LECTURE NOTES OF STATISTICAL MECHANICS

Collection of the lectures notes of professor Orlandini.

Edited By

Alice Pagano

The University of Padua

Compiled: Thursday 26th December, 2019

Contents

Introduction **vii**

1 Recall of Thermodynamics **1**
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

2 Equilibrium phases and thermodynamics of phase transitions **13**
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

3 Recall of statistical mechanics and theory of ensambles **33**
3.1 Statistical ensambles . . . 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

4 Statistical mechanics and phase transitions **41**

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

5 Role of the models in statistical mechanics 49
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 neirest 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 Ω 56
5.4.1 From the continuous to the lattice gas model 57

6 Some exactly solvable models of phase transitions 59
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

7 The role of dimension, symmetry and range of interactions in phase transitions 79
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

8 Mean field theories of phase transitions and variational mean field 91
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

9 Non ideal fluids: Mean field theory, Van der Walls, Virial expansion and Cluster expansion 113
9.1 Mean field theory for fluids . . . 113
9.2 Equation of state and critical point . . . 115
9.2.1 Law of the correspondent states . . . 116
9.3 Region of coexistence and Maxwell construction . . . 116
9.3.1 Equal area or Maxwell construction . . . 117
9.4 Critical exponents of Van der Walls . . . 118
9.5 Theory of weakly interacting fluids . . . 120
9.6 Van der Walls and virial expansion . . . 121
9.6.1 Boyle's temperature T_B . . . 121
9.7 Cluster expansion technique for weakly interacting gases . . . 122
9.7.1 Computation of B_2 for given Φ . . . 124
9.7.2 Higher order terms in the cluster expansion . . . 126

10 Landau theory of phase transition for homogeneous systems 131
10.1 Landau theory of phase transition (uniform systems) . . . 131
10.2 Symmetries . . . 131
10.3 Equilibrium phases . . . 132
10.4 Critical exponents in Landau's theory . . . 134
10.5 Cubic term and first transition . . . 136
10.6 Phase stability and behaviour of $\chi_T \equiv \frac{\partial \eta}{\partial h}$ . . . 139
10.6.1 Computation of T^{**} . . . 139
10.6.2 Computation of T^* . . . 140
10.7 Landau theory and multicritical points . . . 140
10.8 Landau-de Gennes theory of liquid crystals . . . 141
10.8.1 LC-phases . . . 142
10.8.2 Order parameter . . . 142
10.8.3 Landau free energy . . . 143

11 Role of fluctuations in critical phenomena: Ginzburg criterium, Coarse-graining and Ginzburg-Landau theory of phase 145
11.1 Importance of fluctuations: the Ginzburg criterium . . . 145
11.2 Fluctuation-dissipation relation . . . 147
11.2.1 Computation of E_R . . . 147
11.2.2 Estimation of E_{TOT} as $t \to 0^-$ . . . 148
11.3 Spatial variations: coarse graining procedure and the Ginzburg-Landau model . . . 148
11.3.1 Computation of $\mathcal{H}_{eff}[m]$ . . . 150
11.3.2 Magnetic non homogeneous field . . . 152
11.4 Saddle point approximation and Landau theory for non-homogeneous systems . . . 153
11.5 Nono homogeneous systems and correlation functions within a mean field approximation . . . 155
11.5.1 Solution by Fourier transform . . . 156
11.6 Including fluctuations at the Gaussian level (non interacting fields) . . 158
11.7 Ginzburg-Landau in the Gaussian approximation . . . 159
11.7.1 Useful relations . . . 159
11.8 Gaussian Hamiltonian in Fourier space . . . 160
11.8.1 Free energy in Gaussian approximation . . . 161
11.8.2 Specific heat in the Gaussian approximation . . . 162

11.8.3 How about ν? . . . 163

12 Widom's scaling theory. Block-spin Kadanoff's transformation 165
12.1 Static scaling hypothesis (Widom) . . . 165
12.1.1 Homogeneous functions . . . 166
12.2 Generalized homogeneous functions . . . 167
12.3 Widom's theory of static scaling . . . 167
12.3.1 Scaling of the magnetization . . . 168
12.3.2 Equation of state . . . 169
12.3.3 Magnetic susceptibility . . . 169
12.3.4 Scaling of the specific heat . . . 170
12.3.5 Second form of the static scaling hypothesis . . . 170
12.4 Kadanoff's block spin and scaling of the correlation function . . . 170
12.4.1 Kadanoff's argument . . . 171
12.5 Kadanoff's argument applied to the two-point correlation . . . 173

13 Renormalization group theory. Universality 175
13.1 Renormalization group (RG) . . . 175
13.1.1 Main goals of RG . . . 175
13.1.2 Singular behaviour in RG . . . 177
13.1.3 Zoology of the fixed points . . . 177
13.1.4 Linearization of RG close to the fixed points and critical exponents . . . 178
13.2 RG and scaling . . . 179
13.3 Real space renormalization group on 1D Ising model ($H = 0$) . . . 181
13.4 Decimation procedure for $D > 1$ (proliferation of interactions) . . . 183

14 Spontaneous symmetry breaking 185
14.1 Spontaneous symmetry breaking . . . 185
14.2 Spontaneous breaking of continuous symmetries and the anset of Goldstone particles . . . 187
14.2.1 Quantum relativistic case (field theory) . . . 187
14.3 Interpretation in relativistic quantum mechanics . . . 189
14.4 Spontaneous symmetry breaking in Gauge symmetries . . . 191
14.5 Fiedl theory analog . . . 191
14.5.1 Non abelian gauge theories . . . 194
14.6 Extension of Higgs mechanism to non abelian theories . . . 194
14.6.1 GWS model . . . 194
14.6.2 Higgs mechanism . . . 195

Conclusions 197

Bibliography 199

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 thermodinamics.
2. Equilibrium phases and thermodynamics of the phase transitions.
3. Statistical mechanics and theory of ensambles.
4. Thermodinamic 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 *Geinzburg criterium* and the notion of the *upper critical dimension*.
14. The Ginzburg-Landau model.
15. Landau theory for non-homogeneous system. The ν exponent.
16. Gaussian fluctuations in the G-L theory.
17. Widom's scaling theory.
18. Kadauoff'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

Lecture 1.
Wednesday 9th October, 2019.
Compiled: Thursday 26th December, 2019.

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 dreedom. 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 *assiomatic 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 caracterize the system at equilibrium are the internal energy U, volume V, number of particles N and magnetization $\vec{\mathbf{M}}$ that "scale with the system". In general, the extensive variable are additive.

In thermodynamic, it is important the concept of walls and thermodynamic constrains 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 thermodynamcs 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 $\vec{\mathbf{M}}$, etc .... If these variables are time indipendent, 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.

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 Δt the system is at equilibrium. Therefore, consider a sequence of equilibrium states, the **quasi-static transormation** are described by the 1st *Law of Thermodynamic*:

$$\mathrm{d}U = \delta Q - \delta W \tag{1.1}$$

The variation of the internal energy of the systems depends by two factors, δw that is the work done by the system during a *quasi-static process* (infinitively slow), and δQ that is the heat absorbed by the system during the process. Remember that we write $\mathrm{d}U$ because it is a differential quantity, while the other quantities with the δ are only small quanties. Therefore, $\mathrm{d}U$ 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 $\delta W = P\,\mathrm{d}V$. For a magnetized system, we have $\delta W = -\vec{\mathbf{H}} \cdot \mathrm{d}\vec{\mathbf{M}}$.

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 constrains.

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

$$S = S(U, V, N) \tag{1.2}$$

The new values taken by the extensive parameters when a constraint has been removed are the ones that *maximize* S. It means $\mathrm{d}S = 0$ and $\mathrm{d}^2S < 0$, 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:
$$S = \sum_{\alpha} S^{(\alpha)} \tag{1.3}$$
2. S is differentiable and *monotonically increasing* with respect to the internal energy U. It means that $\left(\frac{\partial S}{\partial U}\right)_{V,N} > 0$.
3. For each subsystem (α) we have:
$$S^{(\alpha)} = S^{(\alpha)}(U^{(\alpha)}, V^{(\alpha)}, N^{(\alpha)}) \tag{1.4}$$
This fundamental relation holds for each subsystem.
4. S is an *homogeneous function* of 1st order with respect to the extensive parameters, namely:
$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N), \quad \forall \lambda > 0 \tag{1.5}$$
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

$$U = U(S, V, N) \tag{1.6}$$

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

By taking the differential of the fundamental relation

$$U = U(S, V, N_1, \dots, N_r) \tag{1.7}$$

one gets

$$\mathrm{d}U = \underbrace{\left(\frac{\partial U}{\partial S}\right)_{V,N_j}}_{\substack{T \\ \text{absolute} \\ \text{temperature}}} \mathrm{d}S + \underbrace{\left(\frac{\partial U}{\partial V}\right)_{S,N_j}}_{\substack{-P \\ \text{pressure}}} \mathrm{d}V + \sum_{j=1}^{r} \underbrace{\left(\frac{\partial U}{\partial N_j}\right)_{S,V}}_{\substack{\mu_j \\ \text{elettrochemical} \\ \text{potential}}} \mathrm{d}N_j \tag{1.8}$$

1.3 Equations of states

Now, we define another set of variables that are called **intensive variables**. The term *intensive* means that it is indipendent 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:

$$T = T(S, V, N_1, \dots, N_r) \tag{1.9a}$$
$$P = P(S, V, N_1, \dots, N_r) \tag{1.9b}$$
$$\mu_j = \mu_j(S, V, N_1, \dots, N_r) \tag{1.9c}$$

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

$$PV = NK_bT \tag{1.10}$$

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

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

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

$$M = \frac{CH}{T} \tag{1.12}$$

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

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

$$T(\lambda S, \lambda V, \lambda N) \overset{0}{=} T(S, V, N) \tag{1.13}$$

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

$$P(\lambda S, \lambda V, \lambda N) = P(S, V, N) \tag{1.14}$$

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

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

We define:

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

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

The differential is written as the following:

$$\mathrm{d}U = T\,\mathrm{d}S + \sum_{j=1}^{r+1} P_j\,\mathrm{d}X_j \tag{1.17}$$

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

From the equilibrium condition, $\mathrm{d}U = 0$, one can get a relation between intensive variables in differential form as the **Gibbo-Duhem relation**:

$$\boxed{S\,\mathrm{d}T + \sum_{j=1}^{r+1} X_j\,\mathrm{d}P_j = 0} \tag{1.18}$$

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

$$S\,\mathrm{d}T - V\,\mathrm{d}P + N\,\mathrm{d}\mu = 0 \tag{1.19}$$

and dividing by the number of moles N

$$\mathrm{d}\mu = -s\,\mathrm{d}T + v\,\mathrm{d}P \tag{1.20}$$

that is the Gibbs-Duhem relation in *molar form*.
For a magnetic system we have $\mathrm{d}U = T\,\mathrm{d}S + \vec{\mathbf{H}} \cdot \mathrm{d}\vec{\mathbf{M}} + \mu\,\mathrm{d}N$.

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

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

1.4 Legendre transform and thermodynamic potentials

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

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

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

$$X_k \to P_k \equiv \frac{\partial Y}{\partial X_k} \tag{1.22}$$

i.e. P_k substitutes X_k as a new indipendent 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 conjiugate derivative

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

One starts form the fundamental relation

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

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

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

By differentiating A we get

$$\mathrm{d}A = \mathrm{d}U - T\,\mathrm{d}S - S\,\mathrm{d}T$$

On the other hand

$$\mathrm{d}U = T\,\mathrm{d}S + \sum_j P_j\,\mathrm{d}X_j$$

It implies that

$$\mathrm{d}A = -S\,\mathrm{d}T + \sum_j P_j\,\mathrm{d}X_j$$

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 $A = A(T, \vec{\mathbf{M}}, N_1, \dots, N_r)$.

Helmholtz free energy

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

$$A \equiv U - TS \tag{1.23}$$

In terms of heat and mechanical work, since $\mathrm{d}U = \delta Q - \delta W$:

$$\mathrm{d}A = \mathrm{d}U - \mathrm{d}(TS) = \delta Q - T\,\mathrm{d}S - S\,\mathrm{d}T - \delta W \tag{1.24}$$

Hence,

$$\delta W = (\delta Q - T\,\mathrm{d}S) - S\,\mathrm{d}T - \mathrm{d}A \tag{1.25}$$

On the other hand, for a *reversible transformation* we have $\delta Q = T\,\mathrm{d}S$, which implies

$$\delta W = -S\,\mathrm{d}T - \mathrm{d}A \tag{1.26}$$

If the reversible transformation is also *isothermal*, $\mathrm{d}T = 0$ and we obtain $\mathrm{d}A = \mathrm{d}W$. It is reminiscent of a potential energy.

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

$$\delta Q \leq T\,\mathrm{d}S \tag{1.27}$$

which implies

$$(\delta W)_{irr} = \delta Q - T\,\mathrm{d}S - \mathrm{d}A \leq -\,\mathrm{d}A\,.$$

Hence, if $\delta W = 0$ and $\mathrm{d}T = 0$, we have $\mathrm{d}A \leq 0$. Therefore, in a spontaneous (irreversible) process, the thermodynamc system, as a function of T, V, N etc, evolves towards a *minimum* of the Helmoltz free energy $A = A(T, V, N_1, \ldots, N_r)$.

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

$$\mathrm{d}A = -S\,\mathrm{d}T - P\,\mathrm{d}V + \sum_j \mu_j\,\mathrm{d}N_j \tag{1.28}$$

where

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

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

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

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

$$\mathrm{d}A = -S\,\mathrm{d}T + \vec{\mathbf{H}} \cdot \mathrm{d}\vec{\mathbf{M}} + \sum_j \mu_j\,\mathrm{d}N_j \tag{1.30}$$

with

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

Heltalpy

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

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

$$H = U + PV \tag{1.32}$$

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

We have:

$$\begin{aligned} \mathrm{d}H &= \mathrm{d}U + P\,\mathrm{d}V + V\,\mathrm{d}P \\ &= T\,\mathrm{d}S - P\,\mathrm{d}V + \sum_j \mu_j\,\mathrm{d}N_j + P\,\mathrm{d}V + V\,\mathrm{d}P \\ &= T\,\mathrm{d}S + V\,\mathrm{d}P + \sum_j \mu_j\,\mathrm{d}N_j \end{aligned} \tag{1.33}$$

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

Gibbs potential

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

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

$$G = U - TS + PV = A + PV \tag{1.34}$$

For a simple fluid system

$$\begin{aligned} \mathrm{d}G &= \mathrm{d}U - T\,\mathrm{d}S - S\,\mathrm{d}T + P\,\mathrm{d}V + V\,\mathrm{d}P \\ &= T\,\mathrm{d}S - P\,\mathrm{d}V + \sum_j \mu_j\,\mathrm{d}N_j - T\,\mathrm{d}S - S\,\mathrm{d}T + P\,\mathrm{d}V + V\,\mathrm{d}P \\ &= -S\,\mathrm{d}T + V\,\mathrm{d}P + \sum_j \mu_j\,\mathrm{d}N_j \end{aligned} \tag{1.35}$$

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

For a magnetic system, the Gibbs potential is defined as

$$G = A - \vec{\mathbf{M}} \cdot \vec{\mathbf{H}} \tag{1.36}$$

and

$$\begin{aligned} \mathrm{d}G = \mathrm{d}A - \mathrm{d}\vec{\mathbf{M}} \cdot \vec{\mathbf{H}} &= \mathrm{d}(U - TS) - \mathrm{d}(-\vec{\mathbf{M}} \cdot \vec{\mathbf{H}}) \\ &= \mathrm{d}U - T\,\mathrm{d}S - S\,\mathrm{d}T - \mathrm{d}\vec{\mathbf{M}} \cdot \vec{\mathbf{H}} - \vec{\mathbf{M}} \cdot \mathrm{d}\vec{\mathbf{H}} \\ &= T\,\mathrm{d}S + \vec{\mathbf{H}} \cdot \mathrm{d}\vec{\mathbf{M}} - T\,\mathrm{d}S - S\,\mathrm{d}T - \mathrm{d}\vec{\mathbf{M}} \cdot \vec{\mathbf{H}} - \vec{\mathbf{M}} \cdot \mathrm{d}\vec{\mathbf{H}} \\ &= -S\,\mathrm{d}T - \vec{\mathbf{M}} \cdot \mathrm{d}\vec{\mathbf{H}} \end{aligned} \tag{1.37}$$

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

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

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

Grancanonical potential

Lecture 2. Friday 11th October, 2019. Compiled: Thursday 26th December, 2019.

The two intensive variables to became indipendent are T and μ. The corresponding Legendre transform is

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

Differentiating this relation we obtain:

$$\begin{aligned} \mathrm{d}\Omega &= \mathrm{d}U - S\,\mathrm{d}T - T\,\mathrm{d}S - \sum_{ij} \mathrm{d}\mu_j\, N_j - \sum_{i=1}^{r} \mu_i\,\mathrm{d}N_i \\ &= (\delta Q - T\,\mathrm{d}S) - \delta W - S\,\mathrm{d}T - \sum_{j=1}^{r} \mathrm{d}\mu_j\, N_j - \sum_{j=1}^{r} \mu_j\,\mathrm{d}N_j \end{aligned} \tag{1.40}$$

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

1.5 Maxwell relations

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

$$U = TS - PV + \sum_j \mu_j N_j \tag{1.41}$$

Instead, the **Maxwell relations** are relations between the mixed derivatives of the thermodynamic potentials. They can be obtained from the expressions of $\mathrm{d}U$, $\mathrm{d}H$, $\mathrm{d}A$, $\mathrm{d}G$ and $\mathrm{d}\Omega$ 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 $\frac{t(t+1)}{2}$ indipendent mixed derivatives.

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

$$\mathrm{d}U = T\,\mathrm{d}S - P\,\mathrm{d}V + \mu\,\mathrm{d}N \tag{1.42}$$

where

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

It implies that

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

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 Maxweel 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

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

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

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

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

$$\mathrm{d}A = -S\,\mathrm{d}T - P\,\mathrm{d}V + \mu\,\mathrm{d}N \tag{1.44}$$

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

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

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

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

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

$$\mathrm{d}G = -S\,\mathrm{d}T - V\,\mathrm{d}P + \mu\,\mathrm{d}N \tag{1.46}$$

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

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

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

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

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 variatious of thermodynamic variables. Any osservation is just the pertubation 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.*

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

- *Adiabatic compressibility.*

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

- *Isothermal compressibility.*

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

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

- *Molar heat capacity at constant pressure.*

$$c_P = \left(\frac{\delta Q}{\mathrm{d}T}\right)_{P,N} = T\left(\frac{\partial S}{\partial T}\right)_{P,N} \underset{-S=\left(\frac{\partial G}{\partial T}\right)_{P,N}}{=} -T\left(\frac{\partial^2 G}{\partial T^2}\right)_{P,N} \tag{1.51}$$

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

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

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

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

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

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

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

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

$$\mathrm{d}G = -S\,\mathrm{d}T + V\,\mathrm{d}P$$

and the response function α_P permit to write

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

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

$$\mathrm{d}A = -S\,\mathrm{d}T - P\,\mathrm{d}V + \mu\,\mathrm{d}N$$

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

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

Hence

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

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

$$\begin{cases} c_V \equiv \left(\frac{\delta Q}{\mathrm{d}T}\right)_V \geq 0 \\ c_P \equiv \left(\frac{\delta Q}{\mathrm{d}T}\right)_P \geq 0 \end{cases} \tag{1.58}$$

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

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

Similar considerations for a magnetic system, gives

$$c_H \geq 0, \quad c_M \geq 0, \quad \chi_M \geq 0 \tag{1.60}$$

Remark. In diamangetic systems χ_M can also be negative.

Exercise 1. By using Maxwell relations show that

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

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

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

$$\mathrm{d}S = \left(\frac{\partial S}{\partial T}\right)_V \mathrm{d}T + \left(\frac{\partial S}{\partial V}\right)_T \mathrm{d}V$$

Dividing by $\mathrm{d}T$ 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

$$\begin{cases} c_P \geq c_V \geq 0 \\ c_H \geq c_M \geq 0 \end{cases} \tag{1.62}$$

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

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

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 α the phase of a one-component system (say α = gas or liquid). Therfore, 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 α and β (for example liquid and solid). Consider the surface g_α and g_β, 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 α. Therefore there will be regions in (T, P) space were the most stable phase will be α and others in which it will be β. If we now plot the values of g as a function of T and T 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:

$$g_\alpha(T,P) = g_\beta(T,P) \tag{2.2}$$

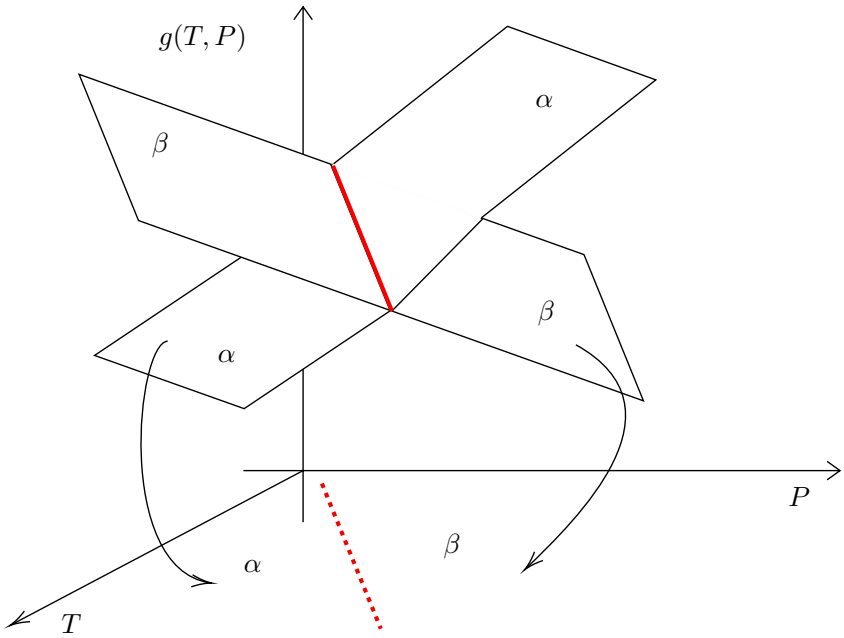


Figure 2.1: Phase diagram: stability of phases.

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.

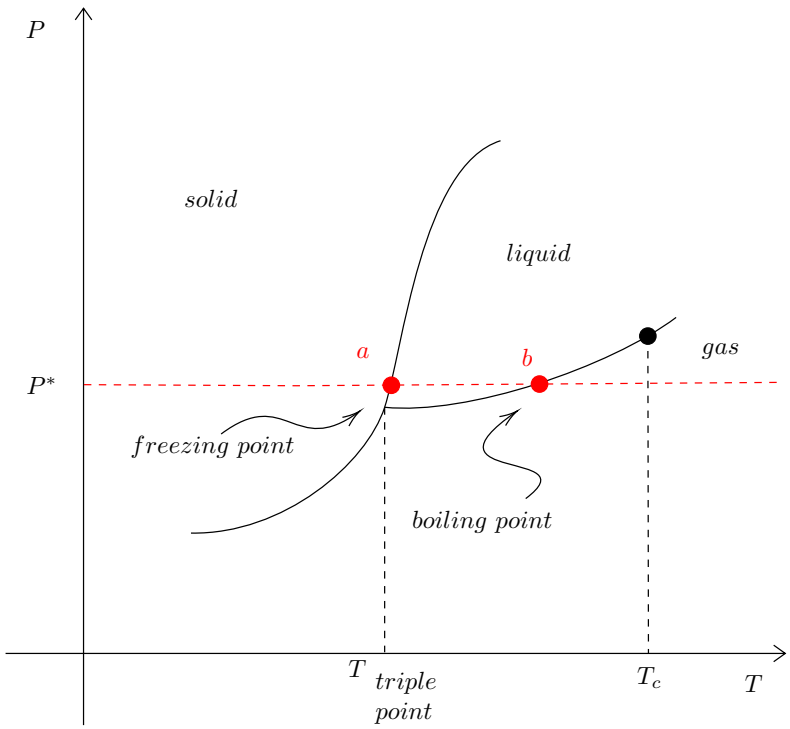


Figure 2.2: (T, P) projection.

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:

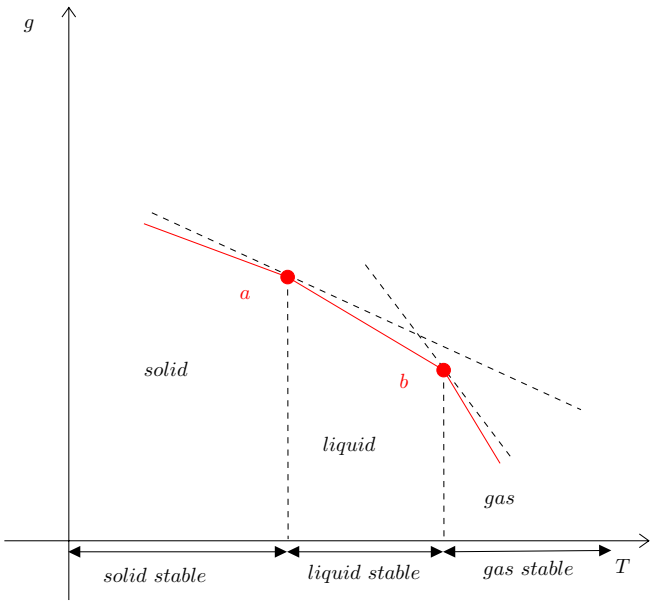


Figure 2.3: (g, T) projection at a fixed pression $P = P^*$.

- 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 ΔsT 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.

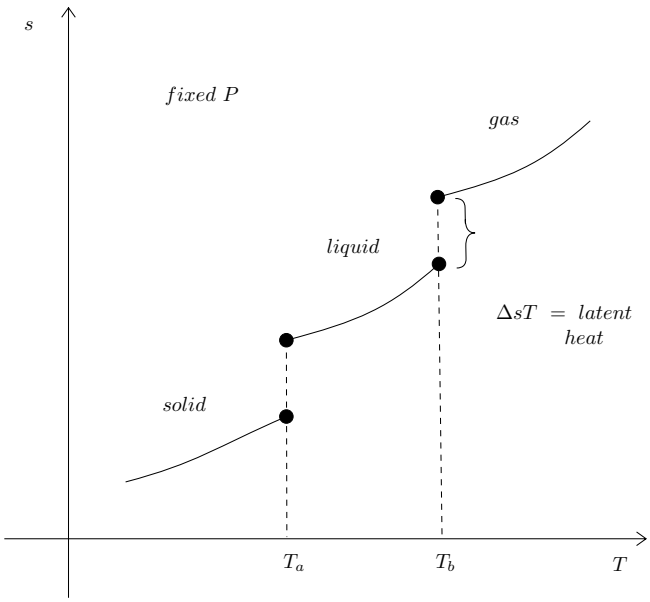


Figure 2.4: (s, T) projection.

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

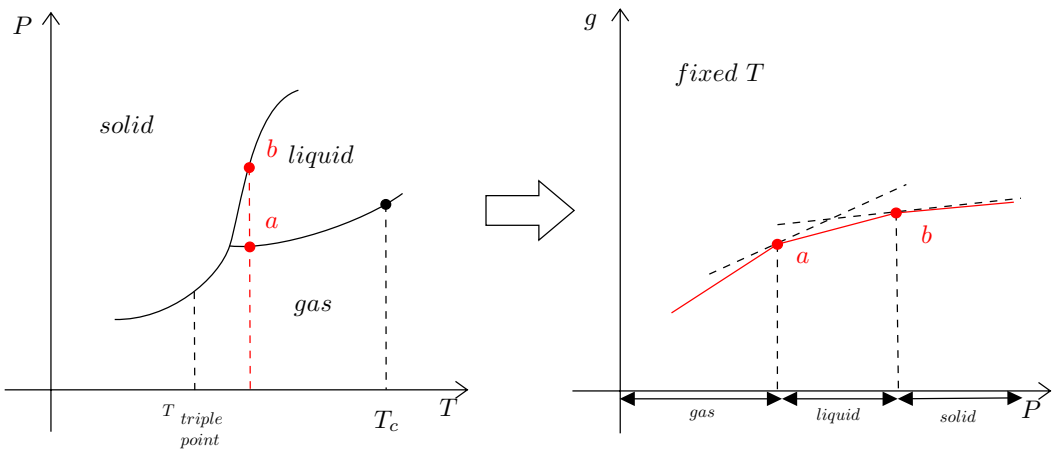


Figure 2.5: Left: (T, P) projection. Right: (g, P) projection at a fixed temperature $T = T^*$.

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

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

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.

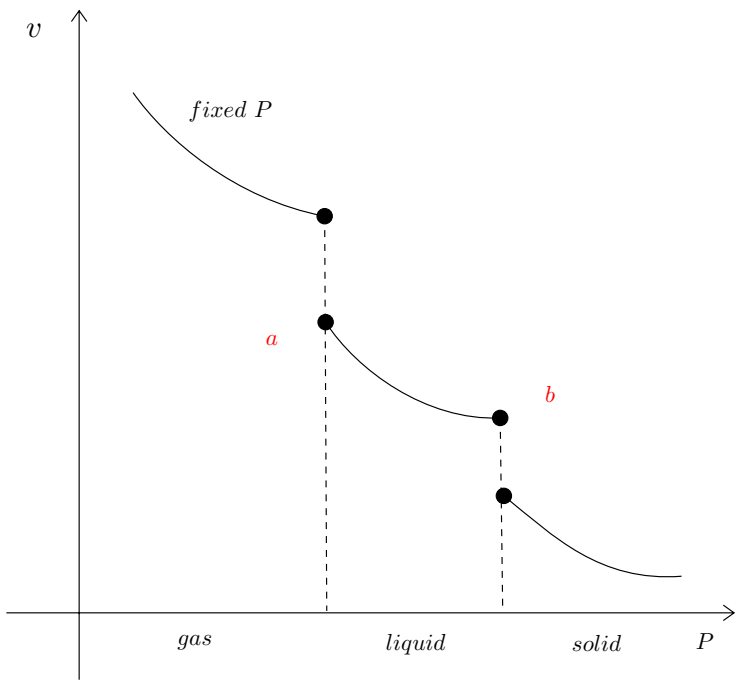


Figure 2.6: (v, P) projection.

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.

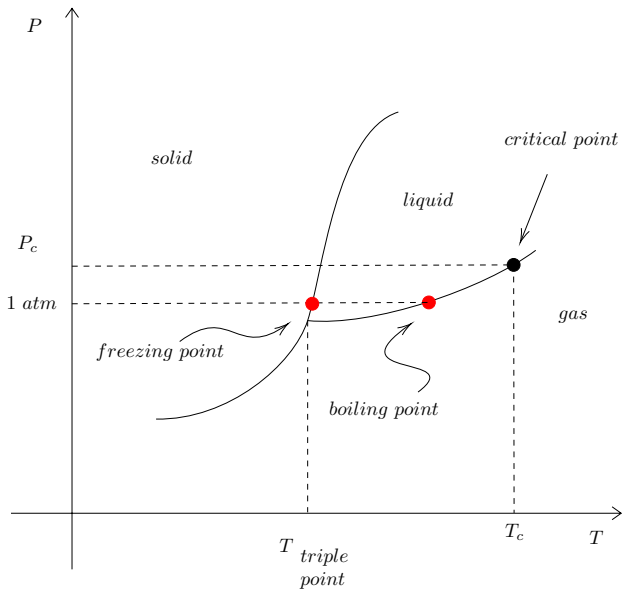


Figure 2.7: Phase diagram of a fluid. All the phase transition are first-order except at the critical point C. Beyond C it is possible to move continuously from liquid to a gas. The boundary between the solid and liquid phases is thought to be always first-oder and not to terminate in a critical point.

2.2.2 Ferromagnetic system

A similar behaviour can be encountered in magnetic systems. We can have a magnetization different from 0 even when the 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 that 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.

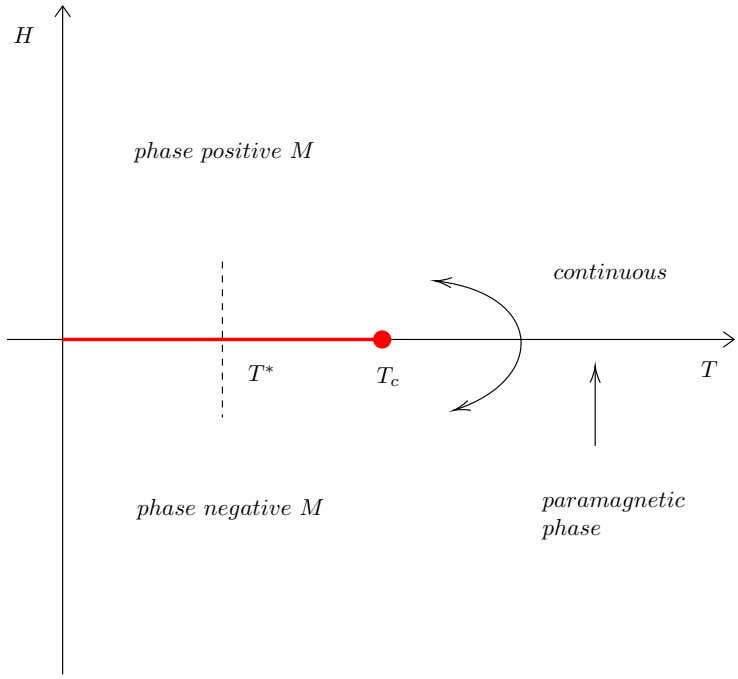


Figure 2.8: Phase diagram for a magnetic system in (T, H) space. A line of first-order transitions at zero field ends in a critical point at a temperature T_c.

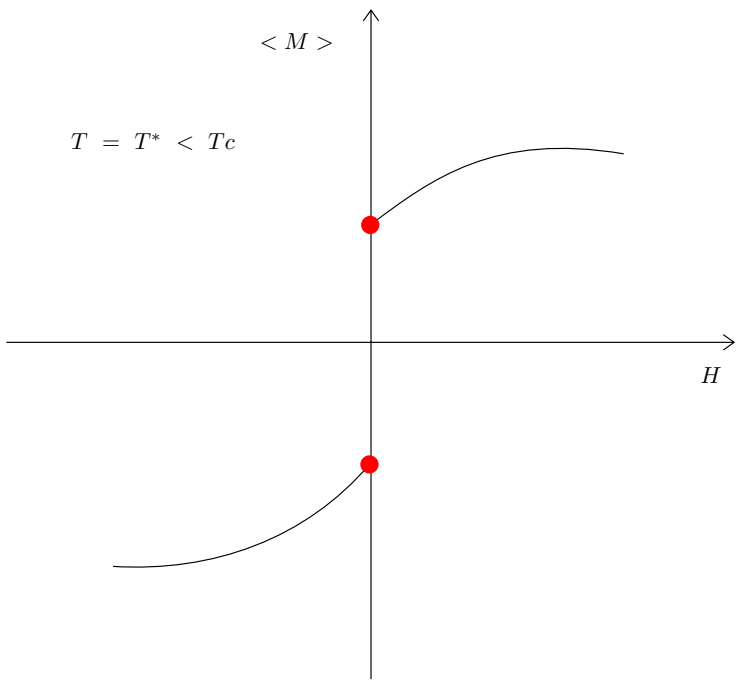


Figure 2.9: Plot of the Magnetization for $T = T^* < T_c$.

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

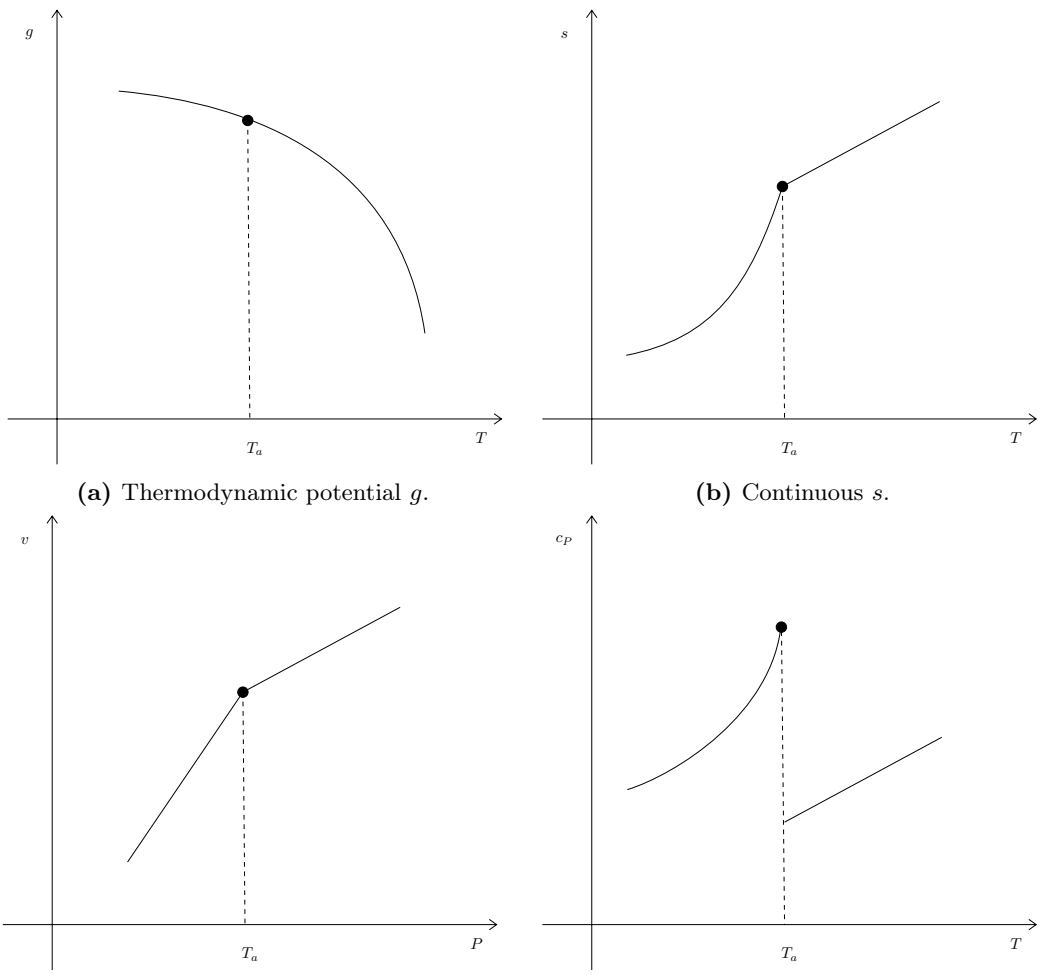
$$\left(\frac{\partial g}{\partial T}\right)_P = -s \tag{2.4a}$$

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

are continuous. This implies:

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

that is discontinuous. An example is *superconductivity*.



(a) Thermodynamic potential g.

(b) Continuous s.

(c) Continuous v.

(d) Discontinuous c_P.

Figure 2.10: Example of a second order phase transition.

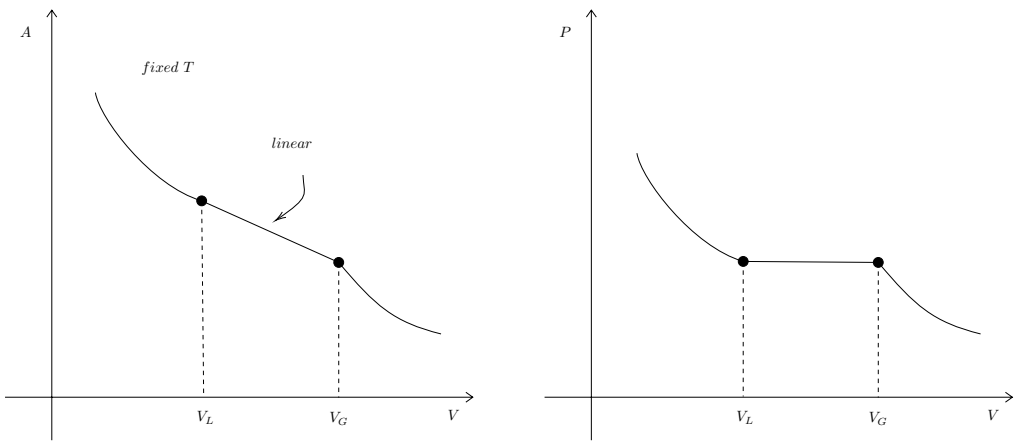
If we look for example at the specific heat c_P in Figure 2.10d, it represent the transition from superconduction.

The critical point is special beacause 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 pression remains constant. At the coexistence line we see bubbles. It is the density that is changing locally, the bubbles becames bigger and bigger and at the V_G , becames a liquid.

2.3.1 Helmholtz free-energy



(a) (V, A) projection for fixed T. **(b)** (P, V) projection.

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

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

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

Lecture 3. *Wednesday 16th October, 2019. Compiled: Thursday 26th December, 2019.*

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, α and β , which themselves contain two elements, A and B, the lever rule states that the mass fraction of the α phase is

$$w^\alpha = \frac{w_B - w_B^\beta}{w_B^\alpha - w_B^\beta} \tag{2.7}$$

where

- w_B^α: is the mass fraction of element B in the α phase.
- w_B^β: is the mass fraction of element B in the β 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 mixsture 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:

$$\frac{x_A}{x_B} = \frac{v_B - v_D}{v_D - v_A} \tag{2.8}$$

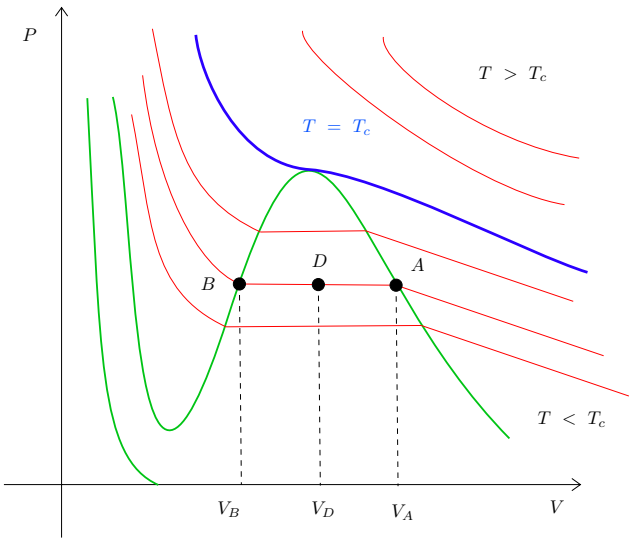


Figure 2.12: (V, P) projection. In the region between A and B the gas and the liquid phase coexist by keeping the pressure constant.

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 μ_1, μ_2. The equilibrium condition is given by the maximum of the total entropy $S = S_1 + S_2$ and gives the conditions

$$T_1 = T_2, \quad P_1 = P_2, \quad \mu_1 = \mu_2 \tag{2.9}$$

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

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

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

Therefore, on the coexistence line it should hold the relation

$$g_1(T, P) = g_2(T, P) \tag{2.11}$$

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 $\mathrm{d}P\,/\,\mathrm{d}T$ 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 ΔP were the change in pressure, then the change in temperature would be $\Delta T = \Delta P/(\mathrm{d}P\,/\,\mathrm{d}T)_{coex}$, 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.

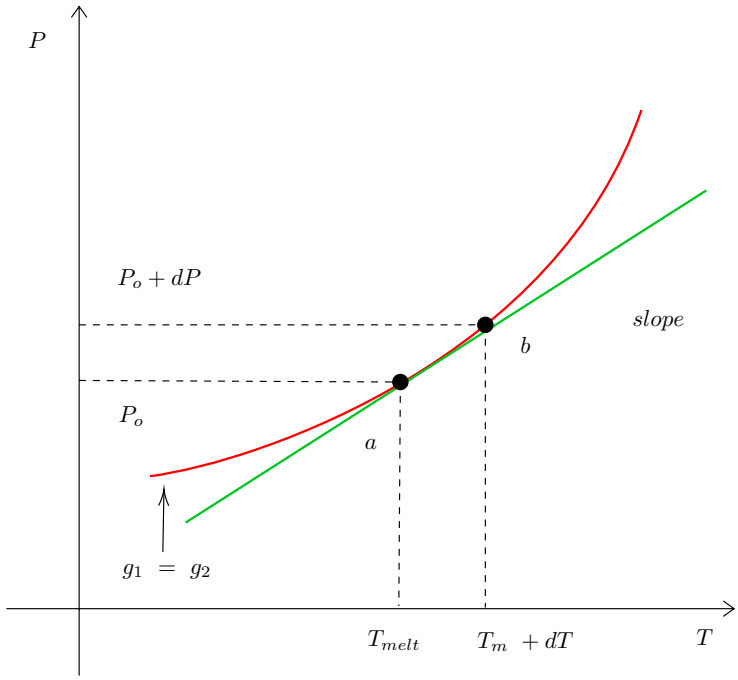


Figure 2.13: (T, P) projection. The coexistence line is represented in red, while in green the slope between the two points a and b.

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. $(\mathrm{d}P\,/\,\mathrm{d}T)$. 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

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

Hence, if a and b are very close:

$$\begin{cases} \mathrm{d}g_1 = g_1^{(b)} - g_1^{(a)} \\ \mathrm{d}g_2 = g_2^{(b)} - g_2^{(a)} \end{cases} \tag{2.13}$$

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

$$\Rightarrow \mathrm{d}g_1 = \mathrm{d}g_2 \tag{2.14}$$

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

$$\begin{cases} \mathrm{d}g_1 = -s_1\,\mathrm{d}T + v_1\,\mathrm{d}P = \mathrm{d}\mu_1 \\ \mathrm{d}g_2 = -s_2\,\mathrm{d}T + v_2\,\mathrm{d}P = \mathrm{d}\mu_2 \end{cases} \tag{2.15}$$

taking the difference, one obtains

$$-(s_2 - s_1)\,\mathrm{d}T + (v_2 - v_1)\,\mathrm{d}P = 0$$

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

$$\boxed{\left(\frac{\mathrm{d}P}{\mathrm{d}T}\right)_{coex} = \frac{(s_2 - s_1)}{(v_2 - v_1)} = \frac{\Delta s}{\Delta v}} \tag{2.16}$$

Remark. Since $(\mathrm{d}P\,/\,\mathrm{d}T)_{coex}$ is finite, the equation explains why a first order transition is characterised by discontinuous changes in entropy and volume (or density). ΔS gives the latent heat L_{12} [1]:

$$L_{12} = T\Delta s \tag{2.17}$$

whence, the Clapeyron equation is

$$\frac{\mathrm{d}P}{\mathrm{d}T} = \frac{L_{12}}{T\Delta v} \tag{2.18}$$

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:

$$\left(\frac{\mathrm{d}P}{\mathrm{d}T}\right)_{coex} = \frac{s_2 - s_1}{v_2 - v_1} \tag{2.19}$$

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

$$\left(\frac{\mathrm{d}P}{\mathrm{d}T}\right)_{coex} > 0 \quad \Rightarrow \frac{s_2 - s_1}{v_2 - v_1} > 0 \tag{2.20}$$

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

[1] The latent heat of fusion is the quantity of heat required to melt one mole of solid.

[2] "When a settled system is disturbed, it will adjust to diminish the change that has been made to it".

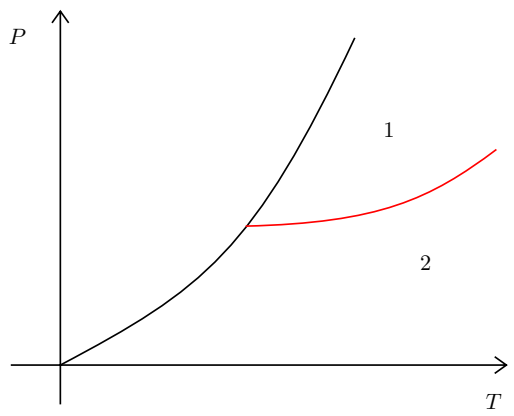


Figure 2.14: (T, P) projection. Region 1: liquid. Region 2: gas.

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 ΔV 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:

$$\left(\frac{\mathrm{d}P}{\mathrm{d}T}\right)_{coex} = \frac{\delta Q_{melt}}{T_{melt}\Delta v_{melt}} \qquad \delta Q_{melt} = Q_{liq} - Q_{solid} > 0 \tag{2.21}$$

In general, $\Delta v_m = v_{liq} - v_{solid} > 0$ which implies $(\mathrm{d}P\,/\,\mathrm{d}T)_{coex} > 0$. There are cases, however, where $\Delta v_m = v_{liq} - v_{solid} < 0$ because $\rho_{liq} > \rho_{solid}$ (for instance the H_20, 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 coxistence $(0 < T < 4)$. This implies that $\mathrm{d}P\,/\,\mathrm{d}T < 0$.

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:

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

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

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

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

Consider $P_{\text{Everest}} = 0.36atm$, $\rho(T = 100°\text{C}) = 0.598kg/m^3$, $L_{ge} = 2.257 \cdot 10^3 J/g$. 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:

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

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

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:

$$O_p = \begin{cases} \neq 0 & T < T_c \\ = 0 & T \to T_c^- \end{cases} \tag{2.22}$$

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 $\to 0$ when $T \to T_c$ (see Figure 2.15).

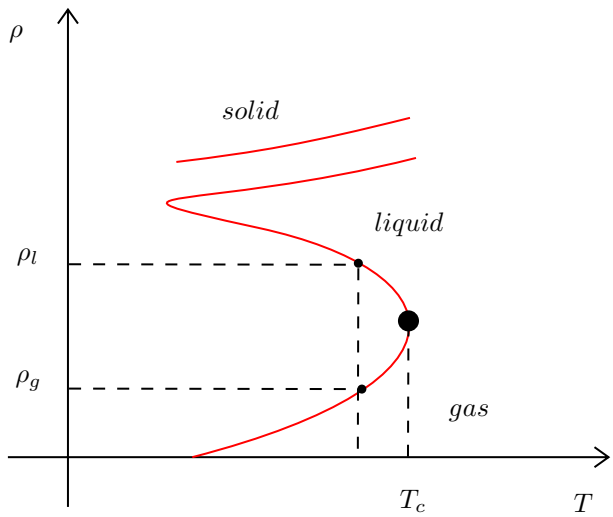


Figure 2.15: (T, ρ) projection of the (P, V, T) system, where $\rho = N/V$.

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:

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

Variable conjugate to O_P

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

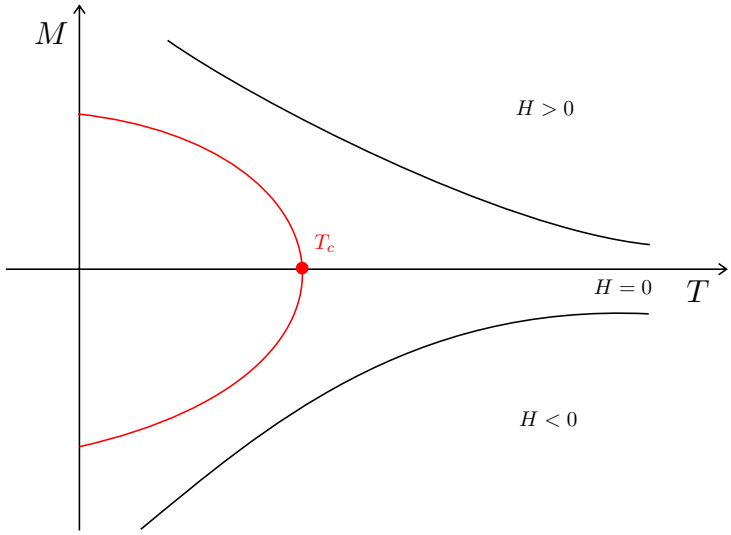


Figure 2.16: Magnetization of a ferromagnet. In red: zero-field magnetization. Below the critical temperature there is a spontaneous magnetization.

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 Eherenfest classification

The *Eherenfest 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 nth 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 ρ is discontinuous ($v = (\partial G/\partial P)_T$).

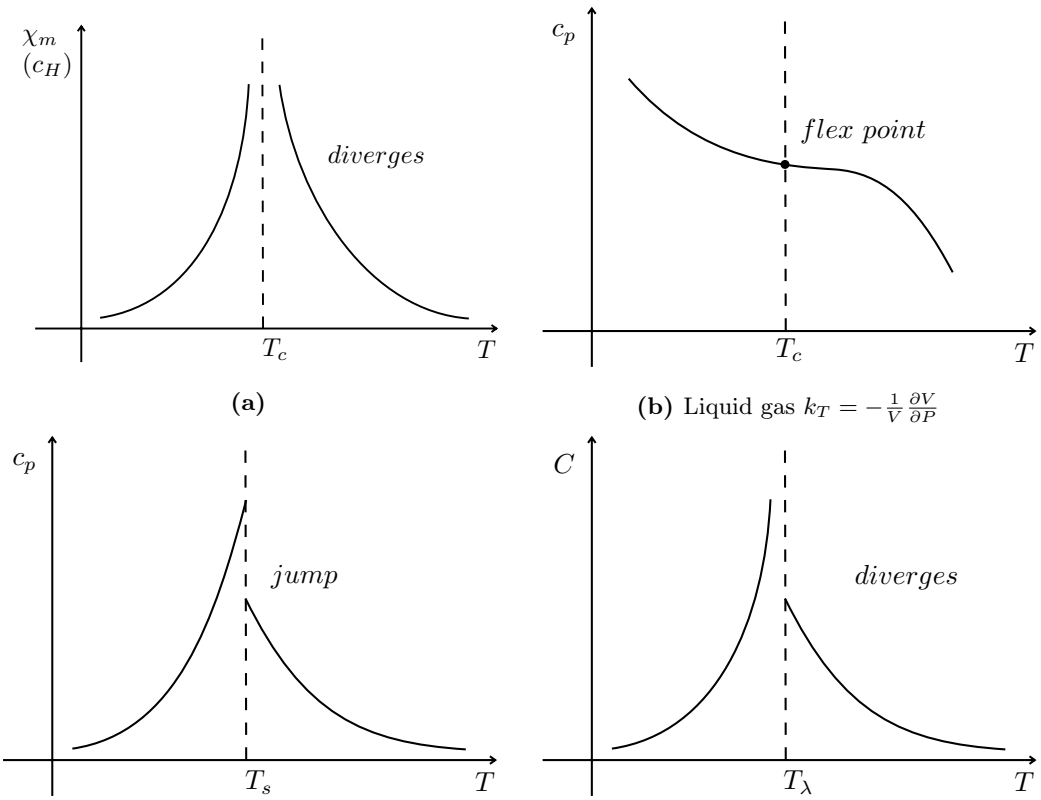
Example 12. Second order transition. The specific heat displays a finite jump, see Figure 2.17c in the conductor-superconductor transition.

