the neighborhood system $N$, then $X$ is Markovian with respect to the same neighborhood system $N$.
Moreover, its local specification is given by the formula

$$\pi^s(x)=\frac{e^{-\sum_{C\ni s}V_C(x)}}{\sum_{\lambda\ni\Lambda}e^{-\sum_{C\ni s}V_C(\lambda,x(S\backslash s))}}=\frac{e^{-\varepsilon_s(x)}}{\sum_{\lambda\ni\Lambda}e^{-\varepsilon_s(\lambda,x(S\backslash s)}}, \tag{3.7}$$

where $\varepsilon_s(x)=\sum_{C\ni s}V_C(x)$ is the *local energy* at site $s$ of configuration $x$.

**Remark 8.** The right-hand side of (3.7), by the point (ii) of the Definition 3.2.2, depends on $x$ only through $x(s)$ and $x(\mathcal{N}_s)$.
Indeed, $V_C(x)$ depends only on $(x(t), t\in C)$ and if $t,s\in C$ clique, it follows that $t=s$ or $t\in\mathcal{N}_s$.

*Proof.* Thanks to the remark 8, if one can show that $P(X(s)=x(s)\mid X(S\backslash s)=x(S\backslash s))$ equals the right-hand side of (3.7), the Markov property and the equality (3.7) will be

proven.
By definition and from the law of total probability,

$$P(X(s) = x(s) \mid X(S\backslash s) = x(S\backslash s)) = \frac{\pi(x)}{\sum_{\lambda\in\Lambda}\pi(\lambda, x(S\backslash s))}. \tag{3.8}$$

Moreover, we can write

$$\pi(x) = \frac{1}{Z}e^{-(\sum_{C\ni s}V_C(x)+\sum_{C\not\ni s}V_C(x))}$$

and similarly

$$\begin{aligned}\pi(\lambda, x(S\backslash s)) &= \frac{1}{Z}e^{-(\sum_{C\ni s}V_C(\lambda,x(S\backslash s))+\sum_{C\not\ni s}V_C(\lambda,x(S\backslash s)))}\\ &= \frac{1}{Z}e^{-(\sum_{C\ni s}V_C(\lambda,x(S\backslash s))+\sum_{C\not\ni s}V_C(x))},\end{aligned}$$

where the last equality follows from the fact that if $C$ is a clique and $s \notin C$, then $V_C(\lambda, x(S\backslash s)) = V_C(x)$ and in particular is independent of $\lambda \in \Lambda$.
Therefore, the formula (3.7) follows factoring out $e^{-\sum_{C\not\ni s}V_C(x)}$.

□

This result admits a converse part, which is important from a theoretical point of view, because we can conclude that the two object apparently different, MRFs and Gibbs fields, are essentially the same object.

**Theorem 3.3.2** (Hammerslet-Clifford Theorem)**.** Let $\pi$ be the distribution of a MRF with respect to a topology $(S,N)$ satisfying the positivity condition. Then

$$\pi(x) = \frac{1}{Z}e^{-\varepsilon(x)}$$

for some energy function $\varepsilon(x)$ deriving from a Gibbs potential associated with the same topology $(S,N)$.

For the proof see [5].

# Chapter 4

# 2D Ising Model

In statistical physics the following model is regarded as a qualitatively correct idealization of a piece of ferromagnetic material, indeed it was introduced by Lenz in 1920, solved in 1D by Ising in 1925, for understanding the phenom of phase transition.

## 4.1 Ising's finite model

Let us defined
$S = \mathbb{Z}_m^2 = \{(i,j) \in \mathbb{Z}^2 : i,j \in [1,m]\}$ the set of the *sites* which represents the piece of material whose sites represent particles;
$\Lambda = \{+1,-1\}$ the *phase space* which represents the orientation of the magnetic spin at given site and
$N = \{\mathcal{N}_s\}_{s\in S}$ the *neighborhood system*, where
$\mathcal{N}_s = \{(s_1,y_2),(y_1,s_2) \in S : |s_2 - y_2| = 1, |y_1 - s_1| = 1\}$, it will be useful to describe the local interactions of the particles.

