

---

---

LECTURE NOTES  
OF  
STATISTICAL MECHANICS

---

---

COLLECTION OF THE LECTURES NOTES OF PROFESSOR ORLANDINI.

EDITED BY  
ALICE PAGANO  
*The University of Padua*

COMPILED: WEDNESDAY 11<sup>TH</sup> DECEMBER, 2019



# Contents

<b>Introduction</b>	<b>vii</b>
<b>1 Recall of Thermodynamics</b>	<b>1</b>
1.1 A short recap of thermodynamics definitions . . . . .	1
1.2 Equilibrium states . . . . .	1
1.3 Equations of states . . . . .	3
1.4 Legendre transform and thermodynamic potentials . . . . .	4
1.5 Maxwell relations . . . . .	8
1.6 Response functions . . . . .	9
1.6.1 Response functions and thermodynamic stability . . . . .	11
<b>2 Equilibrium phases and thermodynamics of phase transitions</b>	<b>13</b>
2.1 Equilibrium phases as minima of Gibbs free energy . . . . .	13
2.2 First order phase transition and phase coexistence . . . . .	13
2.2.1 Critical points . . . . .	17
2.2.2 Ferromagnetic system . . . . .	17
2.3 Second order phase transition . . . . .	19
2.3.1 Helmholtz free-energy . . . . .	20
2.4 Thermodynamic of phase coexistence . . . . .	20
2.4.1 Lever Rule . . . . .	20
2.4.2 Phase coexistence (one component system) . . . . .	21
2.4.3 Clausius-Clapeyron equation . . . . .	22
2.4.4 Application of C-C equation to the liquid-gas coexistence line . . . . .	23
2.5 Order parameter of a phase transition . . . . .	25
2.6 Classification of the phase transitions . . . . .	26
2.6.1 Thermodynamic classification . . . . .	26
2.6.2 Eherenfest classification . . . . .	26
2.6.3 Modern classification . . . . .	26
2.7 Critical exponents . . . . .	27
2.7.1 Divergence of the response functions at the critical point . . . . .	27
2.7.2 Critical exponents definition . . . . .	27
2.7.3 Law of the corresponding states . . . . .	29
2.7.4 Thermodynamic inequalities between critical exponents . . . . .	30
<b>3 Recall of statistical mechanics and theory of ensembles</b>	<b>33</b>
3.1 Statistical ensembles . . . . .	33
3.2 The canonical ensemble . . . . .	34
3.2.1 Energy fluctuations in the canonical ensemble . . . . .	37
3.3 Isothermal and isobaric ensemble . . . . .	38
3.3.1 Saddle point approximation . . . . .	39
3.4 Gran canonical ensemble . . . . .	40
<b>4 Statistical mechanics and phase transitions</b>	<b>41</b>

4.1	Statistical mechanics of phase transitions . . . . .	41
4.1.1	Magnetic system (canonical) . . . . .	42
4.1.2	Fluid system (gran canonical) . . . . .	42
4.1.3	Thermodynamic limit with additional constraints . . . . .	43
4.1.4	Statistical mechanics and phase transitions . . . . .	43
4.2	Critical point and correlations of fluctuations . . . . .	44
4.3	Finite size effects and phase transitions . . . . .	46
4.4	Numerical simulations and phase transitions . . . . .	47
<b>5</b>	<b>Role of the models in statistical mechanics</b>	<b>49</b>
5.1	Role of the models . . . . .	49
5.2	The Ising model . . . . .	49
5.2.1	$d$ -dimensional Ising model . . . . .	50
5.2.2	Mathematical properties of the Ising model with nearest neighbours interactions . . . . .	51
5.2.3	Ising model and $\mathbb{Z}^2$ symmetry. . . . .	54
5.3	Lattice gas model . . . . .	55
5.4	Fluid system in a region $\Omega$ . . . . .	56
5.4.1	From the continuous to the lattice gas model . . . . .	57
<b>6</b>	<b>Some exactly solvable models of phase transitions</b>	<b>59</b>
6.1	$1$ -dimensional Ising model . . . . .	59
6.1.1	Recursive method . . . . .	59
6.1.2	Transfer Matrix method . . . . .	62
6.2	General transfer matrix method . . . . .	64
6.2.1	The free energy . . . . .	65
6.2.2	The correlation function . . . . .	65
6.2.3	Results for the $1$ -dimensional Ising model . . . . .	68
6.3	Classical Heisenberg model for $d=1$ . . . . .	71
6.4	Zipper model . . . . .	72
6.4.1	Transfer matrix method for the Kittel model . . . . .	75
6.5	Transfer matrix for $2D$ Ising . . . . .	77
<b>7</b>	<b>The role of dimension, symmetry and range of interactions in phase transitions</b>	<b>79</b>
7.1	Energy-entropy argument . . . . .	79
7.1.1	Ising $D = 1$ . . . . .	79
7.1.2	$D$ dimension . . . . .	81
7.2	Role of the symmetry . . . . .	83
7.3	XY model . . . . .	84
7.4	Continuous symmetries and phase transitions . . . . .	85
7.5	Role of the interaction range . . . . .	86
7.6	Ising model with infinite range . . . . .	86
<b>8</b>	<b>Mean field theories of phase transitions and variational mean field</b>	<b>91</b>
8.1	Mean field theories . . . . .	91
8.1.1	Mean field for Ising model . . . . .	91
8.1.2	Free-energy expansion for $m \simeq 0$ . . . . .	94
8.1.3	Mean field critical exponents . . . . .	95
8.2	Mean field variational method . . . . .	97
8.2.1	Mean field approximation for the variational approach . . . . .	99
8.2.2	First approach: Bragg-Williams approximation . . . . .	100
8.2.3	Second approach: Blume-Emery-Griffith model . . . . .	104

8.2.4 Mean field again . . . . .	108
<b>Conclusions</b>	<b>111</b>
<b>Bibliography</b>	<b>113</b>



# Introduction

The goal of statistical mechanics [1] is to predict the macroscopic properties of bodies, most especially their thermodynamic properties, on the basis of their microscopic structure.

The macroscopic properties of greatest interest to statistical mechanics are those relating to thermodynamic equilibrium. As a consequence, the concept of thermodynamic equilibrium occupies a central position in the field.

The microscopic structure of systems examined by statistical mechanics can be described by means of mechanical models: for example, gases can be represented as systems of particles that interact by means of a phenomenologically determined potential. Other examples of mechanical models are those that represent polymers as a chain of interconnected particles, or the classical model of crystalline systems, in which particles are arranged in space according to a regular pattern, and oscillate around the minimum of the potential energy due to their mutual interaction. The models to be examined can be, and recently increasingly are, more abstract, however, and exhibit only a faint resemblance to the basic mechanical description (more specifically, to the quantum nature of matter). The explanation of the success of such abstract models is itself the topic of one of the more interesting chapters of statistical mechanics: the theory of universality and its foundation in the renormalization group.

The models of systems dealt with by statistical mechanics have some common characteristics. We are in any case dealing with systems with a large number of degrees of freedom: the reason lies in the corpuscular (atomic) nature of matter. The degrees of freedom that one considers should have more or less comparable effects on the global behavior of the system. This state of affairs excludes the application of the methods of statistical mechanics to cases in which a restricted number of degrees of freedom “dominates” the others—for example, in celestial mechanics, although the number of degrees of freedom of the planetary system is immense, an approximation in which each planet is considered as a particle is a good start. In this case, we can state that the translational degrees of freedom (three per planet)—possibly with the addition of the rotational degrees of freedom, also a finite number—dominate all others. These considerations also make attempts to apply statistical concepts to the human sciences problematic because, for instance, it is clear that, even if the behavior of a nation’s political system includes a very high number of degrees of freedom, it is possible to identify some degrees of freedom that are disproportionately important compared to the rest. On the other hand, statistical methods can also be applied to systems that are not strictly speaking mechanical—for example, neural networks (understood as models of the brain’s components), urban thoroughfares (traffic models), or problems of a geometric nature (percolation).

The simplest statistical mechanical model is that of a large number of identical particles, free of mutual interaction, inside a container with impenetrable and perfectly elastic walls. This is the model of the ideal gas, which describes the behavior of real gases quite well at low densities, and more specifically allows one to derive the well-known equation of state.

The introduction of pair interactions between the particles of the ideal gas allows us to obtain the standard model for simple fluids. Generally speaking, this model cannot be resolved exactly and is studied by means of perturbation or numerical techniques. It allows one to describe the behavior of real gases (especially noble gases), and the liquid–vapor transition (boiling and condensation).

The preceding models are of a classical (nonquantum) nature and can be applied only when the temperatures are not too low. The quantum effects that follow from the inability to distinguish particles are very important for phenomenology, and they can be dealt with at the introductory level if one omits interactions between particles.

In many of the statistical models we will describe, however, the system’s fundamental elements will not be “particles,” and the fundamental degrees of freedom will not be mechanical (position and velocity or impulse). If we want to understand the origin of ferromagnetism, for example, we should isolate only those degrees of freedom that are relevant to the phenomenon being examined (the orientation of the electrons’ magnetic moment) from all those that are otherwise pertinent to the material in question.

The simplest case is that in which there are only two values—in this fashion, we obtain a simple model of ferromagnetism, known as the Ising model, which is by far the most studied model in statistical mechanics. The ferromagnetic solid is therefore represented as a regular lattice in space, each point of which is associated with a degree of freedom, called spin, which can assume the values  $+1$  and  $-1$ . This model allows one to describe the paramagnet-ferromagnet transition, as well as other similar transitions.

In this course, classical statistical mechanics of system at equilibrium is treated. The exam is divided into two parts: first, common oral exam (same exercise and question for everyone, it is a written part), second part, oral.



Outline of the course:

1. Brief recap of thermodynamics.
2. Equilibrium phases and thermodynamics of the phase transitions.
3. Statistical mechanics and theory of ensembles.
4. Thermodynamic limit and phase transitions in statistical mechanics.
5. Order parameter and critical point.
6. The role of modelling in the physics of phase transitions.
7. The Ising model.
8. Exact solutions of the Ising model.
9. Transfer matrix method.
10. Role of dimension and range of interactions in critical phenomena (lower critical dimension).
11. Approximations: *Meanfield theory Weiss* and *variational mean field*.
12. Landau theory of phase transitions: the role of symmetries.
13. Relevance of fluctuations: the *Ginzburg criterium* and the notion of the *upper critical dimension*.
14. The Ginzburg-Landau model.
15. Landau theory for non-homogeneous system. The  $\nu$  exponent.
16. Gaussian fluctuations in the G-L theory.
17. Widom's scaling theory.
18. Kadanoff's theory of scaling.
19. The theory of renormalisations group and the origin of *universality* in critical phenomena.
20. Spontaneous symmetry breaking.



# Chapter 1

## Recall of Thermodynamics

### 1.1 A short recap of thermodynamics definitions

The systems we are considering are

1. In equilibrium with an external bath at fixed temperature  $T$ .
2. Made by a (large) number  $N$  of degrees of freedom. For instance, we remind that  $1 \text{ mol} \approx N_A \sim 10^{23}$  elementary units.

Thermodynamic is a macroscopic theory of matter at equilibrium. It starts either from experimental observations or from *axiomatic assumptions* and establishes rigorous relations between macroscopic variables (**observables**) to describe systems at equilibrium. One of the first important concept is the one of **extensive variables**. For instance, the extensive variables that characterize the system at equilibrium are the internal energy  $U$ , volume  $V$ , number of particles  $N$  and magnetization  $\vec{M}$  that "scale with the system". In general, the extensive variable are additive.

In thermodynamic, it is important the concept of walls and thermodynamic constraints that are necessary for a complete definition of a thermodynamic system. With their presence or absence it is possible to control and redistribute the thermodynamic variables for changing the system. The typical walls are:

- **Adiabatic walls:** no heat flux. If it is removed we obtain a *diathermic walls*.
- **Rigid walls:** no mechanical work. If it is removed we obtain a *flexible or mobile walls*.
- **Impermeable walls:** no flux of particles (the number of particles remain constraints). If it is removed we obtain a *permeable walls*.

### 1.2 Equilibrium states

Consider a system in an equilibrium state, if the system changes our aim is to study the next equilibrium state of the system. Therefore we move from a system in equilibrium to another. The fundamental problem of thermodynamics is how to characterize the new system.

Now, we define the concept of **equilibrium states**. Consider macroscopic states that are fully described by *extensive* variables such as the internal energy  $U$ , the volume  $V$ , the number of particles  $N$ , the magnetization  $\vec{M}$ , etc .... If these variables are time independent, the system is in a *steady state*. If moreover there are no macroscopic currents, the system is at *equilibrium*. Therefore, we describe a system by characterizing all the extensive variables at equilibrium.

**Lecture 1.**  
Wednesday 9<sup>th</sup>  
October, 2019.  
Compiled:  
Wednesday 11<sup>th</sup>  
December, 2019.

Suppose that the system changes slow in time, it goes from an equilibrium state to another one and the transformation is so slow that in each  $\Delta t$  the system is at equilibrium. Therefore, consider a sequence of equilibrium states, the **quasi-static transformation** are described by the 1<sup>st</sup> Law of Thermodynamic:

$$dU = \delta Q - \delta W \quad (1.1)$$

The variation of the internal energy of the systems depends by two factors,  $\delta w$  that is the work done by the system during a *quasi-static process* (infinitively slow), and  $\delta Q$  that is the heat absorbed by the system during the process. Remember that we write  $dU$  because it is a differential quantity, while the other quantities with the  $\delta$  are only small quantities. Therefore,  $dU$  is a function of state, the other are not.

*Remark.* The convention is  $\delta Q > 0$  if the heat is absorbed by the system, and  $\delta W > 0$  if the work is done by the system.

For example, for a simple fluid with a given pressure if we change the volume, the work done by the systems is  $\delta W = P dV$ . For a magnetized system, we have  $\delta W = -\vec{H} \cdot d\vec{M}$ .

In conclusion, starting from an equilibrium state and removing some constraints (i.e. wall properties), we want to find the new equilibrium state compatible with the new constraints.

Suppose a system with adiabatic rigid impermeable constraints. The system on the left is characterized by  $V_1, N_1, U_1$ , the one on the right by  $V_2, N_2, U_2$ . There are many ways for solving this problem. We use the most general way, that is by using the *maxium entropy principle*. If exists a function  $S$  of the extensive variables of the system that is defined for all equilibrium states, we call it **entropy** and the fundamental relation is

$$S = S(U, V, N) \quad (1.2)$$

The new values taken by the extensive parameters when a constraint has been removed are the ones that *maximize*  $S$ . It means  $dS = 0$  and  $d^2S < 0$ , given the remaining constraint.

The properties of  $S$  are:

1.  $S$  is an additive function with respect to the subsystems in which the system is partitioned:

$$S = \sum_{\alpha} S^{(\alpha)} \quad (1.3)$$

2.  $S$  is differentiable and *monotonically increasing* with respect to the internal energy  $U$ . It means that  $(\frac{\partial S}{\partial U})_{V,N} > 0$ .

3. For each subsystem  $(\alpha)$  we have:

$$S^{(\alpha)} = S^{(\alpha)}(U^{(\alpha)}, V^{(\alpha)}, N^{(\alpha)}) \quad (1.4)$$

This fundamental relation holds for each subsystem.

4.  $S$  is an *homogeneous function* of 1<sup>st</sup> order with respect to the extensive parameters, namely:

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N), \quad \forall \lambda > 0 \quad (1.5)$$

It means that  $S$  is an extensive quantity.

*Remark.* Since  $S$  is monotonically increasing in  $U$ , the following inequality holds:

$$\left(\frac{\partial S}{\partial U}\right)_{V,N} > 0$$

Therefore, we have  $\left(\frac{\partial S}{\partial U}\right)_{V,N} \neq 0$  and it can be inverted locally.

Afterwards,  $S = S(U, V, N)$  inverted in  $U$  gives the second fundamental relation

$$U = U(S, V, N) \quad (1.6)$$

It means that, we can look at  $S$  or  $U$  and, when these quantities are known, all the informations about the system can be obtained.

By taking the differential of the fundamental relation

$$U = U(S, V, N_1, \dots, N_r) \quad (1.7)$$

one gets

$$dU = \underbrace{\left(\frac{\partial U}{\partial S}\right)_{V,N_j}}_{\substack{T \\ \text{absolute} \\ \text{temperature}}} dS + \underbrace{\left(\frac{\partial U}{\partial V}\right)_{S,N_j}}_{\substack{-P \\ \text{pressure}}} dV + \sum_{j=1}^r \underbrace{\left(\frac{\partial U}{\partial N_j}\right)_{S,V}}_{\substack{\mu_j \\ \text{electrochemical} \\ \text{potential}}} dN_j \quad (1.8)$$

### 1.3 Equations of states

Now, we define another set of variables that are called **intensive variables**. The term *intensive* means that it is independent of the size of the system, namely that the value of the variable relative to a subsystem is equal to that of the whole system. The intensive variables are themselves functions of  $S, V, N$ , and examples of intensive variables are the pressure,  $P$ , and the temperature of the system,  $T$ .

The **state equations** are defined as:

$$T = T(S, V, N_1, \dots, N_r) \quad (1.9a)$$

$$P = P(S, V, N_1, \dots, N_r) \quad (1.9b)$$

$$\mu_j = \mu_j(S, V, N_1, \dots, N_r) \quad (1.9c)$$

*Remark.* If all the state equations are known, the fundamental relation is determined a part from a constant. It means that the coefficients of the differential (1.8) are known.

**Example 1.** Suppose an *ideal* gas, for which is valid the relation

$$PV = Nk_bT \quad (1.10)$$

The *Van-Der Waals equation* of the state is

$$\left(P + \frac{\alpha N^2}{V^2}\right)(V - Nb) = Nk_bT \quad (1.11)$$

Another equation of state for magnetic systems is the *Curie Law*:

$$M = \frac{CH}{T} \quad (1.12)$$

We compute  $\left(\frac{\partial U}{\partial M}\right)_{S,N} = H$ .

The state equations are homogenous functions of *zero* degree:

$$T(\lambda S, \lambda V, \lambda N) \stackrel{0}{=} T(S, V, N) \quad (1.13)$$

At equilibrium the temperature of a subsystem is equal to the one of the whole system. Similarly

$$P(\lambda S, \lambda V, \lambda N) = P(S, V, N) \quad (1.14)$$

Now, we keep  $S$  parameter separates from the other that are substituted by *generalized displacements*, as  $(V, N_1, \dots, N_r) \rightarrow X_j$ . The fundamental relation is

$$U = U(S, X_1, \dots, X_{r+1}) \quad (1.15)$$

We define:

$$\left( \frac{\partial U}{\partial S} \right) \equiv T \quad (1.16a)$$

$$\left( \frac{\partial U}{\partial X_j} \right) \equiv P_j \quad (1.16b)$$

The differential is written as the following:

$$dU = T dS + \sum_{j=1}^{r+1} P_j dX_j \quad (1.17)$$

where  $P_1 = -P$  is the *pressure* and  $X_1 = V$  is the volume.

From the equilibrium condition,  $dU = 0$ , one can get a relation between intensive variables in differential form as the **Gibbo-Duhem relation**:

$$S dT + \sum_{j=1}^{r+1} X_j dP_j = 0 \quad (1.18)$$

For a one-component simple fluid system, the equation (1.18) simplifies into

$$S dT - V dP + N d\mu = 0 \quad (1.19)$$

and dividing by the number of moles  $N$

$$d\mu = -s dT + v dP \quad (1.20)$$

that is the Gibbs-Duhem relation in *molar form*.

For a magnetic system we have  $dU = T dS + \vec{\mathbf{H}} \cdot d\vec{\mathbf{M}} + \mu dN$ .

*Remark.* Note that  $\mu = \mu(T, P)$  is a relation between intensive variables.

To summarize, the fundamental relations are  $S = S(U, V, N_1, \dots, N_r)$  or  $S = S(U, \vec{\mathbf{M}}, N_1, \dots, N_r)$  for magnetic systems. In the energy representation we have  $U = U(S, V, N_1, \dots, N_r)$  or  $U = U(S, \vec{\mathbf{M}}, N_1, \dots, N_r)$ .

## 1.4 Legendre transform and thermodynamic potentials

In many situations it is convenient to change some extensive variables with their conjugate intensive ones that became independent and free to vary. We have new thermodynamic potentials. It works as following; suppose we have a function as

$$Y = Y(X_0, X_1, \dots, X_k, \dots, X_{r+1}) \quad (1.21)$$

such that  $Y$  is *strictly convex* in say,  $X_k$  ( $\frac{\partial^2 Y}{\partial X_k^2} > 0$ ) and smooth. The idea is to find a transformation such that  $Y = Y(X_0, X_1, \dots, P_k, \dots, X_{r+1})$ , where

$$X_k \rightarrow P_k \equiv \frac{\partial Y}{\partial X_k} \quad (1.22)$$

i.e.  $P_k$  substitutes  $X_k$  as a new independent variable. In mathematic this is called *Legendre transform*.

*Thermodynamic potentials* are extremely useful tools, whose name derives from an analogy with mechanical potential energy: as we will later see, in certain circumstances the work obtainable from a macroscopic system is related to the change of an appropriately defined function, the thermodynamic potential. They are useful because they allow one to define quantities which are experimentally more easy to control and to rewrite the fundamental thermodynamic relations in terms of them.

Mathematically all the thermodynamic potentials are the result of a Legendre transformation of the internal energy, namely they are a rewriting of the internal energy so that a variable has been substituted with another.

**Example 2.** Suppose we want to replace the entropy  $S$  with its conjugate derivative

$$T = \frac{\partial U}{\partial S}$$

One starts from the fundamental relation

$$U = U(S, V, N_1, \dots)$$

and transforms  $U$  such that  $S$  is replaced by  $T$  as new independent variable. Let us consider the transformation

$$A \equiv U - S \frac{\partial U}{\partial S} = U - TS$$

By differentiating  $A$  we get

$$dA = dU - T dS - S dT$$

On the other hand

$$dU = T dS + \sum_j P_j dX_j$$

It implies that

$$dA = -S dT + \sum_j P_j dX_j$$

For such a system we have  $A = A(T, V, N_1, \dots, N_r)$ . It is a function of  $T$  instead of  $S$ , as wanted. Similarly for a magnetic system  $A = A(T, \vec{M}, N_1, \dots, N_r)$ .

### Helmholtz free energy

The **Helmholtz free energy** is defined as:

$$A \equiv U - TS \quad (1.23)$$

In terms of heat and mechanical work, since  $dU = \delta Q - \delta W$ :

$$dA = dU - d(TS) = \delta Q - T dS - S dT - \delta W \quad (1.24)$$

Hence,

$$\delta W = (\delta Q - T dS) - S dT - dA \quad (1.25)$$

On the other hand, for a *reversible transformation* we have  $\delta Q = T dS$ , which implies

$$\delta W = -S dT - dA \quad (1.26)$$

If the reversible transformation is also *isothermal*,  $dT = 0$  and we obtain  $dA = dW$ . It is reminiscent of a potential energy.

*Remark.* For an isothermal but *not reversible* (spontaneous) process we know the 2<sup>nd</sup> Law of Thermodynamics

$$\delta Q \leq T dS \quad (1.27)$$

which implies

$$(\delta W)_{irr} = \delta Q - T dS - dA \leq -dA.$$

Hence, if  $\delta W = 0$  and  $dT = 0$ , we have  $dA \leq 0$ . Therefore, in a spontaneous (irreversible) process, the thermodynamic system, as a function of  $T, V, N$  etc, evolves towards a *minimum* of the Helmholtz free energy  $A = A(T, V, N_1, \dots, N_r)$ .

In the case of a system with  $(P, V, T)$  we have:

$$dA = -S dT - P dV + \sum_j \mu_j dN_j \quad (1.28)$$

where

$$-S = \left( \frac{\partial A}{\partial T} \right)_{V, N_j} \quad (1.29a)$$

$$-P = \left( \frac{\partial A}{\partial V} \right)_{T, N_j} \quad (1.29b)$$

$$\mu_j = \left( \frac{\partial A}{\partial N_j} \right)_{T, V} \quad (1.29c)$$

For a magnetic system  $(\vec{H}, \vec{M}, T)$ :

$$dA = -S dT + \vec{H} \cdot d\vec{M} + \sum_j \mu_j dN_j \quad (1.30)$$

with

$$H_\alpha = \left( \frac{\partial A}{\partial M_\alpha} \right)_{T, \{N_j\}} \quad (1.31)$$

## Hentalpy

The **Hentalpy** is the partial Legendre transform of  $U$  that replaces the volume  $V$  with the pressure  $P$  as independent variable.

Consider  $U = U(S, V, N_1, \dots, N_r)$  and  $-P = \frac{\partial U}{\partial V}$ , we define the hentalpy as

$$H = U + PV \quad (1.32)$$

*Remark.* Note that the plus sign in the definition of the hentalpy is just because the minus of the  $P$ .

We have:

$$\begin{aligned} dH &= dU + P dV + V dP \\ &= T dS - P dV + \sum_j \mu_j dN_j + P dV + V dP \\ &= T dS + V dP + \sum_j \mu_j dN_j \end{aligned} \quad (1.33)$$

Finally, we obtain the relation  $H = H(S, P, N_1, \dots, N_r)$ .



### Gibbs potential

The **Gibbs potential** is obtained by performing the Legendre transform of  $U$  to replace  $S$  and  $V$  with  $T$  and  $P$ .

Consider again  $U = U(S, V, N_1, \dots, N_r)$  and  $T = \frac{\partial U}{\partial S}$ ,  $-P = \frac{\partial U}{\partial V}$ , then we have:

$$G = U - TS + PV = A + PV \quad (1.34)$$

For a simple fluid system

$$\begin{aligned} dG &= dU - T dS - S dT + P dV + V dP \\ &= T dS - P dV + \sum_j \mu_j dN_j - T dS - S dT + P dV + V dP \\ &= -S dT + V dP + \sum_j \mu_j dN_j \end{aligned} \quad (1.35)$$

Hence,  $G = G(T, P, N_1, \dots, N_r)$ .

For a magnetic system, the Gibbs potential is defined as

$$G = A - \vec{\mathbf{M}} \cdot \vec{\mathbf{H}} \quad (1.36)$$

and

$$\begin{aligned} dG &= dA - d\vec{\mathbf{M}} \cdot \vec{\mathbf{H}} = d(U - TS) - d(-\vec{\mathbf{M}} \cdot \vec{\mathbf{H}}) \\ &= dU - T dS - S dT - d\vec{\mathbf{M}} \cdot \vec{\mathbf{H}} - \vec{\mathbf{M}} \cdot d\vec{\mathbf{H}} \\ &= T dS + \vec{\mathbf{H}} \cdot d\vec{\mathbf{M}} - T dS - S dT - d\vec{\mathbf{M}} \cdot \vec{\mathbf{H}} - \vec{\mathbf{M}} \cdot d\vec{\mathbf{H}} \\ &= -S dT - \vec{\mathbf{M}} \cdot d\vec{\mathbf{H}} \end{aligned} \quad (1.37)$$

and finally  $G = G(T, \vec{\mathbf{H}})$  and also

$$S = -\left(\frac{\partial G}{\partial T}\right)_{\vec{\mathbf{H}}} \quad (1.38a)$$

$$\vec{\mathbf{M}} = -\left(\frac{\partial G}{\partial \vec{\mathbf{H}}}\right)_T \quad (1.38b)$$

### Grancanonical potential

The two intensive variables to become independent are  $T$  and  $\mu$ . The corresponding Legendre transform is

$$\Omega = U - TS - \sum_{i=1}^r \mu_i N_i = A - \sum_{i=1}^r \mu_i N_i \quad (1.39)$$

Differentiating this relation we obtain:

$$\begin{aligned} d\Omega &= dU - S dT - T dS - \sum_{ij} d\mu_j N_j - \sum_{i=1}^r \mu_i dN_i \\ &= (\delta Q - T dS) - \delta W - S dT - \sum_{j=1}^r d\mu_j N_j - \sum_{j=1}^r \mu_j dN_j \end{aligned} \quad (1.40)$$

Hence,  $\Omega = \Omega(T, P, \{\mu_j\})$ .

**Lecture 2.**

Friday 11<sup>th</sup>

October, 2019.

Compiled:

Wednesday 11<sup>th</sup>

December, 2019.

## 1.5 Maxwell relations

Internal energy  $U$  and entropy  $S$  are homogeneous function of the first order. A consequence of this fact is the relation called **Euler equation**:

$$U = TS - PV + \sum_j \mu_j N_j \quad (1.41)$$

Instead, the **Maxwell relations** are relations between the mixed derivatives of the thermodynamic potentials. They can be obtained from the expressions of  $dU, dH, dA, dG$  and  $d\Omega$  and from the Schwarz theorem on mixed partial derivatives.

Due to Schwarz theorem, if a thermodynamic potential depends on  $t + 1$  variables there will be  $\frac{t(t+1)}{2}$  independent mixed derivatives.

**Example 3** (Internal energy  $U = U(S, V, N)$ ).

$$dU = T dS - P dV + \mu dN \quad (1.42)$$

where

$$T = \left( \frac{\partial U}{\partial S} \right)_{V,N} \quad -P = \left( \frac{\partial U}{\partial V} \right)_{S,N}$$

It implies that

$$\frac{\partial^2 U}{\partial V \partial S} = \left( \frac{\partial T}{\partial V} \right)_{S,N} \underset{\text{from Schwarz inequality}}{=} \frac{\partial^2 U}{\partial S \partial V} = - \left( \frac{\partial P}{\partial S} \right)_{V,N}$$

therefore, we have the *1° Maxwell relation* :

$$\left( \frac{\partial T}{\partial V} \right)_{S,N} = - \left( \frac{\partial P}{\partial S} \right)_{V,N}$$

All the *3 Maxwell relations* obtained by the differential (1.42) with  $t = 2$ , for which we have  $t + 1 = 3$  and  $\frac{t(t+1)}{2} = 3$  ( $[S, V, N]$ ), are

$$(S, V) : \quad \left( \frac{\partial T}{\partial V} \right)_{S,N} = - \left( \frac{\partial P}{\partial S} \right)_{V,N} \quad (1.43a)$$

$$(S, N) : \quad \left( \frac{\partial T}{\partial N} \right)_{V,S} = \left( \frac{\partial \mu}{\partial S} \right)_{V,N} \quad (1.43b)$$

$$(V, N) : \quad - \left( \frac{\partial P}{\partial N} \right)_{S,V} = \left( \frac{\partial \mu}{\partial V} \right)_{S,N} \quad (1.43c)$$

**Example 4** (Helmholz  $A = A(T, V, N)$ ).

$$dA = -S dT - P dV + \mu dN \quad (1.44)$$

In this case the *3 Maxwell relations* ( $[T, V, N]$ ) are

$$(T, V) : \quad \left( \frac{\partial S}{\partial V} \right)_{T,N} = \left( \frac{\partial P}{\partial T} \right)_{V,N} \quad (1.45a)$$

$$(T, N) : \quad - \left( \frac{\partial S}{\partial N} \right)_{T,V} = \left( \frac{\partial \mu}{\partial T} \right)_{V,N} \quad (1.45b)$$

$$(V, N) : \quad - \left( \frac{\partial P}{\partial N} \right)_{V,T} = \left( \frac{\partial \mu}{\partial V} \right)_{T,N} \quad (1.45c)$$

**Example 5** (Gibbs  $G = G(T, P, N)$  ).

$$dG = -S dT - V dP + \mu dN \quad (1.46)$$

In this case the 3 *Maxwell relations* ( $[T, P, N]$ ) are

$$(T, P) : \quad -\left(\frac{\partial S}{\partial P}\right)_{T,N} = \left(\frac{\partial V}{\partial T}\right)_{P,N} \quad (1.47a)$$

$$(T, N) : \quad -\left(\frac{\partial S}{\partial N}\right)_{T,P} = \left(\frac{\partial \mu}{\partial T}\right)_{P,N} \quad (1.47b)$$

$$(P, N) : \quad \left(\frac{\partial V}{\partial N}\right)_{P,T} = \left(\frac{\partial \mu}{\partial P}\right)_{T,N} \quad (1.47c)$$

## 1.6 Response functions

Response functions are quantities that express how a system reacts when some external parameters are changed.

In fact, aim of most experiments is to measure the response of a thermodynamic system write respect to controlled variations of thermodynamic variables. Any observation is just the perturbation of a system and looking for the response. A list of the commonly used response functions is the following:

- *Thermal expansion coefficient at constant pressure.*

$$\alpha_P \equiv \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_{P,N} \quad (1.48)$$

- *Adiabatic compressibility.*

$$k_S = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_{S,N} \stackrel{V=\left(\frac{\partial H}{\partial P}\right)_{S,N}}{=} -\frac{1}{V} \left( \frac{\partial^2 H}{\partial P^2} \right)_{S,N} \quad (1.49)$$

- *Isothermal compressibility.*

$$k_T = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_{T,N} \stackrel{V=\left(\frac{\partial G}{\partial P}\right)_{T,N}}{=} -\frac{1}{V} \left( \frac{\partial^2 G}{\partial P^2} \right)_{T,N} \quad (1.50)$$

*Remark.* Remember that  $k_T$  it is the second derivative of the Gibbs potential write respect to pressure.

- *Molar heat capacity at constant pressure.*

$$c_P = \left( \frac{\delta Q}{dT} \right)_{P,N} = T \left( \frac{\partial S}{\partial T} \right)_{P,N} \stackrel{S=\left(\frac{\partial G}{\partial T}\right)_{P,N}}{=} -T \left( \frac{\partial^2 G}{\partial T^2} \right)_{P,N} \quad (1.51)$$

- *Specific heat at constant volume.* Consider a quasi static transformation.

$$c_V = \left( \frac{\delta Q}{dT} \right)_{V,N} = T \left( \frac{\partial S}{\partial T} \right)_{V,N} = \left( \frac{\partial(-\partial A/\partial T)_{V,N}}{\partial T} \right)_{V,N} = -T \left( \frac{\partial^2 A}{\partial T^2} \right)_{V,N} \quad (1.52)$$

- *Magnetic suscettibility* ( $d=1$ ) for a magnetic system  $(\vec{\mathbf{M}}, \vec{\mathbf{H}}, T)$ .

$$\chi_T = \left( \frac{\partial M}{\partial H} \right)_T \underset{M=-\frac{\partial G}{\partial H}|_T}{=} - \left( \frac{\partial^2 G}{\partial H^2} \right)_T \quad (1.53)$$

More generals,  $\vec{\mathbf{M}}, \vec{\mathbf{H}}$  we have

$$\chi_{\alpha\beta} = \left( \frac{\partial M_\alpha}{\partial H_\beta} \right)_T, \quad M_\alpha = - \left. \frac{\partial G}{\partial H_\alpha} \right|_T \Rightarrow \chi_{\alpha\beta} = \left. \frac{\partial^2 G}{\partial H_\beta \partial H_\alpha} \right|_T \quad (1.54)$$

Note that the response functions, when used with the Maxwell relations, allow to express observables usually inaccessible to experiments with measurable quantities.

Let us illustrate a lemma useful for calculation:

**Lemma 1.6.1.** *Let  $x, y, z$  be quantities that satisfy the relation  $f(x, y, z) = 0$ . If  $w$  is a function of two any variables chosen between  $x, y, z$ , then:*

1.  $\left( \frac{\partial x}{\partial y} \right)_w \left( \frac{\partial y}{\partial z} \right)_w = \left( \frac{\partial x}{\partial z} \right)_w$
2.  $\left( \frac{\partial x}{\partial y} \right)_z = \frac{1}{\left( \frac{\partial y}{\partial x} \right)_z}$
3.  $\left( \frac{\partial x}{\partial y} \right)_z \left( \frac{\partial y}{\partial z} \right)_x \left( \frac{\partial z}{\partial x} \right)_y = -1$  (*concatenation relation or triple product rule*).

**Example 6.** The Maxwell relation

$$\left( \frac{\partial S}{\partial P} \right)_{T,N} = - \left( \frac{\partial V}{\partial T} \right)_{P,N}$$

obtained from

$$dG = -S dT + V dP$$

and the response function  $\alpha_P$  permit to write

$$\underbrace{\left( \frac{\partial S}{\partial P} \right)_{T,N}}_{\substack{\text{inaccessible} \\ \text{to experiments}}} = \underbrace{-V\alpha_P}_{\text{measurable}} \quad (1.55)$$

**Example 7.** Let us start with the Maxwell relation

$$\left( \frac{\partial S}{\partial V} \right)_{T,N} = \left( \frac{\partial P}{\partial T} \right)_{V,N}$$

obtained from

$$dA = -S dT - P dV + \mu dN$$

From some property of multi-variable differential calculus one has the *triple product rule*:

$$\left( \frac{\partial P}{\partial T} \right)_{V,N} \left( \frac{\partial V}{\partial P} \right)_{T,N} \left( \frac{\partial T}{\partial V} \right)_{P,N} = -1 \quad (1.56)$$

Hence

$$\begin{aligned} \left( \frac{\partial P}{\partial T} \right)_{V,N} &= \frac{-1}{\left( \frac{\partial V}{\partial P} \right)_{T,N} \left( \frac{\partial T}{\partial V} \right)_{P,N}} = - \frac{\left( \frac{\partial V}{\partial T} \right)_{P,N}}{\left( \frac{\partial V}{\partial P} \right)_{T,N}} \\ &= \frac{-V\alpha_P}{-Vk_T} = \frac{\alpha_P}{k_T} \end{aligned} \quad (1.57)$$

### 1.6.1 Response functions and thermodynamic stability

Now, we analyze the concept of *thermal stability*. If one injects heat in a system either at constant volume or at constant pressure, its temperature will inevitably increase

$$\begin{cases} c_V \equiv \left( \frac{\delta Q}{dT} \right)_V \geq 0 \\ c_P \equiv \left( \frac{\delta Q}{dT} \right)_P \geq 0 \end{cases} \quad (1.58)$$

*Remark.* The thermal capacities are *non-negative functions*!

It is useful also the concept of *mechanical stability*. If one compress a system by keeping  $T$  constant, we would expect that it shrinks

$$k_T = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_T \geq 0 \quad (1.59)$$

Similar considerations for a magnetic system, gives

$$c_H \geq 0, \quad c_M \geq 0, \quad \chi_M \geq 0 \quad (1.60)$$

*Remark.* In diamagnetic systems  $\chi_M$  can also be negative.

**Exercise 1.** By using Maxwell relations show that

$$c_P - c_V = \frac{TV\alpha_P^2}{k_T} = \frac{1}{Vk_T} T \left( \frac{\partial V}{\partial T} \right)_P^2 \quad (1.61a)$$

$$c_H - c_M = \frac{T}{\chi_T} \left( \frac{\partial M}{\partial T} \right)_H^2 \quad (1.61b)$$

*Solution.* Let us start considering a system with a fixed number of particles (namely  $dN = 0$ ) and such that  $S$  is explicitly expressed in terms of  $T$  and  $V$ . Then:

$$dS = \left( \frac{\partial S}{\partial T} \right)_V dT + \left( \frac{\partial S}{\partial V} \right)_T dV$$

Dividing by  $dT$  both sides keeping the pressure constant, and then multiplying by  $T$ :

$$T \left( \frac{\partial S}{\partial T} \right)_P - T \left( \frac{\partial S}{\partial T} \right)_V = T \left( \frac{\partial S}{\partial V} \right)_T \left( \frac{\partial V}{\partial T} \right)_P$$

it implies

$$c_P - c_V = T \left( \frac{\partial S}{\partial V} \right)_T \left( \frac{\partial V}{\partial T} \right)_P$$

Now, using the Maxwell relation  $\left( \frac{\partial S}{\partial V} \right)_T = \left( \frac{\partial P}{\partial T} \right)_V$  and using the triple product rule we obtain

$$\left( \frac{\partial P}{\partial T} \right)_V = - \left( \frac{\partial P}{\partial V} \right)_T \left( \frac{\partial V}{\partial T} \right)_P$$

we get:

$$c_P - c_V = -T \left( \frac{\partial P}{\partial V} \right)_T \left( \frac{\partial V}{\partial T} \right)_P^2 = -T \alpha_P^2 V^2 \left( \frac{\partial P}{\partial V} \right)_T = \frac{TV}{k_T} \alpha_P^2$$

It can be shown similarly for magnetic systems.

A consequence is that since the right hand terms are non negative it follows that

$$\begin{cases} c_P \geq c_V \geq 0 \\ c_H \geq c_M \geq 0 \end{cases} \quad (1.62)$$

For reasuming, we have seen the thermodynamic of a phase, where the equilibrium state can be described by the maximum of the entropy. If we have a given phase, we can look for the Gibbs function. If we have more phases, we want to change between these phases.



## Chapter 2

# Equilibrium phases and thermodynamics of phase transitions

### 2.1 Equilibrium phases as minima of Gibbs free energy

Experimentally, any element or compound can be found, depending on the thermodynamic conditions in which it is, in different phases. When we say that a system is in a particular phase we mean that its physical properties (like density or magnetization) are uniform.

Equilibrium states are given by *maxima* of the entropy and *minima* of internal energy, or by *minima* of thermodynamics potentials such as  $A$  and  $G$ . Let us consider for example the Gibbs potential per particle of a fluid system

$$\frac{G}{N} \equiv g = g(T, P) \quad (2.1)$$

that depends on two intensive variables  $T$  and  $P$  and is not anymore a function of  $N$  because we have divided for  $N$ . Let us define  $\alpha$  the phase of a one-component system (say  $\alpha = \text{gas or liquid}$ ). Therefore, the thermodynamic properties are described by surfaces of function  $g_\alpha(T, P)$  and for all equilibrium phase we have a surface on the space  $(T, P, g)$ . For each value of  $T$  and  $P$  the thermodynamically stable phase is the one for which  $g_\alpha(T, P)$  is minimum.