Another example is a second order transition but with divergence. Consider the fluid-superfluid transition (or λ transition) of the He_4 (Figure 2.17d).

Remark. λ 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 λ.

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



(a)

(b) Liquid gas $k_T = -\frac{1}{V}\frac{\partial V}{\partial P}$

(c) 2° order phase transition

(d) Superfluid in λ-transition

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 singolarities. 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 $T \to T_c^-$, 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 $k_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \to \infty$. Similarly, in a magnetic since the curve is like Figure 2.16, we have $\chi_T = \left(\frac{\partial M}{\partial H}\right)_T \underset{T\to T_c}{\to} \infty$.

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:

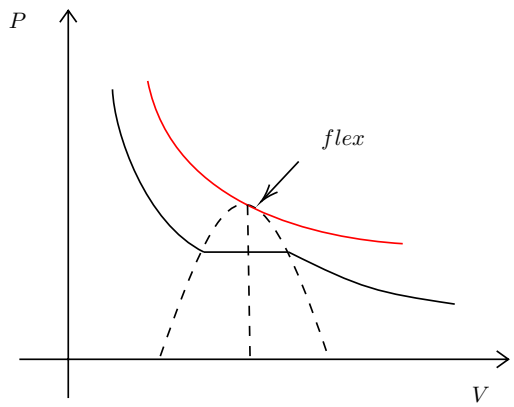


Figure 2.18: (V, T) projection.

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* λ associated to the function $F(t)$ is defined as:

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

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

$$F(t) \overset{t\to 0^\pm}{\sim} |t|^{\lambda_\pm} \tag{2.24}$$

More generally, for $t \ll 1$:

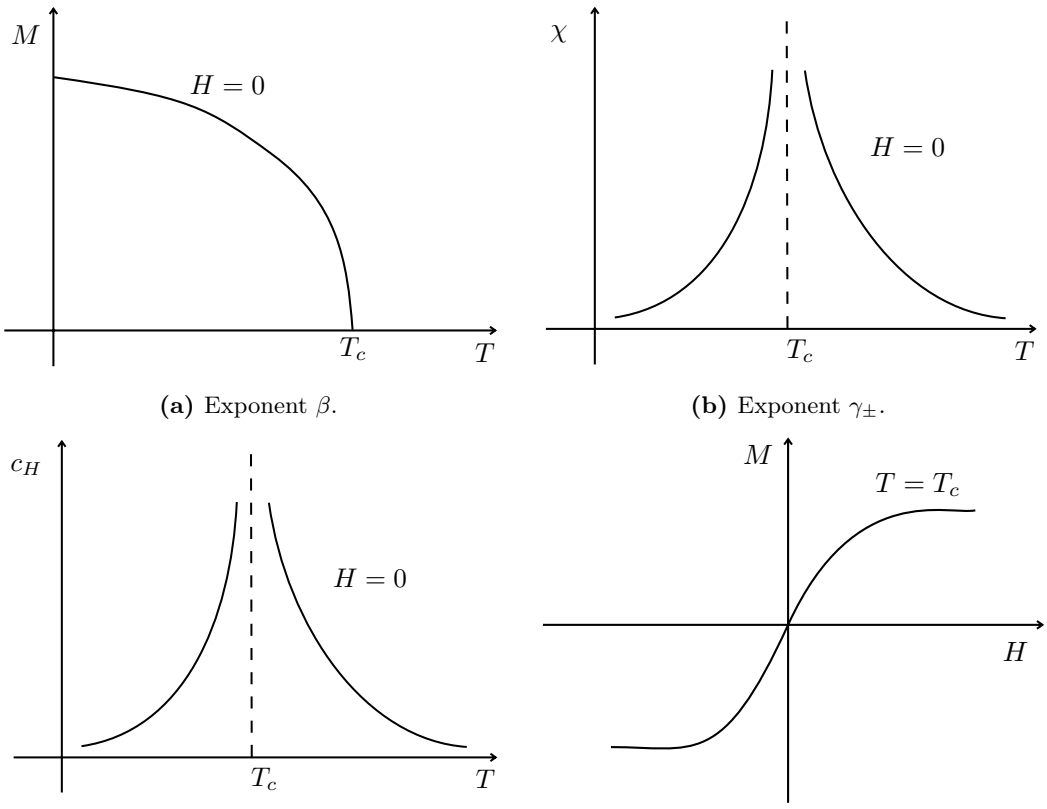
$$F(t) = A|t|^{\lambda_\pm}(1 + bt^{\lambda_1} + \dots), \quad \lambda_1 > 0 \tag{2.25}$$

where all other terms are less important.

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

- **Exponent** β: tells how the order parameter goes to zero. Consider Figure 2.19a, we have $M \overset{t\to 0^-}{\sim} (-t)^\beta$. No sense in going from above where it stays 0.
- **Exponent** $\gamma_\pm$ (suscettibility): related to the response function. Consider Figure 2.19b, we have $\chi_T \overset{t\to 0^\pm}{\sim} |t|^{-\gamma_\pm}$. In principle, the value of γ 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** δ: 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, $H \sim |M|^\delta \mathrm{sign}(M)$.

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 $T \to T_c$ from above or below.



(a) Exponent β.

(b) Exponent $\gamma_{\pm}$.

(c) Exponent $\alpha_{\pm}$.

(d) Exponent δ.

Figure 2.19

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

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

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

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 $T \to T_c$. 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.

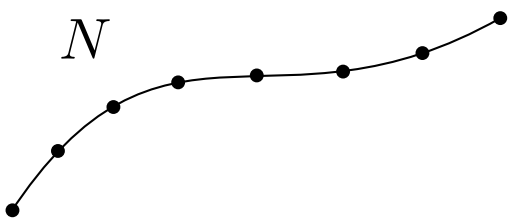
To summurize, the critical exponents are more interesting than T_c since their values do not depend on microscopic details but only on few parameters such as the space dimension d and the symmetry of the system.

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, χ_e, Ne, N_2, CO_2 and O_2. By plotting T/T_c versus ρ/ρ_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 $t \to 0$:

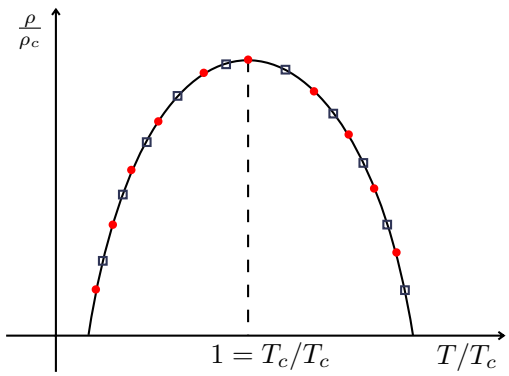
$$(\rho_l - \rho_c) \sim (-t)^\beta \tag{2.26}$$

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 β. 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.



(a) N-Polymer.



(b) Coexistence curve of different fluids plotted in reduced variables.

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.

Rushbrocke 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:

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

For magnetic systems one has

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

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

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

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

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

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

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

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

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

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

$$\alpha + 2\beta + \gamma \geq 2 \tag{2.33}$$

Griffith inequality

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

$$\Rightarrow \alpha + \beta(1 + \delta) \geq 2 \tag{2.34}$$

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 ensambles

Lecture 4.
Friday 18th October, 2019.
Compiled: Thursday 26th December, 2019.

3.1 Statistical ensambles

Statistical mechanics roughly speaking was born as a sort of theory from microscopic and try to compute the macroscopic lenght using thermodynamics. The problem is going from the countinuous 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 equilibirum 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:

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

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:

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

One can shot 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

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

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

$$\frac{\mathrm{d}\ln\Omega_1}{\mathrm{d}E_1} = \frac{\mathrm{d}\ln\Omega_2}{\mathrm{d}E_2} \Rightarrow T_1 = T_2 \tag{3.4a}$$

$$\frac{\mathrm{d}\ln\Omega_1}{\mathrm{d}V_1} = \frac{\mathrm{d}\ln\Omega_2}{\mathrm{d}V_2} \Rightarrow P_1 = P_2 \tag{3.4b}$$

$$\frac{\mathrm{d}\ln\Omega_1}{\mathrm{d}N_1} = \frac{\mathrm{d}\ln\Omega_2}{\mathrm{d}N_2} \Rightarrow \mu_1 = \mu_2 \tag{3.4c}$$

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

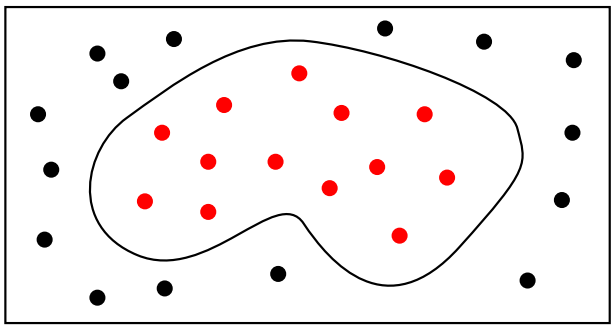


Figure 3.1: Isolated system. There are two subsystems, S constituted by red points and B constituted by the black one.

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 ensambles). Note that V_S, N_S, V_B, N_B are fixed (no exchange of volume and particles). For reasuming, other quantitites 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* (neglet interaction energy):

$$E_T = E_S + E_B = const \qquad E_B \gg E_S$$

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:

$$\Omega_T(E_T, E_S) = \Omega(E_S)\Omega_B(E_T - E_S) \tag{3.5}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Therefore:

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

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

$$S_B \to N_B s_B \qquad E_B \to N_B e_B$$

$$s_B N_B = N_B s_B(E_T) - E_{\mathcal{C}} \left(\frac{\partial s_B}{\partial e_B}\right)_{e_B=e_T} + \frac{E_{\mathcal{C}}^2}{2N_B} \left(\frac{\partial^2 s_B}{\partial e_B^2}\right) \tag{3.13}$$

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

$$\lim_{N_B\to\infty} \frac{E_T}{N_B} = \frac{E_S + N_B e_B}{N_B} \to e_B \tag{3.14a}$$

$$\lim_{N_B\to\infty} k_B \ln \Omega_B(E_T - E_{\mathcal{C}}) \to N_B s_B - E_{\mathcal{C}} \frac{\mathrm{d}s_B}{\mathrm{d}e_B} \tag{3.14b}$$

On the other hand

$$\frac{\mathrm{d}s_B}{\mathrm{d}e_B} \equiv \frac{1}{T_B} = \frac{1}{T} \tag{3.15}$$

which implies

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

Since the first therm 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

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

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:

$$\beta \equiv \frac{1}{k_B T} \tag{3.18}$$

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

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

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

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

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

$$\boxed{A(T,V,N) = -k_B T \ln Q(T,V,N)} \tag{3.21}$$

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,

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

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

$$U = \sum_{\mathcal{C}} E_{\mathcal{C}} P_{\mathcal{C}} = \frac{\sum_{\mathcal{C}} E_{\mathcal{C}} e^{-\beta E_{\mathcal{C}}}}{\sum_{\mathcal{C}} e^{-\beta E_{\mathcal{C}}}} \tag{3.22}$$

or

$$U = -\frac{\partial}{\partial \beta} \ln Q \tag{3.23}$$

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.

$$\left\langle (\delta E)^2 \right\rangle = \left\langle (E - \langle E \rangle)^2 \right\rangle = \left\langle E^2 \right\rangle - \langle E \rangle^2 \tag{3.24}$$

On the other hand,

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

$$\left\langle E^2 \right\rangle = \sum_{\mathcal{C}} P_{\mathcal{C}} E_{\mathcal{C}}^2 = \frac{\left(\frac{\partial^2 Q}{\partial \beta^2} \right)}{Q} \tag{3.26}$$

Therefore,

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

Since

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

we have

$$\left\langle (\delta E)^2 \right\rangle = k_B T^2 C_v \tag{3.29}$$

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

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

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 ensamble is *isothermal* and *isobaric*. All the assumputions done before are valid, in particular assuming as before *weak coupling* between the degrees of freedom of the bath and those of the system

$$E_T = E + E_B$$
$$V_T = V + V_B$$

we look for the partition function that describes this isothermal and isobaric ensamble. 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 $N_B \to \infty$.

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

Recalling that

$$\frac{\mathrm{d}S}{\mathrm{d}V}\bigg|_E = \frac{P}{T} \qquad \text{with} \begin{cases} P_B \to P \\ T_B \to T \end{cases} \quad (3.34)$$

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

If we normalize:

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

where

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

is called the **Gibbs partition function**.

Note that

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

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

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

Remark.

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

Classical systems (fluids)

$$\Delta(T,P,N) = \int_0^\infty \mathrm{d}V\, e^{-\beta PV} \left[\frac{1}{h^{3N} N!} \int \mathrm{d}\vec{\mathbf{p_1}} \ldots \mathrm{d}\vec{\mathbf{p_N}}\, e^{-\beta \mathcal{H}(p^N, r^N)} \right] \tag{3.41}$$

which implies

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

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

$$\beta P \equiv \frac{P}{T} = \frac{\partial S}{\partial V} \tag{3.43}$$

Magnetic system

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

$$P_{\mathcal{C}} \propto e^{S_B(E_T, M_T)/k_B - \frac{E_{\mathcal{C}}}{k_B}\frac{\mathrm{d}S_B}{\mathrm{d}E_B} - \frac{M_{\mathcal{C}}}{k_B}\frac{\mathrm{d}S_B}{\mathrm{d}M_B}} \tag{3.44}$$

Since $\frac{\mathrm{d}S_B}{\mathrm{d}M_B} = -\frac{H_B}{T_B}$ and $\frac{\mathrm{d}S_B}{\mathrm{d}E_B} = \frac{1}{T_B}$:

$$\Rightarrow P_{\mathcal{C}} \propto \exp[-\beta(E_{\mathcal{C}} - HM_{\mathcal{C}})], \qquad T_B \to T, H_B \to H \tag{3.45}$$

Normalization function:

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

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):

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

where

$$\left(\frac{\mathrm{d}S(E^*,V^*,N)}{\mathrm{d}E}\right)_{V,N} = \frac{1}{T}, \qquad \left(\frac{\mathrm{d}S(E^*,V^*,N)}{\mathrm{d}V}\right)_{E,N} = \frac{P}{T} \tag{3.48}$$

this implies

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

We define the **Gibbs free energy**:

$$G(T,P,N) = -k_B T \ln \Delta(T,P,N) \tag{3.50}$$

3.4 Gran canonical ensemble

In this case N varies instead than V.

$$\begin{aligned} P_{\mathcal{C}} &= \exp\left[\frac{S_B(E_T - E_{\mathcal{C}}, N_T - N_{\mathcal{C}})}{k_B}\right] \\ &\sim \exp\left[\frac{S_B(E_T, N_T)}{k_B} - \frac{E_{\mathcal{C}}}{k_B}\frac{\mathrm{d}S_B}{\mathrm{d}E_B} - \frac{N_{\mathcal{C}}}{k_B}\frac{\mathrm{d}S_B}{\mathrm{d}N_B} + \left(\text{terms of order} \leq \frac{1}{V_B}\right)\right] \\ &= \frac{\exp[-\beta E_{\mathcal{C}} + \beta\mu N_{\mathcal{C}}]}{\Theta(T, V, \mu)} \end{aligned} \tag{3.51}$$

where

$$\boxed{\Theta(T, V, \mu) = \sum_N \sum_{\substack{\mathcal{C} \\ V,N\text{fixed}}} e^{-\beta(E_{\mathcal{C}} - \mu N)}} \tag{3.52}$$

is the **grancanonical partition function**.

Remark. Remember that

$$\frac{\mathrm{d}S}{\mathrm{d}E} = \frac{1}{T}, \qquad \frac{\mathrm{d}S}{\mathrm{d}N} = \frac{\mu}{T} \tag{3.53}$$

The *fugacity* is defined as:

$$z \equiv e^{\beta\mu} \tag{3.54}$$

and we rewrite

$$\Theta(T, V, \mu) = \sum_{N=0}^{\infty} z^N \Big(\sum_{\substack{\mathcal{C} \\ V,N\text{fixed}}} e^{-\beta E_{\mathcal{C}}} \Big) \tag{3.55}$$

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

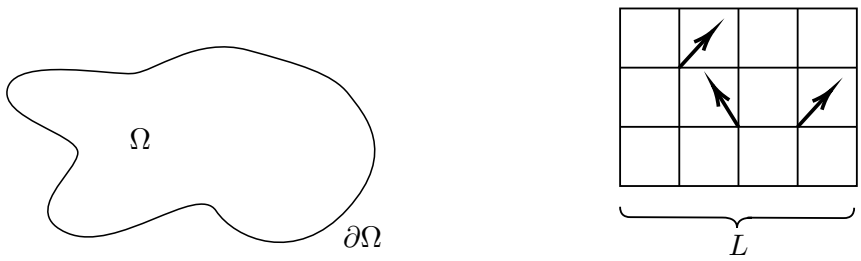
Chapter 4

Statistical mechanics and phase transitions

Lecture 5. *Wednesday 23rd October, 2019. Compiled: Thursday 26th December, 2019.*

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.



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

(b) Magnetic system with characteristic length L.

Figure 4.1

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

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

where d is the spatial dimension.

Remark. Space Ω can be either *discrete* or *continuous*.

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

$$\mathcal{H}_\Omega = -\sum_n k_n \Theta_n \tag{4.1}$$

where

- k_n: are the coupling constants. In general, but not always, they are *intensive thermodynamic variables*.
- Θ_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 $\mathcal{C} = \{\vec{\mathbf{S_1}}, \ldots, \vec{\mathbf{S_N}}\}$. We have:

$$\Theta_1 = \sum_i \vec{\mathbf{S_i}} \tag{4.2a}$$

$$\Theta_2 = \sum_{ij} \vec{\mathbf{S_i}} \cdot \vec{\mathbf{S_j}} \tag{4.2b}$$

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

$$\mathrm{Tr} \equiv \sum_{\{\mathcal{C}\}} \equiv \sum_{\vec{\mathbf{S_1}}} \sum_{\vec{\mathbf{S_2}}} \cdots \sum_{\vec{\mathbf{S_N}}} \tag{4.3}$$

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

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

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

$$\{\mathcal{C}\} = \{(\vec{\mathbf{x_i}}, \vec{\mathbf{p_i}})_{i=1,\ldots,N}\}$$

and

$$\Theta_1 = \sum_i \left[\frac{\vec{\mathbf{p_i}}^2}{2m_i} + U_1(\vec{\mathbf{x_i}})\right] \tag{4.5a}$$

$$\Theta_2 = \frac{1}{2}\sum_{i>j} U(|\vec{\mathbf{x_i}} - \vec{\mathbf{x_j}}|) \tag{4.5b}$$

The trace operation is

$$\mathrm{Tr} \equiv \sum_{\{\mathcal{C}\}} = \sum_{N=0}^{\infty} \frac{1}{N!} \int \prod_{i=1}^{N} \frac{(\mathrm{d}\vec{\mathbf{p_i}})(\mathrm{d}\vec{\mathbf{x_i}})}{h^{dN}} \tag{4.6}$$

The *gran canonical partition function* is:

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

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

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

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

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

In general, one can write

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

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

Definition 3. We define

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

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

For a system defined on a lattice we have

$$L(\Omega) \propto N(\Omega)^{1/d}, \quad V(\Omega) \propto N(\Omega)$$

$$f_b[T, \{k_n\}] = \lim_{N(\Omega)\to\infty} \frac{1}{N(\Omega)} F_N[T, \{k_n\}] \tag{4.11}$$

To get information on surface property of the system

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

4.1.3 Thermodynamic limit with additional constraints

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

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

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

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

These are analytic functions everywhere (it converges), therefore Z_Ω is analytic for Ω 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 extactly 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 fluctuactions 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

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

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

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

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

Hence,

$$\chi_T = \frac{1}{k_BT}\Big(\langle M^2\rangle - \langle M\rangle^2\Big) \tag{4.17}$$

The thermodynamic response function χ_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

$$M(\mathcal{C}) = \int \mathrm{d}^3\vec{\mathbf{r}}\, m(\vec{\mathbf{r}}) \tag{4.18}$$

Hence

$$k_BT\chi_T = \int \mathrm{d}\vec{\mathbf{r}}\,\mathrm{d}\vec{\mathbf{r}'}\left[\Big\langle m(\vec{\mathbf{r}})m(\vec{\mathbf{r}'})\Big\rangle - \langle m(\vec{\mathbf{r}})\rangle\Big\langle m(\vec{\mathbf{r}'})\Big\rangle\right] \tag{4.19}$$

Let us assume the *translational symmetry*:

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

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

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

Given the translational inviariance on can centre the system such that its centre of mass coincides with the origin

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

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

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

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

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

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

meaning that for $|\vec{\mathbf{r}}| > \xi$ the fluctuations are uncorrelated, where ξ 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 ($\xi \to \infty$). When ξ will diverge, there will not be anymore the exponential and the integral cannot be keeped finite.

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

$$k_B T \chi_T \leq V g \xi^3 \tag{4.25}$$

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

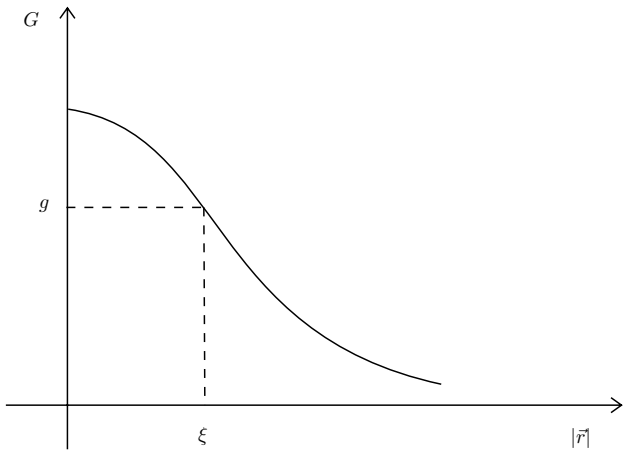


Figure 4.2: Plot of the two-point correlation function, G.

Rearranging the terms we obtain

$$\frac{k_B T \chi_T}{V} < g\xi^3 \tag{4.26}$$

Hence, if χ_T diverges at the critical point it implies $\xi \to \infty$.

In particular, one can see that for $H = 0$ and $T \to T_C^{\pm}$:

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

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. $t = 0$)

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

where μ 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 ξ 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 ρ with respect to ρ_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 ξ, where it is the typical size of the liquid droplets. Clearly $\xi = \xi[T]$ and, in proximity of the critical point $\xi \overset{t\to 0}{\sim} |t|^{-\nu}$.

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

As $T \to T_c$, where ξ 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 $L \to \infty$. 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 $L = 1\,\text{cm}$ and

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

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

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

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 ξ_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.

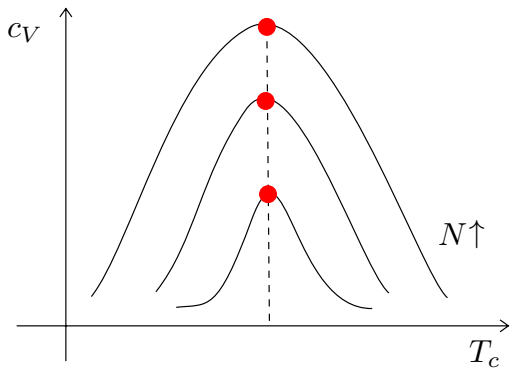


Figure 4.3: (T_c, C_V) plot at different N.

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

Lecture 6.
Friday 25th October, 2019.
Compiled: Thursday 26th December, 2019.

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.

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 anysotropic 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

$$J_{ij} = f(|\vec{\mathbf{r}}_\mathbf{i} - \vec{\mathbf{r}}_\mathbf{j}|)$$

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:

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

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

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

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 dipends on the sign of J.

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

$$\mathrm{Tr} \equiv \sum_{S_1=\pm} \sum_{S_2=\pm} \cdots \sum_{S_N=\pm} \equiv \sum_{\{S\}} \tag{5.3}$$

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

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

and the corresponding *free-energy*,

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

The *bulk limiting free energy* is:

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

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

$$\sum_{j \neq i} |J_{ij}| < \infty \tag{5.7}$$

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

$$J_{ij} = A|\vec{\mathbf{r}}_\mathbf{i} - \vec{\mathbf{r}}_\mathbf{j}|^{-\sigma} \tag{5.8}$$

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

$$\sigma > d$$

Remark. If the interaction is *dipolar* since it decades as $1/r^3$, for the case $d = 3$ the limit does not exists. 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 neirest 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:

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

Now, the model is very simple:

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

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

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

where $\langle \ldots \rangle$ 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 ffects and hence no phase transition.

Since

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

it is easy to show that:

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

where

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

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

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

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:
$$c_H = T \left(\frac{\partial S}{\partial T} \right)_H = -T \left(\frac{\partial^2 f_b}{\partial T^2} \right)_H \geq 0 \tag{5.16}$$
6. $\frac{\partial f_b}{\partial H}$ is a *monotonic non increasing* function of H. That is
$$\frac{\partial^2 f_b}{\partial H^2} \leq 0 \tag{5.17}$$
This implies that
$$\chi_T = \left(\frac{\partial M}{\partial H} \right)_T = -\left(\frac{\partial^2 f_b}{\partial H^2} \right)_T \geq 0 \tag{5.18}$$

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

$$\begin{aligned} -\frac{\partial F_\Omega}{\partial T} &= k_B \ln \left(\operatorname{Tr} e^{-\beta \mathcal{H}_\Omega} \right) + k_B T \frac{1}{k_B T^2} \frac{\operatorname{Tr}\left(\mathcal{H}_\Omega e^{-\beta \mathcal{H}_\Omega} \right)}{\operatorname{Tr}(e^{-\beta \mathcal{H}_\Omega})} \\ &= k_B \left[\ln Z + \frac{\operatorname{Tr}\left(\beta \mathcal{H}_\Omega e^{-\beta \mathcal{H}_\Omega} \right)}{Z_\Omega} \right] \underset{to\,do}{=} -k_B T \operatorname{Tr}(\rho_\Omega \ln \rho_\Omega) \end{aligned} \tag{5.19}$$

where

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

is the probability distribution.

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

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

□

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 α_1, α_2 such that $\alpha_1 + \alpha_2 = 1$, the following inequality holds

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

Now, consider the partiction function:

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

It implies that

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

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

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

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

If we now apply the Hölder inequality we get

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

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

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

It implies

$$f_b(H_1\alpha_1 + H_2\alpha_2) \geq \alpha_1 f_b(H_1) + \alpha_2 f_b(H_2) \tag{5.29}$$

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 Φ of the configuration $\{S_i\}$, the following relation holds:*

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

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,

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

This is a spontaneous broken symmetry.

Hence,

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

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

$$F_\Omega(T, J, H) = F_\Omega(T, J, -H) \tag{5.33}$$

If we take the thermodynamic limit, we have

$$\Rightarrow f_b(T, J, H) = f_b(T, J, -H) \tag{5.34}$$

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,

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

Therefore:

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

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 $a \sim$ Å).

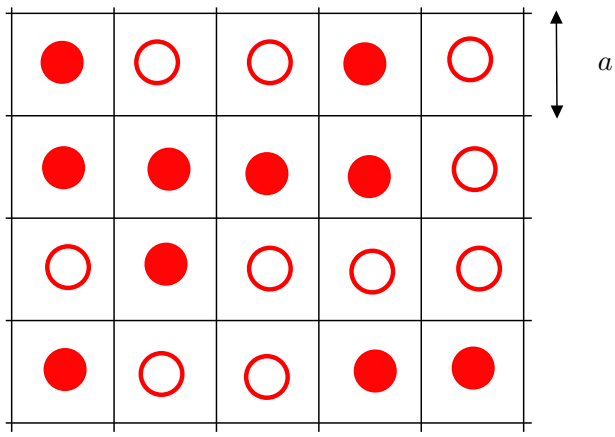


Figure 5.1: d-dimensional lattice with lattice spacing a.

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

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

We have:

$$N_\Omega = \sum_{i=1}^{N_c} n_i \tag{5.38}$$

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

The hamiltonian of the model is

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

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,

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

and we put $U_1 = 0$ for convenience.

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

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

The one body term becames:

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

while the two bodies term is equal to:

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

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

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

It implies

$$\frac{1}{2}\sum_{ij} U_2(i,j)\left[\frac{1}{4}(1 + S_i)(1 + S_j)\right] = \frac{1}{4}zU_2\sum_i^{N_c} S_i + \frac{U_2}{4}\sum_{\langle ij\rangle} S_iS_j + \frac{1}{8}U_2zN_c \tag{5.45}$$

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

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

where

$$E_0 = -\frac{1}{2}\mu N_c + \frac{z}{8}U_2N_c \tag{5.47a}$$

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

$$-J = \frac{U_2}{4} \tag{5.47c}$$

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:

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

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 Ω

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

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

In the gran-canonical ensemble we have:

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

and the gran-canonical potential is

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

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

Indeed, if U_2 is an hard-core repulsion, each particle has a finite volume and, within a finite Ω, 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

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

with the contraint

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

Remember also that

$$\mathrm{d}\omega_b(T, \mu) = -\sigma\, \mathrm{d}T - \rho\, \mathrm{d}\mu = -P \tag{5.53}$$

Now:

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

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

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

where

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

Therefore,

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

with

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

5.4.1 From the continuous to the lattice gas model

Let us divied Ω 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 $\{\vec{\mathbf{r}}_\mathbf{i}\}$.

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

$$\sum_\alpha n_\alpha = N = \int \mathrm{d}^d\vec{\mathbf{r}} \sum_{i=1}^N \delta(\vec{\mathbf{r}}_\mathbf{i} - \vec{\mathbf{r}}) = \int \mathrm{d}\vec{\mathbf{r}}\, \rho(\vec{\mathbf{r}}) \tag{5.59}$$

where

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

Moreover,

$$\sum_i U_1(\vec{\mathbf{r}}_\mathbf{i}) = \sum_i \int_\Omega \mathrm{d}^d\vec{\mathbf{r}}\, U_1(\vec{\mathbf{r}})\delta(\vec{\mathbf{r}}_\mathbf{i} - \vec{\mathbf{r}}) = \int_\Omega \mathrm{d}\vec{\mathbf{r}}\, U_1(\vec{\mathbf{r}})\rho(\vec{\mathbf{r}}) \tag{5.61}$$

We have $U(\{\vec{\mathbf{r}}_\mathbf{i}\}) \to U(\{n_\alpha\})$:

$$Q_N \propto \int \prod_{i=1}^N \mathrm{d}^d\vec{\mathbf{r}}_\mathbf{i} \to \sum_{\{n_\alpha\}} \tag{5.62}$$

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

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

Remark. The symbol $\sum^{'}$ means that the sum has the contraint that the toal number of particles is fixed to N.

Therefore,

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

and

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

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

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

$$\sum_{N=0}^\infty \sum_{\{n_\alpha\}}^{'} f(n_\alpha) = \sum_{\substack{\{n_\alpha\} \\ \sum_\alpha n_\alpha = 0}}^{'} f(n_\alpha) + \sum_{\substack{\{n_\alpha\} \\ \sum_\alpha n_\alpha = 1}}^{'} f(n_\alpha) \cdots + \sum_{\substack{\{n_\alpha\} \\ \sum_\alpha n_\alpha = \infty}}^{'} f(n_\alpha) = \sum_{\{n_\alpha\}} f(n_\alpha) \tag{5.66}$$

with no restriction.

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

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

$$\Rightarrow Z_\Omega = \mathrm{Tr}\, e^{-\beta(\mathcal{H}_\Omega - \widetilde{\mu}N)} = Z_{LG}(\widetilde{\mu}) \tag{5.68}$$

where

$$\widetilde{\mu} = \mu_{LG} = \mu_{phys} + Dk_BT\log\frac{a}{\Lambda} \tag{5.69}$$

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

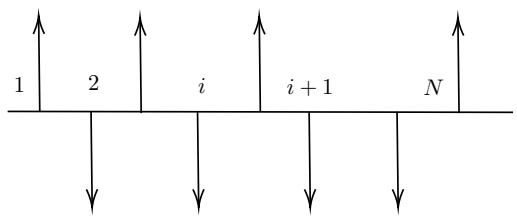


Figure 6.1: One dimensional Bravais Lattice.

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:

$$Z_N(T) = \sum_{S_1=\pm1} \sum_{S_2=\pm1} \cdots \sum_{S_N=\pm1} \exp\left[\overbrace{\beta J}^{K} \sum_{i=1}^{N-1} S_i S_{i+1}\right] \tag{6.1}$$

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

$$K \equiv \beta J, \quad h \equiv \beta H \tag{6.2}$$

Making explicit the sum in the exponential:

$$Z_N(K) = \sum_{S_1=\pm1} \sum_{S_2=\pm1} \cdots \sum_{S_N=\pm1} e^{K(S_1S_2+S_2S_3+\cdots+S_{N-1}S_N)} \tag{6.3}$$

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

$$Z_{N+1}(K) = \sum_{S_{N+1}=\pm1} \sum_{S_1=\pm1} \sum_{S_2=\pm1} \cdots \sum_{S_N=\pm1} e^{K(S_1S_2+S_2S_3+\cdots+S_{N-1}S_N)} e^{KS_NS_{N+1}} \tag{6.4}$$

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

$$\sum_{S_{N+1}=\pm1} e^{KS_NS_{N+1}} = e^{KS_N} + e^{-KS_N} = 2\cosh(KS_N) = 2\cosh(K) \tag{6.5}$$

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

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

By performing a backward iteration

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

Since $Z_1 = \sum_{S_1=\pm1} 1 = 2$, we have

$$Z_N(T) = 2(2\cos(K))^{N-1} \tag{6.8}$$

The free energy is:

$$F_N(K) = -k_BT\ln Z_N(K) = -k_BT\ln 2 - k_BT(N-1)\ln(2\cosh(K)) \tag{6.9}$$

taking the thermodynamic limit it becames

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

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

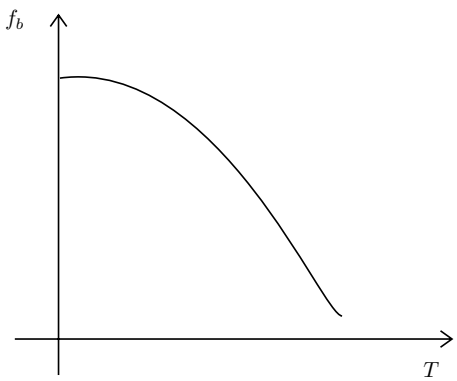


Figure 6.2: Free energy function in thermodynamic limit for the one dimension Ising model.

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:

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

It means that

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

by rearranging,

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

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 addiction 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:

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

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.

Lecture 7.
Wednesday 30th October, 2019.
Compiled: Thursday 26th December, 2019.

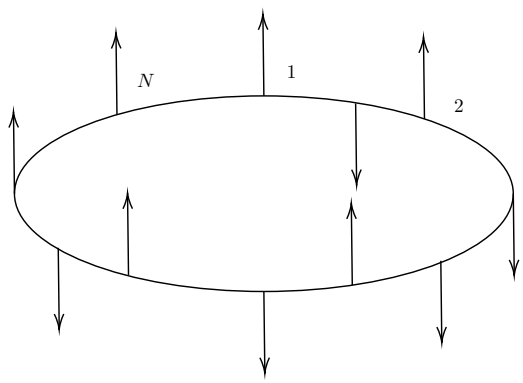


Figure 6.3: One dimensional lattice ring.

6.1.2 Transfer Matrix method

Given the Hamiltonian discussed in (6.1.1) [1] we can write the corresponding partition function in the following symmetric form:

$$Z_N(k,h) = \sum_{S_1=\pm1}\sum_{S_2=\pm1}\cdots\sum_{S_N=\pm1}\left[e^{KS_1S_2+\frac{h}{2}(S_1+S_2)}\right]\left[e^{KS_2S_3+\frac{h}{2}(S_2+S_3)}\right]\cdots\left[e^{KS_NS_1+\frac{h}{2}(S_N+S_1)}\right] \tag{6.21}$$

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

$$\begin{aligned} Z_N(h,k) &= \sum_{S_1=\pm1}\cdots\sum_{S_N=\pm1}\prod_{i=1}^{N}\exp\left[KS_iS_{i+1}+\frac{h}{2}(S_i+S_{i+1})\right] \\ &= \sum_{S_1=\pm1}\cdots\sum_{S_N=\pm1}\langle S_1|\,\mathbb{T}\,|S_2\rangle\langle S_2|\,\mathbb{T}\,|S_3\rangle\ldots\langle S_N|\,\mathbb{T}\,|S_1\rangle \end{aligned} \tag{6.22}$$

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

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

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 ($k_{LL}\sum S_iS_{i+1}S_{i+2}S_{i+3}$).

For Ising $S_i = \pm1$ and neirest neighbour interaction implies that we have two values and that $\mathbb{T}$ is a 2×2 matrix whose components are

$$\langle+1|\,\mathbb{T}\,|+1\rangle = \exp[K+h] \tag{6.24a}$$
$$\langle+1|\,\mathbb{T}\,|-1\rangle = \langle-1|\,\mathbb{T}\,|+1\rangle = \exp[K-h] \tag{6.24b}$$
$$\langle-1|\,\mathbb{T}\,|-1\rangle = \exp[-K] \tag{6.24c}$$

The explicit representation is

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

[1] The choice of boundary conditions becomes irrelevant in the thermodynamic limit, $N\to\infty$.

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

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

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

The identity relation is:

$$\sum_{S_i=\pm1} |S_i\rangle \langle S_i| = \left|S_i^{(+)}\right\rangle\left\langle S_i^{(+)}\right| + \left|S_i^{(-)}\right\rangle\left\langle S_i^{(-)}\right| = \mathbb{1} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} \tag{6.27}$$

By using the identity property we can write

$$\begin{aligned} Z_N(K,h) &= \sum_{S_1=\pm1} \cdots \sum_{S_N=\pm1} \langle S_1| \mathbb{T} |S_2\rangle \langle S_2| \mathbb{T} |S_3\rangle \ldots |S_i\rangle \langle S_i| \mathbb{T} |S_{i+1}\rangle \ldots \\ &= \sum_{S_1=\pm1} \langle S_1| \mathbb{T}^N |S_1\rangle = \mathrm{Tr}\left[\mathbb{T}^N\right] \end{aligned} \tag{6.28}$$

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

$$\mathbb{T}_D = \mathbb{P}^{-1}\mathbb{T}\mathbb{P} \tag{6.29}$$

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

$$\begin{aligned} \mathrm{Tr}\left[\mathbb{T}^N\right] &= \mathrm{Tr}\left[\underbrace{\mathbb{T}\mathbb{T}\mathbb{T}\ldots\mathbb{T}}_{N}\right] = \mathrm{Tr}\left[\mathbb{P}\mathbb{P}^{-1}\mathbb{T}\mathbb{P}\mathbb{P}^{-1}\mathbb{T}\mathbb{P}\ldots\mathbb{P}^{-1}\mathbb{T}\mathbb{P}\mathbb{P}^{-1}\right] \\ &= \mathrm{Tr}\left[\mathbb{P}\mathbb{T}_D^N\mathbb{P}^{-1}\right] \underset{\substack{\text{ciclyc property}\\ \text{of the trace}}}{=} \mathrm{Tr}\left[\mathbb{T}_D^N\mathbb{P}^{-1}\mathbb{P}\right] \\ &= \mathrm{Tr}\left[\mathbb{T}_D^N\right] \end{aligned} \tag{6.30}$$

where

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

with $\lambda_\pm$ are the eigenvalues with $\lambda_+ > \lambda_-$.

Remark. $\mathbb{P}$ is the similitude matrix whose colums are given by the eigenvectors of $\lambda_\pm$.

We finally have:

$$Z_N(K,h) = \mathrm{Tr}\left[\mathbb{T}_{\mathbb{D}}^{\mathbb{N}}\right] = \lambda_+^N + \lambda_-^N \tag{6.32}$$

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 ($S_i = \pm1$) with neirest neighbour and next neirest neighbour interactions. The hamiltonian is:

$$\mathcal{H} = k_1 \sum_i S_i S_{i+1} + k_2 \sum_i S_i S_{i+1} S_{i+2} S_{i+3} \tag{6.33}$$

Because of the second term now there are $2^4 = 16$ possible configurations that can be described by using a 4×4 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 mary 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:

$$\left\langle S_i^{(3)}\right| = (0,0,1^*,0,\dots,0) \qquad \left|S_i^{(3)}\right\rangle = \begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix}$$

these are the bra-ket at the k-esim position.

Similarly to the 2×2 Ising case, it is easy to show the identity property

$$\sum_{S_i} |S_i\rangle\langle S_i| = \mathbb{1}, \quad \mathbb{1}\in(n+2)\times(n+2) \tag{6.34}$$

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)

$$\mathbb{S}_i \equiv \sum_{S_i} |S_i\rangle\, S_i\, \langle S_i| \tag{6.35}$$

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 σ_z.

Remark. By constrution $\langle S_i|$ 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

$$\mathbb{T} = \mathbb{P}\mathbb{T}_D\mathbb{P}^{-1} = \sum_i |t_i\rangle\, \lambda_i\, \langle t_i| \tag{6.36}$$

Indeed,

$$\mathbb{T}\,|t_j\rangle = \sum_i |t_i\rangle\, \lambda_i\, \langle t_i|t_j\rangle = \sum_i |t_i\rangle\, \lambda_i\delta_{ij} = \lambda_j\,|t_j\rangle \tag{6.37}$$

Given the set of λ described above, the N particle partition function is given by

$$Z_N(\{K\},h) = \lambda_+^N + \lambda_-^N + \sum_{i=1}^{n} \lambda_i^N \tag{6.38}$$

6.2.1 The free energy

Now, we are interested in the limit of the bulk free energy

$$F_N() = -k_B T \log Z_N() \tag{6.39}$$

In general, looking at the thermodynamic limit $N \to \infty$, by factorizing λ_+

$$f_b(\{k\}, h) = \lim_{N\to\infty} \frac{1}{N} F_N = \lim_{N\to\infty} \frac{1}{N}(-k_B T) \log \left[\lambda_+^N + \lambda_-^N + \sum_{i=1}^{n} \lambda_i^N \right] \tag{6.40}$$

by rearringing,

$$f_b = \lim_{N\to\infty} \frac{-k_B T}{N} \log \left[\lambda_+^N \left(1 + \frac{\lambda_-^N}{\lambda_+^N} + \sum_{i=1}^{n} \left(\frac{\lambda_i}{\lambda_+} \right)^N \right) \right] \tag{6.41}$$

Since $\lambda_+ > \lambda_- > \lambda_1 > \dots \lambda_n$,

$$\left(\frac{\lambda_-}{\lambda_+} \right)^N \overset{N\to\infty}{\longrightarrow} 0, \quad \left(\frac{\lambda_i}{\lambda_+} \right)^N \overset{N\to\infty}{\longrightarrow} 0 \quad \forall i \tag{6.42}$$

we obtain:

$$f_b = -k_B T \log \lambda_+ \tag{6.43}$$

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 λ_+ exists.

Theorem 6.2.1 (Perron-Frobenius). *Let $\mathbb{A}$ be a $n \times n$ matrix. If $\mathbb{A}$ is finite ($n < \infty$) and $\mathbb{A}_{ij} > 0, \forall i, j$, ($\mathbb{A}_{ij} = \mathbb{A}_{ij}(\vec{\mathbf{x}})$), therefore its largest eigenvalue λ_+ 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. λ_+ *is a analytic function of the parameters of* $\mathbb{A}$.

Remark. Since in our case $\mathbb{A} \leftrightarrow \mathbb{T}$, λ_+ 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 ∞ 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:

$$\Gamma_R \equiv \langle S_1 S_R \rangle - \langle S_1 \rangle \langle S_R \rangle \tag{6.44}$$

Since

$$\Gamma_R \underset{R\to\infty}{\sim} \exp[-R/\xi] \tag{6.45}$$

we can define the correlation length ξ as

$$\xi^{-1} \equiv \lim_{R\to\infty} \left[-\frac{1}{R} \log |\langle S_1 S_R \rangle - \langle S_1 \rangle \langle S_R \rangle| \right] \tag{6.46}$$

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

$$\langle S_1 S_R \rangle_N = \frac{1}{Z_N} \sum_{\{S\}} S_1 S_R \exp[-\beta \mathcal{H}_N] \tag{6.47}$$

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

$$\langle S_1 S_R \rangle_N = \frac{1}{Z_N} \sum_{\{S\}} S_1 \langle S_1 | \mathbb{T} | S_2 \rangle \ldots \langle S_{R-1} | \mathbb{T} | S_R \rangle S_R \langle S_R | \mathbb{T} | S_{R+1} \rangle \ldots \langle S_N | \mathbb{T} | S_1 \rangle \tag{6.48}$$

Summing over the free spins,

$$\langle S_1 S_R \rangle_N = \frac{1}{Z_N} \sum_{S_1, S_R} S_1 \langle S_1 | \mathbb{T}^{R-1} | S_R \rangle S_R \langle S_R | \mathbb{T}^{N-R+1} | S_1 \rangle \tag{6.49}$$

On the other hand, since

$$\mathbb{T} = \sum_{i=1}^{n+2} |t_i\rangle \lambda_i \langle t_i| \tag{6.50}$$

we have

$$\mathbb{T}^{R-1} = \sum_{i=1}^{n+2} |t_i\rangle \lambda_i^{R-1} \langle t_i| \tag{6.51a}$$

$$\mathbb{T}^{N-R+1} = \sum_{i=1}^{n+2} |t_i\rangle \lambda_i^{N-R+1} \langle t_i| \tag{6.51b}$$

Hence,

$$\langle S_1 | \mathbb{T}^{R-1} | S_R \rangle = \sum_{i=1}^{n+2} \langle S_1 | t_i \rangle \lambda_i^{R-1} \langle t_i | S_R \rangle \tag{6.52}$$

and plugging this expression in (6.49) one gets

$$\sum_{\{S\}} S_1 S_R e^{-\beta \mathcal{H}_N} = \sum_{S_1 S_R} S_1 \sum_{i=1}^{n+2} \langle S_1 | t_i \rangle \lambda_i^{R-1} \langle t_i | S_R \rangle S_R \sum_{j=1}^{n+2} \langle S_R | t_j \rangle \lambda_j^{N-R+1} \langle t_j | S_1 \rangle \tag{6.53}$$

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

$$\mathbb{S}_1 = \sum_{S_1} |S_1\rangle S_1 \langle S_1| \tag{6.54a}$$

$$\mathbb{S}_R = \sum_{S_R} |S_R\rangle S_R \langle S_R| \tag{6.54b}$$

one gets

$$\sum_{\{S\}} S_1 S_R e^{-\beta\mathcal{H}_N} = \sum_{ij} \langle t_j | \mathbb{S}_1 | t_i \rangle \lambda_i^{R-1} \langle t_i | \mathbb{S}_R | t_j \rangle \lambda_j^{N-R+1} \tag{6.55}$$

Lecture 8. Wednesday 6th November, 2019. Compiled: Thursday 26th December, 2019.

Since $\sum_k \lambda_k^N = Z_N$ for $k = +, -, 1, \ldots, n$:

$$\langle S_1 S_R \rangle_N = \frac{\sum_{ij} \langle t_j | \mathbb{S}_1 | t_i \rangle \lambda_i^{R-1} \langle t_i | \mathbb{S}_R | t_j \rangle \lambda_j^{N-R+1}}{\sum_{k=1}^n \lambda_k^N} \tag{6.56}$$

If we now multiply and divide by λ_+^N, we get

$$\langle S_1 S_R \rangle_N = \frac{\sum_{ij} \langle t_j | \mathbb{S}_1 | t_i \rangle (\lambda_i/\lambda_+)^{R-1} \langle t_i | \mathbb{S}_R | t_j \rangle (\lambda_j/\lambda_+)^{N-R+1}}{\sum_{k=1}^n (\lambda_k/\lambda_+)^N} \tag{6.57}$$

Remark. In the thermodynamic limit $N \to \infty$, only the terms with $j = +$ and $k = +$ survive in the sum. Remind that R is fixed.

$$\langle S_1 S_R \rangle_N = \lim_{N\to\infty} \langle S_1 S_R \rangle_N = \sum_{i=\pm,1\ldots n} \left(\frac{\lambda_i}{\lambda_+}\right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_i \rangle \langle t_i | \mathbb{S}_R | t_+ \rangle \tag{6.58}$$

Rembember that $\lambda_+ > \lambda_T \geq \lambda_1 \ldots \lambda_n$:

$$\langle S_1 S_R \rangle_N = \langle t_+ | \mathbb{S}_1 | t_+ \rangle \langle t_+ | \mathbb{S}_R | t_+ \rangle + \sum_{i\neq+}^n \left(\frac{\lambda_i}{\lambda_+}\right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_i \rangle \langle t_i | \mathbb{S}_R | t_+ \rangle \tag{6.59}$$

Since one can prove, by a method entirely analogous to that followed above, that

$$\lim_{N\to\infty} \langle S_R \rangle_N = \langle t_+ | \mathbb{S}_R | t_+ \rangle \tag{6.60}$$

we obtain

$$\langle S_1 S_R \rangle = \langle S_1 \rangle \langle S_R \rangle + \sum_{i\neq+} \left(\frac{\lambda_i}{\lambda_+}\right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_i \rangle \langle t_i | \mathbb{S}_R | t_+ \rangle \tag{6.61}$$

The correlation function then follows immediately as

$$\Gamma_R = \langle S_1 S_R \rangle - \langle S_1 \rangle \langle S_R \rangle = \sum_{i\neq+}^n \left(\frac{\lambda_i}{\lambda_+}\right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_i \rangle \langle t_i | \mathbb{S}_R | t_+ \rangle \tag{6.62}$$

Remark. Γ_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 $R \to \infty$ the ratio (λ_-/λ_+) dominates the sum and hence

$$\begin{aligned}
\xi^{-1} &= \lim_{R\to\infty} \left\{ -\frac{1}{R-1} \log |\langle S_1 S_R \rangle - \langle S_1 \rangle \langle S_R \rangle| \right\} \\
&= \lim_{R\to\infty} \left\{ -\frac{1}{R-1} \log \left[\left(\frac{\lambda_-}{\lambda_+}\right)^{R-1} \langle t_+ | \mathbb{S}_1 | t_- \rangle \langle t_- | \mathbb{S}_R | t_+ \rangle \right] \right\} \\
&= -\log \left[\left(\frac{\lambda_-}{\lambda_+}\right) \right] - \lim_{R\to\infty} \frac{1}{R-1} \log \langle t_+ | \mathbb{S}_1 | t_- \rangle \langle t_+ | \mathbb{S}_R | t_+ \rangle \\
&= -\log \left(\frac{\lambda_-}{\lambda_+}\right)
\end{aligned} \tag{6.63}$$

The important result is

$$\boxed{\xi^{-1} = -\log \left(\frac{\lambda_-}{\lambda_+}\right)} \tag{6.64}$$

It means that the correlation length does depend only on the ratio between the two largest eigenvalues of the transfer matrix $\mathbb{T}$.

6.2.3 Results for the *1-dimensional* Ising model

The transfer matrix is given by

$$\mathbb{T} = \begin{pmatrix} \exp(K+h) & \exp(-K) \\ \exp(-K) & \exp(K-h) \end{pmatrix} \tag{6.65}$$

Calculate the eigenvalues:

$$|\mathbb{T} - \lambda \mathbb{1}| = (e^{K+h} - \lambda)(e^{K-h} - \lambda) - e^{-2K} = 0 \tag{6.66}$$

The solutions are

$$\lambda_{\pm} = e^K \cosh(h) \pm \sqrt{e^{2K} \sinh^2(h) + e^{-2K}} \tag{6.67}$$

The free energy

The free energy is

$$\begin{aligned} f_b &\equiv \lim_{N\to\infty} \frac{-k_B T}{N} \log Z_N(K,h) \\ &= -k_B T \lim_{N\to\infty} \frac{1}{N} \log \left[\lambda_+^N \left(1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right) \right] \\ &= -k_B T \log \lambda_+ \end{aligned} \tag{6.68}$$

and inserting the explicit expression of λ_+ for the Ising model, we get

$$\begin{aligned} f_b &= -k_B T \log \left(e^K \cosh h + \sqrt{e^{2K} \sinh^2(h) + e^{-2K}} \right) \\ &= -K k_B T - k_B T \log \left(\cosh(h) + \sqrt{\sinh^2(h) + e^{-4K}} \right) \end{aligned} \tag{6.69}$$

Remark. Rembember 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 $T \to 0$ and $T \to \infty$ by keeping H fixed and J fixed.

- Case: $T \to 0 \Rightarrow K \to \infty, h \to \infty$.

$$e^{-4K} \overset{K\to\infty}{\longrightarrow} 0 \tag{6.70a}$$

$$\sqrt{\sinh^2 h} \overset{h\to\infty}{\sim} \sinh(h) \tag{6.70b}$$

This implies that

$$\cosh(h) + \sinh h \sim \frac{2e^h}{2} \simeq e^h \tag{6.71}$$

and

$$f \overset{\substack{h\to\infty \\ K\to\infty}}{\sim} -K k_B T - k_B T \log e^h \sim -J - H \quad const \tag{6.72}$$

Therefore, as $T \to 0^+$, f goes to a constant that depends on J and H.

- Case: $T \to \infty \Rightarrow K \to 0, h \to 0$. In this case we suppose also that H and J that are fixed, are also finite.

$$e^{-4K} \simeq 1 \tag{6.73a}$$

$$\sqrt{\sinh^2 h + e^{-4K}} \sim \sqrt{1} \tag{6.73b}$$

Since $\cosh h \overset{h\to 0}{\sim} 1$:

$$f_B \sim -K k_B T - k_B T \log(1+1) \sim -J - k_B T \ln 2 \tag{6.74}$$

Therefore, as $T \to \infty$, the free energy goes linearly to zero, as in figure 6.4.