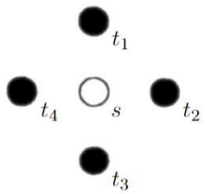


**Figure 4.1:** $\mathcal{N}_s = \{t_1,t_2,t_3,t_4\}$.

The *Gibbs potential* is

$$V_{\{s\}}(x) = -\frac{H}{k}x(s)$$

$$V_{<s,t>}(x) = -\frac{J}{k}x(t)x(s)$$

where $< s,t >$ is the 2-element clique, $t \in \mathcal{N}_s$. Here $k$ is the Boltzmann constant, $H$ is the external magnetic field and $J$ is the internal energy of an elementary magnetic dipole. Therefore, by (3.6) the *Energy function* associated to this potential is:

$$\varepsilon(x) = -\frac{J}{k}\sum_{<s,t>} x(s)x(t) - \frac{H}{k}\sum_{s\in S} x(s) \tag{4.1}$$

As a consequence of the Theorem 3.3.1, for the Ising model the *local energy* at site s is

$$\varepsilon_s(x) = -\frac{1}{k}\Big(J\sum_{t;|s,t|=1} x(t) + H\Big)x(s) \tag{4.2}$$

and the *local characteristics* are

$$\pi_T^s(x) = \frac{e^{\frac{1}{kT}\left(J\sum_{t;|s,t|=1} x(t)+H\right)x(s)}}{e^{+\frac{1}{kT}\left(J\sum_{t;|s,t|=1} x(t)+H\right)} + e^{-\frac{1}{kT}\left(J\sum_{t;|s,t|=1} x(t)+H\right)}} \tag{4.3}$$

## 4.2 Phase Transitions

In a slightly generalization of the Ising model, with the distribution $\pi_T$ associated to the spins that is a Gibbs distribution of the form:

$$\pi_T(x) = \frac{1}{Z_T}e^{-\frac{1}{T}\varepsilon(x)}, \tag{4.4}$$

let us consider the energy function

$$\varepsilon(x) = \varepsilon_0(x) - \frac{H}{k}\sum_{i=1}^{N} x(i) \tag{4.5}$$

where the term $\varepsilon_0(x)$ is considered symmetric on configurations $x \in \Lambda^S =: E$, with S enumerated as $S = \{1,..,N\}$.
The *magnetic moment* of configuration $x$ is

$$m(x) = \sum_{i=1}^{N} x(i) \tag{4.6}$$

and the *magnetization* is the average magnetic moment per site

$$M(H,T) = \frac{1}{N}\sum_{x\in E} \pi_T(x)m(x) = \frac{1}{N} < m > . \tag{4.7}$$

Let us now see some some facts about the magnetization:

1. is an odd function of H:

$$M(-H,T) = -M(H,T). \tag{4.8}$$

Indeed

$$\begin{aligned}\varepsilon(x) &= \varepsilon_0(x) - \frac{H}{k}\sum_{i=1}^{N} x(i) \\ &= \varepsilon_0(x) - \frac{(-H)}{k}\big(-\sum_{i=1}^{N} x(i)\big) \\ &= \varepsilon_0(x) - \frac{(-H)}{k}(-m(x)),\end{aligned}$$

therefore

$$\begin{aligned}M(-H,T) &= \frac{1}{N}\sum_{x\in E}\pi_T(x)\big(-m(x)\big) \\ &= -\frac{1}{N}\sum_{x\in E}\pi_T(x)m(x) \\ &= -M(H,T).\end{aligned}$$

2.

$$-1 \leqslant M(H,T) \leqslant +1 \tag{4.9}$$

3. is a nondecreasing function of H:

$$\frac{\partial M(H,T)}{\partial H} \geqslant 0 \tag{4.10}$$

Indeed, take $E(x) = E_0(x) + E_1(x)$, where $E_0(x) = k\varepsilon_0(x)$ is the interaction energy (assumed symmetric) and $E_1(x) = -Hm(x)$.
The partition function $Z$, is now a function of $T$ and $H$.
The free energy per site is $f(H,T) = -kT\frac{1}{N}ln(Z)$.