### 2.2 First order phase transition and phase coexistence

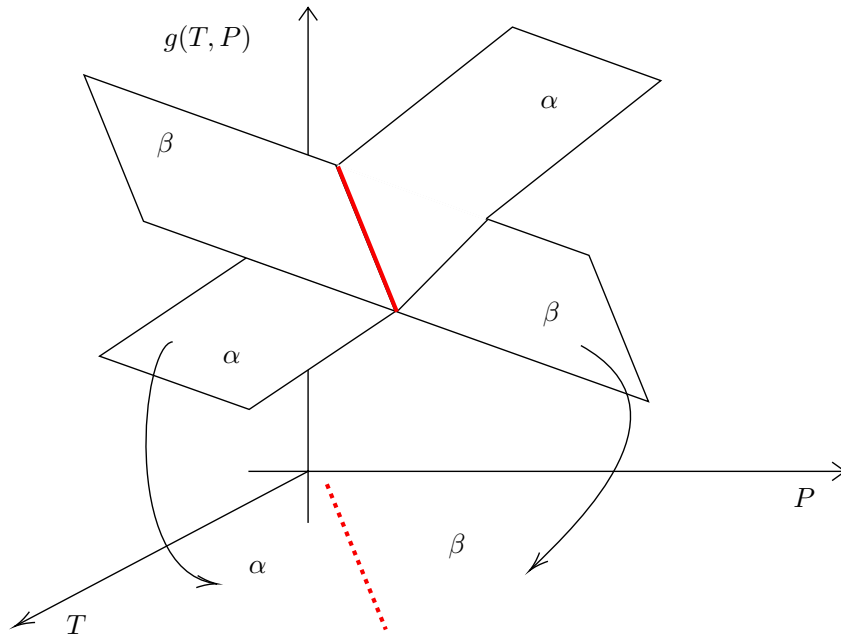
Let us suppose for example that the system can be found in two phases  $\alpha$  and  $\beta$  (for example liquid and solid). Consider the surface  $g_\alpha$  and  $g_\beta$ , we are looking for the lower one.

For given values of  $T$  and  $P$  the stable phase will be that with the lowest value of  $g$ : for example, if we have  $g_\alpha(T, P) < g_\beta(T, P)$  then the system will be in phase  $\alpha$ . Therefore there will be regions in  $(T, P)$  space where the most stable phase will be  $\alpha$  and others in which it will be  $\beta$ . If we now plot the values of  $g$  as a function of  $T$  and  $P$  in  $(g, P, T)$  space for every phase of the system, we can determine the regions where the two phases will be the stable ones, namely we can determine the phase diagram of the system, as illustrated in Figure 2.1.

The very interesting region of this space (and the one on which we will focus our attention in this section) is the line where the surfaces of the two phases intersect: along this the two phases coexist, and when the system crosses it we say that it undergoes a *phase transition*. The coexistence line is the projection on the  $(T, P)$

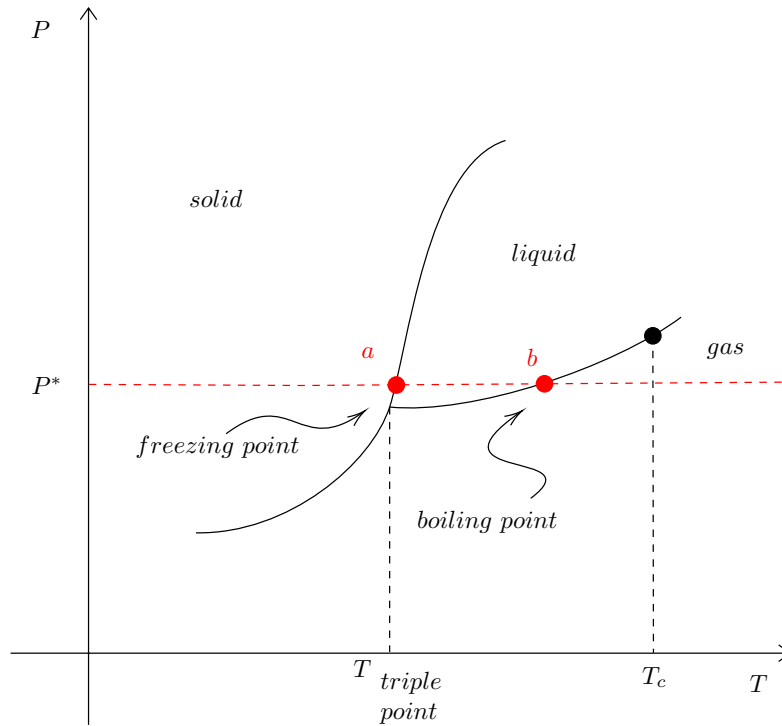
plane of the intersection between different surfaces, so the *coexistence condition* is:

$$g_\alpha(T, P) = g_\beta(T, P) \quad (2.2)$$



**Figure 2.1:** Phase diagram: stability of phases.

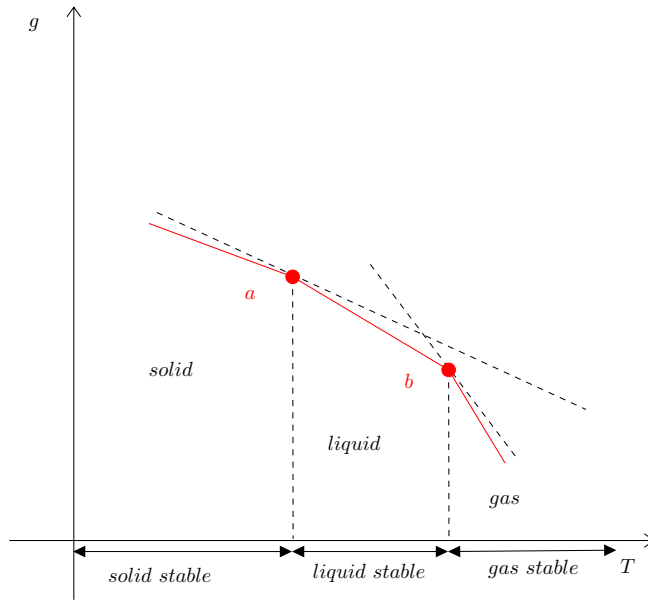
To fix the ideas, let us choose a given value of pressure  $P = P^*$  and study the behaviour of  $g(T, P^*)$  as a function of  $T$  when we go from solid to gas, as illustrated in Figure 2.2.



**Figure 2.2:**  $(T, P)$  projection.

At the triple point  $g_{\text{solid}}(T_a, P^*) = g_{\text{liq}}(T_a)$  and  $g_{\text{liq}}(T_b) = g_{\text{gas}}(T_b, P^*)$ , as shown in Figure 2.3. Note also that:



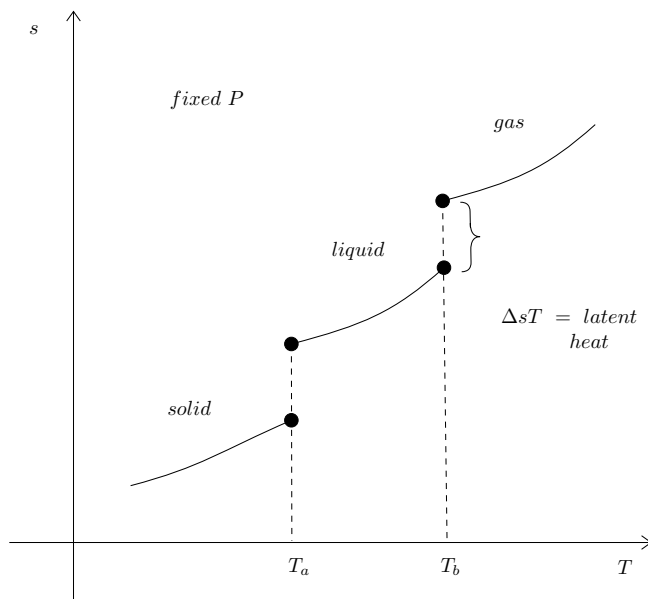


**Figure 2.3:**  $(g, T)$  projection at a fixed pressure  $P = P^*$ .

- At the coexistence points  $a$  and  $b$  of the two phases, one has  $g_\alpha(T) = g_\beta(T)$ .
- $g(T)$  is a continuous function of  $T$ .
- Note that,  $S = -\left(\frac{\partial G}{\partial T}\right)_V$  and  $c_P = -T\left(\frac{\partial^2 G}{\partial T^2}\right) > 0$ . This implies that  $g(T)$  is concave in  $T$  at fixed  $P$ .

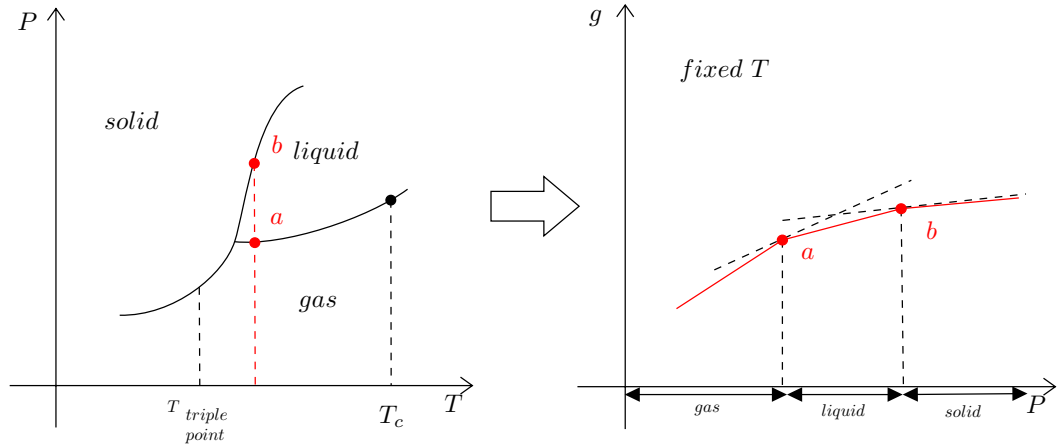
How about its derivatives? Since  $P$  is fixed we can vary  $T$  and look for  $s = -\left(\frac{\partial g}{\partial T}\right)_P$ . As we cross different phases we have discontinuities, where  $\Delta s T$  is called the *latent heat*. It is illustrated in Figure 2.4.

If there is a finite discontinuity in one or more of the first derivatives of the appropriate thermodynamic potential, the transition is called *first order transition*. In general, a phase transition is signalled by a singularity in a thermodynamic potential.



**Figure 2.4:**  $(s, T)$  projection.

We can also fix the temperature  $T = T^*$  and look at the variation of  $P$ , as shown in Figure 2.5.

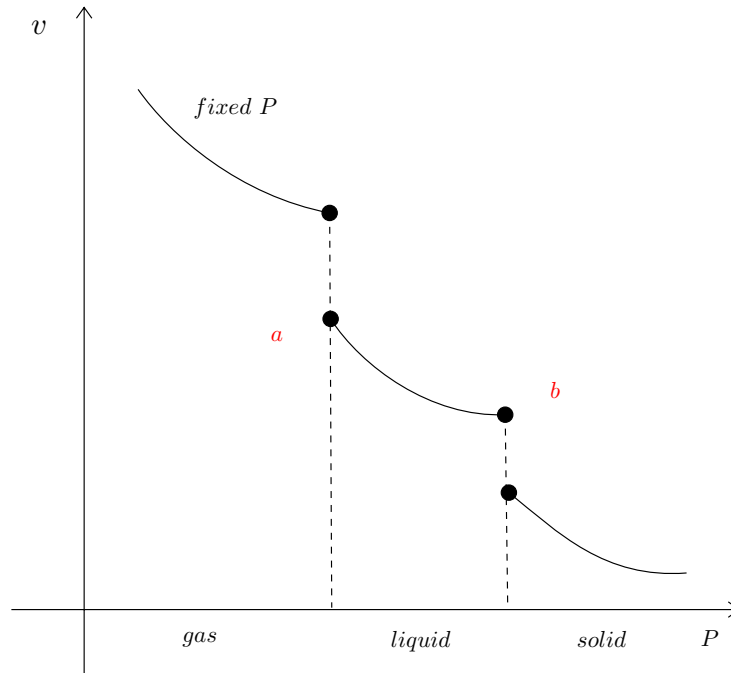


**Figure 2.5:** Left:  $(T, P)$  projection. Right:  $(g, P)$  projection at a fixed temperature  $T = T^*$ .

Note that, we have  $v = \left(\frac{\partial g}{\partial P}\right)_T > 0$  :

$$\left(\frac{\partial^2 g}{\partial P^2}\right) = \left(\frac{\partial U}{\partial P}\right)_T = -vk_T < 0 \quad (2.3)$$

so, also in this case we had a jump of the first order derivative of the thermodynamic potential  $g$ . It is illustrated in Figure 2.6.



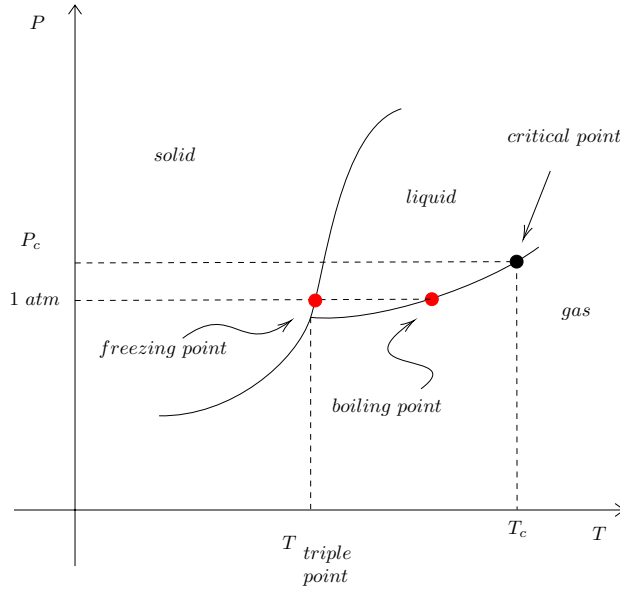
**Figure 2.6:**  $(v, P)$  projection.

### 2.2.1 Critical points

At the critical point  $(P_c, T_c)$  the system can pass from the liquid to the gas phase (and viceversa) in a continuous way

$$\Delta s = \Delta v = 0$$

Usually critical points are end point of first order transition phases. Why there is no critical point between solid and liquid? The crossover between phases having the same symmetry define the Landau point. There is a break of symmetry, for instance we can think about the structure of the bravais lattice. Instead, from gas to liquid symmetries are not broken.

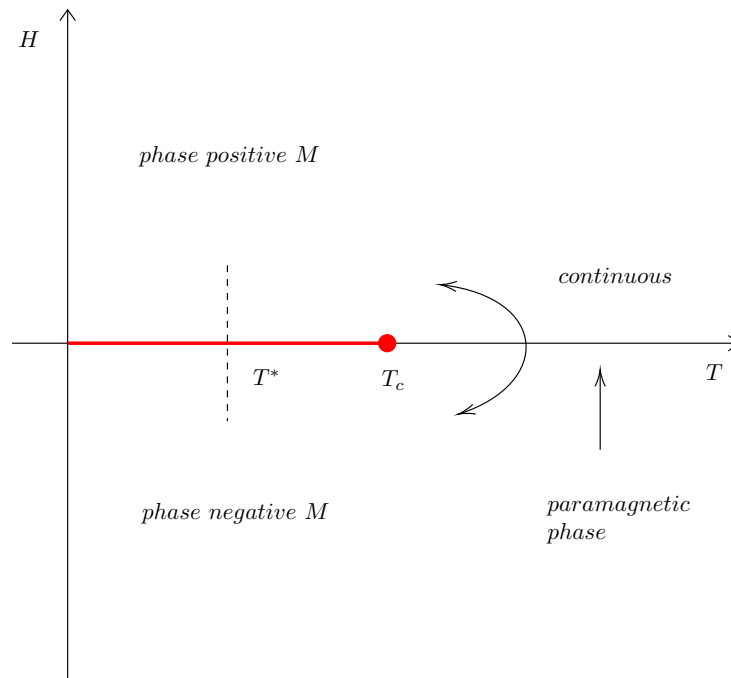


**Figure 2.7:** Phase diagram of a fluid. All the phase transition are first-order except at the critical point  $C$ . Beyond  $C$  it is possible to move continuously from liquid to a gas. The boundary between the solid and liquid phases is thought to be always first-order and not to terminate in a critical point.

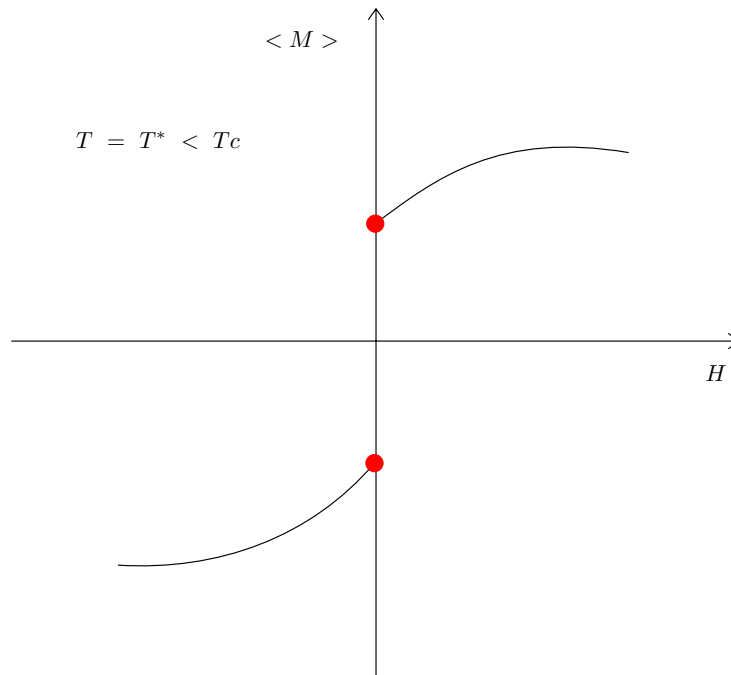
### 2.2.2 Ferromagnetic system

A similar behaviour can be encountered in magnetic systems. We can have a magnetization different from 0 even when there is no magnetic field. Supposing  $P \leftrightarrow H$ ,  $V \leftrightarrow M$ , we have  $(P, T) \leftrightarrow (H, T)$ .

The magnetization  $M$  has a jump at  $H = 0$  for temperatures lower than the critical one; in this case since  $M = -\frac{\partial F}{\partial H}$  we see that the first derivative of the free energy  $F$  with respect to  $H$  has a jump discontinuity. For instance, consider Figure 2.9. At the critical point the magnetization would pass through zero.



**Figure 2.8:** Phase diagram for a magnetic system in  $(T, H)$  space. A line of first-order transitions at zero field ends in a critical point at a temperature  $T_c$ .



**Figure 2.9:** Plot of the Magnetization for  $T = T^* < T_c$ .

## 2.3 Second order phase transition

The transitions are classified in the first order transition and continuous transition. If the first derivatives are continuous but second derivatives are discontinuous or infinite the transition will be described as higher order, continuous or critical. This is different from the previous situation in which we had a jump for the first order derivative of a thermodynamic potential. Some examples are illustrated in Figure 2.10.

Let us suppose that

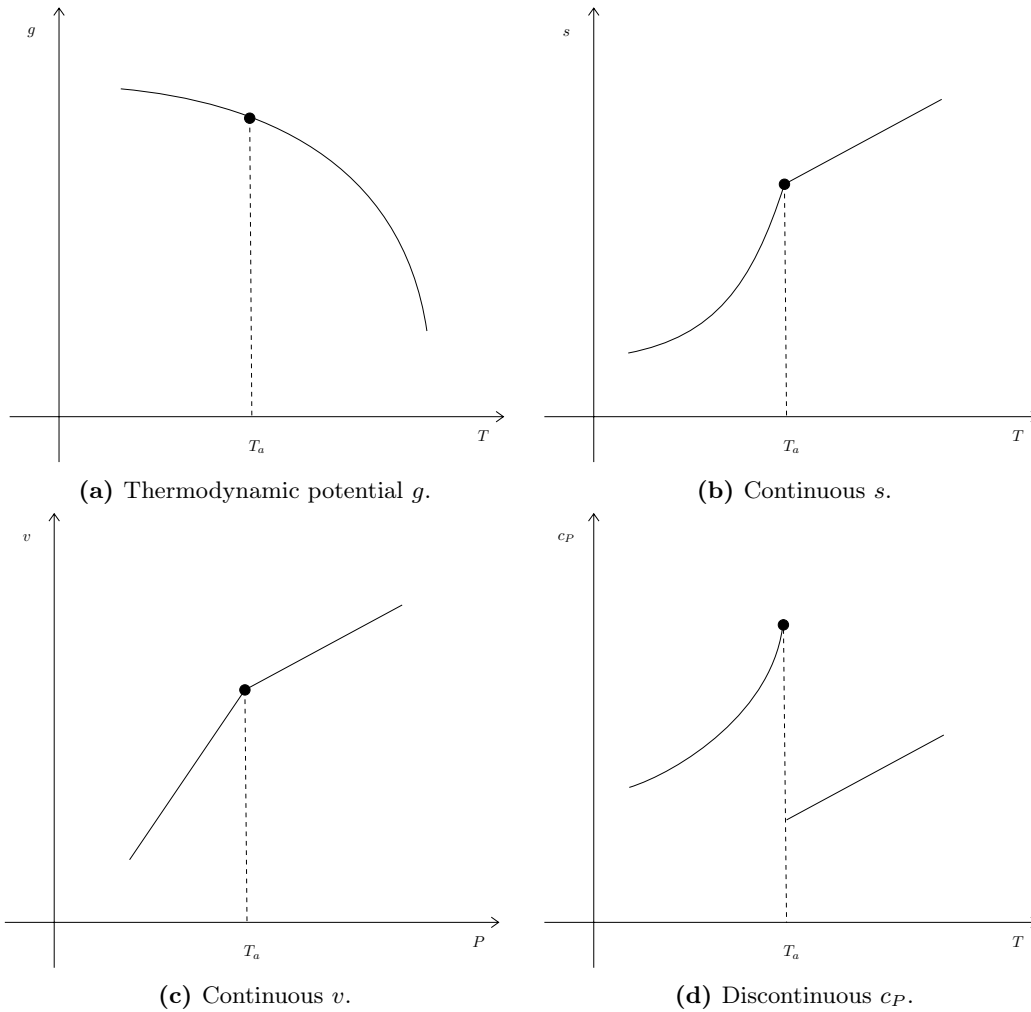
$$\left(\frac{\partial g}{\partial T}\right)_P = -s \quad (2.4a)$$

$$\left(\frac{\partial g}{\partial P}\right)_T = v \quad (2.4b)$$

are continuous. This implies:

$$\left(\frac{\partial^2 g}{\partial T \partial P}\right) = \left(\frac{\partial v}{\partial T}\right)_P = v_{\alpha P} \quad (2.5)$$

that is discontinuous. An example is *superconductivity*.



**Figure 2.10:** Example of a second order phase transition.

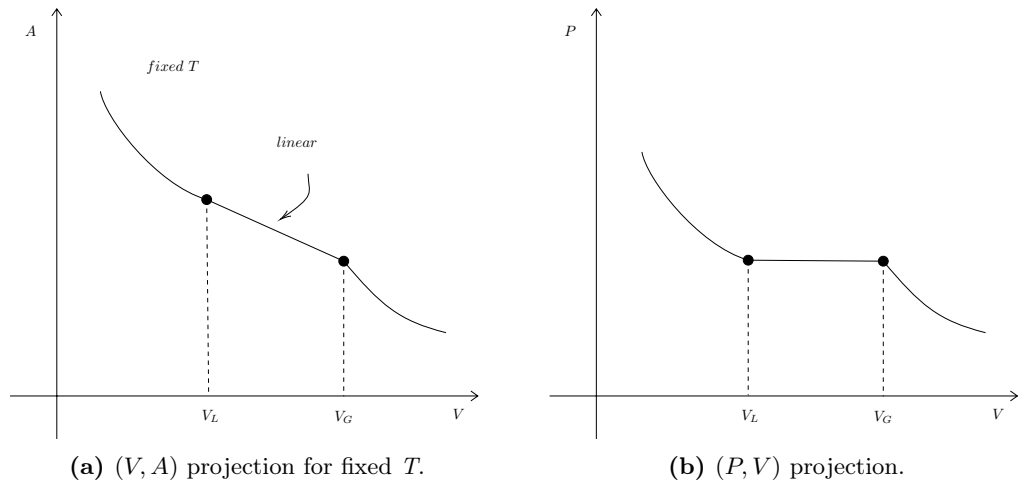
If we look for example at the specific heat  $c_P$  in Figure 2.10d, it represents the transition from superconduction.

The critical point is special because there is not a jump, so we can go continuously from gas to liquid. The response function when we plot this point shows that the specific heat diverges.

The superfluid transition is a transition where the second derivative of the thermodynamic potential diverges. There are many phase transitions that can be classified in different ways.

*Remark.* Note that at the coexistence line we increase  $V$ , but the pressure remains constant. At the coexistence line we see bubbles. It is the density that is changing locally, the bubbles become bigger and bigger and at the  $V_G$ , becomes a liquid.

### 2.3.1 Helmholtz free-energy



**Figure 2.11:** Helmholtz free-energy and phase transition.

Consider  $A = A(T, V, N)$ , here the derivative with respect to  $P$  is replaced by  $V$  which is discontinuous at the first order transition. Moreover  $P > 0$  implies  $\partial A / \partial V < 0$  and

$$k_T = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_T = -\frac{1}{V} \left( \frac{\partial P}{\partial V} \right)_T = \frac{1}{V} \left( \frac{\partial^2 A}{\partial V^2} \right)_T > 0 \quad (2.6)$$

so,  $A$  is an overall convex function of  $V$ . The behaviour of  $A$  when there is a first order phase transition is as in Figure 2.11a. The linear sector becomes an horizontal one in the  $P = -(\partial A / \partial V)_T = P(V)$  curve (Figure 2.11b).

## 2.4 Thermodynamic of phase coexistence

### 2.4.1 Lever Rule

The lever rule [2] is a rule used to determine the mole fraction of each phase of a binary equilibrium phase diagram. For instance, it can be used to determine the fraction of liquid and solid phases for a given binary composition and temperature that is between the liquidus and solidus line.

In an alloy or a mixture with two phases,  $\alpha$  and  $\beta$ , which themselves contain two elements,  $A$  and  $B$ , the lever rule states that the mass fraction of the  $\alpha$  phase is

$$w^\alpha = \frac{w_B - w_B^\beta}{w_B^\alpha - w_B^\beta} \quad (2.7)$$

where

**Lecture 3.**  
Wednesday 16<sup>th</sup>  
October, 2019.  
Compiled:  
Wednesday 11<sup>th</sup>  
December, 2019.

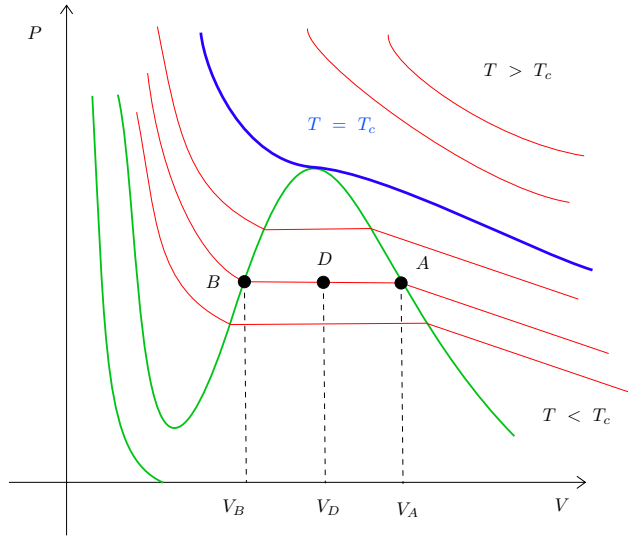
- $w_B^\alpha$ : is the mass fraction of element  $B$  in the  $\alpha$  phase.
- $w_B^\beta$ : is the mass fraction of element  $B$  in the  $\beta$  phase.
- $w_B$ : is the mass fraction of element  $B$  in the entire alloy or mixture.

**Example 8.** Consider the Figure 2.12; at all points between  $A$  and  $B$  the system is a mixture of gas and liquid. Points  $D$  has global density  $\rho_D = \rho_A + \rho_B$  and therefore  $v_D = \frac{1}{\rho_D}, v_A = \frac{1}{\rho_A}, v_B = \frac{1}{\rho_B}$  which implies:

$$v_D = \frac{N_A}{N}v_A + \frac{N_B}{N}v_B = x_A v_A + x_B v_B$$

Since  $x_A + x_B = 1$  we have  $(x_A + x_B)v_D = x_A v_A + x_B v_B$  and finally by rearranging, one finds the *Lever Rule*. It shows that the relative concentration of the liquid-gas mixture changes with  $V$ :

$$\frac{x_A}{x_B} = \frac{v_B - v_D}{v_D - v_A} \quad (2.8)$$



**Figure 2.12:**  $(V, P)$  projection. In the region between  $A$  and  $B$  the gas and the liquid phase coexist by keeping the pressure constant.

### 2.4.2 Phase coexistence (one component system)

Consider a  $(P, V, T)$  system as a mixture of two species  $(1, 2)$  at temperature  $T_1, T_2$ , pressure  $P_1, P_2$  and chemical potentials  $\mu_1, \mu_2$ . The equilibrium condition is given by the maximum of the total entropy  $S = S_1 + S_2$  and gives the conditions

$$T_1 = T_2, \quad P_1 = P_2, \quad \mu_1 = \mu_2 \quad (2.9)$$

this is the *coexistence condition* of the two phases.

In terms of the Gibbs potential  $G = U - TS + PV$ , where  $U$  is given by the Euler equation  $U = TS - PV + \mu_1 N_1 + \mu_2 N_2$ , the Gibbs per mole is

$$g_1(T, P) \equiv \frac{G_1}{N_1} = \mu_1 \quad (2.10a)$$

$$g_2(T, P) \equiv \frac{G_2}{N_2} = \mu_2 \quad (2.10b)$$

Therefore, on the coexistence line it should hold the relation

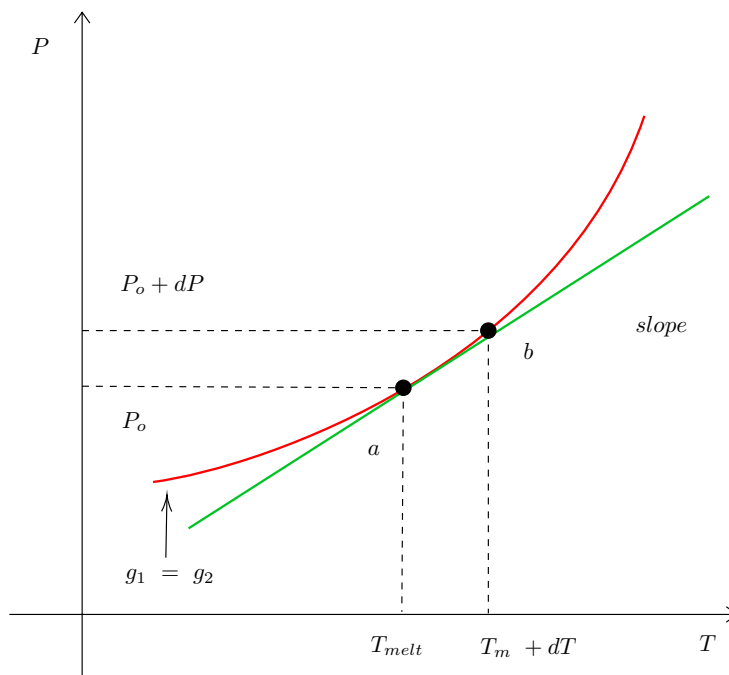
$$g_1(T, P) = g_2(T, P) \quad (2.11)$$

### 2.4.3 Clausius-Clapeyron equation

The coexistence curves [3], as the one illustrated in Figure 2.13, are less arbitrary than is immediately evident; the slope  $dP/dT$  of a coexistence curve is fully determined by the properties of the two coexisting phases.

The slope of a coexistence curve is of direct physical interest. Consider cubes of ice at equilibrium in a glass of water. Given the ambient pressure, the temperature of the mixed system is determined by the liquid-solid coexistence curve of water; if the temperature were not on the coexistence curve some ice would melt, or some liquid would freeze, until the temperature would again lie on the coexistence curve (or one phases would become depleted). If the ambient pressure were to decrease-perhaps by virtue of a change in altitude-then the temperature of the glass of water would appropriately adjust to a new point on the coexistence curve. If  $\Delta P$  were the change in pressure, then the change in temperature would be  $\Delta T = \Delta P / (dP/dT)_{coex}$ , where the derivative in the denominator is the slope of the coexistence curve.

*Remark.* Ice skating presents another interesting example. The pressure applied to the ice directly beneath the blade of the skate shifts the ice across the solid-liquid coexistence curve, providing a lubricating film of liquid on which the skate slides. The possibility of ice skating depends on the negative slope of the liquid-solid coexistence curve of water.



**Figure 2.13:**  $(T, P)$  projection. The coexistence line is represented in red, while in green the slope between the two points  $a$  and  $b$ .

Now, suppose to know the position on the coexistence line (for example the melt temperature  $T_m$  at the atmospheric pressure  $P_0$ ). Is it possible to find other points on the curve? For example  $T_m$  at lower or higher pressure?

The answer is yes for small deviations of  $T$  and  $P$  from  $a$ . The idea is to compute the slope of the tangent of the coexistence curve, i.e.  $(dP/dT)$ . This is given by the



Clausius-Clapeyron equation. Both at  $a$  and  $b$  the two phases 1 and 2 coexist. This means that at the coexistence line

$$\begin{cases} g_1^{(a)} = g_2^{(a)} \\ g_1^{(b)} = g_2^{(b)} \end{cases} \quad (2.12)$$

Hence, if  $a$  and  $b$  are very close:

$$\begin{cases} dg_1 = g_1^{(b)} - g_1^{(a)} \\ dg_2 = g_2^{(b)} - g_2^{(a)} \end{cases} \quad (2.13)$$

Therefore, the *starting point* for *Clausius-Clapeyron* is

$$\Rightarrow dg_1 = dg_2 \quad (2.14)$$

From the molar version of the Gibbs-Duhem relation we have

$$\begin{cases} dg_1 = -s_1 dT + v_1 dP = d\mu_1 \\ dg_2 = -s_2 dT + v_2 dP = d\mu_2 \end{cases} \quad (2.15)$$

taking the difference, one obtains

$$-(s_2 - s_1) dT + (v_2 - v_1) dP = 0$$

The slope is called **Clausius-Clapeyron equation**:

$$\left( \frac{dP}{dT} \right)_{coex} = \frac{(s_2 - s_1)}{(v_2 - v_1)} = \frac{\Delta s}{\Delta v} \quad (2.16)$$

*Remark.* Since  $(dP/dT)_{coex}$  is finite, the equation explains why a first order transition is characterised by discontinuous changes in entropy and volume (or density).  $\Delta S$  gives the latent heat  $L_{12}$ <sup>1</sup>:

$$L_{12} = T \Delta s \quad (2.17)$$

whence, the Clapeyron equation is

$$\frac{dP}{dT} = \frac{L_{12}}{T \Delta v} \quad (2.18)$$

#### 2.4.4 Application of C-C equation to the liquid-gas coexistence line

Now, we go from gas (region 2) to liquid (region 1), we have:

$$\left( \frac{dP}{dT} \right)_{coex} = \frac{s_2 - s_1}{v_2 - v_1} \quad (2.19)$$

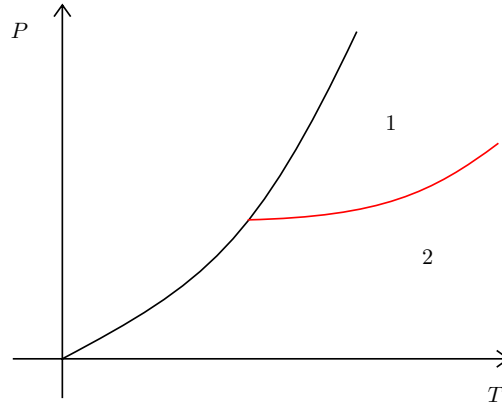
The Clapeyron equation embodies the *Le Chatelier principle*<sup>2</sup> Consider a liquid-solid transition (the coexistence curves are shown in Figure 2.14):

$$\left( \frac{dP}{dT} \right)_{coex} > 0 \Rightarrow \frac{s_2 - s_1}{v_2 - v_1} > 0 \quad (2.20)$$

and since  $v_2 > v_1$ , we have  $s_2 > s_1$ . The gas has more entropy as it should be. The slope of the phase curve is positive, then an increase in pressure at constant

<sup>1</sup>The latent heat of fusion is the quantity of heat required to melt one mole of solid.

<sup>2</sup>"When a settled system is disturbed, it will adjust to diminish the change that has been made to it".



**Figure 2.14:**  $(T, P)$  projection. Region 1: liquid. Region 2: gas.

temperature tends to drive the system to the more dense (solid) phase, and an increase in temperature tends to drive the system to the more entropic (liquid) phase.

When going from a low-temperature phase to a high-temperature phase entropy always increases  $\Delta S > 0$ , because  $c_P \equiv T(\partial S/\partial T)_P > 0$ .

The sign of  $\Delta V$  is more uncertain though. To see this point let us consider the C-C equation at the solid-liquid (now solid is region 1 and liquid region 2) coexistence curve. At the melt temperature:

$$\left(\frac{dP}{dT}\right)_{coex} = \frac{\delta Q_{melt}}{T_{melt} \Delta v_{melt}} \quad \delta Q_{melt} = Q_{liq} - Q_{solid} > 0 \quad (2.21)$$

In general,  $\Delta v_m = v_{liq} - v_{solid} > 0$  which implies  $(dP/dT)_{coex} > 0$ . There are cases, however, where  $\Delta v_m = v_{liq} - v_{solid} < 0$  because  $\rho_{liq} > \rho_{solid}$  (for instance the  $H_2O$ , or also Silicon and Germanium). The paradigmatic example is the freezing of water where  $v_{ice} > v_{liq}$  since ice is less dense than liquid water at the coexistence ( $0 < T < 4$ ). This implies that  $dP/dT < 0$ .

**Example 9** (Melting point on Everest).

Consider  $T = 237K$  and  $P = P_0$ . If  $\delta Q_m = 6.01kJ/mol$  and  $\Delta v = -1.7cm^3/mol$  we have:

$$\frac{dP}{dT} = \frac{\delta Q_m}{T \Delta v} = \frac{6.0110^3 J/mol}{273 \cdot (-1.7cm^3/mol)} = -1.29 \cdot 10^4 J/m^3 = -1.29bar/K$$

$$\Delta T = \frac{\Delta P}{(-1.29Pa/K)} = \frac{(P_0 - P_{Everest})}{(-1.29Pa/K)} = \frac{(1 - 0.36)atm}{(-1.29Pa/K)} = -0.5C$$

$$\Rightarrow T_m(Everest) = T_m(P_0) + 0.5C$$

**Example 10** (Boiling point on Everest).

Consider  $P_{Everest} = 0.36atm$ ,  $\rho(T = 100^\circ C) = 0.598kg/m^3$ ,  $L_{ge} = 2.257 \cdot 10^3 J/g$ . The density of the vapour is about 1000 less than water, it implies that:  $\Delta V = V_g - V_e \approx V_g = \frac{1}{\rho_g}$ . We have:

$$\frac{dP}{dT} = \frac{L_{ge}}{T \Delta V} = \frac{L_{ge} \rho_g}{T} = \frac{2.25 \cdot 10^3 J/g \cdot 0.593kg/m^3}{373K} = \frac{3.6 \cdot 10^3 J}{K} \frac{kg}{g} \frac{1}{m^3} = 3.6 \cdot 10^3 Pa/K$$

$$\Rightarrow \Delta T \approx \Delta P/(3.610^3 Pa/K) = 18^\circ C \rightarrow T_0 - T_{Everest} = 18^\circ C \Rightarrow T_{Everest} \approx 80^\circ C$$

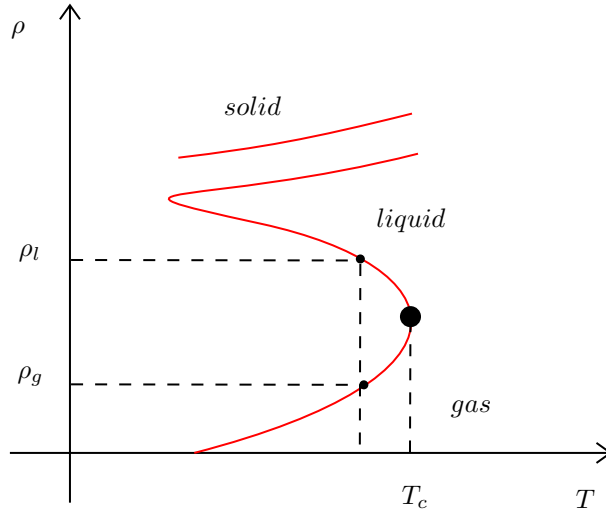
## 2.5 Order parameter of a phase transition

An order parameter is a measure of the degree of order across the boundaries in a phase transition system. In particular, *order parameters* are macroscopic observable that are equal to zero above the critical temperature, and different from zero below:

$$O_p = \begin{cases} \neq 0 & T < T_c \\ = 0 & T \rightarrow T_c^- \end{cases} \quad (2.22)$$

When a phase transition implies a breaking of a phase symmetry, the order parameter is related to this symmetry. Therefore, the order parameter reflects the symmetry of the system. Recall that, at  $T_c$  the system has a symmetry broken.