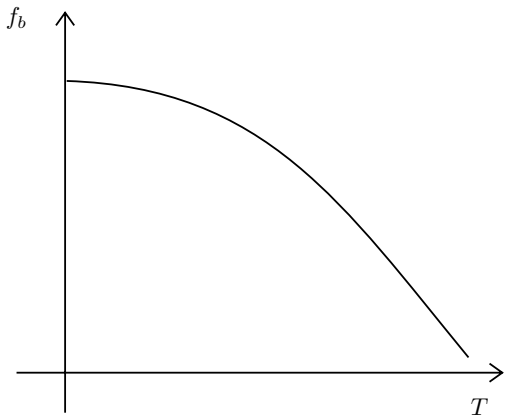


Figure 6.4: Plot of the free energy f_b in function of the temperature T. For $T \to 0$, the free energy becomes constant, while for $T \to \infty$ it goes linearly to zero.

The magnetization

This can be obtained by differentiating the negative of the free energy with respect to the magnetic field H:

$$m = -\frac{\partial f_b}{\partial H} = -\frac{1}{k_B T}\frac{\partial f_b}{\partial h} = \frac{\partial}{\partial h}\left[\log\left(\cosh(h) + \sqrt{\sinh^2(h) + e^{-4K}}\right)\right] \tag{6.75}$$

The result is

$$m = \frac{\sinh h + \frac{\sinh h \cosh h}{\sqrt{\sinh^2 h + e^{-4K}}}}{\cosh h + \sqrt{\sinh^2 h + e^{-4K}}} \tag{6.76}$$

- Case: $T > 0$ fixed, $H \to 0 \leftrightarrow h \to 0$.

$$\sinh h \sim h \sim 0, \tag{6.77a}$$

$$\cosh h \sim 1 \tag{6.77b}$$

  In zero field $h \to 0$, we have $m \to 0$ for all $T > 0$. It means that there is no spontaneous magnetization!

The magnetic susceptibility

$$\chi_T \equiv \frac{\partial m}{\partial H} = \frac{1}{k_B T}\frac{\partial m}{\partial h} \tag{6.78}$$

If we consider the case $h \ll 1$, it is convenient first expand the (6.76) for $h \to 0$ and take the derivative to get χ_T.

Since $\sinh(h) \sim h + h^3$ and $\cosh(h) \sim 1 + h^2$, we have

$$m \overset{h \ll 1}{\sim} \frac{h(1 + e^{2K})}{1 + e^{-2K}} \tag{6.79}$$

If we now derive with respect to h

$$\chi_T = \frac{1}{k_B T}\frac{\partial m}{\partial h} \overset{h \ll 1}{\approx} \frac{1}{k_B T}\frac{(1 + e^{2K})}{(1 + e^{-2K})} \tag{6.80}$$

- Case: $T \to \infty \Leftrightarrow K \to 0$.

$$e^{2K} \simeq e^{-2K} \simeq 1 \tag{6.81}$$

  The *Curie's Law* for paramagnetic systems is:

$$\boxed{\chi_T \sim \frac{1}{k_B T}} \tag{6.82}$$

- Case: $T \to 0 \Leftrightarrow K \to \infty$.

$$e^{-2K} \simeq 0 \tag{6.83}$$

The *Curie's Law* for paramagnetic systems is:

$$\chi_T \sim \frac{1}{k_B T} e^{2K} \sim \frac{1}{k_B T} e^{2J/k_B T} \tag{6.84}$$

The correlation length

$$\xi^{-1} = -\log\left(\frac{\lambda_-}{\lambda_+}\right) = -\log\left[\frac{\cosh h - \sqrt{\sinh^2 h + e^{-4K}}}{\cosh h + \sqrt{\sinh^2 h + e^{-4K}}}\right] \tag{6.85}$$

For for $h = 0$, we have $\cosh h \to 1, \sinh h \to 0$:

$$\xi^{-1} = -\log\left[\frac{1 - e^{-2K}}{1 + e^{-2K}}\right] = -\log\left[\frac{1}{\coth K}\right] \tag{6.86}$$

Therefore:

$$\xi = \frac{1}{\log(\coth K)} \tag{6.87}$$

- Case: $T \to 0 \Leftrightarrow K \to \infty$.

$$\coth K = \frac{e^K + e^{-K}}{e^K - e^{-K}} \overset{K\to\infty}{\simeq} 1 + 2e^{-2K} + \dots \overset{K\to\infty}{\longrightarrow} 1 \tag{6.88}$$

It implies

$$\xi \overset{K\gg 1}{\sim} \frac{1}{\ln(1 + 2e^{-2K})} \sim \frac{e^{2K}}{2} \tag{6.89}$$

Hence

$$\xi \overset{T\to 0}{\sim} \frac{1}{2} e^{J/k_B T} \tag{6.90}$$

It diverges exponentially $\xi \to \infty$, as $T \to 0$.

- Case: $T \to \infty \Leftrightarrow K \to 0$.

$$\coth K = \frac{e^K + e^{-K}}{e^K - e^{-K}} \overset{K\to 0}{\simeq} \frac{1 + K + \frac{K^2}{2} + 1 - K + \frac{K^2}{2}}{1 + K + \frac{K^2}{2} - 1 + K - \frac{K^2}{2}} \sim \frac{2 + 2\frac{K^2}{2}}{2K} \sim \frac{1 + K^2}{K} \tag{6.91}$$

$$\xi^{-1} = \log(\coth K) \overset{K\to 0}{\sim} \ln\frac{1}{K} + \ln(1 + K^2) \sim +\infty \tag{6.92}$$

Therefore:

$$\xi \overset{K\to 0}{\longrightarrow} 0 \tag{6.93}$$

More precisely,

$$\xi \overset{K\to 0}{\sim} \frac{1}{\ln(1/K) + \ln(1 + K^2)} \overset{K\to 0}{\sim} -\frac{1}{\ln K} \tag{6.94}$$

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. $\vec{\mathbf{S}}_i \in \mathbb{R}^3$, $\left|\vec{\mathbf{S}}_i\right|^2 = 1$ (continuous values on the unit sphere):

$$\vec{\mathbf{S}}_i = (S_i^x, S_i^y, S_i^z) \tag{6.95}$$

with periodic boundary condition: $\vec{\mathbf{S}}_{N+1} = \vec{\mathbf{S}}_1$.

Assuming $H = 0$, the model is defined through the following Hamiltonian::

$$-\beta\mathcal{H}(\{\vec{\mathbf{S}}\}) = K\sum_{i=1}^{N} \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_{i+1} \quad (\longrightarrow \sum_i \vec{\mathbf{h}} \cdot \vec{\mathbf{S}}_i) \tag{6.96}$$

This model satisfies $O(3)$ symmetry. In the transfer matrix formalism:

$$Z_N(K) = \sum_{\{\vec{\mathbf{S}}\}} e^{-\beta\mathcal{H}} = \sum_{\{\vec{\mathbf{S}}\}} e^{K\sum_{i=1}^{N} \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_{i+1}} = \mathrm{Tr}\left(\mathbb{T}^N\right) \tag{6.97}$$

where $\left\langle \vec{\mathbf{S}}_i \middle| \mathbb{T} \middle| \vec{\mathbf{S}}_{i+1} \right\rangle = e^{K\vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_{i+1}}$.

Similarly to the Ising case:

$$\mathbb{T} = \sum_i |t_i\rangle \lambda_i \langle t_i| \tag{6.98}$$

and

$$\mathbb{T}_D = \mathbb{P}^{-1}\mathbb{T}\mathbb{P} \tag{6.99}$$

The problem is computing the eigenvalues λ_i of $\mathbb{T}$. Formally, we should find

$$\exp\left[K\vec{\mathbf{S}}_1 \cdot \vec{\mathbf{S}}_2\right] = \left\langle \vec{\mathbf{S}}_1 \middle| \mathbb{T} \middle| \vec{\mathbf{S}}_2 \right\rangle = \sum_{i \in \text{eigenvalues}} \lambda_i \left\langle \vec{\mathbf{S}}_1 \middle| t_i \right\rangle \left\langle t_i \middle| \vec{\mathbf{S}}_2 \right\rangle = \sum_i \lambda_i f_i(\vec{\mathbf{S}}_1) f^*(\vec{\mathbf{S}}_2) \tag{6.100}$$

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

$$e^{i\vec{\mathbf{q}}\cdot\vec{\mathbf{r}}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (i)^l j_l(qr) Y_{lm}^*(\hat{\mathbf{q}}) Y_{lm}(\hat{\mathbf{r}}) \tag{6.101}$$

where

$$j_l(qr) = -\frac{(i)^l}{2} \int_0^{\pi} \sin(\theta) e^{iqr\cos(\theta)} P_l(\cos(\theta))\, \mathrm{d}\theta \tag{6.102}$$

are the *spherical Bessel functions*, while the $P_l(\cos(\theta))$ are the *Legendre polynomial* of order l.

From a formal comparison we have

$$\vec{\mathbf{S}}_1 \leftrightarrow \hat{\mathbf{S}}_1, \qquad \begin{cases} i\vec{\mathbf{q}} \cdot \vec{\mathbf{r}} = iqr \\ K\vec{\mathbf{S}}_1 \cdot \vec{\mathbf{S}}_2 = K\left|\vec{\mathbf{S}}_1\right|\left|\vec{\mathbf{S}}_2\right| = K \end{cases} \tag{6.103}$$

multiplying by $(-i)$ we can write

$$qr = -iK\left|\vec{\mathbf{S}}_1\right|\left|\vec{\mathbf{S}}_2\right| = -iK \tag{6.104}$$

In our case we have $\hat{\mathbf{q}} = \vec{\mathbf{S}}_1, \hat{\mathbf{r}} = \vec{\mathbf{S}}_2$. Hence,

$$e^{K\vec{\mathbf{S}}_1\cdot\vec{\mathbf{S}}_2} = 4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{l}(i)^l j_l(-iK)Y^*_{lm}(\vec{\mathbf{S}}_1)Y_{lm}(\vec{\mathbf{S}}_2) = \sum_i \lambda_i f_i(\vec{\mathbf{S}}_1)f^*(\vec{\mathbf{S}}_2) \tag{6.105}$$

where

$$\lambda_i = \lambda_{lm}(K) = 4\pi(i)^l j_l(-iK) \tag{6.106}$$

Remark. Note that λ_i does not depend on m!

If $l = 0$, the largest eigenvalue is:

$$\lambda_+ = \lambda_0(K) = 4\pi j_0(-iK) = 4\pi\frac{\sin K}{K} \tag{6.107}$$

and

$$\lambda_- = \lambda_1(K) = 4\pi i j_1(-iK) = 4\pi\left[\frac{\cosh K}{K} - \frac{\sinh K}{K^2}\right] \tag{6.108}$$

Exercise 3. Given the largest eigenvalue λ_+:

$$\lambda_+ = 4\pi\frac{\sin(K)}{K} \tag{6.109}$$

find the bulk free energy density of the model and discuss its behaviour in the limits of low $(T \to 0)$ and high $(T \to \infty)$ 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 8th November, 2019. Compiled: Thursday 26th 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*.

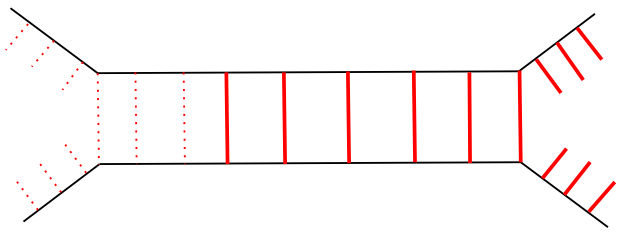


Figure 6.5: Sequential unzipping from the ends.

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.

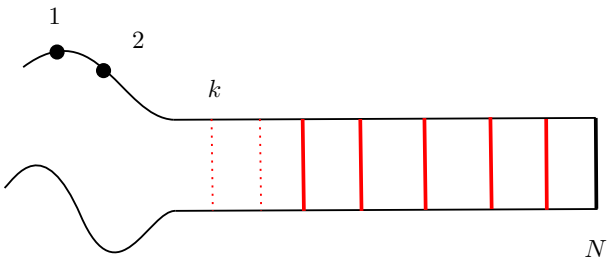


Figure 6.6

If the first k bonds are open (unbounded pairs) the energy to open the $k+1$ is ε_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! One a bond is open it can orient itself in G different ways. In other words there is an entropy

$$S_0 = k_B \log G \tag{6.110}$$

associated to each open band. If k bands ($i = 1, \ldots, k$) are open the associated weight is given by

$$G^k e^{-k\varepsilon_0/k_B T} \tag{6.111}$$

By summing over the possible values of k, the partition function is

$$Z_N(T, G, \varepsilon_0) = \sum_{k=0}^{N-1} G^k e^{-k\varepsilon_0/k_B T} = \sum_{k=0}^{N-1} e^{k(S_0 T - \varepsilon_0)/k_B T} \tag{6.112}$$

Let us call

$$\chi \equiv G e^{-\varepsilon_0/k_B T} \tag{6.113}$$

and simplify the previous expression

$$Z_N = \sum_{k=0}^{N-1} \chi^k = \frac{1-\chi^N}{1-\chi} \tag{6.114}$$

We see immidiately there is a single pole singularity. The free energy is

$$F_N = -k_B T \ln Z_N = -k_B T \ln \left[\frac{1-\chi^N}{1-\chi}\right] \tag{6.115}$$

We can now compute some observables of interest.

Calculate average number of open links

The equal of our magnetization in this case is

$$\langle k \rangle_N \equiv \frac{\sum_{k=0}^{N-1} k\chi^k}{\sum_{k=0}^{N-1} \chi^k} = \chi \frac{\mathrm{d}}{\mathrm{d}\chi} \ln Z_N = \frac{N\chi^N}{\chi^N - 1} - \frac{\chi}{\chi - 1} \tag{6.116}$$

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 $\chi \equiv 1 + \varepsilon$:

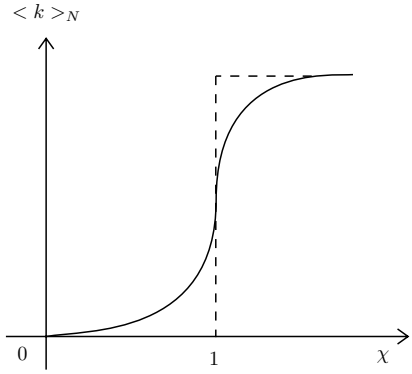


Figure 6.7: Description.

$$
\begin{aligned}
\log Z_N(\chi) &= \log\left[\frac{1-(1+\varepsilon)^N}{1-(1+\varepsilon)}\right] \\
&= \log\left[\frac{1-(1+\varepsilon N+\frac{N(N-1)}{2!}\varepsilon^2+\frac{N(N-1)(N-2)}{3!}\varepsilon^3+O(\varepsilon^4))}{\varepsilon}\right] \\
&= \log\left[N+\frac{N(N-1)}{2}\varepsilon+\frac{N(N-1)(N-2)}{6}\varepsilon^2+\dots\right] \\
&= \log N+\log\left[1+\frac{N-1}{2}\varepsilon+\frac{N(N-1)}{6}\varepsilon^2\right] \\
&= \log N+\log\left[1+\frac{N\varepsilon}{2}+\frac{N^2\varepsilon^2}{6}+\dots\right] \\
&= \log N+\left(\frac{N\varepsilon}{2}+\frac{N^2\varepsilon^2}{6}+\dots\right)+\frac{1}{2}\left(\frac{N\varepsilon}{2}+\frac{N^2\varepsilon^2}{6}+\dots\right)^2+\dots \\
&= \log N+\frac{N\varepsilon}{2}+\frac{N^2\varepsilon^2}{24}+\dots
\end{aligned}
\tag{6.117}
$$

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 $N \gg 1, \varepsilon \ll 1$. A 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):

$$
\frac{1}{N}\frac{\mathrm{d}\langle k\rangle}{\mathrm{d}\varepsilon} \simeq \frac{N}{12}-\frac{N^3\varepsilon^3}{240}+\dots \tag{6.120}
$$

It is max at $\varepsilon = 0$ and diverges as $N \to \infty$ (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

$$
Ge^{-\varepsilon_0/k_BT_C} = 1 \tag{6.121}
$$

Hence

$$T_C = \frac{\varepsilon_0}{k_B \log G} \tag{6.122}$$

Note that as $G \to 1$, $T_c \to 0$. 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

$$E(S_i, S_{i-1}) = (\varepsilon_0 + V_0 \delta_{S_{i-1},0})(1 - \delta_{S_i,0}) \tag{6.123}$$

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):

$$\mathcal{H}_N = \varepsilon_0(1 - \delta_{S_1,0}) + \sum_{i=2}^{N-1} (\varepsilon_0 + V_0 \delta_{S_{i-1},0})(1 - \delta_{S_i,0}) \tag{6.124}$$

THe Kittel's version is obtained by assuming $V_0 = \infty$.

The partition function is

$$Z_N = \sum_{\{S\}} \exp(-\beta \mathcal{H}_N) \tag{6.125}$$

In order to implement the transfer matrix formalism we rewrite Z_N as follows (DUBBIIOOOOOO per la prossima equation)

$$Z_N = \sum_{\{S\}} e^{-\beta\varepsilon_0(1-\delta_{S_1,0})} \prod_{i=1}^{N-2} e^{-\beta\varepsilon_0(1-\delta_{S_{i+1},0})(\dots)\left(1+(e^{-\beta V_0}-1)\delta_{S_i,0}(1-\delta_{S_{i+1},0})\right)} \tag{6.126}$$

Consider the Kittel model, $V_0 = \infty$ which implies $\exp(-\beta V_0) = 0$. We can define the transfer matrix as

$$\mathbb{T} = \{\langle S | \mathbb{T} | S' \rangle \equiv t_{S,S'}\} \tag{6.127}$$

where

$$t_{S,S'} = e^{-\beta\varepsilon_0(1-\delta_{S',0})}[1 - \delta_{S,0}(1 - \delta_{S',0})] \tag{6.128}$$

or in matrix form

$$\mathbb{T} = \begin{bmatrix} 1 & 0 & \ldots\ldots 0 \\ 1 & a & \ldots\ldots a \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & a & \ldots\ldots a \end{bmatrix} \tag{6.129}$$

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

$$\lambda_1 = Ga, \quad \lambda_2 = 1, \quad \lambda_1 = 0 \tag{6.130}$$

The partition function can be written as

$$Z_N = (1, a, \ldots, a)\mathbb{T}^{N-2} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{6.131}$$

$$\lambda_1 \to \vec{\mathbf{v}}_1 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \qquad \lambda_2 \to \vec{\mathbf{v}}_2 = \begin{pmatrix} 1 - Ga \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{6.132}$$

We can then write

$$\begin{pmatrix} 1 \\ a \\ \vdots \\ a \end{pmatrix} = \frac{a(1 - Ga) - 1}{1 - Ga}\vec{\mathbf{v}}_1 + \frac{1}{1 - Ga}\vec{\mathbf{v}}_2 \tag{6.133a}$$

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \frac{-Ga}{1 - Ga}\vec{\mathbf{v}}_1 + \frac{1}{1 - Ga}\vec{\mathbf{v}}_2 \tag{6.133b}$$

Therefore

$$Z_N = \frac{1 - (Ga)^N}{1 - Ga} = \frac{1 - (Ge^{-\beta\varepsilon_0})^N}{1 - Ge^{-\beta\varepsilon_0}} \tag{6.134}$$

or

$$Z_N = \frac{1}{1 - Ge^{-\beta\varepsilon_0}}(-\lambda_1^N + \lambda_2^N) \tag{6.135}$$

Since in the thermodynamic limit only the contribution of the largest eigenvalue matters for f_b we have

$$f_b = -k_B T \ln \max(\lambda_1, \lambda_2) \tag{6.136}$$

Remark. Given that the λ_1 and λ_2 are positive, analytic function of T ($\lambda_1 = Ga, \lambda_2 = 1$). In order to have a phase transition (i.e. non analiticity of f_b) the two eigenvalues must cross for a given value of T. It is true if and only if:

$$Ga_c = 1 \Leftrightarrow Ge^{-\beta_c\varepsilon_0} = 1 \Leftrightarrow T_c = \frac{\varepsilon_0}{k_B \ln G} \tag{6.137}$$

It is agree with previous calculation.

6.5 Transfer matrix for $2D$ Ising

Dimensional reducation from d to *d-1*. You can to the same for the 3D Ising model. In order to go smoothly to the one dimensional to the two dimensional the idea is to solve the Ising model to surfaces.

Consider a square lattice with N rows and M colums, as in Figure 6.8, with periodic boundary conditions (wrapped around a torus). The spin in a site is identifed by $S_{\text{site}} = S_{m,n}$. The Hamiltonian is (if $H \neq 0$ we have also the term with the h):

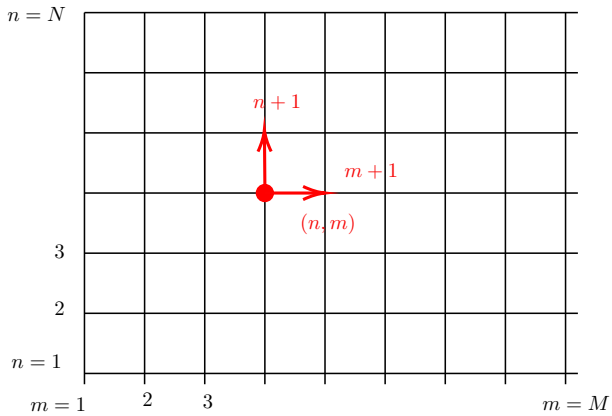


Figure 6.8: Description.

$$\begin{aligned} -\beta\mathcal{H}_\Omega(\{S\}) &= k\sum_{\langle ij\rangle} S_i S_j + h\sum_i S_i \\ &= k\sum_{n=1}^{N}\sum_{m=1}^{M}(S_{m,n}S_{m+1,n} + S_{m,n}S_{m,n+1}) + h\sum_{n=1}^{N}\sum_{m=1}^{M} S_{m,n} \end{aligned} \tag{6.138}$$

in fact

$$\sum_{\langle i,j\rangle} \to \sum_{i,j\in nn(i)} \quad \Rightarrow S_{m,n} \to S_{m+1,n} \tag{6.139}$$

The Hamiltonian can be rewritten as follows:

$$-\beta\mathcal{H}_\Omega(\{S\}) = \sum_{m=1}^{M}\left[E[\mu_m,\mu_{m+1}] + E[\mu_m]\right] \tag{6.140}$$

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 μ is a m dimensional vector, in particular each μ_m represents the set of N spins along column m:

$$\mu_m = \{S_{m,1}, S_{m,2}, \ldots, S_{m,N}\} \tag{6.141}$$

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 neirest 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 suppoe $h = 0$ (so the energy does not depend on h):

$$\langle \mu_m | \, \mathbb{T} \, | \mu_{m+1} \rangle = \exp[k(E[\mu_m, \mu_{m+1}] + E[\mu_m])] \tag{6.143}$$

Now we have to diagonalize. In the 2x2 tranfer 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

$$Z_N(k, h) = \mathrm{Tr}\big(\mathbb{T}^N\big) \tag{6.144}$$

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 diffucult to be diagonalized. This was first achived by Onsanger in 1944 for the case $H = 0$ and in the $N \to \infty$ limit. The result is given by

$$f_b(T) = -k_B T \log\left(2\cosh(2\beta J)\right) - \frac{k_B T}{2\pi} \int_0^{2\pi} \log\left[\frac{1}{2}\left(1 + \sqrt{1 - g^2 \sin^2(\Phi)}\right)\right] \mathrm{d}\Phi \tag{6.145}$$

that is the density bulk free energy, where

$$g = \frac{2}{\cosh(2\beta J)\coth(2\beta J)} \tag{6.146}$$

The equation for the critical T is

$$2\tanh^2\left(\frac{2J}{k_B T_c}\right) = 1 \quad \Rightarrow T_c \simeq 2,264 J/k_B \neq 0! \tag{6.147}$$

One can also show that

$$c \propto A\left[-\ln\left(1 - \frac{T}{T_c}\right) + B\right] \tag{6.148}$$

The specifi 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

Lecture 10. *Wednesday 13th November, 2019. Compiled: Thursday 26th December, 2019.*

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 Onsanger 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.

7.1 Energy-entropy argument

$$\mathrm{d}F = \underbrace{\mathrm{d}U}_{\text{energy}} - T \underbrace{\mathrm{d}S}_{\text{entropy}} \tag{7.1}$$

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

$$E_G = -JN \tag{7.2}$$

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 dimentional, but is valid also in many dimensional.

Which is the variation in energy ΔE respect the ground state? For each excitation there is an energy penalty $\Delta E = 2J$, indeed

$$E_G = -JN, \qquad E^* = -J(N-1) + J \qquad \Rightarrow \Delta E = 2J \tag{7.3}$$

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:

$$\Delta E_M = 2MJ \tag{7.4}$$

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

$$\# = \binom{N}{M} = \binom{N}{xN} \tag{7.5}$$

We have:

$$S_M = k_B \log \binom{N}{M} \tag{7.6}$$

the difference is

$$\Delta S = S_M - S_G = S_M = -k_B \ln \binom{N}{xN} \tag{7.7}$$

Consider:

$$\begin{aligned} \Delta F = F_M - F_G &= \Delta E - T\Delta S \\ &= 2MJ - k_B T \ln \binom{N}{M} \\ &= 2xNJ - k_B T \ln \binom{N}{xN} \\ &= N\{2xJ + k_B T[x \ln x + (1-x)\ln(1-x)]\} \end{aligned} \tag{7.8}$$

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:

$$\Delta f_{b_N} = \frac{\Delta F_N}{N}, \qquad \frac{\partial \Delta f_b}{\partial x} = 0 \tag{7.9}$$

this gives

$$\begin{aligned} \frac{\partial}{\partial x}\{2xJ + k_B T[x \ln x + (1-x)\ln(1-x)]\} &= 2J + k_B T[\ln x + 1 - \ln(1-x) - 1] \\ &= 2J + k_B T[\ln x - \ln(1-x)] = 0 \end{aligned} \tag{7.10}$$

therefore

$$\ln \frac{x}{1-x} = -\frac{2J}{k_B T} \quad \Rightarrow \frac{x}{1-x} = e^{-2J/k_B T} \tag{7.11}$$

and finally the results is

$$x = \frac{1}{1 + e^{2J/k_B T}} \tag{7.12}$$

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}

$$\Delta E \propto 2JL^{D-1} \tag{7.13}$$

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

$$S^* \geq k_B \ln L \tag{7.14}$$

the number of ways to place a straight wall within a system of linear size L. The ΔS is just S because the entropy of the ground state is again zero.

Remark. Understimate of S

$$\Delta F = 2JL^{D-1} - k_B T \ln L \tag{7.15}$$

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 perturbe the system using an external magnetic field as perturbation (it is very small h). We are breaking explicity the symmetry, but then you take the limit $h \to 0$ and swich 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

$$\langle m \rangle_N = 0 \tag{7.16}$$

If, as $N \to \infty$ 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:

$$\langle m \rangle_\infty = \lim_{h \to 0^+} \lim_{N \to \infty} \langle m \rangle_N^{(h)} \tag{7.17}$$

Remark. It is crucial to take first the thermodynamic limit and then the limit $h \to 0^+$!

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.

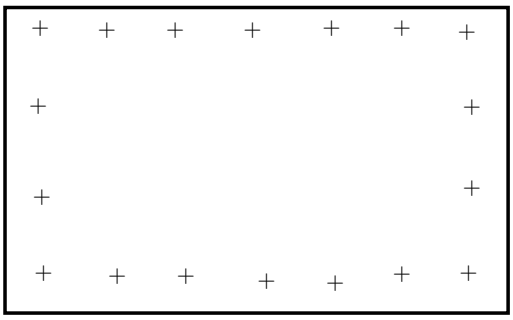


Figure 7.1

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 perturbe 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

$$\langle m \rangle_N = \frac{\langle N_+ \rangle - \langle N_- \rangle}{N} = 1 - 2\frac{\langle N_- \rangle}{N} \tag{7.18}$$

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 $\forall N$

$$\frac{\langle N_- \rangle}{N} < \frac{1}{2} - \varepsilon \tag{7.19}$$

where $\varepsilon > 0$ and indipendent on N. Indeed, if (7.19) holds

$$\langle m \rangle_N \geq 2\varepsilon \quad \forall N \tag{7.20}$$

Pierles was indeed able to show that

$$\frac{\langle N_- \rangle}{N} \leq f_D(x) \tag{7.21}$$

where f_D is a continuous function of x indipendent on N and so

$$x = qe^{-4J\beta} \tag{7.22}$$

and such that $\lim_{x\to 0} f_D(x) = 0$.

In particular for T sufficiently small

$$\frac{\langle N_- \rangle}{N} < \frac{1}{2} - \varepsilon \tag{7.23}$$

More precisely one has

$$\frac{\langle N_- \rangle}{N} \leq \frac{x^2}{36}\frac{2-x}{(1-x)^2} \tag{7.24}$$

where $x = qe^{-4J\beta} < 1$.

Note that above bound gives also a lower bound on the critical temperature

$$\frac{\langle N_- \rangle}{N} \leq \frac{x^2}{36} \frac{2-x}{(1-x)^2} < \frac{1}{2} - \varepsilon \tag{7.25}$$

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

$$\frac{x^2}{36} \frac{2-x}{(1-x)^2} = \frac{1}{2} \tag{7.26}$$

From the solution $x_{1/2}$ and the condition $x_c > x_{1/2}$ one has

$$J\beta_c \leq J\beta_{1/2} \tag{7.27}$$

where $J\beta_{1/2} = \frac{1}{4} \log q / x_{1/2}$ which implies $T_c > T_{1/2}$.

Exercise 4. The equation

$$x^3 + 16x^2 - 36x + 18 = 0 \tag{7.28}$$

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)**.**

$$\mathcal{H}_{\text{Ising}} = -\sum_{i<j} J_{ij} \sigma_i \sigma_j \tag{7.29}$$

where $\sigma_i \in \{-1, 1\}$. The symmetry gropu of this Hamiltonian is $\mathbb{Z}^2$, which has two elements $\{\mathbb{1}, \mu\}$. We have

$$\mathbb{1}: \text{ identity}, \quad \mu\sigma_i = -\sigma_i, \quad \mu^2 = \mathbb{1} \tag{7.30}$$

Example 19 (Potts model)**.**

$$\mathcal{H}_{q-\text{Potts}} = -\sum_{i<j} J_{ij} \delta_{\sigma_i, \sigma_j} \tag{7.31}$$

where $\sigma_i \in [1, 2, 3, \ldots, q]$. $\mathcal{H}_{q-\text{Potts}}$ is invariant under the permutation group of the sequence $\{1, 2, 3, \ldots, q\}$. There are $q!$ elemtns, for example $\{2, 1, 3, \ldots, q\}$. 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*

$$\mu = \begin{pmatrix} 1 & 2 & \ldots & q-1 & q \\ 2 & 3 & \ldots & q & 1 \end{pmatrix} \tag{7.32}$$

and its powers μ^l with $l = 0, \ldots, q-1$. 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 $\theta_i \to \theta_i + \alpha$. Indeed the Hamiltonian of this model is

$$\mathcal{H}_{XY} = -\sum_{i<j} J_{ij} \vec{\mathbf{S}}_i \cdot \vec{\mathbf{S}}_j \tag{7.33}$$

where $\vec{\mathbf{S}}_i$ is a $2D$ spin vector

$$\vec{\mathbf{S}}_i = (S_{x_i}, S_{y_i}) \tag{7.34}$$

that can assume values on the unit circle ($\left|\vec{\mathbf{S}}_i\right| = 1$).

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.

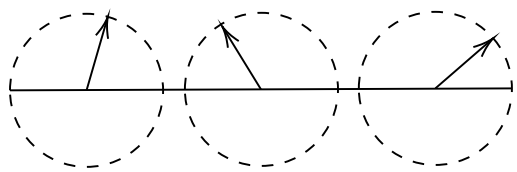


Figure 7.2

The simplest way to parametrize the Hamiltonian is by the angle. Denoting by θ_i the direction angle of spins $\vec{\mathbf{S}}_i$, the $\mathcal{H}_{XY}$ can be written as

$$\mathcal{H}_{XY} = -\sum_{i<j} J_{ij} \cos(\theta_i - \theta_j) \tag{7.35}$$

with $\theta_i \in [0, 2\pi]$.

Remark. The interaction term $\cos(\theta_i - \theta_j)$ can be written also as

$$\frac{1}{2}\left(Z_i^* Z_j + Z_i Z_j^*\right) \tag{7.36}$$

where $Z_j = \exp(i\theta_j)$.

The model is invariant under the global transformation

$$Z_i \to e^{i\alpha} Z_i \tag{7.37}$$

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

$$\hat{\mathbf{\Omega}}_i \cdot \hat{\mathbf{\Omega}}_j \tag{7.38}$$

where $\hat{\mathbf{\Omega}}_i = (\cos\theta_i, \sin\theta_i)$.

Remark. In n-dimensions $\hat{\mathbf{\Omega}}$ has n components $\hat{\mathbf{\Omega}} = \{\Omega^1, \Omega^2, \ldots, \Omega^n\}$ and the corresponding Hamiltonian

$$\mathcal{H} = -\sum_{i>j} J_{ij} \hat{\mathbf{\Omega}}_i \cdot \hat{\mathbf{\Omega}}_j \tag{7.39}$$

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?

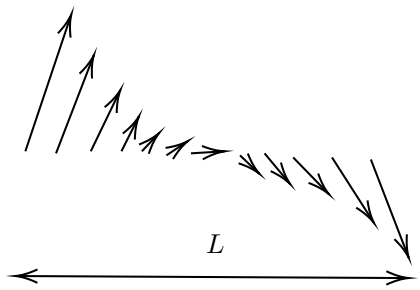


Figure 7.3

7.4 Continuous symmetries and phase transitions

The questions is: when it is not continuous which is the boundary? When the symmetry is continuous the domain walls interpolate smootly 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 neirest neighbours sites is very small, i.e. $(\theta_i - \theta_j) \ll 1$ for i, j neirest neighbours. Now we can diluite the energy, in other words weak the energy term.

Let us do a Taylor expainding of the interaction term

$$\cos(\theta_i - \theta_j) \simeq 1 - \frac{1}{2}(\theta_i - \theta_j)^2 \Rightarrow \sum_{\langle ij \rangle} \left(1 - \frac{1}{2}(\theta_i - \theta_j)^2\right) \tag{7.40}$$

The Hamiltonian can be written as

$$\mathcal{H} \simeq -J \sum_{\langle ij \rangle} \left(1 - \frac{1}{2}(\theta_i - \theta_j)^2\right) \tag{7.41}$$

The (7.40) corresponds to the discrete differential operator where $\theta_i - \theta_j = \partial_x \theta$, therefore

$$\mathcal{H} = E_0 + \underbrace{\frac{J}{2} \int \mathrm{d}\vec{\mathbf{r}}\, (\boldsymbol{\nabla}\theta)^2}_{E \equiv \text{Stifness energy}} \tag{7.42}$$

where $E_0 = 2JN$ is the energy corresponding to the case in which all the spins are oriented along a given direction.

Definition 5 (Stifness energy)**.** The Stifness energy is defined as

$$E = \frac{J}{2} \int \mathrm{d}\vec{\mathbf{r}}\, (\boldsymbol{\nabla}\theta)^2 \tag{7.43}$$

where $\theta(\vec{\mathbf{r}})$ is the angle of a local rotation around an axis and J is the *spin rigidity*. For an ordered phase $\theta(\vec{\mathbf{r}}) = \theta_0$.

Let us now imagine a domain wall where $\theta(\vec{\mathbf{r}})$ rotates by 2π (or $2\pi m$) by using the entire length of the system (see again Figure 7.3):

$$\theta(\vec{\mathbf{r}}) = \frac{2\pi n x}{L} \tag{7.44}$$

where n is the total number of 2π turn of θ in L. Note that there is no variation along the other $D-1$ dimensions, therefore we just doing over one dimension.

Consider only the term called E

$$E = \frac{J}{2} L^{D-1} \int_0^L \mathrm{d}x \left(\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{2\pi n x}{L}\right)\right)^2 = \frac{J}{2} L^{D-1} \int_0^L \mathrm{d}x \left(\frac{2\pi n}{L}\right)^2 \approx 2\pi^2 n^2 J L^{D-2} \tag{7.45}$$

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 dound 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 dos not display an ordered phase, there exist at $T \neq 0$ a special phase transition known as the *Kosterlitz-thouless transition.* This transition dows 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

$$J_{ij} = \frac{J}{|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|^\alpha} \tag{7.46}$$

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 indipendently on their distance. No metric is involed (instead of previously where the definition of J of before is a metric.). It can be solve exactly and later we will seen why.

7.6 Ising model with infinite range

The Hamiltonian is the following:

$$-\mathcal{H}_N(\{S\}) = \frac{J_0}{2} \sum_{i,j}^{N} S_i S_j + H \sum_i S_i \tag{7.47}$$

with $S_i \in [-1, +1]$. The problem is the double sum over i,j, indeed

$$\sum_{i,j} S_i S_j \propto O(N^2) \tag{7.48}$$

and the thermodynamic limit is ill-defined. To circumvent this problem *Kac* suggested to consider a strength

$$J_0 = \frac{J}{N} \tag{7.49}$$

this is called the *kac* approximation

$$-\mathcal{H}_N(\{S\}) = \frac{J}{2N}\sum_{i,j}^{N} S_i S_j + H\sum_i S_i \tag{7.50}$$

with this choice we recover $E \sim O(N)$. In this Hamiltonian since you have no metric you have no dimension.

Lecture 11. Wednesday 20th November, 2019. Compiled: Thursday 26th December, 2019.

The partition function is

$$Z_N(T,J,H) = \sum_{\{S\}} \exp\left[\frac{\beta J}{2N}\sum_{ij} S_i S_j + \beta H \sum_i S_i\right] \tag{7.51}$$

Since there are no restrictions on the double sum we can write

$$\sum_{ij} S_i S_j = \left(\sum_i S_i\right)\left(\sum_j S_j\right) = \left(\sum_i S_i\right)^2 \tag{7.52}$$

Rewriting the partition function:

$$Z_N(T,J,H) = \sum_{\{S\}} \exp\left[\frac{K}{2N}\left(\sum_i S_i\right)^2 + h\sum_i S_i\right] \tag{7.53}$$

In order to transform the quadratic term into al inear 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:

$$x \equiv \sum_i S_i \tag{7.54}$$

The following identity holds

$$e^{\frac{Kx^2}{2N}} = \sqrt{\frac{NK}{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{NK}{2}y^2 + Kxy}\,\mathrm{d}y\,, \quad \text{with} \quad \mathrm{Re}(K) > 0 \tag{7.55}$$

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))

$$-\frac{NK}{2}y^2 + Kxy = -\frac{NK}{2}\left(y - \frac{x}{N}\right)^2 + \frac{Kx^2}{2N} \tag{7.56}$$

Then we integrate:

$$e^{\frac{Kx^2}{2N}}\int_{-\infty}^{+\infty} e^{-\frac{NK}{2}\left(y-\frac{x}{N}\right)^2}\,\mathrm{d}y \overset{(a)}{=} e^{\frac{Kx^2}{2N}}\sqrt{\frac{2\pi}{NK}} \tag{7.57}$$

where in (a) we have considered $z \equiv \left(y - \frac{x}{N}\right)$, $\mathrm{d}z = \mathrm{d}y$:

$$\to \int_{-\infty}^{+\infty} e^{-\alpha z^2}\,\mathrm{d}z = \sqrt{\frac{\pi}{\alpha}} \tag{7.58}$$

with $\alpha \equiv \frac{NK}{2}$. □

By using (7.55) we have

$$Z_N(K,h) = \sqrt{\frac{NK}{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}y\, e^{-\frac{NK}{2}y^2} \underbrace{\left[\sum_{\{S\}} e^{(h+Ky)\sum_i S_i} \right]}_{Q_y} \tag{7.59}$$

Sometimes y is called *auxiliary field*. If y is a fluctuating external field with Gaussian distribution

$$\begin{aligned} Q_y = \sum_{\{S\}} e^{(h+Ky)\sum_i S_i} &= \prod_i^N \left(\sum_{S_i=\pm 1} \exp[(h+Ky)S_i] \right) \\ &= (2\cosh(h+Ky))^N \end{aligned} \tag{7.60}$$

The partition function becomes

$$Z_N(K,h) = \sqrt{\frac{NK}{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}y\, e^{-\frac{NK}{2}y^2} (2\cosh(h+Ky))^N = \sqrt{\frac{NK}{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}y\, e^{N\alpha(K,h,y)} \tag{7.61}$$

where

$$\alpha(K,h,y) = \ln\left[2\cosh(h+Ky)\right] - \frac{K}{2}y^2 \tag{7.62}$$

Remark. In the limit $N \to \infty$ 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 loosing information, it depends on the form of the function. For example, for a delta function it works better. In general:

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}y \to f(\bar{x}) \tag{7.63}$$

where $\bar{x} = \max_x f(x)$.

Indeed as $N \to \infty$, since the integrand is $\exp(N\alpha(K,h,y))$, the integral is dominated by the global maximum in y of the function $\alpha(K,h,y)$.

$$Z_N(K,h) \overset{N\gg 1}{\approx} \sqrt{\frac{NK}{2\pi}} \max_y \left[e^{N\alpha(K,h,y)} \right] \tag{7.64}$$

Let y_s be the value of y at which

$$\alpha(K,h,y_s) = \max_y \alpha(K,h,y) \tag{7.65}$$

therefore

$$Z_N(K,h) \overset{N\gg 1}{\approx} \sqrt{\frac{NK}{2\pi}} e^{N\alpha(K,h,y_s)} \tag{7.66}$$

When we are ample to compute the y_s we can do this approximation and

$$f_b(K,h) = \lim_{N\to\infty} \frac{1}{N}(-k_B T \log Z_N) = -k_B T \alpha(K,h,y_s) \tag{7.67}$$

Looking for y_s we consider the condition $\frac{\partial \alpha}{\partial y} = 0$:

$$\frac{\partial \alpha}{\partial y} = \frac{\sinh(h+Ky)K}{\cosh(h+Ky)} - Ky = 0 \quad \Rightarrow y_s = \tanh(h+Ky_s) \tag{7.68}$$

The last one is an implicit equation that can be solved graphically as a function of K and h.

The magnetization in the $N \to \infty$ limit is given by

$$\begin{aligned} m &= -\left(\frac{\partial f}{\partial H}\right)_T = \lim_{N\to\infty} \frac{1}{\beta N} \frac{\partial \ln Z_N(K,h)}{\partial H} \\ &= \frac{\partial \alpha(K,h,y_s)}{\partial h} + \frac{O(\log N)}{N} = \frac{2\sinh(Ky_s+h)}{2\cosh(ky_s+h)} \\ &= \tanh(Ky_s+h) \end{aligned} \tag{7.69}$$

At the end $m \equiv y_s$

$$m = \tanh(h + Km) \tag{7.70}$$

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 negletic the correlation between the fluctuations of the order parameter. It is equivalent to a statistical indipendence of the microscopic degrees of freedom.

The mean field approximation has different names depending on the system considered:

- Magnetic: Weiss theory.
- Fluids: Van der Walls.
- Polymers: Flory's theory.

8.1.1 Mean field for Ising model

Let us start from the following generic model

$$\mathcal{H}[\{S\}] = -\frac{1}{2}\sum_{ij} J_{ij} S_i S_j - H \sum_i S_i \tag{8.1}$$

where the double sum over i and j have no restrictions, while H is homogeneous.

$$Z_N(T, H, \{J_{ij}\}) = \sum_{\{S\}} e^{-\beta\mathcal{H}[\{S\}]} = \exp(-\beta F_N(T, H, \{J_{ij}\})) \tag{8.2}$$

Since H is uniform, the magnetization per spin is

$$\langle S_i \rangle = \langle S \rangle \equiv m \tag{8.3}$$

Let us now consider the idenity

$$\begin{aligned} S_iS_j &= (S_i - m + m)(S_j - m + m) \\ &= (S_i - m)(S_j - m) + m^2 + m(S_j - m) + m(S_i - m) \end{aligned} \tag{8.4}$$

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,

$$S_iS_j \approx m^2 + m(S_i - m) + m(S_j - m) \tag{8.5}$$

we obtain

$$\frac{1}{2}\sum_{i,j} J_{ij}S_iS_j \overset{MF}{\approx} \frac{1}{2}\sum_{i,j} J_{ij}\left[-m^2 + m(S_i + S_j)\right] \tag{8.6}$$

Let us focus on the term

$$\frac{1}{2}\sum_{i,j} J_{ij}m(S_i + S_j) = 2\frac{1}{2}m\sum_{i,j} J_{ij}S_i \tag{8.7}$$

If we do not many assumption on J_{ij} the mean field hamiltonian is

$$\mathcal{H}_{MF}[\{S\}] = \frac{1}{2}m^2\sum_{ij} J_{ij} - m\sum_{ij} J_{ij}S_i - H\sum_i S_i \tag{8.8}$$

and by calling

$$\bar{J}_i \equiv \sum_j J_{ij} \tag{8.9}$$

we get

$$\mathcal{H}_{MF}[\{S\}] = \frac{1}{2}m^2\sum_i \bar{J}_i - \frac{m}{2}\sum_i \bar{J}_iS_i - H\sum_i S_i \tag{8.10}$$

Remark. Note the coefficient 1/2 that is needed to avoid the double counting of bonds.

If moreover we suppose that

$$\bar{J}_i \to \bar{J}$$

we have

$$\mathcal{H}_{MF}[\{S\}] = \frac{1}{2}m^2N\bar{J} - \left(\frac{m}{2}\bar{J} + H\right)\sum_i S_i \tag{8.11}$$

In the standard Ising model, where

$$\frac{1}{2}\sum_{ij} J_{ij}S_iS_j \to \sum_{\langle ij\rangle} J_{ij}S_iS_j \tag{8.12}$$

the term $2m\sum_{\langle ij\rangle} J_{ij}S_i$ can be written as follows; let

$$\sum_{j\in n.n. \text{ of } i} J_{ij} = z\hat{J}_i \tag{8.13}$$

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) becames

$$2m \sum_{\langle ij \rangle} J_{ij} S_i = 2m \frac{1}{2} z \hat{J} \sum_{i=1}^{N} S_i \tag{8.14}$$

Hence,

$$\mathcal{H}_{MF}[\{S\}] = \frac{1}{2} m^2 N z \hat{J} - (m z \hat{J} + H) \sum_{i=1}^{N} S_i \tag{8.15}$$

$$\begin{aligned} Z_N(T, H, \hat{J}) &= e^{-N\beta \hat{J} \frac{z}{2} m^2} \sum_{\{S\}} e^{\beta m (\hat{J} z m + H) \sum_i S_i} \\ &= e^{-N\beta \hat{J} \frac{z}{2} m^2} \left(\sum_{S=\pm 1} \exp\Big(\beta^m \Big(\hat{J} z m + H \Big) \Big) \right)^N \\ &= e^{-N\beta \hat{J} \frac{z}{2} m^2} \Big(2 \cosh \Big[\beta \Big(\hat{J} z m + J \Big) \Big] \Big)^N \end{aligned} \tag{8.16}$$

Remark. We are replacing the interaction of the J with a field close to the S_i. We called $\hat{J} z m = H_{eff}$, the mean field!

The free energy per spin is

$$\begin{aligned} \frac{F_N(T, H, \hat{J})}{N} &= \frac{1}{N} \Big(-k_B T \ln Z_N(T, H, \hat{J}) \Big) \\ &= \frac{1}{2} \hat{J} z m^2 - k_B T \ln \Big[\cosh \Big(\beta (\hat{J} z m + H) \Big) \Big] - k_B T \ln 2 \end{aligned} \tag{8.17}$$

Sometimes it is useful to use the *dimensionless variables*

$$\bar{f} \equiv \frac{F_N}{N z \hat{J}}, \quad \theta \equiv \frac{k_B T}{z \hat{J}}, \quad \bar{H} \equiv \frac{H}{z \hat{J}} \tag{8.18}$$

Hence,

$$\bar{f}(m, \bar{H}, \theta) = \frac{1}{2} m^2 - \theta \ln \big(2 \cosh \big(\theta^{-1} (m + \bar{H}) \big) \big) \tag{8.19}$$

In order to be a self-consistent, the last equation has to satisfy the thermodynamic relation:

$$m = -\left(\frac{\partial f}{\partial H} \right)_T \quad \Rightarrow m = \tanh \Big(\beta (\hat{J} z m + H) \Big) \tag{8.20}$$

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

$$m = \tanh \Big(\beta (\hat{J} z m) \Big) \tag{8.21}$$

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

$$\frac{z \hat{J}}{k_B T_c} = 1 \Rightarrow T_c = \frac{z \hat{J}}{k_B} \neq 0! \tag{8.22}$$

Remark. T_c depends on z and hence on D!

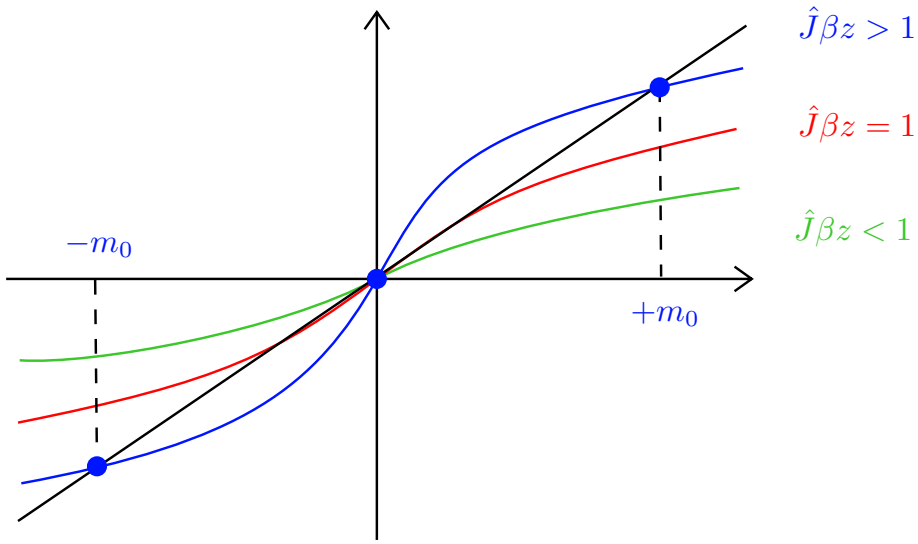


Figure 8.1

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 $H = 0$:

$$f(m, 0, T, \hat{J}) = \frac{1}{2}\hat{J}zm^2 - k_B T \ln\left[\cosh\Big(\beta\hat{J}zm\Big)\right] \tag{8.23}$$

Define $x \equiv \beta\hat{J}zm \simeq 0$, hence

$$\cosh(x) \simeq 1 + \underbrace{\frac{x^2}{2} + \frac{x^4}{4!}}_{t\simeq 0} + \dots \tag{8.24}$$

$$\log(1+t) \simeq t - \frac{1}{2}t^2 \tag{8.25}$$

$$\log(\cosh x) \simeq \frac{x^2}{2} + \frac{x^4}{4!} - \frac{1}{2}\frac{x^4}{4} + O(x^6) = \frac{x^2}{2} - \frac{x^4}{12} + O(x^6) \tag{8.26}$$

This gives

$$f(m, 0, T, \hat{J}) \simeq const + \frac{A}{2}m^2 + \frac{B}{4}m^4 + O(m^6) \tag{8.27}$$

with

$$A \equiv \hat{J}z\Big(1 - \beta\hat{J}z\Big) \tag{8.28a}$$

$$B \equiv \beta^2\frac{(\hat{J}z)^4}{3} > 0 \tag{8.28b}$$

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 $m = 0$.

Lecture 12.
Friday 22nd November, 2019.
Compiled: Thursday 26th December, 2019

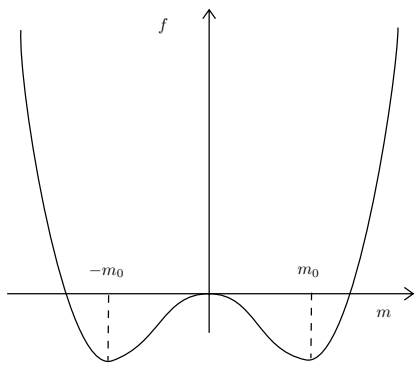


Figure 8.2: Plot of the free energy: case $\beta\hat{J}z > 1 \Rightarrow A < 0$.

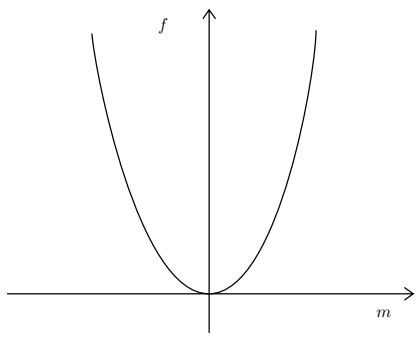


Figure 8.3: Plot of the free energy: case $\beta\hat{J}z < 1 \Rightarrow A > 0$.

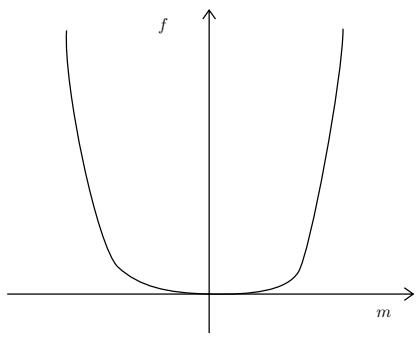


Figure 8.4: Plot of the free energy: case $\beta\hat{J}z = 1 \Rightarrow A = 0$.

8.1.3 Mean field critical exponents

Let us consider

$$f(m, T, 0) \approx const + \frac{A}{2}m^2 + \frac{B}{4}m^4 + O(m^6) \tag{8.29}$$

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.

β exponent

The β exponential observe the order parameter. Consider $H = 0$, $t \equiv \frac{T-T_c}{T_c}$ and $m \overset{t\to 0^-}{\sim} -t^\beta$. The condition of equilibrium is

$$\frac{\partial f}{\partial m} = 0 \tag{8.30}$$

which implies

$$\left.\frac{\partial f}{\partial m}\right|_{m=m_0} = Am_0 + Bm_0^3 = \left[\hat{J}z(1-\beta\hat{J}z) + Bm_0^2\right]m_0 = 0 \tag{8.31}$$

Since $T_c = \frac{\hat{J}z}{k_B}$:

$$0 = \frac{k_B T_c}{T}(T-T_c)m_0 + Bm_0^3 \tag{8.32}$$

The solution are $m_0 = 0$ and

$$m_0 \simeq (T_c - T)^{1/2} \tag{8.33}$$

The mean field value is so $\beta = 1/2$.