Observing that,

$$\begin{aligned}
-\partial_H f(H,T) &= kT\frac{1}{N}\frac{1}{Z}\partial_H Z \\
&= kT\frac{1}{N}\frac{1}{Z}\sum_{x\in E} e^{-\frac{E_0(x)-Hm(x)}{kT}}\frac{m(x)}{kT} \\
&= \frac{1}{N}\sum_{x\in E}\frac{1}{Z}e^{-\frac{E(x)}{kT}}m(x) \\
&= \frac{1}{N} < m > \\
&= M(H,T),
\end{aligned}$$

it holds true

$$\begin{aligned}
\partial_H M(H,T) &= -\partial_H^2 f(H,T) \\
&= kT\frac{1}{N}\left(-\frac{1}{Z^2}(\partial_H Z)^2 + \frac{1}{Z}\partial_H^2 Z\right) \\
&= \frac{1}{NkT}\left(- < m >^2 + < m^2 >\right) \geqslant 0.
\end{aligned}$$

4.

$$M(0,T) = 0 \tag{4.11}$$

since $\pi_T(-x) = \pi_T(x)$ when $H = 0$ and $-m(x) = m(-x)$,
therefore the sum in (4.7) is null.

5. moreover, the magnetization is an *analytic* function of $H$.

However, the experiments seem to contradict the previous two last points. Indeed if an iron bar is placed in a strong magnetic field $H$ parallel to the axis, it is completely magnetized $\big(M(H,T) = +1\big)$, and if $H$ is slowly decreased to 0 one can be observe that M decreases, but tends to a limit $M(0,T) = M_0 > 0$, called *spontaneous magnetization*, in contradiction with the point 4. Moreover, from the point 1 we have a discontinuity of $M$ at $H = 0$, in disagreement with the analyticity of M with respect to $H$.

This discontinuity, called *phase transition* by physicists, occurs at room temperature and if the temperature is increased, the spontaneous magnetization $M_0$ decreases until it reaches the value 0 at a certain temperature $T_c$, called *critical temperature* (see Figure 4.2 below).

This gap that can be noted between the experience and the theory, below the critical temperature, is due to the fact that the experimental results describe the situation in a *ther-*

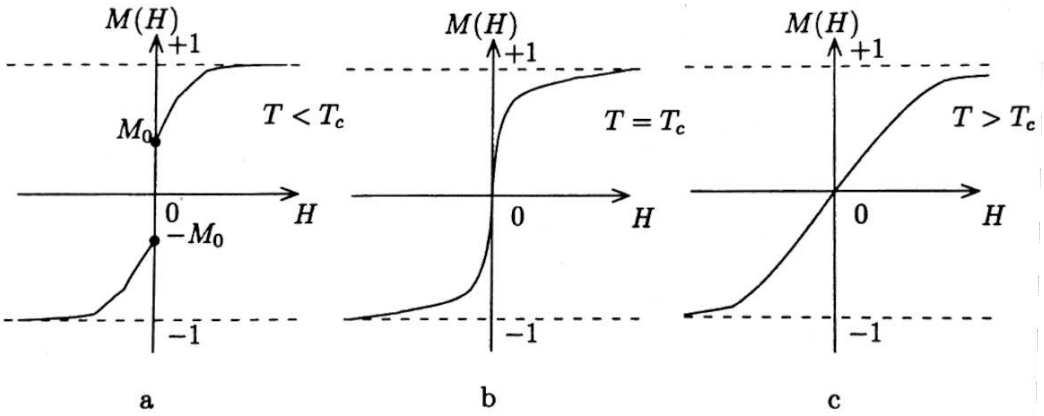


**Figure 4.2:** Magnetization and critical temperature [5]

*modynamical limit* $N = +\infty$. For fixed but sufficient large $N$ the theoretic magnetization curve is analytic, but it present for all practical purposes the same aspect as if it had a discontinuity.

Therefore, looking at the experimental results, it seems that, below the critical temperature, the spontaneous magnetization has two possible "choices" when no external magnetic field is applied.
We shall now explain this phenomenon within the Ising model with the Peierls's Argument.