For instance, consider the densities of liquid and gas and the related order parameter of the gas-liquid transition  $\Delta\rho = \rho_l - \rho_g$ , that is  $\neq 0$  for  $T \neq T_c$  but  $\rightarrow 0$  when  $T \rightarrow T_c$  (see Figure 2.15).



**Figure 2.15:**  $(T, \rho)$  projection of the  $(P, V, T)$  system, where  $\rho = N/V$ .

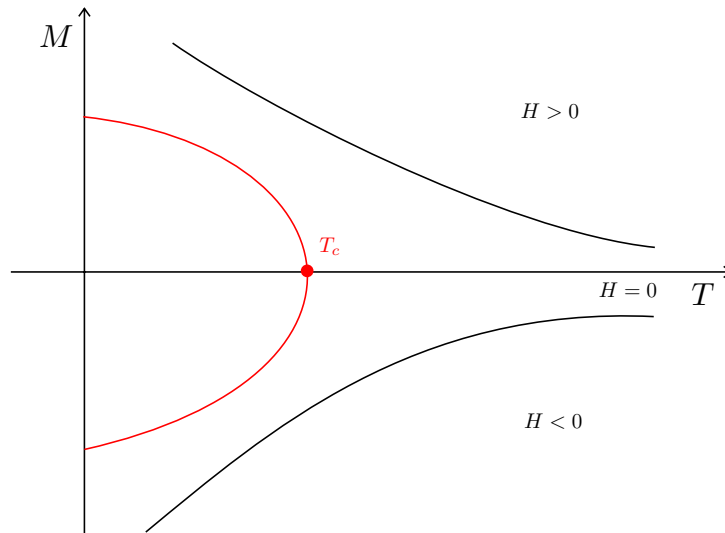
*Remark.* Note that  $\rho = \frac{N}{V} = \frac{1}{v}$  hence either  $N$  or  $V$  varies.

In Figure 2.16 is shown the behaviour for a ferromagnetic system. Clearly  $M \neq 0$  if  $H \neq 0$ . Recall that  $M$  is the order parameter of the paramagnetic-ferromagnetic phase transition:

$$H = 0 \Rightarrow \begin{cases} M \neq 0 & T < T_c \\ M \rightarrow 0 & T \rightarrow T_c^- \end{cases}$$

### Variable conjugate to $O_P$

- *Ferromagnetic system:*  $\vec{M} \rightarrow \vec{H}$  (magnetic field).
- *Ferroelectric:*  $\vec{P} \rightarrow \vec{E}$  (electric field).
- *Liquid crystals:*  $Q_{\alpha\beta} \rightarrow \vec{E}, \vec{H}$ .
- *Fluid:*  $V \rightarrow P$  (pressure), or  $\rho \rightarrow \mu$ .



**Figure 2.16:** Magnetization of a ferromagnet. In red: zero-field magnetization. Below the critical temperature there is a spontaneous magnetization.

## 2.6 Classification of the phase transitions

### 2.6.1 Thermodynamic classification

Thermodynamically one can distinguish two kinds of phase transitions:

1. Ones who develop latent heat.
2. Ones who do not develop latent heat. The entropy changes continuously at the transition.

### 2.6.2 Ehrenfest classification

The *Ehrenfest classification* is based on the behaviour of the derivatives of the thermodynamic potentials.

A phase transition is of order  $n$  if all the  $(n - 1)$  derivatives are continuous and the  $n^{th}$  derivative displays a finite discontinuity.

**Example 11.** For instance, the first order transition  $S = -(\partial G/\partial T)_P$  has finite discontinuity.

*Remark.* There are first order transitions where  $S$  is continuous (no latent heat) but  $\rho$  is discontinuous ( $v = (\partial G/\partial P)_T$ ).

**Example 12.** Second order transition. The specific heat displays a finite jump, see Figure 2.17c in the conductor-superconductor transition.

Another example is a second order transition but with divergence. Consider the fluid-superfluid transition (or  $\lambda$  transition) of the  $\text{He}_4$  (Figure 2.17d).

*Remark.*  $\lambda$  transition: a second-order or higher-order transition, in which the heat capacity shows either a discontinuity (second-order) or a vertex (higher-order) at the transition temperature. It is so named because the shape of the specific heat versus temperature curve resembles the Greek letter  $\lambda$ .

### 2.6.3 Modern classification

A phase transition is of the first order if exists a finite discontinuity in either one or more partial derivatives of the thermodynamic potentials. If instead the first

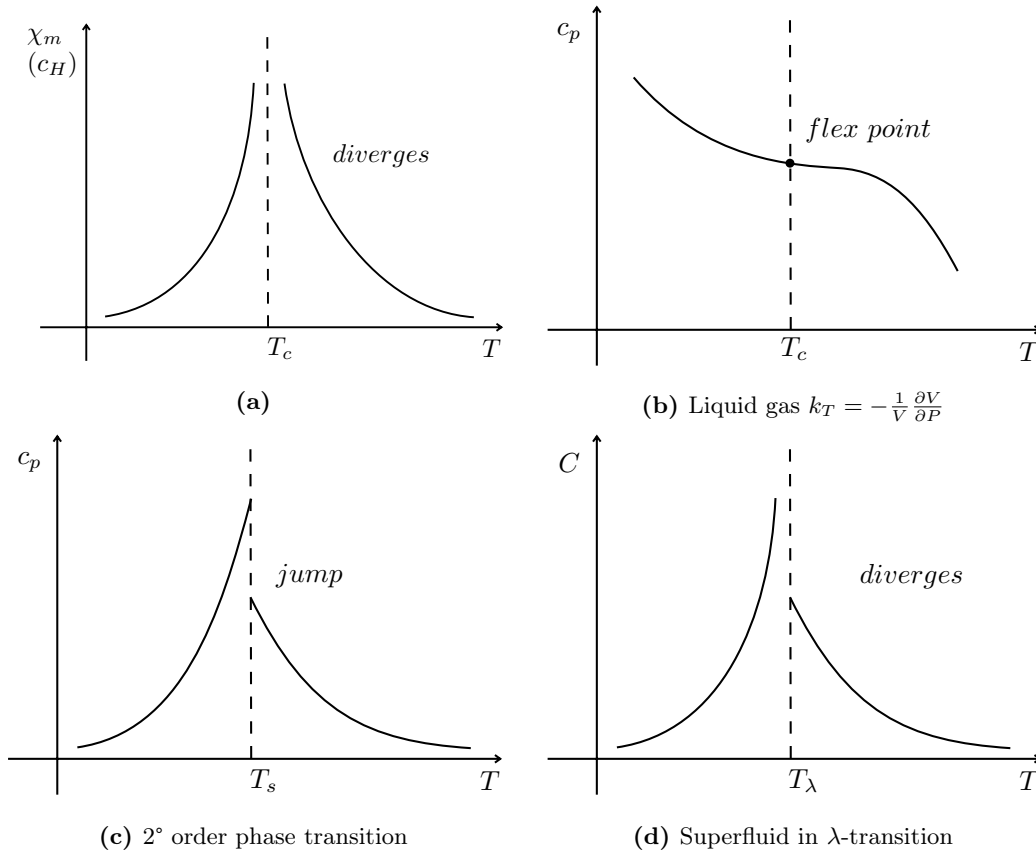


Figure 2.17: Description

derivatives are all continuous but the second are either discontinuous or infinite one talks of continuous transitions. A critical point is a continuous transition.

## 2.7 Critical exponents

At the critical point response functions may diverge. How are these divergence? In general, when you are close to  $T_c$ , there are singularities. Now, we can ask, how the curve diverges? What is the behaviour close to the critical point? Power law, so which are the values of these critical exponents?

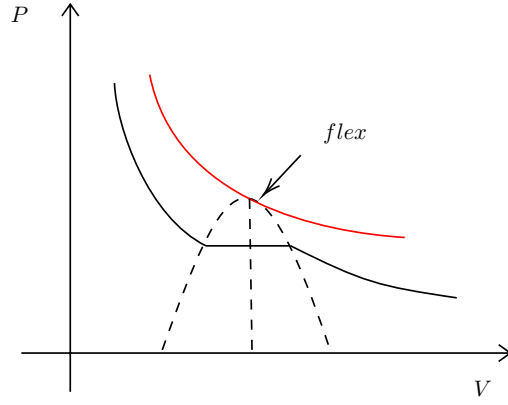
### 2.7.1 Divergence of the response functions at the critical point

While at the critical point the order parameter goes to zero continuously as  $T \rightarrow T_c^-$ , the response function may develop divergences.

**Example 13.** In a fluid system since at  $T = T_c$  the curve  $P = P(V)$  develops an horizontal flex (Figure 2.18), we have  $k_T = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_T \rightarrow \infty$ . Similarly, in a magnetic since the curve is like Figure 2.16, we have  $\chi_T = \left( \frac{\partial M}{\partial H} \right)_T \xrightarrow{T \rightarrow T_c} \infty$ .

### 2.7.2 Critical exponents definition

The notion of *critical exponent* describes the behaviour of the order parameter and the response functions in proximity of the critical point. In order to answer to these questions, let us define:



**Figure 2.18:**  $(V, T)$  projection.

**Definition 1** (*Critical Exponent, or Scale Exponent*). Let us define the adimensional parameter measuring the distance from the critical point  $t \equiv \frac{T-T_c}{T_c}$ , the *Critical Exponent*  $\lambda$  associated to the function  $F(t)$  is defined as:

$$\lambda_{\pm} = \lim_{t \rightarrow 0^{\pm}} \frac{\ln |F(t)|}{\ln |t|} \quad (2.23)$$

We note that it behaves like a power law and that one can also write the *power law*:

$$F(t) \stackrel{t \rightarrow 0^{\pm}}{\sim} |t|^{\lambda_{\pm}} \quad (2.24)$$

More generally, for  $t \ll 1$ :

$$F(t) = A|t|^{\lambda_{\pm}}(1 + bt^{\lambda_1} + \dots), \quad \lambda_1 > 0 \quad (2.25)$$

where all other terms are less important.

**Definition 2** (*Thermodynamic critical exponents*).

- **Exponent**  $\beta$ : tells how the order parameter goes to zero. Consider Figure 2.19a, we have  $M \stackrel{t \rightarrow 0^-}{\sim} (-t)^{\beta}$ . No sense in going from above where it stays 0.
- **Exponent**  $\gamma_{\pm}$  (susceptibility): related to the response function. Consider Figure 2.19b, we have  $\chi_T \stackrel{t \rightarrow 0^{\pm}}{\sim} |t|^{-\gamma_{\pm}}$ . In principle, the value of  $\gamma$  can depend on the sign of  $t$  i.e.  $\gamma^+ \neq \gamma^-$ , but they are the same in reality and we have  $\gamma^+ = \gamma^- = \gamma$ .
- **Exponent**  $\alpha_{\pm}$ : how specific heat diverges (second order derivative in respect of  $T$ ). For instance see Figure 2.19c, we have  $c_H \sim |t|^{-\alpha_{\pm}}$ .
- **Exponent**  $\delta$ : in this case one consider the isotherm  $T = T_c$  and look for the behaviour of  $M$  at the critical point at small  $H$  (or viceversa). The result is  $M \sim H^{1/\delta}$ . In Figure 2.19d,  $H \sim |M|^{\delta} \text{sign}(M)$ .

*Remark.* In compiling Table 2.1 and 2.2 we have made the as yet totally unjustified assumption that the critical exponent associated with a given thermodynamic variable is the same as  $T \rightarrow T_c$  from above or below.

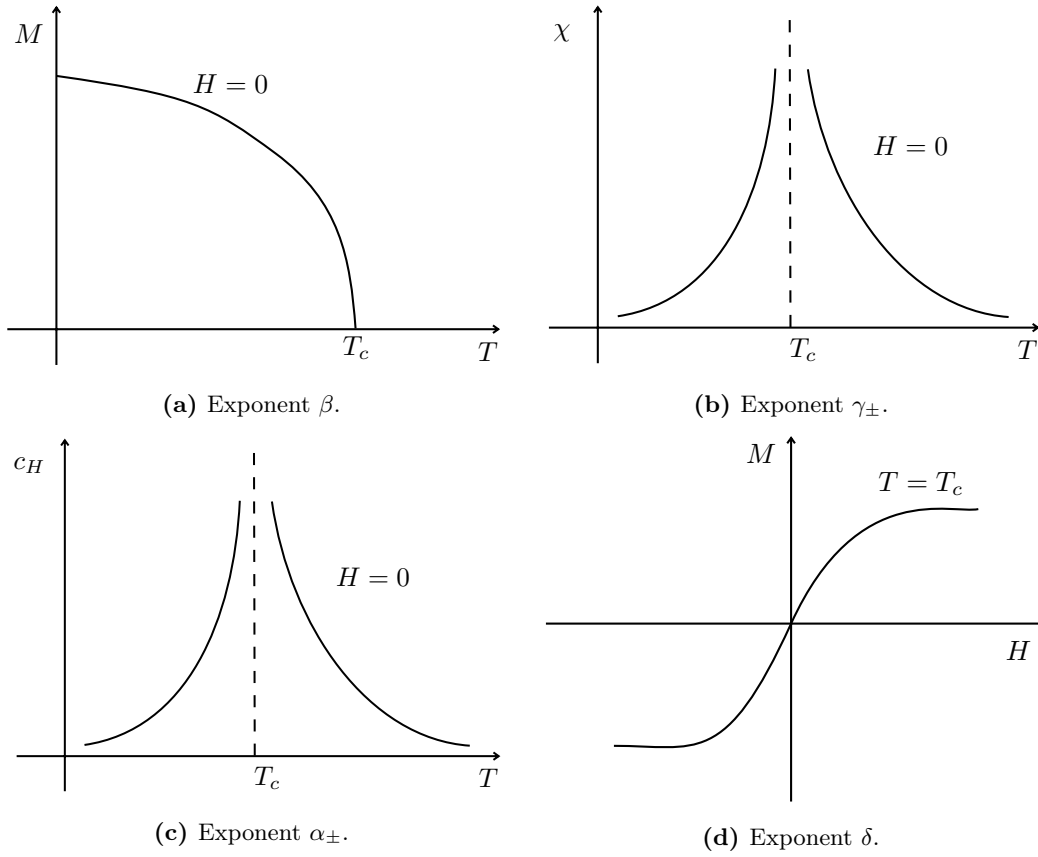


Figure 2.19

Zero-field specific heat	$C_H \sim  t ^{-\alpha}$
Zero-field magnetization	$M \sim (-t)^{\beta}$
Zero-field isothermal susceptibility	$\chi_T \sim  t ^{-\gamma}$
Critical isotherm ( $t = 0$ )	$H \sim  M ^{\delta} \text{sign}(M)$
Correlation length	$\xi \sim  t ^{-\nu}$
Pair correlation function at $T_c$	$G(\vec{r}) \sim \frac{1}{r^{d-2+\eta}}$

Table 2.1: Definitions of the most commonly used critical exponents for a magnetic system [4].

Specific heat at constant volume $V_c$	$C_V \sim  t ^{-\alpha}$
Liquid-gas density difference	$(\rho_l - \rho_g) \sim (-t)^{\beta}$
Isothermal compressibility	$k_T \sim  t ^{-\gamma}$
Critical isotherm ( $t = 0$ )	$P - P_c \sim  \rho_l - \rho_g ^{\delta} \text{sign}(\rho_l - \rho_g)$
Correlation length	$\xi \sim  t ^{-\nu}$
Pair correlation function at $T_c$	$G(\vec{r}) \sim \frac{1}{r^{d-2+\eta}}$

Table 2.2: Definitions of the most commonly used critical exponents for a fluid system [4].

### 2.7.3 Law of the corresponding states

The system displays correlation at very long distance, these goes to the size of the system when  $T \rightarrow T_c$ . We are talking about long range correlation. The *correlation function* is  $\xi \sim t^{-\nu}$ . For instance, consider a polymer as in Figure 2.20a.

Having defined the critical exponents we need to justify why they are interesting. And indeed, why they are more interesting than the critical temperature  $T_c$  itself. It turns out that, whereas  $T_c$  depends sensitively on the details of the interatomic interactions, the critical exponents are to a large degree *universal* depending only on a few fundamental parameters.

To summarize, the critical exponents are more interesting than  $T_c$  since their values do not depend on microscopic details but only on few parameters such as the space dimension  $d$  and the symmetry of the system.

One of the first experimental evidence of this universality was given by the work of Guggenheim on the coexistence curves of  $g$  different fluids: A, Kn,  $\chi_e$ , Ne,  $N_2$ ,  $CO_2$  and  $O_2$ . By plotting  $T/T_c$  versus  $\rho/\rho_c$  (Figure 2.20b) he found that all the data collapse on the same curve, i.e. different sets of data fit the same function. Moreover for  $t \rightarrow 0$ :

$$(\rho_l - \rho_c) \sim (-t)^\beta \quad (2.26)$$

and  $\beta \sim 1/3 \approx 0.335$ . Therefore, close to the critical point all the data lie on the same curve and hence can be described by the same exponent  $\beta$ . A further test of universality is to compare this value to that obtained for a phase transition in a completely different system with a scalar order parameter. For instance, if we do the same for a string ferromagnetic the result is  $\beta = 1/3$  too.

*Remark.* The law of corresponding states gives a universal liquid-gas coexistence curve.

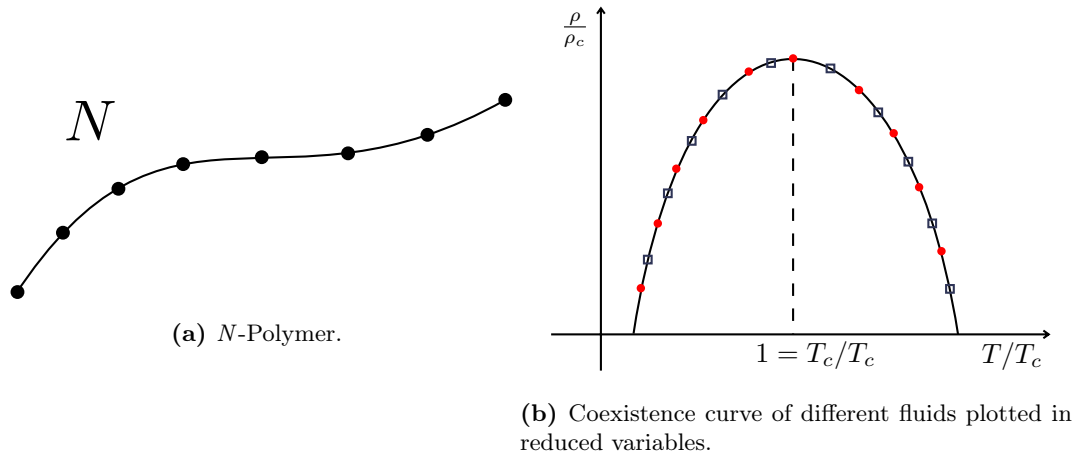


Figure 2.20

#### 2.7.4 Thermodynamic inequalities between critical exponents

It is possible to obtain several rigorous inequalities between the critical exponents. The easiest to prove is due to Rushbrooke.

##### Rushbrooke inequality

It follows from the well known thermodynamic relation between the specific heats at constant field and constant magnetization. Remember the relation between response functions:

$$k_T(c_p - c_v) = Tv\alpha^2 = Tv\frac{1}{v^2}\left(\frac{\partial v}{\partial T}\right)_P^2 = T\frac{1}{v}\left(\frac{\partial v}{\partial T}\right)_P^2 \quad (2.27)$$



For magnetic systems one has

$$\chi_T(c_H - c_M) = T \left( \frac{\partial M}{\partial T} \right)_H^2 \quad (2.28)$$

from the thermodynamic stability we have  $c_M \geq 0, \chi_T \geq 0$ . Hence from the previous relation we have

$$c_H \geq \frac{T}{\chi_T} \left( \frac{\partial M}{\partial T} \right)_H^2 \quad (2.29)$$

On the other hand, for  $T \rightarrow T_c^-$  ( $t \rightarrow 0^-$ ) and  $H = 0$  (zero field) we have

$$\begin{cases} c_H \sim (-t)^{-\alpha} \\ \chi_T \sim (-t)^{-\gamma} \end{cases} \quad (2.30)$$

Moreover,  $M \sim (-t)^\beta$ , implies  $\left( \frac{\partial M}{\partial T} \right)_{H=0} \sim (-t)^{\beta-1}$ .

Since the inequality (2.29) is valid for all temperature  $T$ , it follows that can only be obeyed if

$$B(T_c - T)^{-\alpha} \geq B'T \frac{[(T_c - T)^{\beta-1}]^2}{(T_c - T)^{-\gamma}} \quad (2.31)$$

with  $B, B' > 0$ . Take the limit  $T \rightarrow T_c^-$  we have:

$$\lim_{T \rightarrow T_c^-} (T_c - T)^{2-\alpha-2\beta-\gamma} \geq \frac{B'T}{B} > 0 \quad (2.32)$$

Since the left hand side must be strictly greater than zero, we have the *RushBrook inequality*:

$$\alpha + 2\beta + \gamma \geq 2 \quad (2.33)$$

### Griffith inequality

The *Griffith inequality* is obtained from the convexity property (in  $T$  and  $V$ ) of the Helmholtz free energy and from  $A \sim t^{2-\alpha}$ :

$$\Rightarrow \alpha + \beta(1 + \delta) \geq 2 \quad (2.34)$$

We have introduced two very new ideas, universality and inequalities between the critical exponents which appear to hold as equalities.

In the intervening chapters we look at models of systems which undergo phase transitions and how to calculate their critical exponents and other properties.



## Chapter 3

# Recall of statistical mechanics and theory of ensembles

### 3.1 Statistical ensembles

Statistical mechanics roughly speaking was born as a sort of theory from microscopic and try to compute the macroscopic length using thermodynamics. The problem is going from the continuous problems to the macroscopic problems. In origin was statistical mechanics of equilibrium system. Each microstate with a given energy fixed, will have the same probability, this is the equal probability statement.

In general, if we consider a system with  $N, V$  (number of particles and volume) fixed and also the total energy  $E$  fixed, we call  $\Omega(E, V, N)$  the number of microstate with total energy  $E$ , volume  $V$  and number of particles  $N$ .

If the system is *isolated* and in *equilibrium* the rule of **equal probability** of the microstates holds:

If the system is isolated and in equilibrium with energy  $E$  it visits each microstate consistent with energy  $E$  with equal probability.

Another way to say is: the system spends the same amount of time in each of the  $\Omega(E, V, N)$  microstates.

Therefore, we call a single configuration of a given microstate  $\mathcal{C}$ . A configuration is just when you have the spatial part, because momentum can be obtained by integrating. Suppose you want to compute the probability of a given configuration  $\mathcal{C}$ ,  $P_{\mathcal{C}}$ , because there is equal probability:

$$P_{\mathcal{C}} = \frac{1}{\Omega(E, V, N)} \quad (3.1)$$

Let us now consider two subsystem 1 and 2 that can exchange energy, volume and/or particles. The number of microstates of the combined system of total energy  $E_T = E_1 + E_2$ , total volume  $V_T = V_1 + V_2$  and  $N_T = N_1 + N_2$  is given by:

$$\Omega(E_T, V_T, N_T) = \sum_{E_1, V_1, N_1} \Omega_1(E_1, V_1, N_1) \Omega_2(E_T - E_1, V_T - V_1, N_T - N_1) \quad (3.2)$$

One can show that in the thermodynamic limit at equal  $\Omega(E_T, V_T, N_T)$  is strongly peaked around a given point  $(E_1^*, V_1^*, N_1^*)$  and the fluctuations around this value are rare and small. Writing  $\Omega(E_T, V_T, N_T)$  as

$$\Omega(E_T, V_T, N_T) \propto e^{\frac{S(E_T, V_T, N_T)}{k_B}} = \sum_{E_1, V_1, N_1} \exp \left[ \frac{1}{k_B} (S_1(E_1, V_1, N_1) + S_2(E_2, V_2, N_2)) \right] \quad (3.3)$$

**Lecture 4.**

Friday 18<sup>th</sup>

October, 2019.

Compiled:

Wednesday 11<sup>th</sup>

December, 2019.

The values  $(E_1^*, V_1^*, N_1^*)$  are obtained by the max entropy condition that can be written as

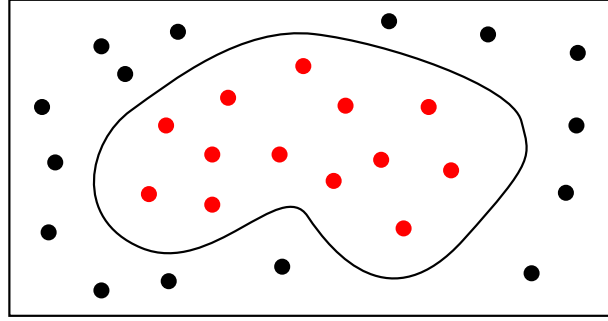
$$\frac{d \ln \Omega_1}{dE_1} = \frac{d \ln \Omega_2}{dE_2} \Rightarrow T_1 = T_2 \quad (3.4a)$$

$$\frac{d \ln \Omega_1}{dV_1} = \frac{d \ln \Omega_2}{dV_2} \Rightarrow P_1 = P_2 \quad (3.4b)$$

$$\frac{d \ln \Omega_1}{dN_1} = \frac{d \ln \Omega_2}{dN_2} \Rightarrow \mu_1 = \mu_2 \quad (3.4c)$$

We next consider these properties to the case in which 1 is the system we want to study and 2 is a much larger system than 1 (a bath). This setup will bring us to the *canonical ensemble*.

### 3.2 The canonical ensemble



**Figure 3.1:** Isolated system. There are two subsystems,  $S$  constituted by red points and  $B$  constituted by the black one.

Let us consider an isolate system made by two subsystems, one  $S$  and one much larger,  $B$ , that we call *thermal bath* (Figure 3.1). The total number of particles is given by  $N_T = N_B + N_S$  with  $N_B \gg N_S \gg 1$  (they are both large but  $B$  is much larger than  $S$ ), where  $N_B$  are the particles in the thermal bath and  $N_S$  the particle of the system. Let  $E_T$  be the energy of the composite system. The two subsystems can exchange energy but the whole system has constant energy  $E_T$ . Therefore, let the energy to be free to fluctuate in time at fixed temperature  $T_B$  (isothermal ensembles). Note that  $V_S, N_S, V_B, N_B$  are fixed (no exchange of volume and particles). For reasuming, other quantitties fixed are the temperature of the bath  $T_B$ , the number of the total particles of the system  $N_T$ , and also the total volume  $V_T$ . We have also  $V_T = V_B + V_S$ , with  $V_B \gg V_S$ .

The key to the canonical formalism is the determination of the probability distribution of the system among its microstates. And this problem is solved by the realization that the system plus the bath constitute a closed system, with fixed temperature, to which the principle of equal probability of microstates applies.

If one assumes that the system and the bath are *weakly coupled* (neglect interaction energy):

$$E_T = E_S + E_B = \text{const} \quad E_B \gg E_S$$

Let  $\mathcal{C}$  by the microstate of the system  $S$ , and  $\mathcal{G}$  the microstate of the heat bath  $B$ . A given microstate of the isolated composite system  $B$ - $S$  is given from a pair  $(\mathcal{C}, \mathcal{G})$  of microstate  $\mathcal{C} \in S$  and  $\mathcal{G} \in B$ . The number of microstates of the isolated system with total energy  $E_T$  and system energy  $E_S$  is given by:

$$\Omega_T(E_T, E_S) = \Omega(E_S)\Omega_B(E_T - E_S) \quad (3.5)$$

*Remark.* In this analysis  $V$  and  $N$  are fixed. Since  $E_T$  is fixed

$$\Omega_T(E_T) = \sum_{E_S} \Omega(E_S) \Omega_B(E_T - E_S) \quad (3.6)$$

From the principle of equal probability for microstates at equilibrium, the probability of a composed microstate ( $\mathcal{C} \circ \mathcal{G}$ ) is given by:

$$P_{\mathcal{C} \circ \mathcal{G}} = \begin{cases} \frac{1}{\Omega_T(E_T)} & E_{\mathcal{C}} + E_{\mathcal{G}} = E_T \\ 0 & \text{otherwise} \end{cases} \quad (3.7)$$

Since we are not interested to the microstates of the heat bath

$$P_{\mathcal{C}} = \sum_{\substack{\text{all } \mathcal{G} \\ \text{such that} \\ g(E_T - E_{\mathcal{C}} - E_{\mathcal{G}})}} P_{\mathcal{C} \circ \mathcal{G}} = \sum_{\mathcal{G}} \frac{1}{\Omega_T(E_T)} = \frac{1}{\Omega_T} \sum_{\mathcal{G}} 1 \quad (3.8)$$

The number of microstates  $\mathcal{G}$  with energy  $E_{\mathcal{G}} = E_T - E_{\mathcal{C}}$  is given by:

$$\Omega_B(E_{\mathcal{G}}) = \Omega_B(E_T - E_{\mathcal{C}}) \quad (3.9)$$

This implies that the probability of a given configuration is related to the number of microstate of the bath:

$$\Rightarrow P_{\mathcal{C}} = \frac{\Omega_B(E_T - E_{\mathcal{C}})}{\Omega_T(E_T)} \propto \Omega_B(E_T - E_{\mathcal{C}}) \quad (3.10)$$

It is more convenient to deal with the logarithmic of  $P_{\mathcal{C}}$  that is smoother

$$\Rightarrow k_B \ln \Omega_B(E_T - E_{\mathcal{C}}) = S_B \quad (3.11)$$

This is the entropy of  $B$  and is a function of  $N_B$ . Since  $E_{\mathcal{C}} \ll E_B \simeq E_T$  we can expand  $S_B(E_T - E_{\mathcal{C}})$  around  $E_T$  by the small amount

$$\Delta \equiv x - x_0 = -E_{\mathcal{C}}$$

$$f(E_B) = f(E_T) + \left. \frac{df}{dE_B} \right|_{E_B=E_T} (E_B - E_T) + \dots$$

Therefore:

$$k_B \ln \Omega_B(E_B) = S_B(E_B) = S_B(E_T) - E_{\mathcal{C}} \left( \frac{\partial S_B}{\partial E_B} \right)_{E_B=E_T} + \frac{E_{\mathcal{C}}^2}{2} \left( \frac{\partial^2 S_B}{\partial E_B^2} \right)_{E_B=E_T} + \dots \quad (3.12)$$

To make explicit the  $N_B$  dependence let us consider the molar version

$$\begin{aligned} S_B &\rightarrow N_B s_B & E_B &\rightarrow N_B e_B \\ s_B N_B &= N_B s_B(E_T) - E_{\mathcal{C}} \left( \frac{\partial s_B}{\partial e_B} \right)_{e_B=e_T} + \frac{E_{\mathcal{C}}^2}{2 N_B} \left( \frac{\partial^2 s_B}{\partial e_B^2} \right) \end{aligned} \quad (3.13)$$

Let us consider the limit in which the system size is fixed while the one of the heat bath is going to  $\infty$ :

$$\lim_{N_B \rightarrow \infty} \frac{E_T}{N_B} = \frac{E_S + N_B e_B}{N_B} \rightarrow e_B \quad (3.14a)$$

$$\lim_{N_B \rightarrow \infty} k_B \ln \Omega_B(E_T - E_{\mathcal{C}}) \rightarrow N_B s_B - E_{\mathcal{C}} \frac{ds_B}{de_B} \quad (3.14b)$$

On the other hand

$$\frac{ds_B}{de_B} \equiv \frac{1}{T_B} = \frac{1}{T} \quad (3.15)$$

which implies

$$P_{\mathcal{C}} \propto \Omega_B(E_T - E_{\mathcal{C}}) = \exp\left(\frac{N_B s_B}{k_B} - \frac{E_{\mathcal{C}}}{k_B T}\right) \quad (3.16)$$

Since the first term does not depend on  $\mathcal{C}$ , it can be absorbed in the constant and what we get by expanding considering the huge number of particles

$$P_{\mathcal{C}} \propto \exp(-E_{\mathcal{C}}/k_B T) \quad (3.17)$$

*Observation 1.* Since the energy of the system fluctuates, its microstates are not anywhere equiprobable but are visited with probability given by (3.17).

*Remark.* Since the bath is very large  $T$  is the only property of the bath that affects the system. The *Boltzmann factor* is defined as:

$$\beta \equiv \frac{1}{k_B T} \quad (3.18)$$

The normalization consists in dividing by the normalization factor that is the sum of all microstates

$$P_{\mathcal{C}} = \frac{e^{-\beta E_{\mathcal{C}}}}{\sum_{\mathcal{C}} e^{-\beta E_{\mathcal{C}}}} \quad (3.19)$$

Finally, the **canonical partition function** is defined as

$$Q(T, V, N) \equiv \sum_{\substack{\text{all } \mathcal{C} \\ \text{with } V, N \\ \text{fixed}}} \exp(-\beta E_{\mathcal{C}}) \quad (3.20)$$

Given  $Q(T, V, N)$  one gets the *Helmholtz free energy*:

$$A(T, V, N) = -k_B T \ln Q(T, V, N) \quad (3.21)$$

that is the free energy describing the isothermal (or canonical) ensemble at fixed  $T$ , volume  $V$  and number of particles  $N$ .

*Remark.*

$$Q(T, V, N) = \sum_{\substack{\mathcal{C} \\ V, N \text{ fixed}}} e^{-\beta E(\mathcal{C})} = \sum_E e^{-\beta E} \Omega(E, V, N)$$

What we have done is a *foliation* in energy of the space, that is a sum over the energy (keeping  $\{V, N\}$  fixed):

$$Q(T, V, N) = \sum_E e^{-\beta E} \Omega(E, V, N) = \sum_E e^{-\beta E} e^{S/k_B} = \sum_E e^{-\beta(E - TS)}$$

so,

$$Q(T, V, N) = e^{-\beta A} \Rightarrow A = -k_B T \ln Q(T, V, N)$$

We have now formulated a complete algorithm for the calculation of a fundamental relation in the canonical formalism. Given a list of states of the system, and their energies  $E_{\mathcal{C}}$ , we calculate the partition function (3.20). The partition function is thus obtained as a function of temperature and of the parameters that determine the energy levels. The fundamental relation is (3.21) that determines the Helmholtz potential.

The probability of a configuration can be written as (3.19), that is a very useful form. Indeed, the average energy is expected to be

$$U = \sum_{\mathbf{e}} E_{\mathbf{e}} P_{\mathbf{e}} = \frac{\sum_{\mathbf{e}} E_{\mathbf{e}} e^{-\beta E_{\mathbf{e}}}}{\sum_{\mathbf{e}} e^{-\beta E_{\mathbf{e}}}} \quad (3.22)$$

or

$$U = -\frac{\partial}{\partial \beta} \ln Q \quad (3.23)$$

### 3.2.1 Energy fluctuations in the canonical ensemble

Despite energy in the canonical ensemble fluctuates while in the microcanonical one is constant this does not contradict the equivalence principle of the ensemble (in the thermodynamic limit). The reason is that the relative size of the energy fluctuation decreases in the large system limit. Remember that thermodynamic assume that the number of number of freedom is related to the number of Avogadro. To see it let us compute the average square fluctuations of  $E$ .

$$\langle (\delta E)^2 \rangle = \langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 \quad (3.24)$$

On the other hand,

$$\langle E \rangle = \sum_{\mathbf{e}} P_{\mathbf{e}} E_{\mathbf{e}} = \sum_{\mathbf{e}} E_{\mathbf{e}} \frac{e^{-\beta E_{\mathbf{e}}}}{\sum_{\mathbf{e}} e^{-\beta E_{\mathbf{e}}}} = - \left( \frac{\frac{\partial Q(T, V, N)}{\partial \beta}}{Q} \right) = - \left( \frac{\partial}{\partial \beta} \ln Q \right)_{N, V} \quad (3.25)$$

$$\langle E^2 \rangle = \sum_{\mathbf{e}} P_{\mathbf{e}} E_{\mathbf{e}}^2 = \frac{\left( \frac{\partial^2 Q}{\partial \beta^2} \right)}{Q} \quad (3.26)$$

Therefore,

$$\begin{aligned} \langle (\delta E)^2 \rangle &= \langle (E - \langle E \rangle)^2 \rangle = \frac{1}{Q} \left( \frac{\partial^2 Q}{\partial \beta^2} \right)_{N, V} - \frac{1}{Q^2} \left( \frac{\partial Q}{\partial \beta} \right)_{N, V}^2 \\ &= \left( \frac{\partial^2 \ln Q}{\partial \beta^2} \right)_{N, V} = - \left( \frac{\partial \langle E \rangle}{\partial \beta} \right)_{N, V} \end{aligned} \quad (3.27)$$

Since

$$C_v = \left( \frac{\partial E}{\partial T} \right)_{N, V} \quad (3.28)$$

we have

$$\langle (\delta E)^2 \rangle = k_B T^2 C_v \quad (3.29)$$

Both  $c_V$  and  $\langle E \rangle$  are extensive

$$\frac{\sqrt{\langle (\delta E)^2 \rangle}}{\langle E \rangle} = \frac{\sqrt{k_B T^2 C_v}}{\langle E \rangle} \sim O\left(\frac{1}{\sqrt{N}}\right) \Rightarrow 0 \quad (3.30)$$

because  $N \sim 10^{23}$ .

### 3.3 Isothermal and isobaric ensemble

Now the system is coupled both to a thermal and a volumic bath at temperature  $T_B$  and pressure  $P_B$ . The idea is now: consider the same system with the bath; the difference is that in this case the system can exchange energy but also volume (we continue to keep the temperature of the bath fixed). At this point the ensemble is *isothermal* and *isobaric*. All the assumptions done before are valid, in particular assuming as before *weak coupling* between the degrees of freedom of the bath and those of the system

$$\begin{aligned} E_T &= E + E_B \\ V_T &= V + V_B \end{aligned}$$

we look for the partition function that describes this isothermal and isobaric ensemble. Similarly to the previous case one can write

$$P_{\mathcal{C}} \propto \Omega_B(E_B, V_B) \propto \Omega_B(E_T - E_{\mathcal{C}}, V_T - V_{\mathcal{C}}) \propto \exp[S_B(E_T - E_{\mathcal{C}}, V_T - V_{\mathcal{C}})/k_B] \quad (3.32)$$

*Remark.* Now  $\mathcal{C}$  is specified both by its volume  $V$  and energy  $E$ . As before one can expand  $\log \Omega_B$  both in  $E_B$  and in  $V_B$  (around  $E_T$  and  $V_T$ ) and take the limit  $N_B \rightarrow \infty$ .

$$P_{\mathcal{C}} \propto \exp \left[ \frac{S_B(E_T, V_T)}{k_B} - \frac{E_{\mathcal{C}}}{k_B} \frac{\partial S_B}{\partial E_B} \Big|_{E_T, V_T} - \frac{V_{\mathcal{C}}}{k_B} \frac{\partial S_B}{\partial V_B} \Big|_{V_T, E_T} + \left( \text{term} \simeq \frac{1}{N_B} \right) \right] \quad (3.33)$$

Recalling that

$$\frac{dS}{dV} \Big|_E = \frac{P}{T} \quad \text{with} \quad \begin{cases} P_B \rightarrow P \\ T_B \rightarrow T \end{cases} \quad (3.34)$$

$$\Rightarrow P_{\mathcal{C}} \propto \exp \left[ -\frac{E_{\mathcal{C}}}{k_B T} - \frac{P V_{\mathcal{C}}}{k_B T} \right] \quad (3.35)$$

If we normalize:

$$P_{\mathcal{C}} = \frac{e^{-\beta(E_{\mathcal{C}} + P V_{\mathcal{C}})}}{\Delta(T, P, N)} \quad (3.36)$$

where

$$\Delta(T, P, N) = \sum_{\mathcal{C}} e^{-\beta(E(\mathcal{C}) + P V(\mathcal{C}))} \quad (3.37)$$

is called the **Gibbs partition function**.

Note that

$$\begin{aligned} \Delta(T, P, N) &= \sum_V e^{-\beta P V} \left( \sum_{\substack{\mathcal{C} \\ V, N \text{ fixed}}} e^{-\beta E_{\mathcal{C}}} \right) \stackrel{\text{Laplace transform}}{=} \sum_V e^{-\beta P V} Q(T, V, N) \\ &= \underbrace{\sum_V \sum_E}_{\text{fluctuating variables}} e^{-\beta(E + P V)} \Omega(E, V, N) \end{aligned} \quad (3.38)$$

By summing over all the microstates compatible with  $E$  and  $V$ :

$$P(E, V) = \frac{\Omega(E, V, N)}{\Delta(T, P, N)} e^{-\beta(E + P V)} \quad (3.39)$$



*Remark.*

$$\Delta(T, P, N) = \underbrace{\sum_E \sum_V e^{-\beta E - \beta PV} \Omega(E, V, N)}_{\text{Laplace transform}} = \sum_{E, V} e^{-\beta E - \beta PV + S(E, V, N)/k_B} \quad (3.40)$$

### Classical systems (fluids)

$$\Delta(T, P, N) = \int_0^\infty dV e^{-\beta PV} \left[ \frac{1}{h^{3N} N!} \int d\vec{p}_1 \dots d\vec{p}_N e^{-\beta \mathcal{H}(p^N, r^N)} \right] \quad (3.41)$$

which implies

$$\Delta(T, P, N) = \int_0^\infty dV e^{-\beta PV} Q(T, V, N) \quad (3.42)$$

that is the Laplace transform of the canonical partition function  $Q$ .

$$\beta P \equiv \frac{P}{T} = \frac{\partial S}{\partial V} \quad (3.43)$$

### Magnetic system

Ensamble in which both  $E$  and  $M$  can fluctuate  $\Omega(E, M)$  (both with  $T_B$  and  $H_B$ ).

$$P_{\mathcal{C}} \propto e^{S_B(E_{\mathcal{C}}, M_{\mathcal{C}})/k_B - \frac{E_{\mathcal{C}}}{k_B} \frac{dS_B}{dE_B} - \frac{M_{\mathcal{C}}}{k_B} \frac{dS_B}{dM_B}} \quad (3.44)$$

Since  $\frac{dS_B}{dM_B} = -\frac{H_B}{T_B}$  and  $\frac{dS_B}{dE_B} = \frac{1}{T_B}$ :

$$\Rightarrow P_{\mathcal{C}} \propto \exp[-\beta(E_{\mathcal{C}} - H M_{\mathcal{C}})], \quad T_B \rightarrow T, H_B \rightarrow H \quad (3.45)$$

Normalization function:

$$\Delta(T, H, N) = \sum_{\mathcal{C}} e^{-\beta(E_{\mathcal{C}} - H M_{\mathcal{C}})} = \sum_{E, M} e^{-\beta E + \beta M H} \Omega(E, M) \quad (3.46)$$

that is the *Gibbs partition function* for magnetic systems.