δ exponent

Now, let us concentrate in the δ 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

$$m = \tanh\Big(\beta(\hat{J}zm + H)\Big) \tag{8.34}$$

Inverting it

$$\beta(\hat{J}zm + H) = \tanh^{-1} m \tag{8.35}$$

On the other hand, for $m \sim 0$

$$\tanh^{-1} m \simeq m + \frac{m^3}{3} + \frac{m^5}{5} + \dots \tag{8.36}$$

Therefore, by substituting

$$\begin{aligned} H = k_B T\left(m + \frac{m^3}{3} + \dots\right) - \hat{J}zm &= \Big(k_B T - \hat{J}z\Big)m + k_B T\frac{m^3}{3} + \dots \\ &\simeq k_B(T-T_c)m + \frac{k_B T}{3}m^3 \end{aligned} \tag{8.37}$$

At $T = T_c = \frac{\hat{J}z}{k_B}$, we have

$$H \sim k_B T_c \frac{m^3}{3} \tag{8.38}$$

The mean field value is $\delta = 3$.

α exponent

Consider the α 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 $m_0 = 0$,

$$f(m,H) = \frac{\hat{J}zm^2}{2} - \frac{1}{\beta}\ln\left(2\cosh\Big(\beta(\hat{J}zm + H)\Big)\right) \tag{8.39}$$

If $m = 0$, $\cosh 0 = 1$ and

$$f = -k_B T \ln 2 \tag{8.40}$$

therefore

$$c_H = -T\left(\frac{\partial^2 f}{\partial T^2}\right) = 0 \tag{8.41}$$

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 $A = const$. We obtain $\alpha = 0$ also in this case.

$$m_0 = \pm\sqrt{-\frac{\hat{J}z}{2T_c}(T-T_c)} \tag{8.42}$$

γ exponent

Now we consider the γ exponent, for $H = 0$, $\chi \sim t^{-\gamma}$. Starting again from equation (8.34):

$$m = \tanh\Big(\beta(\hat{J}zm + H)\Big)$$

and developing it around $m \simeq 0$, as shown before we get

$$H = mk_B(T - T_c) + \frac{k_B T}{3} m^3 \tag{8.43}$$

$$\Rightarrow \chi_T = \frac{\partial m}{\partial H} = \frac{1}{\frac{\partial H}{\partial m}} \tag{8.44}$$

Since $\frac{\partial H}{\partial m} \simeq k_B(T - T_c) + K_B T m^2$, as $m \to 0$

$$\chi \sim (T - T_c)^{-1} \tag{8.45}$$

The mean field value is $\gamma = 1$.

Summary

The mean field critical exponents are

$$\beta = \frac{1}{2}, \quad \gamma = 1, \quad \delta = 3, \quad \alpha = 0 \tag{8.46}$$

Remark. In the mean field critical exponents the dimension D does not appear. T_c instead depends on the number of z of neirest neighbours and hence on the embedding lattice!

Remark. (lesson) The ν 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 negleting correlation. How we can compute the correlation function within the meanfield theory with thermal fluctuations? We look at the response of the system. Esperimentally 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 happend 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 flexibile to deal with complex systems.

The method is similar to the one used in quantum mechanics, namely it is based on the following inequality

$$E_\alpha = \langle \psi_\alpha | \hat{H} | \psi_\alpha \rangle \geq E_0 \tag{8.47}$$

valid for all trial function ψ_α.

Remark. E_0 is the ground state energy.

Example 20. In many mody problem we have Hartee and Hartree-Fock.

The closest bound to E_0 is the one that is obtained by minimizing E_α, 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 Φ be a random variable (either discrete or continuous) and let $f(\Phi)$ be a function of it.

   For all function f of Φ, the mean value with respect to a distribution function $p(\Phi)$ is given by
   $$\langle f(\Phi)\rangle_p \equiv \mathrm{Tr}(p(\Phi)f(\Phi)) \tag{8.48}$$
   If we consider the function
   $$f(\Phi) = \exp[-\lambda\Phi] \tag{8.49}$$
   it is possible to show the inequality
   $$\boxed{\left\langle e^{-\lambda\Phi}\right\rangle_p \geq e^{-\lambda\langle\Phi\rangle_p}, \quad \forall p} \tag{8.50}$$

   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
   $$\rightarrow \left\langle e^{-\lambda\Phi}\right\rangle_p \geq \left\langle (1-\lambda(\Phi-\langle\Phi\rangle))e^{-\lambda\langle\Phi\rangle}\right\rangle_p = e^{-\lambda\langle\Phi\rangle_p}$$
   □

2. The second inequality refers to the free energy. Let $\rho(\Phi)$ be a probability distribution, i.e. such that
   $$\mathrm{Tr}(\rho(\Phi)) = 1 \quad \rho(\Phi)\geq 0 \quad \forall\Phi \tag{8.51}$$
   Hence,
   $$e^{-\beta F_N} = Z_N = \mathrm{Tr}_{\{\Phi\}}\, e^{-\beta\mathcal{H}[\{\Phi\}]} = \mathrm{Tr}_{\{\Phi\}}\, \rho e^{-\beta\mathcal{H}-\ln\rho} = \left\langle e^{-\beta\mathcal{H}-\ln\rho}\right\rangle_\rho \tag{8.52}$$
   From the inequality (8.50),
   $$e^{-\beta F_N} = \left\langle e^{-\beta\mathcal{H}-\ln\rho}\right\rangle_\rho \geq e^{-\beta\langle\mathcal{H}\rangle_\rho-\langle\ln\rho\rangle_\rho} \tag{8.53}$$
   Taking the logs one has
   $$F \leq \langle\mathcal{H}\rangle_\rho + k_BT\langle\ln\rho\rangle_\rho = \mathrm{Tr}(\rho\mathcal{H}) + k_BT\,\mathrm{Tr}(\rho\ln\rho) \equiv F_\rho \tag{8.54}$$
   Whenever we are able to write the last equation by using a ρ, then we will minimize. This is the variational approach of statistical mechanics. The question is: which is the ρ that minimize?

   The functional F_ρ will reach its minimum value with respect to the variation of ρ with the consrtraint $\mathrm{Tr}(\rho)=1$, when
   $$\boxed{\bar{\rho} = \rho_{eq} = \frac{1}{Z}e^{-\beta\mathcal{H}}} \tag{8.55}$$
   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 ρ_{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 everithing is exact. The idea is to choose a functional form of ρ and then minimize F_ρ with respect to ρ. Note that ρ is the $N-$point probability density function (it is a function of all the degrees of freedom):

$$\rho = \rho(\Phi_1, \dots, \Phi_N) \tag{8.56}$$

it is a $N-$body problem. The Φ_α 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:

$$\psi_\alpha(\vec{\mathbf{r}}_1, \vec{\mathbf{P}}_1, \dots, \vec{\mathbf{r}}_N, \vec{\mathbf{P}}_N) \tag{8.57}$$

The mean-field approximation consists in factorising ρ into a product of $1-$point distribution function:

$$\boxed{\rho(\Phi_1, \dots, \Phi_N) \overset{MF}{\simeq} \prod_{\alpha=1}^{N} \rho^{(1)}(\Phi_\alpha) \equiv \prod_{\alpha=1}^{N} \rho_\alpha} \tag{8.58}$$

where we have used the short-hand notation $\rho^{(1)}(\Phi_\alpha) \to \rho_\alpha$.

Remark. Approximation (8.58) is equivalent to assume statistical indipendence between particles (or more generally between different degrees of freedom). The indipendence of the degree of freedom is a very strong assumption!

Example 21. Let us consider the spin model on a lattice; what is the Φ_α? We have:

$$\Phi_\alpha \to S_i$$

Hence, $\rho = \rho(S_1, S_2, \dots, S_N)$ and (8.58) becomes

$$\rho \overset{MF}{\simeq} \prod_{i=1}^{N} \rho^{(1)}(S_i) \equiv \prod_{i=1}^{N} \rho_i \tag{8.59}$$

With (8.58) and $\mathrm{Tr}(\rho_\alpha) = 1$, we compute the two averages in the (8.54) given the field. We have:

$$\mathrm{Tr}_{\{\Phi\}}(\rho \ln \rho) = \mathrm{Tr}\left(\prod_\alpha \rho_\alpha \left(\sum_\alpha \ln \rho_\alpha\right)\right) \overset{\text{to do}}{=} \sum_\alpha \mathrm{Tr}^{(\alpha)}(\rho_\alpha \ln \rho_\alpha) \tag{8.60}$$

where $\mathrm{Tr}^{(\alpha)}$ means sum over all possible values of the random variable Φ_α (with α fixed).

Remark. We have $\mathrm{Tr}^{(\alpha)} \rho_\alpha = 1$.

We end up that

$$F_{\rho_{MF}} = \langle \mathcal{H} \rangle_{\rho_{MF}} + k_B T \sum_\alpha \mathrm{Tr}^{(\alpha)}(\rho_\alpha \ln \rho_\alpha) \tag{8.61}$$

Remark. $F_{\rho_{MF}} = F(\{\rho_\alpha\})$ and we have to minimize it with respect to ρ_α.

How can we parametrixe ρ_α? There are two approaches that are mostly used:

1. Parametrize $\rho_\alpha \equiv \rho^{(1)}(\Phi_\alpha)$ by the average of Φ_α with respect to ρ_α, $\langle \Phi_\alpha \rangle_{\rho_\alpha}$ (in general is the local order parameter):

$$\rho_\alpha = \rho^{(1)}(\Phi_\alpha) \to \langle \Phi_\alpha \rangle_{\rho_\alpha}$$

This means that there are two constraints in the minimization procedure:

$$\mathrm{Tr}^{(\alpha)} \rho_\alpha = 1, \quad \mathrm{Tr}^{(\alpha)}(\rho_\alpha \Phi_\alpha) = \langle \Phi_\alpha \rangle \tag{8.62}$$

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 ρ_α itself the variational parameter. $F_{\rho_{MF}}$ is minimized by varying ρ_α. 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:

$$\mathcal{H}[\{S\}] = -J \sum_{\langle ij \rangle} S_i S_j - \sum_i H_i S_i \tag{8.63}$$

it means that

$$\Phi_\alpha \to S_i = \pm 1 \tag{8.64}$$

and that the order parameter (variational parameter) becomes

$$\langle \Phi_\alpha \rangle \to \langle S_i \rangle \equiv m_i \tag{8.65}$$

Remark. Note that this time $H \to H_i$ (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

$$\rho^{(1)} \equiv \rho^{(1)}(S_i) \to \begin{cases} \mathrm{Tr}\, \rho_i^{(1)} = 1 \\ \mathrm{Tr}\, \rho_i^{(1)} S_i = m_i \end{cases} \tag{8.66}$$

Since we have to satisdy these two constraints, we need two free parameters. A linear functional form is sufficient. Denoting by:

- a: stastistical 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

$$\rho_i \equiv \rho^{(1)}(S_i) = a(1 - \delta_{S_i,1}) + b\delta_{S_i,1} \tag{8.67}$$

Lecture 13. *Wednesday 27th November, 2019. Compiled: Thursday 26th December, 2019.*

Using the constraints

$$\begin{cases} \mathrm{Tr}^{(i)}(\rho_i) = 1 & \to a + b = 1 \\ \mathrm{Tr}^{(i)}((\rho_i S_i) = m_i & \to a - b = m_i \end{cases} \tag{8.68}$$

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:

$$\begin{cases} a = \frac{1-m_i}{2} \\ b = \frac{1+m_i}{2} \end{cases} \tag{8.69}$$

Hence,

$$\rho_i = \frac{1-m_i}{2}(1-\delta_{S_i,-1}) + \frac{1+m_i}{2}\delta_{S_i,-1} \tag{8.70}$$

In matrix form

$$\begin{pmatrix} \frac{(m_i+1)}{2} & 0 \\ 0 & \frac{(1-m_i)}{2} \end{pmatrix} \tag{8.71}$$

Mean-field energy term

Let us consider the Hamiltonian

$$\langle \mathcal{H} \rangle_{\rho_{MF}} = \left\langle -J \sum_{\langle ij \rangle} S_i S_j - \sum_i H_i S_i \right\rangle_{\rho_{MF}} = -J \sum_{\langle ij \rangle} \langle S_i S_j \rangle_{\rho_{MF}} - \sum_i H_i \langle S_i \rangle_{\rho_{MF}} \tag{8.72}$$

Since

$$\rho_{MF} = \prod_{i=1}^{N} \rho_i \tag{8.73}$$

the term $\langle S_i S_j \rangle_{\rho_{MF}}$ will trasform into

$$\langle S_i S_j \rangle_{\rho_{MF}} = \langle S_i \rangle_{\rho_{MF}} \langle S_j \rangle_{\rho_{MF}} \tag{8.74}$$

Moreover, for all function g of S_i we can write

$$\begin{aligned} \langle g(S_i) \rangle_{\rho_{MF}} &= \mathrm{Tr}^{(i)}(g(S_i)\rho_i) = \sum_{S_i=\pm 1} g(S_i)\rho_i \\ &= \sum_{S_i=\pm 1} g(S_i)\left[\frac{1+m_i}{2}\delta_{S_i,1} + \frac{1-m_i}{2}(1-\delta_{S_i,1})\right] \\ &= \frac{1+m_i}{2} g(1) + \frac{1-m_i}{2} g(-1) \end{aligned} \tag{8.75}$$

Note that, if $g(S_i) = S_i$,

$$\langle S_i \rangle_{\rho_{MF}} = m_i$$

as expected.

Hence,

$$\langle \mathcal{H} \rangle_{\rho_{MF}} = -J \sum_{\langle ij \rangle} m_i m_j - \sum_i H_i m_i \tag{8.76}$$

Remark. This has the form of the original Hamiltonian where S_i have been replaced by their statistical averages.

The entropy term is:

$$\begin{aligned} \langle \ln \rho \rangle_{\rho_{MF}} &= \mathrm{Tr}(\rho \ln \rho) \overset{MF}{=} \sum_i \mathrm{Tr}^{(i)}(\rho_i \ln \rho_i) \\ &= \sum_i \left[\frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2}\right] \end{aligned} \tag{8.77}$$

The total free energy becames:

$$\begin{aligned} F_{\rho_{MF}} &= \langle \mathcal{H} \rangle_{\rho_{MF}} + k_B T \langle \ln \rho \rangle_{\rho_{MF}} \\ &= -J \sum_{\langle ij \rangle} m_i m_j - \sum_i H_i m_i + k_B T \sum_i \left[\frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \right] \end{aligned} \tag{8.78}$$

We now look for the values $m_i = \bar{m}_i$, that minimizes $F_{\rho_{MF}}$ (equilibrium phases):

$$\left. \frac{\partial F_{\rho_{MF}}}{\partial m_i} \right|_{m_i = \bar{m}_i} = 0 \tag{8.79}$$

This gives:

$$0 = -J \sum_{j \in \, n.n. \, \text{of} \, i} \bar{m}_j - H_i + \frac{k_B T}{2} \ln \left[\frac{1+\bar{m}_i}{1-\bar{m}_i} \right] \tag{8.80}$$

To solve it, remember that

$$\tanh^{-1}(x) = \frac{1}{2} \ln \frac{1+x}{1-x} \quad (TO\,DO) \tag{8.81}$$

Hence,

$$k_B T \tanh^{-1}(\bar{m}_i) = J \sum_{j \in \, n.n. \, \text{of} \, i} \bar{m}_j + H_i \tag{8.82}$$

which implies

$$\bar{m}_i = \tanh \left[(k_B T)^{-1} \left(J \sum_{j \in \, n.n. \, \text{of} \, i} \bar{m}_j + H_i \right) \right] \tag{8.83}$$

Defining

$$z \bar{m}_i \equiv \sum_{j \in \, n.n. \, \text{of} \, i} \bar{m}_j \tag{8.84}$$

we get

$$\bar{m}_i = \tanh \left[\beta (J z \bar{m}_i + H_i) \right] \tag{8.85}$$

this is the Bragg-William approximation.

Example 22 (Ising antiferromagnet in an external field)**.** Consider the model

$$\mathcal{H} = + J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i, \tag{8.86}$$

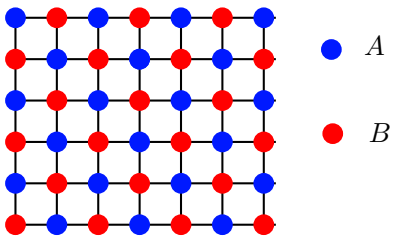
Remark. Note the + before J. This means that the interactions are antiferromagnetic.

- 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.

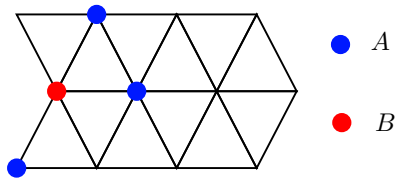
  Remark. FCC is not bipartite, while BCC it is. See Figure 8.5.

  If the lattic 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

$$S'_j \begin{cases} +S_j & j \in A \\ -S_j & j \in B \end{cases} \tag{8.87}$$



(a) Square lattice is bipartite.



(b) Triangular lattice is not bipartite.

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:

$$\mathcal{H}^* = -J \sum_{\langle ij \rangle} S'_i S'_j \tag{8.88}$$

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 $H_B = -H$.

$$\mathcal{H}^*[S] = -J \sum_{\langle r_A r_B \rangle} S(r_A)S(r_B) - H \sum_{r_A} S(r_A) + H \sum_{r_B} S(r_B), \quad J > 0, H > 0 \tag{8.89}$$

  The average magnetization per spin is

$$m \equiv \frac{1}{2}(m_A + m_B) \tag{8.90}$$

  while

$$m_S = \frac{1}{2}(m_A - m_B) \tag{8.91}$$

  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

$$\begin{cases} \rho_A^{(1)}(S) = \frac{1+m_A}{2}\delta_{S,1} + \frac{1-m_A}{2}\delta_{S,-1} \\ \rho_B^{(1)}(S) = \frac{1+m_B}{2}\delta_{S,1} + \frac{1-m_B}{2}\delta_{S,-1} \end{cases} \tag{8.92}$$

  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

$$\langle \mathcal{H} \rangle_{\rho_{MF}} = -J \sum_{\langle ij \rangle} \langle S_i S_j \rangle_{\rho_{MF}} - H \sum_i \langle S_i \rangle_{\rho_{MF}} \tag{8.93}$$

$$\langle \ln \rho \rangle_{\rho_{MF}} = \sum_i \mathrm{Tr}^{(1)}(\rho_i \ln \rho_i) \tag{8.94}$$

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

$$\frac{\partial(F/N)}{\partial m_A} = 0 \qquad \Rightarrow m_B = \frac{H}{z\hat{J}} - \frac{k_B T}{z\hat{J}} \ln\left(\frac{1+m_A}{1-m_A}\right) \quad (8.97a)$$

$$\frac{\partial(F/N)}{\partial m_B} = 0 \qquad \Rightarrow m_A = \frac{H}{z\hat{J}} - \frac{k_B T}{z\hat{J}} \ln\left(\frac{1+m_B}{1-m_B}\right) \quad (8.97b)$$

As before, since

$$\tanh^{-1}(x) = \frac{1}{2} \ln \frac{1+x}{1-x} \quad (8.98)$$

these self-consisten equations can be written as

$$\begin{cases} m_A = \tanh\Big(\beta\Big(H - z\hat{J} m_B\Big)\Big) \\ m_B = \tanh\Big(\beta\Big(H - z\hat{J} m_A\Big)\Big) \end{cases} \quad (8.99)$$

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 $He^3 - He^4$ mixture undergoing a fluid-superfluid transition.

Remark. He^4 is a non radiative isotope with two protons and two neutrons. Roughly 1/4 of the universe matter is He^4! From a quantum statistical point of view He^4 is a *boson*.

A gas of 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.

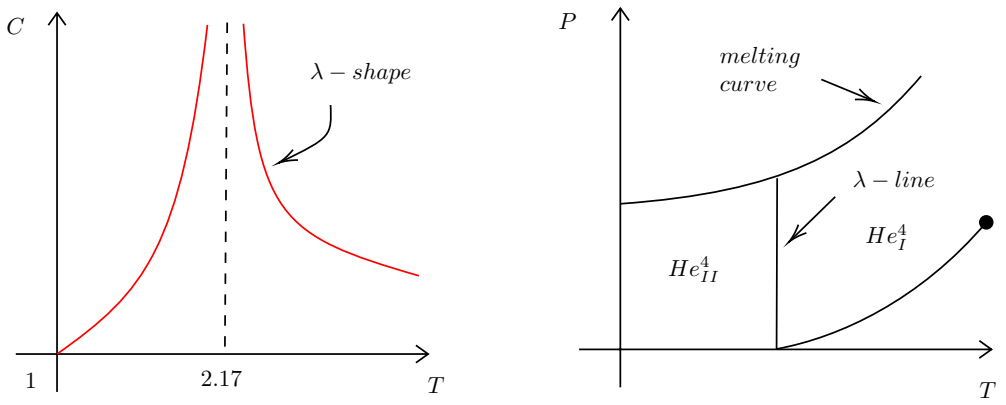
Remark. The $\lambda-$transition is a genuine critical point (second order). For $T < T_\lambda$, He^4 is in the superfluid phase and it can be described by a two-fluids model in which one component has zero viscosity and zero entropy.

Now, the question is: what happens to the system if a given amount of He^3 is inserted to form a $He^3 - He^4$ mixture? He^3 is a non-radiative isotope with 2 protons and 1 neutron. From a quantum statistical point of vieq is a *fermion*.

Hence, if inserted in a system of He^4 it will "dilute" its bosonic property. Then, one expects that T_λ decreases. Indeed, denoting by x the concentration of He^3 one observes

$$T_\lambda = T_\lambda(x) \quad (8.100)$$

with $T_\lambda(x)$ that decreases as x increases.



(a) d

(b) (P, T) phase diagram.

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

$$x > x_t = \frac{n_3}{n_3 + n_4} \sim 0.67 \tag{8.101}$$

the fluid-superfluid transition becames first order! It is accompained 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 diluited 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

$$x \equiv 1 - \langle S_i^2 \rangle \tag{8.102}$$

The chemical potentials difference is

$$\Delta \propto \mu_3 - \mu_4 \tag{8.103}$$

and controls the number of He^3 atoms.

If:

- $\Delta \to -\infty \quad \Rightarrow x \to 0.$
- $\Delta \to +\infty \quad \Rightarrow x \to 1.$

and the order parameter for the $\lambda-$transition becomes

$$\langle S_i \rangle = \begin{cases} 0 & T > T_\lambda \\ m & T < T_\lambda \end{cases} \tag{8.104}$$

The minimal version of the model is:

$$\mathcal{H} = -J \sum_{\langle ij \rangle}^{N} S_i S_j + \Delta \sum_{i=1}^{N} S_i^2 - \Delta N \tag{8.105}$$

Remark. The ΔN term is a typical term for a gas in gran canonical ensemble.

Variational mean field approach to BEG

Since $\rho_{MF} = \prod_i \rho_i$,

$$G(T, J, \Delta) = \langle \mathcal{H} \rangle_{\rho_{MF}} + k_B T \sum_i \mathrm{Tr}(\rho_i \ln \rho_i) \tag{8.106}$$

where the first term can be written as

$$\begin{aligned} \langle \mathcal{H} \rangle_{\rho_{MF}} &= -J \sum_{\langle ij \rangle} \langle S_i S_j \rangle + \Delta \sum_i \left\langle S_i^2 \right\rangle - N\Delta \\ &\overset{MF}{\simeq} -J \sum_{\langle ij \rangle} \langle S_i \rangle \langle S_j \rangle + \Delta \sum_i \left\langle S_i^2 \right\rangle - \Delta N \end{aligned} \tag{8.107}$$

where

$$\langle S_i \rangle = \langle S_j \rangle \equiv m \tag{8.108}$$

Hence,

$$G(T, J, \Delta)_{MF} = -\frac{1}{2} N J z (\mathrm{Tr}_{S_i}(\rho_i S_i))^2 + N\Delta \, \mathrm{Tr}_{S_i}(\rho_i S_i^2) - N\Delta + N k_B T \, \mathrm{Tr}_{S_i}(\rho_i \ln \rho_i) \tag{8.109}$$

We noe minimize $G(T, J, \Delta)_{MF}$ with respect to the function ρ_i with constraint $\mathrm{Tr}_{S_i}(\rho_i) = 1$:

$$\frac{\mathrm{d}G}{\mathrm{d}\rho_i} = 0 \tag{8.110}$$

Let us consider each term

$$\frac{\mathrm{d}}{\mathrm{d}p_i} (\mathrm{Tr}(\rho_i S_i))^2 = 2(\mathrm{Tr}(\rho_i S_i)) S_i = 2 \langle S_i \rangle S_i = 2 m S_i \tag{8.111a}$$

$$\frac{\mathrm{d}}{\mathrm{d}p_i} \left(\mathrm{Tr}\left(\rho_i S_i^2\right)\right) = S_i^2 \tag{8.111b}$$

$$\frac{\mathrm{d}}{\mathrm{d}p_i} (\mathrm{Tr}(\rho_i \ln \rho_i)) = \ln \rho_i + 1 \tag{8.111c}$$

Remark. Remind that $\rho_i = \rho^{(1)}(S_i)$.

$$0 = -JNzmS_i + \Delta N S_i^2 + N k_B T \ln \rho_i + N k_B T \tag{8.112}$$

Dividing by $N k_B T$,

$$\ln \rho_i \equiv \ln \rho^{(1)}(S_i) = \beta J z m S_i - \beta \Delta S_i^2 - 1 \tag{8.113}$$

which implies

$$\rho^{(1)}(S_i) = A^{-1} e^{\beta(zJmS_i - \Delta S_i^2)} \tag{8.114}$$

Remark. In A^{-1} it is included the term e^{-1}.

The constant A can be found by imposing $\mathrm{Tr}_{S_i} \rho^{(1)}(S_i) = 1$, (TO DO)

$$A = 1 + 2e^{-\beta\Delta} \cosh(\beta z J m) \tag{8.115}$$

Given $\rho^{(1)}(S_i)$ it is possible to show (TO DO)

$$\left\langle S_i^2 \right\rangle = \mathrm{Tr}_{S_i}(\rho_i S_i^2) = \frac{1}{A} 2e^{-\beta\Delta} \cosh(\beta z J m) \tag{8.116}$$

and

$$x = 1 - \left\langle S_i^2 \right\rangle = \frac{A - 2e^{-\beta\Delta} \cosh(\beta z J m)}{A} \quad \Rightarrow x = \frac{1}{A} \tag{8.117}$$

Hence,

$$\frac{G(T, \Delta, m, J)}{N} = \frac{z}{2} J m^2 - \Delta - k_B T \ln A \tag{8.118}$$

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

$$\cosh(t) = 1 + \frac{t^2}{2} + \frac{t^4}{24}, \quad \ln(1+t) = t - \frac{t^2}{2} \tag{8.119}$$

(TO DO)

$$G(T, \Delta, J, m) = a_0(T, \Delta) + a(T, \Delta)m^2 + b(T, \Delta)m^4 + \frac{c(T, \Delta)}{6}m^6 \tag{8.120}$$

where

$$\begin{cases} a(T, \Delta) = \frac{zJ}{2}\left(1 - \frac{zJ}{\delta k_B T}\right) \\ b(T, \Delta) = k\left(1 - \frac{\delta}{3}\right) \\ c(T, \Delta) > 0 \end{cases} \tag{8.121}$$

and the parameter

$$\delta \equiv 1 + \frac{e^{\beta\Delta}}{2} = \delta(T, \Delta) \tag{8.122}$$

is related to the concentration of $\mathrm{He}^3 - x$. This can be seen as follows.

Since

$$x(T, \Delta, J) \equiv 1 - \langle S_i^2 \rangle = \frac{1}{A} = \left(1 + 2e^{-\beta\Delta}\cosh(\beta z J m)\right)^{-1} \tag{8.123}$$

In the disordered phase ($m = 0$) one has

$$x(T, \Delta, J) = \left(1 + 2e^{-\beta\Delta}\right)^{-1} = \frac{\delta - 1}{\delta} \tag{8.124}$$

By combining this result with the order-disorder transition

$$a(T_c(\Delta)) = \frac{zJ}{2}\left(1 - \frac{zJ}{\delta k_B T_c}\right) = 0, \quad \left[\delta = \frac{zJ}{k_B T_c}\right] \tag{8.125}$$

one obtains (TO DO)

$$T_c(x) = T_c(0)(1 - x) \tag{8.126}$$

There is a dependence of the critical temperature λ on the He^3 concentration x.

The tricritical point is the one that satisfies the conditions

$$\begin{cases} a(T_t, \Delta_t) = 0 \\ b(T_t, \Delta_t) = 0 \end{cases} \Rightarrow \begin{cases} \delta_t = \frac{zJ}{k_B T_t} \\ \delta_t = 3 \end{cases} \tag{8.127}$$

$$x(T_t, \Delta_t) = \frac{\delta_t - 1}{\delta_t} = \frac{2}{3} \tag{8.128}$$

Remark. Experimental estimate of x_t is ~ 0.67.

Exercise 5. Expand the free-energy per site

$$\frac{G}{N} = \frac{z}{2}Jm^2 - \Delta - k_B T \ln A \tag{8.129}$$

where $A = 1 + 2e^{-\beta\Delta}\cosh(\beta z J m)$ for small values of m.

$$x \equiv \beta z J m, \quad B \equiv 2e^{-\beta\Delta} \tag{8.130}$$

Since $\cosh x \simeq 1 + \frac{x^2}{2} + \frac{x^4}{24}$,

$$A = 1 + B\cosh x \simeq 1 + B\left(1 + \frac{x^2}{2} + \frac{x^4}{24}\right) \tag{8.131}$$

$$\begin{aligned} \ln A &= \ln\left(1 + B + \frac{Bx^2}{2} + \frac{Bx^4}{24}\right) \\ &\simeq \ln\left[(1+B)\left(1 + \frac{B}{2(1+B)}x^2 + \frac{B}{24(1+B)}x^4\right)\right] \\ &= \ln(1+B) + \ln(1+t) \end{aligned} \tag{8.132}$$

where

$$t \equiv \frac{B}{2(1+B)}x^2 + \frac{B}{24(1+B)}x^4 \tag{8.133}$$

Let us first consider the term

$$\frac{B}{1+B} = \frac{2e^{-\beta\Delta}}{1+2e^{-\beta\Delta}} = \frac{2}{2+e^{\beta\Delta}} = \delta^{-1} \tag{8.134}$$

$$\Rightarrow \ln A = \ln(1+B) + \frac{x^2}{2\delta} + \left(\frac{1}{24\delta} - \frac{1}{8\delta^2}\right)x^4 - \frac{1}{24\delta^2}x^6 \tag{8.135}$$

Remark. We have that $x \equiv \beta z J m$.

$$\begin{aligned} -\frac{\ln A}{B} + \frac{z}{2}Jm^2 - \Delta \simeq & a_0(T,\Delta) + \left(\frac{z}{2}J - \frac{\beta z^2 J^2}{2\delta}\right)m^2 \\ & + \left(\frac{1}{8\delta} - \frac{1}{24\delta}\right)\beta^3 z^4 J^4 m^4 + \frac{1}{24\delta^2}\beta^5 z^6 J^6 m^6 \end{aligned} \tag{8.136}$$

$$G(T,\Delta,J,m) = a_0(T,\Delta) + a(T,\Delta)m^2 + b(T,\Delta)m^4 + c(T,\Delta)m^6 \tag{8.137}$$

where

$$a(T,\Delta) = \frac{zJ}{2}\left(1 - \frac{\beta z J}{\delta}\right) \tag{8.138a}$$

$$b(T,\Delta) = \frac{\beta^3 z^4 J^4}{8\delta}\left(\frac{1}{\delta} - \frac{1}{3}\right) = \frac{\beta^3 z^4 J^4}{8\delta^2}\left(1 - \frac{\delta}{3}\right) \tag{8.138b}$$

$$c(T,\Delta) = \frac{\beta^5 z^6 J^6}{24\delta^2} > 0 \tag{8.138c}$$

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

$$F_{var} = \langle \mathcal{H} \rangle_{\rho_{TR}} + k_B T \langle \ln \rho_{TR} \rangle_{\rho_{TR}} \tag{8.139}$$

We have to choose a family of distribution. If one assumes that the family of trial distribution is of the Gibbs-Boltzmann form

$$\rho_{TR} = \frac{e^{-\beta\mathcal{H}_{TR}}}{Z_{TR}} \tag{8.140}$$

with

$$Z_{TR} = e^{-\beta F_{TR}} = \sum_{\{\Phi_i\}} e^{-\beta\mathcal{H}_{TR}(\{\Phi_i\})} \tag{8.141}$$

If we are able to choose hamiltonian trial as here, we can do this sum, this is the idea. In principle, if we are able to find an Hamiltonian that we can solve without choosing a mean filed it is ok.

Then, since

$$\ln \rho_{TR} = -\beta \mathcal{H}_{TR} - \ln Z_{TR} \tag{8.142}$$

we have

$$k_B T \left\langle \ln \rho_{TR} \right\rangle_{\rho_{TR}} = k_B T \left\langle \frac{-\mathcal{H}_{TR}}{k_B T} \right\rangle + k_B T \underbrace{\left\langle -\ln Z_{TR} \right\rangle}_{\beta F_{TR}} \tag{8.143}$$

$$k_B T \left\langle \ln \rho_{TR} \right\rangle_{\rho_{TR}} = \left\langle -\mathcal{H}_{TR} \right\rangle + F_{TR} \tag{8.144}$$

$$\Rightarrow F_{var} = \left\langle \mathcal{H} \right\rangle_{\rho_{TR}} - \left\langle \mathcal{H}_{TR} \right\rangle_{\rho_{TR}} + F_{TR} = \left\langle \mathcal{H} - \mathcal{H}_{TR} \right\rangle_{\rho_{TR}} + F_{TR} \tag{8.145}$$

Clearly, $F \leq F_{var}$ and one has to look for the minima of F_{var} by varying ρ_{TR}. Within this approach, the mean field approximation is still given by

$$\rho_{TR}^{MF}(\Phi_1, \ldots, \Phi_N) = \prod_{i=1}^{N} \rho_{TR}^{(1)}(\Phi_i) \tag{8.146}$$

that in this case becomes

$$\prod_i \rho_{TR}^{(1)}(\Phi_i) = \frac{1}{Z_{TR}^{MF}} e^{-\beta \sum_i b_i \Phi_i} \tag{8.147}$$

and

$$Z_{TR} = \sum_{\{\Phi\}} e^{-\beta \sum_i b_i \Phi_i} \tag{8.148}$$

where b_i are the variational parameters.

$$\mathcal{H}_{TR} = -\sum_i b_i \Phi_i \tag{8.149}$$

If we consider again the Ising model

$$\mathcal{H} = -J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i \tag{8.150}$$

$$\begin{aligned} F_{var} &= \left\langle \mathcal{H} - \mathcal{H}_{TR} \right\rangle_{\rho_{TR}} + F_{TR} \\ &= F_{TR} + \left\langle \left(-J \sum_{\langle ij \rangle} S_i S_j - H \sum_i S_i \right) - \left(-\sum_i b_i S_i \right) \right\rangle_{\rho_{TR}} \\ &= F_{TR} + \left\langle -J \sum_{\langle ij \rangle} S_i S_j + \sum_i (b_i - H) S_i \right\rangle_{\rho_{TR}} \\ &= F_{TR} - J \sum_{\langle ij \rangle} \left\langle S_i S_j \right\rangle_{\rho_{TR}} + \sum_i (b_i - H) \left\langle S_i \right\rangle_{\rho_{TR}} \end{aligned} \tag{8.151}$$

Since $\rho_{TR} = \prod_i \rho_i$,

$$\left\langle S_i S_j \right\rangle_{\rho_{TR}} = \left\langle S_i \right\rangle_{\rho_{TR}} \left\langle S_j \right\rangle_{\rho_{TR}} \tag{8.152}$$

$$F_{var} = F_{TR} - J \sum_{\langle ij \rangle} \left\langle S_i \right\rangle_{\rho_{TR}} \left\langle S_j \right\rangle_{\rho_{TR}} + \sum_i (b_i - H) \left\langle S_i \right\rangle_{\rho_{TR}} \tag{8.153}$$

$$\frac{\partial F_{var}}{\partial b_i} = 0, \quad \forall i \tag{8.154}$$

$$0 = \frac{\partial F_{var}}{\partial b_i} = \left[-J \sum_{j\, n.n.\, i} \langle S_i \rangle_{\rho_{TR}} + b_i - H \right] \frac{\partial \langle S_i \rangle}{\partial b_i} \tag{8.155}$$

or

$$b_i = J \sum_{j\, n.n.\, i} \langle S_j \rangle_{\rho_{TR}} + H \tag{8.156}$$

$$\begin{aligned} \langle S_i \rangle_{\rho_{TR}} &= \frac{1}{Z_{TR}} \sum_{\{S\}} S_i e^{\beta \sum_k S_k b_k} = \frac{\prod_k \sum_{S_k} S_i e^{\beta S_k b_k}}{\prod_k \sum_{S_k} e^{\beta S_k b_k}} \\ &= \frac{\sum_{S_i=\pm 1} S_i e^{\beta S_i b_i}}{\sum_{S_i=\pm 1} e^{\beta S_i b_i}} = \frac{\cosh(\beta b_i)}{\sinh(\beta b_i)} \\ &= \tanh(\beta b_i) \end{aligned} \tag{8.157}$$

$$b_i = J \sum_{j\, n.n.\, i} \tanh(\beta b_j) + H \tag{8.158}$$

Remark. The main step to understand is how to derive F_{var} from a ρ_{TR}. This is nice to see a variation with respect to the real hamiltonian. Consider a bunch of data, for instance a million of configuration, which is the distribution of the configuration? Usually, we build up a model with a distribution that depends on parameters and what we want to do is statistical inference. Starting from the model and the data we have to obtain the real distribution. (lesson)

Exercise 6. Consider again the antiferromagnetic Ising model

$$\mathcal{H}[\{S\}] = -J \sum_{\langle \vec{\mathbf{r}}_A \vec{\mathbf{r}}_B \rangle} S(\vec{\mathbf{r}}_A) S(\vec{\mathbf{r}}_B) - H \sum_{\vec{\mathbf{r}}_A} S(\vec{\mathbf{r}}_A) + H \sum_{\vec{\mathbf{r}}_B} S(\vec{\mathbf{r}}_B) \tag{8.159}$$

where $J > 0$ and $H > 0$.

- $\vec{\mathbf{r}}_A$ denotes the site on the A sublattice.
- $\vec{\mathbf{r}}_B$ denotes the site on the B sublattice.

Let us find again the mean-field solution but now using the variational ansatz

$$F \leq F_{var} = \langle \mathcal{H} \rangle_{\rho_{TR}} - \langle \mathcal{H}_{TR} \rangle_{\rho_{TR}} + F_{TR} = \langle \mathcal{H} - \mathcal{H}_{TR} \rangle_{\rho_{TR}} + F_{TR} \tag{8.160}$$

Remark. Since the problem can be splitted in two sublattices, it is convenient to use

$$\mathcal{H}_{TR} = -H_A \sum_{r_A} S(r_A) - H_B \sum_{r_B} S(r_B) \tag{8.161}$$

- Show that F_{var} has the following expression:

$$\begin{aligned} F_{var} =& F_{TR}(\beta H_A, \beta H_B) - 4NJ \langle S_A \rangle_{\rho_{TR}} \langle S_B \rangle_{\rho_{TR}} \\ &- \frac{1}{2} NH \Big(\langle S_A \rangle_{\rho_{TR}} - \langle S_B \rangle_{\rho_{TR}} \Big) + \frac{1}{2} N \Big(H_A \langle S_A \rangle_{\rho_{TR}} + \langle S_B \rangle_{\rho_{TR}} \Big) \end{aligned} \tag{8.162}$$

where

$$\langle S_A \rangle_{\rho_{TR}} \equiv m_A + n \tag{8.163a}$$

$$\langle S_B \rangle_{\rho_{TR}} \equiv m_B - n \tag{8.163b}$$

with $m = m_A + m_B$, and

$$m_A = \tanh(\beta H - 4\beta J m_B) \tag{8.164a}$$

$$m_B = \tanh(\beta H - 4\beta J m_A) \tag{8.164b}$$

- Expand the free energy F_{var} in powrs of m of the form

$$F_{var} = A + Bm^2 + cm^4 + O(m^6) \tag{8.165}$$

and find the explicit expression of A, B and C as a function of T, H and n.

Chapter 9

Non ideal fluids: Mean field theory, Van der Walls, Virial expansion and Cluster expansion

Lecture 14.
Friday 29th November, 2019.
Compiled: Thursday 26th December, 2019.

9.1 Mean field theory for fluids

Fluid system of N particles with position vectors $\{\vec{\mathbf{r}}_i\}_{i=1,\dots,N}$. The configurational (grancanonical) partition function is (we can do Gaussian integrals)

$$Q_N(T) = \int_V \prod_{i=1}^{N} \mathrm{d}\vec{\mathbf{r}}_i \exp\left[-\beta\Phi(\{\vec{\mathbf{r}}\}) - \beta\sum_{i=1}^{N}\psi_{ext}(\vec{\mathbf{r}}_i)\right] \tag{9.1}$$

where ψ_{ext} is a one body potential, but we do not consider it because is not the aim of our problem.

In general,

$$\Phi(\{\vec{\mathbf{r}}_i\}) = \sum_{i,j>i} U_2(\vec{\mathbf{r}}_i, \vec{\mathbf{r}}_j) + \sum_{i,j,\mu} U_3(\vec{\mathbf{r}}_i, \vec{\mathbf{r}}_j, \vec{\mathbf{r}}_\mu) + \dots \tag{9.2}$$

but we forgot about U_3 that is the three body interaction.

We consider

$$U_2(\vec{\mathbf{r}}_i, \vec{\mathbf{r}}_j) \to U_2(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|) \tag{9.3}$$

Therefore,

$$Q_N(T) = \int_V \prod_{i=1}^{N} \mathrm{d}\vec{\mathbf{r}} \exp\left[-\beta\sum_{i,j>i} U_2(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|)\right] \tag{9.4}$$

Now, we replace all this story with just a field, it is a sort of average of the interactions. Doing the mean field assumption for U_2, we obtain

$$\sum_{i,j>1} U_2(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|) \to \sum_i \Phi_{MF}(\vec{\mathbf{r}}_i) \tag{9.5}$$

In particular, the **mean field approximation** consists in substituting the multi-body interaction potential $\Phi(\{\vec{\mathbf{r}}_i\})$ with an effective one body potential $\Phi(\vec{\mathbf{r}})$ withing which all the particles move.

$$\Phi(\{\vec{\mathbf{r}}_i\}) = \sum_i \Phi_{MF}(\vec{\mathbf{r}}_i) \tag{9.6}$$

As said, for simplicity consider $\psi_{ext} = 0$, hence

$$Q_N(T) \simeq \left[\int_V \prod_{i=1}^{N} \mathrm{d}\vec{\mathbf{r}} \exp[-\beta \Phi_{MF}(\vec{\mathbf{r}})] \right]^N \tag{9.7}$$

Remark. The integral depends on the form of $\Phi_{MF}(\vec{\mathbf{r}})$.

If one assumes the *spatial isotropy*, what it is important is not anymore the vector but only the distance; hence, it is important just the integral over the modulus:

$$\Phi_{MF}(\vec{\mathbf{r}}) = \Phi_{MF}(|\vec{\mathbf{r}}|) = \Phi_{MF}(r) \tag{9.8}$$

In general

$$\Phi_{MF}(r) = \begin{cases} \infty & r < r_0 \quad \text{repulsion} \\ u < 0 & r > r_0 \quad \text{attraction} \end{cases} \tag{9.9}$$

as plotted in Figure 9.1. Moreover,

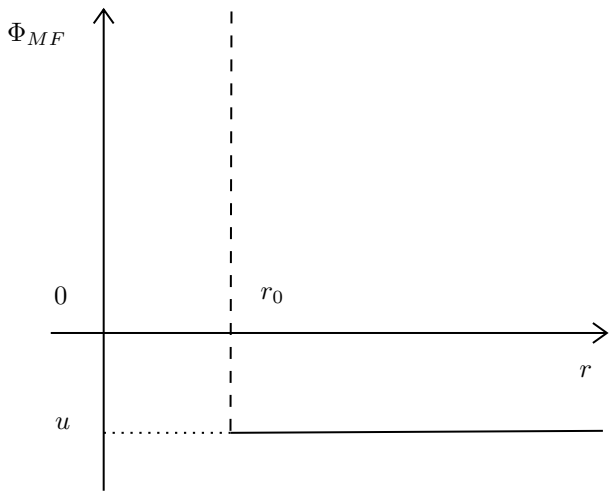


Figure 9.1: Plot of the potential $\Phi_{MF}(r)$.

$$Q_N^{MF}(T) = \left[V_{ex} e^{-\infty} + (V - V_{ex}) e^{-\beta u} \right]^N \tag{9.10}$$

where $V_{ex} \simeq r_0^3$ is the volume not accessible by the particle.

$$Q_N^{MF}(T) = \left[(V - V_{ex}) e^{-\beta u} \right]^N \tag{9.11}$$

The free energy is

$$F_N^{MF}(T) = -N k_B T [\ln (V - V_{ex}) - \beta u] \tag{9.12}$$

The pressure is

$$P_N^{MF} = -\left. \frac{\partial F_N^{MF}}{\partial V} \right|_T = \frac{N k_B T}{V - V_{ex}} - N \left(\frac{\partial u}{\partial V} \right)_T \tag{9.13}$$

Remark. In general, the deep u can go up and down depending on the V: $u = u(V)$. This is becous u is the attractive well of the mean field potential and, for $r \geq r_0$ must be proportional to the fluid density

$$u \sim -N/V$$

where the minus sign means attraction. On the other hand, also V_{ex}, the volume not accessible, must be proportional to N,

$$V_{ex} = bN \quad \Rightarrow u = -a \frac{N}{V}$$

where b is the volume of a single particle.

Inserting the last term in (9.13), we obtain the *Van der Walls equation of state*:

$$P_N^{MF}(V,T) = \frac{Nk_BT}{V-bN} - a\left(\frac{N}{V}\right)^2 \tag{9.14}$$

9.2 Equation of state and critical point

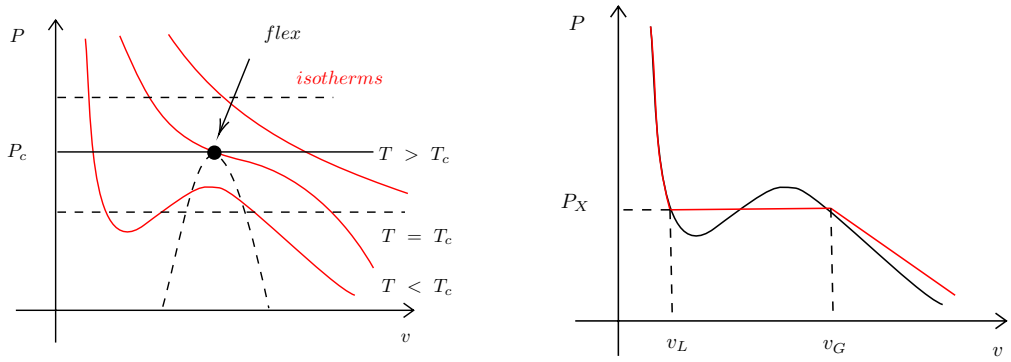
Let us define the specific volume

$$v \equiv \frac{1}{\rho} = \frac{V}{N} \tag{9.15}$$

Hence,

$$P = \frac{K_BT}{v-b} - \frac{a}{v^2} \tag{9.16}$$

We have this effect because it is a mean field, so the curve in Figure 9.2a it is replaced by the curve in Figure 9.2b.



(a) Description

(b) Description

Figure 9.2

For $T < T_c$ the equation $P(v) = const$ has 3 distinct solutions. For $T > T_c$ only one solution $\in \mathbb{R}$.

- The *inflection point condition* at $T = T_c$ is

$$\frac{\partial P}{\partial v} = \frac{\partial^2 P}{\partial v^2} = 0 \tag{9.17}$$

  and these are the critical point conditions. The second in particular means that there is a flex point. Let us pay attention to it, indeed it is a standard way to find critical points. We obtain $v_c = 3b$:

$$P_c = \frac{a}{27b^2}, \quad k_BT_c = \frac{8a}{27b} \tag{9.18}$$

  and

$$\frac{P_cv_c}{k_BT_c} = \frac{3}{8} \approx 0.375 \tag{9.19}$$

- Another way to find the critical point consists in noticing that at $T = T_c$, the 3 solutions coincide. Let us rewrite the equation of state

$$P = \frac{v^2k_BT - a(v-b)}{v^2(v-b)} \tag{9.20}$$

as

$$v^3 - \left(b + \frac{k_B T}{P}\right) v^2 + \frac{a}{P} v - \frac{ab}{P} = 0 \tag{9.21}$$

Since at $T = T_c$ the 3 solutions coincide, equation (9.21) must be of the form

$$(v - v_c)^3 = 0 \quad \Rightarrow v^3 - 3v^2 v_c + 3v v_c^2 - v_c^3 = 0 \tag{9.22}$$

and, by identifying its coefficients with the ones in (9.21) one gets

$$v_c^3 = \frac{ab}{P_c}, \quad 3v_c^2 = \frac{a}{P_c}, \quad 3v_c = b + \frac{k_B T_c}{P_c} \tag{9.23}$$

Giving

$$v_c = 3b, \quad P_c = \frac{a}{27b^2}, \quad k_B T_c = \frac{8a}{27b} \tag{9.24}$$

From the relations (9.24), it is clear that it is sufficient to estimate the a and b coefficients of a gas using the equation of state at T sufficiently high to estimate its critical values, v_c, T_c and P_c. Moreover, one can show that

$$\frac{P_c v_c}{k_B T_c} = \frac{3}{8} \approx 0.375 \tag{9.25}$$

It is an universal ratio whose value does not depends on either a or b.

9.2.1 Law of the correspondent states

The universal value of the ratio $\frac{P_c v_c}{k_B T_c}$ suggests a deeper correspondence between different fluid systems.

Let us rewrite the Van der Walls equation of state (9.14) in adimensional form

$$\pi \equiv \frac{P}{P_c}, \quad \nu \equiv \frac{v}{v_c}, \quad \tau \equiv \frac{T}{T_c} \tag{9.26}$$

The result is

$$\left(\pi + \frac{3}{\nu^2}\right)(3\nu - 1) = 8\tau \tag{9.27}$$

Remark. When rescaled by P_c, v_c and T_c, the thermodynamic variables P, v and T of the Van der Walls fluids, follow the same equation of state! The Van der Walls theory describes the law of correspondent states found experimentally.

9.3 Region of coexistence and Maxwell construction

In real fluids, for $T < T_c$ ($\tau < 1$), there is a first order liquid-gas transition with coexistence between vapor and liquid phase and non analiticity of the thermodynamic potential. In particular, an isotherm for $T < T_c$ is the one in Figure 9.2b. How this is described by the mean-field (i.e. Van der Walls) theory?

The Van der Walls isotherm is given by the graphic in Figure 9.3. The liquid phase goes into a phase region that is not thermodinamycally stable.

How can we remove the non physical regions of the Van der Walls equation of state and describe coexistence? The solution is the Maxwell (or equal area) construction.

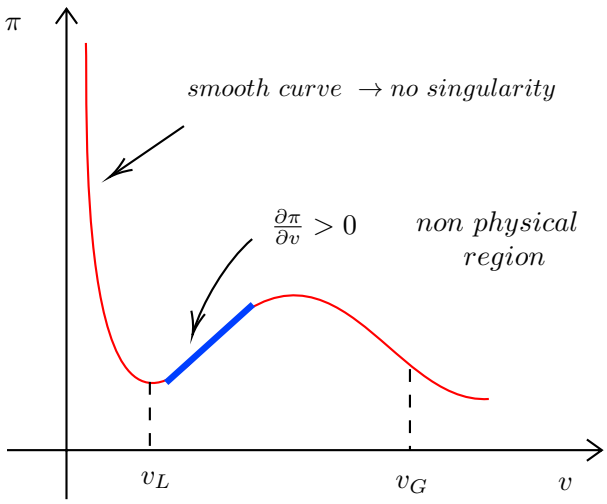


Figure 9.3: Description.

9.3.1 Equal area or Maxwell construction

Remark. The coexistence regions are characterised by having the same pressure (mechanical equilibrium) and same temperature (thermal equilibrium).

Since

$$\mu \equiv \frac{G}{N} = \frac{F + PV}{N} \tag{9.28}$$

we have

$$\mathrm{d}\mu = -\frac{S}{N}\,\mathrm{d}T + \frac{V}{N}\,\mathrm{d}P \tag{9.29}$$

Along an isotherm ($T = const$), $\mathrm{d}T = 0$, which implies

$$\mathrm{d}\mu = \frac{V}{N}\,\mathrm{d}P \tag{9.30}$$

At the coexistence $\mathrm{d}P_{coex} = 0$, hence

$$\mathrm{d}\mu = 0 \tag{9.31}$$

is the physical condition.

Now, let us consider a Van der Walls isotherm at $T < T_c$, as in Figure 9.4.

Remark. For Van der Wall $\mathrm{d}P \neq 0$!

The physical coexistence condition $\mathrm{d}\mu_{phy}$ implies

$$0 = \int_1^2 \mathrm{d}\mu = \mu(2) - \mu(1) \overset{Van\,der\,Walls}{=} \frac{1}{N}\int_{P_G}^{P_L} \mathrm{d}P\,V \tag{9.32}$$

The integral can be partitioned in two parts

$$0 = \int_{P_G}^{P_L} \mathrm{d}P\,V \quad \Rightarrow \int_{P_G}^{P_x} \mathrm{d}P\,V = -\int_{P_x}^{P_L} \mathrm{d}P\,V \tag{9.33}$$

Remark. The equal area condition gives the value of P_x of the coexistence line!