### 4.2.1 Peierls's Argument

In the Ising model, considering the situation when no external magnetic field is applied, let us set $H = 0$. In this case the energy function is

$$\varepsilon(x) = -J \sum_{<s,t>} x(s)x(t) \tag{4.12}$$

where $< s,t >$ represents an unordered pair of neighbors.
It is clear that the previous sum is well defined when $S = \{1,..,N\}$ is finite, but when the cardinal of $S$ is infinite, in general is not defined for all configurations, and therefore is not possible to define the Gibbs distribution $\pi_T$ on $\Lambda^S$ by the formula (3.4). Nevertheless, the local specification, according to the formula (4.3)

$$\pi_T^s(x) = \frac{e^{\beta \sum_{|s,t|=1} x(t)x(s)}}{e^{+\beta \sum_{|s,t|=1} x(t)} + e^{-\beta \sum_{|s,t|=1} x(t)}} \tag{4.13}$$

where $\beta$ is ,up to a multiplicative factor, the inverse temerature,
is well-defined $\forall x \in \Lambda^S$, since the sum is finite even $S$ is not.

A probablity distribution $\pi_T$ on $\Lambda^S$, which admits a local specification as in (4.13), is called a *solution of DRL problem* (Dobrushin (1965), Ruelle and Lanfrod (1969)).

Therefore, when $S$ is finite, we know that there exists a unique solution given by (4.4), since the sum in (4.12) is well-defined.
When $S = \mathbb{Z}^2$, the existence of at least one solution of the DRL problem has be proven by Dobrushin. One way of constructing a solution is to start from a solution on a finite configuration space, for example taking as the set of sites $K_N = \mathbb{Z}^2 \cap [-N, +N]^2$, and then let $N$ tend to infinity. More precisely:
let us take an arbitrary configuration $z$ and the probability distribution $\pi_T^{(N)}$ that has the local specification (4.13) on $\Lambda^{K_{N-1}}$, then we extend it to $\Lambda^S$ considering the field frozen at the configuration $z$ outside $K_{N-1}$, i.e. $\pi_T^{(N)}(z(S \backslash K_{N-1})) = 1$.
Now we can obtain a solution $\pi_T$ of the DRL problem as follows:
$\forall\, x \in \Lambda^S$ and $\forall\, A \subset S, |A| < +\infty$, exists the following limit

$$\lim_{N \to +\infty} \pi_T^{(N)}(x(A)) = \pi_T(x(A)) \tag{4.14}$$

and moreover, there exists a unique MRF, let us it call $X$, with local specification (4.13) and characterized by the probability distribution

$$P(X(A) = x(A)) = \pi_T(x(A)) \tag{4.15}$$

$\forall\, x \in \Lambda^S$ and $\forall\, A \subset S, |A| < +\infty$.

**Remark 9.** Note that $\pi_T^{(N)}$ depends on $z$ only through the restriction of $z$ to the boundary $\partial K_N = K_N \backslash K_{N-1}$.

If the DRL problem admits more than one solution, we will say that a *phase transition* occurs.

In the sequel, it will be clear as the method that we have seen before to construct a solution, can be useful also to prove the existence of the phase transition when it occurs. It is sufficient to show that for two different configurations $z_1$, $z_2$ and for a given finite subset $A \subset S$, the left-hand of (4.14) is not uniquely determined since is different for $z = z_1$ and for $z = z_2$.

Peierls applied the above program with $z_1 = \{+1\}^S$, $z_2 = \{-1\}^S$ and $A = \{0_{\mathbb{Z}^2}\}$ to prove that for sufficiently small values of $T$, there is a phase transition.

Let us introduce a few notations and definitions that will be useful to our aim .
We will indicate with $\pi_+^{(N)}$ (risp. $\pi_-^{(N)}$) the restriction to $K_N$ of the probability distribution $\pi_T^{(N)}$ when $z = z_1$ (risp. $z = z_2$).
We will call *even* (risp. *odd*) *bound* of a configuration $x \in \Lambda^{K_N}$ a clique $< s,t >$ such that $x(s)x(t) = 1$ (risp. $x(s)x(t) = -1$) and we will indicate with $n_e(x)$ (risp. $n_o(x)$) the number of even (risp. odd) bounds.
For a fixed configuration $x$, we will denote with $C(x;0)$ the circuit that is the boundary of the largest connected, in the sense of the topology, group of sites with negative phase, containing site 0. If $x(0) = +1$ then $C(x;0) = \varnothing$.
We will denote with $\tilde{x}$ the configuration obtained by $x$ reversing all the phases inside the circuit $C(x;0)$.