#### 3.3.1 Saddle point approximation

The sum (3.40) can be approximated by the maximum of the integrand (this is fair for highly peaked functions):

$$\sum_{E, V} \exp[-\beta E - \beta PV + S(E, V, N)/k_B] \approx e^{-\beta E^* - \beta PV^* + S(E^*, V^*, N)/k_B} \quad (3.47)$$

where

$$\left( \frac{dS(E^*, V^*, N)}{dE} \right)_{V, N} = \frac{1}{T}, \quad \left( \frac{dS(E^*, V^*, N)}{dV} \right)_{E, N} = \frac{P}{T} \quad (3.48)$$

this implies

$$-k_B T \ln \Delta(T, P, N) \simeq E^* + PV^* - TS \quad (3.49)$$

We define the **Gibbs free energy**:

$$G(T, P, N) = -k_B T \ln \Delta(T, P, N) \quad (3.50)$$

### 3.4 Gran canonical ensemble

In this case  $N$  varies instead than  $V$ .

$$\begin{aligned}
 P_{\mathfrak{c}} &= \exp \left[ \frac{S_B(E_T - E_{\mathfrak{c}}, N_T - N_{\mathfrak{c}})}{k_B} \right] \\
 &\sim \exp \left[ \frac{S_B(E_T, N_T)}{k_B} - \frac{E_{\mathfrak{c}}}{k_B} \frac{dS_B}{dE_B} - \frac{N_{\mathfrak{c}}}{k_B} \frac{dS_B}{dN_B} + \left( \text{terms of order } \leq \frac{1}{V_B} \right) \right] \quad (3.51) \\
 &= \frac{\exp[-\beta E_{\mathfrak{c}} + \beta \mu N_{\mathfrak{c}}]}{\Theta(T, V, \mu)}
 \end{aligned}$$

where

$$\Theta(T, V, \mu) = \sum_N \sum_{\substack{\mathfrak{c} \\ V, N \text{ fixed}}} e^{-\beta(E_{\mathfrak{c}} - \mu N)} \quad (3.52)$$

is the **grancanonical partition function**.

*Remark.* Remember that

$$\frac{dS}{dE} = \frac{1}{T}, \quad \frac{dS}{dN} = \frac{\mu}{T} \quad (3.53)$$

The *fugacity* is defined as:

$$z \equiv e^{\beta \mu} \quad (3.54)$$

and we rewrite

$$\Theta(T, V, \mu) = \sum_{N=0}^{\infty} z^N \left( \sum_{\substack{\mathfrak{c} \\ V, N \text{ fixed}}} e^{-\beta E_{\mathfrak{c}}} \right) \quad (3.55)$$

In principle, if one is able to compute the partition function is able to compute the thermodynamic quantities.

## Chapter 4

# Statistical mechanics and phase transitions

### 4.1 Statistical mechanics of phase transitions

From the microscopic degrees of freedom one compute the partition function in the appropriate ensemble, then the corresponding thermodynamic potential and from it all the thermodynamic properties of the system as *equilibrium phases* and, if present, *phase transitions*. Actually, until the '30 there were strong concerns about the possibility that statistical mechanics could describe phase transitions.

**Lecture 5.**

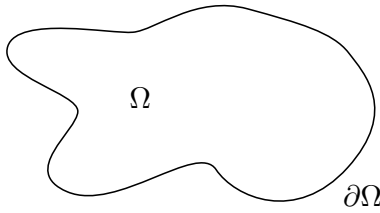
Wednesday 23<sup>rd</sup>

October, 2019.

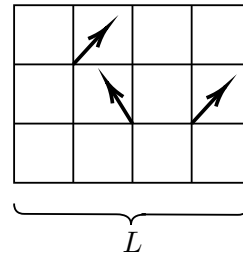
Compiled:

Wednesday 11<sup>th</sup>

December, 2019.



(a) Region  $\Omega$  with boundary  $\partial\Omega$ .



(b) Magnetic system with characteristic length  $L$ .

**Figure 4.1**

Let us consider a system with a region  $\Omega$  of volume  $V(\Omega)$  and boundary  $\partial\Omega$  of area  $S(\Omega)$  (Figure 4.1a). Denoting by  $L$  a characteristic length of the system

$$V(\Omega) \propto L^d, \quad S(\partial\Omega) \propto L^{d-1}$$

where  $d$  is the spatial dimension.

*Remark.* Space  $\Omega$  can be either *discrete* or *continuous*.

Suppose that the system is *finite*. Formally we can write

$$\mathcal{H}_\Omega = - \sum_n k_n \Theta_n \quad (4.1)$$

where

- $k_n$ : are the coupling constants. In general, but not always, they are *intensive thermodynamic variables*.
- $\Theta_n$ : is a linear, or higher order, combination of the dynamical microscopic degrees of freedom (local operators in quantum statistical mechanics).

- $k_n \Theta_n$ : must obey the symmetry of the system. It is important that in principle the term satisfies the symmetry of the system. This is a master rule!

To fix the idea let us consider two classical examples: the magnetic system and the fluid system.

#### 4.1.1 Magnetic system (canonical)

The degrees of freedom are the *spins* lying on a Bravais lattice  $\vec{\mathbf{S}}_i$  with  $1 \leq i \leq N(\Omega)$ , where the  $N(\Omega)$  are the number of lattice sites (Figure 4.1b). A configuration is the orientation of the spin in each site  $\mathcal{C} = \{\vec{\mathbf{S}}_1, \dots, \vec{\mathbf{S}}_N\}$ . We have:

$$\Theta_1 = \sum_i \vec{\mathbf{S}}_i \quad (4.2a)$$

$$\Theta_2 = \sum_{ij} \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_j \quad (4.2b)$$

We consider the trace operation that is the sum over all possible values that each degree of freedom can assume:

$$\text{Tr} \equiv \sum_{\{\mathcal{C}\}} \equiv \sum_{\vec{\mathbf{S}}_1} \sum_{\vec{\mathbf{S}}_2} \cdots \sum_{\vec{\mathbf{S}}_N} \quad (4.3)$$

where  $\sum$  can also indicate an integration if values are continuous. The *canonic partition function* is

$$Q_\Omega(T, \{k_n\}) = \text{Tr} \left( e^{-\beta \mathcal{H}_\Omega} \right) \quad (4.4)$$

with  $\beta \equiv \frac{1}{k_B T}$ .

#### 4.1.2 Fluid system (gran canonical)

Consider  $N$  particles in a volume  $V$ , with number density  $\rho = N/V$ . The  $2dN$  degrees of freedom are

$$\{\mathcal{C}\} = \{(\vec{\mathbf{x}}_i, \vec{\mathbf{p}}_i)_{i=1, \dots, N}\}$$

and

$$\Theta_1 = \sum_i \left[ \frac{\vec{\mathbf{p}}_i^2}{2m_i} + U_1(\vec{\mathbf{x}}_i) \right] \quad (4.5a)$$

$$\Theta_2 = \frac{1}{2} \sum_{i>j} U(|\vec{\mathbf{x}}_i - \vec{\mathbf{x}}_j|) \quad (4.5b)$$

The trace operation is

$$\text{Tr} \equiv \sum_{\{\mathcal{C}\}} = \sum_{N=0}^{\infty} \frac{1}{N!} \int \prod_{i=1}^N \frac{(d\vec{\mathbf{p}}_i)(d\vec{\mathbf{x}}_i)}{h^{dN}} \quad (4.6)$$

The *gran canonical partition function* is:

$$\mathcal{F}_\Omega = \text{Tr} \left( e^{-\beta(\mathcal{H}_\Omega - \mu N)} \right) \quad (4.7)$$

For a generic partition function  $Q_\Omega(T, \{k_n\})$  we can define the finite system *free energy* as

$$F_\Omega[T, \{k_n\}] = -k_B T \ln Q_\Omega(T, \{k_n\}) \quad (4.8)$$

The relation with thermodynamic is through the *thermodynamic limit*. Since the free energy is an extensive function

$$F_\Omega \propto V(\Omega) \sim L^d$$

In general, one can write

$$F_\Omega[T, \{k_n\}] = V(\Omega)f_b[T, \{k_n\}] + S(\Omega)f_s[T, \{k_n\}] + O(L^{d-2}) \quad (4.9)$$

where  $f_b[T, \{k_n\}]$  is the bulk free energy density.

**Definition 3.** We define

$$f_b[T, \{k_n\}] \equiv \lim_{V(\Omega) \rightarrow \infty} \frac{F_\Omega[T, \{k_n\}]}{V(\Omega)} \quad (4.10)$$

If the limit exists (to prove for each system) and does not depend on  $\Omega$ .

For a system defined on a lattice we have

$$\begin{aligned} L(\Omega) &\propto N(\Omega)^{1/d}, \quad V(\Omega) \propto N(\Omega) \\ f_b[T, \{k_n\}] &= \lim_{N(\Omega) \rightarrow \infty} \frac{1}{N(\Omega)} F_N[T, \{k_n\}] \end{aligned} \quad (4.11)$$

To get information on surface property of the system

$$f_s[T, \{k_n\}] \equiv \lim_{S(\Omega) \rightarrow \infty} \frac{F_\Omega[T, \{k_n\}] - V(\Omega)f_b}{S(\Omega)} \quad (4.12)$$

#### 4.1.3 Thermodynamic limit with additional constraints

For a fluid we cannot simply take the limit  $V(\Omega) \rightarrow \infty$  by keeping  $N$  fixed, otherwise we will always get a infinite system with *zero density*. One has to take also the limit  $N(\Omega) \rightarrow \infty$  such that:

$$\frac{N(\Omega)}{V(\Omega)} \equiv \rho = \text{const}$$

In general is not so easy to prove the existence of the limit and it depends on the range of the particle-particle interactions.

#### 4.1.4 Statistical mechanics and phase transitions

Since all the thermodynamic information of a system can be obtained by the partition function, in principle also the ones concerning the existence and nature of the phase transition must be contained in  $Z$  (or  $Q$ ). On the other hand, we know from thermodynamic that phase transitions are characterized by singularities in the derivation of  $F$ . Also  $Z$  must display these singularities. Moreover,  $Z$  is a sum of exponentials

$$Z_\Omega = \text{Tr} \left( e^{-\beta \mathcal{H}_\Omega} \right) \quad (4.13)$$

These are analytic functions everywhere (it converges), therefore  $Z_\Omega$  is analytic for  $\Omega$  finite! The question is: where the singularities came from? It is only in the thermodynamic limit that singularities in  $F$  and hence points describing phase transitions can arise!

For summarizing, there is no way out of this for producing singularities. The singularities will develop in the thermodynamic limits. For reach singularities we have to reach so precision in thermodynamic that we are not able to go exactly into the critical point. How can we relate singularities in the behaviour of the system geometrically?

## 4.2 Critical point and correlations of fluctuations

From thermodynamics we know that, at the critical point, some response functions may diverge. Now, we show that this is a consequence of the onset of microscopic fluctuations that are spatially correlated over long distances. To see this let us compute the response of a ferromagnetic in presence of an external magnetic field  $H$ . The Gibbs partition function of a generic magnetic system is

$$Z_{\text{Gibbs}}[T, \{k_n\}] = \text{Tr} \left( e^{-\beta(\mathcal{H}(\mathcal{C}) - HM(\mathcal{C}))} \right) = \sum_{M, E} e^{-\beta E + \beta HM} \Omega(E, M) \quad (4.14)$$

*Remark.* The term  $(-HM)$  is the work done by the system against the external field  $H$  to maintain a given magnetization  $M$ .

$$\langle M \rangle = \frac{\partial \ln Z_G}{\partial(\beta H)} \Big|_T = \frac{1}{Z_G} \text{Tr} \left[ M(\mathcal{C}) e^{-\beta(\mathcal{H}(\mathcal{C}) - HM(\mathcal{C}))} \right] \quad (4.15)$$

$$\chi_T = \frac{\partial \langle M \rangle}{\partial H} = \left\{ \frac{\beta}{Z_G} \text{Tr} \left[ M^2(\mathcal{C}) e^{-\beta\mathcal{H} + \beta HM} \right] - \frac{\beta}{Z_G^2} \left[ \text{Tr} \left[ M(\mathcal{C}) e^{-\beta\mathcal{H} + \beta HM} \right] \right]^2 \right\} \quad (4.16)$$

Hence,

$$\chi_T = \frac{1}{k_B T} \left( \langle M^2 \rangle - \langle M \rangle^2 \right) \quad (4.17)$$

The thermodynamic response function  $\chi_T$  in statistical mechanics is related to the variance of the magnetization.

We can relate the above expression with the correlation of the microscopic by performing a coarse-graining of the system where the magnetization  $M(\mathcal{C})$  can be computed as an integral

$$M(\mathcal{C}) = \int d^3 \vec{r} m(\vec{r}) \quad (4.18)$$

Hence

$$k_B T \chi_T = \int d\vec{r} d\vec{r}' \left[ \langle m(\vec{r}) m(\vec{r}') \rangle - \langle m(\vec{r}) \rangle \langle m(\vec{r}') \rangle \right] \quad (4.19)$$

Let us assume the *translational symmetry*:

$$\begin{cases} \langle m(\vec{r}) \rangle = m & \text{homogeneous} \\ \langle m(\vec{r}) m(\vec{r}') \rangle \equiv G(\vec{r} - \vec{r}') & \text{two-point correlation function} \end{cases} \quad (4.20)$$

Let us consider instead the *connected correlation function*, i.e. the correlation function of the fluctuations  $\delta m = m - \langle m \rangle$ :

$$\langle m(\vec{r}) m(\vec{r}') \rangle_c \equiv \langle (m(\vec{r}) - \langle m(\vec{r}) \rangle) (m(\vec{r}') - \langle m(\vec{r}') \rangle) \rangle = G(\vec{r} - \vec{r}') - m^2 \quad (4.21)$$

Given the translational invariance one can centre the system such that its centre of mass coincides with the origin

$$\begin{aligned} \vec{r}_{CM} &\Rightarrow \vec{r}_0 \equiv \vec{0} \\ &\Rightarrow \int d\vec{r} \int d\vec{r}' [G(\vec{r} - \vec{r}_0) - m^2] \end{aligned} \quad (4.22)$$

The integration over  $\vec{r}'$  gives the volume  $V(\Omega)$  of the system:

$$\underbrace{k_B T \chi_T}_{\text{response function}} = V(\Omega) \int d\vec{r} \underbrace{\langle m(\vec{r}) m(\vec{r}_0) \rangle_c}_{\substack{\text{correlation function} \\ \text{of the fluctuations} \\ \text{of the local magnetization}}} \quad (4.23)$$

The equation (4.23) is called the **Fluctuation-Dissipation relation**.

How  $G_c(\vec{r})$  behaves? In general one has

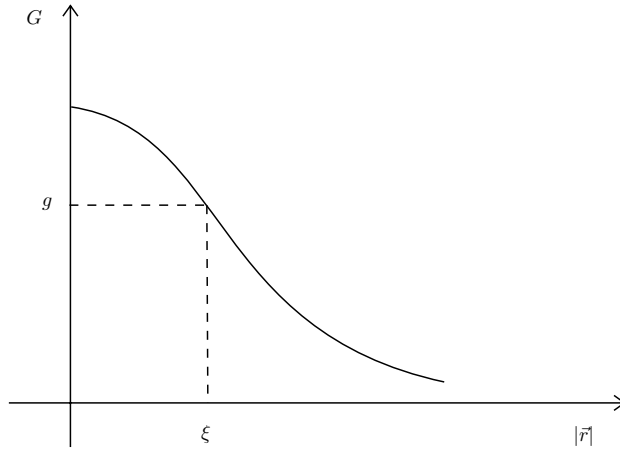
$$G_c(\vec{r}) \sim e^{-|\vec{r}|/\xi} \quad (4.24)$$

meaning that for  $|\vec{r}| > \xi$  the fluctuations are uncorrelated, where  $\xi$  is the *correlation length*. The correlation length it is related to the correlation function. In general is finite but if you approach  $T_c$ , it diverges. In fact, at the critical point this correlation will expand in the whole space and reaches the size of all the system, in other words, it goes to infinity ( $\xi \rightarrow \infty$ ). When  $\xi$  will diverge, there will not be anymore the exponential and the integral cannot be kept finite.

Let  $g$  be the value of  $G_c$  for  $|\vec{r}| < \xi$ :

$$k_B T \chi_T \leq V g \xi^3 \quad (4.25)$$

where there is an inequality because we are underestimating the integral (Figure 4.2).



**Figure 4.2:** Plot of the two-point correlation function,  $G$ .

Rearranging the terms we obtain

$$\frac{k_B T \chi_T}{V} < g \xi^3 \quad (4.26)$$

Hence, if  $\chi_T$  diverges at the critical point it implies  $\xi \rightarrow \infty$ .

In particular, one can see that for  $H = 0$  and  $T \rightarrow T_c^\pm$ :

$$\xi_\pm(T, H = 0) \sim |t|^{-\nu_\pm} \quad (4.27)$$

where  $\nu_+ = \nu_- = \nu$  is the correlation length critical exponent.

*Remark.* It does not derive from thermodynamic considerations.

Scaling (4.27) is often used as the most general definition of a critical point. One can also show that at  $T = T_c$  (i.e.  $t = 0$ )

$$G_c(r) \sim \frac{1}{r^{d-2+\mu}} \quad (4.28)$$

where  $\mu$  is the correlation critical exponent.

*Remark.* The formula is a power law decay instead than exponential.

### 4.3 Finite size effects and phase transitions

Actually, the thermodynamic limit is a mathematical trick and in real systems it is never reached. Is it then physically relevant?

If we had instruments with *infinite* precision each change of the physical properties of a system would occur within a finite range, therefore we would observe a smooth crossover instead than a singularity. In this respect the notion of correlation length  $\xi$  is extremely important.

To illustrate this point let us consider the gas-liquid system in proximity of its critical point ( $T \sim T_c$ ). If we approach  $T_c$  from the gas phase, there will be fluctuations of  $\rho$  with respect to  $\rho_G$ ,  $\Delta\rho = \rho - \rho_G$ , due to the presence of denser droplets (liquid) in the continuum gas phase. These droplets will have different diameters, but the average size would be  $\xi$ , where it is the typical size of the liquid droplets. Clearly  $\xi = \xi[T]$  and, in proximity of the critical point  $\xi \xrightarrow{t \rightarrow 0} |t|^{-\nu}$ .

On the other hand, in a finite system,  $\xi$  cannot diverge since is bounded above,  $\xi \leq L$ , where  $L$  is the linear system size.

As  $T \rightarrow T_c$ , where  $\xi$  should be larger than the system size, the behaviour of the system should deviate from the one expected by the theory that is obtained in the limit  $L \rightarrow \infty$ . How far the real system would be from the critical point  $t = 0$  where singularities develop? Let us try to give an estimate of this deviation.

Let us consider a system of size  $L = 1$  cm and

$$t \equiv (T - T_c)/T_c, \quad \xi \sim \xi_0 t^{-\nu}$$

Let us assume that the lattice distance is  $\xi_0 = 10 \text{ \AA}$ . Hence,

$$t \sim \left( \frac{\xi}{\xi_0} \right)^{-1/\nu} \sim \left( \frac{L}{10 \text{ \AA}} \right)^{-1/\nu} \sim (10^{10})^{-1/\nu} \quad (4.29)$$

In the next chapters, we will see that  $\nu < 1$  and close to  $1/2$ , hence:

$$t \sim (10^{10})^{-2} = 10^{-20}$$

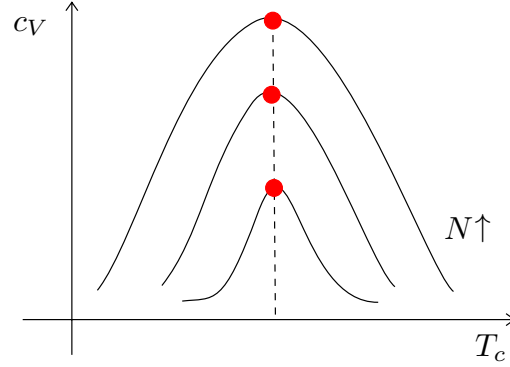
Therefore we have  $t \approx 10^{-20}$  as distance from  $T_c$ .

This estimate suggests that the experimental instrument that measures temperature must have a precision of  $10^{-20}$  to see deviations from the results obtained in the thermodynamic limit.



## 4.4 Numerical simulations and phase transitions

In this case the size  $L$  of the simulated system is few multiples of  $\xi_0$  and the finite-size effects of the simulated data can strongly affect the location and the scaling laws of the phase transition under numerical investigation. Finite size scaling analysis of the numerical data is needed.



**Figure 4.3:**  $(T_c, C_V)$  plot at different  $N$ .

We can find the critical point by doing Montecarlo simulation. Supposing a Montecarlo simulation of a Ising model, for which there is no an analitic solution and compute the energy. Try to estrapolate for example the position of the peak as  $N$  increases. If we start to see the behaviour as in Figure 4.3, something is happening. There are two approaches we can use.

The first approach is studying the system by looking for all the details. An example could be a protein, that interact with other proteins; in this case we can look at all the electrons (or atoms). Nevertheless, even if we thought at the simple protein that exist, there would be a lot of degrees of freedom.

For doing a simulation, if we are interested in long time behaviour and in large scale behaviur, details are not important. What it is important are symmetries, ranges of interaction. Therefore, we can forget about all the details. We can introduce the effective potentials as Van der Waals or Lenard Jones potential and studying collective effects. This is the second approach.



## Chapter 5

# Role of the models in statistical mechanics

### 5.1 Role of the models

Which is the role of models in statistical mechanics? There are two possible approaches:

1. The model must describe the real system in a very detailed way. The maximum number of details and parameters to be tuned are included. The *pro* is the closer to the real specific system (faithfull description). The *drawback* is that the model is so complicated that no analytical solution is possible. Moreover, even numerically, these models can be studied for very short times and small sizes. An example is the simulation of the folding dynamics that can be performed for few nanoseconds. On the other hand the introduction of many details are often not crucial if one is interested in large scale properties.
2. Try to introduce the most simple model that satisfies few essential properties of the real system such as its symmetries, dimensionality, range of interactions etc. Since most of the microscopic details are integrated, these models cannot describe the full physics of a specific system but they can reproduce its main features. Moreover these models can be studied numerically and, to some extent also analitically (exact solution).

It is the latter approach that we shall take here. Let us start by introducing what is, perhaps, the most paradigmatic model in the statistical mechanics of phase transition, the *Ising model*.

### 5.2 The Ising model

Suggested by Lenz to Ising for his PHD thesis (1925), it is supposed to describe a magnetic system that undergoes a transition between a paramagnetic and a ferromagnetic phase. In  $d = 1$  the model was solved exactly by Ising. Unfortunately, he found that for  $T > 0$  the model does not display a phase transition.

The wrong conclusion was that this model was not able to describe a phase transition. In fact, it turns out that, for  $d > 1$ , the model does display a paramagnetic-ferromagnetic phase transition.

Let us first discuss some general feature of the model for any dimension  $d$ .

**Lecture 6.**  
Friday 25<sup>th</sup>  
October, 2019.  
Compiled:  
Wednesday 11<sup>th</sup>  
December, 2019.

### 5.2.1 $d$ -dimensional Ising model

For hypercubic lattice with given  $N(\Omega)$  sites  $\{i\}_{i=1,\dots,N(\Omega)}$  and linear size  $L(\Omega)$ , we have

$$N(\Omega) = L^d$$

The microscopic degrees of freedom are the spins  $S_i$ , defined at each  $i$ -esim lattice site. Each spin can assume the values  $S_i = \pm 1$ , that means that at each site the possible values are the spin up or down. For a lattice with  $N(\Omega)$  spins, there are  $2^{N(\Omega)}$  possible configurations.

*Remark.* Since we do not consider the spin as a vector, this is a model for a strongly anisotropic ferromagnet (along a given direction).

The minimal model that can try to capture the interaction between the spin is the following. Suppose to have also an external magnetic field  $H_i$  (it values depends on the site  $i$ ). One can consider interactions between spins whose strength are described by functions  $J_{ij}, k_{ijk}, \dots$ . For instance, there is a coupling that derives from electrons coupling

$$J_{ij} = f(|\vec{r}_i - \vec{r}_j|)$$

The physical origin is the overlap between the electronic orbitals of the neighbouring atoms forming the Bravais lattice. Remember that a term as  $\sum_i S_i$  is not correlated, while we need an interaction for describing the model.

A general Hamiltonian of the model can be written as:

$$\mathcal{H}_\Omega(\{S_i\}) = \sum_{ij} J_{ij} S_i S_j - \sum_i H_i S_i - \sum_{ijk} S_i S_j S_k + \dots \quad (5.1)$$

Standard Ising model one keeps only the two-body interactions:

$$\mathcal{H}_\Omega(\{S_i\}) = -\frac{1}{2} \sum_{i \neq j}^N J_{ij} S_i S_j - \sum_{i=1}^N H_i S_i \quad (5.2)$$

where the first term represents a two body interaction that is a quadratic term, while the second term is a one body interaction. We have put the minus because we want to minimize the energy, but it depends on the sign of  $J$ .

For this model the sum over all configurations on trace is given by

$$\text{Tr} \equiv \sum_{S_1=\pm} \sum_{S_2=\pm} \cdots \sum_{S_N=\pm} \equiv \sum_{\{S\}} \quad (5.3)$$

Our problem is to find the partition function with  $N$  sites, which depends on  $T$  and in principle depends in the configuration given (it is fixed both for  $H$  and  $J$ !). Hence, the canonical partition function is given by

$$Z_\Omega(T, \{H_i\}, \{J_{ij}\}) = \text{Tr} e^{-\beta \mathcal{H}_\Omega(\{S\})} \quad (5.4)$$

and the corresponding *free-energy*,

$$F_\Omega(T, \{H_i\}, \{J_{ij}\}) = -k_B T \ln Z_\Omega \quad (5.5)$$

The *bulk limiting free energy* is:

$$f_b(T, \{H_i\}, \{J_{ij}\}) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln F_\Omega \quad (5.6)$$

How do we know that the above limit does exist? It must be proven. The surface is not important in the bulk limit. Note that we are assuming that the interaction between the spin is a short range force, it is not as the size of the system.

For this model it is possible to show that the limit exists if

$$\sum_{j \neq i} |J_{ij}| < \infty \quad (5.7)$$

*Remark.* In general what determines the existence of the limit of these spin models are the dimension  $d$  and the range of the spins interactions.

For example it is possible to show that, if

$$J_{ij} = A |\vec{r}_i - \vec{r}_j|^{-\sigma} \quad (5.8)$$

so it is a long range interaction, the limit exists when

$$\sigma > d$$

*Remark.* If the interaction is *dipolar* since it decays as  $1/r^3$ , for the case  $d = 3$  the limit does not exist. However, it is still possible to prove the existence of the limit for this case if one assumes that not all dipoles are fully aligned.

Assuming that the thermodynamic limit exists, we now look at some additional rigorous results on the limiting free energy and its derivatives.

### 5.2.2 Mathematical properties of the Ising model with nearest neighbours interactions

For simplicity let us consider the case in which the external magnetic field is homogeneous, i.e.  $H_i \equiv H$ , and the spin-spin interaction is only between spins that are nearest-neighbours (n.n.) on the lattice:

$$J_{ij} = \begin{cases} J & \text{if } i \text{ and } j \text{ are n.n.} \\ 0 & \text{otherwise} \end{cases} \quad (5.9)$$

Now, the model is very simple:

$$-\mathcal{H}_\Omega(\{S\}) = J \sum_{\langle ij \rangle}^{N(\Omega)} S_i S_j + H \sum_i^{N(\Omega)} S_i \quad (5.10)$$

where the notation  $\langle ij \rangle$  means a double sum over  $i$  and  $j$ , with the constraint that  $i$  and  $j$  are nearest-neighbours.

Since  $H$  is uniform, the average magnetization per spin is

$$\langle m \rangle = \frac{1}{N(\Omega)} \sum_{i=1}^{N(\Omega)} \langle S_i \rangle \quad (5.11)$$

where  $\langle \dots \rangle$  means average over the chosen ensemble.

*Remark.* For  $J = 0$ , (5.10) is the hamiltonian of a paramagnet. The only influence ordering the spins is the field  $H$ . They do not interact, there are no cooperative effects and hence no phase transition.

Since

$$\sum_{i=1}^N \langle S_i \rangle = \frac{1}{Z} \text{Tr} \left[ \sum_i S_i e^{-\beta \mathcal{H}_\Omega(\{S_i\})} \right] = \frac{1}{Z} \text{Tr} \left[ \sum_i S_i \exp \left( \beta J \sum_{\langle ij \rangle}^{N(\Omega)} S_i S_j + \beta H \sum_i^{N(\Omega)} S_i \right) \right] \quad (5.12)$$

it is easy to show that:

$$\langle m \rangle = -\frac{1}{N} \frac{\partial F_\Omega}{\partial H} \quad (5.13)$$

where

$$F_\Omega(T, J, H) = -k_B T \ln Z_N(T, J, H) \quad (5.14)$$

Let us now consider the properties of the limiting free-energy

$$f_b = \lim_{N \rightarrow \infty} \frac{1}{N} (-k_B T \ln Z_N) \quad (5.15)$$

It is possible to prove the following properties:

1.  $f_b < 0$ .
2.  $f_b(T, J, H)$  is a *continuous* function of  $T, J$  and  $H$ .
3. The right and left derivatives of  $f_b(T, J, H)$  exist and are equal almost everywhere.
4. The molar entropy  $s = -\frac{\partial f_b}{\partial T} \geq 0$  almost everywhere.
5.  $\frac{\partial f_b}{\partial T}$  is a *monotonic non increasing* function of  $T$ . That is  $\frac{\partial^2 f_b}{\partial T^2} \leq 0$ . This implies that:

$$c_H = T \left( \frac{\partial S}{\partial T} \right)_H = -T \left( \frac{\partial^2 f_b}{\partial T^2} \right)_H \geq 0 \quad (5.16)$$

6.  $\frac{\partial f_b}{\partial H}$  is a *monotonic non increasing* function of  $H$ . That is

$$\frac{\partial^2 f_b}{\partial H^2} \leq 0 \quad (5.17)$$

This implies that

$$\chi_T = \left( \frac{\partial M}{\partial H} \right)_T = - \left( \frac{\partial^2 f_b}{\partial H^2} \right)_T \geq 0 \quad (5.18)$$

*Remark.* The above properties have been postulated in thermodynamics, but here they have been rigorously proved for the Ising model using statistical mechanics.

*Proof of property (4).* Almost everywhere we have to prove that:

$$s \equiv -\frac{\partial f_b}{\partial T} \geq 0$$

Let us consider a finite system

$$\begin{aligned} -\frac{\partial F_\Omega}{\partial T} &= k_B \ln (\text{Tr} e^{-\beta \mathcal{H}_\Omega}) + k_B T \frac{1}{k_B T^2} \frac{\text{Tr}(\mathcal{H}_\Omega e^{-\beta \mathcal{H}_\Omega})}{\text{Tr}(e^{-\beta \mathcal{H}_\Omega})} \\ &= k_B \left[ \ln Z + \frac{\text{Tr}(\beta \mathcal{H}_\Omega e^{-\beta \mathcal{H}_\Omega})}{Z_\Omega} \right] \stackrel{\text{todo}}{=} -k_B T \text{Tr}(\rho_\Omega \ln \rho_\Omega) \end{aligned} \quad (5.19)$$

where

$$\rho_\Omega = \frac{e^{-\beta \mathcal{H}_\Omega}}{Z_\Omega} \quad (5.20)$$

is the probability distribution.

Since  $\rho_\Omega \leq 1$  it implies  $\ln \rho_\Omega \leq 0$  and so  $-\text{Tr}(\rho_\Omega \ln \rho_\Omega)$  is positive. Then, let us divide by  $N(\Omega)$  and take the thermodynamic limit:

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \frac{\partial F_\Omega}{\partial T} = -k_B T \lim_{N \rightarrow \infty} \frac{1}{N} \underbrace{\text{Tr}(\rho_\Omega \ln \rho_\Omega)}_{S_\Omega} = T s \geq 0 \quad \Rightarrow \quad s \geq 0 \quad (5.21)$$

□

All the other properties listed before (except (1)) are consequences of the *convexity* property of  $f_b$ .

**Theorem 5.2.1.**  $f_b(T, J, H)$  is an upper convex (i.e. concave) function of  $H$ .

*Proof.* The proof is based on the Hölder inequality for two sequences  $\{g_k\}, \{h_k\}$ :

**Definition 4** (Hölder inequality). Given  $\{g_k\}, \{h_k\}$  with  $g_k, h_k \geq 0, \forall k$  and two non negative real numbers  $\alpha_1, \alpha_2$  such that  $\alpha_1 + \alpha_2 = 1$ , the following inequality holds

$$\sum_k (g_k)^{\alpha_1} (h_k)^{\alpha_2} \leq \left( \sum_k g_k \right)^{\alpha_1} \left( \sum_k h_k \right)^{\alpha_2} \quad (5.22)$$

Now, consider the partition function:

$$Z_\Omega(H) = \text{Tr} \left[ \exp \left( \beta H \sum_i S_i \right) \underbrace{\exp \left( \beta J \sum_{\langle ij \rangle} S_i S_j \right)}_{G(S)} \right] = \text{Tr} \left[ \exp \left( \beta H \sum_i S_i \right) G(S) \right] \quad (5.23)$$

It implies that

$$Z_\Omega(H_1 \alpha_1 + H_2 \alpha_2) = \text{Tr} \left( \exp \left\{ \beta \alpha_1 H_1 \sum_i S_i + \beta \alpha_2 H_2 \sum_i S_i \right\} G(S) \right) \quad (5.24)$$

On the other hand, since  $\alpha_1 + \alpha_2 = 1$ :

$$G(S) = G(S)^{\alpha_1} G(S)^{\alpha_2} \quad (5.25)$$

$$Z_\Omega(H_1 \alpha_1 + H_2 \alpha_2) = \text{Tr} \left[ (e^{\beta H_1 \sum_i S_i} G(S))^{\alpha_1} (e^{\beta H_2 \sum_i S_i} G(S))^{\alpha_2} \right] \quad (5.26)$$

If we now apply the Hölder inequality we get

$$\begin{aligned} Z_\Omega(H_1 \alpha_1 + H_2 \alpha_2) &\leq \left( \text{Tr} \left( e^{\beta H_1 \sum_i S_i} G(S) \right)^{\alpha_1} \right) \left( \text{Tr} \left( e^{\beta H_2 \sum_i S_i} G(S) \right)^{\alpha_2} \right) \\ &= Z_\Omega(H_1)^{\alpha_1} Z_\Omega(H_2)^{\alpha_2} \end{aligned} \quad (5.27)$$

If we now take the logs and multiply by  $-k_B T$  both sides we have

$$\lim_{N \rightarrow \infty} -\frac{1}{N} k_B T \ln Z_\Omega(H_1 \alpha_1 + H_2 \alpha_2) \geq -\lim_{N \rightarrow \infty} \frac{\alpha_1}{N} k_B T \ln Z_\Omega(H_1) - \lim_{N \rightarrow \infty} \frac{\alpha_2}{N} k_B T \ln Z_\Omega(H_2) \quad (5.28)$$

It implies

$$f_b(H_1 \alpha_1 + H_2 \alpha_2) \geq \alpha_1 f_b(H_1) + \alpha_2 f_b(H_2) \quad (5.29)$$

That is a concave function of  $H$ . □

### 5.2.3 Ising model and $\mathbb{Z}^2$ symmetry.

The symmetry of the system in sense of the Hamiltonian is: we can invert the value of the  $S$  and the Hamiltonian does not change. It is valid when  $H = 0$ , otherwise is not true. Let us see the  $\mathbb{Z}^2$  symmetry and the following interesting relation:

**Lemma 5.2.2.**  $\forall$  function  $\Phi$  of the configuration  $\{S_i\}$ , the following relation holds:

$$\sum_{\{S_i=\pm 1\}} \Phi(\{S_i\}) = \sum_{\{S_i=\pm 1\}} \Phi(\{-S_i\}) \quad (5.30)$$

this is true for all function of the spin.

Now, we consider the hamiltonian of the Ising model:

$$-\mathcal{H}_\Omega = J \sum_{\langle ij \rangle}^{N(\Omega)} S_i S_j + H \sum_i^{N(\Omega)} S_i$$

Clearly,

$$\mathcal{H}(H, J, \{S_i\}) = \mathcal{H}_\Omega(-H, J, \{-S_i\}) \quad (5.31)$$

This is a spontaneous broken symmetry.

Hence,

$$\begin{aligned} Z_\Omega(-H, J, T) &= \sum_{\{S_i=\pm 1\}} \exp[-\beta \mathcal{H}_\Omega(-H, J, \{S_i\})] \stackrel{(5.30)}{=} \sum_{\{S_i=\pm 1\}} \exp[-\beta \mathcal{H}_\Omega(-H, J, \{-S_i\})] \\ &\stackrel{(5.31)}{=} \sum_{\{S_i=\pm 1\}} \exp[-\beta \mathcal{H}_\Omega(H, J, \{S_i\})] = Z_\Omega(H, J, T) \end{aligned} \quad (5.32)$$

Taking  $-k_B T \log$  and the  $\lim_{N \rightarrow \infty} \frac{1}{N}$  we got:

$$F_\Omega(T, J, H) = F_\Omega(T, J, -H) \quad (5.33)$$

If we take the thermodynamic limit, we have

$$\Rightarrow f_b(T, J, H) = f_b(T, J, -H) \quad (5.34)$$

and it means that the free energy density is an even function of  $H$ !

*Remark.* From the finite-size relation (5.33), one can show that a *finite-size* Ising model does not display a transition to a ferromagnetic phase (for all dimension  $d$ ). Indeed,

$$N(\Omega)M(H) = -\frac{\partial F(H)}{\partial H} \stackrel{(5.33)}{=} -\frac{\partial F(-H)}{\partial(H)} = \frac{\partial F(-H)}{\partial(-H)} = -N(\Omega)M(-H) \quad (5.35)$$

Therefore:

$$M(H) = -M(-H), \quad \forall H \quad (5.36)$$

If  $H = 0$ , we have  $M(0) = -M(0)$ , that is valid if and only if  $M(0) = 0$ !

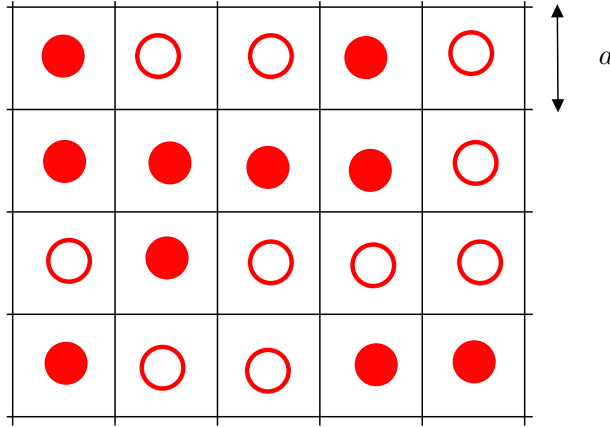
The magnetization of a finite system is, at  $H = 0$ , always zero. This is simply consequence of the symmetry argument shown above. It is only in the thermodynamic limit, where the symmetry is spontaneously broken.



### 5.3 Lattice gas model

Even if we had not seen any transition, the Ising model is interesting because we can use this model to solve other problems that seems different but are not. In fact, the importance of the Ising model relies also on the fact that it can be mapped into other discrete systems. Despite its simplicity the Ising model is widely applicable because it describes any interacting two-state system. One of these applications is the *lattice gas model*, where a gas is put in a lattice.

What is a lattice gas model in more details? The archetypal lattice gas is a model where each lattice site can either be occupied by an atom or vacant. Let us consider a  $d$ -dimensional lattice with coordination number  $z$  and lattice spacing  $a$ , divided into cell as in Figure 5.1. Let us suppose that each cell is either empty or occupied by a single particle (this is more true if  $a \sim \text{\AA}$ ).



**Figure 5.1:**  $d$ -dimensional lattice with lattice spacing  $a$ .

The  $n_i$  is the occupation of the  $i$ -esim cell and it is:

$$n_i = \begin{cases} 0 & \text{if empty} \\ 1 & \text{if occupied} \end{cases} \quad (5.37)$$

We have:

$$N_\Omega = \sum_{i=1}^{N_c} n_i \quad (5.38)$$

where  $N_c$  is the number of the lattice cells. In particular,  $N_c > N_\Omega$ .