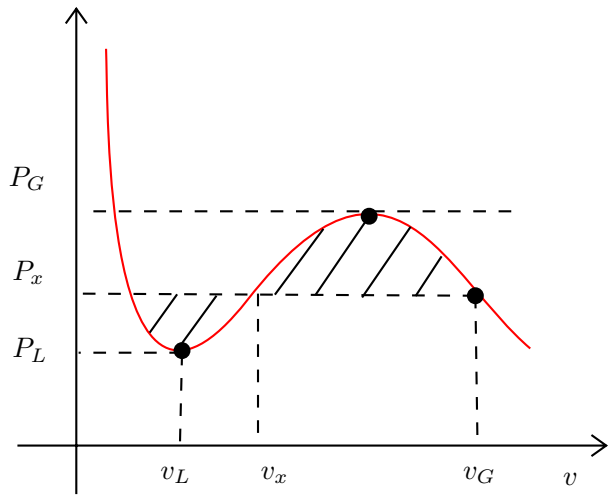


Figure 9.4: Description.

9.4 Critical exponents of Van der Walls

β exponent

The equation of state is

$$\left(\pi + \frac{3}{\nu^2}\right)(3\nu - 1) = 8\tau \tag{9.34}$$

where

$$\pi = \frac{P}{P_c}, \quad \nu = \frac{v}{v_c}, \quad \tau = \frac{T}{T_c} \tag{9.35}$$

Let us consider

$$\begin{cases} t \equiv \tau - 1 = \frac{T - T_c}{T_c} \\ \Phi = \nu - 1 = \frac{v - v_c}{v_c} \end{cases} \tag{9.36}$$

we want the deviation from the critical point. Close to the critical point $\tau \sim \nu \sim 1$ and $t \sim \Phi \sim 0$.

We now expand the equation of state with respect to t and Φ

$$\left(\pi + \frac{3}{(1+\Phi)^2}\right)(3(\Phi - 1) - 1) = 8(t+1) \tag{9.37}$$

$$\Rightarrow \pi = \frac{8(t+1)}{2(\Phi+1) - 1} - \frac{3}{(1+\Phi)^2} \tag{9.38}$$

Expanding for $\Phi \sim 0$,

$$(1+\Phi)^\alpha \simeq 1 + \Phi \tag{9.39}$$

To do:

$$\pi \simeq 1 + 4t - 6t\Phi - \frac{3}{2}\Phi^3 + O(t\Phi^2, \Phi^4) \tag{9.40}$$

Remark. Since we will get $\Phi \sim t^{1/2}$ the truncation of the expansion is justified a posteriori.

In order to get the values of $v_G(P)$ and $v_L(P)$ at coexistence, we use the Maxweel construction

$$\int_{P_G}^{P_L} v \, \mathrm{d}P = 0 \tag{9.41}$$

$$\int_{v_L}^{v_G} v \, \mathrm{d}P = 0 \tag{9.42}$$

Rewrite it with respect to the rescaled variables $\pi = \frac{P}{P_c}$ and $\nu = \frac{v}{v_c}$, if $\Phi = \nu - 1$ we have

$$\mathrm{d}P = P_c \, \mathrm{d}\pi \tag{9.43}$$

Let us consider $T < T_c$ fixed, if and only if $t < 0$ (but small)

$$\pi = \pi(v) = \pi(\Phi) \tag{9.44}$$

$$\pi \simeq 1 + 4t - 6t\Phi - \frac{3}{2}\Phi^3 + O(t\Phi^2, \Phi^4) \tag{9.45}$$

$$\Rightarrow \mathrm{d}\pi \simeq -6t \, \mathrm{d}\Phi - \frac{q}{2}\Phi^2 \, \mathrm{d}\Phi \tag{9.46}$$

$$\Rightarrow \mathrm{d}P = P_c\Big[-6t \, \mathrm{d}\Phi - \frac{q}{2}\Phi^2 \, \mathrm{d}\Phi\Big] \tag{9.47}$$

The equation (9.42) implies

$$\int_{\Phi_c}^{\Phi_g} \Phi\Big(-6t - \frac{q}{2}\Phi^2\Big) \, \mathrm{d}\Phi = 0 \tag{9.48}$$

It must be true for all $t \ll 1$.

Remark.

$$\int_c^g 1 \, \mathrm{d}P = 0 \tag{9.49}$$

since we are at coexistence!

$$\Rightarrow -3\Phi_g^2\left[t + \frac{\Phi_g^2}{g}\right] + 3\Phi_c^2\left[t + \frac{\Phi_c^2}{g}\right] = 0 \tag{9.50}$$

if and only if $\Phi_g^2 = \Phi_c^2$, hence if and only if

$$\Phi_g = \pm\Phi_c \tag{9.51}$$

On the other hand,

$$\Phi_g = \frac{v_g - v_c}{v_c}, \quad \Phi_l = \frac{v_l - v_c}{v_c} \tag{9.52}$$

$$\Rightarrow \Phi_g = -\Phi_l \tag{9.53}$$

If we now insert Φ_g and Φ_l in the expression of π, we get

$$\Phi_g \to \pi = 1 + 4t - 6t\Phi_g - \frac{3}{2}\Phi_g^3 \tag{9.54a}$$

$$\Phi_l \to \pi = 1 + 4t + 6t\Phi_g + \frac{3}{2}\Phi_g^3 \tag{9.54b}$$

The two expression of π must be equal since we are at coexistence. Solving with respect to Φ_g we get

$$\Phi_g = 2\sqrt{-t} \sim \left(\frac{T_c - T}{T_c}\right)^{1/2} \tag{9.55}$$

which implies

$$\beta = \frac{1}{2} \tag{9.56}$$

9.5 Theory of weakly interacting fluids

If the gas is not ideal but made by weakly interacting particles, it is possible to follow a *perturbative approach.* Let us consider N particles in region Ω of volume V. Particles interact by a two-body interacting potential

$$U_2(\vec{\mathbf{r}}_i, \vec{\mathbf{r}}_j) = \Phi(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|) \tag{9.57}$$

hence,

$$\Rightarrow U(\{\vec{\mathbf{r}}\}) = \frac{1}{2}\sum_{i,j} \Phi(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|) \tag{9.58}$$

The Hamiltonian we are considering is

$$\mathcal{H}_\Omega(\{\vec{\mathbf{r}}\}) = \sum_{i=1}^{N} \frac{\vec{\mathbf{p}}_i^2}{2m} + \sum_{i,j>i} \Phi(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|) \tag{9.59}$$

In the canonical ensemble

$$Z_\Omega(N, V, T) = \frac{1}{N!\Lambda^{3N}} Q_N(V, T) \tag{9.60}$$

where

$$Q_N(V, T) = \int_V \mathrm{d}\vec{\mathbf{r}}_1 \int_V \mathrm{d}\vec{\mathbf{r}}_2 \cdots \int_V \mathrm{d}\vec{\mathbf{r}}_N \exp[-\beta U(\{\vec{\mathbf{r}}\})] \tag{9.61}$$

Remark. For ideal gases $U = 0$, so

$$Q_N(V, T) = V^N \quad \to Z_N^{ideal} = \frac{V^N}{N!\Lambda^{3N}} \tag{9.62}$$

and the dependence on T is exclusively due to $\Lambda = \Lambda(T)$ (i.e. kinetic energy).

Now, suppose $U \neq 0$, but small! We can say that we can assume that our $Q_N(V, T)$ it would be the ideal version times a new function

$$Q_N(V, T) \simeq V^N \chi(N, V, T) \tag{9.63}$$

Remark. If Φ is only repulsive, χ does not depend on T.

In conclusion, we obtain

$$F_N = F_N^{ideal} - k_B T \ln \chi_N \tag{9.64}$$

remaThe correction χ due to particle-particle interaction depends on the particle density ρ.

$$\begin{cases} \rho_{small} \Rightarrow U = 0 \\ \rho_{high} \Rightarrow U \neq 0 \text{ and not negligible} \end{cases} \tag{9.65}$$

This suggests that the equation of state of a weakly interacting gas can be expanded formally in powers of ρ. This is known as *virial expansion.*

For the ideal gas:

$$\frac{P}{k_B T} = \rho \tag{9.66}$$

This is the ideal gas, now start to add the other terms of the expansion

$$\to \frac{P}{k_B T} = \rho + B_2(T)\rho^2 + B_3(T)\rho^3 + \cdots + O(\rho^n) \tag{9.67}$$

this is a *virial expansion* and it is one of the most used. The coefficient B are called the *virial coefficients.* The equation (9.67) was first introduced as a formula to fit

experimental data. Indeed, making a fit, you will obtain the virial coefficients. This is what physicist have done for years. Then, mapping the coefficient with the real world experiments, we can find some macroscopical parameters.

The formula (9.67) can be also obtained rigorously from a perturbation approach to the partition function (see later).

Which is the virial expansion of a Van der Walls (i.e. mean field) gas?

9.6 Van der Walls and virial expansion

Consider the Van der Walls equation of state

$$\frac{P}{k_B T} = \frac{N}{V - bN} - \frac{aN^2}{k_B T V^2} \tag{9.68}$$

Let us factorize N/V,

$$\frac{P}{k_B T} = \left(\frac{N}{V}\right)\left(1 - b\frac{N}{V}\right)^{-1} - \frac{a}{k_B T}\left(\frac{N}{V}\right)^2 \tag{9.69}$$

Expanding in power of (N/V),

$$\begin{aligned} \Rightarrow \frac{P}{k_B T} &= \left(\frac{N}{V}\right) + \left(\frac{N}{V}\right)^2\left(b - \frac{a}{k_B T}\right) + \left(\frac{N}{V}\right)^3 b^2 + \left(\frac{N}{V}\right)^4 b^3 + \dots \\ &= \rho + \left(b - \frac{a}{k_B T}\right)\rho^2 + b^2\rho^3 + b^3\rho^4 + \dots \end{aligned} \tag{9.70}$$

Therefore,

$$B_2(T)^{VdW} = b - \frac{a}{k_B T} \qquad B_3^{VdW} = b^2 \tag{9.71}$$

where in $B_2(T)^{VdW}$ the first term is repulsive on excluded volume and the second one is the attraction term. In B_3^{VdW} we note that the term is always positive.

9.6.1 Boyle's temperature T_B

The Boyle's temperature is the T at which the second coefficient is zero:

$$B_2^{VdW}(T_B) = 0 \tag{9.72}$$

so we have removed the most important coefficient. The competiting effects of repulsion and attraction are cancelled out.

We have the Van der Walls temperature T_B^{VdW}

$$T_B^{VdW} = \frac{a}{bk_B} \tag{9.73}$$

to be compared with T_c^{VdW} that is

$$T_c^{VdW} = \frac{8a}{27b^3} \tag{9.74}$$

We notice that $T_c^{VdW} \ll T_B$. It is clear that the Boyle's temperature must be much greater than the critical one.

We now obtain the formal virial expansion by starting from the microscopic system and performing a perturbation expansion of the Boltzmann weights for small values of U.

Remark. Consider a polymer, the transition point called the θ point is when the second coefficient is zero, as the case described above, but it is interesting in polymer kind of system (lesson).

9.7 Cluster expansion technique for weakly interacting gases

Now, we try to do some calculation of this virial coefficients, starting from the model microscopical. Let us start from the partition function

$$\begin{aligned} Q_N &= \int_V \mathrm{d}\vec{\mathbf{r}}_1 \cdots \int_V \mathrm{d}\vec{\mathbf{r}}_N \, e^{-\beta U(\{\vec{\mathbf{r}}\})} \\ &= \int_V \mathrm{d}\vec{\mathbf{r}}_1 \cdots \int_V \mathrm{d}\vec{\mathbf{r}}_N \, e^{-\beta \sum_{i,j>i} \Phi_{ij}} \end{aligned} \tag{9.75}$$

where

$$\Phi_{ij} \equiv \Phi(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|) \tag{9.76}$$

Definition 7 (Mayer function)**.** The *Mayer function* is something that is smaller in that given point

$$f(|\vec{\mathbf{r}}|) \equiv \exp[-\beta\Phi(|\vec{\mathbf{r}}|)] - 1 \tag{9.77}$$

with $f_{ij} \equiv f(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|)$.

Remark. Note: if $\beta\Phi(r) \ll 1$, we have $f(r) \ll 1$.

Hence,

$$\begin{aligned} \Rightarrow e^{-\beta \sum_{i,j>i} \Phi_{ij}} &= \prod_i \left(\prod_{j>i} (1 + f_{ij}) \right) \\ &= \underbrace{(1+f_{12})(1+f_{13})\dots(1+f_{1N})}_{i=1} \dots \underbrace{(1+f_{23})(1+f_{24})\dots(1+f_{2N})}_{i=2} \dots \\ &= (1 + f_{12} + f_{13} + f_{12}f_{13})(1+f_{14})\dots(1+f_{23}) \\ &= 1 + \sum_i \sum_{k>i} f_{ij} + \cancel{\sum_i^N \sum_{k\geq i} \sum_{l>k} \sum_{j>i} \sum_{(ij)\neq(lk)} f_{ik} f_{kl}} + O(f^3) \end{aligned} \tag{9.78}$$

where

$$f_{ij} \equiv e^{-\beta\Phi_{ij}} - 1 \tag{9.79}$$

Higher order terms contain products of $3, 4, \dots f_{ij}$ terms. For simplicity, let us consider first only linear terms. Hence, the solution is given by considering only the linear term. This is the cluster expansion.

Lecture 15. *Wednesday 4th December, 2019. Compiled: Thursday 26th December, 2019.*

This first approximation is reasonable if either

1. ρ is small enough. It implies that $|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j| \gg 1$ and hence $\Phi_{ij} \ll 1$.
2. Sufficiently high T such that $\Phi(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|)/k_B T \ll 1$. What is important it is the ration between β and Φ_{ij}.

In either cases we have $\exp(-\beta\Phi_{ij}) \to 1$ and $f_{ij} \to 0$. By keeping only linear terms the partition function is

$$\begin{aligned} Q_N(V,T) &= \int_V \mathrm{d}\vec{\mathbf{r}}_1 \dots \mathrm{d}\vec{\mathbf{r}}_N \left(1 + \sum_{i,j>i} f_{ij} + \dots \right) \\ &= V^N + \sum_{i,j>i} \int_V \mathrm{d}\vec{\mathbf{r}}_1 \cdots \int_V \mathrm{d}\vec{\mathbf{r}}_N \, f_{ij} \\ &= V^N + V^{N-2} \sum_{i,j>i} \int_V \mathrm{d}\vec{\mathbf{r}}_i \, \mathrm{d}\vec{\mathbf{r}}_j \, f_{ij} + \dots \end{aligned} \tag{9.80}$$

We are summing up over all configurations ij. Let us try to compute the double integral:

$$\int_V \mathrm{d}\vec{\mathbf{r}}_i\,\mathrm{d}\vec{\mathbf{r}}_j\, f_{ij}(|\vec{\mathbf{r}}_i-\vec{\mathbf{r}}_j|) \underset{\substack{\text{translational}\\ \text{symmetry}}}{=} \int \mathrm{d}\vec{\mathbf{r}}_i\,\mathrm{d}\vec{\mathbf{r}}_j\, f(\vec{\mathbf{r}}) = V\int_V \mathrm{d}\vec{\mathbf{r}}\, f(|\vec{\mathbf{r}}|) \equiv -2B_2V \tag{9.81}$$

where

$$B_2 \equiv -\frac{1}{2}\int_V \mathrm{d}\vec{\mathbf{r}}\, f(|\vec{\mathbf{r}}|) \tag{9.82}$$

so, what is important it is the relative distance. $\vec{\mathbf{r}}$ gives us the position from the center we have choosen. Rewriting again the partition function we obtain:

$$Q_N(V,T) = V^N - V^{N-1}N(N-1)B_2(T) \tag{9.83}$$

Remark. The factor $\frac{N(N-1)}{2}$ comes out becouse in the double sum are considered all the possible connections (bonds) between pairs of particles (i,j) with $j>i$.

$$Z_N(V,T) = \left(\frac{V^N}{N!\Lambda^{3N}}\right)\left(1 - \overset{N(N-1)\approx N^2}{\frac{N^2}{V}} B_2(T) + \dots\right) \tag{9.84}$$

Remark. We do not care about the $(N-1)$ term, because N is big enough!

The free energy is:

$$F_N = F_N^{ideal} - k_BT\ln\left[1 - \frac{N^2}{V}B_2(T) + \dots\right] \tag{9.85}$$

Hence,

$$P_N = -\left(\frac{\partial F_N}{\partial V}\right)_{T,N} = \frac{Nk_BT}{V}\left(1 + \frac{\frac{N}{V}B_2}{1-\frac{N^2}{V}B_2}\right) = \frac{Nk_BT}{V}\left(\frac{1-\frac{N^2}{V}B_2+\frac{N}{V}B_2}{1-\frac{N^2}{V}B_2}\right) \tag{9.86}$$

Expanding the denominator for $\frac{N}{V}B_2 \ll 1$ $\rho \ll 1$ one gets

$$P_N \simeq \frac{Nk_BT}{V}\left(1 + \frac{N}{V}B_2 + \dots\right) \tag{9.87}$$

here we see the ideal gas and the correction to the ideal gas.

Remark. The equation (9.87) gives an important relation between experimentally accessible observables as P_N and microscopic quantities such as $f(\vec{\mathbf{r}})$ (and hence $\Phi(\vec{\mathbf{r}})$) trough the estimate of B_2.

Therefore, it is important computing B_2, because one time we have this we have the expansion. Or if we wish, by doing the fit of data at different temperature we obtain B_2 from the experiment and we can see f_{ij}.

Remark. The virial expansion obtained in (9.87) is valid for small densities.

To consider higher order terms in the virial expansion we need to consider higher order products of the f_{ij}.

Before doing this, however, one can show that an higher order expansion can be obtained by using the following (rather rude) trick.

Since $(1-x)^{-1} = 1 + x + \dots$, ny going backward it is possible to write

$$\frac{PV}{Nk_BT} \approx 1 + \rho B_2 + \dots \simeq \frac{1}{1-B_2\rho} \tag{9.88}$$

this is the Clausius equation. On the other hand,

$$\frac{1}{1-B_2\rho} \simeq 1 + B_2\rho + \underbrace{(B_2)^2\rho^2}_{B_3\approx B_2^2} + \underbrace{(B_2)^3\rho^3}_{B_4\approx B_2^3} + \dots \tag{9.89}$$

Identifying the coefficients for each power we get, for example

$$B_3 \approx (B_2)^2, \quad B_4 \approx (B_2)^3 \tag{9.90}$$

This is the approximation of higher order virial coefficients with powers of B_2.

Example 23. Exam: let us compute virial expansion of a gas in a potential.

9.7.1 Computation of B_2 for given Φ

Gas of hard spheres

The particles are interacting (it is not ideal!) and there is a size that is the range of the potential.

$$\Phi(r) = \begin{cases} \infty & r < \sigma \\ 0 & r \geq \sigma \end{cases} \tag{9.91}$$

Hence,

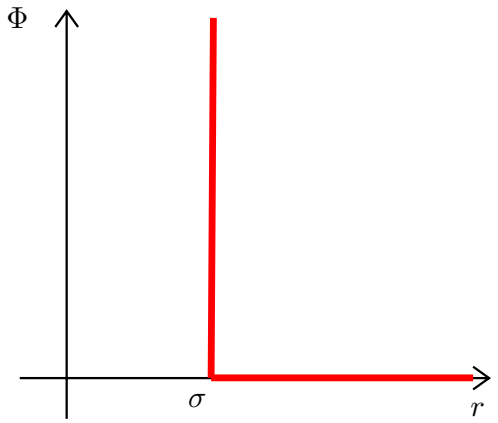


Figure 9.5: Plot of the potential $\Phi(r)$.

$$e^{-\beta\Phi(r)} = \begin{cases} 0 & r < \sigma \\ 1 & r \geq \sigma \end{cases} \tag{9.92}$$

This implies that

$$B_2(T) = -\frac{1}{2}\int_V \mathrm{d}\vec{\mathbf{r}}\, f(|\vec{\mathbf{r}}|) = -\frac{1}{2}4\pi\int_V \mathrm{d}r\, r^2\left[e^{-\beta\Phi(r)} - 1\right] = 2\pi\int_0^\sigma \mathrm{d}r\, r^2 = \frac{2}{3}\pi\sigma^3 \tag{9.93}$$

$$\Rightarrow B_2^{HS}(T) = \frac{2}{3}\pi\sigma^3 \tag{9.94}$$

this is the second virial coefficient for a hard sphere gas. There is no condensation in the gas spheres.

Remark. As expected B_2^{HS} does not depend on temperature (purely repulsive interaction).

For hard spheres we have:

$$PV = Nk_BT\left(1 + \frac{2}{3}\pi\sigma^3\frac{N}{V}\right) \tag{9.95}$$

Remark. The excluded volume interaction (hard sphere) increases the product PV with respect to the ideal gas.

Let us say, that the potential is not anymore zero but is $-\varepsilon$ when it is in the case $r \geq \sigma$. Or consider a case in which it is $-\varepsilon$ between $[\sigma, 2\sigma]$, then it goes to zero.

Gas with Lennard-Jones interaction

We can consider a Lennard-Jones potential. The 2-body Lennard-Jones potential energy is

$$\Phi = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{9.96}$$

where the first one is the repulsive term (Pauli excluded principle) and the second is the attractive term (fluctuation of the elecrtic dipole moment).

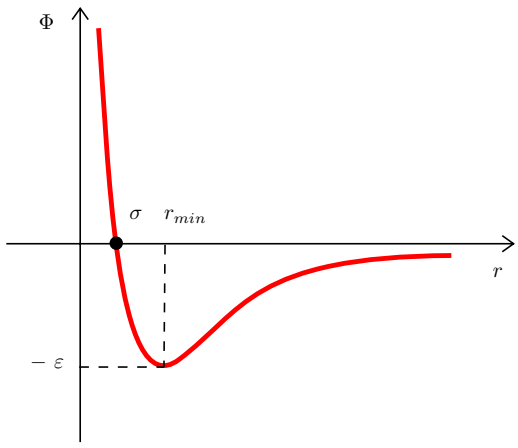


Figure 9.6: Plot of the Lennard-Jones potential.

The minimum is in $r_{min} = 2^{1/\sigma}$. You can play with the range of attraction by changing σ or by changing the ε. What it is important is that for the Lennard-Jones we have

$$B_2 \overset{LJ}{=} B_2(T) \tag{9.97}$$

$$B_2(T) = -2\pi \int_0^\infty r^2 \left[e^{-\frac{4\varepsilon}{k_B T}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]} - 1\right] \mathrm{d}r \tag{9.98}$$

It is not solvable exactly but by series expansion. Let us consider the following change of variables

$$x = \frac{r}{\sigma}, \quad \tau = \frac{k_B T}{\varepsilon} \tag{9.99}$$

Integrating by parts $\int f'g = fg - \int g'f$ where $f' = x^2 g = \exp[-()]$, we obtain

$$\begin{aligned} B_2(T^*) &= \frac{2}{3}\pi\sigma^3 \frac{4}{\tau} \int_0^\infty x^2 \left(\frac{12}{x^{12}} - \frac{6}{x^6}\right) e^{-\frac{4}{\tau}\left(\frac{1}{x^{12}} - \frac{1}{x^6}\right)} \mathrm{d}x \\ &= A \int_0^\infty \left(\frac{12}{x^{16}} - \frac{6}{x^4}\right) e^{-\frac{4}{\tau}\left(\frac{1}{x^{12}} - \frac{1}{x^6}\right)} \mathrm{d}x \end{aligned} \tag{9.100}$$

Expand the exponential and then integrate term by term. One gets a serie in inverse power of τ.

$$B_2(\tau) = -2A' \sum_{n=0}^{\infty} \frac{1}{4n!} \Gamma\left(\frac{2n-1}{4}\right) \left(\frac{1}{\tau}\right)^{\frac{2n+1}{4}} \tag{9.101}$$

where Γ is the Euler function.

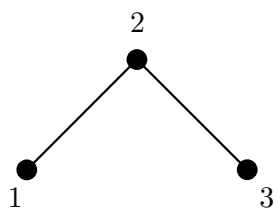
Remark. Note that, because the Lennard-Jones potential has an attractive interaction term $\left[-\left(\frac{\sigma}{r}\right)^6\right]$, the second virial coefficient depends on temperature.

9.7.2 Higher order terms in the cluster expansion

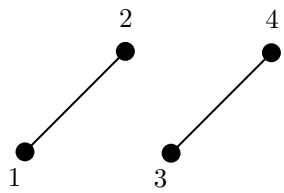
Let us consider again the formal expansion

$$\prod_i \left(\prod_{j>i} (1 + f_{ij}) \right) = 1 + \sum_{i,j>i} f_{ij} + \sum_{\substack{i \\ j>i \\ l>k \\ k\geq i \\ (ij)\neq(kl)}} f_{ij} f_{kl} + \dots \tag{9.102}$$

The problem with this expansion is that it groups terms quite different from one another. Fro example the terms $f_{12}f_{23}$ and $f_{12}f_{34}$. Indeed the first term correspond to a diagram as in Figure 9.7a, while the second to two disconnected diagrams as in Figure 9.7b.



(a) Description



(b) Description

Figure 9.7

Another problem of the above expansion is that it does not recognize identical clusters formed by different particles. For example the terms $f_{12}f_{23}$ and $f_{12}f_{14}$ contribute in the same way to the partition function. It is then convenient to follow a diagrammatic approach similar to the Feymann approach in the reciprocal space.

For the linear term f_{ij} the only diagram is given by Figure 9.8. As we have seen this has multeplicity $\frac{N(N-1)}{2}$ and the integral is of the form

$$\int f_{12}\, \mathrm{d}\vec{\mathbf{r}}_1\, \mathrm{d}\vec{\mathbf{r}}_2 = V \int f(\vec{\mathbf{r}})\, \mathrm{d}\vec{\mathbf{r}} = -2VB_2 \tag{9.103}$$

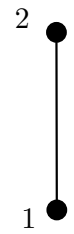


Figure 9.8

For the term $f_{ij}f_{kl}$ we can have the case as in Figure 9.9, that has molteplicity

$$\frac{N(N-1)}{2} \frac{(N-1)(N-3)}{2} \frac{1}{2} \tag{9.104}$$

and the integral is of the form

$$\int f_{12} f_{34}\, \mathrm{d}\vec{\mathbf{r}}_1\, \mathrm{d}\vec{\mathbf{r}}_2\, \mathrm{d}\vec{\mathbf{r}}_3\, \mathrm{d}\vec{\mathbf{r}}_4 \tag{9.105}$$

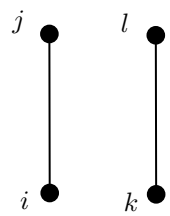


Figure 9.9

i.e. involving 4-particles

$$\begin{aligned} &\int f(|\vec{\mathbf{r}}_1 - \vec{\mathbf{r}}_2|) f(|\vec{\mathbf{r}}_3 - \vec{\mathbf{r}}_4|)\, \mathrm{d}\vec{\mathbf{r}}_1\, \mathrm{d}\vec{\mathbf{r}}_2\, \mathrm{d}\vec{\mathbf{r}}_3\, \mathrm{d}\vec{\mathbf{r}}_4 = \\ &= V^2 \left(\int f(\vec{\mathbf{r}})\, \mathrm{d}\vec{\mathbf{r}} \right)^2 = 4V^2 B_2^2 \end{aligned} \tag{9.106}$$

The next case if for instance as in Figure 9.10. This involves 3 particles. The multiplicity of this diagram is

$$\frac{N(N-1)(N-2)}{3!} \times 3 \tag{9.107}$$

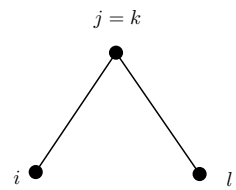


Figure 9.10

The integral is of the form

$$\begin{aligned}
&\int f_{12} f_{23}\, \mathrm{d}\vec{\mathbf{r}}_1\, \mathrm{d}\vec{\mathbf{r}}_2\, \mathrm{d}\vec{\mathbf{r}}_3 \simeq V\left(\int \mathrm{d}r\, f(r)\right)^2 = \\
&= \int f(|\vec{\mathbf{r}}_1 - \vec{\mathbf{r}}_2|) f(|\vec{\mathbf{r}}_2 - \vec{\mathbf{r}}_3|)\, \mathrm{d}\vec{\mathbf{r}}_1\, \mathrm{d}\vec{\mathbf{r}}_2\, \mathrm{d}\vec{\mathbf{r}}_3 = \\
&= V\left(\int f(\vec{\mathbf{r}})\, \mathrm{d}\vec{\mathbf{r}}\right)^2 = 4VB_2^2
\end{aligned} \tag{9.108}$$

Another interesting diagram is the one in Figure 9.11. Its molteplicity is

$$\frac{N(N-1)(N-2)}{3!} \tag{9.109}$$

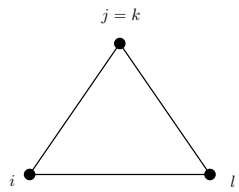


Figure 9.11

The associated integral involves 3 particles and it is of the form

$$\begin{aligned}
&\int f_{12} f_{23} f_{31}\, \mathrm{d}\vec{\mathbf{r}}_1\, \mathrm{d}\vec{\mathbf{r}}_2\, \mathrm{d}\vec{\mathbf{r}}_3 = \\
&= \int f(|\vec{\mathbf{r}}_1 - \vec{\mathbf{r}}_2|) f(|\vec{\mathbf{r}}_2 - \vec{\mathbf{r}}_3|) f(|\vec{\mathbf{r}}_3 - \vec{\mathbf{r}}_1|)\, \mathrm{d}\vec{\mathbf{r}}_1\, \mathrm{d}\vec{\mathbf{r}}_2\, \mathrm{d}\vec{\mathbf{r}}_3 \\
&= \int f(|\vec{\mathbf{r}}_1 - \vec{\mathbf{r}}_2|) f(|\vec{\mathbf{r}}_2 - \vec{\mathbf{r}}_3|) f(|\vec{\mathbf{r}}_3 - \vec{\mathbf{r}}_1|)\, \mathrm{d}\vec{\mathbf{r}}_2\, \mathrm{d}\vec{\mathbf{r}}_{21}\, \mathrm{d}\vec{\mathbf{r}}_{23}
\end{aligned} \tag{9.110}$$

On the other hand $\vec{\mathbf{r}}_{13} = \vec{\mathbf{r}}_{23} - \vec{\mathbf{r}}_{21}$, which implies

$$f(|\vec{\mathbf{r}}_3 - \vec{\mathbf{r}}_1|) = f(|\vec{\mathbf{r}}_{23} - \vec{\mathbf{r}}_{21}|) \tag{9.111}$$

hence,

$$\int f(|\vec{\mathbf{r}}_{12}|) f(|\vec{\mathbf{r}}_{23}|) f(|\vec{\mathbf{r}}_{31}|)\, \mathrm{d}\vec{\mathbf{r}}_{21}\, \mathrm{d}\vec{\mathbf{r}}_{23}\, \mathrm{d}\vec{\mathbf{r}}_2 = \int f(|\vec{\mathbf{r}}_{12}|) f(|\vec{\mathbf{r}}_{23}|) f(|\vec{\mathbf{r}}_{23} - \vec{\mathbf{r}}_{21}|)\, \mathrm{d}\vec{\mathbf{r}}_{21}\, \mathrm{d}\vec{\mathbf{r}}_{23}\, \mathrm{d}\vec{\mathbf{r}}_2 \tag{9.112}$$

Let us call this integral

$$\int f_{12} f_{23} f_{31}\, \mathrm{d}\vec{\mathbf{r}}_1\, \mathrm{d}\vec{\mathbf{r}}_2\, \mathrm{d}\vec{\mathbf{r}}_3 \equiv 3! V\left(B_3 - 2B_2^2\right) \tag{9.113}$$

The configurational partition function with these terms becomes

$$\begin{aligned}
Q_N(V,T) =& V^N - V^N \frac{N(N-1)}{V} B_2 + V^N \frac{N(N-1)(N-2)(N-3)}{8V^2}(4B_2^2) \\
&+ V^N \frac{N(N-1)(N-2)}{2V^2} 4B_2^2 \\
&= V^N \left(1 + \frac{N(N-1)}{V} B_2 + \frac{N(N-1)(N-2)(N-3)}{2V^2} B_2^2 + \frac{N(N-1)(N-3)}{V^2} B_3\right)
\end{aligned} \tag{9.114}$$

Let us now face the problem in a slightly different ways. Let us remind that

$$Q_N(V,T) = \sum_{diagrams} \int \prod_{kl} f_{kl}\, \mathrm{d}^{3N} r \tag{9.115}$$

i.e. sum over all possible diagrams i.e. all possible ways in which a can draw edges between pairs of points (k, l). For each such diagrams I have to product between all edge and then integrate over the configurational space (N points).

Let us now consider only *connected* diagrams for i sites. In other words given i points (i particles) from a system of N points and I consider all the possible ways I can connect these i points (an example is shown in Figure 9.12).

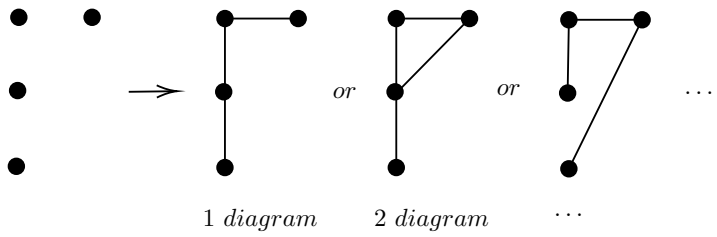


Figure 9.12: Example for $i = 4$.

For each diagram we take the product $\prod_{kl} f_{kl}$ and then integrate over the position of the i points (i particles). For a fixed diagram:

$$\int \prod_{lk \in \, diagram} f_{kl} \, \mathrm{d}\vec{\mathbf{r}}_1 \ldots \mathrm{d}\vec{\mathbf{r}}_i \tag{9.116}$$

Example 24 ($i = 4$)**.** For example the diagram 1 in Figure 9.12 gives the contribution

$$\int f_{12} f_{13} f_{34} \, \mathrm{d}\vec{\mathbf{r}}_1 \, \mathrm{d}\vec{\mathbf{r}}_2 \, \mathrm{d}\vec{\mathbf{r}}_3 \, \mathrm{d}\vec{\mathbf{r}}_4 \tag{9.117}$$

The diagram 2 gives

$$\int f_{12} f_{13} f_{23} f_{34} \, \mathrm{d}\vec{\mathbf{r}}_1 \, \mathrm{d}\vec{\mathbf{r}}_2 \, \mathrm{d}\vec{\mathbf{r}}_3 \, \mathrm{d}\vec{\mathbf{r}}_4 \tag{9.118}$$

and so on.

Finally we sum over all these connected diagrams of i points.

$$\sum_{\substack{\text{connected} \\ \text{diagrams}}} \int \prod_{lk \in \, diagram} f_{kl} \, \mathrm{d}\vec{\mathbf{r}}_1 \ldots \mathrm{d}\vec{\mathbf{r}}_i \tag{9.119}$$

This is what we call $i! V B_i$ and defines B_i.

- Case $i = 1$: clearly $B_1 = 1$.
- Case $i = 2$: just one edge, one connected diagram

$$\int f_{12} \, \mathrm{d}\vec{\mathbf{r}}_1 \, \mathrm{d}\vec{\mathbf{r}}_2 = -2 V B_2 \tag{9.120}$$

- Case $i = 3$: the connected diagrams are shown in Figure 9.13.

$$\begin{aligned} &= \underbrace{\int f_{12} f_{23} \, \mathrm{d}\vec{\mathbf{r}}_1 \, \mathrm{d}\vec{\mathbf{r}}_2 \, \mathrm{d}\vec{\mathbf{r}}_3 + \int f_{12} f_{13} \, \mathrm{d}\vec{\mathbf{r}}_1 \, \mathrm{d}\vec{\mathbf{r}}_2 \, \mathrm{d}\vec{\mathbf{r}}_3 + \int f_{13} f_{23} \, \mathrm{d}\vec{\mathbf{r}}_1 \, \mathrm{d}\vec{\mathbf{r}}_2 \, \mathrm{d}\vec{\mathbf{r}}_3}_{3V\left(\int f(\vec{\mathbf{r}}) \mathrm{d}\vec{\mathbf{r}}\right)^2} \\ &+ \underbrace{\int f_{12} f_{23} f_{13} \, \mathrm{d}\vec{\mathbf{r}}_1 \, \mathrm{d}\vec{\mathbf{r}}_2 \, \mathrm{d}\vec{\mathbf{r}}_3}_{3! V (B_3 - 2B_2^2)} \end{aligned} \tag{9.121}$$

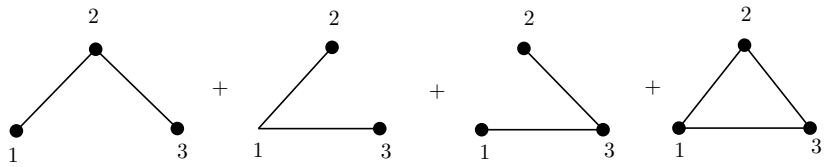


Figure 9.13: Description.

Hence,

$$\sum_{\substack{\text{connected diagrams}\\ \text{of } i=3 \text{ points}}} \int \prod_{kl \in \, diagram} f_{kl} \, \mathrm{d}\vec{\mathbf{r}}_1 \, \mathrm{d}\vec{\mathbf{r}}_2 \, \mathrm{d}\vec{\mathbf{r}}_3 = 3V(-2B_2)^2 + 6V(B_3 - 2B_2)$$
$$= 6VB_3 = 3!VB_3 \tag{9.122}$$

For the partition function we have to sum over all possible clusters. Possible procedure:

1. Given the N points we can partition the into connected clusters. For all i points we can make m_i clusters of that size i.

$$\sum_i i m_i = N \tag{9.123}$$

   For each cluster of size i we have a term $(i!VB_i)$. If there are m_i of them we have a weight $(i!VB_i)^{m_i}$.

2. Now we have to count in how many ways we can make the partition of N in a set of $\{m_i\}$ clusters. Clearly if we permute the label of the N vertices we have possible different clusters. In principle this degenerancy is proportional to $N!$

   On the other hand, if one changes the order of the labels within a cluster (in $i!$ ways) this does not change the cliuster and since there are m_i clusters of size i we have to divide by $(i!)^{m_i}$.

   Moreover, since there are m_i clusters one can swap them (in $m_i!$ ways). THe degenerancy is $\frac{N!}{m_i!(i!)^{m_i}}$. Therefore,

$$Q_N(V,T) = \sum_{\{m_i\}} \prod_i \frac{N!}{m_i!(i!)^{m_i}} (i!VB_i)^{m_i} \tag{9.124}$$

Example 25 ($N = q$)**.** Consider the point in Figure 9.14.

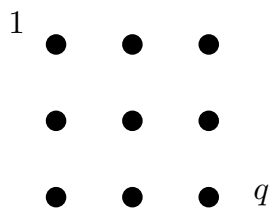


Figure 9.14: Description.

1. Parition these points into clusters.

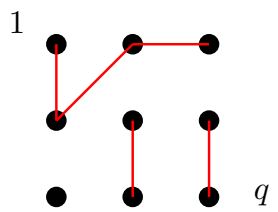


Figure 9.15: Description.

2. For this partition $\{m_i\}$ we have $m_4 = 1, m_2 = 2, m_1 = 1$. Now, the cluster of size 4 can be connected in a given different ways $(4!VB_4)^1$.

Exercise 7. TO DO: compute the degenerancy of this case.

Chapter 10

Landau theory of phase transition for homogeneous systems

10.1 Landau theory of phase transition (uniform systems)

It is a phenomenological mean field theory of phase transitions for uniform systems (no spatial variation of the order parameter). It is based on the following assumptions:

1. Existence of an unfirom order parameter η. Remember the definition of the order parameter:

$$\eta = \begin{cases} 0 & T \geq \bar{T} \text{ (disordered or symmetric phase)} \\ \neq 0 & T < \bar{T} \text{ (ordered symmetry is broken)} \end{cases} \tag{10.1}$$

   Well known examples are

$$\begin{cases} \eta \to m \\ \eta \to \rho_L - \rho_G \end{cases} \tag{10.2}$$

2. The free energy is an *analyc function* of the order parameter η. It is because you are doing the expansion close to...etc etc. Therefore, $\mathcal{L} = \mathcal{L}(\eta)$.

3. The form of $\mathcal{L}$ must satisfy the underlying symmetry of the system.

4. Equilibrium states correspond to the absolute minima of $\mathcal{L}$.

Remark. Since $\mathcal{L}$ os analytic it can be formally expanded in power of η, for $\eta \sim 0$.

$$\mathcal{L}(\eta) \approx a_0 + a_1\eta + a_2\eta^2 + a_3\eta^3 + \dots \tag{10.3}$$

10.2 Symmetries

To fix the ideas let us consider the theory for the Ising model. In this case η is a scalar (magnetization).

For $T > \bar{T}$ (critical point) we expect a paramagnetic phase. $\mathcal{L}$ has a minimum at $\eta = 0$, hence

$$\frac{\partial \mathcal{L}}{\partial \eta} = a_1 + 2a_2\eta + 3a_3\eta^2 + \dots = 0 \tag{10.4}$$

$\eta = 0$ is a solution if and only if $a_1 = 0$.

Remark. No linear term must be present!

Since Ising has $\mathbb{Z}^2$ symmetry, we should require

$$\mathcal{L}(-\eta) = \mathcal{L}(\eta) \tag{10.5}$$

which implies

$$a_k = 0 \quad \forall k \text{ odd} \tag{10.6}$$

Moreover, since $\mathcal{L}$ is analytic, terms proportional to $|\eta|$ are excluded. The minimal expression for $\mathcal{L}(\eta)$ that describes the equilibrium phase diagram of an Ising-like system is

$$\mathcal{L}(\eta) \simeq a_0(J,T) + a_2(J,T)\eta^2 + a_4(J,T)\eta^4 + O(\eta^6) \tag{10.7}$$

The coefficients of the expansion are functions of the physical parameters, J and T.

Since for $T > \bar{T}$, $\eta = \bar{\eta} = 0$ and

$$\mathcal{L}(\eta = 0) = a_0 \tag{10.8}$$

$a_0(T, J)$ value of $\mathcal{L}$ in the paramagnetic phase. Since what matters is the free-energy difference we can put $a_0 = 0$ identically.

Moreover, in order to have $\eta = \bar{\eta} \neq 0 < \infty$ for $T < \bar{T}$ (thermodynamic stability) we should impose that the coefficient of the highest power of η is always positive. In this case

$$a_4(J,T) > 0 \tag{10.9}$$

Indeed if this condition is violated $\mathcal{L}$ reaches it s absolute minimum for $|\eta| \to \infty$! Therefore,

$$\mathcal{L}(\eta) \simeq a_2\eta^2 + a_4\eta^4 \tag{10.10}$$

where he term a_4 it is positive and fixed.

If we now fix J and expand the coefficients a_2 and a_4 as a function of $t \equiv \frac{T-\bar{T}}{\bar{T}}$,

$$a_2 \sim a_2^0 + \frac{T-\bar{T}}{\bar{T}}\frac{a}{2} \tag{10.11a}$$

$$a_4 \sim \frac{b}{4} + \dots \tag{10.11b}$$

By choosing $a_2^0 = 0$ the sign of a_2 is determined by the one of

$$t \equiv \frac{T-\bar{T}}{\bar{T}} \tag{10.12}$$

In particular, at $T = \bar{T}$, one has $a_2 = 0$. Hence, for scalar, Ising like systems the minimal Landau free energy is given by

$$\mathcal{L} = \frac{a}{2}t\eta^2 + \frac{b}{4}\eta^4 + O(\eta^6) \tag{10.13}$$

Remark. Does not matter the coefficient in green in front, so in the next part of the course we will change it. If it is written in this way we have always $a > 0$. We have also $b > 0$.

10.3 Equilibrium phases

Now, the equilibrium states

$$\frac{\partial \mathcal{L}}{\partial \eta} = 0 \quad \Rightarrow at\eta + b\eta^3 = 0 \tag{10.14}$$

Hence,

$$\bar{\eta} = \begin{cases} 0 & T > \bar{T} \\ \pm\sqrt{\frac{-at}{b}} & T < \bar{T} \end{cases} \tag{10.15}$$

At $T = \bar{T}$ the 3 solutions coincide!

- Case $T > \bar{T}$ ($t > 0$): only one solution $\bar{\eta} = 0$.

$$\frac{\partial^2 \mathcal{L}}{\partial \eta^2} = at + 3b\eta^2 \geq 0 \tag{10.16}$$

  for $\bar{\eta} = 0$ and $t > 0$ implying that $\eta = \bar{\eta}$ is a global minimum, as in Figure 10.1.

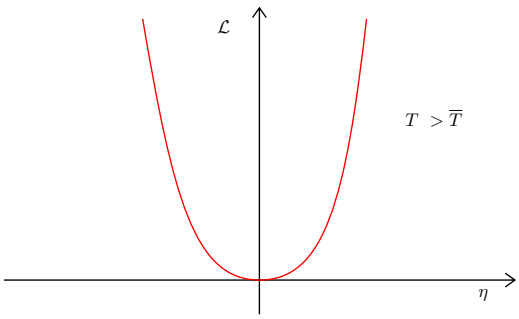


Figure 10.1: Description.

- Case $T < \bar{T}$ ($t < 0$): 3 solutions $\bar{\eta} = 0$ and $\bar{\eta} = \pm\sqrt{-\frac{at}{b}}$. Let us see wheter they are minima or local maxima.

$$\left.\frac{\partial^2 \mathcal{L}}{\partial \eta^2}\right|_{\bar{\eta}=0} = at < 0 \quad \Rightarrow \bar{\eta} = 0 \text{ local maximum (no equilibrium)} \tag{10.17}$$

$$\left.\frac{\partial^2 \mathcal{L}}{\partial \eta^2}\right|_{\bar{\eta}=\pm\sqrt{-\frac{at}{b}}} = at + 3b\left(-\frac{at}{b}\right) = -2at \tag{10.18}$$

  since $t < 0$, $-2at > 0$ and hence $\bar{\eta} = \pm\sqrt{-\frac{at}{b}}$ are 2 minima!

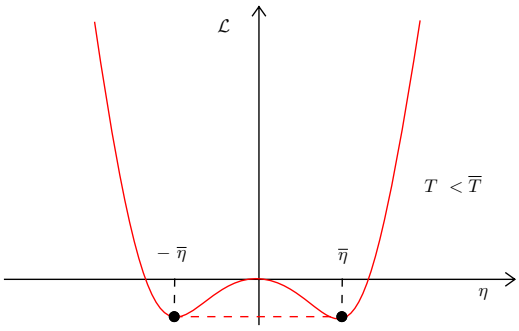


Figure 10.2: Description.

$$\mathcal{L}\left(\eta = \pm\sqrt{-\frac{at}{b}}\right) = -\frac{a^2t^2}{2b} + \frac{a^2t^2}{4b} = -\frac{a^2t^2}{4b} < 0 \tag{10.19}$$

The 2 minima are related by the group symmetry $\mathbb{Z}^2$ ($\bar{\eta} \to -\bar{\eta}$).

Remark. Note that, in presence of an external magnetic field h, one should consider the Legendre transform of $\mathcal{L}$ obtaining its Gibbs version:

$$\mathcal{L} = \frac{a}{2}t\eta^2 + \frac{b}{4}\eta^4 - h\eta \tag{10.20}$$

we have inserted a field coupled with the order parameter.

10.4 Critical exponents in Landau's theory

Consider $t \equiv \frac{T-\bar{T}}{\bar{T}}$.

Exponent β

We have $\eta \sim t^\beta$ for $h = 0$, $t \to 0^-$. Since $t < 0$, the minima of $\mathcal{L}$ are

$$\bar{\eta} = \pm\sqrt{-\frac{at}{b}} \quad \Rightarrow \beta = \frac{1}{2} \tag{10.21}$$

as expected.

Exponent α

We have $C \sim t^{-\alpha}$ for $h = 0$, $|t| \to 0$. Two cases:

- $t > 0$: $\bar{\eta} = 0$ and $\mathcal{L}(\bar{\eta}) = 0$.
- $t < 0$: $\mathcal{L}_{min} = \mathcal{L}\left(\bar{\eta} = \pm\sqrt{-\frac{at}{b}}\right) = -\frac{a^2t^2}{4b}$, that implies

$$\mathcal{L}_{min} = \begin{cases} 0 & t > 0 \\ -\frac{a^2t^2}{4b} & t < 0 \end{cases} \tag{10.22}$$

Hence,

$$c_V = -T\frac{\partial^2\mathcal{L}}{\partial T^2} = -T\frac{\partial^2}{\partial T^2}\left(-\frac{a^2}{4b\bar{T}^2}(T-\bar{T})^2\right) \tag{10.23}$$

We have

$$\frac{\partial}{\partial T} = -\frac{a^2}{2b\bar{T}^2}(T-\bar{T}) \tag{10.24}$$

$$\frac{\partial^2}{\partial T^2} = \frac{\partial}{\partial T}\left[-\frac{a^2}{2b\bar{T}^2}(T-\bar{T})\right] = -\frac{a^2}{2b\bar{T}^2} \tag{10.25}$$

$$c_V \begin{cases} 0 & T > \bar{T} \\ \frac{a^2}{2b\bar{T}^2}T & T < \bar{T} \end{cases} \tag{10.26}$$

We have $t \to 0^-$ if and only if $T \to \bar{T}^-$, which implies $c_V \to \frac{a^2}{2b\bar{T}}$ that is constant.

In obth cases $\alpha = 0$.

Exponent δ

We have $h \sim \eta^\delta$ at $T = \bar{T}$. Let us start from the equation of state. This is obtained by computing the $\frac{\partial}{\partial \eta}$ of the Gibbs version.

$$\frac{\partial \mathcal{L}_G}{\partial \eta} = at\eta + b\eta^3 - h = 0 \tag{10.27}$$

that is the condition for equilibrium. The equation of state is

$$h = at\eta + b\eta^3 \tag{10.28}$$

Equation (10.28) tells us that, for fixed h, the extreme points of $\mathcal{L}$ are given by the values of η that satisfies (10.28).

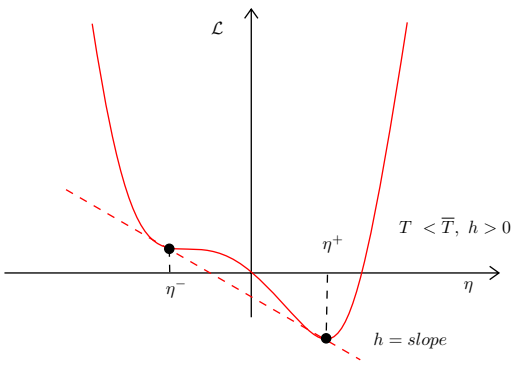


Figure 10.3: Description.

At $T = \bar{T}$ $(t = 0)$ we have $h \sim \eta^3$, hence $\delta = 3$.

Exponent γ

$\chi_T \sim t^{-\gamma}$ for $h = 0$, $|t| \to 0$. Let us derive the equation of state (10.28) with respect to h:

$$at\frac{\partial \eta}{\partial h} + 3b\eta^3\frac{\partial \eta}{\partial h} = 1 \tag{10.29}$$

and since

$$\chi = \frac{\partial \eta}{\partial h} \tag{10.30}$$

we have

$$\chi = \frac{1}{at + 3b\eta^2} \tag{10.31}$$

- Case $t > 0$, $\bar{\eta} = 0$: $\chi_T = -\frac{1}{at}$.
- Case $t < 0$, $\bar{\eta} = \pm\left(-\frac{at}{b}\right)^{1/2}$: $\chi_T = -\frac{1}{2at}$.

In both cases $\chi_T \sim 1/t$ and this gives

$$\gamma = \gamma' = 1 \tag{10.32}$$

Summary

In summary the Landau theory gives the following (mean field) values of the critical exponents

$$\beta = \frac{1}{2}, \quad \alpha = 0, \quad \delta = 3, \quad \gamma = 1 \tag{10.33}$$

Landau theory does not depend on the system dimension D (as expected since isa mean field theory) but only on its symmetries.

Remark. For a $O(n)$ (vector) model the order parameter η becomes a vector field $\vec{\eta}$ with n compnents and

$$\mathcal{L}_G(\vec{\eta}) = \frac{a}{2} t\vec{\eta} \cdot \vec{\eta} + \frac{b}{4}(\vec{\eta} \cdot \vec{\eta})^2 - \vec{\mathbf{h}} \cdot \vec{\eta} + O\Big((\vec{\eta} \cdot \vec{\eta})^3\Big) \tag{10.34}$$

Lecture 16. *Friday 6th December, 2019. Compiled: Thursday 26th December, 2019.*

10.5 Cubic term and first transition

Let us consider more general forms of the Landay free energy. For example, in the case in which the symmetry is not violated, one can consider also odd temrs such as the cubic one. In fact, we want to eneralize to include multicritical points, or phase transitions. The phase transitions can be obtained by introducing a cubic term in the Landau expansion. Remember that in the Ising model we have phase transition derived by symmetry breaking. Now, we have another type of phase transition.

The simplest Landau free energy that depends on a particular field is:

$$\mathcal{L}(\eta, t, h) = at\eta^2 - w\eta^3 + \frac{b}{4}\eta^4 - h\eta \tag{10.35}$$

where $t \equiv \frac{T-T^*}{2}$ and w is an additional parameter that we fix to be positive, $w > 0$.

Remark. For $w < 0$ the results are the same, but in the $\eta < 0$ diagram.