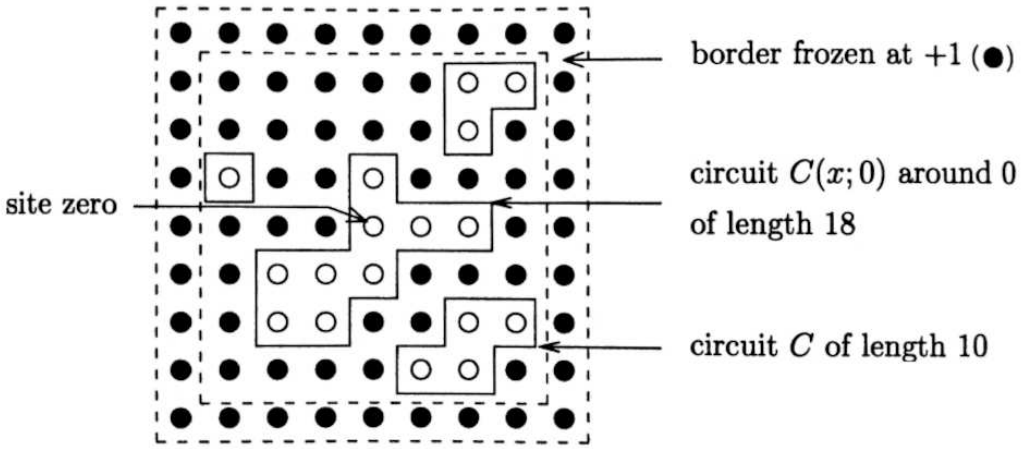


**Figure 4.3:** Circuits in the Ising model [5]

Let us then proceed more formally with Peierls's argument to demonstrate how the phase transition occurs if $T$ is large enough.
It is sufficient to prove that

$$\pi_+^{(N)}(x(0) = -1) < \frac{1}{3} \tag{4.16}$$

$\forall\, N \in \mathbb{N}$.
In fact, by symmetry of the problem, it holds true also that $\pi_-^{(N)}(x(0) = +1) < \frac{1}{3}$ and therefore $\pi_-^{(N)}(x(0) = -1) > \frac{2}{3}$.
Since the previous estimations are preserved passing to the limit $N \nearrow +\infty$, it implies

that $\pi_+(x(0)=-1) < \frac{1}{3}$ and $\pi_-(x(0)=-1) > \frac{2}{3}$, i.e. the limiting distributions are not identical.

*Proof.* Let $x$ be a configuration in $\Lambda^N$.
Observing that $-\sum_{<s,t>} x(s)x(t) = n_o(x) - n_e(x)$ and $n_e(x) = M - n_o(x)$, where $M$ is the total number of two-elements cliques, we can write

$$\pi_+^{(N)}(x) = \frac{e^{-2\beta n_0(x)}}{Z_+^{(N)}}, \tag{4.17}$$

where $Z_+^{(N)}$ is the normalization factor.
For a given circuit $C$ around 0, by the law of total probability we have

$$\pi_+^{(N)}(C(x;0)=C) = \frac{\sum_{x;C(x;0)=C} e^{-2\beta n_0(x)}}{\sum_y e^{-2\beta n_0(y)}} \leqslant \frac{\sum_{x;C(x;0)=C} e^{-2\beta n_0(x)}}{\sum_{x;C(x;0)=C} e^{-2\beta n_0(\tilde{x})}}.$$

Now, if $x$ is such that $C(x;0)=C$, it follows that $n_0(\tilde{x}) = n_0(x) - L$, where $L$ is the length of $C$.
Therefore

$$\pi_+^{(N)}(C(x;0)=C) \leqslant e^{-2\beta L}.$$

In particular

$$\pi_+^{(N)}(x(0)=-1) \leqslant \sum_{\substack{\text{C around 0}\\ \text{C}\neq\varnothing}} \pi_+^{(N)}(C(x;0)=C) \leqslant \sum_{\substack{L=\\ 4,6,..,2f(N)}} r(L)e^{-2\beta L}$$

where $r(L)$ is the number of nonempty circuits around 0 of length $L$ and $f(N)$ is s.t. $\lim_{N\to+\infty} f(N) = +\infty$.
Observe that

$$r(L) \leqslant 4L^2 3^L.$$

Indeed a circuit around 0 of length $L$ must have at least one point at distance $L/2$, or less, from 0; there are $L^2$ way of selecting such point and at most 4 ways of choosing the segment of $C$ starting from this point, then at most 3 ways of choosing the next connect segment and so on..., see [8] for more additional observations.