The hamiltonian of the model is

$$\mathcal{H}_\Omega = \sum_{i=1}^{N_c} U_1(i)n_i + \frac{1}{2} \sum_{ij} U_2(i,j)n_i n_j + O(n_i n_j n_k) \quad (5.39)$$

where  $U_1$  is for instance an external field, while  $U_2$  is a many body interaction.

Since we want to work in the gran-canonical ensemble,

$$\mathcal{H}_\Omega - \mu N_\Omega = \sum_{i=1}^{N_c} (U_1(i) - \mu)n_i + \frac{1}{2} \sum_{ij} U_2(i,j)n_i n_j + \dots \quad (5.40)$$

and we put  $U_1 = 0$  for convenience.

A formal relation with the Ising model can be obtained by choosing

$$n_i = \frac{1}{2}(1 + S_i) \quad \text{with } S_i = \pm 1 \quad (5.41)$$

The one body term becomes:

$$\sum_i (U_1(i) - \mu) \frac{1}{2} (1 + S_i) = \frac{1}{2} \sum_i (U_1(i) - \mu) + \frac{1}{2} \sum_i S_i (U_1(i) - \mu) \quad (5.42)$$

while the two bodies term is equal to:

$$\frac{1}{2} \sum_{ij} U_2(i, j) \left[ \frac{1}{4} (1 + S_i)(1 + S_j) \right] = \frac{1}{8} 2 \sum_{ij}^{N_c} U_2(i, j) S_i + \frac{1}{8} \sum_{ij}^{N_c} U_2(i, j) S_i S_j + \frac{1}{8} \sum_{ij}^{N_c} U_2(i, j) \quad (5.43)$$

Let us consider only short-range interactions, i.e.

$$U_2(i, j) = \begin{cases} U_2 & i, j \text{ are n.n.} \\ 0 & \text{otherwise} \end{cases} \quad (5.44)$$

It implies

$$\frac{1}{2} \sum_{ij} U_2(i, j) \left[ \frac{1}{4} (1 + S_i)(1 + S_j) \right] = \frac{1}{4} z U_2 \sum_i S_i + \frac{U_2}{4} \sum_{\langle ij \rangle} S_i S_j + \frac{1}{8} U_2 z N_c \quad (5.45)$$

Remember that we put  $U_1 = 0$  for simplicity:

$$\mathcal{H}_\Omega - \mu N_\Omega = E_0 - H \sum_{i=1}^{N_c} S_i - J \sum_{\langle ij \rangle} S_i S_j \quad (5.46)$$

where

$$E_0 = -\frac{1}{2} \mu N_c + \frac{z}{8} U_2 N_c \quad (5.47a)$$

$$-H = -\frac{1}{2} \mu + \frac{z}{4} U_2 \quad (5.47b)$$

$$-J = \frac{U_2}{4} \quad (5.47c)$$

and remember that  $z$  is the coordination number of neighbours.  $J$  is a nearest neighbour interaction which favours neighbouring sites being occupied.

The last equation implies:

$$Z_{LG} = \text{Tr}_{\{n\}} (e^{-\beta(\mathcal{H}_\Omega - \mu N)}) = e^{-\beta E_0} Z_{\text{Ising}}(H, J, N_c) \quad (5.48)$$

We have seen that the Ising model is something more general than the magnetization transition. In the next section, we show how to pass from the partition  $Z$  of a fluid, in the continuum, to the  $Z_{LG}$  of the lattice gas model.

## 5.4 Fluid system in a region $\Omega$

We can consider the system with periodic boundary condition, or within a box, or confined by an external one-body potential.

The hamiltonian for  $N$  particles in  $d$ -dimension is

$$\mathcal{H}_\Omega = \sum_{i=1}^N \left[ \frac{p_i^2}{2m} + U_1(\vec{r}_i) \right] + \frac{1}{2} \sum_{i \neq j} U_2(\vec{r}_i, \vec{r}_j) + \frac{1}{3!} \sum_{i \neq j \neq k} U_3(\vec{r}_i, \vec{r}_j, \vec{r}_k) \quad (5.49)$$

In the gran-canonical ensemble we have:

$$Z_\Omega = \text{Tr} \left( e^{-\beta(\mathcal{H}_\Omega - \mu N)} \right) = \sum_{N=0}^{\infty} \frac{1}{N!} \int \prod_{i=1}^N \frac{d^d \vec{\mathbf{p}}_i d^d \vec{\mathbf{r}}_i}{h^{dN}} \left( e^{-\beta(\mathcal{H}_\Omega - \mu N)} \right) \quad (5.50)$$

and the gran-canonical potential is

$$\omega_\Omega(T, \mu, U_1, U_2, \dots) = -k_B T \ln Z_\Omega \quad (5.51)$$

*Remark.*  $\omega_\Omega(\dots)$  even if it contains an infinite sum is not singular if  $\Omega$  is finite!

Indeed, if  $U_2$  is an hard-core repulsion, each particle has a finite volume and, within a finite  $\Omega$ , only  $N_{max}$  particles can fit in

$$\Rightarrow \sum_{N=0}^{\infty} \sim \sum_{N=0}^{N_{max}}$$

In the thermodynamic limit it corresponds to

$$\omega_b(T, \mu, U_1, U_2, \dots) = \lim_{V(\Omega) \rightarrow \infty} \frac{\omega_\Omega}{V(\Omega)} \quad (5.52)$$

with the constraint

$$\rho = \lim_{V \rightarrow \infty} \frac{\langle N \rangle}{V(\Omega)} = \text{const}$$

Remember also that

$$d\omega_b(T, \mu) = -\sigma dT - \rho d\mu = -P \quad (5.53)$$

Now:

$$Z_N = \sum_{N=0}^{\infty} \frac{1}{N!} \left[ \prod_{i=1}^N \left\{ \int_{-\infty}^{+\infty} d^d \vec{\mathbf{p}} \frac{1}{h^{dN}} e^{-\beta \vec{\mathbf{p}}_i^2 / 2m} \right\} Q_N(T) \right] \quad (5.54)$$

On the other hand since  $\int dx e^{-\alpha x^2} = \sqrt{2\pi/\alpha}$ ,

$$\int_{-\infty}^{+\infty} d^d \vec{\mathbf{p}} \frac{1}{h^{dN}} e^{-\beta \vec{\mathbf{p}}_i^2 / 2m} = \frac{1}{\Lambda(T)^d} \quad (5.55)$$

where

$$\Lambda(T) = \frac{h}{\sqrt{2\pi m k_B T}} \quad (5.56)$$

Therefore,

$$Z_\Omega = \sum_{N=0}^{\infty} \frac{1}{N!} \left( \frac{e^{\beta \mu}}{\Lambda(T)^d} \right)^N Q_N \quad (5.57)$$

with

$$Q_N(T) = \int \prod_{i=1}^N d\vec{\mathbf{r}}_i e^{-\beta(U(\vec{\mathbf{r}}))} \quad (5.58)$$

### 5.4.1 From the continuous to the lattice gas model

Let us divided  $\Omega$  in discrete cells of size  $a$ . If  $a$  is approximate a repulsive range between particles we have that the probability that there is more than a particles sits in a cell is  $\ll 1$ . The potentials of the continuoum model depend on  $\{\vec{\mathbf{r}}_i\}$ .

Consider  $n_\alpha = n_\alpha(\vec{\mathbf{r}}_i)$  the occupation numbers. We have:

$$\sum_{\alpha} n_{\alpha} = N = \int d^d \vec{\mathbf{r}} \sum_{i=1}^N \delta(\vec{\mathbf{r}}_i - \vec{\mathbf{r}}) = \int d\vec{\mathbf{r}} \rho(\vec{\mathbf{r}}) \quad (5.59)$$

where

$$\rho(\vec{\mathbf{r}}) = \sum_{i=1}^N \delta(\vec{\mathbf{r}}_i - \vec{\mathbf{r}}) \quad (5.60)$$

Moreover,

$$\sum_i U_1(\vec{\mathbf{r}}_i) = \sum_i \int_{\Omega} d^d \vec{\mathbf{r}} U_1(\vec{\mathbf{r}}) \delta(\vec{\mathbf{r}}_i - \vec{\mathbf{r}}) = \int_{\Omega} d\vec{\mathbf{r}} U_1(\vec{\mathbf{r}}) \rho(\vec{\mathbf{r}}) \quad (5.61)$$

We have  $U(\{\vec{\mathbf{r}}_i\}) \rightarrow U(\{n_{\alpha}\})$  :

$$Q_N \propto \int \prod_{i=1}^N d^d \vec{\mathbf{r}}_i \rightarrow \sum_{\{n_{\alpha}\}} \quad (5.62)$$

Indeed, for each configuration specified by the set  $\{n_{\alpha}\}$  there are  $N!$  possible configurations of  $\{\vec{\mathbf{r}}_i\}$ . This is because the particles can exchange position between occupied cells. Hence,

$$Q_N \propto \int \prod_{i=1}^N d^d \vec{\mathbf{r}}_i \simeq N! (a^d)^{N_c} \sum'_{\{n_{\alpha}=0,1\}} \dots \quad (5.63)$$

*Remark.* The symbol  $\sum'$  means that the sum has the constraint that the total number of particles is fixed to  $N$ .

Therefore,

$$Q_N \propto N! (a^d)^N \sum'_{\{n_{\alpha}\}} e^{-\beta U(\{n_{\alpha}\})} \quad (5.64)$$

and

$$Z_{\Omega} = \sum_{N=0}^{\infty} \frac{1}{N!} \left( \frac{e^{\beta \mu}}{\Lambda^d(T)} \right)^N Q_N = \sum_{N=0}^{\infty} \left[ \left( e^{\beta \mu} \frac{a}{\Lambda(T)} \right)^d \right]^N \sum'_{\{n_{\alpha}\}} e^{-\beta U(\{n_{\alpha}\})} \quad (5.65)$$

where  $\sum' = \sum_{\{n_{\alpha}\}}$  with the constraint  $\sum_{\alpha} n_{\alpha} = N$ .

*Remark.* In general it is difficult to perform sum with constraints. Fortunately, we are considering the gran-canonical ensemble. Indeed, we can write

$$\sum_{N=0}^{\infty} \sum'_{\{n_{\alpha}\}} f(n_{\alpha}) = \sum'_{\substack{\{n_{\alpha}\} \\ \sum_{\alpha} n_{\alpha}=0}} f(n_{\alpha}) + \sum'_{\substack{\{n_{\alpha}\} \\ \sum_{\alpha} n_{\alpha}=1}} f(n_{\alpha}) \dots + \sum'_{\substack{\{n_{\alpha}\} \\ \sum_{\alpha} n_{\alpha}=\infty}} f(n_{\alpha}) = \sum_{\{n_{\alpha}\}} f(n_{\alpha}) \quad (5.66)$$

with no restriction.

*Remark.* In the final sum all the  $2^N$  possible microscopic states are included (considering  $U_1 = 0$ ):

$$Z_{\Omega}^{GC} \propto \sum_{\{n_{\alpha}\}} \exp \left[ -\beta \left( -\mu - \frac{D}{\beta} \log \frac{a}{\Lambda} \right) \sum_{\alpha=0}^{N_c} n_{\alpha} + \beta U_2 \sum_{\langle \alpha \beta \rangle} n_{\alpha} n_{\beta} + \dots \right] \quad (5.67)$$

$$\Rightarrow Z_{\Omega} = \text{Tr} e^{-\beta(\mathcal{H}_{\Omega} - \tilde{\mu}N)} = Z_{LG}(\tilde{\mu}) \quad (5.68)$$

where

$$\tilde{\mu} = \mu_{LG} = \mu_{phys} + Dk_B T \log \frac{a}{\Lambda} \quad (5.69)$$

## Chapter 6

# Some exactly solvable models of phase transitions

The bibliography of the entire chapter is taken by [5], [6], [7].

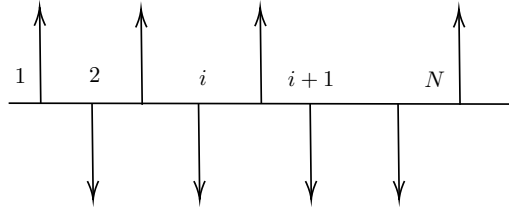
### 6.1 1-dimensional Ising model

In this section, we arrive at the exact solution of the one dimensional Ising model. There are two techniques for solving the model:

1. the *recursive method*;
2. the *transfer matrix method*.

#### 6.1.1 Recursive method

Case with  $H = 0$  and free boundary conditions



**Figure 6.1:** One dimensional Bravais Lattice.

Consider a Bravais lattice in the one dimensional case, that is just a one dimensional lattice, as in Figure 6.1.

The canonical partition function of such a system is:

$$Z_N(T) = \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} \cdots \sum_{S_N=\pm 1} \exp \left[ \overbrace{\beta J}^K \sum_{i=1}^{N-1} S_i S_{i+1} \right] \quad (6.1)$$

The two body interaction is the sum in all the neighbours that in that case are  $(i-1)$  and  $(i+1)$ , but you have only to consider the one after, because the one behind is yet taken by the behind site.

Solve now this partition function. If we consider *free boundary* condition, the  $N$  does not have a  $N+1$ , almost for the moment. Let us define

$$K \equiv \beta J, \quad h \equiv \beta H \quad (6.2)$$

Making explicit the sum in the exponential:

$$Z_N(K) = \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} \cdots \sum_{S_N=\pm 1} e^{K(S_1 S_2 + S_2 S_3 + \cdots + S_{N-1} S_N)} \quad (6.3)$$

What if we just add another spin at the end  $S_{N+1}$  ? Which is the partition function with that new spin? We obtain:

$$Z_{N+1}(K) = \sum_{S_{N+1}=\pm 1} \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} \cdots \sum_{S_N=\pm 1} e^{K(S_1 S_2 + S_2 S_3 + \cdots + S_{N-1} S_N)} e^{K S_N S_{N+1}} \quad (6.4)$$

On the other hand, this sum is just involve this term:

$$\sum_{S_{N+1}=\pm 1} e^{K S_N S_{N+1}} = e^{K S_N} + e^{-K S_N} = 2 \cosh(K S_N) = 2 \cosh(K) \quad (6.5)$$

where the last equivalence derive from the fact that  $\cosh$  is an even function and it does not depend on  $\pm 1$ . Therefore:

$$Z_{N+1}(K) = (2 \cosh(K)) Z_N(K) \quad \text{and} \quad Z_N(K) = (2 \cosh(K)) Z_{N-1}(K) \quad (6.6)$$

By performing a backward iteration

$$Z_N(K) = Z_1 (2 \cosh(K))^{N-1} \quad (6.7)$$

Since  $Z_1 = \sum_{S_1=\pm 1} 1 = 2$ , we have

$$Z_N(K) = 2 (2 \cosh(K))^{N-1} \quad (6.8)$$

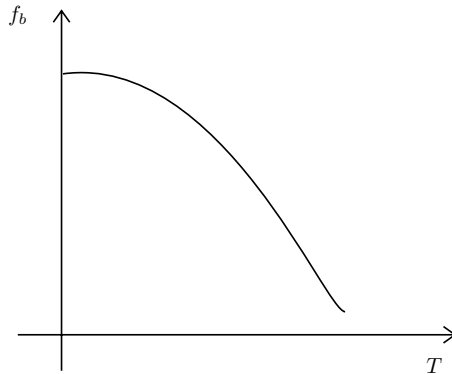
The free energy is:

$$F_N(K) = -k_B T \ln Z_N(K) = -k_B T \ln 2 - k_B T (N-1) \ln (2 \cosh(K)) \quad (6.9)$$

taking the thermodynamic limit it becomes

$$f_b(T) \equiv \lim_{N \rightarrow \infty} \frac{1}{N} F_N(K) = -k_B T \ln \left( 2 \cosh \left( \frac{J}{k_B T} \right) \right) \quad (6.10)$$

As one can see  $f_b(T)$  is an analytic function of  $T$ , so we have no phase transition at  $T \neq 0$ .



**Figure 6.2:** Free energy function in thermodynamic limit for the one dimension Ising model.

We can compute the magnetization (the average over the spin  $\langle S_j \rangle$ ) for a generic site  $j$ . This can be done in many ways. Assume again that  $S_i = \pm 1$ . Here, we choose one that consider another way to compute  $Z$  for the 1 – dimensional Ising model. This method can be useful for other calculations. It is based on the following identity:

$$\exp[K S_i S_{i+1}] \underset{to\ do}{=} \cosh(K) + S_i S_{i+1} \sinh(K) = \cosh(K)[1 + S_i S_{i+1} \tanh(K)] \quad (6.11)$$

It means that

$$Z_N(K) = \sum_{\{S\}} \exp \left[ K \sum_{i=1}^{N-1} S_i S_{i+1} \right] = \sum_{\{S\}} \prod_{i=1}^{N-1} [\cosh(K)(1 + S_i S_{i+1} \tanh(K))] \quad (6.12)$$

by rearranging,

$$Z_N(K) = (\cosh K)^{N-1} \sum_{\{S\}} \prod_{i=1}^{N-1} (1 + S_i S_{i+1} \tanh K) \quad (6.13)$$

If we now expand the products, we get terms of the following form:

$$\sum_{\substack{S_{i_e} = \pm 1 \\ e=1, \dots, M}} (\tanh K)^M S_{i_1} S_{i_{1+1}} S_{i_2} S_{i_{2+1}} \dots S_{i_M} S_{i_{M+1}} = 0 \quad (6.14)$$

where  $i_1 \dots i_m$  is a set of  $M$  sites of the lattice.

*Remark.* The terms above, when summed over  $\{S\}$  are zero, except the term with  $M = 0$  that is equal to 1 and, when summed over  $\{S\}$ , gives  $2^N$ .

Therefore:

$$Z_N(K) = 2^N (\cosh K)^{N-1} \quad (6.15)$$

that coincides with the result obtained before.

If we now compute the average  $\langle S_j \rangle$ , the procedure is similar but now there will be terms as (6.14) but with the addition of an  $S_j$ :

$$(\tanh K)^M S_{i_1} S_{i_{1+1}} S_{i_2} S_{i_{2+1}} \dots S_{i_M} S_{i_{M+1}} S_j \quad (6.16)$$

that, when one sums over  $\{S\}$  are all zero, included the term with  $H = 0$  that now is equal to  $S_j$  and  $\sum_{S_j = \pm 1} = 0$ . Hence, we have the result

$$\langle S_j \rangle = 0 \quad \forall j \quad (6.17)$$

The magnetization is always zero  $\forall j \neq \infty$ !

### Case with $H \neq 0$ and periodic boundary conditions

Consider the spins sitting on a 1D lattice ring as in Figure 6.3. The periodic boundary conditions are:

$$S_{N+1} = S_1 \quad (6.18)$$

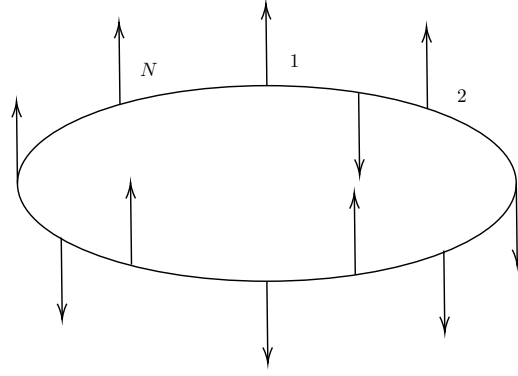
We have:

$$-\beta \mathcal{H}_\Omega(\{S\}) = K \sum_{i=1}^N S_i S_{i+1} + h \sum_{i=1}^N S_i \quad (6.19)$$

where

$$K \equiv \beta J, \quad h \equiv \beta H \quad (6.20)$$

The 1 – dimensional Ising model with this setup can be solved in several ways. Here we will use the method of the transfer matrix. This is a quite general technique that we will discuss within the Ising model.



**Figure 6.3:** One dimensional lattice ring.

### 6.1.2 Transfer Matrix method

Given the Hamiltonian discussed in (6.1.1) <sup>1</sup> we can write the corresponding partition function in the following symmetric form:

$$Z_N(k, h) = \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} \cdots \sum_{S_N=\pm 1} \left[ e^{KS_1 S_2 + \frac{h}{2}(S_1 + S_2)} \right] \left[ e^{KS_2 S_3 + \frac{h}{2}(S_2 + S_3)} \right] \cdots \left[ e^{KS_N S_1 + \frac{h}{2}(S_N + S_1)} \right] \quad (6.21)$$

We want to write the partition function in a form similarly to  $\sum_j M_{ij} P_{jk}$ . Note that, in the previous form  $Z_N$  can be written as a product of matrices

$$\begin{aligned} Z_N(h, k) &= \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \prod_{i=1}^N \exp \left[ KS_i S_{i+1} + \frac{h}{2}(S_i + S_{i+1}) \right] \\ &= \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \langle S_1 | \mathbb{T} | S_2 \rangle \langle S_2 | \mathbb{T} | S_3 \rangle \cdots \langle S_N | \mathbb{T} | S_1 \rangle \end{aligned} \quad (6.22)$$

where  $\mathbb{T}$  is a  $2 \times 2$  matrix defined as

$$\langle S | \mathbb{T} | S' \rangle = \exp \left[ KSS' + \frac{h}{2}(S + S') \right] \quad (6.23)$$

*Remark.* Note that the labels of the matrix corresponds to the values of  $S_i$ . Hence its dimension depends on the number of possible values a spin  $S_i$  can assume. It can also depends on how many spins are involved in the interacting terms that are present in the hamiltonian ( $k_{LL} \sum S_i S_{i+1} S_{i+2} S_{i+3}$ ).

For Ising  $S_i = \pm 1$  and nearest neighbour interaction implies that we have two values and that  $\mathbb{T}$  is a  $2 \times 2$  matrix whose components are

$$\langle +1 | \mathbb{T} | +1 \rangle = \exp[K + h] \quad (6.24a)$$

$$\langle +1 | \mathbb{T} | -1 \rangle = \langle -1 | \mathbb{T} | +1 \rangle = \exp[K - h] \quad (6.24b)$$

$$\langle -1 | \mathbb{T} | -1 \rangle = \exp[-K] \quad (6.24c)$$

The explicit representation is

$$\mathbb{T} = \begin{pmatrix} e^{K+h} & e^{-K} \\ e^{-K} & e^{K-h} \end{pmatrix} \quad (6.25)$$

<sup>1</sup>The choice of boundary conditions becomes irrelevant in the thermodynamic limit,  $N \rightarrow \infty$ .



Let us introduce some useful notations and relations using the bra-ket formalism:

$$|S_i^{(+)}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_i \quad |S_i^{(-)}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_i \quad (6.26a)$$

$$\langle S_i^{(+)}| = (1^*, 0)_i \quad \langle S_i^{(-)}| = (0, 1^*)_i \quad (6.26b)$$

The identity relation is:

$$\sum_{S_i=\pm 1} |S_i\rangle \langle S_i| = |S_i^{(+)}\rangle \langle S_i^{(+)}| + |S_i^{(-)}\rangle \langle S_i^{(-)}| = \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (6.27)$$

By using the identity property we can write

$$\begin{aligned} Z_N(K, h) &= \sum_{S_1=\pm 1} \cdots \sum_{S_N=\pm 1} \langle S_1 | \mathbb{T} | S_2 \rangle \langle S_2 | \mathbb{T} | S_3 \rangle \cdots \langle S_N | \mathbb{T} | S_1 \rangle \cdots \\ &= \sum_{S_1=\pm 1} \langle S_1 | \mathbb{T}^N | S_1 \rangle = \text{Tr}[\mathbb{T}^N] \end{aligned} \quad (6.28)$$

this is exactly the trace of the matrix, which is most usefully expressed in terms of the eigenvalues. Being  $\mathbb{T}$  symmetric, we can diagonalize it by an unitary transformation

$$\mathbb{T}_D = \mathbb{P}^{-1} \mathbb{T} \mathbb{P} \quad (6.29)$$

with  $\mathbb{P} \mathbb{P}^{-1} = \mathbb{1}$ . Hence,

$$\begin{aligned} \text{Tr}[\mathbb{T}^N] &= \text{Tr} \left[ \underbrace{\mathbb{T} \mathbb{T} \mathbb{T} \cdots \mathbb{T}}_N \right] = \text{Tr}[\mathbb{P} \mathbb{P}^{-1} \mathbb{T} \mathbb{P} \mathbb{P}^{-1} \mathbb{T} \mathbb{P} \mathbb{P}^{-1} \cdots \mathbb{P}^{-1} \mathbb{T} \mathbb{P} \mathbb{P}^{-1}] \\ &= \text{Tr}[\mathbb{P} \mathbb{T}_D^N \mathbb{P}^{-1}] \stackrel{\text{cyclic property of the trace}}{=} \text{Tr}[\mathbb{T}_D^N \mathbb{P}^{-1} \mathbb{P}] \\ &= \text{Tr}[\mathbb{T}_D^N] \end{aligned} \quad (6.30)$$

where

$$\mathbb{T}_D = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \Rightarrow \mathbb{T}_D^N = \begin{pmatrix} \lambda_+^N & 0 \\ 0 & \lambda_-^N \end{pmatrix} \quad (6.31)$$

with  $\lambda_{\pm}$  are the eigenvalues with  $\lambda_+ > \lambda_-$ .

*Remark.*  $\mathbb{P}$  is the similitude matrix whose columns are given by the eigenvectors of  $\lambda_{\pm}$ .

We finally have:

$$Z_N(K, h) = \text{Tr}[\mathbb{T}_D^N] = \lambda_+^N + \lambda_-^N \quad (6.32)$$

*Remark.* As mentioned previously the dimension of the transfer matrix  $\mathbb{T}$  and hence the number of eigenvalues  $\{\lambda\}$  depend both on the possible values of  $S_i$  and on the number of sites involved in terms of the Hamiltonian (range of interaction).

**Example 14.** For example consider the Ising ( $S_i = \pm 1$ ) with nearest neighbour and next nearest neighbour interactions. The hamiltonian is:

$$\mathcal{H} = k_1 \sum_i S_i S_{i+1} + k_2 \sum_i S_i S_{i+1} S_{i+2} S_{i+3} \quad (6.33)$$

Because of the second term now there are  $2^4 = 16$  possible configurations that can be described by using a  $4 \times 4$  transfer matrix that we can write formally as

$$\langle S_i S_{i+1} | \mathbb{T} | S_{i+2} S_{i+3} \rangle$$

**Example 15.** For example, suppose  $S_i = +1, 0, -1$ , therefore the spin can assume three different values. This is a *deluted* ising model.

Let us now consider the transfer matrix formalism in a more general setting.

## 6.2 General transfer matrix method

The aim of this section is to describe how transfer matrices can be used to solve classical spin models. The idea is to write down the partition function in terms of a matrix, the transfer matrix. The thermodynamic properties of the model are then wholly described by the eigenspectrum of the matrix. In particular, the free energy per spin in the thermodynamic limit depends only on the largest eigenvalue and the correlation length only on the two largest eigenvalues through simple formulae.

Let  $\mathbb{T}$  be a square matrix  $(n+2) \times (n+2)$  that, for example, it is built if the spin variables may assume  $(n+2)$  possible values. The  $k$ -esim value can be defined by the bra-ket notation where the two vectors are given by a sequence of "0" and a single "1" at the  $k$ -esim position.

**Example 16.** If  $k = 3$  and there are  $(n+2)$  possible values:

$$\langle S_i^{(3)} | = (0, 0, 1^*, 0, \dots, 0) \quad | S_i^{(3)} \rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

these are the bra-ket at the  $k$ -esim position.

Similarly to the  $2 \times 2$  Ising case, it is easy to show the identity property

$$\sum_{S_i} |S_i\rangle \langle S_i| = \mathbb{1}, \quad \mathbb{1} \in (n+2) \times (n+2) \quad (6.34)$$

where now the sum is over  $(n+2)$  values.

Let us consider the *diagonal matrix*  $\mathbb{S}_i$  where the elements along the diagonal are all the  $(n+2)$  possible values of the  $i$ -esim spin (or of some of their combination if longer interaction terms are considered)

$$\mathbb{S}_i \equiv \sum_{S_i} |S_i\rangle S_i \langle S_i| \quad (6.35)$$

**Example 17.** Ising model  $n+2 = 2$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} S^{(1)}(1^*, 0) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} S^{(2)}(0, 1^*) = \begin{pmatrix} S^{(1)} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & S^{(2)} \end{pmatrix} = \begin{pmatrix} S^{(1)} & 0 \\ 0 & S^{(2)} \end{pmatrix}$$

Ising:  $S^{(1)} = +1, S^{(2)} = -1$ .

*Remark.* Note that in this case the matrix  $\mathbb{S}_i$  is equal to the Pauli matrix  $\sigma_z$ .

*Remark.* By construction  $\langle S_i |$  and  $| S_i \rangle$  are the eigenvectors related to the eigenvalues  $S_i = S^{(1)}, S^{(2)}, \dots, S^{(n+2)}$ .

Similarly let  $\langle t_i |$  and  $| t_i \rangle$  be the eigenvectors related to the  $(n+2)$  eigenvalues of the transfer matrix  $\mathbb{T}$ :  $\{\lambda_+, \lambda_-, \lambda_1, \dots, \lambda_n\}$ , with  $\lambda_+ > \lambda_- \geq \lambda_1 \geq \dots \geq \lambda_n$ .

Clearly

$$\mathbb{T} = \mathbb{P} \mathbb{T}_D \mathbb{P}^{-1} = \sum_i |t_i\rangle \lambda_i \langle t_i| \quad (6.36)$$

Indeed,

$$\mathbb{T} |t_j\rangle = \sum_i |t_i\rangle \lambda_i \langle t_i | t_j \rangle = \sum_i |t_i\rangle \lambda_i \delta_{ij} = \lambda_j |t_j\rangle \quad (6.37)$$

Given the set of  $\lambda$  described above, the  $N$  particle partition function is given by

$$Z_N(\{K\}, h) = \lambda_+^N + \lambda_-^N + \sum_{i=1}^n \lambda_i^N \quad (6.38)$$

### 6.2.1 The free energy

Now, we are interested in the limit of the bulk free energy

$$F_N() = -k_B T \log Z_N() \quad (6.39)$$

In general, looking at the thermodynamic limit  $N \rightarrow \infty$ , by factorizing  $\lambda_+$

$$f_b(\{k\}, h) = \lim_{N \rightarrow \infty} \frac{1}{N} F_N = \lim_{N \rightarrow \infty} \frac{1}{N} (-k_B T) \log \left[ \lambda_+^N + \lambda_-^N + \sum_{i=1}^n \lambda_i^N \right] \quad (6.40)$$

by rearranging,

$$f_b = \lim_{N \rightarrow \infty} \frac{-k_B T}{N} \log \left[ \lambda_+^N \left( 1 + \frac{\lambda_-^N}{\lambda_+^N} + \sum_{i=1}^n \left( \frac{\lambda_i}{\lambda_+} \right)^N \right) \right] \quad (6.41)$$

Since  $\lambda_+ > \lambda_- > \lambda_1 > \dots \lambda_n$ ,

$$\left( \frac{\lambda_-}{\lambda_+} \right)^N \xrightarrow{N \rightarrow \infty} 0, \quad \left( \frac{\lambda_i}{\lambda_+} \right)^N \xrightarrow{N \rightarrow \infty} 0 \quad \forall i \quad (6.42)$$

we obtain:

$$f_b = -k_B T \log \lambda_+ \quad (6.43)$$

The limiting free-energy depends only on the largest eigenvalue of the transfer matrix  $\mathbb{T}$ ! This is important since sometimes it is much simpler to compute only the largest eigenvalue than the whole spectrum of  $\mathbb{T}$ . Also an important theorem about  $\lambda_+$  exists.

**Theorem 6.2.1** (Perron-Frobenius). *Let  $\mathbb{A}$  be a  $n \times n$  matrix. If  $\mathbb{A}$  is finite ( $n < \infty$ ) and  $\mathbb{A}_{ij} > 0, \forall i, j$ , ( $\mathbb{A}_{ij} = \mathbb{A}_{ij}(\vec{x})$ ), therefore its largest eigenvalue  $\lambda_+$  has the following properties:*

1.  $\lambda_+ \in \mathbb{R}^+$
2.  $\lambda_+ \neq$  from  $\{\lambda_i\}_{i=1, \dots, n-1}$ . It means there is no degeneracy.
3.  $\lambda_+$  is a analytic function of the parameters of  $\mathbb{A}$ .

*Remark.* Since in our case  $\mathbb{A} \leftrightarrow \mathbb{T}$ ,  $\lambda_+$  is related to  $f_b$  from the theorem. This means that  $f_b$  is an analytic function!

*Remark.* This is true for  $T > 0$  since for  $T = 0$  some  $\mathbb{T}_{ij}$  can be either 0 or  $\infty$  violating the hypothesis of the theorem.

*Remark.* If  $\mathbb{T}$  has infinite dimension (see  $d > 1$ ) the hypothesis of the theorem are not valid any more and  $f_b$  can be non-analytic.

If the conditions of the Perron-Frobenius theorem are satisfied by  $\mathbb{T}$ , the model described by  $\mathbb{T}$  cannot display a phase transition!

### 6.2.2 The correlation function

A second important quantity which is simply related to the eigenvalues of the transfer matrix is the correlation length. To calculate this we need the spin-spin correlation function which serves as an example of how to obtain averages of products of spins using transfer matrices.

Let us consider the two point correlation between two spins at distance  $R$  to another. The fluctuation respect to the average is:

$$\Gamma_R \equiv \langle S_1 S_R \rangle - \langle S_1 \rangle \langle S_R \rangle \quad (6.44)$$

Since

$$\Gamma_R \underset{R \rightarrow \infty}{\sim} \exp[-R/\xi] \quad (6.45)$$

we can define the correlation length  $\xi$  as

$$\xi^{-1} \equiv \lim_{R \rightarrow \infty} \left[ -\frac{1}{R} \log |\langle S_1 S_R \rangle - \langle S_1 \rangle \langle S_R \rangle| \right] \quad (6.46)$$

Now, let us compute the terms  $\langle S_1 S_R \rangle_N$  and  $\langle S_1 \rangle_N \langle S_R \rangle_N$ . From the definition of average we obtain

$$\langle S_1 S_R \rangle_N = \frac{1}{Z_N} \sum_{\{S\}} S_1 S_R \exp[-\beta \mathcal{H}_N] \quad (6.47)$$

*Remark.* The subscript  $N$  denotes that we are again considering a ring of  $N$  spins.  $Z_N$  is known from equation (6.38).

Writing this expression by using the transfer matrix formalism one obtains

$$\langle S_1 S_R \rangle_N = \frac{1}{Z_N} \sum_{\{S\}} S_1 \langle S_1 | \mathbb{T} | S_2 \rangle \dots \langle S_{R-1} | \mathbb{T} | S_R \rangle S_R \langle S_R | \mathbb{T} | S_{R+1} \rangle \dots \langle S_N | \mathbb{T} | S_1 \rangle \quad (6.48)$$

Summing over the free spins,

$$\langle S_1 S_R \rangle_N = \frac{1}{Z_N} \sum_{S_1, S_R} S_1 \langle S_1 | \mathbb{T}^{R-1} | S_R \rangle S_R \langle S_R | \mathbb{T}^{N-R+1} | S_1 \rangle \quad (6.49)$$

On the other hand, since

$$\mathbb{T} = \sum_{i=1}^{n+2} |t_i\rangle \lambda_i \langle t_i| \quad (6.50)$$

we have

$$\mathbb{T}^{R-1} = \sum_{i=1}^{n+2} |t_i\rangle \lambda_i^{R-1} \langle t_i| \quad (6.51a)$$

$$\mathbb{T}^{N-R+1} = \sum_{i=1}^{n+2} |t_i\rangle \lambda_i^{N-R+1} \langle t_i| \quad (6.51b)$$

Hence,

$$\langle S_1 | \mathbb{T}^{R-1} | S_R \rangle = \sum_{i=1}^{n+2} \langle S_1 | t_i \rangle \lambda_i^{R-1} \langle t_i | S_R \rangle \quad (6.52)$$

and plugging this expression in (6.49) one gets

$$\sum_{\{S\}} S_1 S_R e^{-\beta \mathcal{H}_N} = \sum_{S_1 S_R} S_1 \sum_{i=1}^{n+2} \langle S_1 | t_i \rangle \lambda_i^{R-1} \langle t_i | S_R \rangle S_R \sum_{j=1}^{n+2} \langle S_R | t_j \rangle \lambda_j^{N-R+1} \langle t_j | S_1 \rangle \quad (6.53)$$

Since the term  $\langle t_j | S_1 \rangle$  is a scalar it can be moved at the beginning of the product. Remembering the notations

$$\mathbb{S}_1 = \sum_{S_1} |S_1\rangle S_1 \langle S_1| \quad (6.54a)$$

$$\mathbb{S}_R = \sum_{S_R} |S_R\rangle S_R \langle S_R| \quad (6.54b)$$

one gets

$$\sum_{\{S\}} S_1 S_R e^{-\beta \mathcal{H}_N} = \sum_{ij} \langle t_j | \mathbb{S}_1 | t_i \rangle \lambda_i^{R-1} \langle t_i | \mathbb{S}_R | t_j \rangle \lambda_j^{N-R+1} \quad (6.55)$$

Since  $\sum_k \lambda_k^N = Z_N$  for  $k = +, -, 1, \dots, n$ :

$$\langle S_1 S_R \rangle_N = \frac{\sum_{ij} \langle t_j | \mathbb{S}_1 | t_i \rangle \lambda_i^{R-1} \langle t_i | \mathbb{S}_R | t_j \rangle \lambda_j^{N-R+1}}{\sum_{k=1}^n \lambda_k^N} \quad (6.56)$$

If we now multiply and divide by  $\lambda_+^N$ , we get

$$\langle S_1 S_R \rangle_N = \frac{\sum_{ij} \langle t_j | \mathbb{S}_1 | t_i \rangle (\lambda_i / \lambda_+)^{R-1} \langle t_i | \mathbb{S}_R | t_j \rangle (\lambda_j / \lambda_+)^{N-R+1}}{\sum_{k=1}^n (\lambda_k / \lambda_+)^N} \quad (6.57)$$

*Remark.* In the thermodynamic limit  $N \rightarrow \infty$ , only the terms with  $j = +$  and  $k = +$  survive in the sum. Remind that  $R$  is fixed.

$$\langle S_1 S_R \rangle_N = \lim_{N \rightarrow \infty} \langle S_1 S_R \rangle_N = \sum_{i=\pm, 1, \dots, n} \left( \frac{\lambda_i}{\lambda_+} \right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_i \rangle \langle t_i | \mathbb{S}_R | t_+ \rangle \quad (6.58)$$

Remember that  $\lambda_+ > \lambda_T \geq \lambda_1 \dots \lambda_n$ :

$$\langle S_1 S_R \rangle_N = \langle t_+ | \mathbb{S}_1 | t_+ \rangle \langle t_+ | \mathbb{S}_R | t_+ \rangle + \sum_{i \neq +}^n \left( \frac{\lambda_i}{\lambda_+} \right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_i \rangle \langle t_i | \mathbb{S}_R | t_+ \rangle \quad (6.59)$$

Since one can prove, by a method entirely analogous to that followed above, that

$$\lim_{N \rightarrow \infty} \langle S_R \rangle_N = \langle t_+ | \mathbb{S}_R | t_+ \rangle \quad (6.60)$$

we obtain

$$\langle S_1 S_R \rangle = \langle S_1 \rangle \langle S_R \rangle + \sum_{i \neq +} \left( \frac{\lambda_i}{\lambda_+} \right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_i \rangle \langle t_i | \mathbb{S}_R | t_+ \rangle \quad (6.61)$$

The correlation function then follows immediately as

$$\Gamma_R = \langle S_1 S_R \rangle - \langle S_1 \rangle \langle S_R \rangle = \sum_{i \neq +}^n \left( \frac{\lambda_i}{\lambda_+} \right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_i \rangle \langle t_i | \mathbb{S}_R | t_+ \rangle \quad (6.62)$$

*Remark.*  $\Gamma_R$  depends only on the eigenvalues and eigenvectors of the transfer matrix  $\mathbb{T}$  and by the values of the spins  $S_1$  and  $S_R$ .

A much simpler formula is obtained for the correlation length (6.46). Taking the limit  $R \rightarrow \infty$  the ratio  $(\lambda_- / \lambda_+)$  dominates the sum and hence

$$\begin{aligned} \xi^{-1} &= \lim_{R \rightarrow \infty} \left\{ -\frac{1}{R-1} \log |\langle S_1 S_R \rangle - \langle S_1 \rangle \langle S_R \rangle| \right\} \\ &= \lim_{R \rightarrow \infty} \left\{ -\frac{1}{R-1} \log \left[ \left( \frac{\lambda_-}{\lambda_+} \right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_- \rangle \langle t_- | \mathbb{S}_R | t_+ \rangle \right] \right\} \\ &= -\log \left[ \left( \frac{\lambda_-}{\lambda_+} \right) \right] - \lim_{R \rightarrow \infty} \frac{1}{R-1} \log \langle t_+ | \mathbb{S}_1 | t_- \rangle \langle t_+ | \mathbb{S}_R | t_+ \rangle \\ &= -\log \left( \frac{\lambda_-}{\lambda_+} \right) \end{aligned} \quad (6.63)$$

The important result is

$$\xi^{-1} = -\log \left( \frac{\lambda_-}{\lambda_+} \right) \quad (6.64)$$

It means that the correlation length does depend only on the ratio between the two largest eigenvalues of the transfer matrix  $\mathbb{T}$ .