To satisfy thermodynamic stability we require $b > 0$, while

$$at = \frac{a}{2}(T - T^*) \quad \text{if} \begin{cases} > 0 & T > T^* \\ < 0 & T < T^* \end{cases} \tag{10.36}$$

The equation of state for $h \neq 0$ is:

$$\frac{\partial \mathcal{L}_G}{\partial \eta} = 0 \quad \Rightarrow h = 2at\eta - 3w\eta^2 + b\eta^3 \tag{10.37}$$

Let us consider the equilibrium states when $h = 0$:

$$h = 0 \quad \Rightarrow 0 = \eta(2at - 3w\eta + \beta\eta^2) \tag{10.38}$$

The possible solutions are

$$\begin{cases} \bar{\eta} = 0 & \text{disordered phase} \\ \bar{\eta} = \frac{1}{4b}\Big(3w \pm \sqrt{qw^2 - 16abt}\Big) & \text{ordered phases} \end{cases} \tag{10.39}$$

Let us rewrite the 'ordered' solutions as

$$\bar{\eta}_\pm = c \pm \sqrt{c^2 - \frac{at}{b}} \tag{10.40}$$

where

$$c = \frac{3w}{4b} \tag{10.41}$$

Note that

$$\bar{\eta}_\pm \in \mathbb{R} \iff c^2 > \frac{at}{b} \quad \Rightarrow \frac{T - T^*}{2} < \frac{c^2 b}{a} \equiv t^{**} \equiv \frac{T^{**} - T^*}{2} \tag{10.42}$$

It implies:

$$T^{**} = T^* + \frac{2c^2 b}{a} \tag{10.43}$$

- If $t > t^{**}$ ($\iff$ $T > T^{**}$), we have $\bar{\eta}_\pm \notin \mathbb{R}$. The only real solution is $\bar{\eta} = 0$ that is also the absolute minimum of $\mathcal{L}$. The plot is shown in Figure 10.4.

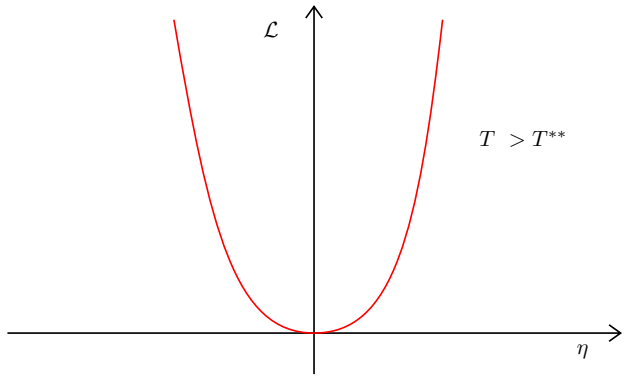


Figure 10.4: Description.

- If $t \leq t^{**}$ ($\iff$ $T \leq T^{**}$), we have $\bar{\eta}_\pm = c \pm \sqrt{c^2 - \frac{at}{b}} \in \mathbb{R}$ are both possible solutions. One will be a local maximum and the other a local minimum.

  - At $T = T^{**}$, $\bar{\eta}_- = \bar{\eta}_+$ (flex point), as shown in Figure 10.5.

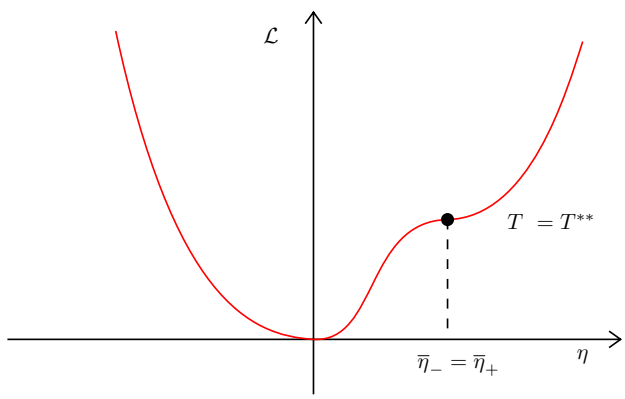


Figure 10.5: Description.

  - For $T_t < T < T^{**}$, $\mathcal{L}(\bar{\eta}_+) > 0$. Since $\mathcal{L}(\bar{\eta} = 0)$, the solution $\bar{\eta}_+$ is a local minimum, as shown in Figure 10.6.
  - By decreasing T firther one eventually reaches the value $T = T_t$ at which the local minimum $\mathcal{L}(\bar{\eta}_+)$ becomes zero, as in Figure 10.7. T_t is given by the coexistence condition

$$\mathcal{L}(\bar{\eta}_+) = \mathcal{L}(0) \tag{10.44}$$

    that is the coexistence between the disordered and ordered phases!

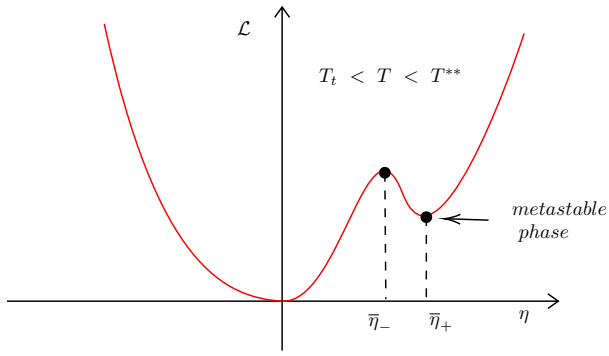


Figure 10.6: Description.

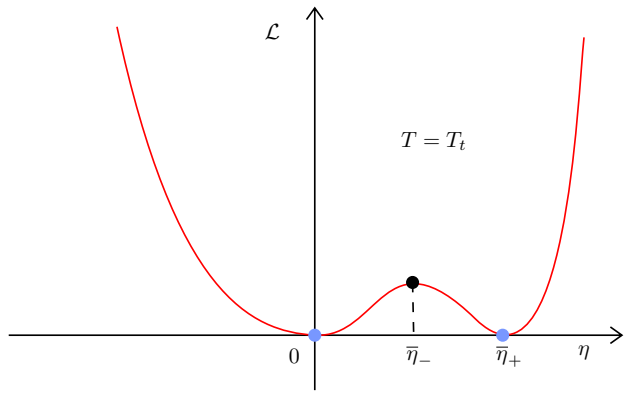


Figure 10.7: Description.

In the plot of Figure 10.7 we see that there are two minima in the same line, this is a first order transition. At $T = T_t$ the system undergoes a first order transition. To determine T_t we consider

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \eta} = 0 = \eta\big(2at - 3w\eta + 2b\eta^2\big) & \text{extreme condition} \\ \mathcal{L}(0) = \mathcal{L}(\eta_+) & \text{coexistence condition} \end{cases} \tag{10.45}$$

$$\Rightarrow \begin{cases} 2at - 3w\eta + 2b\eta^2 = 0 \\ at - w\eta + \frac{b}{2}\eta^2 = 0 \end{cases} \tag{10.46}$$

Solving with respect to η and t we get

$$\begin{cases} \bar{\eta}_{+t} = +\frac{w}{b} > 0 \\ t_t = \frac{w^2}{2ba} \equiv \frac{1}{2}(T_t - T^*) \end{cases} \tag{10.47}$$

$$T_t = \frac{w^2}{ba} + T^* \tag{10.48}$$

Remark. Note that $T_t > T^*$.

Since at $T = T_t$ there is a first order transition does the system display latent heat?

$$s = -\frac{\partial \mathcal{L}}{\partial T}\bigg|_{\eta_t} = -\frac{1}{2}a\bar{\eta}_t^2 = -\frac{a}{2}\Big(\frac{w}{b}\Big)^2 \tag{10.49}$$

Remark. There is an entropy jump.

The latent heat adsorbed to go from the ordered to the disordred phase is

$$q = -T_t s = \frac{a}{2}T_t\Big(\frac{w}{b}\Big)^2 \tag{10.50}$$

– Finally for $T^* < T < T_t$, $\eta = \bar{\eta}_+$ becomes the global minimum ordered phase is the only stable one (Figure 10.8).

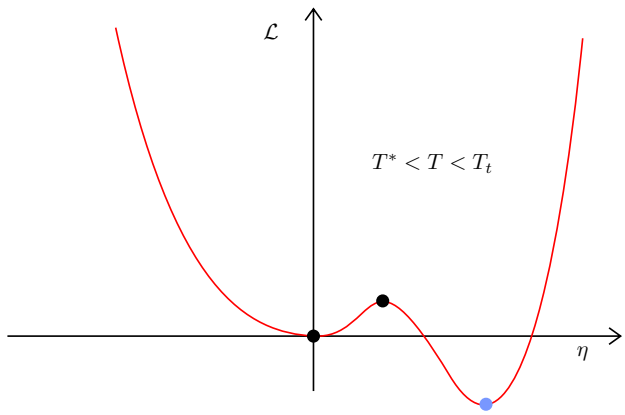


Figure 10.8: Description.

10.6 Phase stability and behaviour of $\chi_T \equiv \frac{\partial \eta}{\partial h}$

Let us derive the equation of state with respect to h

$$\frac{\partial}{\partial h}\left(\frac{\partial \mathcal{L}_G}{\partial \eta} = 0\right) = \frac{\partial}{\partial h}\left(2at\eta - 3w\eta^2 + 2b\eta^3 = h\right) \tag{10.51}$$

$$\Rightarrow \chi\left(2at - 6w\eta + 6b\eta^2\right) = 1 \tag{10.52}$$

The results is

$$\chi_T = \frac{1}{2at - 6w\eta + 6b\eta^2} \tag{10.53}$$

We now make use of equation (10.53) to compute the limit of stability of the phases we have found.

10.6.1 Computation of T^{**}

$T = T^{**}$ is the value below which the ordred phase ($\bar{\eta} = \bar{\eta}_+$) becomes a metastable state (local minima). Since this coincides with the flex point

$$\frac{\partial^2 \mathcal{L}}{\partial \eta^2} = 0 \quad \Rightarrow \chi^{-1} = 2at - 6w\eta + 6b\eta^2 = 0 \tag{10.54}$$

Remember that at $T = T{*}{*}$ the two solutions $\eta_\pm$ do coincide

$$c^2 - \frac{at}{b} = 0 \Rightarrow \eta_\pm = \eta_2 = \frac{3w}{4b} \tag{10.55}$$

Inserting in the expression for χ^{-1} since $\chi^{-1} = 0$ we have

$$\chi^{-1} = 0 = 2at^{**} - 6w\bar{\eta}_2 + 6\bar{\eta}_2^2 \tag{10.56}$$

$$\Longleftrightarrow \; t^{**} = \frac{qw^2}{16ab} = \frac{1}{2}(T^{**} - T^*) \tag{10.57}$$

For $T_t < T < T^{**}$ the ordered phase is metastable.

10.6.2 Computation of T^*

The instability of the disordered phase ($\eta = 0$) is when $\mathcal{L}$ presents a flex point at $\eta = 0$, as in Figure 10.9.

$$\chi^{-1}(\eta = 0) = 0 \iff \chi^{-1}(\eta = 0) = 2at = 0 \iff t = 0 \quad \Rightarrow T = T^* \tag{10.58}$$

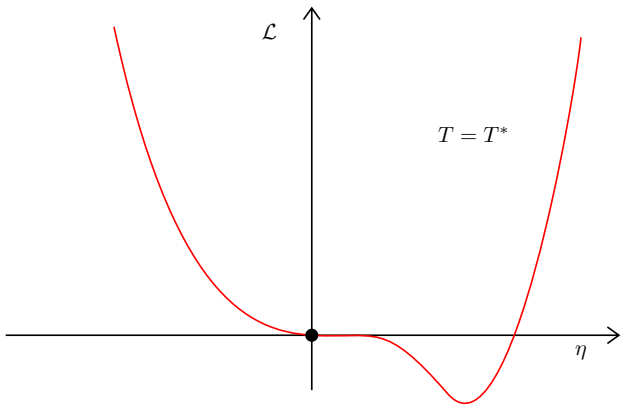


Figure 10.9: Description.

10.7 Landau theory and multicritical points

Tricritical point is a critical point that separates a line of first transition points from a line of critical points.

Remark. The introduction of a cubic term is not the only way to obtain a first order transition. Let allow the coefficient of η^4 to change signs. We need the η^6 term (See Blume-Emery-Griffith model).

$$\mathcal{L}_h(T, \Delta, \eta) = \frac{a(t, \Delta)}{2}\eta^2 + \frac{b(t, \Delta)}{4}\eta^4 + \frac{c}{6}\eta^6 - h\eta \tag{10.59}$$

where we have two parameteres (T, Δ) and $c > 0$, with Δ that is the disordered field (% He^3 in the BEG model).

Now, consider Δ (Δ_c is a critical value). The phenomenology is

- $\Delta < \Delta_c$: as T decreases, $a(T, \Delta)$ decreases and at $T = T_c(\Delta)$ becomes zero. In this region $b(T, \Delta) > 0$ and the system displays the standard (η^4) critical point.

$$T = T_c(\Delta) \Rightarrow \begin{cases} a(T_c, \Delta) = 0 \\ b(T_c, \Delta) > 0 \end{cases} \tag{10.60}$$

- $\Delta > \Delta_c$: as T decreases, $b(T, \Delta)$ becomes zero 'before' $a(T, \Delta)$. In this case one can show the existence of a line of first transition points:

$$b(\bar{T}, \Delta) = 0 \quad \Rightarrow \mathcal{L} = a\eta^2 + c\eta^6, \quad a > 0 \tag{10.61}$$

$$\frac{\partial \mathcal{L}}{\partial \eta} = 2a\eta + 6c\eta^5 \Rightarrow \begin{cases} \eta = 0 \\ \eta_{1,2,3,4} \end{cases} \tag{10.62}$$

If $b > 0$, η^6 can be neglected.

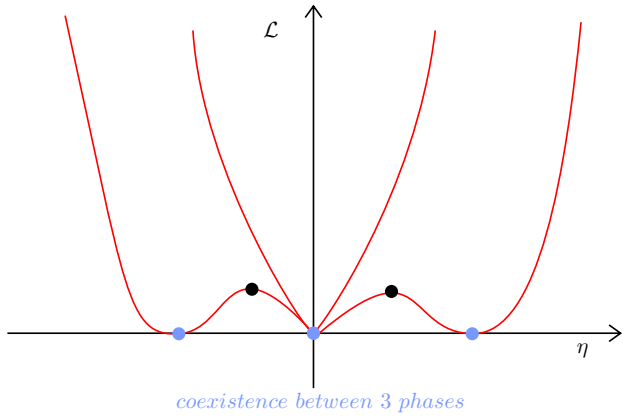


Figure 10.10: Description.

- The Tricritical point is given by the values of $\Delta = \Delta_t$ and $T = T_c$ such that

$$a(\Delta_t, T_t) = b(\Delta_t, T_t) = 0 \tag{10.63}$$

At this point the system is described by the following Landau free-energy

$$\mathcal{L}_t = c\eta^6 - h\eta \tag{10.64}$$

and equation of state

$$h = 6c\eta^5 \tag{10.65}$$

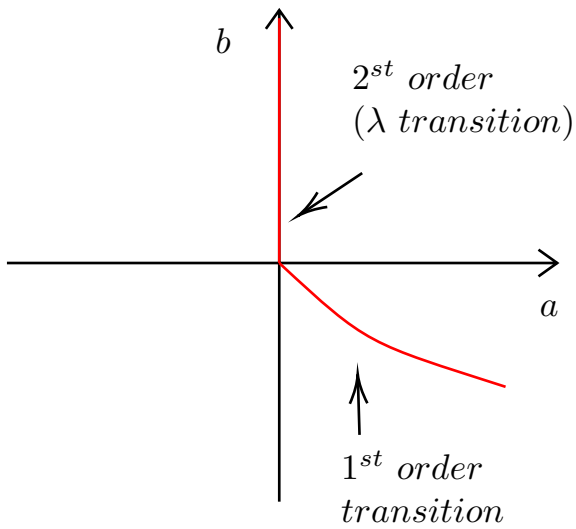


Figure 10.11: Description.

10.8 Landau-de Gennes theory of liquid crystals

A liquid crystal phase (LC) can be seen as an intermediate phase (mesophase) between a solid and a liquid phase. In this phase the system flows as a fluid but it also displays an orientational order typical of a crystal. Because of this order this phase displays *anisotropic* optical, magnetic and electrical properties.

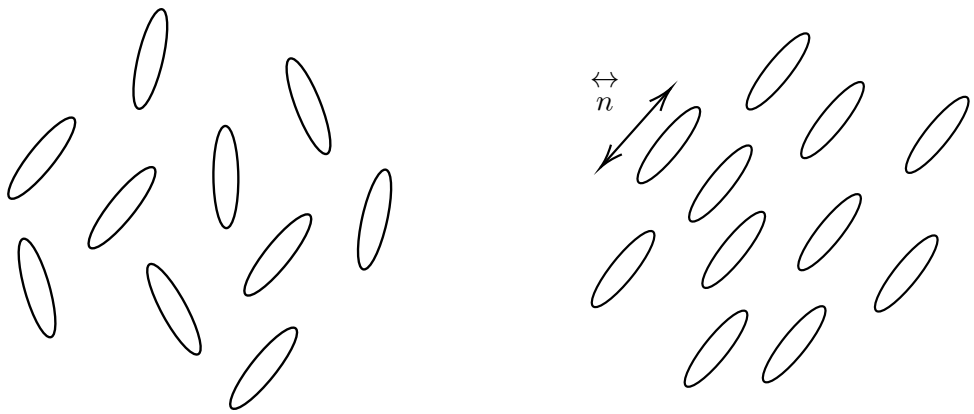
Typical structural properties of the elementary constituents (molecules) giving rise to LC phase are the following

- Anisotropic (elangated) shape.
- The long axis can be approximated as a rigid backbane.
- Existence of strong dipoles and groups easy to polarize.

10.8.1 LC-phases

There are many possible LC phase, from *nematic*, *cholesteric*, *smectic*, *columnar*, etc. Here we gocus on the most common one the *nematic phase*.

The nematic phase is characterized by a strong and long range orientational order. As a measure of this order, one consider the director $\overleftrightarrow{n}(\vec{\mathbf{r}})$. This is a 'two arrow vector' that gives the local average alignment of the elementary constituents. In this description the amplitude of $\overleftrightarrow{n}(\vec{\mathbf{r}})$ is irrelevant and one takes $\overleftrightarrow{n}(\vec{\mathbf{r}})$ such that $\left|\overleftrightarrow{n}(\vec{\mathbf{r}})\right| = 1$. Since there is no head-tail symmetry (apolar order), $\overleftrightarrow{n} = -\overleftrightarrow{n}$. From optical point of view the neumatic phase is birifrangent, i.e. two refraction indices, one parallel to $\overleftrightarrow{n}$ and one perpendicular to $\overleftrightarrow{n}$ (special index).



(a) Isotropic phase

(b) Nematic phase

Figure 10.12

10.8.2 Order parameter

A macroscopic definition of an order parameter for LC phase is based on the system response when subject to magnetic or electric fields. For instance, given an external magnetic field $\vec{\mathbf{H}}$ the (diamagnetic) response of the system can be written as

$$\vec{\mathbf{M}} = \bar{\bar{\chi}}\vec{\mathbf{H}} \tag{10.66}$$

where the matrix $\bar{\bar{\chi}}$ is the response function.

In components

$$M_\alpha = \chi_{\alpha\beta} H_\beta \tag{10.67}$$

with $\alpha, \beta = x, y, z$. For a static $\vec{\mathbf{H}}$, $\chi_{\alpha\beta}$ is symmetric.

Clearly, for a isotropic fluid

$$\chi_{\alpha\beta} = \chi\delta_{\alpha\beta} \tag{10.68}$$

For a LC in the neumatic phase one has

$$\chi_{\alpha\beta} = \begin{pmatrix} \chi_\perp & 0 & 0 \\ 0 & \chi_\perp & 0 \\ 0 & 0 & \chi_\parallel \end{pmatrix} \tag{10.69}$$

One can then define an order parameter based on $\chi_{\alpha\beta}$ as

$$Q_{\alpha\beta} = A\left(\chi_{\alpha\beta} - \frac{1}{3}\delta_{\alpha\beta}\,\mathrm{Tr}\,\bar{\bar{\chi}}\right) \tag{10.70}$$

The order parameter is a second rank traceless tensor. It is possible to show that $Q_{\alpha\beta}$ can be written in terms of the local average orientiational order of the elementary constituents, $\overleftrightarrow{n}(\vec{\mathbf{r}})$ and the degree of local order given by a scalar $S(\vec{\mathbf{r}})$.

$$Q_{\alpha\beta}(\vec{\mathbf{r}}) = S(\vec{\mathbf{r}})\left(n_\alpha(\vec{\mathbf{r}})n_\beta(\vec{\mathbf{r}}) - \frac{1}{3}\delta_{\alpha\beta}\right) \tag{10.71}$$

Note that by construtction $\bar{\bar{Q}}$ is symmetric and traceless.

Its more general expression is

$$Q_{\alpha\beta} = \begin{pmatrix} q_1 & q_2 & q_3 \\ q_2 & q_4 & q_5 \\ q_3 & q_5 & -q_1 - q_4 \end{pmatrix} \tag{10.72}$$

10.8.3 Landau free energy

The free energy must be invariant to rotations of the system. Since $Q_{\alpha\beta}$ transforms as a tensor under rotations, $\mathcal{L}$ must be a combination of terms as $\mathrm{Tr}\,\bar{\bar{Q}}^P$. By keeping terms up to fourth order

$$\mathcal{L} = \mathcal{L}_0 + \frac{1}{2}A(T)\,\mathrm{Tr}\,\bar{\bar{Q}}^2 + \frac{1}{3}B(T)\,\mathrm{Tr}\,\bar{\bar{Q}}^3 + \frac{1}{4}C(T)\left[\left(\mathrm{Tr}\,\bar{\bar{Q}}^2\right)^2 + \mathrm{Tr}\,\bar{\bar{Q}}^4\right] \tag{10.73}$$

In components,

$$\mathcal{L} = \mathcal{L}_0 + \frac{1}{2}A(T)Q_{\alpha\beta}Q_{\beta\alpha} + \frac{1}{3}B(T)Q_{\alpha\beta}Q_{\beta\gamma}Q_{\gamma\alpha} + \frac{1}{4}C(T)(Q_{\alpha\beta}Q_{\beta\alpha})^2 \tag{10.74}$$

Remark. Since each 3×3 matrix satisfies the relation

$$\mathrm{Tr}\,\bar{\bar{Q}}^4 = \frac{1}{2}\left(\mathrm{Tr}\,\bar{\bar{Q}}^2\right)^2 \tag{10.75}$$

the term proportional to $C(T)$ can be written as $\frac{1}{2}C(T)\,\mathrm{Tr}\,\bar{\bar{Q}}^4$.

Remark. The cubic term is here allowed since the rod like molecules have quadrupolar symmetry and the order parameter is a tensor for which the rotational invariance does not imply the absence of the odd powers.

For the most general case of a *biaxial nematic phase* $\bar{\bar{Q}}$ can be diagonalized giving

$$Q_{\alpha\beta} = \begin{pmatrix} \frac{2}{3}s & 0 & 0 \\ 0 & -\frac{s+\eta}{3} & 0 \\ 0 & 0 & -\frac{(s-\eta)}{3} \end{pmatrix} \tag{10.76}$$

where $\eta = 0$ corresponds to the standard uniaxial nematic phase.

In this cases

$$Q_{\alpha\beta} = \begin{pmatrix} \frac{2}{3}s & 0 & 0 \\ 0 & -\frac{s}{3} & 0 \\ 0 & 0 & -\frac{(s)}{3} \end{pmatrix} \tag{10.77}$$

and

$$\mathcal{L} = \mathcal{L}_0 + A(T)\frac{s^2}{3} + \frac{2}{27}B(T)s^3 + \frac{3}{81}C(T)s^4 \tag{10.78}$$

Assuming $A(T) \simeq A(T - T^*)$, $B(T) = B$ and $C(T) = C$:

$$\mathcal{L} = \mathcal{L}_0 + \frac{A}{3}(T - T^*)s^2 + \frac{2}{27}Bs^3 + \frac{C}{9}s^4 \tag{10.79}$$

that has the same form of the one studied before for the first order phase transition.

In particular

$$a = \frac{2}{3}A, \quad -w = \frac{2}{27}B, \quad b = \frac{2}{9}C \tag{10.80}$$

It is then easy to see that the isotropic-nematic transition is of the first order and occur at

$$T_t = \frac{w^2}{ba} + T^* = \frac{B^2}{27AC} + T^* \tag{10.81}$$

Chapter 11

Role of fluctuations in critical phenomena: Ginzburg criterium, Coarse-graining and Ginzburg-Landau theory of phase

11.1 Importance of fluctuations: the Ginzburg criterium

Maybe mean field is not a very good approximation in proximity of the critical point. Question: how bad is the mean field approximation in proximity of the critical point?

Mean field approximations neglect the fluctuations of the order parameter in the computation of Z. This is a rather strong assumption since the correlation length

$$\xi \overset{T\to T_c}{\sim} |T - T_c|^{-\nu} \tag{11.1}$$

diverges. Let us estimate the error one does in using the MF approximation.

To fix the ideas let us consider the Ising model. Since

$$\langle S_i S_j \rangle \xrightarrow{MF} \langle S_i \rangle \langle S_j \rangle \tag{11.2}$$

For each pair of spin (S_i, S_j) the relative error is

$$E_{ij} = \frac{|\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle|}{\langle S_i \rangle \langle S_j \rangle} \tag{11.3}$$

We need to compute the correlation function

$$G_c(i,j) \equiv \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle = \langle (S_i - \langle S_i \rangle)(S_j - \langle S_j \rangle) \rangle \tag{11.4}$$

Assuming *translational invariance*

$$G_c(i,j) \to G_c(|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|) \to G_c(r) \tag{11.5}$$

Remark. In order to compute $G_c(r)$ we cannot assume omogeneity since $\langle S_i \rangle = \langle S_j \rangle = m$. It implies that $G_c = 0$ identically in MF. We have to allow the system to display disomogeneities in space!

Remark. If we want to compute the error in the mean field, is always zero. So, if we want calculate the average with respect to fluctuations it does not work. We can either look at the variation in which the field is the internal one, or we can somehow try to make a variation not because of thermal fluctuations but because we control it. We do this by using an external field. This is the response theory with a variation of the field. (lesson)

Let us discuss an important point. The function G_c does not only descrive the spatial correlation of the fluctuations but also, through the linear response theory, the way in which m varies in response to an external non-homogeneous magnetic field. Within a mean field theory this is the only way to compute G_c!

Let us see how it works. In order to do that, instead using an H we use an H_i.

$$Z[H_i] = \mathrm{Tr}_{\{S\}} \left(e^{-\beta\left(-J\sum_{\langle ij\rangle} S_i S_j - \sum_i H_i S_i\right)} \right) \tag{11.6}$$

The definition of the thermal average is

$$\langle S_i \rangle = \frac{\mathrm{Tr}_{\{S\}} \left(S_i e^{-\beta\left(-J\sum_{\langle ij\rangle} S_i S_j - \sum_i H_i S_i\right)} \right)}{Z[H_i]} = \beta^{-1} \frac{\partial \ln Z}{\partial H_i} = -\frac{\partial F}{\partial H_i} \tag{11.7}$$

Similarly one can show that

$$\langle S_i S_j \rangle = \frac{\beta^{-1}}{Z} \frac{\partial^2 Z}{\partial H_i \partial H_j} \quad (to\, do) \tag{11.8}$$

Hence,

$$\begin{aligned} G_c(i,j) &= \frac{\beta^{-2}}{Z} \frac{\partial^2 Z}{\partial H_i \partial H_j} - \left(\frac{\beta^{-1}}{Z} \frac{\partial Z}{\partial H_i} \right) \left(\frac{\beta^{-1}}{Z} \frac{\partial Z}{\partial H_j} \right) \\ &= \beta^{-1} \frac{\partial^2 \ln Z}{\partial H_i \partial H_j} = -\frac{\partial^2 F(\{H_i\})}{\partial H_i \partial H_j} \end{aligned} \tag{11.9}$$

Remark. The response in i due to a variation of H in j is

$$\frac{\partial}{\partial H_j} \langle S_i \rangle = \frac{\partial}{\partial H_j} \left[\beta^{-1} \frac{\partial \ln Z}{\partial H_i} \right] = \frac{\partial}{\partial H_j} \left[-\frac{\partial F}{\partial H_i} \right] = \beta G_c(i,j) \tag{11.10}$$

The *generating functions* are:

- $Z[H_i]$: generating function of $G(i,j)$.
- $\ln Z[H_i] = -\beta F$: generating function of $G_c(i,j)$.

Remark. (lesson)

$$\frac{\partial M}{\partial H_j} = \sum_i \frac{\partial \langle S_i \rangle}{\partial H_j} = \sum_i G_c(i,j) \tag{11.11}$$

$$H_j = H_j(H) \tag{11.12}$$

$$\frac{\partial M}{\partial H} = \sum_j \frac{\partial M}{\partial H_j} \frac{\partial H_j}{\partial H} = \beta \sum_{ij} G_c(i,j) \tag{11.13}$$

the last one is the susceptibility χ_T. Therefore,

$$\chi_T = \beta \sum_{i,j} G_c(i,j) \tag{11.14}$$

$$G_c(i,j) \to G_c(|i-j|) \sim G(|\vec{\mathbf{r}}|) \tag{11.15}$$

11.2 Fluctuation-dissipation relation

$$Z_N = \mathrm{Tr}\exp\left[\beta J \sum_{\langle ij\rangle} S_i S_j + H \sum_i S_i\right] \tag{11.16}$$

where $H_i = H\,\forall i$.

$$M = \sum_i \langle S_i\rangle = \frac{1}{Z}\,\mathrm{Tr}\sum_i S_i \exp\left[\beta J \sum_{\langle ij\rangle} S_i S_j + H \sum_i S_i\right] = \frac{1}{\beta Z_N}\frac{\partial Z_N}{\partial H} \tag{11.17}$$

Similarly

$$\sum_{ij} \langle S_i S_j\rangle \overset{todo}{=} \frac{1}{\beta^2 Z_n}\frac{\partial^2 Z_N}{\partial H^2} \tag{11.18}$$

$$\begin{aligned}
\chi_T = \frac{\partial m}{\partial H} &= \frac{\partial}{\partial H}\left[-\frac{1}{N}\frac{\partial F}{\partial H}\right] = \frac{\partial}{\partial H}\left[\frac{1}{N}k_B T\frac{\partial \ln Z}{\partial H}\right] \\
&= \frac{1}{N}k_B T\left[\frac{\partial^2 \ln Z}{\partial H^2}\right] = \frac{1}{N}k_B T\left[\frac{1}{Z}\frac{\partial^2 Z}{\partial H^2} - \frac{1}{Z^2}\left(\frac{\partial Z}{\partial H}\right)^2\right] \\
&= \frac{1}{N\beta}\left[\beta^2 \sum_{ij}\langle S_i S_j\rangle - \beta^2\left(\sum_i \langle S_i\rangle\right)^2\right] \\
&= \frac{\beta}{N}\sum_{ij} G_c(i,j) = \frac{\beta}{N}\sum_{ij} G_c(\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j) \\
&= \beta \sum_i G_c(\vec{\mathbf{x}_\mathbf{i}})
\end{aligned} \tag{11.19}$$

where $\vec{\mathbf{x}}_i \equiv \vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j$.

$$\Rightarrow \chi_T = (a^d k_B T)^{-1}\int_\Omega \mathrm{d}^d\vec{\mathbf{r}}\, G_c(\vec{\mathbf{r}}) \tag{11.20}$$

this is the fluctuation dissipation relation.

11.2.1 Computation of E_R

The total relative error is the $E_R(r)$ integrated over the region of radius $R \le \xi$, i.e. where correlations are not negligible.

$$E_{TOT} = \frac{\int_0^\xi G_c(r)\,\mathrm{d}^D\vec{\mathbf{r}}}{\int_0^\xi \langle S_i\rangle\langle S_j\rangle\,\mathrm{d}^D\vec{\mathbf{r}}} \tag{11.21}$$

If $T < T_c$, there is a solution $\eta(r) = \eta \neq 0$

$$\Rightarrow \langle S_i\rangle\langle S_j\rangle \approx \eta^2 \tag{11.22}$$

is uniform in the region $|\vec{\mathbf{r}}| < \xi$.

$$E_{TOT} \sim \frac{\int_0^\xi G_c(\vec{\mathbf{r}})\,\mathrm{d}^D\vec{\mathbf{r}}}{\int_0^\xi \eta^2\,\mathrm{d}^D\vec{\mathbf{r}}} \ll 1 \tag{11.23}$$

This is the Ginzburg criterium. If satisfied the mean field theory is a valid approximation.

11.2.2 Estimation of E_{TOT} as $t \to 0^-$

The numerator of the (11.23) can be approximated as

$$\int_{V_\xi} G_c(r)\,\mathrm{d}^D r \overset{\substack{\text{fluctuation}\\ \text{dissipation}}}{\sim} k_B T_c \chi_T \sim t^{-\gamma} \tag{11.24}$$

On the other hand, the denominator

$$\int_{V_\xi} \eta^2\,\mathrm{d}^D r \sim \xi^D |t|^{2\beta} \sim t^{2\beta-\nu D} \tag{11.25}$$

To satisfy the Ginzburg criterium

$$E_{TOT} \overset{t\to 0^-}{\sim} t^{-\gamma+\nu D-2\beta} \ll 1 \tag{11.26}$$

As $t \to 0^-$ this is possible only if $-\gamma + \nu D - 2\beta > 0$, hence

$$D > \frac{\gamma + 2\beta}{\nu} \equiv D_c \tag{11.27}$$

is called the upper critical dimension.

- For $D < D_c$: fluctuations are relevant and mean field is no a good approximation.
- For $D > D_c$: fluctuations are less important and mean field describes properly the critical point.
- For $D = D_c$: mean field critical exponents ok but strong correction to the scaling expected. For a Ising-like systems (in the mean field)

  $$\beta = \frac{1}{2}, \quad \gamma = 1 \quad \Rightarrow D_c = \frac{2}{\nu} \tag{11.28}$$

  In order to compute D_c we need to compute ν withing the mean field approximation. We have to consider the system with spatial disomogeneitis!

  Remark. We have $\nu_{MF} = 1/2$. Therefore, the dimension is $D > 4$ for the mean field.

Now we have a lower critical dimension (remember the last lessons!) and an upper one.

Lecture 17. *Wednesday 11th December, 2019. Compiled: Thursday 26th December, 2019.*

11.3 Spatial variations: coarse graining procedure and the Ginzburg-Landau model

Since in proximity of the critical point the correlation length ξ diverges, there is no point in which we can see small scales. It is convenient to rewrite the microscopic partition function as an effective partition function obtained by integrating out the degrees of freedom over regions of linear size $L \gg a$ but still $\ll \xi$. This *coarse graining procedure* is formally given by

$$Z = \mathrm{Tr}_{\{S\}} \exp[-\beta\mathcal{H}[\{S\}]] = \int \mathrm{d}[m(\vec{\mathbf{r}})] \left[\sum_{\substack{\{S\}\\ \text{compatible with the}\\ \text{profile } m(\vec{\mathbf{r}})}} \exp[-\beta\mathcal{H}[\{S\}]] \right] \tag{11.29}$$

Defining

$$\sum_{\substack{\{S\} \\ \text{compatible with the} \\ \text{profile } m(\vec{\mathbf{r}})}} \exp[-\beta\mathcal{H}[\{S\}]] = e^{-\beta\mathcal{H}_{eff}(m(\vec{\mathbf{r}}))} \tag{11.30}$$

Remark. We trace over the configuration $\{S\}$ that give a profile $m(\vec{\mathbf{r}})$.

How, in pratice, can we perform the coarse graining procedure and obtain $\mathcal{H}_{eff}$? The idea is to partition the system into cells of size l and to integrate the microscopic degrees of freedom within each cell. We partition the configuration according to the magnetization profile. For example, if we have a configuration with half spin up and half down, we obtain a profile with 1 and -1. Suppose to consider the two dimensional system, we have many spins, in each square there is a huge number of spins (Figure 11.1). Once we have l, we replace what is inside with average:

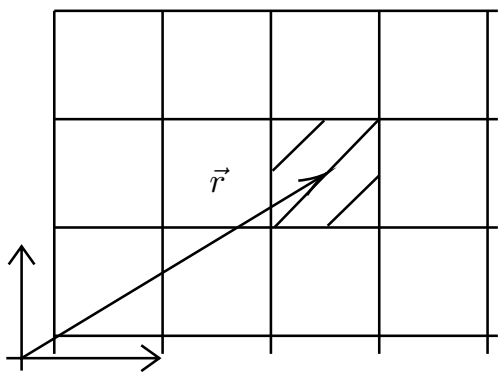


Figure 11.1: Description.

$$m_l(\vec{\mathbf{r}}) = \frac{1}{N_l}\sum_{i\in\vec{\mathbf{r}}} S_i \tag{11.31}$$

where $N_l = (l/a)^D$ is the number of spins in each cell, if D is the dimension of the system.

Remark. This approach works if one assumes

$$a \ll l \ll \xi(T) < L \tag{11.32}$$

Remark. Since close to T_c we have $\xi \gg a$, we can always choose $l \ll \xi$ but still $l \gg a$ such that the number N_l is large enough. In this way, $m(\vec{\mathbf{r}})$ can be made to be a regular function of $\vec{\mathbf{r}}$. The little l cannot go to zero, because it is a physical quantity.

Remark. In the reciprocal space (Fourier transform), bound (11.32) implies the following cut off on the wave vector $\vec{\mathbf{q}}$:

$$|\vec{\mathbf{q}}| > \Lambda = l^{-1} \tag{11.33}$$

No ultraviolet divergences can occur.

We have

$$Z = \sum_{m_l(\vec{\mathbf{r}})}\left(\sum_{\{S^*\}} e^{-\beta\mathcal{H}(\{S\})}\right) = \sum_{m_l(\vec{\mathbf{r}})} e^{-\beta\mathcal{H}_{eff}[m(\vec{\mathbf{r}})]} \tag{11.34}$$

If $m(\vec{\mathbf{r}})$ is regular, the sum converges to a functional integral

$$Z_{GL} = \int \mathrm{D}[m_l(\vec{\mathbf{r}})]e^{-\beta\mathcal{H}_{eff}[m]} \tag{11.35}$$

Let us now compute $\mathcal{H}_{eff}[m]$.

11.3.1 Computation of $\mathcal{H}_{eff}[m]$

First let us notice that

$$e^{-\beta\mathcal{H}_{eff}[m]} = \sum_{\{S^*\}} e^{-\beta\mathcal{H}(\{S\})} \tag{11.36}$$

is proportional to the probability that the system displays a configuration with a profile $m_l(\vec{\mathbf{r}})$. $\beta\mathcal{H}_{eff}[m]$ is made by two contributions:

1. Bulk contribution: terms relative to each cell. We would expect that the hamiltonian will be very similar to the Landau. It is quite reasonable to say that in this case, *inside* a given cell $l \ll \xi$: the boltzmann weight it is related to the exponent.

   Since within each cell the system is uniform, we can consider the Landau free energy

$$\beta\mathcal{H}^b_{eff}[m] = \bar{a}tm^2 + \frac{\bar{b}}{2}m^4 \tag{11.37}$$

$$P^{cell}(m_l(\vec{\mathbf{r}})) \simeq \exp\left[-\bar{a}tm^2 - \frac{b}{l}m^4\right] \tag{11.38}$$

2. Surface term: term due to the interactions between adjiacant cells. Let us consider a cell (Figure 11.2, we can thing about different interactions. We have

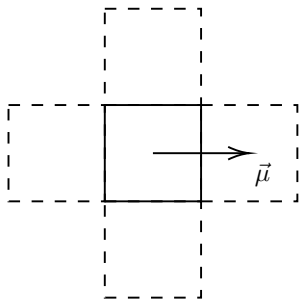


Figure 11.2: Description.

   $|\vec{\mu}| = l$, where $\vec{\mu}$ is a vector pointing towards the 2D neighbours. If we consider a simple quadratic interaction we can write

$$-\beta\mathcal{H}^s_{eff} = \sum_{\vec{\mu}} \frac{\bar{K}}{2}\left[(m_l(\vec{\mathbf{r}}+\vec{\mu}) - m_l(\vec{\mathbf{r}}))^2\right] + O\left[(m_l(\vec{\mathbf{r}}+\vec{\mu}) - m_l(\vec{\mathbf{r}}))^4\right] \tag{11.39}$$

   Remark. If l is small enough respect to ξ, we have other terms in this expansion ($O[\dots]$).

The total energy is obtained by summing over all the cells: is just given by the sum of the two term. Now we have to sum over all the cell if we want to have the energy of the whole system. We also look for the continuum limit:

$$\sum_{\vec{\mathbf{r}}} \xrightarrow{\frac{l}{L}\ll 1} \frac{1}{l^D}\int \mathrm{d}^D\vec{\mathbf{r}} \tag{11.40}$$

Remark. A lezione avevo scritto per $\frac{l^D}{r} \ll 1$, è la stessa cosa?????

Let us consider the continuum limit for the

- Surface term.

$$\begin{aligned}
\frac{\bar{k}}{2}\sum_{\vec{\mu}}\sum_{\vec{\mathbf{r}}}(m_l(\vec{\mathbf{r}}+\vec{\mu})-m_l(\vec{\mathbf{r}}))^2 &\xrightarrow{\frac{l}{L}\ll 1} \frac{\bar{k}}{2l^{D-2}}\sum_{\vec{\mu}}\int \mathrm{d}^D\vec{\mathbf{r}}\left(\frac{m_l(\vec{\mathbf{r}}+\vec{\mu})-m_l(\vec{\mathbf{r}})}{l}\right)^2 \\
&= \frac{\bar{k}}{2l^{D-2}}\int \mathrm{d}^D\vec{\mathbf{r}}\sum_{\vec{\mu}}\left(\frac{m_l(\vec{\mathbf{r}}+\vec{\mu})-m_l(\vec{\mathbf{r}})}{l}\right)^2 \\
&\overset{\frac{a}{L}\ll\frac{l}{L}\ll 1}{\simeq} \int \frac{\bar{k}}{2l^{D-2}}\sum_{\vec{\mu}}\left(\frac{\partial m_l}{\partial \chi_\mu}\right)^2 \mathrm{d}^D\vec{\mathbf{r}} \\
&\xrightarrow{\frac{a}{L}\ll\frac{l}{L}\ll 1} \int \mathrm{d}^D\vec{\mathbf{r}}\,\frac{k}{2}\left(\vec{\boldsymbol{\nabla}} m_l(\vec{\mathbf{r}})\right)^2
\end{aligned} \tag{11.41}$$

where

$$k \equiv \frac{\bar{k}}{l^{D-2}} \tag{11.42}$$

so rescaling the elastic constant by l^{D-2}. In this way the result is indipendent on l.

- Bulk term. We have

$$\sum_{\vec{\mathbf{r}}} -\bar{a}tm^2(\vec{\mathbf{r}}) \xrightarrow{\frac{l}{L}\ll 1} \frac{1}{l^D}\int -\bar{a}tm^2(\vec{\mathbf{r}})\,\mathrm{d}^D\vec{\mathbf{r}} = \int -atm^2(\vec{\mathbf{r}})\,\mathrm{d}^D\vec{\mathbf{r}} \tag{11.43}$$

with

$$a \equiv \frac{\bar{a}}{l^D} \tag{11.44}$$

Similarly we have

$$b \equiv \frac{\bar{b}}{l^D} \tag{11.45}$$

Consider

$$\beta\mathcal{H}_{eff}[m] = \int \mathrm{d}^D\vec{\mathbf{r}}\left[atm^2(\vec{\mathbf{r}}) + \frac{b}{2}m^4(\vec{\mathbf{r}}) + \frac{k}{2}(\vec{\boldsymbol{\nabla}} m(\vec{\mathbf{r}}))^2\right] \tag{11.46}$$

$$Z_{GL} = \int \mathrm{D}[m(\vec{\mathbf{r}})]e^{-\beta\mathcal{H}_{eff}} \tag{11.47}$$

Some considerations:

1. If $m(\vec{\mathbf{r}}) = m$ (uniform system) we go back to the Landau theory.

2. For non uniform system there is a new term given by $\left(\vec{\boldsymbol{\nabla}} m(\vec{\mathbf{r}})^2\right)$. This term can be also added directly to the Landau theory by simply assuming that, whe $m \to m(\vec{\mathbf{r}})$, one has to consider an additioned energy cost due to small variation of m.

Why we take $\left(\vec{\boldsymbol{\nabla}} m\right)^2$ and not something else? The choise is first of all a consequence of the isotropy of the system (all directions are equivalent). Among all the possible combinations of partial derivatives that are invariant by spatial rotations the term, $\left(\vec{\boldsymbol{\nabla}} m\right)^2$ is the simplest one.

Remark. When $m \to \vec{\mathbf{m}}$ ($O(n)$ models) we have

$$(\boldsymbol{\nabla}\vec{\mathbf{m}})^2 = \sum_{i=1}^{n}\sum_{\alpha=1}^{D}\sum_{\beta=1}^{D} \partial_\alpha m_i \partial_\beta m_i \tag{11.48}$$

Higher order terms are

$$\left(\boldsymbol{\nabla}^2\vec{\mathbf{m}}\right)^2 = \sum_{i=1}^{n}\sum_{\alpha=1}^{D}\sum_{\beta=1}^{D} (\partial_\alpha\partial_\alpha m_i)(\partial_\beta\partial_\beta m_i) \tag{11.49}$$

and

$$\vec{\mathbf{m}}^2(\boldsymbol{\nabla}\vec{\mathbf{m}})^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\alpha=1}^{D} m_i m_i \partial_\alpha m_j \partial_\alpha m_j \tag{11.50}$$

In most cases it is sufficient to consider only the lowest order term.

11.3.2 Magnetic non homogeneous field

$$\vec{\mathbf{h}}(\vec{\mathbf{r}}) \equiv \beta\vec{\mathbf{H}}(\vec{\mathbf{r}}) \tag{11.51}$$

Additional term (Legendre transform)

$$-\int \mathrm{d}^D\vec{\mathbf{r}}\,\vec{\mathbf{h}}(\vec{\mathbf{r}})\cdot\vec{\mathbf{m}}(\vec{\mathbf{r}}) \tag{11.52}$$

Going back to the scalar case, the usual partition function is

$$Z^h_{GL} = \int \mathrm{D}[m(\vec{\mathbf{r}})] e^{-\int \mathrm{d}^D\vec{\mathbf{r}}\left[atm^2(\vec{\mathbf{r}})+\frac{b}{2}m^4(\vec{\mathbf{r}})+\frac{k}{2}(\bar{\boldsymbol{\nabla}}m(\vec{\mathbf{r}}))^2-h(\vec{\mathbf{r}})m(\vec{\mathbf{r}})\right]} \tag{11.53}$$

and we define

$$F_{eff}[m] = -\int \mathrm{d}^D\vec{\mathbf{r}}\left[atm^2(\vec{\mathbf{r}}) + \frac{b}{2}m^4(\vec{\mathbf{r}}) + \frac{k}{2}(\bar{\boldsymbol{\nabla}}m(\vec{\mathbf{r}}))^2 - h(\vec{\mathbf{r}})m(\vec{\mathbf{r}})\right] \tag{11.54}$$

Functional derivatives

$$\frac{\delta G(h(\vec{\mathbf{r}}))}{\delta h(\vec{\mathbf{r}})} \equiv \lim_{\varepsilon\to 0}\frac{G(h[\vec{\mathbf{r}}+G]) - G(h[\vec{\mathbf{r}}])}{\varepsilon} \tag{11.55}$$

Example 26.

$$\frac{\delta f(\vec{\mathbf{r}}')}{\delta f(\vec{\mathbf{r}})} = \delta^D(\vec{\mathbf{r}}' - \vec{\mathbf{r}}) \tag{11.56}$$

$$\frac{\delta}{\delta m(\vec{\mathbf{r}})}\left[\int \mathrm{d}^D\vec{\mathbf{r}}'\,\frac{k}{2}\left(\boldsymbol{\nabla}m(\vec{\mathbf{r}}')\right)^2\right] \overset{todo}{=} -k\boldsymbol{\nabla}^2 m \tag{11.57}$$

where

$$\delta G[\boldsymbol{\nabla}m] = \left[\int \mathrm{d}^D\vec{\mathbf{r}}'\,\frac{k}{2}(\boldsymbol{\nabla}m(\vec{\mathbf{r}}))^2\right] \tag{11.58}$$

$$\delta G = G[\boldsymbol{\nabla}[m+\delta m]] - G[\boldsymbol{\nabla}m] \tag{11.59}$$

The average is

$$\langle m(\vec{\mathbf{r}})\rangle = -\frac{\delta F}{\delta h(\vec{\mathbf{r}})} = -\frac{\delta \ln Z[h]}{\delta h(\vec{\mathbf{r}})} \tag{11.60}$$

and one can show that the magnetic suscpetibility is

$$\begin{aligned}\chi(\vec{\mathbf{r}},\vec{\mathbf{r}}') &= \frac{\delta^2 F}{\delta h(\vec{\mathbf{r}})\delta h(\vec{\mathbf{r}}')} = \beta^{-1}\frac{\delta^2 \ln Z[h]}{\delta h(\vec{\mathbf{r}})\delta h(\vec{\mathbf{r}}')} \\ &= \beta^{-1}\left[\left\langle m(\vec{\mathbf{r}})m(\vec{\mathbf{r}}')\right\rangle - \left\langle m(\vec{\mathbf{r}})\right\rangle\left\langle m(\vec{\mathbf{r}}')\right\rangle\right] \\ &= \beta^{-1}G_c(\vec{\mathbf{r}},\vec{\mathbf{r}}')\end{aligned} \tag{11.61}$$

The problem is again try to approximate this term as much as we can. Let us compute the approximation.

11.4 Saddle point approximation and Landau theory for non-homogeneous systems

$$Z_{GL}[h] = \int \mathrm{D}[m] e^{-\int \mathrm{d}^D\vec{\mathbf{r}}\left[\beta\mathcal{H}_{eff}[m] - h(\vec{\mathbf{r}})m(\vec{\mathbf{r}})\right]} \tag{11.62}$$

where $h(\vec{\mathbf{r}}) = \beta H(\vec{\mathbf{r}})$ and

$$\beta\mathcal{H}_{eff}[m] = atm^2 + \frac{b}{2}m^4 + \frac{k}{2}(\boldsymbol{\nabla} m)^2 \tag{11.63}$$

We noe approximate the functional integral with its dominant term, i.e. with the one for which the exponent

$$L(m,h) = \int \mathrm{d}^D\vec{\mathbf{r}}\,[\beta\mathcal{H}_{eff} - hm] \tag{11.64}$$

is minimum. Let $m_0(\vec{\mathbf{r}})$ be the profile for which $L(m_0(\vec{\mathbf{r}}), h(\vec{\mathbf{r}}))$ is minimum. It implies

$$Z_{GL}[h] \overset{\substack{\text{Saddle}\\ \text{point}}}{\simeq} Z^0_{GL}[h] = e^{-L[m_0]} \tag{11.65}$$

In order to find the minimum, one has to consider the stationarity condition

$$\delta L = 0 \tag{11.66}$$

where

$$L[m,h] = \int \mathrm{d}^D\vec{\mathbf{r}}\,[\beta\mathcal{H}_{eff}[m] - hm] \tag{11.67}$$

By considering δL with respect to the variations δm and $\delta(\boldsymbol{\nabla} m)$ one gets the equation of state

$$h(\vec{\mathbf{r}}) = -\left[\boldsymbol{\nabla}\left(\frac{\partial\mathcal{H}_{eff}}{\partial(\boldsymbol{\nabla} m)}\right) - \frac{\partial\mathcal{H}_{eff}}{\partial m}\right] \tag{11.68}$$

Remark. For $h = 0$, equation (11.68) reduces to the Euler-Lagrande equation

$$\frac{\partial\mathcal{H}_{eff}}{\partial m} = \boldsymbol{\nabla}\left(\frac{\partial\mathcal{H}_{eff}}{\partial(\boldsymbol{\nabla} m)}\right) \tag{11.69}$$

Indeed, let

$$L[m,\boldsymbol{\nabla} m,h] = \int \mathrm{d}^D\vec{\mathbf{r}}\,\mathcal{L}(m,\boldsymbol{\nabla} m,h) \tag{11.70}$$

Let us consider $h = 0$ and look for the variation of L with respect to the variations of m, δm and the variation of $\boldsymbol{\nabla} m$, $\delta(\boldsymbol{\nabla} m)$,

$$\begin{aligned}\delta L[m, \boldsymbol{\nabla} m, 0] &= \int \mathrm{d}^D\vec{\mathbf{r}}\, \mathcal{L}(m + \delta m, \boldsymbol{\nabla} m + \delta(\boldsymbol{\nabla} m)) - \int \mathrm{d}^D\vec{\mathbf{r}}\, \mathcal{L}(m, \boldsymbol{\nabla} m) \\ &= \int \mathrm{d}^D\vec{\mathbf{r}}\, [\mathcal{L}(m + \delta m, \boldsymbol{\nabla} m + \delta(\boldsymbol{\nabla} m)) - \mathcal{L}(m, \boldsymbol{\nabla} m)] \\ &= \int \mathrm{d}^D\vec{\mathbf{r}} \left[\frac{\partial \mathcal{L}}{\partial m} \delta m + \underset{f}{\frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla} m)}} \underset{g'}{\delta(\boldsymbol{\nabla} m)} \right]\end{aligned} \tag{11.71}$$

We now integrate by parts the second term

$$\delta L = \int \mathrm{d}^D\vec{\mathbf{r}} \left[\frac{\partial \mathcal{L}}{\partial m} \delta m - \boldsymbol{\nabla} \frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla} m)} \delta m \right] + \underbrace{\int_V \mathrm{d}^D\vec{\mathbf{r}}\, \boldsymbol{\nabla} \left[\frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla} m} \delta m \right]} \tag{11.72}$$

Reming the divergence theorem $\int_{\partial\Omega} F = \int_V \boldsymbol{\nabla} F$, the second term vanishes at the boundary of the integration.

$$\delta L = \int \mathrm{d}^D\vec{\mathbf{r}} \left[\frac{\partial \mathcal{L}}{\partial m} - \boldsymbol{\nabla} \frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla} m)} \right] \delta m \tag{11.73}$$

Remind the stationarity condition (11.66), $\forall \delta m \neq 0$, if and only if the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial m} - \boldsymbol{\nabla} \left(\frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla} m)} \right) = 0 \tag{11.74}$$

In our case, when $h = 0$ we have $\mathcal{L} \to \beta \mathcal{H}_{eff}$. Since

$$\beta \mathcal{H}_{eff} = atm^2 + \frac{b}{2} m^4 + \frac{k}{2} (\boldsymbol{\nabla} m)^2 \tag{11.75}$$

the Euler-Lagrange equation become

$$h(\vec{\mathbf{r}}) = -k \boldsymbol{\nabla}^2 m_0(\vec{\mathbf{r}}) + 2atm_0(\vec{\mathbf{r}}) + 2bm_0^3(\vec{\mathbf{r}}) \tag{11.76}$$

this is the mean field solution of the Gibbs-Landau. It is more general than the Landau we have used. It is just the Landau with the additional term $\boldsymbol{\nabla}$.