Therefore

$$\pi_+^{(N)}(x(0)=-1) \leqslant \sum_{\substack{L=\\ 4,6,..,2f(N)}} 4L^2(3e^{-2\beta})^L.$$

Now, $\sum_{L=4,6,..} L^2 x^L \leqslant \sum_{L=1}^{+\infty} L^2 x^L$ and the latter series has radius of convergence 1;
so we can conclude that
if $3e^{-2\beta}$ is small enough,or equivalently T is large enough, it holds

$$\pi_+^{(N)}(x(0) = -1) < \frac{1}{3}$$

$\forall\, N \in \mathbb{N}$.

□

## 4.3 Simulations via the Metropolis-Hastings Algorithm

The aim of this section is to show, with a more practical approach, that in the 2-D Ising model the phase transition occurs. Let us give some motivations that should better clarify the simulation that is presented below.

### 4.3.1 Motivations

In the 2D square-lattice Ising model, with the same notations introduced in the Section 4.1 and in the absence of the external magnetic field, let us set $J = 1$ for a simplified version of the Energy function, that results to be

$$\varepsilon(x) = -\sum_{<s,t>} x(s)x(t), \tag{4.18}$$

according to the formula (4.1).
Therefore, we can define the Gibbs distribution $\pi_\beta$ associated to the spins as

$$\pi_\beta(x) = \frac{1}{Z_\beta} e^{\beta \sum_{<s,t>} x(s)x(t)}, \tag{4.19}$$

where $\beta$ is up to a multiplicative factor the inverse temperature.

Our aim is to show that a phase transition occurs, with respect to the parameter $\beta$, estimating the mean of the square of the observable

$$\tilde{m}(x) = \frac{\sum_i x(i)}{N^2}, \qquad x \in \Lambda^S, \tag{4.20}$$

where $N$ in this case is the dimension of the set of the sites $S = \mathbb{Z}_N^2$.
Indeed, thinking of a piece of ferromagnetic material, represented by the lattice-square,

where the spins are distributed according to the probability distribution $\pi_\beta$, for large enough values of $\beta$, i.e. low temperature values, the system, for any initially possible configuration of the states, via only interactions between its pairs of spin, tends over time to minimize internal energy by arranging its configuration in the most orderly possible way, showing a complete magnetization.
On the other hand, for sufficient low values of $\beta$, i.e. large temperature values, the system is in a state of high excitement, it completely loses its magnetic properties and fails to arrange its configuration as in the opposite situation previously presented.

### 4.3.2 Simulations

Before simulating the model, we would point out our choice to work on the Ising model, where the set of sites is the torus-lattice-square $S = \mathbb{Z}_N^2$.

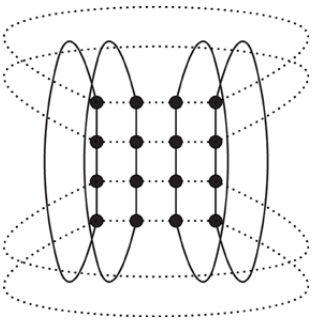

**Figure 4.4:** torus-lattice-square $\mathbb{Z}_4^2$

The idea behind the simulation is to generate a Markov chain with values in the state space $\Lambda^S$, $\Lambda = \{+1,-1\}$, which definitively is homogeneous and it has the stationary Gibbs distribution $\pi_\beta$ of the form (4.19), let us call $X = \{X_n\}_{n\geqslant 1}$ this homogeneous chain.
The first obstacle that we may expect, is to numerically calculate the partition function $Z_\beta$.
To overcome this aspect, one of the possible choices, is to simulate via the Metropolis-Hasting algorithm, we have seen in the Chapter 2 that is able to cope with this apparent problem.
We are now able, with the chain $X$ and thanks to the Ergodic Theorem 1.5.2, to estimate the mean of $\tilde{m}^2$.
Therefore, starting for any initial distribution, i.e. for any initial configuration of the spins, we obtain that

$$E_{\pi_\beta}\left(\tilde{m}^2\right) \approx \frac{1}{N}\sum_{k=1}^{N} \tilde{m}^2(X_k) \tag{4.21}$$

Let us illustrate the results of the simulations.