**Lecture 8.**

Wednesday 6<sup>th</sup>

November, 2019.

Compiled:

Wednesday 11<sup>th</sup>

December, 2019.

### 6.2.3 Results for the 1-dimensional Ising model

The transfer matrix is given by

$$\mathbb{T} = \begin{pmatrix} \exp(K+h) & \exp(-K) \\ \exp(-K) & \exp(K-h) \end{pmatrix} \quad (6.65)$$

Calculate the eigenvalues:

$$|\mathbb{T} - \lambda \mathbb{1}| = (e^{K+h} - \lambda)(e^{K-h} - \lambda) - e^{-2K} = 0 \quad (6.66)$$

The solutions are

$$\lambda_{\pm} = e^K \cosh(h) \pm \sqrt{e^{2K} \sinh^2(h) + e^{-2K}} \quad (6.67)$$

#### The free energy

The free energy is

$$\begin{aligned} f_b &\equiv \lim_{N \rightarrow \infty} \frac{-k_B T}{N} \log Z_N(K, h) \\ &= -k_B T \lim_{N \rightarrow \infty} \frac{1}{N} \log \left[ \lambda_+^N \left( 1 + \left( \frac{\lambda_-}{\lambda_+} \right)^N \right) \right] \\ &= -k_B T \log \lambda_+ \end{aligned} \quad (6.68)$$

and inserting the explicit expression of  $\lambda_+$  for the Ising model, we get

$$\begin{aligned} f_b &= -k_B T \log \left( e^K \cosh h + \sqrt{e^{2K} \sinh^2(h) + e^{-2K}} \right) \\ &= -K k_B T - k_B T \log \left( \cosh(h) + \sqrt{\sinh^2(h) + e^{-4K}} \right) \end{aligned} \quad (6.69)$$

*Remark.* Remember that  $K \equiv \beta J, h \equiv \beta H$ .

**Exercise 2.** Check that if  $h = 0$  we get back the expression found previously with the iterative method (what is the important of boundary conditions?).

Let us now consider the limits  $T \rightarrow 0$  and  $T \rightarrow \infty$  by keeping  $H$  fixed and  $J$  fixed.

- Case:  $T \rightarrow 0 \Rightarrow K \rightarrow \infty, h \rightarrow \infty$ .

$$e^{-4K} \xrightarrow{K \rightarrow \infty} 0 \quad (6.70a)$$

$$\sqrt{\sinh^2 h} \xrightarrow{h \rightarrow \infty} \sinh(h) \quad (6.70b)$$

This implies that

$$\cosh(h) + \sinh h \sim \frac{2e^h}{2} \simeq e^h \quad (6.71)$$

and

$$f \xrightarrow{K \rightarrow \infty} -K k_B T - k_B T \log e^h \sim -J - H \quad \text{const} \quad (6.72)$$

Therefore, as  $T \rightarrow 0^+$ ,  $f$  goes to a constant that depends on  $J$  and  $H$ .

- Case:  $T \rightarrow \infty \Rightarrow K \rightarrow 0, h \rightarrow 0$ . In this case we suppose also that  $H$  and  $J$  that are fixed, are also finite.

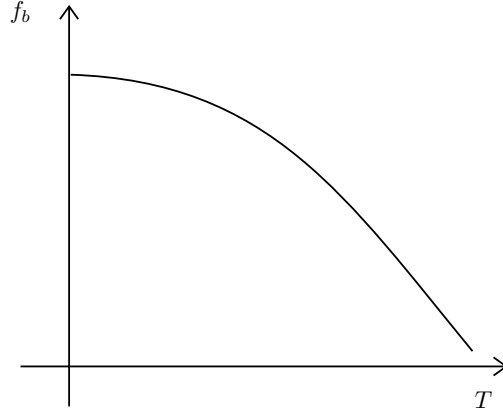
$$e^{-4K} \simeq 1 \quad (6.73a)$$

$$\sqrt{\sinh^2 h + e^{-4K}} \sim \sqrt{1} \quad (6.73b)$$

Since  $\cosh h \xrightarrow{h \rightarrow 0} 1$ :

$$f_B \sim -K k_B T - k_B T \log(1+1) \sim -J - k_B T \ln 2 \quad (6.74)$$

Therefore, as  $T \rightarrow \infty$ , the free energy goes linearly to zero, as in figure 6.4.



**Figure 6.4:** Plot of the free energy  $f_b$  in function of the temperature  $T$ . For  $T \rightarrow 0$ , the free energy becomes constant, while for  $T \rightarrow \infty$  it goes linearly to zero.

### The magnetization

This can be obtained by differentiating the negative of the free energy with respect to the magnetic field  $H$ :

$$m = -\frac{\partial f_b}{\partial H} = -\frac{1}{k_B T} \frac{\partial f_b}{\partial h} = \frac{\partial}{\partial h} \left[ \log \left( \cosh(h) + \sqrt{\sinh^2(h) + e^{-4K}} \right) \right] \quad (6.75)$$

The result is

$$m = \frac{\sinh h + \frac{\sinh h \cosh h}{\sqrt{\sinh^2 h + e^{-4K}}}}{\cosh h + \sqrt{\sinh^2 h + e^{-4K}}} \quad (6.76)$$

- Case:  $T > 0$  fixed,  $H \rightarrow 0 \Leftrightarrow h \rightarrow 0$ .

$$\sinh h \sim h \sim 0, \quad (6.77a)$$

$$\cosh h \sim 1 \quad (6.77b)$$

In zero field  $h \rightarrow 0$ , we have  $m \rightarrow 0$  for all  $T > 0$ . It means that there is no spontaneous magnetization!

### The magnetic susceptibility

$$\chi_T \equiv \frac{\partial m}{\partial H} = \frac{1}{k_B T} \frac{\partial m}{\partial h} \quad (6.78)$$

If we consider the case  $h \ll 1$ , it is convenient first expand the (6.76) for  $h \rightarrow 0$  and take the derivative to get  $\chi_T$ .

Since  $\sinh(h) \sim h + h^3$  and  $\cosh(h) \sim 1 + h^2$ , we have

$$m \underset{h \ll 1}{\sim} \frac{h(1 + e^{2K})}{1 + e^{-2K}} \quad (6.79)$$

If we now derive with respect to  $h$

$$\chi_T = \frac{1}{k_B T} \frac{\partial m}{\partial h} \underset{h \ll 1}{\approx} \frac{1}{k_B T} \frac{(1 + e^{2K})}{(1 + e^{-2K})} \quad (6.80)$$

- Case:  $T \rightarrow \infty \Leftrightarrow K \rightarrow 0$ .

$$e^{2K} \simeq e^{-2K} \simeq 1 \quad (6.81)$$

The *Curie's Law* for paramagnetic systems is:

$$\chi_T \sim \frac{1}{k_B T} \quad (6.82)$$

- Case:  $T \rightarrow 0 \Leftrightarrow K \rightarrow \infty$ .

$$e^{-2K} \simeq 0 \quad (6.83)$$

The *Curie's Law* for paramagnetic systems is:

$$\chi_T \sim \frac{1}{k_B T} e^{2K} \sim \frac{1}{k_B T} e^{2J/k_B T} \quad (6.84)$$

### The correlation length

$$\xi^{-1} = -\log \left( \frac{\lambda_-}{\lambda_+} \right) = -\log \left[ \frac{\cosh h - \sqrt{\sinh^2 h + e^{-4K}}}{\cosh h + \sqrt{\sinh^2 h + e^{-4K}}} \right] \quad (6.85)$$

For  $h = 0$ , we have  $\cosh h \rightarrow 1, \sinh h \rightarrow 0$ :

$$\xi^{-1} = -\log \left[ \frac{1 - e^{-2K}}{1 + e^{-2K}} \right] = -\log \left[ \frac{1}{\coth K} \right] \quad (6.86)$$

Therefore:

$$\xi = \frac{1}{\log(\coth K)} \quad (6.87)$$

- Case:  $T \rightarrow 0 \Leftrightarrow K \rightarrow \infty$ .

$$\coth K = \frac{e^K + e^{-K}}{e^K - e^{-K}} \xrightarrow{K \rightarrow \infty} 1 + 2e^{-2K} + \dots \xrightarrow{K \rightarrow \infty} 1 \quad (6.88)$$

It implies

$$\xi \xrightarrow{K \gg 1} \frac{1}{\ln(1 + 2e^{-2K})} \sim \frac{e^{2K}}{2} \quad (6.89)$$

Hence

$$\xi \xrightarrow{T \rightarrow 0} \frac{1}{2} e^{J/k_B T} \quad (6.90)$$

It diverges exponentially  $\xi \rightarrow \infty$ , as  $T \rightarrow 0$ .

- Case:  $T \rightarrow \infty \Leftrightarrow K \rightarrow 0$ .

$$\coth K = \frac{e^K + e^{-K}}{e^K - e^{-K}} \xrightarrow{K \rightarrow 0} \frac{1 + K + \frac{K^2}{2} + 1 - K + \frac{K^2}{2}}{1 + K + \frac{K^2}{2} - 1 + K - \frac{K^2}{2}} \sim \frac{2 + 2\frac{K^2}{2}}{2K} \sim \frac{1 + K^2}{K} \quad (6.91)$$

$$\xi^{-1} = \log(\coth K) \xrightarrow{K \rightarrow 0} \ln \frac{1}{K} + \ln(1 + K^2) \sim +\infty \quad (6.92)$$

Therefore:

$$\xi \xrightarrow{K \rightarrow 0} 0 \quad (6.93)$$

More precisely,

$$\xi \xrightarrow{K \rightarrow 0} \frac{1}{\ln(1/K) + \ln(1 + K^2)} \xrightarrow{K \rightarrow 0} -\frac{1}{\ln K} \quad (6.94)$$



### 6.3 Classical Heisenberg model for $d=1$

Suppose to study something different from the Ising model, we do not anymore assume spin that can assume values as  $-1$  or  $+1$ , but spin that can assume a continuous value. This is the classical Heisenberg model.

Take a  $d = 1$  dimensional lattice. In the classical Heisenberg model the spins are unit length vectors  $\vec{\mathbf{S}}_i$ , i.e.  $\vec{\mathbf{S}}_i \in \mathbb{R}^3$ ,  $|\vec{\mathbf{S}}_i|^2 = 1$  (continuous values on the unit sphere):

$$\vec{\mathbf{S}}_i = (S_i^x, S_i^y, S_i^z) \quad (6.95)$$

with periodic boundary condition:  $\vec{\mathbf{S}}_{N+1} = \vec{\mathbf{S}}_1$ .

Assuming  $H = 0$ , the model is defined through the following Hamiltonian::

$$-\beta\mathcal{H}(\{\vec{\mathbf{S}}\}) = K \sum_{i=1}^N \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_{i+1} \quad (\longrightarrow \sum_i \vec{\mathbf{h}} \cdot \vec{\mathbf{S}}_i) \quad (6.96)$$

This model satisfies  $O(3)$  symmetry. In the transfer matrix formalism:

$$Z_N(K) = \sum_{\{\vec{\mathbf{S}}\}} e^{-\beta\mathcal{H}} = \sum_{\{\vec{\mathbf{S}}\}} e^{K \sum_{i=1}^N \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_{i+1}} = \text{Tr}(\mathbb{T}^N) \quad (6.97)$$

where  $\langle \vec{\mathbf{S}}_i | \mathbb{T} | \vec{\mathbf{S}}_{i+1} \rangle = e^{K \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_{i+1}}$ .

Similarly to the Ising case:

$$\mathbb{T} = \sum_i |t_i\rangle \lambda_i \langle t_i| \quad (6.98)$$

and

$$\mathbb{T}_D = \mathbb{P}^{-1} \mathbb{T} \mathbb{P} \quad (6.99)$$

The problem is computing the eigenvalues  $\lambda_i$  of  $\mathbb{T}$ . Formally, we should find

$$\exp[K \vec{\mathbf{S}}_1 \cdot \vec{\mathbf{S}}_2] = \langle \vec{\mathbf{S}}_1 | \mathbb{T} | \vec{\mathbf{S}}_2 \rangle = \sum_{i \in \text{eigenvalues}} \lambda_i \langle \vec{\mathbf{S}}_1 | t_i \rangle \langle t_i | \vec{\mathbf{S}}_2 \rangle = \sum_i \lambda_i f_i(\vec{\mathbf{S}}_1) f_i^*(\vec{\mathbf{S}}_2) \quad (6.100)$$

*Remark.* We start by noticing that the term  $e^{K \vec{\mathbf{S}}_1 \cdot \vec{\mathbf{S}}_2}$  is similar to the plane wave  $e^{i \vec{\mathbf{q}} \cdot \vec{\mathbf{r}}}$ , that in scattering problems is usually expanded in spherical coordinates. Plane wave can be expanded as a sum of spherical harmonics

$$e^{i \vec{\mathbf{q}} \cdot \vec{\mathbf{r}}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l (i)^l j_l(qr) Y_{lm}^*(\hat{\mathbf{q}}) Y_{lm}(\hat{\mathbf{r}}) \quad (6.101)$$

where

$$j_l(qr) = -\frac{(i)^l}{2} \int_0^\pi \sin(\theta) e^{iqr \cos(\theta)} P_l(\cos(\theta)) d\theta \quad (6.102)$$

are the *spherical Bessel functions*, while the  $P_l(\cos(\theta))$  are the *Legendre polynomial* of order  $l$ .

From a formal comparison we have

$$\vec{\mathbf{S}}_1 \leftrightarrow \hat{\mathbf{S}}_1, \quad \begin{cases} i \vec{\mathbf{q}} \cdot \vec{\mathbf{r}} = iqr \\ K \vec{\mathbf{S}}_1 \cdot \vec{\mathbf{S}}_2 = K |\vec{\mathbf{S}}_1| |\vec{\mathbf{S}}_2| = K \end{cases} \quad (6.103)$$

multiplying by  $(-i)$  we can write

$$qr = -iK \left| \vec{\mathbf{S}}_1 \right| \left| \vec{\mathbf{S}}_2 \right| = -iK \quad (6.104)$$

In our case we have  $\hat{\mathbf{q}} = \vec{\mathbf{S}}_1, \hat{\mathbf{r}} = \vec{\mathbf{S}}_2$ . Hence,

$$e^{K\vec{\mathbf{S}}_1 \cdot \vec{\mathbf{S}}_2} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l (i)^l j_l(-iK) Y_{lm}^*(\vec{\mathbf{S}}_1) Y_{lm}(\vec{\mathbf{S}}_2) = \sum_i \lambda_i f_i(\vec{\mathbf{S}}_1) f_i^*(\vec{\mathbf{S}}_2) \quad (6.105)$$

where

$$\lambda_i = \lambda_{lm}(K) = 4\pi (i)^l j_l(-iK) \quad (6.106)$$

*Remark.* Note that  $\lambda_i$  does not depend on  $m$ !

If  $l = 0$ , the largest eigenvalue is:

$$\lambda_+ = \lambda_0(K) = 4\pi j_0(-iK) = 4\pi \frac{\sin K}{K} \quad (6.107)$$

and

$$\lambda_- = \lambda_1(K) = 4\pi i j_1(-iK) = 4\pi \left[ \frac{\cosh K}{K} - \frac{\sinh K}{K^2} \right] \quad (6.108)$$

**Exercise 3.** Given the largest eigenvalue  $\lambda_+$ :

$$\lambda_+ = 4\pi \frac{\sin(K)}{K} \quad (6.109)$$

find the bulk free energy density of the model and discuss its behaviour in the limits of low ( $T \rightarrow 0$ ) and high ( $T \rightarrow \infty$ ) temperatures.

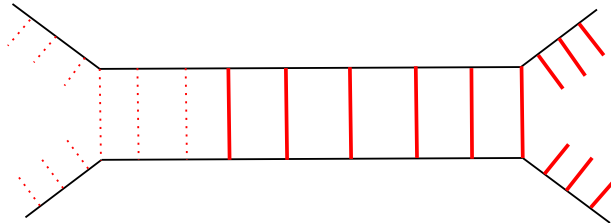
How can we violate the hypothesis of the Perron-Frobenius theorem hoping to find a phase transition also in a  $d = 1$  model? One of the hypothesis of the Perron-Frobenius theorem is the one in which  $A_{ij} > 0$  for all  $i, j$ . Hence, one possibility is to build a model in which its transfer matrix has same  $A_{ij}$  that are equal to zero also for  $T \neq 0$ .

## 6.4 Zipper model

**Lecture 9.**  
Friday 8<sup>th</sup>  
November, 2019.  
Compiled:  
Wednesday 11<sup>th</sup>  
December, 2019.

It is a model introduced by Kittel to describe oligomers undergoing *denaturation transition*. Simplest model of DNA thermal denaturation transition (no bubbles). Better model for the denaturation of short oligomers.

The hypothesis are: the binding energy between two bases located at the end of the molecule is smaller than the one for pairs away from the ends. The unbinding starts and develops from the ends as a *zipper*.



**Figure 6.5:** Sequential unzipping from the ends.

In this denaturation transition we do not allow bubbles. Let us consider first the zingle-ended zipper, i.e. a molecular zipper of  $N$  parallel links that can be opened only from one end as in Figure 6.6.

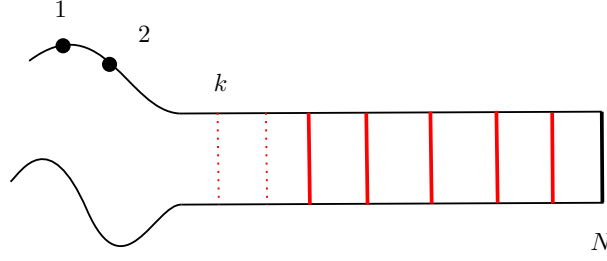


Figure 6.6

If the first  $k$  bonds are open (unbounded pairs) the energy to open the  $k+1$  is  $\varepsilon_0$ . Note that if at least one of the previous  $k$  bond is closed the energy needed to open the  $k+1$  band is infinite! Once a bond is open it can orient itself in  $G$  different ways. In other words there is an entropy

$$S_0 = k_B \log G \quad (6.110)$$

associated to each open band. If  $k$  bands ( $i = 1, \dots, k$ ) are open the associated weight is given by

$$G^k e^{-k\varepsilon_0/k_B T} \quad (6.111)$$

By summing over the possible values of  $k$ , the partition function is

$$Z_N(T, G, \varepsilon_0) = \sum_{k=0}^{N-1} G^k e^{-k\varepsilon_0/k_B T} = \sum_{k=0}^{N-1} e^{k(S_0 T - \varepsilon_0)/k_B T} \quad (6.112)$$

Let us call

$$\chi \equiv G e^{-\varepsilon_0/k_B T} \quad (6.113)$$

and simplify the previous expression

$$Z_N = \sum_{k=0}^{N-1} \chi^k = \frac{1 - \chi^N}{1 - \chi} \quad (6.114)$$

We see immediately there is a single pole singularity. The free energy is

$$F_N = -k_B T \ln Z_N = -k_B T \ln \left[ \frac{1 - \chi^N}{1 - \chi} \right] \quad (6.115)$$

We can now compute some observables of interest.

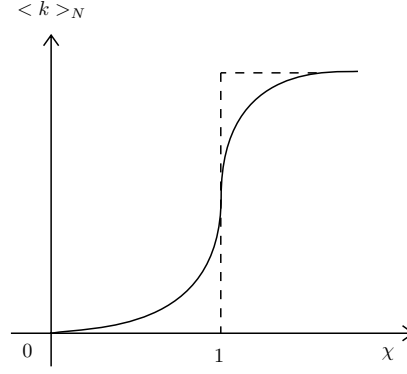
### Calculate average number of open links

The equal of our magnetization in this case is

$$\langle k \rangle_N \equiv \frac{\sum_{k=0}^{N-1} k \chi^k}{\sum_{k=0}^{N-1} \chi^k} = \chi \frac{d}{d\chi} \ln Z_N = \frac{N \chi^N}{\chi^N - 1} - \frac{\chi}{\chi - 1} \quad (6.116)$$

There is a pole in  $\chi = 1$  and the last quantity can be drawn as in Figure 6.7.

In order to analyze what happens near 1, we expand  $\chi \equiv 1 + \varepsilon$ :



**Figure 6.7:** Description.

$$\begin{aligned}
\log Z_N(\chi) &= \log \left[ \frac{1 - (1 + \varepsilon)^N}{1 - (1 + \varepsilon)} \right] \\
&= \log \left[ \frac{1 - (1 + \varepsilon N + \frac{N(N-1)}{2!}\varepsilon^2 + \frac{N(N-1)(N-2)}{3!}\varepsilon^3 + O(\varepsilon^4))}{\varepsilon} \right] \\
&= \log \left[ N + \frac{N(N-1)}{2}\varepsilon + \frac{N(N-1)(N-2)}{6}\varepsilon^2 + \dots \right] \\
&= \log N + \log \left[ 1 + \frac{N-1}{2}\varepsilon + \frac{N(N-1)}{6}\varepsilon^2 \right] \\
&= \log N + \log \left[ 1 + \frac{N\varepsilon}{2} + \frac{N^2\varepsilon^2}{6} + \dots \right] \\
&= \log N + \left( \frac{N\varepsilon}{2} + \frac{N^2\varepsilon^2}{6} + \dots \right) + \frac{1}{2} \left( \frac{N\varepsilon}{2} + \frac{N^2\varepsilon^2}{6} + \dots \right)^2 + \dots \\
&= \log N + \frac{N\varepsilon}{2} + \frac{N^2\varepsilon^2}{24} + \dots
\end{aligned} \tag{6.117}$$

By doing the same for  $\langle k \rangle_N = \frac{N\chi^N}{\chi^N - 1} - \frac{\chi}{\chi - 1}$ , one gets

$$\langle k \rangle_N = \frac{N}{2} \left( 1 + \frac{N\varepsilon}{6} - \frac{N^3\varepsilon^3}{360} + \dots \right) \tag{6.118}$$

this is true for  $N \gg 1, \varepsilon \ll 1$ . At the transition point  $\chi_C = 1$ , where  $\varepsilon = 0$ :

$$\langle k \rangle_N \simeq \frac{N}{2} \tag{6.119}$$

We can define the variation (slope per site) as a response function (the derivative with respect to the parameter):

$$\frac{1}{N} \frac{d \langle k \rangle}{d \varepsilon} \simeq \frac{N}{12} - \frac{N^3\varepsilon^3}{240} + \dots \tag{6.120}$$

It is max at  $\varepsilon = 0$  and diverges as  $N \rightarrow \infty$  (linearly). The response function diverges linearly to  $N$ . This is a good signal that we have a transition. The temperature  $T_c$  corresponding to the pole  $\chi = 1$  is given

$$Ge^{-\varepsilon_0/k_B T_c} = 1 \tag{6.121}$$

Hence

$$T_C = \frac{\varepsilon_0}{k_B \log G} \quad (6.122)$$

Note that as  $G \rightarrow 1$ ,  $T_C \rightarrow 0$ . For  $G = 1$  there is no solution and hence the model does not display a phase transition for any finite  $T$ ! This is telling you that if  $G = 1$  what is important it is the energy, you have no entropy as disorder. At that point everything can happen.

*Remark.* Despite the model is  $1D$ , for  $G > 1$  there is a phase transition. This is due to two contributions:

1. Existence of forbidden configuration (infinite energy). Necessary condition for a phase transition in  $d = 1$  with finite range interactions.
2. Degeneracy of the excited state ( $G$ ).

#### 6.4.1 Transfer matrix method for the Kittel model

The idea is: we want to map this model to an Ising model. The spin like model consists on associating to each bond a spin such that  $S_i = 0$  if the  $i$ -esim bond is *closed*, while  $S_i = 1, \dots, G$  if the  $i$ -esim bond is *open* with  $G$  possible orientations.

- Case:  $S_i \neq 0$  open. We have two subcases:
  - $S_{i-1}$  open:  $S_{i-1} \neq 0 \Rightarrow E(S_i \neq 0 | S_{i-1} \neq 0) = \varepsilon_0$ .
  - $S_{i-1}$  closed:  $S_{i-1} = 0 \Rightarrow E(S_i \neq 0 | S_{i-1} = 0) = \varepsilon_0 + V_0$
- Case:  $S_i = 0$  closed. We have  $E(S_i = 0) = 0$  irrespective of  $S_{i-1}$ .

Therefore

$$E(S_i, S_{i-1}) = (\varepsilon_0 + V_0 \delta_{S_{i-1}, 0})(1 - \delta_{S_i, 0}) \quad (6.123)$$

The boundary condition is  $S_N = 0$  (always closed). The full Hamiltonian of the model can be written as (it could be also a function of delta, but it is not a problem):

$$\mathcal{H}_N = \varepsilon_0(1 - \delta_{S_1, 0}) + \sum_{i=2}^{N-1} (\varepsilon_0 + V_0 \delta_{S_{i-1}, 0})(1 - \delta_{S_i, 0}) \quad (6.124)$$

The Kittel's version is obtained by assuming  $V_0 = \infty$ .

The partition function is

$$Z_N = \sum_{\{S\}} \exp(-\beta \mathcal{H}_N) \quad (6.125)$$

In order to implement the transfer matrix formalism we rewrite  $Z_N$  as follows (DUB-BIIOOOOOO per la prossima equation )

$$Z_N = \sum_{\{S\}} e^{-\beta \varepsilon_0(1 - \delta_{S_1, 0})} \prod_{i=1}^{N-2} e^{-\beta \varepsilon_0(1 - \delta_{S_{i+1}, 0})} \dots \left(1 + (e^{-\beta V_0} - 1) \delta_{S_i, 0}(1 - \delta_{S_{i+1}, 0})\right) \quad (6.126)$$

Consider the Kittel model,  $V_0 = \infty$  which implies  $\exp(-\beta V_0) = 0$ . We can define the transfer matrix as

$$\mathbb{T} = \{\langle S | \mathbb{T} | S' \rangle \equiv t_{S, S'}\} \quad (6.127)$$

where

$$t_{S, S'} = e^{-\beta \varepsilon_0(1 - \delta_{S', 0})} [1 - \delta_{S, 0}(1 - \delta_{S', 0})] \quad (6.128)$$

or in matrix form

$$\mathbb{T} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & a & \dots & a \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & a & \dots & a \end{bmatrix} \quad (6.129)$$

where  $a \equiv e^{-\beta\varepsilon_0}$ .

The first think to notice that the constraint that the bond  $S_{i+1}$  cannot be open if bond  $S_i$  is closed ( $S_i = 0$ ) yields the null entries in the first row of  $\mathbb{T}$ . This violates the hypothesis of the Perron-Frobenius theorem!

The matrix  $\mathbb{T}$  has three differen eigenvalues

$$\lambda_1 = Ga, \quad \lambda_2 = 1, \quad \lambda_3 = 0 \quad (6.130)$$

The partition function can be written as

$$Z_N = (1, a, \dots, a) \mathbb{T}^{N-2} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad (6.131)$$

$$\lambda_1 \rightarrow \vec{v}_1 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \lambda_2 \rightarrow \vec{v}_2 = \begin{pmatrix} 1 - Ga \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad (6.132)$$

We can then write

$$\begin{pmatrix} 1 \\ a \\ \vdots \\ a \end{pmatrix} = \frac{a(1 - Ga) - 1}{1 - Ga} \vec{v}_1 + \frac{1}{1 - Ga} \vec{v}_2 \quad (6.133a)$$

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \frac{-Ga}{1 - Ga} \vec{v}_1 + \frac{1}{1 - Ga} \vec{v}_2 \quad (6.133b)$$

Therefore

$$Z_N = \frac{1 - (Ga)^N}{1 - Ga} = \frac{1 - (Ge^{-\beta\varepsilon_0})^N}{1 - Ge^{-\beta\varepsilon_0}} \quad (6.134)$$

or

$$Z_N = \frac{1}{1 - Ge^{-\beta\varepsilon_0}} (-\lambda_1^N + \lambda_2^N) \quad (6.135)$$

Since in the thermodynamic limit only the contribution of the largest eigenvalue matters for  $f_b$  we have

$$f_b = -k_B T \ln \max(\lambda_1, \lambda_2) \quad (6.136)$$

*Remark.* Given that the  $\lambda_1$  and  $\lambda_2$  are positive, analytic function of  $T$  ( $\lambda_1 = Ga, \lambda_2 = 1$ ). In order to have a phase transition (i.e. non analyticity of  $f_b$ ) the two eigenvalues must cross for a given value of  $T$ . It is true if and only if:

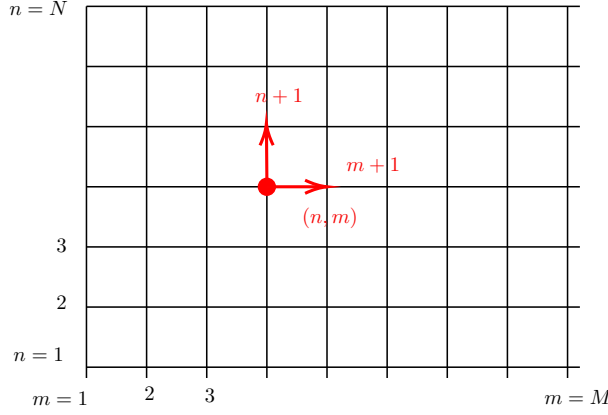
$$Ga_c = 1 \Leftrightarrow Ge^{-\beta_c \varepsilon_0} = 1 \Leftrightarrow T_c = \frac{\varepsilon_0}{k_B \ln G} \quad (6.137)$$

It is agree with previous calculation.

## 6.5 Transfer matrix for 2D Ising

Dimensional reduction from  $d$  to  $d-1$ . You can do the same for the 3D Ising model. In order to go smoothly from the one dimensional to the two dimensional the idea is to solve the Ising model on surfaces.

Consider a square lattice with  $N$  rows and  $M$  columns, as in Figure 6.8, with periodic boundary conditions (wrapped around a torus). The spin in a site is identified by  $S_{\text{site}} = S_{m,n}$ . The Hamiltonian is (if  $H \neq 0$  we have also the term with the  $h$ ):



**Figure 6.8:** Description.

$$\begin{aligned}
 -\beta\mathcal{H}_\Omega(\{S\}) &= k \sum_{\langle ij \rangle} S_i S_j + h \sum_i S_i \\
 &= k \sum_{n=1}^N \sum_{m=1}^M (S_{m,n} S_{m+1,n} + S_{m,n} S_{m,n+1}) + h \sum_{n=1}^N \sum_{m=1}^M S_{m,n}
 \end{aligned} \tag{6.138}$$

in fact

$$\sum_{\langle i,j \rangle} \rightarrow \sum_{i,j \in nn(i)} \Rightarrow S_{m,n} \rightarrow S_{m+1,n} \tag{6.139}$$

The Hamiltonian can be rewritten as follows:

$$-\beta\mathcal{H}_\Omega(\{S\}) = \sum_{m=1}^M [E[\mu_m, \mu_{m+1}] + E[\mu_m]] \tag{6.140}$$

where the first term is the interaction between columns (two body interaction), the second term is the one body interaction of one column. Moreover, the  $\mu$  is a  $m$  dimensional vector, in particular each  $\mu_m$  represents the set of  $N$  spins along column  $m$ :

$$\mu_m = \{S_{m,1}, S_{m,2}, \dots, S_{m,N}\} \tag{6.141}$$

We have:

$$E[\mu_m, h] = k \sum_{n=1}^N S_{m,n} S_{m,n+1} + h \sum_{n=1}^N S_{m,n} \tag{6.142a}$$

$$E[\mu_m, \mu_{m+1}, h] = k \sum_{n=1}^N S_{m,n} S_{m+1,n} \tag{6.142b}$$

where the first equation is the one body interaction, while the second equation represents the interaction between nearest neighbours columns.

We can write a transfer matrix between these new variables. The transfer matrix permit to transfer along the  $m$ . To make it simpler suppose  $h = 0$  (so the energy does not depend on  $h$ ):

$$\langle \mu_m | \mathbb{T} | \mu_{m+1} \rangle = \exp[k(E[\mu_m, \mu_{m+1}] + E[\mu_m])] \quad (6.143)$$

Now we have to diagonalize. In the 2x2 transfer matrix in the two dimensional we have 2 possible values. Now we have to do the same in principle, but we have to do for all of the (6.141).  $\mathbb{T}$  is a matrix of dimension  $2^N \times 2^N$ . In the thermodynamic limit is an infinite matrix (violation of Perron-Frobenius). According to the formalism

$$Z_N(k, h) = \text{Tr}(\mathbb{T}^N) \quad (6.144)$$

To find the eigenvalues of  $\mathbb{T}$  given by (6.143) is highly non trivial. The big problem it is that in the thermodynamic limit is that the dimension of the transfer matrix goes to infinity, then it is difficult to be diagonalized. This was first achieved by Onsanger in 1944 for the case  $H = 0$  and in the  $N \rightarrow \infty$  limit. The result is given by

$$f_b(T) = -k_B T \log(2 \cosh(2\beta J)) - \frac{k_B T}{2\pi} \int_0^{2\pi} \log \left[ \frac{1}{2} \left( 1 + \sqrt{1 - g^2 \sin^2(\Phi)} \right) \right] d\Phi \quad (6.145)$$

that is the density bulk free energy, where

$$g = \frac{2}{\cosh(2\beta J) \coth(2\beta J)} \quad (6.146)$$

The equation for the critical  $T$  is

$$2 \tanh^2 \left( \frac{2J}{k_B T_c} \right) = 1 \quad \Rightarrow T_c \simeq 2,264 J / k_B \neq 0! \quad (6.147)$$

One can also show that

$$c \propto A \left[ -\ln \left( 1 - \frac{T}{T_c} \right) + B \right] \quad (6.148)$$

The specific heat displays at the transition a logarithmic divergence (no power law!). Therefore the in  $d = 2$  the Ising model as  $\alpha_{\text{Ising } 2D} = 0$ . You do Montecarlo simulation.



## Chapter 7

# The role of dimension, symmetry and range of interactions in phase transitions

Which is the role of dimensionality in phase transition? Consider  $D$ , the dimension of the system. For the Ising model we have seen that in  $D = 1$  there is no phase transition while the Onsager solution tell us that for  $D = 2$  there is a paramagnetic-ferromagnetic transition for  $T_c > 0$ . Therefore, dimensionality seems a crucial parameter! Since in general analytic solutions are not available is there a simple argument to establish the existence of a phase transition? In the case of a para-ferro transition may we establish wheter a phase with long range order exists and is stable withn a range of  $T > 0$ ?

We need an argument that can tell us which kind of system has a phase transition. The idea is to use the entropy energy argument. Indeed, our systems are ruled by a free energy and the previous states are found by making derivative. We have energy and entropy: low energy state can be stable respect thermal fluctuations, but the fluctuations will destroy the long range order. This idea can be generalized.

**Lecture 10.**  
Wednesday 13<sup>th</sup>  
November, 2019.  
Compiled:  
Wednesday 11<sup>th</sup>  
December, 2019.

### 7.1 Energy-entropy argument

$$dF = \underbrace{dU}_{\text{energy}} - T \underbrace{dS}_{\text{entropy}} \quad (7.1)$$

We expect that:

- $T \gg 1$ : entropy should dominates.
- $T \ll 1$ : energy should dominates.

Question: there is a temperature different to zero in which this is compatible?

#### 7.1.1 Ising $D = 1$

Consider  $N$  the size of the system. Study the stability of the states with minimum energy to fluctuations for  $T \neq 0$ . We already know that in the case  $T = 0$ , two ground states exist, either all spins up or all spins down.

For instance, suppose that we have the ground state with all the spin up; the energy of the state is

$$E_G = -JN \quad (7.2)$$

Now consider  $T \neq 0$ , there could be a given number of elementary excitations of the kind spin up/down. What happens if we swap one or more spins? These are defects respect the ground state and they are also called *domain walls*. This is in one dimensional, but is valid also in many dimensional.

Which is the variation in energy  $\Delta E$  respect the ground state? For each excitation there is an energy penalty  $\Delta E = 2J$ , indeed

$$E_G = -JN, \quad E^* = -J(N-1) + J \quad \Rightarrow \Delta E = 2J \quad (7.3)$$

Here we have supposed that we have only one swap. If there is a finite concentration of domains, we have  $x$  such that there are  $M = Nx$  domain walls. Therefore:

$$\Delta E_M = 2MJ \quad (7.4)$$

Now compute the change in entropy. We can compute immediately the entropy of the ground state: this is zero because is the logarithm of the number of configurations, but in this case we have only one configuration, namely  $S_G = \ln 1 = 0$ . Therefore, the different of the entropies is just the entropy of the last state.

Now we estimate the entropy of the states with  $M$  domain walls. The number of possible ways to insert  $M$  domains in  $N$  positions, namely the number of configurations, is

$$\# = \binom{N}{M} = \binom{N}{xN} \quad (7.5)$$

We have:

$$S_M = k_B \log \binom{N}{M} \quad (7.6)$$

the difference is

$$\Delta S = S_M - S_G = S_M = -k_B \ln \binom{N}{xN} \quad (7.7)$$

Consider:

$$\begin{aligned} \Delta F &= F_M - F_G = \Delta E - T\Delta S \\ &= 2MJ - k_B T \ln \binom{N}{M} \\ &= 2xNJ - k_B T \ln \binom{N}{xN} \\ &= N\{2xJ + k_B T[x \ln x + (1-x) \ln (1-x)]\} \end{aligned} \quad (7.8)$$

were we have used the Stirling approximation:  $\ln N! = N \ln N - N$ .

Since equilibrium states are obtained by the minima of  $F$  we can minimize with respect to  $x$ . We are interested in the free energy in the bulk:

$$\Delta f_{bN} = \frac{\Delta F_N}{N}, \quad \frac{\partial \Delta f_b}{\partial x} = 0 \quad (7.9)$$

this gives

$$\begin{aligned} \frac{\partial}{\partial x} \{2xJ + k_B T[x \ln x + (1-x) \ln (1-x)]\} &= 2J + k_B T[\ln x + 1 - \ln (1-x) - 1] \\ &= 2J + k_B T[\ln x - \ln (1-x)] = 0 \end{aligned} \quad (7.10)$$

therefore

$$\ln \frac{x}{1-x} = -\frac{2J}{k_B T} \quad \Rightarrow \quad \frac{x}{1-x} = e^{-2J/k_B T} \quad (7.11)$$

and finally the results is

$$x = \frac{1}{1 + e^{2J/k_B T}} \quad (7.12)$$

It means that  $\forall T \neq 0$  exist a finite concentration  $x$ . The ground state is unstable  $\forall T > 0$ . Indeed, if you have a finite density of  $x$  it means no long range order exist for  $T > 0$ . It tells you that at  $T = 0$ ,  $x = 0$ .

Let us try to do the same for  $D$  dimensions.

### 7.1.2 $D$ dimension

What is a *domain wall* in  $D$  dimension? The domain walls is an hypersurface of size  $L^{D-1}$

$$\Delta E \propto 2JL^{D-1} \quad (7.13)$$

Computing the entropy it is a very difficult problem. In fact, the entropy of a fluctuating hypersurface is difficult to estimate. For a single domain wall we can say

$$S^* \geq k_B \ln L \quad (7.14)$$

the number of ways to place a straight wall within a system of linear size  $L$ . The  $\Delta S$  is just  $S$  because the entropy of the ground state is again zero.

*Remark.* Underestimate of  $S$

$$\Delta F = 2JL^{D-1} - k_B T \ln L \quad (7.15)$$

it means that now energy can win if the temperature is different from zero. Therefore, for  $D = 2$  or greater ( $D > 1$ ) that long range order can survive thermal fluctuations, the system could present an ordered phase!

### Peierls argument

Idea: try to perturb the system using an external magnetic field as perturbation (it is very small  $h$ ). We are breaking explicitly the symmetry, but then you take the limit  $h \rightarrow 0$  and switch off the magnetic field. This is the typical way to see this stability.

An argument due to *Pierls* shows that this is indeed the case. We now that for finite systems, from the  $\mathbb{Z}^2$  symmetry it follows

$$\langle m \rangle_N = 0 \quad (7.16)$$

If, as  $N \rightarrow \infty$  we have  $\langle m \rangle_\infty \neq 0$  with  $T < T_c$ . We have spontaneous symmetry breaking. This can be seen as a thermodynamical instability, namely the value of  $\langle m \rangle_\infty$  in the ferromagnetic phase, is determined by small perturbations. Usually the value of  $\langle m \rangle_\infty$  is determined by using an infinitesimal magnetic field:

$$\langle m \rangle_\infty = \lim_{h \rightarrow 0^+} \lim_{N \rightarrow \infty} \langle m \rangle_N^{(h)} \quad (7.17)$$

*Remark.* It is crucial to take first the thermodynamic limit and then the limit  $h \rightarrow 0^+$ !

That is to see the stability. Another way to do that is instead using a small  $h$ , using periodic boundary condition.

A different infinitesimal perturbation that trigger the instability is to chose appropriate boundary conditions. For example, if we want  $\langle m \rangle_\infty > 0$  one can choose a boundary condition such that at the surface of the system all the spins are up ( $S_i = +1$  with  $i \in \partial\Omega$ ), as in Figure 7.1.

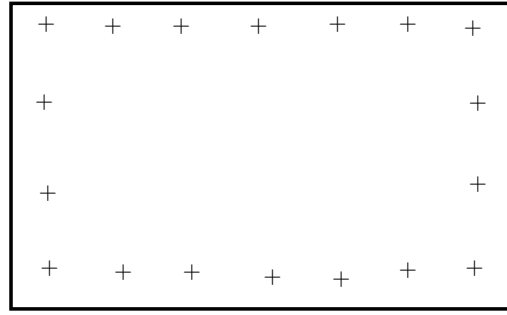


Figure 7.1

It is equivalent, because on the border we have a very high field, by it is at the border so it do not really count for the surface internal. This is a very smart way to do this perturbation. The idea is therefore to perturb the surface. This particular configuration with all the spin up will give you a particular shape, you can do this also in higher dimensional. From this you can also exstimate the temperature.