Remark. If $h(\vec{\mathbf{r}}) = h$ (homogeneous field) and $m_0(\vec{\mathbf{r}}) = m_0$, equation reduces to the equation of state of the Landau theory of uniform system

$$h = 2atm_0 + 2bm_0^3 \tag{11.77}$$

Remark. A mean field theory of systems with spatial disomogeneity can start directly by considering the free energy functional

$$F[m, h] = \int \mathrm{d}^D\vec{\mathbf{r}} \left[atm^2(\vec{\mathbf{r}}) + \frac{b}{2} m^4(\vec{\mathbf{r}}) + \frac{k}{2} (\boldsymbol{\nabla} m)^2 - h(\vec{\mathbf{r}}) m(\vec{\mathbf{r}}) \right] \tag{11.78}$$

lesson

$$Z_{GL}^h = \int \mathrm{D}[m] \exp(-F_{eff}[m]) \tag{11.79}$$

In the frist approximation, we replace this integral with just

$$\simeq \exp[-\beta F_{eff}[\bar{m}]] \tag{11.80}$$

where $\bar{m}(\vec{\mathbf{r}})$ is such that $F_{eff}[\bar{m}]$ is minimal.

$$\bar{m} = \min_m [F_{eff}[m]] \tag{11.81}$$

Taking the variation of δF_{eff} when $m \to m + \delta m$:

$$\delta F_{eff} = F_{eff}[m + \delta m] - F_{eff}[m] \tag{11.82}$$

$$h(\vec{\mathbf{r}}) = -\left[\boldsymbol{\nabla}\left(\frac{\partial \mathcal{H}_{eff}}{\partial(\boldsymbol{\nabla} m)}\right) - \frac{\partial \mathcal{H}_{eff}}{\partial m}\right] \tag{11.83}$$

11.5 Nono homogeneous systems and correlation functions within a mean field approximation

Let us now try to compute the correlation function. Starting from (11.76), we derive it (is a functional derivative!) with respect to $h(\vec{\mathbf{r}}')$.

Remembering that

$$\chi_T(\vec{\mathbf{r}}, \vec{\mathbf{r}}') = \frac{\delta m(\vec{\mathbf{r}})}{\delta h(\vec{\mathbf{r}}')} \tag{11.84}$$

and

$$\frac{\delta h(\vec{\mathbf{r}})}{\delta h(\vec{\mathbf{r}}')} = \delta(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.85}$$

we have

$$\delta(\vec{\mathbf{r}} - \vec{\mathbf{r}}') = \left[-k\boldsymbol{\nabla}^2 + 2at + 6bm_0^2(\vec{\mathbf{r}})\right]\chi_T(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.86}$$

where we have assumed translational invariance.

On the other hand, from fluctuation dissipation relation

$$G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}') = k_B T \chi_T(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.87}$$

we obtain

$$\beta\left[-k\boldsymbol{\nabla}^2 + 2at + 6bm_0^2\right]G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}') = \delta(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.88}$$

The equation (11.88) is also known as fundamental equation and $G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}')$ the fundamental solution or green function of the differential operator

$$D \equiv -k\boldsymbol{\nabla}^2 + 2at + 6bm^2 \tag{11.89}$$

We now look for the solutions of (11.88). We consider two cases, $T > T_c$ and $T < T_c$ and let $h \to 0$.

- Case $T > T_c$ $(t > 0)$: the mean field solution is $m_0(\vec{\mathbf{r}}) = m_0 = 0$. Hence,

$$\left(-k\boldsymbol{\nabla}^2 + 2at\right)G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}') = k_B T \delta(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.90}$$

  By calling

$$\xi_>(t) = \left(\frac{k}{2at}\right)^{1/2} \tag{11.91}$$

  we have

$$\left(-\boldsymbol{\nabla}^2 + \xi_>^{-2}(t)\right)G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}') = \frac{k_B T}{k}\delta(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.92}$$

- Case $T < T_c$ ($t < 0$). Let us conside the omogeneous solution

$$m_0 = \pm\left(-\frac{at}{b}\right)^{1/2} \tag{11.93}$$

This gives

$$\left(-\boldsymbol{\nabla}^2 + \xi_<^{-2}(t)\right)G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}') = \frac{k_B T}{k}\delta(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.94}$$

where

$$\xi_<(t) = \left(-\frac{k}{4at}\right)^{1/2} \tag{11.95}$$

Remark. In both cases $\xi \sim t^{-1/2}$, we have $\nu = \nu' = \frac{1}{2}$, that is the mean field value of ν!

Remark. Since $\nu = 1/2$, the upper critical dimension of a critical point belonging to the Ising universality class is

$$D_c = \frac{2}{\nu} = 4 \tag{11.96}$$

For both cases, the equation for $G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}')$ has the following form

$$\left(-\boldsymbol{\nabla}^2 + \xi^{-2}\right)G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}') = \frac{k_B T}{k}\delta(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.97}$$

This equation can be solved in at least two ways:

- Fourier transform.
- Spherical coordinates.

Lecture 18. *Friday 13th December, 2019. Compiled: Thursday 26th December, 2019.*

11.5.1 Solution by Fourier transform

Let us try to do the Fourier transform and define

$$\vec{\mathbf{x}} \equiv \vec{\mathbf{r}} - \vec{\mathbf{r}}' \tag{11.98}$$

Taking the Fourier transform of

$$\left(-\boldsymbol{\nabla}^2 + \xi^{-2}(t)\right)G_c(\vec{\mathbf{r}} - \vec{\mathbf{r}}') = \frac{k_B T}{k}\delta(\vec{\mathbf{r}} - \vec{\mathbf{r}}') \tag{11.99}$$

and calling $\widetilde{G}(q)$ the Fourier transform of the function G, one gets

$$\widetilde{G}(q) = \int_{-\infty}^{+\infty} \mathrm{d}|\vec{\mathbf{x}}|\, G_c(|\vec{\mathbf{x}}|)e^{-iq|\vec{\mathbf{x}}|} \quad \Rightarrow \widetilde{G}(q) = \frac{k_B T}{k}\frac{1}{q^2 + \xi^{-2}} \tag{11.100}$$

Remark. At $T = T_c$, since $\xi \to \infty$ we have and $\widetilde{G}(q) \simeq \frac{1}{q^2}$ and performing the inverse Fourier transform one gets

$$G_c(|\vec{\mathbf{x}}|) = |\vec{\mathbf{x}}|^{2-D} \tag{11.101}$$

On the other hand, at $T = T_c$ we defined

$$G(r) \sim |\vec{\mathbf{x}}|^{2-D-\eta} \tag{11.102}$$

hence, we have $\eta = 0$.

Let us now obtain the full expression of $G(r)$ by computing the inverse Fourier transform

$$G(\vec{\mathbf{x}}) = \int \mathrm{d}^D\vec{\mathbf{q}}\,\frac{1}{(2\pi)^D}\frac{1}{q^2+\xi^{-2}}e^{i\vec{\mathbf{q}}\cdot\vec{\mathbf{x}}} \tag{11.103}$$

Let us do it for $D = 3$:

$$\begin{aligned} G(\vec{\mathbf{x}}) &= \int \mathrm{d}^3\vec{\mathbf{q}}\,\frac{1}{(2\pi)^3}\frac{1}{q^2+\xi^{-2}}e^{i\vec{\mathbf{q}}\cdot\vec{\mathbf{x}}} \overset{\text{spherical coordinates}}{=} \\ &= \frac{4\pi}{(2\pi)^3}\int_0^\infty \mathrm{d}q\,\frac{q^2}{q^2+\xi^{-2}}\int_{-1}^{+1}\mathrm{d}(\cos\theta)\,e^{iq|\vec{\mathbf{x}}|\cos\theta} \\ &\overset{z\equiv\cos(\theta)}{=} \frac{2}{(2\pi)^2}\int_0^\infty \mathrm{d}q\,\frac{q^2}{q^2+\xi^{-2}}\left[\frac{1}{iq|\vec{\mathbf{x}}|}e^{iq|\vec{\mathbf{x}}|z}\right]_{-1}^{1} \\ &= \frac{4}{(2\pi)^2|\vec{\mathbf{x}}|}\int_0^\infty \frac{q\sin(q|\vec{\mathbf{x}}|)}{q^2+\xi^{-2}}\,\mathrm{d}q \end{aligned} \tag{11.104}$$

Let us now consider the integral

$$I = \int_0^\infty \frac{q\sin(q|\vec{\mathbf{x}}|)}{q^2+\xi^{-2}}\,\mathrm{d}q = \frac{1}{2}\int_{-\infty}^{+\infty}\frac{q\sin(q|\vec{\mathbf{x}}|)}{q^2+\xi^{-2}}\,\mathrm{d}q \tag{11.105}$$

that can be computed in the complex plane, i.e. as

$$I = \frac{1}{2}\,\mathrm{Im}\oint \frac{ze^{iz|\vec{\mathbf{x}}|}}{(z^2+\xi^{-2})}\,\mathrm{d}z \tag{11.106}$$

The integration contour is a semicirle in the positive half plane. There are two poles $z_P = \pm i\xi^{-1}$, but only one is enclosed by the semicirle (Figure 11.3).

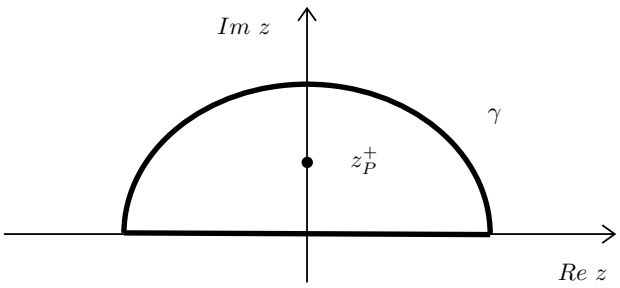


Figure 11.3: Description.

Using the Theorem of residues one has

$$\begin{aligned} I &= \frac{1}{2}\,\mathrm{Im}\oint_\gamma \frac{ze^{iz|\vec{\mathbf{x}}|}}{(z^2+\xi^{-2})}\,\mathrm{d}z = \frac{1}{2}\oint_\gamma \frac{ze^{iz|\vec{\mathbf{x}}|}}{(z+i\xi^{-1})(z-i\xi^{-1})}\,\mathrm{d}z \\ &\overset{th.\,of\,residues}{=} \frac{1}{2}\,\mathrm{Im}\left[2\pi i\mathrm{Res}(i\xi^{-1})\right] \end{aligned} \tag{11.107}$$

and since

$$\mathrm{Res}(i\xi^{-1}) = \frac{i\xi^{-1}e^{-\xi^{-1}|\vec{\mathbf{x}}|}}{2i\xi^{-1}} = \frac{e^{-|\vec{\mathbf{x}}|/\xi}}{2} \tag{11.108}$$

$$\Rightarrow I = \frac{1}{2}\,\mathrm{Im}\left[2\pi i\mathrm{Res}(i\xi^{-1})\right] = \frac{\pi}{2}e^{-|\vec{\mathbf{x}}|/\xi} \tag{11.109}$$

Hence, at the end

$$\Rightarrow G(|\vec{\mathbf{x}}|) = \frac{1}{2\pi}\frac{e^{-\frac{|\vec{\mathbf{x}}|}{\xi}}}{|\vec{\mathbf{x}}|} \tag{11.110}$$

where it is confirmed the exponential behaviour for $t \neq 0$ and $\eta = 0$.

One can also solve the equation for $G(\vec{\mathbf{r}} - \vec{\mathbf{r}}')$ by using the spherical coordinates and use the Bessel functions.

11.6 Including fluctuations at the Gaussian level (non interacting fields)

Can we reach the simple level of fluctuation? The simple level is the one that follow gaussian distribution. Let us introduce fluctuations at the Gaussian level.

Let us consider $h = 0$ and let $m_0(\vec{\mathbf{r}}) = m_0$ be the solution of the saddle point approximation. Let us expand the general expression

$$\beta\mathcal{H}_{eff}[m] = \int \mathrm{d}^D\vec{\mathbf{r}} \left(atm^2 + \frac{b}{2}m^4 + \frac{k}{2}(\boldsymbol{\nabla} m)^2 \right) \tag{11.111}$$

by using

$$m(\vec{\mathbf{r}}) = m_0 + \delta m(\vec{\mathbf{r}}) \tag{11.112}$$

We are assuming that the fluctuations $\delta m(\vec{\mathbf{r}})$ are small:

$$(\boldsymbol{\nabla} m)^2 = (\boldsymbol{\nabla}(m_0 + \delta m))^2 = (\boldsymbol{\nabla}(\delta m))^2 \tag{11.113a}$$

$$m^2 = m_0^2 + 2m_0\delta m + (\delta m)^2 \tag{11.113b}$$

$$m^4 = m_0^4 + 4m_0^3\delta m + 6m_0^2(\delta m)^2 + 4m_0\delta m^3 + (\delta m)^4 \tag{11.113c}$$

Hence,

$$\beta\mathcal{H}_{eff} = V \underbrace{\left(atm_0^2 + \frac{b}{2}m_0^4 \right)}_{A_0} + \int \mathrm{d}^D\vec{\mathbf{r}} \left(\frac{k}{2}(\boldsymbol{\nabla} m)^2 + \left(at + 3bm_0^2\right)\delta m^2 + 2bm_0\delta m^3 + \frac{b}{2}\delta m^4 \right) \tag{11.114}$$

Remark. The term proportional to δm, $\left(2atm_0 + \frac{b}{2}4m_0^3\right)$ is zero since m_0 is the solution of the extremal condition

$$\left.\frac{\delta\mathcal{H}_{eff}}{\delta m}\right|_{m=m_0} = 0 \tag{11.115}$$

For simplicity let us first consider $T > T_c$, we know that $m_0 = 0$ and $m(\vec{\mathbf{r}}) = m_0 + \delta m(\vec{\mathbf{r}}) = \delta m(\vec{\mathbf{r}})$. We have $A_0 = 0$, $3bm_0^2\delta m^2 = 0$ and $2bm_0\delta m^3 = 0$, that implies

$$\beta\mathcal{H}_{eff}^{T>T_c}(\delta m) = \int \mathrm{d}^D\vec{\mathbf{r}} \left(\frac{k}{2}(\boldsymbol{\nabla}\delta m)^2 + at(\delta m)^2 + \cancel{\frac{b}{2}(\delta m)^4} \right) \tag{11.116}$$

Remark. It is important to understand that these are fluctuations with respect to the solution.

The Gaussian approximation consists in neglecting the quartic term $(\delta m)^4$

$$\Rightarrow \beta\mathcal{H}_{eff}^{T>T_c}(\delta m) \simeq \int \mathrm{d}^D\vec{\mathbf{r}} \left(\frac{k}{2}(\boldsymbol{\nabla}\delta m)^2 + at(\delta m)^2 \right) \tag{11.117}$$

Remark. We cannot do again the saddle point, otherwise we do not get too much information. We consider gaussian fluctuations: fluctuations that follow gaussian distribution. Therefore, the term in $(\delta m)^4$ is cancelled. What it is the difference in the exponent respect the mean field? Then we will to the same for $T < T_c$, the story is the same. (lesson)

11.7 Ginzburg-Landau in the Gaussian approximation

Let us solve

$$Z_{GL}^{G} = \int \mathrm{D}[\delta m] e^{-\int \mathrm{d}^D r \left(\frac{k}{2}(\boldsymbol{\nabla}\delta m)^2 + at(\delta m)^2\right)} \tag{11.118}$$

in the Fourier space. Consider a system in a box of volume $V = L^D$ (periodic boundary conditions):

$$\delta m(\vec{\mathbf{r}}) = \frac{1}{V}\sum_{\vec{\mathbf{k}}} e^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}}(\delta m_{\vec{\mathbf{k}}}) \tag{11.119}$$

where

$$\delta m_{\vec{\mathbf{k}}} = \int_V \delta m(\vec{\mathbf{r}}) e^{-i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}}\,\mathrm{d}^D\vec{\mathbf{r}} \tag{11.120}$$

where $\vec{\mathbf{k}} = k_1, \dots, k_D = \frac{2\pi\bar{n}}{L}$ with $k_\alpha = \frac{2\pi}{L}n_\alpha$ ($n_\alpha = 0, \pm 1, \dots$).

Remark. We have $\delta m_{\vec{\mathbf{k}}} \in \mathbb{C}$ but $\delta m(\vec{\mathbf{r}}) \in \mathbb{R}$, hence

$$\delta m_{\vec{\mathbf{k}}} = -\delta m^*_{-\vec{\mathbf{k}}} \tag{11.121}$$

11.7.1 Useful relations

Sometimes it is useful to convert the sum over $\vec{\mathbf{k}}$ by an integral by using the density of states in the $\vec{\mathbf{k}}$ space that is $\frac{V}{(2\pi)^D}$,

$$\boxed{\sum_{\vec{\mathbf{k}}} \to \frac{V}{(2\pi)^D}\int_{\mathbb{R}^D} \mathrm{d}\vec{\mathbf{k}}} \tag{11.122}$$

Hence,

$$\begin{aligned}\frac{1}{V}\sum_{\vec{\mathbf{k}}} e^{i\vec{\mathbf{k}}(\vec{\mathbf{r}}-\vec{\mathbf{r}}')} &\to \frac{1}{V}\frac{V}{(2\pi)^2}\int_{\mathbb{R}^D} \mathrm{d}^D\vec{\mathbf{k}}\, e^{i\vec{\mathbf{k}}(\vec{\mathbf{r}}-\vec{\mathbf{r}}')} \\ &= \frac{1}{(2\pi)^D}\int_{\mathbb{R}^D} e^{i\vec{\mathbf{k}}(\vec{\mathbf{r}}-\vec{\mathbf{r}}')}\,\mathrm{d}^D\vec{\mathbf{k}} = \delta(\vec{\mathbf{r}}-\vec{\mathbf{r}}')\end{aligned} \tag{11.123}$$

so

$$\boxed{\frac{1}{V}\sum_{\vec{\mathbf{k}}} e^{i\vec{\mathbf{k}}(\vec{\mathbf{r}}-\vec{\mathbf{r}}')} \to \delta(\vec{\mathbf{r}}-\vec{\mathbf{r}}')} \tag{11.124}$$

Inserting

$$m(\vec{\mathbf{r}}) = \frac{1}{V}\sum_{\vec{\mathbf{k}}} e^{i\vec{\mathbf{k}}'\vec{\mathbf{r}}} m_{\vec{\mathbf{k}}'} \tag{11.125}$$

In the expression

$$m_{\vec{\mathbf{k}}} = \int_V \mathrm{d}^D\vec{\mathbf{r}}\, e^{-i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}} m(\vec{\mathbf{r}}) \tag{11.126}$$

one gets

$$\boxed{\frac{1}{V}\int \mathrm{d}^D\vec{\mathbf{r}}\, e^{i(\vec{\mathbf{k}}-\vec{\mathbf{k}}')\cdot\vec{\mathbf{r}}} = \delta_{\vec{\mathbf{k}}\vec{\mathbf{k}}'}} \tag{11.127}$$

Finally, since

$$\int \mathrm{d}^D\vec{\mathbf{r}}\, e^{i(\vec{\mathbf{k}}-\vec{\mathbf{k}}')\cdot\vec{\mathbf{r}}} \overset{V\to\infty}{\longrightarrow} (2\pi)^D\delta(\vec{\mathbf{k}}-\vec{\mathbf{k}}') \tag{11.128}$$

From (11.127) one gets,

$$\boxed{V\delta_{\vec{\mathbf{k}}\vec{\mathbf{k}}'} \overset{V\to\infty}{\longrightarrow} (2\pi)^D\delta(\vec{\mathbf{k}}-\vec{\mathbf{k}}')} \tag{11.129}$$

Remark. Since the minimal spatial length of the system is a,

$$a \Rightarrow \left|\vec{\mathbf{k}}\right| \leq \frac{\pi}{a} = \Lambda \tag{11.130}$$

that is the ultraviolet cut-off.

11.8 Gaussian Hamiltonian in Fourier space

For simplicity, let us change notation

$$\delta m(\vec{\mathbf{r}}) \leftrightarrow \varphi(\vec{\mathbf{r}}), \quad k \leftrightarrow c \tag{11.131}$$

obtaining

$$\beta\mathcal{H}_{eff}^{G,>}[\varphi] = \int \mathrm{d}^D\vec{\mathbf{r}}\left[\frac{c}{2}(\boldsymbol{\nabla}\varphi)^2 + at\varphi^2\right] \tag{11.132}$$

where

$$\varphi(\vec{\mathbf{r}}) = \frac{1}{V}\sum_{\vec{\mathbf{k}}} e^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}}\varphi_{\vec{\mathbf{k}}} \tag{11.133}$$

Consider the term of expression (11.132) separately:

- Term $at\varphi^2$.

$$\int \mathrm{d}^D\vec{\mathbf{r}}\, at\varphi^2(\vec{\mathbf{r}}) = \frac{at}{V^2}\sum_{\vec{\mathbf{k}},\vec{\mathbf{k}}'}\int_{\mathbb{R}^D} e^{i(\vec{\mathbf{k}}+\vec{\mathbf{k}}')\cdot\vec{\mathbf{r}}}\varphi_{\vec{\mathbf{k}}}\varphi_{\vec{\mathbf{k}}'}\,\mathrm{d}^D\vec{\mathbf{r}} = \frac{at}{V^2}\sum_{\vec{\mathbf{k}}\vec{\mathbf{k}}'}\varphi_{\vec{\mathbf{k}}}\varphi_{\vec{\mathbf{k}}'}\delta(\vec{\mathbf{k}}+\vec{\mathbf{k}}')(2\pi)^D \tag{11.134}$$

  On the other hand,

$$\frac{(2\pi)^D}{V}\delta(\vec{\mathbf{k}}+\vec{\mathbf{k}}') \overset{V\gg 1}{\longrightarrow} \delta_{\vec{\mathbf{k}},-\vec{\mathbf{k}}'} \tag{11.135}$$

$$\Rightarrow \int \mathrm{d}^D\vec{\mathbf{r}}\, at\varphi^2(\vec{\mathbf{r}}) \overset{V\gg 1}{\longrightarrow} \frac{1}{2V}\sum_{\vec{\mathbf{k}}} 2at\varphi_{\vec{\mathbf{k}}}\varphi_{\vec{\mathbf{k}}'} \tag{11.136}$$

- Term $\int \mathrm{d}^D\vec{\mathbf{r}}\,\frac{c}{2}(\boldsymbol{\nabla}\varphi)^2$.

$$\begin{aligned}\int \mathrm{d}^D r\,\frac{c}{2}(\boldsymbol{\nabla}\varphi)^2 &= \frac{c}{2}\frac{1}{V^2}\int \mathrm{d}^D\vec{\mathbf{r}}\left(\boldsymbol{\nabla}\sum_{\vec{\mathbf{k}}} e^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}}\varphi_{\vec{\mathbf{k}}}\right)\left(\boldsymbol{\nabla}\sum_{\vec{\mathbf{k}}'} e^{i\vec{\mathbf{k}}'\cdot\vec{\mathbf{r}}}\varphi_{\vec{\mathbf{k}}'}\right)\\ &= \frac{c}{2V^2}\sum_{\vec{\mathbf{k}}\vec{\mathbf{k}}'}\left(-\vec{\mathbf{k}}\cdot\vec{\mathbf{k}}'\right)\varphi_{\vec{\mathbf{k}}}\varphi_{\vec{\mathbf{k}}'}\underbrace{\int \mathrm{d}^D\vec{\mathbf{r}}\, e^{i(\vec{\mathbf{k}}+\vec{\mathbf{k}}')\cdot\vec{\mathbf{r}}}}_{(2\pi)^2\delta(\vec{\mathbf{k}}+\vec{\mathbf{k}}')\to V\delta_{\vec{\mathbf{k}},-\vec{\mathbf{k}}'}}\\ &= \frac{c}{2V}\sum_{\vec{\mathbf{k}}}\left|\vec{\mathbf{k}}\right|^2\varphi_{\vec{\mathbf{k}}}\varphi_{-\vec{\mathbf{k}}'}\end{aligned} \tag{11.137}$$

In summary, the exponent is:

$$\boxed{\beta\mathcal{H}_{eff}^{G,>} \to \frac{1}{2V}\sum_{\vec{\mathbf{k}}}\left(2at + ck^2\right)\varphi_{\vec{\mathbf{k}}}\varphi_{-\vec{\mathbf{k}}'}} \tag{11.138}$$

this is the Gaussian model in Fourier space.

Now, the question is: how can we write the measure $\int \mathrm{D}[\varphi]$ in Fourier space?

$$\varphi(\vec{\mathbf{r}}) = \frac{1}{V}\sum_{\vec{\mathbf{k}}}\varphi_{\vec{\mathbf{k}}}e^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}}, \quad \varphi_{\vec{\mathbf{k}}} \in \mathbb{C} \tag{11.139}$$

$$\int \mathrm{D}[\varphi(\vec{\mathbf{r}})] \to \int_{-\infty}^{+\infty} \prod_{|\vec{\mathbf{k}}|<\Lambda} \mathrm{d}(\mathrm{Re}\{\varphi_{\vec{\mathbf{k}}}\})\, \mathrm{d}(\mathrm{Im}\{\varphi_{\vec{\mathbf{k}}}\}) \tag{11.140}$$

with $\varphi(\vec{\mathbf{r}}) \in \mathbb{R}$, that implies $\varphi^*_{\vec{\mathbf{k}}} = \varphi_{-\vec{\mathbf{k}}}$, hence

$$\begin{cases} \mathrm{Re}\{\varphi_{\vec{\mathbf{k}}}\} = \mathrm{Re}\left\{\varphi_{-\vec{\mathbf{k}}}\right\} \\ \mathrm{Im}\{\varphi_{\vec{\mathbf{k}}}\} = -\,\mathrm{Im}\left\{\varphi_{-\vec{\mathbf{k}}}\right\} \end{cases} \tag{11.141}$$

Therefore, if we considered

$$\prod_{|\vec{\mathbf{k}}|<\Lambda} \mathrm{d}\,\mathrm{Re}\{\varphi_{\vec{\mathbf{k}}}\}\, \mathrm{d}\,\mathrm{Im}\{\varphi_{\vec{\mathbf{k}}}\} \tag{11.142}$$

we would count twice. We need to consider one restriction. For example we can assume $k_z > 0$ (hupper half of the $\vec{\mathbf{k}}$ space), hence

$$\mathrm{Tr} \equiv \int_{-\infty}^{+\infty} \prod_{\substack{|\vec{\mathbf{k}}|<\Lambda \\ k_z>0}} \mathrm{d}\,\mathrm{Re}\{\varphi_{\vec{\mathbf{k}}}\}\, \mathrm{d}\,\mathrm{Im}\{\varphi_{\vec{\mathbf{k}}}\} \tag{11.143}$$

or

$$\mathrm{Tr} \equiv \frac{1}{2} \int_{-\infty}^{+\infty} \prod_{|\vec{\mathbf{k}}|<\Lambda} \mathrm{d}\,\mathrm{Re}\{\varphi_{\vec{\mathbf{k}}}\}\, \mathrm{d}\,\mathrm{Im}\{\varphi_{\vec{\mathbf{k}}}\} \tag{11.144}$$

For sake of brevity

$$\mathrm{Tr} \equiv \int_{-\infty}^{+\infty} {\prod_{\vec{\mathbf{k}}}}' \mathrm{d}\varphi_{\vec{\mathbf{k}}} \tag{11.145}$$

where

$$\prod_{\vec{\mathbf{k}}} \mathrm{d}\varphi_{\vec{\mathbf{k}}} \equiv \prod_{\substack{|\vec{\mathbf{k}}|<\Lambda \\ k_z>0}} \mathrm{d}\,\mathrm{Re}\{\varphi_{\vec{\mathbf{k}}}\}\, \mathrm{d}\,\mathrm{Im}\{\varphi_{\vec{\mathbf{k}}}\} \tag{11.146}$$

In summary, we have

$$\widetilde{Z}^{G,>}_{GC} = \int_{-\infty}^{+\infty} \left({\prod_{\vec{\mathbf{k}}}}' \mathrm{d}\varphi_{\vec{\mathbf{k}}} \right) e^{-\beta \widetilde{\mathcal{H}}^G_{eff}[\varphi_{\vec{\mathbf{k}}}]} \tag{11.147}$$

where

$$-\,\beta \widetilde{\mathcal{H}}^G_{eff}[\varphi_{\vec{\mathbf{k}}}] = -\frac{1}{2V} \sum_{\vec{\mathbf{k}}} \left(2at + c\left|\vec{\mathbf{k}}\right|^2 \right) \left|\varphi_{\vec{\mathbf{k}}}\right|^2 \tag{11.148}$$

11.8.1 Free energy in Gaussian approximation

Let us consider the case $T > T_c$ ($m_0 = 0$) and $h = 0$

$$\widetilde{Z}^{T>T_c}_G = {\prod_{\vec{\mathbf{k}}}}' \int_{-\infty}^{+\infty} \mathrm{d}\,\mathrm{Re}\{\varphi_{\vec{\mathbf{k}}}\}\, \mathrm{d}\,\mathrm{Im}\{\varphi_{\vec{\mathbf{k}}}\}\, e^{-\frac{1}{2V}\sum_{\vec{\mathbf{k}}} \left(2at + c|\vec{\mathbf{k}}|^2\right) |\varphi_{\vec{\mathbf{k}}}|^2} \tag{11.149}$$

Change variables

$$\begin{cases} x \equiv \mathrm{Re}\,\varphi_{\vec{\mathbf{k}}} \\ y \equiv \mathrm{Im}\,\varphi_{\vec{\mathbf{k}}} \end{cases} \tag{11.150}$$

and since

$$\int_{-\infty}^{+\infty} \mathrm{d}x\,\mathrm{d}y\, e^{-A(x^2+y^2)} = \frac{\pi}{A} \tag{11.151}$$

with $A \equiv \frac{1}{2V}$.

$$e^{-\beta \widetilde{F}^{>}_{GL}} = \left(\prod_{\substack{\vec{\mathbf{k}} \\ |\vec{\mathbf{k}}|<\Lambda \\ k_z>0}}' \frac{2\pi V}{2at + c\left|\vec{\mathbf{k}}\right|^2} \right) = \exp\left[\frac{1}{2} \sum_{|\vec{\mathbf{k}}|<\Lambda} \log\left(\frac{2\pi V}{2at + c\left|\vec{\mathbf{k}}\right|^2} \right) \right] \tag{11.152}$$

Hence,

$$\widetilde{F}^{G,T>T_c}_{GL} = -\frac{1}{2} k_B T \sum_{|\vec{\mathbf{k}}|<\Lambda} \log\left(\frac{2\pi V}{2at + c\left|\vec{\mathbf{k}}\right|^2} \right) \tag{11.153}$$

Remark. For $T < T_c$ we have $m_0 \neq 0$. In addition, to redefine the quadratic term $(at + 3bm_0^2)$, that for $m_0^2 = -at/b$, becames $-2at$ we have also the term

$$VA_0 = V(atm_0^2 + \frac{b}{2} m_0^4) \tag{11.154}$$

Therefore,

$$\widetilde{F}^{G,T<T_c}_{GL} = \beta A_0 V - \frac{1}{2} k_B T \sum_{|\vec{\mathbf{k}}|<\Lambda} \log\left(\frac{2\pi V}{2at + c\left|\vec{\mathbf{k}}\right|^2} \right) \tag{11.155}$$

11.8.2 Specific heat in the Gaussian approximation

$$c_V^G = -T \frac{\partial^2 F^G/V}{\partial T^2} \overset{todo}{=} \underbrace{\frac{A}{V} \sum_{|\vec{\mathbf{k}}|<\Lambda} \frac{1}{\left(2at + c\left|\vec{\mathbf{k}}\right|^2\right)^2}}_{1^{st}} - \underbrace{\frac{B}{V} \sum_{|\vec{\mathbf{k}}|<\Lambda} \frac{1}{2at + c\left|\vec{\mathbf{k}}\right|^2}}_{2^{st}} \tag{11.156}$$

One can show that

$$1^{st} \propto \begin{cases} \xi^{4-D} \sim t^{-\nu(4-D)} & D < 4 \\ < \infty & D > 4 \end{cases} \tag{11.157}$$

$$2^{nd} \propto \begin{cases} \xi^{2-D} \sim t^{-\nu(2-D)} & D < 2 \\ < \infty & D > 2 \end{cases} \tag{11.158}$$

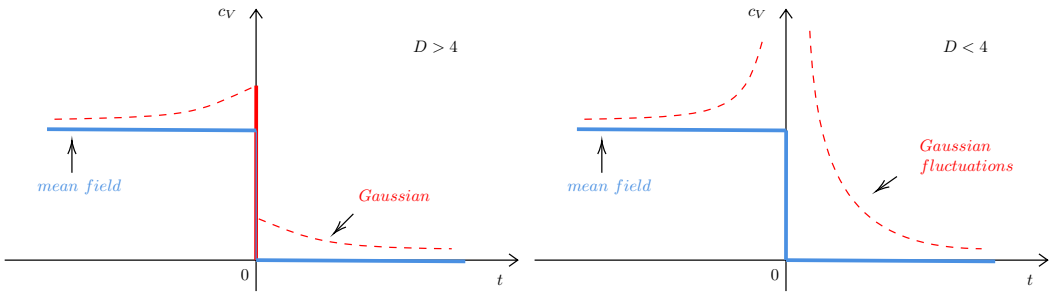
In summary, the behaviour of the specific heat is

$$c_V \sim \begin{cases} t^{-\nu(4-D)} & D < 4 \\ < \infty & D > 4 \end{cases} \tag{11.159}$$

The introduction of the fluctuations even if only at the Gaussian level changes the behaviour of c_V at the transition point.

The introduction of Gaussian fluctuations gives rise to a divergence in $c_V \sim t^{-\alpha}$ with

$$\alpha = \nu(4 - D) \quad \text{for } D < 4 \tag{11.160}$$



(a) Description **(b)** Description

Figure 11.4

11.8.3 How about ν?

We have to compute the 2−point correlation function for $\mathcal{H}^G_{eff}(\varphi)$ that in Fourier space become

$$\langle \varphi_{\vec{\mathbf{k}}} \varphi_{\vec{\mathbf{k}}'} \rangle_G = \frac{\int_{-\infty}^{+\infty} \mathrm{d}\tilde{\varphi}_{\vec{\mathbf{k}}_1} \dots \mathrm{d}\tilde{\varphi}_{\vec{\mathbf{k}}} \, \mathrm{d}\tilde{\varphi}_{\vec{\mathbf{k}}'} \, \tilde{\varphi}_{\vec{\mathbf{k}}} \tilde{\varphi}_{\vec{\mathbf{k}}'} \dots e^{-\beta \mathcal{H}^G_{eff}} \dots}{\int_{-\infty}^{+\infty} \mathrm{d}\tilde{\varphi}_{\vec{\mathbf{k}}_1} \dots \mathrm{d}\tilde{\varphi}_{\vec{\mathbf{k}}} \, \mathrm{d}\tilde{\varphi}_{\vec{\mathbf{k}}'} \dots \mathrm{d}\tilde{\varphi}_{\vec{\mathbf{k}}_n} \dots} \tag{11.161}$$

where

$$\mathcal{H}^G_{eff} = A_0 V + \frac{1}{2V} \sum_{\vec{\mathbf{k}}} \left(2at + c\vec{\mathbf{k}}^2 \right) \left| \varphi_{\vec{\mathbf{k}}} \right|^2 \tag{11.162}$$

It is possible to show:

$$\langle \varphi_{\vec{\mathbf{k}}} \varphi_{\vec{\mathbf{k}}'} \rangle_G = \delta_{\vec{\mathbf{k}}, -\vec{\mathbf{k}}'} \frac{V}{2at + ck^2} \tag{11.163}$$

that, in the limit $V \to \infty$ becomes

$$\langle \varphi_{\vec{\mathbf{k}}} \varphi_{\vec{\mathbf{k}}'} \rangle_G \overset{V \to \infty}{\longrightarrow} \frac{(2\pi)^D}{2at + c\vec{\mathbf{k}}^2} \delta^D(\vec{\mathbf{k}} + \vec{\mathbf{k}}') \tag{11.164}$$

By antitrasforming

$$\begin{aligned} \langle \varphi(\vec{\mathbf{r}}) \varphi(\vec{\mathbf{r}}') \rangle_G &= \frac{1}{V^2} \sum_{\vec{\mathbf{k}}, \vec{\mathbf{k}}'} e^{i(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} + \vec{\mathbf{k}}' \cdot \vec{\mathbf{r}}')} \langle \varphi_{\vec{\mathbf{k}}} \varphi_{\vec{\mathbf{k}}'} \rangle_G \\ &= \frac{1}{V^2} \sum_{\vec{\mathbf{k}}, \vec{\mathbf{k}}'} e^{i(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} + \vec{\mathbf{k}}' \cdot \vec{\mathbf{r}}')} \delta_{\vec{\mathbf{k}}, -\vec{\mathbf{k}}'} \frac{V}{2at + ck^2} \\ &= \frac{1}{V} \sum_{\vec{\mathbf{k}}} \frac{e^{i\vec{\mathbf{k}} \cdot (\vec{\mathbf{r}} - \vec{\mathbf{r}}')}}{2at + ck^2} \end{aligned} \tag{11.165}$$

Using

$$\xi_>(t) = \left(\frac{c}{2at} \right)^{1/2} \tag{11.166}$$

we get

$$\langle \varphi(\vec{\mathbf{r}}) \varphi(\vec{\mathbf{r}}') \rangle_G = \frac{1}{V_c} \sum_{\vec{\mathbf{k}}} \frac{e^{i\vec{\mathbf{k}} \cdot (\vec{\mathbf{r}} - \vec{\mathbf{r}}')}}{k^2 + \xi_>^{-2}} = \frac{1}{V} \sum_{\vec{\mathbf{k}}} e^{i\vec{\mathbf{k}} \cdot \vec{\mathbf{x}}} \hat{G}(\vec{\mathbf{k}}) \tag{11.167}$$

where

$$\hat{G}(\vec{\mathbf{k}}) = \frac{1}{c\left(k^2 + \xi_>^{-2}\right)} \tag{11.168}$$

i.e. is of the same form found in mean field.

$$\Rightarrow \begin{cases} \nu_G = \frac{1}{2} \\ \eta_G = 0 \end{cases} \tag{11.169}$$

There are no changes with Gaussian fluctuation! Interactions between $\varphi_{\vec{\mathbf{k}}}$ are needed!

Chapter 12

Widom's scaling theory. Block-spin Kadanoff's transformation

Lecture 19.
Wednesday 18th December, 2019.
Compiled: Thursday 26th December, 2019.

12.1 Static scaling hypothesis (Widom)

It is used whenever you have a collective behaviour. As $T \to T_c^{\pm}$ we know that $\xi \to L$. In this limit all the microscopic and intermediate scales are irrelevant. The length scale of the problem are a, L, ξ, but ξ is the only relevant length scale in the problem.

One should expect that, for $t \sim 0$, the free energy (better, its critical term) is invariant in form by a change of scale. This hypothesis is also suggested by experimental data such as the ones shown by Guggenheim for the gas phase diagrams and the ones shown for ferromagnetic materials at different temperatures.

Which are the experimental data which gives us this ideas? What you can see from experiment (Figure 12.1) is that data from different temperatures, if scaled properly, collapse into two (one for $t < 0$ and one for $t > 0$) unique curves.

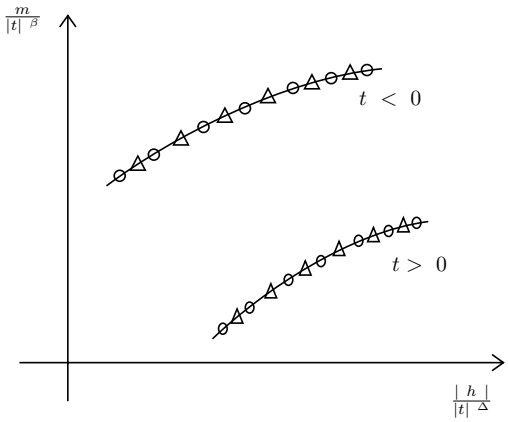


Figure 12.1: Description.

The Widom's static scaling theory is introduced also to explain the collapse of data shown above.

In order to properly define the scaling hypothesis, we should rely on the mathematical concept of homogeneous functions.

12.1.1 Homogeneous functions

Consider a single variable r.

Definition 8 (Homogeneous function)**.** $f(r)$ is homogeneous in r if $\forall\lambda \in \mathbb{R}$,

$$f(\lambda r) = g(\lambda)f(r) \tag{12.1}$$

where $g(\lambda)$ is, for the moment, arbitrary.

Example 27. Parabola:

$$f(r) = Br^2 \tag{12.2}$$

$$f(\lambda r) = B(\lambda r)^2 = \lambda^2 f(r) \quad \Rightarrow g(\lambda) = \lambda^2 \tag{12.3}$$

Remark. An interesting property of an homogeneous function is that, if $f(r_0)$ is known and $g(\lambda)$ is known, hence $f(r)$ is known $\forall r \in \mathbb{R}$. Indeed, we can always write $r = \lambda r_0$,

$$f(r) = f(\lambda r_0) = g(\lambda)f(r_0) \tag{12.4}$$

Theorem 12.1.1. *The function $g(\lambda)$ is not arbitrary but it must be of the form*

$$g(\lambda) = \lambda^p \tag{12.5}$$

where p is the degree of the homogeneity of the function.

Proof. Change of scale with respect to μ and λ.

$$f(\lambda(\mu r)) = g(\lambda)f(\mu r) = g(\lambda)g(\mu)f(r) \tag{12.6}$$

On the other hand,

$$f[(\lambda\mu)r] = g(\lambda\mu)f(r) \tag{12.7}$$

Hence,

$$g(\lambda\mu) = g(\lambda)g(\mu) \tag{12.8}$$

Let us assume that $g(\lambda)$ is differentiable (actually $g(\lambda)$ continuous is sufficient but proof more complicated)

$$\frac{\partial}{\partial\mu}[g(\lambda\mu)] = \frac{\partial}{\partial\mu}[g(\lambda)g(\lambda)] \tag{12.9}$$

$$\Rightarrow \lambda g'(\lambda\mu) = g(\lambda)g'(\mu) \tag{12.10}$$

Let $\mu = 1$ and let $g'(\mu = 1) = p$, equation (12.10) implies

$$\frac{g'(\lambda)}{g(\lambda)} = \frac{p}{\lambda} \tag{12.11}$$

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}(\ln g(\lambda)) = \frac{p}{\lambda} \quad \Rightarrow \ln g(\lambda) = p\ln\lambda + c \tag{12.12}$$

or

$$g(\lambda) = e^c\lambda^p \quad \Rightarrow g'(\lambda) = pe^c\lambda^{p-1} \tag{12.13}$$

and since $g'(1) = p$, we have $p = pe^c$ if and only if $c = 0$ which implies

$$g(\lambda) = \lambda^p \tag{12.14}$$

□

12.2 Generalized homogeneous functions

We can make it for any variable, not only for a single one. We are discussing $f(x, y)$, that is a generalized homogeneous function if as a most general form as

$$f(\lambda^a x, \lambda^b y) = \lambda f(x, y) \tag{12.15}$$

Indeed, if we consider instead

$$f(\lambda^a x, \lambda^b y) = \lambda^p f(x, y) \tag{12.16}$$

we can always choose $\lambda^p \equiv s$ such that

$$f(s^{a/p} x, s^{b/p} y) = s f(x, y) \tag{12.17}$$

and choosing $a' = a/p$ and $b' = b/p$ we are back to (12.15).

Remark. Since λ is arbitrary $\Rightarrow \lambda = y^{-1/b}$, we get

$$f(x, y) = y^{1/b} f\left(\frac{x}{y^{a/b}}, 1\right) \tag{12.18}$$

f depends on x and y only through the ratio $\frac{x}{y^{a/b}}$! Similarly, for x, one can choose $\lambda = x^{-1/a}$.

Example 28. Examples of non-homogeneous functions are

$$f(x) = e^{-x} \tag{12.19}$$

$$f(x) = \log x \tag{12.20}$$

12.3 Widom's theory of static scaling

Close to the critical point, the singular part of the free energy does not change form by a change of the scale.

More precisely, given $t \equiv \frac{T-T_c}{T_c}$ and $h = \frac{H-H_c}{H_c}$ and the free energy density

$$f(T, H) = f_{ana}(T, H) + f_{sing}(t, h) \tag{12.21}$$

where f_{ana} is an analytic term and f_{sing} diverges, has a singularity. Given that the scaling hypothesis is

$$f_s(\lambda^{p_1} t, \lambda^{p_2} h) = \lambda f_s(t, h), \quad \forall \lambda \in \mathbb{R} \tag{12.22}$$

where p_1 and p_2 are the degrees of the homogeneity of the singular part of the free energy.

Remark. The exponents p_1 and p_2 are not specified by the hypothesis but they will depend on the critical exponents.

Remark. Since f_s is a generalized homogeneous function, it is always possible to choose λ to eliminate one dependence.

For example, one can choose $\lambda = h^{-1/p_2}$ to obtain

$$f_s(t, h) = h^{1/p_2} f_s(h^{-p_1/p_2} t, 1) \tag{12.23}$$

where

$$\Delta \equiv \frac{p_1}{p_2} \tag{12.24}$$

is called the *gap exponent*.

Now, let us see how this simple hypothesis allow us, by simple differential calculus, to obtain relations between the thermodynamic critical exponents.

12.3.1 Scaling of the magnetization

Let us start from the scaling hypothesis

$$f_s(\lambda^{p_1}t, \lambda^{p_2}h) = \lambda f_s(t, h) \tag{12.25}$$

Since M is the 1st derivative of f with respect to H, let us derive (12.25) with respect to h

$$\lambda^{p_2}\frac{\partial f_s}{\partial h}(\lambda^{p_1}t, \lambda^{p_2}h) = \lambda\frac{\partial f_s}{\partial h} \tag{12.26}$$

$$\Rightarrow \lambda^{p_2}M_s(\lambda^{p_1}t, \lambda^{p_2}h) = \lambda M_s(t, h) \tag{12.27}$$

On the other hand, we know that, for $h = 0$ and $t \to 0^-$, $M_s(t) \sim (-t)^\beta$. Putting $h = 0$ in (12.27)

$$M_s(t, 0) = \lambda^{p_2-1}M_s(\lambda^{p_1}t, 0) \tag{12.28}$$

Since λ is arbitrary, we can eliminate the dependence on t by choosing

$$\lambda^{p_1}t = -1 \quad \Rightarrow \lambda = (t)^{-1/p_1} \tag{12.29}$$

hence,

$$M_s(t, 0) = -(t)^{\frac{1-p_2}{p_1}}M_s(-1, 0) \tag{12.30}$$

Obtaining the 1st relation:

$$\beta = \frac{1 - p_2}{p_1} \tag{12.31}$$

Exponent δ

At $T = T_c$ $(t = 0)$ we have

$$M_s \overset{h\to 0^+}{\sim} h^{1/\delta} \tag{12.32}$$

$$\Rightarrow M_s(0, h) = \lambda^{p_2-1}M_s(0, \lambda^{p_2}h) \tag{12.33}$$

Now, we want

$$\lambda^{p_1}h = 1 \quad \Rightarrow \lambda = h^{-1/p_2} \tag{12.34}$$

$$\Rightarrow M_s(0, h) = h^{\frac{1-p_2}{p_2}}M_s(0, 1) \tag{12.35}$$

Since $M_s \sim h^{1/\delta}$, we get the 2nd relation:

$$\delta = \frac{p_2}{1 - p_2} \tag{12.36}$$

By getting p_1 and p_2 from the two relations we have

$$p_1 = \frac{1}{\beta(\delta + 1)}, \quad p_2 = \frac{\delta}{\delta + 1} \tag{12.37}$$

and in particular the gap exponent is

$$\frac{p_2}{p_1} \equiv \Delta = \beta\delta \tag{12.38}$$

12.3.2 Equation of state

We can also predict the scaling form of the equation of state and explain the collapse of the experimental data. Let us start from

$$M_s(t,h) = \lambda^{p_2-1} M_s(\lambda^{p_1} t, \lambda^{p_2} h) \tag{12.39}$$

and choose $\lambda = |t|^{-1/p_1}$.

We have

$$M_s(t,h) = |t|^{\frac{1-p_2}{p_1}} M_s(\frac{t}{|t|}, \frac{h}{|t|^{\Delta}}) \tag{12.40}$$

Since $\beta = \frac{1-p_2}{p_1}$, $\delta = \frac{p_2}{p_1}$ we have

$$\frac{M_s(t,h)}{|t|^{\beta}} = M_s(\frac{t}{|t|}, \frac{h}{|t|^{\Delta}}) \tag{12.41}$$

We define the *scaled magnetization* as

$$\bar{m} \equiv |t|^{-\beta} M(t,h) \tag{12.42}$$

and the *scaled magnetic field* as

$$\bar{h} \equiv |t|^{-\Delta} h(t,M) \tag{12.43}$$

Hence, the equation (12.41) becomes

$$\bar{m} = F_{\pm}(\bar{h}) \tag{12.44}$$

universal curves!

12.3.3 Magnetic susceptibility

Compute $\frac{\partial^2}{\partial h^2}$ of the scaling hypothesis

$$f_s(\lambda^{p_1} t, \lambda^{p_2} h) = \lambda f_s(t,h) \tag{12.45}$$

From equation (12.41) we have

$$\lambda^{2p_2} \chi_T(\lambda^{p_1} t, \lambda^{p_2} h) = \lambda \chi_T(t,h) \tag{12.46}$$

Exponent γ

Consider $h = 0$, $t \to 0^-$ (for example). Let $\lambda = -t^{-1/p_1}$, from equation (12.41)

$$\chi_T(t,0) = (-t)^{-\frac{2p_2-1}{p_1}} \chi_T(-1,0) \tag{12.47}$$

and since

$$\chi_T(t,0) \overset{t\to 0^-}{\sim} (-t)^{-\gamma_-} \tag{12.48}$$

we get

$$\gamma_- = \frac{2p_2-1}{p_1} = \beta(\delta-1) \tag{12.49}$$

Remark. For $t \to 0^+$, we have $\lambda = t^{1/p_1}$ and we get

$$\gamma_- = \gamma_+ \equiv \gamma = \beta(\delta-1) \tag{12.50}$$

relation between exponents!

12.3.4 Scaling of the specific heat

Starting again from

$$f_s(\lambda^{p_1}t, \lambda^{p_2}h) = \lambda f_s(t,h) \tag{12.51}$$

and compute $\frac{\partial^2}{\partial t^2}$ of both side, we get

$$\lambda^{2p_1} c_h(\lambda^{p_1}t, \lambda^{p_2}h) = \lambda c_H(t,h) \tag{12.52}$$

For $h = 0$ and $t \to 0^-$, we have $\lambda = (-t)^{-1/p_1}$ which implies

$$-t^{-2}c_H(-1,0) = -t^{-1/p_1}c_H(t,0) \tag{12.53}$$

and since

$$c_H(t,0) \overset{t\to 0^-}{\sim} t^{-\alpha_-} \tag{12.54}$$

we get

$$\alpha_- = 2 - \frac{1}{p_1} \tag{12.55}$$

Also in this case it is easy to show that

$$\alpha_- = \alpha_+ \tag{12.56}$$

Moreover, by inserting in (12.55) the relation $p_1 = \frac{1}{\beta(\delta+1)}$ one gets the *Griffith relation*

$$\alpha + \beta(\delta+1) = 2 \tag{12.57}$$

Moreover, by combining Griffith with the relation $\gamma = \beta(\delta-1)$ one get the *Rushbrooke relation*

$$\alpha + 2\beta + \gamma = 2 \tag{12.58}$$

Remark. By using these hypothesis you can get equalities. All the inequality that you have in thermodynamics becames equalities. (lesson)

12.3.5 Second form of the static scaling hypothesis

$$f_s(\lambda^{p_1}t, \lambda^{p_2}h) = \lambda f_s(t,h) \tag{12.59}$$

The relation $\lambda = t^{-1/p_1}$ implies

$$f_s(1, t^{-p_2/p_1}h) = t^{-1/p_1}f_s(t,h) \tag{12.60}$$

Since $\Delta = \frac{p_2}{p_1}$ and $\alpha = 2 - \frac{1}{p_1}$, one gets

$$f_s(t,h) = t^{2-\alpha}f_s(1, \frac{h}{t^\Delta}) \tag{12.61}$$

that is the 2nd scaling form of the free-energy.

12.4 Kadanoff's block spin and scaling of the correlation function

Widom's theory refers only to the free energy of the system. Moreover it do, not fully explain how the scaling hypothesis derives by ξ being the only relevant length of the system.

The first argument on this point was given 1966 by Kadonoff. The Kadonoff's intuition was: the divergence of ξ implies a relation between the coupling constants of a $\mathcal{H}_{eff}$ and the length scales over which m is defined.

We will that Kadanoff's intuition is correct. However, the relation found by Kadanoff is an approximation of the one proposed later by the theory of renormalization.

Remark. How can we relate the coupling costant of the two hamiltonian. This is the idea of the renormalization group. (lesson)

12.4.1 Kadanoff's argument

It is based on a coarse-grained operation on the system and two basic assumptions.

Coarse graining operation

Let us start from the Hamiltonian

$$-\beta\mathcal{H}_\Omega = k\sum_{\langle ij\rangle}\sigma_i\sigma_j + h\sum_i \sigma_i \tag{12.62}$$

with

$$\sigma_i = \pm 1, \quad k \equiv \frac{J}{k_B T}, \quad h \equiv \frac{H}{k_B T} \tag{12.63}$$

Given xi we know that, for distances $r < \xi$, the spins are correlated.