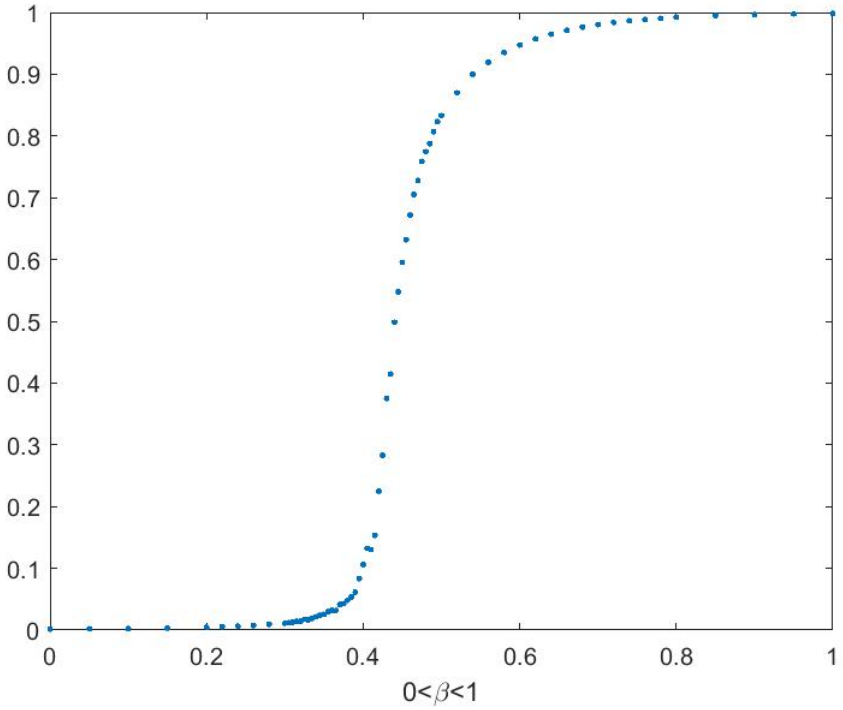


**Figure 4.5:** $S = \mathbb{Z}_{25}^2$

We can see how the behavior of the mean $E_{\pi_\beta}\left(\tilde{m}^2\right)$, as function of $\beta$, changes between the two opposite states 0 and $+1$, in accordance with the reasons we had made explicit above.

Indeed, the distribution $\pi_\beta$, is reduced to being an uniform distribution for small value of $\beta$. Moreover, with high probability, the spins configuration is randomly arranged and the observable $\tilde{m}$ has a null mean.

In the opposite situation, when $\beta$ is large enough, the largest values of the distribution $\pi_\beta$, are on the configurations with spins of the same value and vice versa.

Therefore the mean of $\tilde{m}^2$ gets closer to $+1$ as $\beta$ increases.

## Conclusions

The present thesis studies the Ising model both from a theoretical and simulation perspective. Historically the Ising model was initially proposed by the physicists Ernest Ising and Wilhelm Lenz as a model of spontaneous magnetization of materials, but nowadays possible applications of the Ising model and its variations are almost uncountable, ranging from physics to economics, from neurosciences to quantitative sociology. Indeed the Ising model became a paradigmatic model in complex system science. The reasons of its success are grounded on the fact although simple to define, the Ising model has a fairly complicated behavior, showing the phenomenon of phase transition.

In Chapters 1 and 3 theoretical results about Markov chains and Markov fields that are useful in the sequel of the thesis are collected. Here, an attempt has been made to make this part self-contained, presenting it not only as an instrumental. In this regard, we included almost all the proofs of the results presented in these chapters, as we skipped only few technical demonstrations.

In Chapter 2, we present the Metropolis-Hasting algorithm in its most general form, specifying the peculiarities that give its own importance.

In Chapter 4, we define the Ising model and study the phenomenon of phase transition using a probabilistic language. In particular we focus on the 2D Ising model and we prove the existence of a phase transition using Peierls' contours. This famous Peierls' argument is an elegant and effective geometric-combinatorial approach to the problem, generalizable to more complex models, which is effective to prove the existence of the phase transition. To compute the value of the critical point of the model is a much more difficult task and it is out of the scopes of the present work. However, in the last part of the thesis, we simulated the Ising model through the Metropolis-Hasting algorithm, via the Matlab platform. The results of our simulations are compatible with a critical value of $\beta(\approx 0.44)$, as predicted by the Onsager's solution of the 2D Ising model.