In the paramagnetic phase this constraint at the boundary does not affect the disorder bulk phase. In the ferromagnetic phase, however, this boundary condition has the same effect of considering an infinitesim magnetic field.

This is the boundary condition chosen by Pierls to establish the existence of a  $T_c \neq 0$  for the  $< D$  Ising model.

Let us gives just a qualitative presentation of the (rigorous) result. If one is interested to the full proof refer for example to [8].

Let  $N_+, N_-$  the number of spin up and down respectively. Clearly,  $N = N_+ + N_-$  and

$$\langle m \rangle_N = \frac{\langle N_+ \rangle - \langle N_- \rangle}{N} = 1 - 2 \frac{\langle N_- \rangle}{N} \quad (7.18)$$

In order to show that  $\langle m \rangle_\infty > 0$  (remember that we are considerin boundary conditions with spin ap at  $\partial\Omega$ ) it is sufficient to show that  $\forall N$

$$\frac{\langle N_- \rangle}{N} < \frac{1}{2} - \varepsilon \quad (7.19)$$

where  $\varepsilon > 0$  and indipendent on  $N$ . Indeed, if (7.19) holds

$$\langle m \rangle_N \geq 2\varepsilon \quad \forall N \quad (7.20)$$

Pierles was indeed able to show that

$$\frac{\langle N_- \rangle}{N} \leq f_D(x) \quad (7.21)$$

where  $f_D$  is a continuous function of  $x$  indipendent on  $N$  and so

$$x = qe^{-4J\beta} \quad (7.22)$$

and such that  $\lim_{x \rightarrow 0} f_D(x) = 0$ .

In particular for  $T$  sufficiently small

$$\frac{\langle N_- \rangle}{N} < \frac{1}{2} - \varepsilon \quad (7.23)$$

More precisely one has

$$\frac{\langle N_- \rangle}{N} \leq \frac{x^2}{36} \frac{2-x}{(1-x)^2} \quad (7.24)$$

where  $x = qe^{-4J\beta} < 1$ .

Note that above bound gives also a lower bound on the critical temperature

$$\frac{\langle N_- \rangle}{N} \leq \frac{x^2}{36} \frac{2-x}{(1-x)^2} < \frac{1}{2} - \varepsilon \quad (7.25)$$

As long as  $\frac{\langle N_- \rangle}{N} < \frac{1}{2} - \varepsilon$  the system is in the ferromagnetic phase. The critical value  $x_c \equiv x(\beta_c)$  must be outside the interval  $[0, x_{1/2}]$  where  $x_{1/2}$  is the smallest positive solution of the equation

$$\frac{x^2}{36} \frac{2-x}{(1-x)^2} = \frac{1}{2} \quad (7.26)$$

From the solution  $x_{1/2}$  and the condition  $x_c > x_{1/2}$  one has

$$J\beta_c \leq J\beta_{1/2} \quad (7.27)$$

where  $J\beta_{1/2} = \frac{1}{4} \log q/x_{1/2}$  which implies  $T_c > T_{1/2}$ .

**Exercise 4.** The equation

$$x^3 + 16x^2 - 36x + 18 = 0 \quad (7.28)$$

gives  $x_{1/2}$ . Found  $T_{1/2}$ .

## 7.2 Role of the symmetry

Interacting systems can be classified with respect to their *global symmetry group*.

**Example 18** (Ising model).

$$\mathcal{H}_{\text{Ising}} = - \sum_{i < j} J_{ij} \sigma_i \sigma_j \quad (7.29)$$

where  $\sigma_i \in \{-1, 1\}$ . The symmetry group of this Hamiltonian is  $\mathbb{Z}^2$ , which has two elements  $\{\mathbb{1}, \mu\}$ . We have

$$\mathbb{1} : \text{identity}, \quad \mu \sigma_i = -\sigma_i, \quad \mu^2 = \mathbb{1} \quad (7.30)$$

**Example 19** (Potts model).

$$\mathcal{H}_{q\text{-Potts}} = - \sum_{i < j} J_{ij} \delta_{\sigma_i, \sigma_j} \quad (7.31)$$

where  $\sigma_i \in [1, 2, 3, \dots, q]$ .  $\mathcal{H}_{q\text{-Potts}}$  is invariant under the permutation group of the sequence  $\{1, 2, 3, \dots, q\}$ . There are  $q!$  elements, for example  $\{2, 1, 3, \dots, q\}$ . The symmetry group is denoted by  $S_q$ .

*Remark.* The difference between a  $\mathbb{Z}_q$  and  $S_q$  symmetry is that an Hamiltonian has symmetry  $\mathbb{Z}_q$  if it is invariant with respect to *cyclic permutations*

$$\mu = \begin{pmatrix} 1 & 2 & \dots & q-1 & q \\ 2 & 3 & \dots & q & 1 \end{pmatrix} \quad (7.32)$$

and its powers  $\mu^l$  with  $l = 0, \dots, q-1$ . Both models satisfy a *discrete global symmetry*.

Now, we jump into the case in which we consider *continuous* symmetries.

### 7.3 XY model

This is a spin model that is invariant with respect to the continuous global symmetry  $\theta_i \rightarrow \theta_i + \alpha$ . Indeed the Hamiltonian of this model is

$$\mathcal{H}_{XY} = - \sum_{i < j} J_{ij} \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_j \quad (7.33)$$

where  $\vec{\mathbf{S}}_i$  is a  $2D$  spin vector

$$\vec{\mathbf{S}}_i = (S_{x_i}, S_{y_i}) \quad (7.34)$$

that can assume values on the unit circle ( $|\vec{\mathbf{S}}_i| = 1$ ).

Suppose that you have spins that are sitting in hyper dimensional. Rotate along a circle this spins. They can assume all the value as in Figure 7.2.

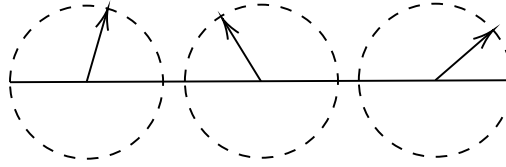


Figure 7.2

The simplest way to parametrize the Hamiltonian is by the angle. Denoting by  $\theta_i$  the direction angle of spins  $\vec{\mathbf{S}}_i$ , the  $\mathcal{H}_{XY}$  can be written as

$$\mathcal{H}_{XY} = - \sum_{i < j} J_{ij} \cos(\theta_i - \theta_j) \quad (7.35)$$

with  $\theta_i \in [0, 2\pi]$ .

*Remark.* The interaction term  $\cos(\theta_i - \theta_j)$  can be written also as

$$\frac{1}{2} (Z_i^* Z_j + Z_i Z_j^*) \quad (7.36)$$

where  $Z_j = \exp(i\theta_j)$ .

The model is invariant under the global transformation

$$Z_i \rightarrow e^{i\alpha} Z_i \quad (7.37)$$

The phase  $\exp(i\alpha)$  form a group under multiplication known as  $U(1)$  that is equivalent to  $O(2)$ . Indeed the interaction term can be written also as

$$\hat{\Omega}_i \cdot \hat{\Omega}_j \quad (7.38)$$

where  $\hat{\Omega}_i = (\cos \theta_i, \sin \theta_i)$ .

*Remark.* In  $n$ -dimensions  $\hat{\Omega}$  has  $n$  components  $\hat{\Omega} = \{\Omega^1, \Omega^2, \dots, \Omega^n\}$  and the corresponding Hamiltonian

$$\mathcal{H} = - \sum_{i > j} J_{ij} \hat{\Omega}_i \cdot \hat{\Omega}_j \quad (7.39)$$

It is symmetric with respect to the global symmetry group  $O(n)$ .

Which are the domain walls for continuous symmetries? Which are the implications for the stability of the ordered phase?

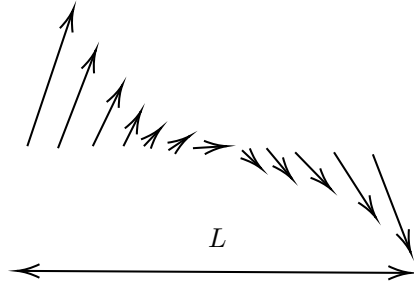


Figure 7.3

## 7.4 Continuous symmetries and phase transitions

The question is: when it is not continuous which is the boundary? When the symmetry is continuous the domain walls interpolate smoothly between two ordered regions (see Figure 7.3).

The energy term that in Ising is proportional to  $2JL^{D-1}$  how does it change here? Suppose that the variation of the direction between two nearest neighbour sites is very small, i.e.  $(\theta_i - \theta_j) \ll 1$  for  $i, j$  nearest neighbours. Now we can dilute the energy, in other words weak the energy term.

Let us do a Taylor expanding of the interaction term

$$\cos(\theta_i - \theta_j) \simeq 1 - \frac{1}{2}(\theta_i - \theta_j)^2 \Rightarrow \sum_{\langle ij \rangle} \left( 1 - \frac{1}{2}(\theta_i - \theta_j)^2 \right) \quad (7.40)$$

The Hamiltonian can be written as

$$\mathcal{H} \simeq -J \sum_{\langle ij \rangle} \left( 1 - \frac{1}{2}(\theta_i - \theta_j)^2 \right) \quad (7.41)$$

The (7.40) corresponds to the discrete differential operator where  $\theta_i - \theta_j = \partial_x \theta$ , therefore

$$\mathcal{H} = E_0 + \underbrace{\frac{J}{2} \int d\vec{r} (\nabla \theta)^2}_{E \equiv \text{Stiffness energy}} \quad (7.42)$$

where  $E_0 = 2JN$  is the energy corresponding to the case in which all the spins are oriented along a given direction.

**Definition 5** (Stiffness energy). The Stiffness energy is defined as

$$E = \frac{J}{2} \int d\vec{r} (\nabla \theta)^2 \quad (7.43)$$

where  $\theta(\vec{r})$  is the angle of a local rotation around an axis and  $J$  is the *spin rigidity*. For an ordered phase  $\theta(\vec{r}) = \theta_0$ .

Let us now imagine a domain wall where  $\theta(\vec{r})$  rotates by  $2\pi$  (or  $2\pi n$ ) by using the entire length of the system (see again Figure 7.3):

$$\theta(\vec{r}) = \frac{2\pi n x}{L} \quad (7.44)$$

where  $n$  is the total number of  $2\pi$  turn of  $\theta$  in  $L$ . Note that there is no variation along the other  $D - 1$  dimensions, therefore we just do over one dimension.

Consider only the term called  $E$

$$E = \frac{J}{2} L^{D-1} \int_0^L dx \left( \frac{d}{dx} \left( \frac{2\pi n x}{L} \right) \right)^2 = \frac{J}{2} L^{D-1} \int_0^L dx \left( \frac{2\pi n}{L} \right)^2 \approx 2\pi^2 n^2 J L^{D-2} \quad (7.45)$$

*Remark.* Unlike the Ising model where  $E \sim L^{D-1}$ , here  $E \sim L^{D-2}$ !

If  $S \geq k_B \ln L$  for a single domain wall,  $S$  should dominate if  $D \leq 2$ , the ordered phase is always unstable and no phase transition is expected for  $T \neq 0$ !

**Definition 6** (Lower critical dimension). The Lower Critical dimension  $D_c$  is the dimension at which (and below which) the system does not display a ordered phase (there is no long range order). In other words if  $D \leq D_c$ , we have  $T_c = 0$ .

**Theorem 7.4.1** (Mermin-Wagner). *For continuous global symmetries the  $D_c = 2$ .*

From what we have found before we can say that

- Discrete global symmetries:  $D_c = 1$ .
- Continuous global symmetries:  $D_c = 2$ .

*Remark.* The  $XY$  model in  $D = 2$  is rather special. Although it does not display an ordered phase, there exist at  $T \neq 0$  a special phase transition known as the *Kosterlitz-thouless transition*. This transition does not imply the spontaneous breaking of the  $O(2)$  symmetry! There is no long range order for  $T < T_{KT}$  (statistic of vortices, topological defects in  $D = 2$ ).

## 7.5 Role of the interaction range

So far we have considered models where the interactions were short range. How things change if long range are considered instead? How does the symmetry broken depends on the range of interactions? One can show, for example, that if

$$J_{ij} = \frac{J}{|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|^\alpha} \quad (7.46)$$

with  $1 \leq \alpha \leq 2$ , phase with long range order is stable for  $0 < T < T_c$  also for  $D = 1$ ! So we have a long range order with  $T_c \neq 0$ , even for  $D = 1$ !

*Remark.* If  $\alpha > 2 + \varepsilon$  we get back the physics found for short range interactions. If  $\alpha < 1$  the thermodynamic limit does not exist.

A limiting case of long range interaction is the infinite range case where all the spins interact one to another with the same intensity independently on their distance. No metric is involved (instead of previously where the definition of  $J$  of before is a metric.). It can be solved exactly and later we will see why.

## 7.6 Ising model with infinite range

The Hamiltonian is the following:

$$-\mathcal{H}_N(\{S\}) = \frac{J_0}{2} \sum_{i,j}^N S_i S_j + H \sum_i S_i \quad (7.47)$$

with  $S_i \in [-1, +1]$ . The problem is the double sum over  $i, j$ , indeed

$$\sum_{i,j} S_i S_j \propto O(N^2) \quad (7.48)$$

and the thermodynamic limit is ill-defined. To circumvent this problem *Kac* suggested to consider a strength

$$J_0 = \frac{J}{N} \quad (7.49)$$



this is called the *kac* approximation

$$-\mathcal{H}_N(\{S\}) = \frac{J}{2N} \sum_{i,j}^N S_i S_j + H \sum_i S_i \quad (7.50)$$

with this choice we recover  $E \sim O(N)$ . In this Hamiltonian since you have no metric you have no dimension.

The partition function is

$$Z_N(T, J, H) = \sum_{\{S\}} \exp \left[ \frac{\beta J}{2N} \sum_{ij} S_i S_j + \beta H \sum_i S_i \right] \quad (7.51)$$

Since there are no restrictions on the double sum we can write

$$\sum_{ij} S_i S_j = \left( \sum_i S_i \right) \left( \sum_j S_j \right) = \left( \sum_i S_i \right)^2 \quad (7.52)$$

Rewriting the partition function:

$$Z_N(T, J, H) = \sum_{\{S\}} \exp \left[ \frac{K}{2N} \left( \sum_i S_i \right)^2 + h \sum_i S_i \right] \quad (7.53)$$

In order to transform the quadratic term into a linear one we make use of the integral identity known as the *Hubbard-stratonovitch transformation* (we can do it in any dimension). The idea is to rewrite something as a square.

Let us define:

$$x \equiv \sum_i S_i \quad (7.54)$$

The following identity holds

$$e^{\frac{Kx^2}{2N}} = \sqrt{\frac{NK}{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{NK}{2}y^2 + Kxy} dy, \quad \text{with } \text{Re}(K) > 0 \quad (7.55)$$

where  $y$  is a random field that follows a random distribution. To show the identity above it is sufficient to complete the square

*Proof.* (of (7.55))

$$-\frac{NK}{2}y^2 + Kxy = -\frac{NK}{2}\left(y - \frac{x}{N}\right)^2 + \frac{Kx^2}{2N} \quad (7.56)$$

Then we integrate:

$$e^{\frac{Kx^2}{2N}} \int_{-\infty}^{+\infty} e^{-\frac{NK}{2}\left(y - \frac{x}{N}\right)^2} dy \stackrel{(a)}{=} e^{\frac{Kx^2}{2N}} \sqrt{\frac{2\pi}{NK}} \quad (7.57)$$

where in (a) we have considered  $z \equiv \left(y - \frac{x}{N}\right)$ ,  $dz = dy$ :

$$\rightarrow \int_{-\infty}^{+\infty} e^{-\alpha z^2} dz = \sqrt{\frac{\pi}{\alpha}} \quad (7.58)$$

with  $\alpha \equiv \frac{NK}{2}$ . □

**Lecture 11.**  
Wednesday 20<sup>th</sup>  
November, 2019.  
Compiled:  
Wednesday 11<sup>th</sup>  
December, 2019.

By using (7.55) we have

$$Z_N(K, h) = \sqrt{\frac{NK}{2\pi}} \int_{-\infty}^{+\infty} dy e^{-\frac{NK}{2}y^2} \underbrace{\left[ \sum_{\{S\}} e^{(h+Ky) \sum_i S_i} \right]}_{Q_y} \quad (7.59)$$

Sometimes  $y$  is called *auxiliary field*. If  $y$  is a fluctuating external field with Gaussian distribution

$$Q_y = \sum_{\{S\}} e^{(h+Ky) \sum_i S_i} = \prod_i^N \left( \sum_{S_i=\pm 1} \exp[(h+Ky)S_i] \right) = (2 \cosh(h+Ky))^N \quad (7.60)$$

The partition function becomes

$$Z_N(K, h) = \sqrt{\frac{NK}{2\pi}} \int_{-\infty}^{+\infty} dy e^{-\frac{NK}{2}y^2} (2 \cosh(h+Ky))^N = \sqrt{\frac{NK}{2\pi}} \int_{-\infty}^{+\infty} dy e^{N\alpha(K, h, y)} \quad (7.61)$$

where

$$\alpha(K, h, y) = \ln [2 \cosh(h+Ky)] - \frac{K}{2}y^2 \quad (7.62)$$

*Remark.* In the limit  $N \rightarrow \infty$  the integral can be computed exactly by the *saddle point method*. We can replace the medium of the integral with the maximum of the integrand, we say that all the information is coming only from a bit of information. Replace the all integral with the integrand computed where it is maximum. That is an approximation and we are losing information, it depends on the form of the function. For example, for a delta function it works better. In general:

$$\int_{-\infty}^{+\infty} f(x) dy \rightarrow f(\bar{x}) \quad (7.63)$$

where  $\bar{x} = \max_x f(x)$ .

Indeed as  $N \rightarrow \infty$ , since the integrand is  $\exp(N\alpha(K, h, y))$ , the integral is dominated by the global maximum in  $y$  of the function  $\alpha(K, h, y)$ .

$$Z_N(K, h) \stackrel{N \gg 1}{\approx} \sqrt{\frac{NK}{2\pi}} \max_y \left[ e^{N\alpha(K, h, y)} \right] \quad (7.64)$$

Let  $y_s$  be the value of  $y$  at which

$$\alpha(K, h, y_s) = \max_y \alpha(K, h, y) \quad (7.65)$$

therefore

$$Z_N(K, h) \stackrel{N \gg 1}{\approx} \sqrt{\frac{NK}{2\pi}} e^{N\alpha(K, h, y_s)} \quad (7.66)$$

When we are able to compute the  $y_s$  we can do this approximation and

$$f_b(K, h) = \lim_{N \rightarrow \infty} \frac{1}{N} (-k_B T \log Z_N) = -k_B T \alpha(K, h, y_s) \quad (7.67)$$

Looking for  $y_s$  we consider the condition  $\frac{\partial \alpha}{\partial y} = 0$ :

$$\frac{\partial \alpha}{\partial y} = \frac{\sinh(h+Ky)K}{\cosh(h+Ky)} - Ky = 0 \Rightarrow y_s = \tanh(h+Ky_s) \quad (7.68)$$

The last one is an implicit equation that can be solved graphically as a function of  $K$  and  $h$ .

The magnetization in the  $N \rightarrow \infty$  limit is given by

$$\begin{aligned}
 m &= -\left(\frac{\partial f}{\partial H}\right)_T = \lim_{N \rightarrow \infty} \frac{1}{\beta N} \frac{\partial \ln Z_N(K, h)}{\partial H} \\
 &= \frac{\partial \alpha(K, h, y_s)}{\partial h} + \frac{O(\log N)}{N} = \frac{2 \sinh(K y_s + h)}{2 \cosh(K y_s + h)} \\
 &= \tanh(K y_s + h)
 \end{aligned} \tag{7.69}$$

At the end  $m \equiv y_s$

$$m = \tanh(h + K m) \tag{7.70}$$

that is a self consistent equation. We have solved analitically this problem.



## Chapter 8

# Mean field theories of phase transitions and variational mean field

### 8.1 Mean field theories

Increasing the dimension of the systems the effort to solve analitically the problems increase , in fact we have seen that

- In  $D = 1$ : many (simple) models can be solved exactly using techniques such as the transfer matrix method.
- In  $D = 2$ : few models can still be solved exactly (often with a lot of effort).
- In  $D = 3$ : almost no model can be exactly solved.

hence, approximations are needed. The most important and most used one is the *mean field approximation*. The idea is to try to simplify the problem by neglect the correlation between the fluctuations of the order parameter. It is equivalent to a statistical independence of the microscopic degrees of freedom.

The mean field approximation has different names depending on the system considered:

- Magnetic: Weiss theory.
- Fluids: Van der Waals.
- Polymers: Flory's theory.

#### 8.1.1 Mean field for Ising model

Let us start from the following generic model

$$\mathcal{H}[\{S\}] = -\frac{1}{2} \sum_{ij} J_{ij} S_i S_j - H \sum_i S_i \quad (8.1)$$

where the double sum over  $i$  and  $j$  have no restrictions, while  $H$  is homogeneous.

$$Z_N(T, H, \{J_{ij}\}) = \sum_{\{S\}} e^{-\beta \mathcal{H}[\{S\}]} = \exp(-\beta F_N(T, H, \{J_{ij}\})) \quad (8.2)$$

Since  $H$  is uniform, the magnetization per spin is

$$\langle S_i \rangle = \langle S \rangle \equiv m \quad (8.3)$$

Let us now consider the identity

$$\begin{aligned} S_i S_j &= (S_i - m + m)(S_j - m + m) \\ &= (S_i - m)(S_j - m) + m^2 + m(S_j - m) + m(S_i - m) \end{aligned} \quad (8.4)$$

*Remark.* The term

$$(S_i - m)(S_j - m) = (S_i - \langle S_i \rangle)(S_j - \langle S_j \rangle)$$

measures correlation between fluctuations. The mean field approximation consists in neglecting this term!

Using the mean field approximation,

$$S_i S_j \approx m^2 + m(S_i - m) + m(S_j - m) \quad (8.5)$$

we obtain

$$\frac{1}{2} \sum_{i,j} J_{ij} S_i S_j \stackrel{MF}{\approx} \frac{1}{2} \sum_{i,j} J_{ij} [-m^2 + m(S_i + S_j)] \quad (8.6)$$

Let us focus on the term

$$\frac{1}{2} \sum_{i,j} J_{ij} m(S_i + S_j) = 2 \frac{1}{2} m \sum_{i,j} J_{ij} S_i \quad (8.7)$$

If we do not many assumption on  $J_{ij}$  the mean field hamiltonian is

$$\mathcal{H}_{MF}[\{S\}] = \frac{1}{2} m^2 \sum_{ij} J_{ij} - m \sum_{ij} J_{ij} S_i - H \sum_i S_i \quad (8.8)$$

and by calling

$$\bar{J}_i \equiv \sum_j J_{ij} \quad (8.9)$$

we get

$$\mathcal{H}_{MF}[\{S\}] = \frac{1}{2} m^2 \sum_i \bar{J}_i - \frac{m}{2} \sum_i \bar{J}_i S_i - H \sum_i S_i \quad (8.10)$$

*Remark.* Note the coefficient 1/2 that is needed to avoid the double counting of bonds.

If moreover we suppose that

$$\bar{J}_i \rightarrow \bar{J}$$

we have

$$\mathcal{H}_{MF}[\{S\}] = \frac{1}{2} m^2 N \bar{J} - \left( \frac{m}{2} \bar{J} + H \right) \sum_i S_i \quad (8.11)$$

In the standard Ising model, where

$$\frac{1}{2} \sum_{ij} J_{ij} S_i S_j \rightarrow \sum_{\langle ij \rangle} J_{ij} S_i S_j \quad (8.12)$$

the term  $2m \sum_{\langle ij \rangle} J_{ij} S_i$  can be written as follows; let

$$\sum_{j \in n.n. \text{ of } i} J_{ij} = z \hat{J}_i \quad (8.13)$$

where  $z$  is the coordination number of the underlying lattice.

*Remark.* For the hypercubic lattice  $z = 2D$ .

By assuming  $\hat{J}_i = \hat{J}$  and inserting the  $1/2$  to avoid double counting, we have that equation (8.7) becomes

$$2m \sum_{\langle ij \rangle} J_{ij} S_i = 2m \frac{1}{2} z \hat{J} \sum_{i=1}^N S_i \quad (8.14)$$

Hence,

$$\mathcal{H}_{MF}[\{S\}] = \frac{1}{2} m^2 N z \hat{J} - (m z \hat{J} + H) \sum_{i=1}^N S_i \quad (8.15)$$

$$\begin{aligned} Z_N(T, H, \hat{J}) &= e^{-N\beta \hat{J} \frac{z}{2} m^2} \sum_{\{S\}} e^{\beta m (z \hat{J} m + H) \sum_i S_i} \\ &= e^{-N\beta \hat{J} \frac{z}{2} m^2} \left( \sum_{S=\pm 1} \exp\left(\beta m (z \hat{J} m + H)\right) \right)^N \\ &= e^{-N\beta \hat{J} \frac{z}{2} m^2} \left( 2 \cosh\left[\beta (z \hat{J} m + H)\right] \right)^N \end{aligned} \quad (8.16)$$

*Remark.* We are replacing the interaction of the  $J$  with a field close to the  $S_i$ . We called  $\hat{J}zm = H_{eff}$ , the mean field!

The free energy per spin is

$$\begin{aligned} \frac{F_N(T, H, \hat{J})}{N} &= \frac{1}{N} \left( -k_B T \ln Z_N(T, H, \hat{J}) \right) \\ &= \frac{1}{2} \hat{J} z m^2 - k_B T \ln \left[ \cosh\left(\beta (z \hat{J} m + H)\right) \right] - k_B T \ln 2 \end{aligned} \quad (8.17)$$

Sometimes it is useful to use the *dimensionless variables*

$$\bar{f} \equiv \frac{F_N}{N z \hat{J}}, \quad \theta \equiv \frac{k_B T}{z \hat{J}}, \quad \bar{H} \equiv \frac{H}{z \hat{J}} \quad (8.18)$$

Hence,

$$\bar{f}(m, \bar{H}, \theta) = \frac{1}{2} m^2 - \theta \ln \left( 2 \cosh(\theta^{-1}(m + \bar{H})) \right) \quad (8.19)$$

In order to be a self-consistent, the last equation has to satisfy the thermodynamic relation:

$$m = - \left( \frac{\partial f}{\partial H} \right)_T \Rightarrow m = \tanh\left(\beta (z \hat{J} m + H)\right) \quad (8.20)$$

*Remark.* The results of  $m$  is similar to the Ising with infinite range ( $\hat{J}z \leftrightarrow J$ ).

Now, let us consider the  $H = 0$  case, we have

$$m = \tanh\left(\beta (z \hat{J} m)\right) \quad (8.21)$$

and the graphical solution is shown in Figure 8.1 (hyperbolic function).

In particular:

- case  $\beta \hat{J} z > 1$ : there are three solutions, one at  $m = 0$  and two symmetric at  $m = \pm m_0$ . Magnetization is  $\neq 0$  ( $= |m_0|$ ) for  $H = 0$  (*ordered phase*). The two solution are symmetric because they are related by the  $\mathbb{Z}^2$  symmetry;
- case  $\beta \hat{J} z < 1$ : single solution at  $m = 0$  (*disordered or paramagnetic phase*);
- case  $\beta \hat{J} z = 1$ : the three solutions coincide at  $m = 0$  (*critical point*).

As said, the condition  $\beta_c \hat{J} z = 1$ , define the critical point. The critical temperature  $T_c$  is given by

$$\frac{z \hat{J}}{k_B T_c} = 1 \Rightarrow T_c = \frac{z \hat{J}}{k_B} \neq 0! \quad (8.22)$$

*Remark.*  $T_c$  depends on  $z$  and hence on  $D$ !

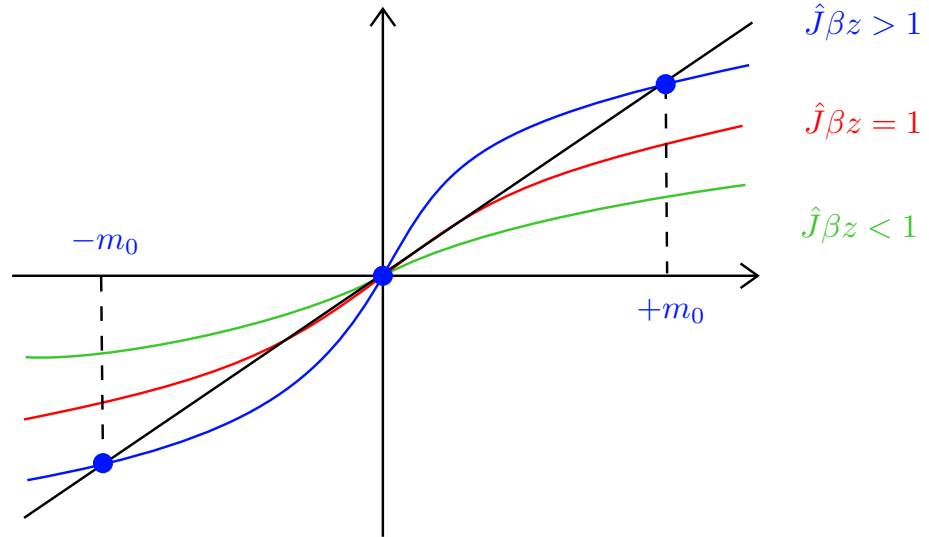


Figure 8.1

### 8.1.2 Free-energy expansion for $m \simeq 0$

The critical point is characterized by the order parameter that is zero. Now we want to expand the free energy around the critical point. Let us put  $H = 0$ :

$$f(m, 0, T, \hat{J}) = \frac{1}{2} \hat{J} z m^2 - k_B T \ln \left[ \cosh(\beta \hat{J} z m) \right] \quad (8.23)$$

Define  $x \equiv \beta \hat{J} z m \simeq 0$ , hence

$$\cosh(x) \simeq 1 + \underbrace{\frac{x^2}{2} + \frac{x^4}{4!}}_{t \simeq 0} + \dots \quad (8.24)$$

$$\log(1+t) \simeq t - \frac{1}{2} t^2 \quad (8.25)$$

$$\log(\cosh x) \simeq \frac{x^2}{2} + \frac{x^4}{4!} - \frac{1}{2} \frac{x^4}{4} + O(x^6) = \frac{x^2}{2} - \frac{x^4}{12} + O(x^6) \quad (8.26)$$

This gives

$$f(m, 0, T, \hat{J}) \simeq \text{const} + \frac{A}{2} m^2 + \frac{B}{4} m^4 + O(m^6) \quad (8.27)$$

with

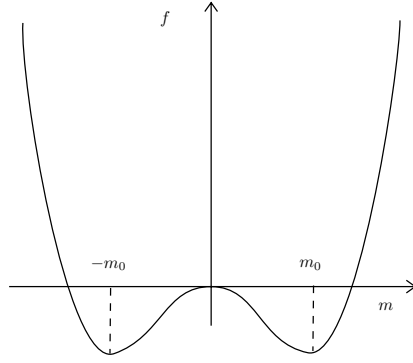
$$A \equiv \hat{J} z (1 - \beta \hat{J} z) \quad (8.28a)$$

$$B \equiv \beta^2 \frac{(\hat{J} z)^4}{3} > 0 \quad (8.28b)$$

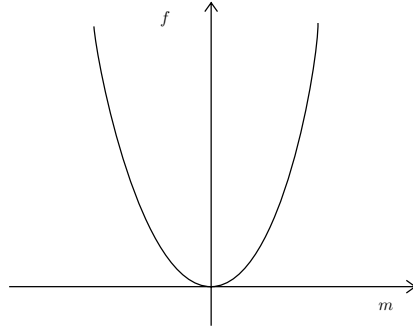
Let us consider:

- case  $\beta \hat{J} z > 1 \Rightarrow A < 0$ : two stable symmetric minima at  $m = \pm m_0$ . Coexistence between the two ordered phases.
- case  $\beta \hat{J} z < 1 \Rightarrow A > 0$ : one minimum at  $m = 0$ .
- case  $\beta \hat{J} z = 1 \Rightarrow A = 0$ : 3 minima coincide at  $m = 0$ .

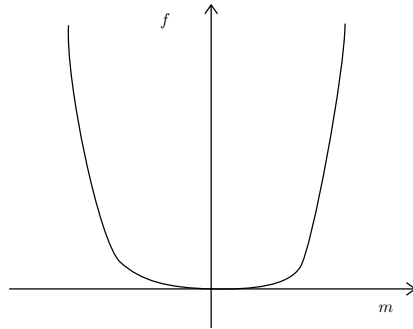




**Figure 8.2:** Plot of the free energy: case  $\beta \hat{J}z > 1 \Rightarrow A < 0$ .



**Figure 8.3:** Plot of the free energy: case  $\beta \hat{J}z < 1 \Rightarrow A > 0$ .



**Figure 8.4:** Plot of the free energy: case  $\beta \hat{J}z = 1 \Rightarrow A = 0$ .

### 8.1.3 Mean field critical exponents

Let us consider

$$f(m, T, 0) \approx \text{const} + \frac{A}{2}m^2 + \frac{B}{4}m^4 + O(m^6) \quad (8.29)$$

with  $B > 0$ , so we do not need more term to find the minima of the solution. This is called stabilization. What is most important is the coefficient  $A = \hat{J}z(1 - \beta \hat{J}z)$ , that means that  $A$  can change sign.

#### $\beta$ exponent

The  $\beta$  exponential observe the order parameter. Consider  $H = 0$ ,  $t \equiv \frac{T - T_c}{T_c}$  and  $m \stackrel{t \rightarrow 0^-}{\sim} -t^\beta$ . The condition of equilibrium is

$$\frac{\partial f}{\partial m} = 0 \quad (8.30)$$

which implies

$$\left. \frac{\partial f}{\partial m} \right|_{m=m_0} = Am_0 + Bm_0^3 = [\hat{J}z(1 - \beta\hat{J}z) + Bm_0^2]m_0 = 0 \quad (8.31)$$

Since  $T_c = \frac{\hat{J}z}{k_B}$  :

$$0 = \frac{k_B T_c}{T} (T - T_c) m_0 + Bm_0^3 \quad (8.32)$$

The solution are  $m_0 = 0$  and

$$m_0 \simeq (T_c - T)^{1/2} \quad (8.33)$$

The mean field value is so  $\beta = 1/2$ .

### $\delta$ exponent

Now, let us concentrate in the  $\delta$  exponent. We are in the only case in which we are in  $T = T_c$  and we want to see how the magnetization decrease:  $H \sim m^\delta$ .

Starting from the self-consistent equation, we have

$$m = \tanh(\beta(\hat{J}zm + H)) \quad (8.34)$$

Inverting it

$$\beta(\hat{J}zm + H) = \tanh^{-1} m \quad (8.35)$$

On the other hand, for  $m \sim 0$

$$\tanh^{-1} m \simeq m + \frac{m^3}{3} + \frac{m^5}{5} + \dots \quad (8.36)$$

Therefore, by substituting

$$\begin{aligned} H = k_B T \left( m + \frac{m^3}{3} + \dots \right) - \hat{J}zm &= (k_B T - \hat{J}z)m + k_B T \frac{m^3}{3} + \dots \\ &\simeq k_B (T - T_c)m + \frac{k_B T}{3} m^3 \end{aligned} \quad (8.37)$$

At  $T = T_c = \frac{\hat{J}z}{k_B}$ , we have

$$H \sim k_B T_c \frac{m^3}{3} \quad (8.38)$$

The mean field value is  $\delta = 3$ .

### $\alpha$ exponent

Consider the  $\alpha$  exponent, for  $H = 0$ ,  $c_H \sim t^{-\alpha}$  and we have  $t = (T - T_c)/T_c$ . Compute the specific heat at  $H = 0$ . Consider first  $T > T_c$ , where  $m_0 = 0$ ,

$$f(m, H) = \frac{\hat{J}zm^2}{2} - \frac{1}{\beta} \ln \left( 2 \cosh(\beta(\hat{J}zm + H)) \right) \quad (8.39)$$

If  $m = 0$ ,  $\cosh 0 = 1$  and

$$f = -k_B T \ln 2 \quad (8.40)$$

therefore

$$c_H = -T \left( \frac{\partial^2 f}{\partial T^2} \right) = 0 \quad (8.41)$$

The mean field value is  $\alpha = 0$ .

*Remark.* For  $T < T_c$ ,  $m = m_0 \neq 0$ . This implies that  $c_H \neq 0$ , but still  $f = -k_B T \ln A$  with  $A = \text{const}$ . We obtain  $\alpha = 0$  also in this case.

$$m_0 = \pm \sqrt{-\frac{\hat{J}z}{2T_c} (T - T_c)} \quad (8.42)$$

**$\gamma$  exponent**

Now we consider the  $\gamma$  exponent, for  $H = 0$ ,  $\chi \sim t^{-\gamma}$ . Starting again from equation (8.34):

$$m = \tanh\left(\beta(\hat{J}zm + H)\right)$$

and developing it around  $m \simeq 0$ , as shown before we get

$$H = mk_B(T - T_c) + \frac{k_B T}{3} m^3 \quad (8.43)$$

$$\Rightarrow \chi_T = \frac{\partial m}{\partial H} = \frac{1}{\frac{\partial H}{\partial m}} \quad (8.44)$$

Since  $\frac{\partial H}{\partial m} \simeq k_B(T - T_c) + K_B T m^2$ , as  $m \rightarrow 0$

$$\chi \sim (T - T_c)^{-1} \quad (8.45)$$

The mean field value is  $\gamma = 1$ .

**Summary**

The mean field critical exponents are

$$\beta = \frac{1}{2}, \quad \gamma = 1, \quad \delta = 3, \quad \alpha = 0 \quad (8.46)$$

*Remark.* In the mean field critical exponents the dimension  $D$  does not appear.  $T_c$  instead depends on the number of  $z$  of nearest neighbours and hence on the embedding lattice!

*Remark.* (lesson) The  $\nu$  exponent define the divergence of the correlation lengths. In order to do that, in principle we should compute the correlation function, but which are the correlation we are talking about? The correlation or the fluctuation with to respect the average? In the ferromagnetic we have infinite correlation lengths, but it is not true, because instead of that we consider the variation correlated! Which is the problem here? In mean field we were neglecting correlation between fluctuation. We thought: let us compute neglecting correlation. How we can compute the correlation function within the meanfield theory with thermal fluctuations? We look at the response of the system. Experimentally what can we do? It is a magnetic field, but we cannot use homogeneous magnetic field. Another way to compute the correlation function without looking at thermal fluctuation it is by considering a non homogeneous magnetic field.

If we make a variation in  $H_i$  in the system, what happen in the  $H_j$ ? This is an important point.

**8.2 Mean field variational method**

The mean field variational method is a general approach to derive a mean field theory. The method is valid for all  $T$  and is sufficiently flexible to deal with complex systems.

The method is similar to the one used in quantum mechanics, namely it is based on the following inequality

$$E_\alpha = \langle \psi_\alpha | \hat{H} | \psi_\alpha \rangle \geq E_0 \quad (8.47)$$

valid for all trial function  $\psi_\alpha$ .

*Remark.*  $E_0$  is the ground state energy.

**Example 20.** In many body problem we have Hartree and Hartree-Fock.

The closest bound to  $E_0$  is the one that is obtained by minimizing  $E_\alpha$ , i.e.  $\langle \psi_\alpha | \hat{H} | \psi_\alpha \rangle$  over  $|\psi_\alpha\rangle$ , where the  $|\psi_\alpha\rangle$  are functions to be parametrized in some convenient way.

The method is based on the following inequalities

1. Let  $\Phi$  be a random variable (either discrete or continuous) and let  $f(\Phi)$  be a function of it.

For all function  $f$  of  $\Phi$ , the mean value with respect to a distribution function  $p(\Phi)$  is given by

$$\langle f(\Phi) \rangle_p \equiv \text{Tr}(p(\Phi)f(\Phi)) \quad (8.48)$$

If we consider the function

$$f(\Phi) = \exp[-\lambda\Phi] \quad (8.49)$$

it is possible to show the inequality

$$\langle e^{-\lambda\Phi} \rangle_p \geq e^{-\lambda\langle\Phi\rangle_p}, \quad \forall p \quad (8.50)$$

*Proof.*  $\forall \Phi \in \mathbb{R}$ ,  $e^\Phi \geq 1 + \Phi$ . Hence,

$$e^{-\lambda\Phi} = e^{-\lambda\langle\Phi\rangle} e^{-\lambda[\Phi - \langle\Phi\rangle]} \geq e^{-\lambda\langle\Phi\rangle} (1 - \lambda(\Phi - \langle\Phi\rangle))$$

Taking the average of both sides, we get

$$\rightarrow \langle e^{-\lambda\Phi} \rangle_p \geq \langle (1 - \lambda(\Phi - \langle\Phi\rangle)) e^{-\lambda\langle\Phi\rangle} \rangle_p = e^{-\lambda\langle\Phi\rangle_p}$$

□

2. The second inequality refers to the free energy. Let  $\rho(\Phi)$  be a probability distribution, i.e. such that

$$\text{Tr}(\rho(\Phi)) = 1 \quad \rho(\Phi) \geq 0 \quad \forall \Phi \quad (8.51)$$

Hence,

$$e^{-\beta F_N} = Z_N = \text{Tr}_{\{\Phi\}} e^{-\beta \mathcal{H}[\{\Phi\}]} = \text{Tr}_{\{\Phi\}} \rho e^{-\beta \mathcal{H} - \ln \rho} = \langle e^{-\beta \mathcal{H} - \ln \rho} \rangle_\rho \quad (8.52)$$

From the inequality (8.50),

$$e^{-\beta F_N} = \langle e^{-\beta \mathcal{H} - \ln \rho} \rangle_\rho \geq e^{-\beta \langle \mathcal{H} \rangle_\rho - \langle \ln \rho \rangle_\rho} \quad (8.53)$$

Taking the logs one has

$$F \leq \langle \mathcal{H} \rangle_\rho + k_B T \langle \ln \rho \rangle_\rho = \text{Tr}(\rho \mathcal{H}) + k_B T \text{Tr}(\rho \ln \rho) \equiv F_\rho \quad (8.54)$$

Whenever we are able to write the last equation by using a  $\rho$ , then we will minimize. This is the variational approach of statistical mechanics. The question is: which is the  $\rho$  that minimize?

The functional  $F_\rho$  will reach its minimum value with respect to the variation of  $\rho$  with the constraint  $\text{Tr}(\rho) = 1$ , when

$$\bar{\rho} = \rho_{eq} = \frac{1}{Z} e^{-\beta \mathcal{H}} \quad (8.55)$$

So far so good but not very useful, since we are back to the known result that the distribution that best approximately the free energy of the canonical ensemble is given by the Gibbs-Boltzmann distribution. To compute  $\rho_{eq}$ , we need some approximation!

### 8.2.1 Mean field approximation for the variational approach

Let us now try to compute the  $Z$  by starting to the inequality (8.54). Up to now everything is exact. The idea is to choose a functional form of  $\rho$  and then minimize  $F_\rho$  with respect to  $\rho$ . Note that  $\rho$  is the  $N$ -point probability density function (it is a function of all the degrees of freedom):

$$\rho = \rho(\Phi_1, \dots, \Phi_N) \quad (8.56)$$

it is a  $N$ -body problem. The  $\Phi_\alpha$  is the random variables associated to the  $\alpha$ -esim degree of freedom. This is in general a very difficult distribution to deal with. This is equivalent exactly when we have:

$$\psi_\alpha(\vec{\mathbf{r}}_1, \vec{\mathbf{P}}_1, \dots, \vec{\mathbf{r}}_N, \vec{\mathbf{P}}_N) \quad (8.57)$$

The mean-field approximation consists in factorising  $\rho$  into a product of 1-point distribution function:

$$\rho(\Phi_1, \dots, \Phi_N) \stackrel{MF}{\simeq} \prod_{\alpha=1}^N \rho^{(1)}(\Phi_\alpha) \equiv \prod_{\alpha=1}^N \rho_\alpha \quad (8.58)$$

where we have used the short-hand notation  $\rho^{(1)}(\Phi_\alpha) \rightarrow \rho_\alpha$ .

*Remark.* Approximation (8.58) is equivalent to assume statistical independence between particles (or more generally between different degrees of freedom). The independence of the degree of freedom is a very strong assumption!

**Example 21.** Let us consider the spin model on a lattice; what is the  $\Phi_\alpha$ ? We have:

$$\Phi_\alpha \rightarrow S_i$$

Hence,  $\rho = \rho(S_1, S_2, \dots, S_N)$  and (8.58) becomes

$$\rho \stackrel{MF}{\simeq} \prod_{i=1}^N \rho^{(1)}(S_i) \equiv \prod_{i=1}^N \rho_i \quad (8.59)$$

With (8.58) and  $\text{Tr}(\rho_\alpha) = 1$ , we compute the two averages in the (8.54) given the field. We have:

$$\text{Tr}_{\{\Phi\}}(\rho \ln \rho) = \text{Tr} \left( \prod_{\alpha} \rho_{\alpha} \left( \sum_{\alpha} \ln \rho_{\alpha} \right) \right) \stackrel{\text{to do}}{=} \sum_{\alpha} \text{Tr}^{(\alpha)}(\rho_{\alpha} \ln \rho_{\alpha}) \quad (8.60)$$

where  $\text{Tr}^{(\alpha)}$  means sum over all possible values of the random variable  $\Phi_\alpha$  (with  $\alpha$  fixed).

*Remark.* We have  $\text{Tr}^{(\alpha)} \rho_\alpha = 1$ .

We end up that

$$F_{\rho_{MF}} = \langle \mathcal{H} \rangle_{\rho_{MF}} + k_B T \sum_{\alpha} \text{Tr}^{(\alpha)}(\rho_{\alpha} \ln \rho_{\alpha}) \quad (8.61)$$

*Remark.*  $F_{\rho_{MF}} = F(\{\rho_\alpha\})$  and we have to minimize it with respect to  $\rho_\alpha$ .

How can we parametrise  $\rho_\alpha$ ? There are two approaches that are mostly used:

1. Parametrize  $\rho_\alpha \equiv \rho^{(1)}(\Phi_\alpha)$  by the average of  $\Phi_\alpha$  with respect to  $\rho_\alpha$ ,  $\langle \Phi_\alpha \rangle_{\rho_\alpha}$  (in general is the local order parameter):

$$\rho_\alpha = \rho^{(1)}(\Phi_\alpha) \rightarrow \langle \Phi_\alpha \rangle_{\rho_\alpha}$$

This means that there are two constraints in the minimization procedure:

$$\text{Tr}^{(\alpha)} \rho_\alpha = 1, \quad \text{Tr}^{(\alpha)}(\rho_\alpha \Phi_\alpha) = \langle \Phi_\alpha \rangle \quad (8.62)$$

where the second is the self-consistent equation.

*Remark.* In this case the variational parameter coincides with the order parameter.

2. In the second approach is  $\rho_\alpha$  itself the variational parameter.  $F_{\rho_{MF}}$  is minimized by varying  $\rho_\alpha$ . It is a more general approach, that involves functional minimization.

### 8.2.2 First approach: Bragg-Williams approximation

We apply this approach to the Ising model with non uniform magnetic field:

$$\mathcal{H}[\{S\}] = -J \sum_{\langle ij \rangle} S_i S_j - \sum_i H_i S_i \quad (8.63)$$

it means that

$$\Phi_\alpha \rightarrow S_i = \pm 1 \quad (8.64)$$

and that the order parameter (variational parameter) becomes

$$\langle \Phi_\alpha \rangle \rightarrow \langle S_i \rangle \equiv m_i \quad (8.65)$$

*Remark.* Note that this time  $H \rightarrow H_i$  (non-uniform), hence  $m_i$  depends on the site  $i$ .

We have to define a 1-particle probability density distribution  $\rho_i \equiv \rho^{(1)}(S_i)$  such that

$$\rho^{(1)} \equiv \rho^{(1)}(S_i) \rightarrow \begin{cases} \text{Tr} \rho_i^{(1)} = 1 \\ \text{Tr} \rho_i^{(1)} S_i = m_i \end{cases} \quad (8.66)$$

Since we have to satisfy these two constraints, we need two free parameters. A linear functional form is sufficient. Denoting by:

- $a$ : statistical weight associated to the value  $S_i = -1$ .
- $b$ : statistical weight associated to all the remaining possible values of  $S_i$  (for an Ising only one value remains, i.e.  $S_i = +1$ ).

The simplest function form with two parameters is the linear function, namely

$$\rho_i \equiv \rho^{(1)}(S_i) = a(1 - \delta_{S_i,1}) + b\delta_{S_i,1} \quad (8.67)$$

Using the constraints

$$\begin{cases} \text{Tr}^{(i)}(\rho_i) = 1 & \rightarrow a + b = 1 \\ \text{Tr}^{(i)}(\rho_i S_i) = m_i & \rightarrow a - b = m_i \end{cases} \quad (8.68)$$

where  $a, b$  are the functions of the order parameter. In that case we have not to write the functions for all the  $i$ . For  $S_i = 1$  we have one value, for all the other values another one.

The results of the previous equation are:

$$\begin{cases} a = \frac{1-m_i}{2} \\ b = \frac{1+m_i}{2} \end{cases} \quad (8.69)$$

Hence,

$$\rho_i = \frac{1-m_i}{2}(1 - \delta_{S_i,-1}) + \frac{1+m_i}{2}\delta_{S_i,-1} \quad (8.70)$$

In matrix form

$$\begin{pmatrix} \frac{(m_i+1)}{2} & 0 \\ 0 & \frac{(1-m_i)}{2} \end{pmatrix} \quad (8.71)$$

### Mean-field energy term

Let us consider the Hamiltonian

$$\langle \mathcal{H} \rangle_{\rho_{MF}} = \left\langle -J \sum_{\langle ij \rangle} S_i S_j - \sum_i H_i S_i \right\rangle_{\rho_{MF}} = -J \sum_{\langle ij \rangle} \langle S_i S_j \rangle_{\rho_{MF}} - \sum_i H_i \langle S_i \rangle_{\rho_{MF}} \quad (8.72)$$

Since

$$\rho_{MF} = \prod_{i=1}^N \rho_i \quad (8.73)$$

the term  $\langle S_i S_j \rangle_{\rho_{MF}}$  will transform into

$$\langle S_i S_j \rangle_{\rho_{MF}} = \langle S_i \rangle_{\rho_{MF}} \langle S_j \rangle_{\rho_{MF}} \quad (8.74)$$

Moreover, for all function  $g$  of  $S_i$  we can write

$$\begin{aligned} \langle g(S_i) \rangle_{\rho_{MF}} &= \text{Tr}^{(i)}(g(S_i)\rho_i) = \sum_{S_i=\pm 1} g(S_i)\rho_i \\ &= \sum_{S_i=\pm 1} g(S_i) \left[ \frac{1+m_i}{2}\delta_{S_i,1} + \frac{1-m_i}{2}(1 - \delta_{S_i,1}) \right] \\ &= \frac{1+m_i}{2}g(1) + \frac{1-m_i}{2}g(-1) \end{aligned} \quad (8.75)$$

Note that, if  $g(S_i) = S_i$ ,

$$\langle S_i \rangle_{\rho_{MF}} = m_i$$

as expected.

Hence,

$$\langle \mathcal{H} \rangle_{\rho_{MF}} = -J \sum_{\langle ij \rangle} m_i m_j - \sum_i H_i m_i \quad (8.76)$$

*Remark.* This has the form of the original Hamiltonian where  $S_i$  have been replaced by their statistical averages.

The entropy term is:

$$\begin{aligned} \langle \ln \rho \rangle_{\rho_{MF}} &= \text{Tr}(\rho \ln \rho) \stackrel{MF}{=} \sum_i \text{Tr}^{(i)}(\rho_i \ln \rho_i) \\ &= \sum_i \left[ \frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \right] \end{aligned} \quad (8.77)$$

The total free energy becomes:

$$\begin{aligned} F_{\rho_{MF}} &= \langle \mathcal{H} \rangle_{\rho_{MF}} + k_B T \langle \ln \rho \rangle_{\rho_{MF}} \\ &= -J \sum_{\langle ij \rangle} m_i m_j - \sum_i H_i m_i + k_B T \sum_i \left[ \frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \right] \end{aligned} \quad (8.78)$$

We now look for the values  $m_i = \bar{m}_i$ , that minimizes  $F_{\rho_{MF}}$  (equilibrium phases):

$$\left. \frac{\partial F_{\rho_{MF}}}{\partial m_i} \right|_{m_i = \bar{m}_i} = 0 \quad (8.79)$$

This gives:

$$0 = -J \sum_{j \in n.n. \text{ of } i} \bar{m}_j - H_i + \frac{k_B T}{2} \ln \left[ \frac{1 + \bar{m}_i}{1 - \bar{m}_i} \right] \quad (8.80)$$

To solve it, remember that

$$\tanh^{-1}(x) = \frac{1}{2} \ln \frac{1+x}{1-x} \quad (TO DO) \quad (8.81)$$

Hence,

$$k_B T \tanh^{-1}(\bar{m}_i) = J \sum_{j \in n.n. \text{ of } i} \bar{m}_j + H_i \quad (8.82)$$

which implies

$$\bar{m}_i = \tanh \left[ (k_B T)^{-1} \left( J \sum_{j \in n.n. \text{ of } i} \bar{m}_j + H_i \right) \right] \quad (8.83)$$

Defining

$$z \bar{m}_i \equiv \sum_{j \in n.n. \text{ of } i} \bar{m}_j \quad (8.84)$$

we get

$$\bar{m}_i = \tanh [\beta (J z \bar{m}_i + H_i)] \quad (8.85)$$

this is the Bragg-William approximation.

**Example 22** (Ising antiferromagnet in an external field). Consider the model

$$\mathcal{H} = +J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i, \quad (8.86)$$

*Remark.* Note the + before  $J$ . This means that the interactions are antiferromagnetic.

- If  $H = 0$  ferromagnetic and antiferromagnetic behave similarly when the interactions are between nearest neighbours on a *bipartite lattice*, i.e. a lattice that can be divided into two sublattices, say  $A$  and  $B$ , such that a  $A$  site has only  $B$  neighbours and a  $B$  site only  $A$  ones.

*Remark.* FCC is not bipartite, while BCC it is. See Figure 8.5.

If the lattice is bipartite and  $J_{ij}$  is non zero only when  $i$  and  $j$  belong to different sublattices (they do not have to be only n.n.!), one can redefine the spins such that

$$S'_j \begin{cases} +S_j & j \in A \\ -S_j & j \in B \end{cases} \quad (8.87)$$



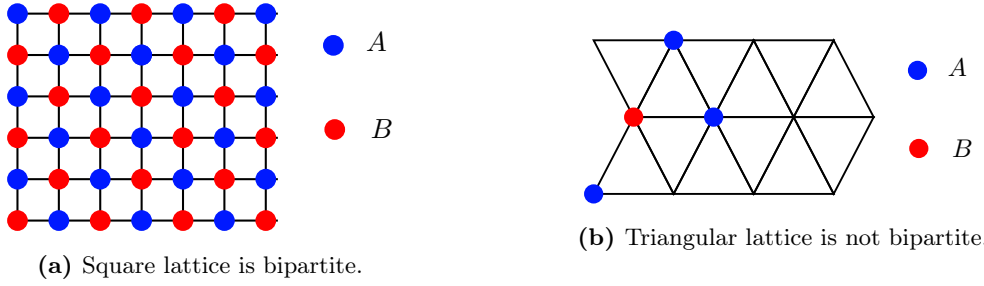


Figure 8.5

Clearly,  $S'_i S'_j = -S_i S_j$ . It is like if the  $J_{ij}$  have changed sign and we are formally back to ferromagnetic model for the two sublattices:

$$\mathcal{H}^* = -J \sum_{\langle ij \rangle} S'_i S'_j \quad (8.88)$$

i.e. a ferromagnetic Ising.

- In presence of a magnetic field  $H$ , we need to reverse its sign when applied to sites  $B$ .

The thermodynamic of a ferromagnetic Ising model on a bipartite lattice in a uniform magnetic field  $H$  is identical to the one of the Ising antiferromagnetic model in presence of the so called *staggered field*, i.e.  $H_A = H$  and  $H_B = -H$ .

$$\mathcal{H}^*[S] = -J \sum_{\langle r_A r_B \rangle} S(r_A) S(r_B) - H \sum_{r_A} S(r_A) + H \sum_{r_B} S(r_B), \quad J > 0, H > 0 \quad (8.89)$$

The average magnetization per spin is

$$m \equiv \frac{1}{2}(m_A + m_B) \quad (8.90)$$

while

$$m_S = \frac{1}{2}(m_A - m_B) \quad (8.91)$$

is the *staggered magnetization*.

In order to use the variational density matrix method for this problem we consider two independent variational parameters  $m_A$  and  $m_B$  for sublattice  $A$  and  $B$  respectively. On each sublattice, the model is like the standard Ising

$$\begin{cases} \rho_A^{(1)}(S) = \frac{1+m_A}{2} \delta_{S,1} + \frac{1-m_A}{2} \delta_{S,-1} \\ \rho_B^{(1)}(S) = \frac{1+m_B}{2} \delta_{S,1} + \frac{1-m_B}{2} \delta_{S,-1} \end{cases} \quad (8.92)$$

*Remark.* Note that, being  $H$  uniform,  $\langle S_i \rangle = m$ , i.e. does not depend on  $i$ . Same for the 1-particle distribution functions  $\rho_A^{(1)}(S)$  and  $\rho_B^{(1)}(S)$ .

By performing the calculation for the terms

$$\langle \mathcal{H} \rangle_{\rho_{MF}} = -J \sum_{\langle ij \rangle} \langle S_i S_j \rangle_{\rho_{MF}} - H \sum_i \langle S_i \rangle_{\rho_{MF}} \quad (8.93)$$

$$\langle \ln \rho \rangle_{\rho_{MF}} = \sum_i \text{Tr}^{(1)}(\rho_i \ln \rho_i) \quad (8.94)$$

as before, but remembering to partition the procedure into the two sublattices  $A$  and  $B$ , one can show (TO DO) that the variational free energy is given by

$$\frac{F(m_A, m_B)}{N} = \frac{z\hat{J}}{2}m_A m_B - \frac{1}{2}H(m_A + m_B) - \frac{1}{2}k_B T s(m_A) - \frac{1}{2}k_B T s(m_B) \quad (8.95)$$

where the entropy is given by

$$s(m) = \left[ \frac{1+m}{2} \ln \left( \frac{1+m}{2} \right) + \frac{1-m}{2} \ln \left( \frac{1-m}{2} \right) \right] \quad (8.96)$$

By differentiating  $\frac{F}{N}$  with respect to  $m_A$  and  $m_B$ , one gets

$$\frac{\partial(F/N)}{\partial m_A} = 0 \quad \Rightarrow m_B = \frac{H}{z\hat{J}} - \frac{k_B T}{z\hat{J}} \ln \left( \frac{1+m_A}{1-m_A} \right) \quad (8.97a)$$

$$\frac{\partial(F/N)}{\partial m_B} = 0 \quad \Rightarrow m_A = \frac{H}{z\hat{J}} - \frac{k_B T}{z\hat{J}} \ln \left( \frac{1+m_B}{1-m_B} \right) \quad (8.97b)$$

As before, since

$$\tanh^{-1}(x) = \frac{1}{2} \ln \frac{1+x}{1-x} \quad (8.98)$$

these self-consistent equations can be written as

$$\begin{cases} m_A = \tanh \left( \beta \left( H - z\hat{J}m_B \right) \right) \\ m_B = \tanh \left( \beta \left( H - z\hat{J}m_A \right) \right) \end{cases} \quad (8.99)$$

The sites  $\in A$  experience an internal field  $H_{A,MF} = -z\hat{J}m_B$  from the  $B$  neighbours and viceversa for the sites  $\in B$ .

### 8.2.3 Second approach: Blume-Emery-Griffith model

We apply this approach to the so called Blume-Emery-Griffith model. This is a spin model with vacancies that describes the phase diagram and the critical properties of an interacting system displaying a *tricritical point*. Perhaps the most famous of these systems is the  $\text{He}^3 - \text{He}^4$  mixture undergoing a fluid-superfluid transition.

*Remark.*  $\text{He}^4$  is a non radiative isotope with two protons and two neutrons. Roughly  $1/4$  of the universe matter is  $\text{He}^4$ ! From a quantum statistical point of view  $\text{He}^4$  is a *boson*.

A gas of  $\text{He}^4$  undergoes a fluid-superfluid transition at  $T_\lambda = 2.17K$  and a  $P = P_0$ . It is known as  $\lambda$ -transition since at  $T \sim T_\lambda$  the specific heat  $C(T)$  behaves as in Figure 8.6a.

*Remark.* The  $\lambda$ -transition is a genuine critical point (second order). For  $T < T_\lambda$ ,  $\text{He}^4$  is in the superfluid phase and it can be described by a two-fluids model in which one component has zero viscosity and zero entropy.

Now, the question is: what happens to the system if a given amount of  $\text{He}^3$  is inserted to form a  $\text{He}^3 - \text{He}^4$  mixture?  $\text{He}^3$  is a non-radiative isotope with 2 protons and 1 neutron. From a quantum statistical point of view it is a *fermion*.

Hence, if inserted in a system of  $\text{He}^4$  it will "dilute" its bosonic property. Then, one expects that  $T_\lambda$  decreases. Indeed, denoting by  $x$  the concentration of  $\text{He}^3$  one observes

$$T_\lambda = T_\lambda(x) \quad (8.100)$$

with  $T_\lambda(x)$  that decreases as  $x$  increases.

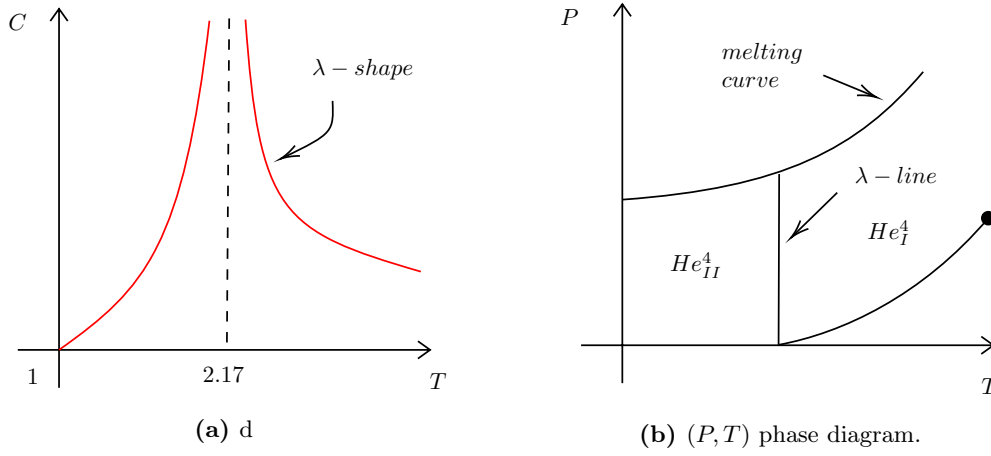


Figure 8.6

At the same time, at a given point the mixture undergoes a separation between a phase rich and a phase poor of  $He^3$ . In particular one observes that, for

$$x > x_t = \frac{n_3}{n_3 + n_4} \sim 0.67 \quad (8.101)$$

the fluid-superfluid transition becomes first order! It is accompanied by a phase separation. The point  $(x_T, T_t)$  is a *tricritical point*, i.e. a critical point that separates a line of second order transition from a line of first order transition.

### BEG Model

The BEG Model is the model of a diluted ferromagnetic system. The spins are  $S_i = \pm 1, 0$  (similar to a lattice gas model), we have  $S_i = \pm 1$  for  $He^4$  atom at site  $i$ ,  $S_i = 0$  for  $He^3$  atom at site  $i$ .

Let us consider:

- $\langle S_i \rangle = m_i$ , order parameter.
- $\langle S_i^2 \rangle$  is the density  $He^4$  atoms.

Let us define the density of  $He^3$  atoms as

$$x \equiv 1 - \langle S_i^2 \rangle \quad (8.102)$$

The chemical potentials difference is

$$\Delta \propto \mu_3 - \mu_4 \quad (8.103)$$

and controls the number of  $He^3$  atoms.

If:

- $\Delta \rightarrow -\infty \Rightarrow x \rightarrow 0$ .
- $\Delta \rightarrow +\infty \Rightarrow x \rightarrow 1$ .

and the order parameter for the  $\lambda$ -transition becomes

$$\langle S_i \rangle = \begin{cases} 0 & T > T_\lambda \\ m & T < T_\lambda \end{cases} \quad (8.104)$$

The minimal version of the model is:

$$\mathcal{H} = -J \sum_{\langle ij \rangle} S_i S_j + \Delta \sum_{i=1}^N S_i^2 - \Delta N \quad (8.105)$$

*Remark.* The  $\Delta N$  term is a typical term for a gas in gran canonical ensemble.

**Variational mean field approach to BEG**

Since  $\rho_{MF} = \prod_i \rho_i$ ,

$$G(T, J, \Delta) = \langle \mathcal{H} \rangle_{\rho_{MF}} + k_B T \sum_i \text{Tr}(\rho_i \ln \rho_i) \quad (8.106)$$

where the first term can be written as

$$\begin{aligned} \langle \mathcal{H} \rangle_{\rho_{MF}} &= -J \sum_{\langle ij \rangle} \langle S_i S_j \rangle + \Delta \sum_i \langle S_i^2 \rangle - N\Delta \\ &\stackrel{MF}{\simeq} -J \sum_{\langle ij \rangle} \langle S_i \rangle \langle S_j \rangle + \Delta \sum_i \langle S_i^2 \rangle - \Delta N \end{aligned} \quad (8.107)$$

where

$$\langle S_i \rangle = \langle S_j \rangle \equiv m \quad (8.108)$$

Hence,

$$G(T, J, \Delta)_{MF} = -\frac{1}{2} N J z (\text{Tr}_{S_i}(\rho_i S_i))^2 + N \Delta \text{Tr}_{S_i}(\rho_i S_i^2) - N \Delta + N k_B T \text{Tr}_{S_i}(\rho_i \ln \rho_i) \quad (8.109)$$

We now minimize  $G(T, J, \Delta)_{MF}$  with respect to the function  $\rho_i$  with constraint  $\text{Tr}_{S_i}(\rho_i) = 1$ :

$$\frac{dG}{d\rho_i} = 0 \quad (8.110)$$

Let us consider each term

$$\frac{d}{dp_i} (\text{Tr}(\rho_i S_i))^2 = 2(\text{Tr}(\rho_i S_i)) S_i = 2 \langle S_i \rangle S_i = 2m S_i \quad (8.111a)$$

$$\frac{d}{dp_i} (\text{Tr}(\rho_i S_i^2)) = S_i^2 \quad (8.111b)$$

$$\frac{d}{dp_i} (\text{Tr}(\rho_i \ln \rho_i)) = \ln \rho_i + 1 \quad (8.111c)$$

*Remark.* Remind that  $\rho_i = \rho^{(1)}(S_i)$ .

$$0 = -J N z m S_i + \Delta N S_i^2 + N k_B T \ln \rho_i + N k_B T \quad (8.112)$$

Dividing by  $N k_B T$ ,

$$\ln \rho_i \equiv \ln \rho^{(1)}(S_i) = \beta J z m S_i - \beta \Delta S_i^2 - 1 \quad (8.113)$$

which implies

$$\rho^{(1)}(S_i) = A^{-1} e^{\beta(z J m S_i - \Delta S_i^2)} \quad (8.114)$$

*Remark.* In  $A^{-1}$  it is included the term  $e^{-1}$ .

The constant  $A$  can be found by imposing  $\text{Tr}_{S_i} \rho^{(1)}(S_i) = 1$ , (TO DO)

$$A = 1 + 2e^{-\beta \Delta} \cosh(\beta z J m) \quad (8.115)$$

Given  $\rho^{(1)}(S_i)$  it is possible to show (TO DO)

$$\langle S_i^2 \rangle = \text{Tr}_{S_i}(\rho_i S_i^2) = \frac{1}{A} 2e^{-\beta \Delta} \cosh(\beta z J m) \quad (8.116)$$

and

$$x = 1 - \langle S_i^2 \rangle = \frac{A - 2e^{-\beta \Delta} \cosh(\beta z J m)}{A} \Rightarrow x = \frac{1}{A} \quad (8.117)$$

Hence,

$$\frac{G(T, \Delta, m, J)}{N} = \frac{z}{2} J m^2 - \Delta - k_B T \ln A \quad (8.118)$$

*Remark.* Now we should minimize  $G(T, \Delta, m, J)$  with respect to  $m$  to obtain the equilibrium phases.

The expansion for small values of  $m$  is

$$\cosh(t) = 1 + \frac{t^2}{2} + \frac{t^4}{24}, \quad \ln(1+t) = t - \frac{t^2}{2} \quad (8.119)$$

(TO DO)

$$G(T, \Delta, J, m) = a_0(T, \Delta) + a(T, \Delta)m^2 + b(T, \Delta)m^4 + \frac{c(T, \Delta)}{6}m^6 \quad (8.120)$$

where

$$\begin{cases} a(T, \Delta) = \frac{zJ}{2} \left(1 - \frac{zJ}{\delta k_B T}\right) \\ b(T, \Delta) = k \left(1 - \frac{\delta}{3}\right) \\ c(T, \Delta) > 0 \end{cases} \quad (8.121)$$

and the parameter

$$\delta \equiv 1 + \frac{e^{\beta\Delta}}{2} = \delta(T, \Delta) \quad (8.122)$$

is related to the concentration of  $\text{He}^3 - x$ . This can be seen as follows.

Since

$$x(T, \Delta, J) \equiv 1 - \langle S_i^2 \rangle = \frac{1}{A} = \left(1 + 2e^{-\beta\Delta} \cosh(\beta z J m)\right)^{-1} \quad (8.123)$$

In the disordered phase ( $m = 0$ ) one has

$$x(T, \Delta, J) = \left(1 + 2e^{-\beta\Delta}\right)^{-1} = \frac{\delta - 1}{\delta} \quad (8.124)$$

By combining this result with the order-disorder transition

$$a(T_c(\Delta)) = \frac{zJ}{2} \left(1 - \frac{zJ}{\delta k_B T_c}\right) = 0, \quad \left[\delta = \frac{zJ}{k_B T_c}\right] \quad (8.125)$$

one obtains (TO DO)

$$T_c(x) = T_c(0)(1 - x) \quad (8.126)$$

There is a dependence of the critical temperature  $\lambda$  on the  $\text{He}^3$  concentration  $x$ .

The tricritical point is the one that satisfies the conditions

$$\begin{cases} a(T_t, \Delta_t) = 0 \\ b(T_t, \Delta_t) = 0 \end{cases} \Rightarrow \begin{cases} \delta_t = \frac{zJ}{k_B T_t} \\ \delta_t = 3 \end{cases} \quad (8.127)$$

$$x(T_t, \Delta_t) = \frac{\delta_t - 1}{\delta_t} = \frac{2}{3} \quad (8.128)$$

*Remark.* Experimental estimate of  $x_t$  is  $\sim 0.67$ .

**Exercise 5.** Expand the free-energy per site

$$\frac{G}{N} = \frac{z}{2} J m^2 - \Delta - k_B T \ln A \quad (8.129)$$

where  $A = 1 + 2e^{-\beta\Delta} \cosh(\beta z J m)$  for small values of  $m$ .

$$x \equiv \beta z J m, \quad B \equiv 2e^{-\beta\Delta} \quad (8.130)$$

Since  $\cosh x \simeq 1 + \frac{x^2}{2} + \frac{x^4}{24}$ ,

$$A = 1 + B \cosh x \simeq 1 + B \left( 1 + \frac{x^2}{2} + \frac{x^4}{24} \right) \quad (8.131)$$

$$\begin{aligned} \ln A &= \ln \left( 1 + B + \frac{Bx^2}{2} + \frac{Bx^4}{24} \right) \\ &\simeq \ln \left[ (1+B) \left( 1 + \frac{B}{2(1+B)} x^2 + \frac{B}{24(1+B)} x^4 \right) \right] \\ &= \ln(1+B) + \ln(1+t) \end{aligned} \quad (8.132)$$

where

$$t \equiv \frac{B}{2(1+B)} x^2 + \frac{B}{24(1+B)} x^4 \quad (8.133)$$

Let us first consider the term

$$\frac{B}{1+B} = \frac{2e^{-\beta\Delta}}{1+2e^{-\beta\Delta}} = \frac{2}{2+e^{\beta\Delta}} = \delta^{-1} \quad (8.134)$$

$$\Rightarrow \ln A = \ln(1+B) + \frac{x^2}{2\delta} + \left( \frac{1}{24\delta} - \frac{1}{8\delta^2} \right) x^4 - \frac{1}{24\delta^2} x^6 \quad (8.135)$$

*Remark.* We have that  $x \equiv \beta z J m$ .

$$\begin{aligned} -\frac{\ln A}{B} + \frac{z}{2} J m^2 - \Delta &\simeq a_0(T, \Delta) + \left( \frac{z}{2} J - \frac{\beta z^2 J^2}{2\delta} \right) m^2 \\ &+ \left( \frac{1}{8\delta} - \frac{1}{24\delta} \right) \beta^3 z^4 J^4 m^4 + \frac{1}{24\delta^2} \beta^5 z^6 J^6 m^6 \end{aligned} \quad (8.136)$$

$$G(T, \Delta, J, m) = a_0(T, \Delta) + a(T, \Delta) m^2 + b(T, \Delta) m^4 + c(T, \Delta) m^6 \quad (8.137)$$

where

$$a(T, \Delta) = \frac{zJ}{2} \left( 1 - \frac{\beta z J}{\delta} \right) \quad (8.138a)$$

$$b(T, \Delta) = \frac{\beta^3 z^4 J^4}{8\delta} \left( \frac{1}{\delta} - \frac{1}{3} \right) = \frac{\beta^3 z^4 J^4}{8\delta^2} \left( 1 - \frac{\delta}{3} \right) \quad (8.138b)$$

$$c(T, \Delta) = \frac{\beta^5 z^6 J^6}{24\delta^2} > 0 \quad (8.138c)$$

### 8.2.4 Mean field again

Another way to introduce the variational approach and the mean field approximation often discussed starts from the general expression of the variational free energy

$$F_{var} = \langle \mathcal{H} \rangle_{\rho_{TR}} + k_B T \langle \ln \rho_{TR} \rangle_{\rho_{TR}} \quad (8.139)$$

If one assumes that the family of trial distribution is of the Gibbs-Boltzmann form

$$\rho_{TR} = \frac{e^{-\beta \mathcal{H}_{TR}}}{Z_{TR}} \quad (8.140)$$

with

$$Z_{TR} = e^{-\beta F_{TR}} = \sum_{\{\Phi_i\}} e^{-\beta \mathcal{H}_{TR}(\{\Phi_i\})} \quad (8.141)$$

then, since

$$\ln \rho_{TR} = -\beta \mathcal{H}_{TR} - \ln Z_{TR} \quad (8.142)$$

we have

$$k_B T \langle \ln \rho_{TR} \rangle_{\rho_{TR}} = k_B T \left\langle \frac{-\mathcal{H}_{TR}}{k_B T} \right\rangle + k_B T \underbrace{\langle -\ln Z_{TR} \rangle}_{\beta F_{TR}} \quad (8.143)$$

$$k_B T \langle \ln \rho_{TR} \rangle_{\rho_{TR}} = \langle -\mathcal{H}_{TR} \rangle + F_{TR} \quad (8.144)$$

$$\Rightarrow F_{var} = \langle \mathcal{H} \rangle_{\rho_{TR}} - \langle \mathcal{H}_{TR} \rangle_{\rho_{TR}} + F_{TR} = \langle \mathcal{H} - \mathcal{H}_{TR} \rangle_{\rho_{TR}} + F_{TR} \quad (8.145)$$

Clearly,  $F \leq F_{var}$  and one has to look for the minima of  $F_{var}$  by varying  $\rho_{TR}$ . Within this approach, the mean field approximation is still given by

$$\rho_{TR}^{MF}(\Phi_1, \dots, \Phi_N) = \prod_{i=1}^N \rho_{TR}^{(1)}(\Phi_i) \quad (8.146)$$

that in this case becomes

$$\prod_i \rho_{TR}^{(1)}(\Phi_i) = \frac{1}{Z_{TR}^{MF}} e^{-\beta \sum_i b_i \Phi_i} \quad (8.147)$$

and

$$Z_{TR} = \sum_{\{\Phi\}} e^{-\beta \sum_i b_i \Phi_i} \quad (8.148)$$

where  $b_i$  are the variational parameters.

$$\mathcal{H}_{TR} = - \sum_i b_i \Phi_i \quad (8.149)$$

If we consider again the Ising model

$$\mathcal{H} = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i \quad (8.150)$$

$$\begin{aligned} F_{var} &= \langle \mathcal{H} - \mathcal{H}_{TR} \rangle_{\rho_{TR}} + F_{TR} \\ &= F_{TR} + \left\langle \left( -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i \right) - \left( - \sum_i b_i S_i \right) \right\rangle_{\rho_{TR}} \\ &= F_{TR} + \left\langle -J \sum_{\langle ij \rangle} S_i S_j + \sum_i (b_i - H) S_i \right\rangle_{\rho_{TR}} \\ &= F_{TR} - J \sum_{\langle ij \rangle} \langle S_i S_j \rangle_{\rho_{TR}} + \sum_i (b_i - H) \langle S_i \rangle_{\rho_{TR}} \end{aligned} \quad (8.151)$$

Since  $\rho_{TR} = \prod_i \rho_i$ ,

$$\langle S_i S_j \rangle_{\rho_{TR}} = \langle S_i \rangle_{\rho_{TR}} \langle S_j \rangle_{\rho_{TR}} \quad (8.152)$$

$$F_{var} = F_{TR} - J \sum_{\langle ij \rangle} \langle S_i \rangle_{\rho_{TR}} \langle S_j \rangle_{\rho_{TR}} + \sum_i (b_i - H) \langle S_i \rangle_{\rho_{TR}} \quad (8.153)$$

$$\frac{\partial F_{var}}{\partial b_i} = 0, \quad \forall i \quad (8.154)$$

$$0 = \frac{\partial F_{var}}{\partial b_i} = \left[ -J \sum_{j \text{ n.n. } i} \langle S_i \rangle_{\rho_{TR}} + b_i - H \right] \frac{\partial \langle S_i \rangle}{\partial b_i} \quad (8.155)$$

or

$$b_i = J \sum_{j \text{ n.n. } i} \langle S_j \rangle_{\rho_{TR}} + H \quad (8.156)$$

$$\begin{aligned} \langle S_i \rangle_{\rho_{TR}} &= \frac{1}{Z_{TR}} \sum_{\{S\}} S_i e^{\beta \sum_k S_k b_k} = \frac{\prod_k \sum_{S_k} S_k e^{\beta S_k b_k}}{\prod_k \sum_{S_k} e^{\beta S_k b_k}} \\ &= \frac{\sum_{S_i} S_i e^{\beta S_i b_i}}{\sum_{S_i=\pm 1} e^{\beta S_i b_i}} = \frac{\cosh(\beta b_i)}{\sinh(\beta b_i)} \\ &= \tanh(\beta b_i) \end{aligned} \quad (8.157)$$

$$b_i = J \sum_{j \text{ n.n. } i} \tanh(\beta b_j) + H \quad (8.158)$$

**Exercise 6.** Consider again the antiferromagnetic Ising model

$$\mathcal{H}[\{S\}] = -J \sum_{\langle \vec{r}_A \vec{r}_B \rangle} S(\vec{r}_A) S(\vec{r}_B) - H \sum_{\vec{r}_A} S(\vec{r}_A) + H \sum_{\vec{r}_B} S(\vec{r}_B) \quad (8.159)$$

where  $J > 0$  and  $H > 0$ .

- $\vec{r}_A$  denotes the site on the  $A$  sublattice.
- $\vec{r}_B$  denotes the site on the  $B$  sublattice.

Let us find again the mean-field solution but now using the variational ansatz

$$F \leq F_{var} = \langle \mathcal{H} \rangle_{\rho_{TR}} - \langle \mathcal{H}_{TR} \rangle_{\rho_{TR}} + F_{TR} = \langle \mathcal{H} - \mathcal{H}_{TR} \rangle_{\rho_{TR}} + F_{TR} \quad (8.160)$$

*Remark.* Since the problem can be splitted in two sublattices, it is convenient to use

$$\mathcal{H}_{TR} = -H_A \sum_{r_A} S(r_A) - H_B \sum_{r_B} S(r_B) \quad (8.161)$$

- Show that  $F_{var}$  has the following expression:

$$\begin{aligned} F_{var} &= F_{TR}(\beta H_A, \beta H_B) - 4NJ \langle S_A \rangle_{\rho_{TR}} \langle S_B \rangle_{\rho_{TR}} \\ &\quad - \frac{1}{2}NH \left( \langle S_A \rangle_{\rho_{TR}} - \langle S_B \rangle_{\rho_{TR}} \right) + \frac{1}{2}N \left( H_A \langle S_A \rangle_{\rho_{TR}} + \langle S_B \rangle_{\rho_{TR}} \right) \end{aligned} \quad (8.162)$$

where

$$\langle S_A \rangle_{\rho_{TR}} \equiv m_A + n \quad (8.163a)$$

$$\langle S_B \rangle_{\rho_{TR}} \equiv m_B - n \quad (8.163b)$$

with  $m = m_A + m_B$ , and

$$m_A = \tanh(\beta H - 4\beta J m_B) \quad (8.164a)$$

$$m_B = \tanh(\beta H - 4\beta J m_A) \quad (8.164b)$$

- Expand the free energy  $F_{var}$  in powers of  $m$  of the form

$$F_{var} = A + Bm^2 + cm^4 + O(m^6) \quad (8.165)$$

and find the explicit expression of  $A, B$  and  $C$  as a function of  $T, H$  and  $n$ .



# Conclusions



# Bibliography

- [1] Luca Peliti *Statistical Mechanics in a Nutshell*.
- [2] Wikipedia [https://en.wikipedia.org/wiki/Lever\\_rule](https://en.wikipedia.org/wiki/Lever_rule).
- [3] Herbert B. Callen *Thermodynamics and an introduction to thermostatistics, second edition*.
- [4] J.M.Yeomans *Statistical Mechanics of Phase Transitions*.
- [5] L. Onsanger. *Phy.Rev.*65, 117 (1944).
- [6] T.D Schultz et al. *Rev.Mad.Phys.*36, 856 (1964).
- [7] R.J. Baxter and I.G. Enting. *J. Phys. A* 35, 5189, (1978).
- [8] Griffiths *Phy.Rev* 65, 117(1944).