A further study on these subjects could be about the latest more practical aspect of the simulations. In order to be able to verify via simulations a sharper phase transition and, thus, a more accurate value of the critical point for which it occurs, optimizing the implemented algorithm could be useful for applications. An other aspect that could be considered, could be the ratio between the size of the model and the number of iterations of the algorithm necessary to produce the results and more generally the study of computational complexity.

## Matlab Code

```
Ndim=%25;
Niter=%9.*10.*^6;
Beta=%0:.02:1;

iter=1:Niter;
X=(-1).^randi(2,Ndim);
M=[];
for beta=Beta
    Y=X;
    m=[];
    for k=iter
        i=randi(Ndim);
        j=randi(Ndim);
        v=[];
            if (((i~=1) && (i~=Ndim)) && ((j~=1) && (j~=Ndim)))
                v(1)=Y(i+1,j);
                v(2)=Y(i,j+1);
                v(3)=Y(i-1,j);
                v(4)=Y(i,j-1);
            elseif ((j~=1) && (j~=Ndim))
                if i==1
                    v(1)=Y(i+1,j);
                    v(2)=Y(i,j+1);
                    v(3)=Y(Ndim,j);
                    v(4)=Y(i,j-1);
                else
                    v(1)=Y(1,j);
                    v(2)=Y(i,j+1);
                    v(3)=Y(i-1,j);
                    v(4)=Y(i,j-1);
                end
            elseif ((i~=1) && (i~=Ndim))
                if j==1
                    v(1)=Y(i+1,j);
                    v(2)=Y(i,j+1);
                    v(3)=Y(i-1,j);
                    v(4)=Y(i,Ndim);
```

```
                else
39                  v(1)=Y(i+1,j);
                    v(2)=Y(i,1);
41                  v(3)=Y(i-1,j);
                    v(4)=Y(i,j-1);
43              end
            else
45              if i==Ndim
                    v(1)=Y(1,j);
47              else
                    v(1)=Y(i+1,j);end
49              if j==Ndim
                    v(2)=Y(i,1);
51              else
                    v(2)=Y(i,j+1);end
53              if i==1
                    v(3)=Y(Ndim,j);
55              else
                    v(3)=Y(i-1,j);end
57              if j==1
                    v(4)=Y(i,Ndim);
59              else
                    v(4)=Y(i,j-1);end
61          end
        s=sum(v);
63      ratdistr=exp((-2).*beta.*Y(i,j).*s);
        a=min(1,ratdistr);
65      u=rand;
        if (u <=a)
67         Y(i,j)=-Y(i,j);
        end
69      m(k)= sum(sum(Y))./(Ndim.^2);
    end
71  M(length(M)+1)=sum(m((900000+1):end).^2)./(Niter-(900000));
end
```

# Bibliography


[1] Lenz, Wilhelm. "Beitrag zum Verständnis der magnetischen Erscheinungen in festen Körpern." Z. Phys. 21 (1920): 613-615.

[2] Ising, Ernst. "Contribution to the theory of ferromagnetism." Z. Phys 31.1 (1925): 253-258.

[3] Peierls, Rudolf. "On Ising's model of ferromagnetism." Mathematical Proceedings of the Cambridge Philosophical Society. Vol. 32. No. 3. Cambridge University Press, 1936.

[4] Onsager, Lars. "Crystal statistics. I. A two-dimensional model with an order-disorder transition." Physical Review 65.3-4 (1944): 117.

[5] Brémaud, Pierre. Markov chains: Gibbs fields, Monte Carlo simulation, and queues. Vol. 31. Springer Science  Business Media, 2013.

[6] Hastings, W. Keith. "Monte Carlo sampling methods using Markov chains and their applications." Biometrika (1970): 97-109.

[7] Metropolis, Nicholas, et al. "Equation of state calculations by fast computing machines." The journal of chemical physics 21.6 (1953): 1087-1092.

[8] Caravenna, Francesco, and Paolo Dai Pra. Probabilità: Un'introduzione attraverso modelli e applicazioni. Vol. 67. Springer Science  Business Media, 2013.