Let partition the system into blocks of size la (a is the lattice unit length while l is a adimensional scale) such that

$$r \ll \xi \quad \Rightarrow a \ll la \ll \xi(t) \tag{12.64}$$

The coarse-graining procedure consits in substituting the spins σ_i in a block of length la with a *superspin* or *block spin* S_I.

The transformation chosen by Kadanoff was the following

$$S_I \equiv \frac{1}{|m_l|}\frac{1}{l^D}\sum_{i\in I}\sigma_i \tag{12.65}$$

where m_l is the average magnetization of the I-esim block

$$m_l \equiv \frac{1}{l^D}\sum_{i\in I}\sigma_i \tag{12.66}$$

with the sum that is over all the sites with a given cell.

Remark. The division by $|m_l|$ in equation (12.65) is crucial because it rescales the new variables S_I to have the original values ± 1 (rescaling of the fields).

1st crucial assumption

The frist assumption is that the Hamiltonian of the new system $\mathcal{H}_l$ is equal in form to $\mathcal{H}_\Omega$, the original one:

$$-\beta\mathcal{H}_l = k_l\sum_{\langle IJ\rangle}^{N_b} S_I S_J + h_l\sum_{I=1}^{N_b} S_I \tag{12.67}$$

Remark. This assumption is in general wrong!

Only the coupling constants k_l and h_l can change by the transformation. Note that the number of block spins N_l is given by $N_l = \frac{N}{l^D}$.

How you measure in the new system the lengths? the correlations lengths. Before we have changed the size of the system. We have increased the ruler, so the number is smaller now. It seems stupid but it is fundamental. An important point is: the

new system has a lattice length equal to la, hence in the new system all the lengths will be measured in units of la. In particular

$$\xi_l = \frac{\xi}{l} \tag{12.68}$$

that is ξ_l has a numerical value smaller than the one measured in the original lattice.

What does it mean? Since $\xi_l < \xi$ the system described by $\mathcal{H}_l$ is more distant from the critical point than the original $\mathcal{H}(\sigma_i)$! We should expect $t_l > t$.

Similarly, h_l will be such that

$$h\sum_i \sigma_i = h\sum_I \sum_{i\in\sigma_i} \sigma_i = h\sum_I |m_l| l^D S_I = \underbrace{h|m_l|l^D}_{h_l} \sum_I S_I \equiv h_l \sum_I S_I \tag{12.69}$$

giving

$$h_l = |m_l| l^D h \tag{12.70}$$

Since $\mathcal{H}_l$ is equal in form to $\mathcal{H}$, so it is Z_l and hence the free energies per spins satisfy the equation

$$N_l f_s(t_l, h_l) = N f_s(t, h) \tag{12.71}$$

Hence,

$$\Rightarrow f_s(t_l, h_l) = l^D f_s(t, h) \tag{12.72}$$

Note that the homogeneity condition is recovered with $\lambda \equiv l^D$. We now should ask how t and h change under the block spin transformation.

2st crucial assumption

The second assumption is

$$\begin{cases} t_l = t l^{Y_t} \\ h_l = h l^{Y_h} \end{cases} \tag{12.73}$$

where Y_t, Y_h are scaling exponents and are unknown! Inserting (12.73) in the free energy equation one obtains

$$f_s(t, h) = l^{-D} f_s(t l^{Y_t}, h l^{Y_h}) \tag{12.74}$$

As the parameter λ, in Widom also l can be chosen such that

$$l = |t|^{-1/Y_t} \tag{12.75}$$

so

$$f_s(t, h) = |t|^{D/Y_t} f_s(1, h|t|^{-Y_h/Y_t}) \tag{12.76}$$

Therefore,

$$\Delta = \frac{Y_h}{Y_t} \tag{12.77}$$

Remark. By comparing with the Widom's scaling

$$f_s(t, h) \sim t^{2-\alpha} f_s\left(1, \frac{h}{t^\Delta}\right) \tag{12.78}$$

we get the new relation

$$2 - \alpha = \frac{D}{Y_t} \tag{12.79}$$

12.5 Kadanoff's argument applied to the two-point correlation

Let us compute the two-point correlation function for the block spins

$$G_{IJ}(\vec{\mathbf{r}}_l, t_l) \equiv \langle S_I S_J \rangle - \langle S_I \rangle \langle S_J \rangle \tag{12.80}$$

Since

$$h_l = h|m_l|l^D \quad \Rightarrow |m_l| = \frac{h_l l^{-D}}{h} \tag{12.81}$$

and using $h_l = hl^{Y_h}$, we obtain

$$|m_l| = l^{\frac{1}{h}-D} \tag{12.82}$$

Since

$$S_I = \frac{1}{|m_l|}\frac{1}{l^D}\sum_{i\in I}\sigma_i \tag{12.83}$$

we have

$$\begin{aligned} G_{IJ}(\vec{\mathbf{r}}_l, t_l) &= \langle S_I S_J \rangle - \langle S_I \rangle \langle S_J \rangle \\ &= \frac{1}{l^{2(Y_h-D)}l^{2D}}\sum_{i\in I}\sum_{j\in J}\left[\underbrace{\langle \sigma_i\sigma_j\rangle - \langle \sigma_i\rangle\langle\sigma_j\rangle}_{(a)}\right] \\ &= \frac{l^D l^D}{l^{2(Y_h-D)}l^{2D}}[\langle \sigma_i\sigma_j\rangle - \langle \sigma_i\rangle\langle\sigma_j\rangle] \end{aligned} \tag{12.84}$$

in fact (a) we have made the assumption that since $la \ll \xi$, G_{ij} inside a block is fairly constant

Remark.

$$|\vec{\mathbf{r}}_l| = \frac{\vec{\mathbf{r}}}{l} \tag{12.85}$$

Hence,

$$G_{IJ}(\vec{\mathbf{r}}_l, t_l) = l^{2(D-Y_h)}G_{ij}(\vec{\mathbf{r}}, t) \tag{12.86}$$

If one includes also the h dependence:

$$G_{IJ}\left(\frac{\vec{\mathbf{r}}}{l}, tl^{Y_t}, hl^{Y_h}\right) = l^{2(D-Y_h)}G_{ij}(\vec{\mathbf{r}}, t, h) \tag{12.87}$$

By choosing $l = t^{-1/Y_t}$ one gets

$$G(\vec{\mathbf{r}}, t, h) = t^{\frac{2(D-Y_h)}{Y_t}}G(\vec{\mathbf{r}}t^{1/Y_t}, 1, ht^{-Y_h/Y_t}) \tag{12.88}$$

Since we are interested at large length scale in proximiti of $T = T_c$, we can choose $|\vec{\mathbf{r}}|$ such that

$$|\vec{\mathbf{r}}|t^{1/Y_t} = 1 \quad \Rightarrow t = |\vec{\mathbf{r}}|^{-Y_t} \tag{12.89}$$

Hence, inserting in (12.88)

$$G(\vec{\mathbf{r}}, t, h) = |\vec{\mathbf{r}}|^{-2(D-Y_h)}F_G\left(ht^{-Y_h/Y_t}\right) \tag{12.90}$$

where

$$F_G\left(ht^{-Y_h/Y_t}\right) \equiv G(1, ht^{-Y_h/Y_t}, 1) \tag{12.91}$$

Remembering the power law behaviour of G in proximity of the critical point, i.e. $G \sim |\vec{\mathbf{r}}|^{2-D-\eta}$ and $\xi \sim t^{-\nu}$. Finally we get

$$\nu = \frac{1}{Y_t}, \quad 2(D - Y_h) = D - 2 + \eta \tag{12.92}$$

Hence, we obtain the hyperscaling relation

$$\Rightarrow 2 - \alpha = \nu D \tag{12.93}$$

Chapter 13

Renormalization group theory. Universality

Lecture 20. *Friday 20th December, 2019. Compiled: Thursday 26th December, 2019.*

13.1 Renormalization group (RG)

Kadonoff: the block spin transformations justifies the Widom scaling with $\lambda \iff l$,

$$\begin{cases} f_s(t,h) = l^{-D} f_s\big(tl^{Y_t}, hl^{Y_h}\big) \\ G(\vec{\mathbf{r}},t,h) = l^{-2(D-Y_h)} G\Big(\frac{\vec{\mathbf{r}}}{l}, tl^{Y_t}, hl^{Y_h}\Big) \\ t_l = tl^{Y_t}, \quad h_l = h^{Y_h} \end{cases} \tag{13.1}$$

We do two crucial assumptions:

1. $$\mathcal{H}_l = \mathcal{H}_1 \tag{13.2}$$

2. $$\begin{cases} t_l = tl^Y \\ h_l = hl^h \end{cases} \tag{13.3}$$

Remark. An open problem is: how an iterative procedure of coarse-graining can produce the second assumption?

How this can gives rise to the singular behaviour of f_s? How can we explain universality of the critical points?

13.1.1 Main goals of RG

1. To fornish an algorithm way to perform systematically the coarse graining procedure.

   Achivede by integrating the degrees of freedom at smaller length scales and obtain a new hamiltonian describing the system at larger length scales.

   This will be equivalent to find a transformation between the coupling constants $k \to k'$.

2. Identify the origin of the critical behaviour.

   The coarse graining procedure will give rise to a system with

   $$\xi_l = \frac{\xi}{l} \tag{13.4}$$

   more distant from criticality.

Let us start from a generic hamiltonian $\mathcal{H} = \mathcal{H}([k])$, where $[k] = (k_1, k_2, \dots k_n)$. For example, for the standard Ising model $k \equiv k_1$ and $h \equiv k_2$. In general we can have n coupling constants.

Suppose we apply a coarse-graining procedure in which we integrate the degree of freedom within distance l with $a \leq la \leq L$.

There are many ways (for example the Kadanoff's block spin)

$$[k'] = \mathcal{R}_l[k], \quad l > 1 \tag{13.5}$$

where $\mathcal{R}_l$ is the transformation of the RG. The relation (13.5) gives rise to a recursive relation.

The properties of $\mathcal{R}_l$ are:

1. $\mathcal{R}_l$ is analitic! (Sum of a finite number of degrees of freedom).

2. The set of $\mathcal{R}_l$, labeled by l forms a semigroup.

   Let $\mathcal{R}_{l_1}$ and $\mathcal{R}_{l_2}$ be two transformations

   $$\begin{cases} [k'] = \mathcal{R}_{l_1}[k] \\ [k''] = \mathcal{R}_{l_2}[k'] = \mathcal{R}_{l_2} \circ \mathcal{R}_{l_1}[k] \end{cases} \tag{13.6}$$

   Hence,

   $$\mathcal{R}_{l_2 l_1}[k] = \mathcal{R}_{l_2} \circ \mathcal{R}_{l_1}[k] \tag{13.7}$$

   Remark. Is not a group since does not exist the inverse element ($l > 1$ always!).

   Remark. There is no a unique way to define $\mathcal{R}_{l_1}$ but all of them result in the integration of the degrees of freedom at smaller length scales.

   The procedure can be carried out either in real space or in the momentum space. $\mathcal{R}_{l_1}$ but all of them result in the integration of the degrees of freedom at smaller negth scales.

   Given $\mathcal{H}[k]$ we can define

   $$Z_N[k] = \operatorname{Tr} e^{-\beta \mathcal{H}[k]} \tag{13.8}$$

   and

   $$f_N[k] = -\frac{k_B T}{N} \log Z_N[k] \tag{13.9}$$

   After the coarse graining between a and la, the number N of degrees of freedom is reduced by

   $$N_l = \frac{N}{l^D} \tag{13.10}$$

3. The new hamiltonian $\mathcal{H}_l[k']$ can be (and in general it is) different from the previous one: $\mathcal{H}[k]$ (main difference with Kadanoff) but its symmetry cannot change!

   For example, if we start from

   $$\mathcal{H}_N = N k_0 + k_2 \sum_{ij} S_i S_j \tag{13.11}$$

   we cannot produce

   $$\mathcal{H}_{N'} = N' k_0' + k_1' \sum_I S_I + k_2' \sum_{IJ} S_I S_J + k_3' \sum_{IJK} S_I S_J S_k \tag{13.12}$$

Remark. If $km = 0$, it is possible that $k'm \neq 0$ as long as the symmetry of $\mathcal{H}'$ is equal to the one of $\mathcal{H}$.

4. The invariance condition is not in $\mathcal{H}$ (as in Kadanoff) but in Z!

$$Z_{N'}[k'] = Z_N[k] \tag{13.13}$$

Condition (13.13) must be satisfied by the coarse-graining procedure. Which is the consequence of (13.13) on the free-energy?

$$f_N[k] \simeq \frac{1}{N} \log Z_N[k] \simeq \frac{l^D}{l^D N} \log Z_{N'}[k'] \simeq l^{-D} \frac{1}{N'} \log Z_{N'}[k'] \tag{13.14}$$

Hence,

$$f[k] \propto l^{-D} f\ [k'] \tag{13.15}$$

Remark. Since a single $\mathcal{R}_l$ involves the integration of a finite number of dregrees of freedom, it cannot develop the singularity behaviour we are looking for.

Question: if $\mathcal{R}_l$ is analitic, what is the origin of the singular behaviour in the RG approach?

13.1.2 Singular behaviour in RG

The point is that, in order to integrate a thermodynamic number of degrees of freedom, one has to apply an infinite number of RG transformations.

After n iterations we have $l \to l^n$ and $[k] \to [k^{(n)}]$. As n changes, the coupling constants perform a trajectory in the parameter space $[k]$.

By varying the initial conditions (i.e. either the starting model or the values of the physical parameters) one obtains a *flux of trajectories.* A part from some pathological cases (cycle limits, attractors) these trajectories are attracted towards fixed points.

The scaling behaviour introduced by Widom is related to the behaviour of the trajectories close to same fixed points.

13.1.3 Zoology of the fixed points

A fixed point k^* of $\mathcal{R}_l[k]$ satisfies

$$[k^*] = \mathcal{R}_l[k^*] \tag{13.16}$$

On the other hand, we know that

$$\xi[k'] = \frac{\xi[k]}{l} \equiv \xi_l \tag{13.17}$$

At the fixed point

$$\xi[k^*] = \frac{\xi[k^*]}{l} \tag{13.18}$$

There are two cases:

$$\xi[k^*] = \begin{cases} 0 & \text{trivial} \\ \infty & \text{critical} \end{cases} \tag{13.19}$$

Each fixed point has its own basin of attraction. This is defined as the set $\{[k]\}$ such that

$$\mathcal{R}_l^{(n)}[k] \stackrel{n\to\infty}{\longrightarrow} [k^*] \tag{13.20}$$

Theorem 13.1.1. *All the points $[k]$ belonging to a basin of attraction of a critical fixed point have their correlation length $\xi = \infty$.*

Proof.

$$\xi[k] = l\xi[k^{(1)}] = \dots l^n \xi[k^{(n)}] \tag{13.21}$$

If $[k]$ is in basin of attraction of $[k^*]$, we have

$$[k^{(n)}] \overset{n\to\infty}{\longrightarrow} [k^*] \tag{13.22}$$

On the other hand, $\xi[k^*] = \infty$, so

$$\Rightarrow \xi[k] = l^n \xi[k^*] = \infty \tag{13.23}$$

□

Definition 9 (Critical manifold)**.** Set of $[k]$ that forms the basin of attraction of a critical fixed point.

Universality

All the critical models that belong to the critical manifold, have the same critical behaviour of the corresponding critical fixed point. Study the behaviour of $\mathcal{R}_l$ close to the fixed points.

13.1.4 Linearization of RG close to the fixed points and critical exponents

Let $[k^*]$ be a fixed point of $\mathcal{R}_l$ and consider

$$\vec{\mathbf{k}} = \vec{\mathbf{k}}^* + \delta\vec{\mathbf{k}} \tag{13.24}$$

In components, the RG transformation is

$$k'_j = (\mathcal{R}_l)_j(\vec{\mathbf{k}}^* + \delta\vec{\mathbf{k}}) = k^*_j + \sum_i \left(\frac{\partial k'_j}{\partial k_i}\right)_{\vec{\mathbf{k}}^*} \delta k_i + O\big(\delta k_i^2\big) \tag{13.25}$$

The linearized transformation is

$$\delta\vec{\mathbf{k}}' = \bar{\pi}\delta\vec{\mathbf{k}} \tag{13.26}$$

where

$$(\bar{\pi})_{ij} = \left(\frac{\partial k'_j}{\partial k_i}\right)_{\vec{\mathbf{k}}^*} \tag{13.27}$$

is a square matrix

1. $\bar{\pi}$ is in general not symmetric. One has to distinguish between left and right eigenvectors.

2. $\bar{\pi}$ is not always diagonalizible or sometimes the eigenvalues are complete. For most of the physical system, however, $\bar{\pi}$ can be diagonalized and the eigenvalues are real.

Let $\lambda_l^{(r)}$ be the σ-esim eigenvalue and $\vec{\mathbf{l}}^{(\sigma)}$ the corresponding eigenvector

$$\Rightarrow \bar{\pi}_{ij}^{(l)} l_j^{(\sigma)} = \lambda_l^{(\sigma)} l_i^{(\sigma)} \tag{13.28}$$

Since the $\mathcal{R}_l$'s form a semi-group

$$\bar{\pi}^{(l)}\bar{\pi}^{(l')} = \bar{\pi}^{(ll')} \tag{13.29}$$

and

$$\lambda_l^{(\sigma)}\lambda_{l'}^{(\sigma)} = \lambda_{ll'}^{(\sigma)} \tag{13.30}$$

By solving (13.30) one can show

$$\lambda_l^{(\sigma)} = l^{(Y_\sigma)} \tag{13.31}$$

where Y_σ is the exponent to be determined.

How $\delta\vec{\mathbf{k}}$ behaves under the linearized transformation $\bar{\pi}$?

Let us expand $\delta\vec{\mathbf{k}}$ in terms of $\vec{\mathbf{l}}^{(\sigma)}$

$$\delta\vec{\mathbf{k}} = \sum_\sigma a^{(\sigma)}\vec{\mathbf{l}}^{(\sigma)} \tag{13.32}$$

where

$$a^{(\sigma)} = \vec{\mathbf{l}}^{(\sigma)} \cdot \delta\vec{\mathbf{k}} \tag{13.33}$$

Remark. Ortonormality is not always true since in general $\bar{\pi}$ is not symmetric!

$$\delta\vec{\mathbf{k}}' = \bar{\pi}\delta\vec{\mathbf{k}} = \sum_\sigma a^{(\sigma)}\bar{\pi}\left(\vec{\mathbf{l}}^{(\sigma)}\right) = \sum_\sigma a^{(\sigma)}\lambda^{(\sigma)}\vec{\mathbf{l}}^{(\sigma)} = \sum_\sigma a^{(\sigma)\prime}\vec{\mathbf{l}}^{(\sigma)} \tag{13.34}$$

where $a^{(\sigma)\prime}$ is the projection of $\delta\vec{\mathbf{k}}'$ on $\vec{\mathbf{l}}^{(\sigma)}$.

Same components of $\delta\vec{\mathbf{k}}$ grow under the action of $\bar{\pi}$ while same others shrink. If we order the eigenvalues in descending order

$$\left|\vec{\lambda}_1\right| \geq \left|\vec{\lambda}_2\right| \geq \left|\vec{\lambda}_3\right| \geq \dots \tag{13.35}$$

threee cases can occur

1. Case $\left|\lambda^{(\sigma)}\right| > 1$ (i.e. $y^\sigma > 0$): implies that $a^{(\sigma)}$ grows under $\bar{\pi}$. There are relevant eigenvalues/eigenvectors.
2. Case $\left|\lambda^{(\sigma)}\right| < 1$ (i.e. $y^\sigma < 0$): implies that $a^{(\sigma)}$ decreases under $\bar{\pi}$. There are irrelevant eigenvalues/eigenvectors.
3. Case $\left|\lambda^{(\sigma)}\right| = 1$ (i.e. $y^\sigma = 0$): implies that $a^{(\sigma)}$ remains constant under $\bar{\pi}$. There are marginal eigenvalues/eigenvectors.

Starting from a point close to a critical fixed point $\vec{\mathbf{k}}^*$ (but not on the critical manifold), the trajectory will abandau $[k^*]$ along the relevant directions whereas it will approach $[k^*]$ along the irrelevant directions.

- *Irrelevant eigenvectors* form the local basis of the basin of attraction of $[k^*]$.
- *Relevant eigenvectors* codimension C of the basin of attraction.

Remark. The eigenvalues and the corresponding eigenvectors depend on the chosen fixed point $[k^*]$.

13.2 RG and scaling

Let

$$\begin{cases} k_1 \to k_1(T) = T \\ k_2 \to k_2(H) = H \end{cases} \tag{13.36}$$

for simplicity. Suppose to perform a RG transformation $\mathcal{R}_l$ such that

$$\begin{cases} T' = \mathcal{R}_l^T(T, H) \\ H' = \mathcal{R}_l^H(T, H) \end{cases} \tag{13.37}$$

The critical fixed point is defined as

$$\begin{cases} T^* = \mathcal{R}_l^T(T^*, H^*) \\ H^* = \mathcal{R}_l^H(T^*, , H^*) \end{cases} \tag{13.38}$$

with $\xi(T^*, H^*) = \infty$. For standard magnetic systems $H^* = 0$. By linearising around the fixed point $\delta\vec{\mathbf{k}} = (t, h)$ where

$$\begin{cases} t = \frac{(T-T^*)}{T^*} \\ h = \frac{(H-H^*)}{H^*} \end{cases} \tag{13.39}$$

we obtain

$$\begin{pmatrix} t' \\ h' \end{pmatrix} = \bar{\pi} \begin{pmatrix} t \\ h \end{pmatrix} \tag{13.40}$$

with

$$\bar{\pi} = \begin{pmatrix} \frac{\partial \mathcal{R}_l^T}{\partial T} & \frac{\partial \mathcal{R}_l^T}{\partial H} \\ \frac{\partial \mathcal{R}_l^H}{\partial T} & \frac{\partial \mathcal{R}_l^H}{\partial H} \end{pmatrix}_{T^*, H^*} \tag{13.41}$$

In general the eigenvectors $\vec{\mathbf{l}}^{(\sigma)}$ are linear combinations of t and h. When $\bar{\pi}$ can be diagonalized, t and h are not mixed. Suppose this is the case and let

$$\lambda_l^{(t)} = l^{Y_t}, \quad \lambda_l^{(h)} = l^{Y_h} \tag{13.42}$$

A linear transformation is

$$\begin{pmatrix} t' \\ h' \end{pmatrix} = \begin{pmatrix} \lambda_l^{(t)} & 0 \\ 0 & \lambda_l^{(h)} \end{pmatrix} \begin{pmatrix} t \\ h \end{pmatrix} \tag{13.43}$$

After n iterations we have

$$\begin{cases} t^{(n)} = \left(l^{Y_t}\right)^n t \\ h^{(n)} = \left(l^{Y_h}\right)^n h \end{cases} \tag{13.44}$$

On the other hand,

$$\xi' \equiv \xi(t', h') = \frac{\xi(t, h)}{l} \tag{13.45}$$

and after n iterations

$$\xi(t, h) = l^n \xi\left(l^{nY_t} t, l^{nY_h} h\right) \tag{13.46}$$

Let $h = 0$ ($H = H^*$), hence

$$\xi(t, 0) = l^n \xi\left(l^{nY_t} t, 0\right) \tag{13.47}$$

Since l is arbitrary, we can choose it such that

$$t l^{nY_t} = b \gg 1, \quad b \in \mathbb{R}^+ \tag{13.48}$$

Remark. l is not an integer any more.

$$\Rightarrow l^n = \left(\frac{b}{t}\right)^{1/Y_t} \tag{13.49}$$

so, for $t \to 0$

$$\xi(t) = \left(\frac{t}{b}\right)^{-1/Y_t} \xi(b) \tag{13.50}$$

On the other hand $\xi \sim t^{-\nu}$, $t \to 0$

$$\Rightarrow \nu = \frac{1}{Y_t} \tag{13.51}$$

where

$$Y_t = \frac{1}{l} \ln \lambda_l^t \tag{13.52}$$

The equation (13.52) is an important result since it gives a recipe, based on $\mathcal{R}_l$, to compute Y_t explicitely!

Similarly,

$$\begin{aligned} f_s(t,h) &= l^{-D} f_s(t',h') = l^{-nD} f_s(t^{(n)}, h^{(n)}) \\ &= l^{-nD} f_s\big(l^{nY_t} t, l^{nY_h} h\big) \end{aligned} \tag{13.53}$$

Let l such that $l^{nY_t} = b$, hence

$$l^n = \left(\frac{b}{t}\right)^{1/Y_t} \tag{13.54}$$

$$f_s(t,h) = t^{D/Y_t} b^{-D/Y_t} f_s\left(b, \frac{b^{Y_h/Y_t} h}{t^{Y_h/Y_t}}\right) \tag{13.55}$$

that is of the form

$$f_s(t,h) = \bar{b} t^{2-\alpha} f_s\left(b, \frac{\bar{b} h}{t^{\Delta}}\right) \tag{13.56}$$

with

$$\begin{cases} 2 - \alpha = D\nu = \frac{D}{Y_t} \\ \Delta = \frac{Y_h}{Y_t} \end{cases} \tag{13.57}$$

$$Y_h = \frac{1}{l} \log \lambda_l^{(h)} \tag{13.58}$$

13.3 Real space renormalization group on 1D Ising model ($H = 0$)

Decimation procedure. $D = 1$ the procedure is clearly exact. Idea: pass from a $N-$ spins system to one with $\frac{N}{b}$ spins by summing the remaining $N - \frac{N}{b}$ spins.

Case b=3 (see Figure 13.1).

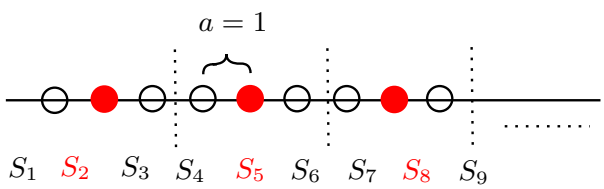


Figure 13.1: Description.

Sum over the spins with the empty circle at the border of each clok, keeping the spins full circle untouched. We obtain Figure 13.2.

Let us see how it works for the two blocks $[S_1, S_2, S_3]$ and $[S_4, S_5, S_6]$. Let us call

$$S_2 \equiv S_1', \quad S_5 = S_2' \quad \text{fixed} \tag{13.59}$$

$$\sum_{S_3=\pm 1} \sum_{S_4=\pm 1} \exp\big[k S_1' S_3 + k S_3 S_4 + k S_4 S_2'\big] \tag{13.60}$$

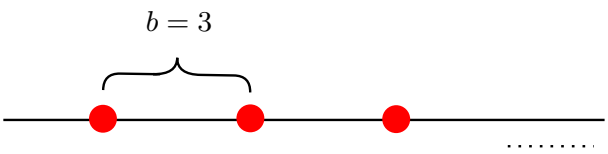


Figure 13.2: Description.

Since

$$e^{kS_3S_4} = \cosh(k)(1 + xS_3S_4) \tag{13.61}$$

with

$$x \equiv \tanh k \tag{13.62}$$

we have

$$\sum_{S_3,S_4} (\cosh k)^3 \big(1 + xS_1'S_3\big)(1 + xS_3S_4)\big(1 + xS_4S_2'\big) \tag{13.63}$$

Performing the expansion and summing over S_3, S_4 it is easy to show that (to do)

$$Z'_{N'}(k') = \mathrm{Tr}_{\{S_I'\}} \left[2^{2N'} (\cosh k)^{3N'} \big(1 + x^3 S_I' S_{I+1}\big)\right] \tag{13.64}$$

This must have the same form of $Z_N(k)$. Hence, we should rewrite equation (13.64) as

$$Z'_{N'}(k') = \mathrm{Tr}_{\{S_I'\}} \exp\big[-\beta\mathcal{H}'(k')\big] \tag{13.65}$$

with

$$-\beta\mathcal{H}' = N' g(k, k') + k' \sum_I S_I' S_{I+1}' \tag{13.66}$$

We note that

$$\begin{aligned} 2^2(\cosh k)^3(1 + x^3 S_I' S_{I+1}') &= 2^2 \frac{\cosh k'}{\cosh k'} (\cosh k)^3 (1 + x^3 S_I' S_{I+1}') \\ &= 2^2 \frac{(\cosh k)^3}{\cosh k'} \big(\cosh k'\big)(1 + x' S_I' S_{I+1}') \\ &= 2^2 \frac{(\cosh k)^3}{\cosh k'} \exp\big(k' S_I' S_{I+1}'\big) \\ &= \exp\left[2\ln 2 + \ln\left[\frac{(\cosh k)^3}{\cosh k'}\right] + k' S_I' S_{I+1}'\right] \end{aligned} \tag{13.67}$$

It is ok with

$$\begin{cases} g(k, k') = 2\ln 2 + \ln\left[\frac{(\cosh k)^3}{\cosh k'}\right] \\ x' = x^3 \iff k' = \tanh^{-1}\left[(\tanh k)^3\right] \Rightarrow k' = R_{b=3}(k) \\ N' = \frac{N}{b} \end{cases} \tag{13.68}$$

For $x' = x^3$, we have two fixed points

$$\begin{cases} x^* = 0 \iff k \to 0 \iff T \to \infty \\ x^* = 1^- \iff k \to \infty \iff T \to 0 \end{cases} \tag{13.69}$$

For $\forall x_0 < 1$, we have $R^{(n)} \overset{n\to\infty}{\longrightarrow} 0^+$, $x = 1$ is an unstable fixed point, as shown in Figure 13.3. We have

$$\xi(x') = \frac{\xi(x)}{b}, \quad \text{with } x' = x^b \tag{13.70}$$

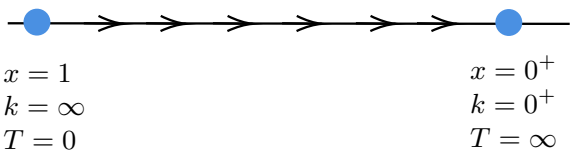


Figure 13.3: Description.

where b is arbitrary. Let us choose $b = c/\ln x$:

$$\xi(x^b) = \xi\left(e^{b\ln x}\right) = \xi(e^c) = \frac{\xi(x)}{b} = \left(\frac{c}{\ln x}\right)^{-1}\xi(x) \tag{13.71}$$

Finally,

$$\xi(e^c) = \frac{\ln x}{c}\xi(x) \quad \Rightarrow \xi(x) = \frac{c\xi(e^c)}{\ln x} = \frac{const}{\ln x} \tag{13.72}$$

$$\xi(k) = \frac{const}{\ln(\tanh k)} \tag{13.73}$$

equal to the exact model in $D = 1$. $k \to \infty$ $(x \to 1)$, we have

$$\xi \sim e^{const/T} \tag{13.74}$$

finite $\forall k \neq \infty$.

13.4 Decimation procedure for $D > 1$ (proliferation of interactions)

Ioing on a square lattice, block transformation (length b). There are problems with the spins at the boundaries. In order to get $Z'_{N'}(k') = Z_N(k)$, at least an additional term in the Hamiltonian must be introduced

$$\mathcal{H} \to \mathcal{H}' = \mathcal{H}'(k', k'_2) \tag{13.75}$$

This unfortunately occurs at each iteration! There is an uncontrolled proliferation and approximations are necessary.

Example 29 (Decimation of Ising on square lattice)**.** x: spin to sum over

$$\begin{cases} k' = \frac{1}{4}\ln\cosh 4k \\ L' = \frac{1}{8}\ln\cosh 4k \\ Q' = \frac{1}{8}\ln\cosh 4k - \frac{1}{2}\ln\cosh 2k \end{cases} \tag{13.76}$$

(4 spins). The approximations are:

1. Neglect Q' term.
2. Omit explicit dependence on L'. Hence, assuming $k' + L' \to k'$. 1 recursion

$$k' = \frac{3}{8}\ln\cosh 4k \tag{13.77}$$

$$\begin{cases} k^{(n)} \overset{n\to\infty}{\longrightarrow} \infty & k_0 > k_c \\ k^{(n)} \overset{n\to\infty}{\longrightarrow} 0 & k_0 < k_c \end{cases} \tag{13.78}$$

where $\infty, 0$ are stable fixed point.

$$\delta k' = \lambda_t \delta k = l^{\lambda_t}\delta k \tag{13.79}$$

where

$$\lambda_t = \left.\frac{\mathrm{d}k'}{\mathrm{d}k}\right|_{k=k_c} \tag{13.80}$$

For $l = \sqrt{2}$,

$$\begin{aligned} Y_t = \frac{\ln \lambda_t}{\ln l} &= \frac{1}{\ln \sqrt{2}} \ln \left(\frac{\mathrm{d}k'}{\mathrm{d}k}\right)_{k=k_c} \\ &= \frac{1}{\ln \sqrt{2}} \ln \left(\frac{\mathrm{d}}{\mathrm{d}k}\left(\frac{3}{8}\ln\left(\cosh 4k\right)\right)\right) = 1.070 \end{aligned} \tag{13.81}$$

we have also

$$\nu = \frac{1}{Y_t} = 0.9345, \quad \alpha = 2 - \frac{d}{Y_t} = 0.1313 \tag{13.82}$$

Chapter 14

Spontaneous symmetry breaking

Lecture n.
Tuesday 0th January, 2020.
Compiled: Thursday 26th December, 2019.

14.1 Spontaneous symmetry breaking

When we talk about a broken symmetry, we oftene refer to a situation as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1 \tag{14.1}$$

where $\mathcal{H}_0$ is invariant under the group $\mathcal{G}$ and $\mathcal{H}_1$ is invariant under a subgroup $\mathcal{G}' \subset \mathcal{G}$.

Example 30 (Ising with magnetic field)**.**

$$\mathcal{H} = J \sum_{\langle ij \rangle} S_i S_j + \sum_i H_i S_i \tag{14.2}$$

The second term, $\mathcal{H}_1$, breaks the $\mathbb{Z}^2$ symmetry satisfied by the 1st alone.

Example 31. In quantum mechanics: hydrogen atom in presence of an electric field $\vec{\mathbf{E}}$ (Stark effect) or a magnetic one, $\vec{\mathbf{B}}$, (Zeeman effect). If $\mathcal{H}_1$ is small, the original symmetry is weakly violated and perturbativ approaches are often used.

In all the above examples, one says that the symmetry is broken explicitly.

Definition 10 (Spontaneous symmetry breaking)**.** The Hamiltonian maintains the original symmetry but the variables used to describe the system become asymmetric.

At this point it is convenient to distinguish between

- Discrete symmetries: examples are $\mathbb{Z}^2$, $\mathbb{Z}_q$.
- Continuous symmetries: examples are xy, $O(n)$.

Let us consider first the discrete ones by focusing on $\mathbb{Z}^2$ (Ising).

If $H = 0$, $\mathcal{H}_{Ising}$ is invariant with respect to the change $S_i \to -S_i$, hence the discrete group is

$$\mathcal{G} = \mathbb{Z}^2 \tag{14.3}$$

A Ginzburg-Landau theory of the Ising is given by

$$-\beta\mathcal{H}(\Phi) = \int \mathrm{d}^D\vec{\mathbf{x}} \left[\frac{1}{2}(\boldsymbol{\nabla}\Phi)^2 + \frac{r_0}{2}\Phi^2 + \frac{u_0}{4}\Phi^4 - h\Phi \right] \tag{14.4}$$

and

$$Z(r_0, u_0, h) = \int \mathrm{D}\Phi e^{-\beta\mathcal{H}(\Phi)} \tag{14.5}$$

The symmetry is $\Phi \to -\Phi$ if $h = 0$. Consider the saddle point equation of state

$$-\boldsymbol{\nabla}^2\Phi + r_0\Phi + u_0\Phi^3 = h \tag{14.6}$$

If h does not depend on $\vec{\mathbf{x}}$, uniform solution ($\boldsymbol{\nabla}\Phi = 0$).

The saddle point is equivalent to find the uniform value Φ_0 that is the extrema of the potential

$$V(\Phi) = \frac{1}{2} r_0 \Phi^2 + \frac{u_0}{4}\Phi^4 - h\Phi \tag{14.7}$$

For $h = 0$, $V' = (r_0 + u_0\Phi^2)\Phi = 0$. Remembering that $r_0 \propto (T - T_c)$, we have two cases

1. Case $T > T_c$ ($r_0 > 0$): there is only one solution $\Phi_0 = 0$.

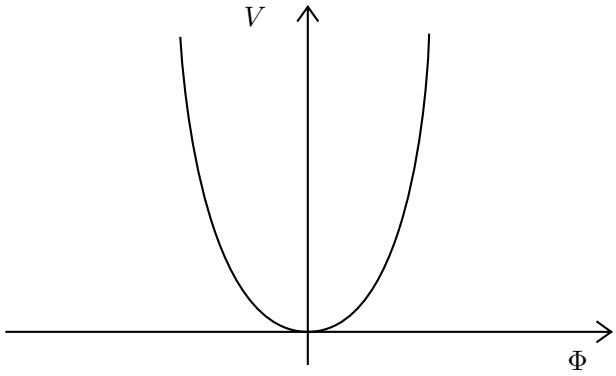


Figure 14.1: Description.

2. Case $T < T_c$ ($r_0 < 0$): there are two solutions $\Phi_0 = \pm\sqrt{-\frac{r_0}{u_0}}$.

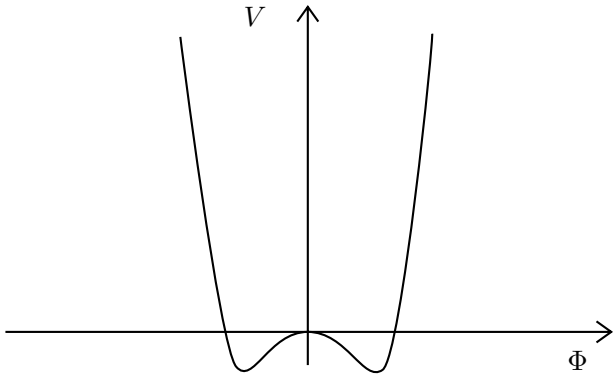


Figure 14.2: Description.

Remark. The two solution $\pm\Phi_0$ are related by the transformation $\in \mathbb{Z}^2$: $\Phi \to -\Phi$.

Remark. For $T < T_c$ the two states (phases) $\pm\Phi_0$ have a lower symmetry than the state $\Phi_0 = 0$.

Remark. If the thermal fluctuations $\delta\Phi$ are sufficiently strong to allow passages between the two states $\pm\Phi_0$ at $T < T_c$, we have $\langle\Phi\rangle = 0$ (preserves states).

However, for $T < T_c$ and $N \to +\infty$, transition between the two states will be less and less probable and the system will be trapped into one of the two states ($\pm\Phi_0$). The system choose spontaneously one of the two less symmetric state. Therefore, its physics is not any more described by Φ but the fluctuations $\delta\Phi$ around the chosen minimum Φ_0. There is spontaneous symmetry breaking.

The variable Φ is not any more symmetric and one has to look at $\Phi \to \Phi_0 + \delta\Phi$, where $\delta\Phi$ is a new variable!

14.2 Spontaneous breaking of continuous symmetries and the anset of Goldstone particles

Let us start with a simple model in which the order parameter is a scalar complex variable

$$\Phi = \frac{\Phi_1 + i\Phi_2}{\sqrt{2}} \tag{14.8}$$

and with an $\mathcal{H}$ that is invariant with respect to a global continuous transformation.

The simplest model in statistical mechanics is the XY model with $O(2)$ symmetry or a GL model for a superfluid or superconductor (no magnetic field)

$$\mathcal{H}_{eff} = \int \mathrm{d}^D\vec{\mathbf{x}} \left[\boldsymbol{\nabla}\Phi \cdot \boldsymbol{\nabla}\Phi^* + \frac{r_0}{2}\Phi^*\Phi + \frac{u_0}{4}(\Phi^*\Phi)^2 \right] \tag{14.9}$$

where

$$\Phi(\vec{\mathbf{x}}) = \frac{1}{\sqrt{2}}[\Phi_1(\vec{\mathbf{x}}) + i\Phi_2(\vec{\mathbf{x}})] \tag{14.10}$$

or

$$\Phi(\vec{\mathbf{x}}) = \psi(\vec{\mathbf{x}})e^{i\alpha(\vec{\mathbf{x}})} \tag{14.11}$$

- Superfluid: Φ macroscopic wave function of the Bose condensate (density of superfluid $n = |\Phi^2|$).
- Superconductor: Φ single particle wave function describing the position of the centre of mass of the Cooper pair.

14.2.1 Quantum relativistic case (field theory)

The analog of $\mathcal{H}$ is the action

$$S(\Phi) = \int \mathrm{d}^D\vec{\mathbf{x}}\, \mathcal{L}(\Phi) \tag{14.12}$$

where

$$\mathcal{L}(\Phi) = -\frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi^* - \frac{r_0}{2}\Phi\Phi^* - \frac{u_0}{4}(\Phi\Phi^*)^2 \tag{14.13}$$

It describes a scalar complex (i.e. charged) muonic field with mass m. Note that in this case $r_0 > 0$ and $m \equiv \sqrt{r_0}$. The term $(\Phi\Phi^*)^2$ means self-interaction with strenght $\lambda \equiv u_0$.

In all cases, the original symmetry is $U(1)$, i.e. both $\mathcal{H}$ and $\mathcal{L}$ are invariant with respect to the transformation

$$\Phi \to e^{i\theta}\Phi, \quad \Phi^* \to e^{-i\theta}\Phi^* \tag{14.14}$$

Remark. The phase θ does not depend on $\vec{\mathbf{x}}$ (global symmetry).

In components (14.14) become

$$\begin{cases} \Phi_1 \to \Phi_1\cos\theta - \Phi_2\sin\theta \\ \Phi_2 \to \Phi_2\cos\theta + \Phi_1\sin\theta \end{cases} \tag{14.15}$$

$$(\Phi_1', \Phi_2') = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} \tag{14.16}$$

Let us focus first on the statistical mechanics model and to the most interesting case of $r_0 < 0$.

In components $\mathcal{H}$ becomes

$$\mathcal{H} = \int \mathrm{d}^D\vec{\mathbf{x}} \left[(\boldsymbol{\nabla}\Phi_1)^2 + (\boldsymbol{\nabla}\Phi_2)^2 \right] + \int \mathrm{d}^D\vec{\mathbf{x}}\, V(\Phi_1, \Phi_2) \tag{14.17}$$

where

$$V(\Phi_1, \Phi_2) = \frac{r_0}{2}\left(\Phi_1^2 + \Phi_2^2\right) + \frac{u_0}{4}\left(\Phi_1^2 + \Phi_2^2\right)^2 \tag{14.18}$$

It is the mexican hat potential, shown in Figure 14.3

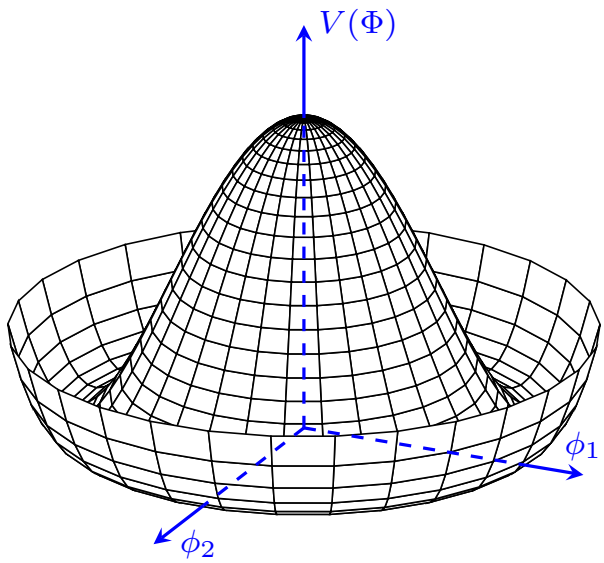


Figure 14.3: Case $r_0 < 0$.

For $r_0 < 0$, there is a uniform solution ($\boldsymbol{\nabla}\Phi_1 = \boldsymbol{\nabla}\Phi_2 = 0$). Let $S = \sqrt{\Phi_1^2 + \Phi_2^2}$,

$$V(S) = \frac{r_0}{2}S^2 + \frac{u_0}{4}S^4 \tag{14.19}$$

$$\frac{\mathrm{d}V(S)}{\mathrm{d}S} = r_0 S + u_0 S^3 = 0 \tag{14.20}$$

There is 1 maximum at $S = 0$ and minima for $S^2 = -\frac{r_0}{u_0}$

For $r_0 < 0$, $\mathcal{H}$ displays minima when

$$\Phi_1^2 + \Phi_2^2 \equiv v^2 = -\frac{r_0}{u_0} \tag{14.21}$$

On the $2D$ plane (Φ_1, Φ_2) the minima lie on the circle of radius

$$v = \sqrt{-\frac{r_0}{u_0}} \tag{14.22}$$

The spontaneous symmetry breaking occurs when the system "chooses" one of the infinite available minima.

In our example, suppose that the chosen minimum is

$$\Phi_1 = v = \sqrt{-\frac{r_0}{u_0}}, \quad \Phi_2 = 0 \tag{14.23}$$

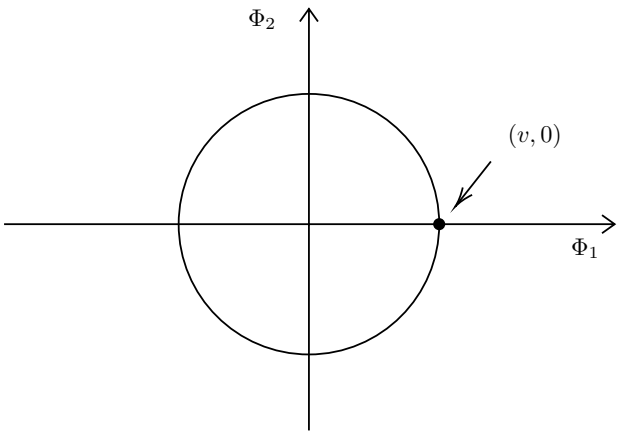


Figure 14.4: Description.

14.3 Interpretation in relativistic quantum mechanics

1. $r_0 < 0$ corresponds to an imaginary mass. This is because to move away from $\Phi = 0$, the system experiences a negative resistence in both directions, being $\Phi = 0$ a relative local minimum.

2. The minimum has the lowest energy and therefore it must correspond to the empty state. In this case, however, there is an infinite number of empty states!

Summarizing: the starting Hamiltonian (or Lagrangian) is invariant with respect to $U(1)$ but the one that describes the fluctuation dynamics around one of the chosen minimum state is not invariant with respect to $U(1)$.

Let us now write the Lagrangian with respect to the fluctuations of Φ_1 and Φ_2 around the chosen state

$$\begin{cases} \Phi_1 = v + \delta\Phi_1 \\ \Phi_2 = 0 + \delta\Phi_2 \end{cases} \tag{14.24}$$

or

$$\Phi = v + (\delta\Phi_1 + i\delta\Phi_2) \tag{14.25}$$

Note that, since

$$\begin{cases} \delta\Phi_1 = \Phi_1 - v \\ \delta\Phi_2 = \Phi_2 \end{cases} \tag{14.26}$$

we have

$$\langle \delta\Phi_1 \rangle_{\Phi_0} = \langle \delta\Phi_2 \rangle_{\Phi_0} = 0 \tag{14.27}$$

As expected the expectation of the empty state is back to be zero.

For the quantum relativistic Lagrangian, let us write

$$r_0 \to m^2, \quad u_0 \to \lambda, \quad v^2 = -\frac{m^2}{\lambda} \tag{14.28}$$

$$
\begin{aligned}
\mathcal{L} = &- \frac{1}{2}\partial_\mu(\delta\Phi_1 + i\delta\Phi_2)\partial_\mu(\delta\Phi_1 - i\delta\Phi_2) \\
&- \frac{m^2}{2}(v + \delta\Phi_1 + i\delta\Phi_2)(v + \delta\Phi_1 - i\delta\Phi_2) \\
&- \frac{\lambda}{4}[(v + \delta\Phi_1 + i\delta\Phi_2)(v + \delta\Phi_1 - i\delta\Phi_2)]^2 \\
= &- \frac{1}{2}(\partial_\mu\delta\Phi_1\partial^\mu\delta\Phi_1) - \frac{1}{2}(\partial_\mu\delta\Phi_2\partial^\mu\delta\Phi_2) \\
&- \frac{m^2}{2}\left(v^2 + 2v\delta\Phi_1 + \delta\Phi_1^2 + \delta\Phi_2^2\right) \\
&- \frac{\lambda}{4}\left(v^2 + 2v\delta\Phi_1 + \delta\Phi_1^2 + \delta\Phi_2^2\right)^2
\end{aligned}
\tag{14.29}
$$

Since $m^2 = -v^2\lambda$,

$$
\begin{aligned}
\mathcal{L} = &- \frac{1}{2}(\partial_\mu\delta\Phi_1\partial^\mu\delta\Phi_1) - \frac{1}{2}(\partial_\mu\delta\Phi_2\partial^\mu\delta\Phi_2) \\
&+ \frac{\lambda v^2}{2}\left(v^2 + 2v\delta\Phi_1 + \delta\Phi_1^2 + \delta\Phi_2^2\right) \\
&- \frac{\lambda}{4}\Big(v^4 + 4v^2\delta\Phi_1^2 + \left(\delta\Phi_1^2 + \delta\Phi_2^2\right)^2 4v^3\delta\Phi_1 + 2v^2\left(\delta\Phi_1^2 + \delta\Phi_2^2\right) + 4v\delta\Phi_1\left(\delta\Phi_1^2 + \delta\Phi_2^2\right)\Big)
\end{aligned}
\tag{14.30}
$$

Neglecting the constant terms in v

$$
\begin{aligned}
\mathcal{L}(\delta\Phi_1, \delta\Phi_2) = &- \frac{1}{2}(\partial_\mu\delta\Phi_1)^2 - \frac{1}{2}(\partial_\mu\delta\Phi_2)^2 \\
&- \lambda v^2\delta\Phi_1^2 - v\lambda\delta\Phi_1\Big((\delta\Phi_1)^2 + (\delta\Phi_2)^2\Big) \\
&- \frac{\lambda}{4}\Big((\delta\Phi_1)^2 + (\delta\Phi_2)^2\Big)^2
\end{aligned}
\tag{14.31}
$$

Remark. The term $-\lambda v^2\delta\Phi_1^2$ indicates that the field $\delta\Phi_1$ (related to the transversal fluctuations) has a null empty state ($\langle\delta\Phi_1\rangle = 0$) and a mass M such that:

$$
M^2 = 2\lambda v^2 = -2r_0 \tag{14.32}
$$

Therefore, it represents a real, massive, mesonic scalar field that is physically accettable.

However, $\mathcal{L}$ is not any more invariant under the transformation $\delta\Phi_1 \to -\delta\Phi_1$.

Remark. The field $\delta\Phi_2$ has no mass! It describes the fluctuations along the circle where the potential V is in its minimum which imples no dynamical inertia, that implies no mass!

So, starting with one complex scalar field $\Phi(\vec{\mathbf{x}})$ having mass m, when $m^2 < 0$one gets a real scalar field $\delta\Phi_1$ with mass $M = \sqrt{-2m^2}$ and a second scalar field $\delta\Phi_2$ that is massless. This is called the *Goldstone boson.*

Theorem 14.3.1. *If a continuous symmetry is spontaneously broken and there are no long range interactions, exists an elementary eccitation with zero momentum or particle of zero mass called Goldstone boson.*

More generally, let $\mathcal{P}$ be a subgroup of $\mathcal{G}$. If $\mathcal{G}$ has N indipendent generators and $\mathcal{P}$ has M indipendent generators, if $\mathcal{P}$ is the new (lower) symmetry, therefore exist $N - M$ Goldstone bosons.

In the previous case $\mathcal{G} = U(1) \Rightarrow N = 1$ whereas $M = 0$ (we have chosen a specific minimum).

Example 32. XY model in statistical mechanics:

- $\delta\Phi_1$: fluctuation of the modulus of m.
- $\delta\Phi_2$: fluctuations of the spin directions $\Rightarrow$ spin waves.

Remark. In particle physics the presence of Goldstone bosons brings a serious problem in field theory since the corresponding particles are not observed!

Higgs-Englert-Brout (1964)

Higgs mechanism gives back the mass to the Goldstone particles. The basic idea is that the Goldstone theorem that works for a continuous global symmetry it can fail for loacl gauge theories!

14.4 Spontaneous symmetry breaking in Gauge symmetries

Statistical mechanics, Gl model for superconductors in presence of a magnetic field (*Meissner effect*, i.e. the magnetic induction $\vec{\mathbf{B}} = 0$ inside the superconductor).

$$\mathcal{H}(\Phi) = \int \mathrm{d}^D\vec{\mathbf{x}} \left[\frac{1}{2}B^2 + \left|\left(\vec{\boldsymbol{\nabla}} - 2i\vec{\mathbf{A}}\right)\Phi\right|^2\right] + \frac{r_0}{2}\Phi^*\Phi + \frac{u_0}{4}(\Phi^*\Phi)^2 - \vec{\mathbf{B}}\cdot\vec{\mathbf{H}} \tag{14.33}$$

where $\frac{B^2}{2}$ is the energy of the magnetic field $\vec{\mathbf{B}}$ and $\vec{\boldsymbol{\nabla}} \to \left[\vec{\boldsymbol{\nabla}} + iq\vec{\mathbf{A}}\right]$ is the minimal coupling. Consider $\vec{\mathbf{H}}$ the external magnetic field

$$\vec{\mathbf{B}} = \vec{\mathbf{H}} + \vec{\mathbf{M}} \tag{14.34}$$

is the induction field.

Normal conductor corresponding to $\Phi_0 = 0$, that implies $\vec{\mathbf{B}} = \vec{\mathbf{H}}$. For a superconfuctor we have $\Phi \neq 0$, a spontaneous symmetry breaking.

14.5 Fiedl theory analog

Scalar charged mesonic fields selfinteracting and in presence of an electromagnetic field with potential quadrivector $A_\mu(\vec{\mathbf{x}})$.

$$\partial_\mu \to D_\mu = [\partial_\mu + iqA_\mu] \tag{14.35}$$

$$\Phi = \frac{1}{\sqrt{2}}(\Phi_1 + i\Phi_2), \quad \Phi^* = \frac{1}{\sqrt{2}}(\Phi_1 - i\Phi_2) \tag{14.36}$$

Remark. Because of the presence of $A_\mu(\vec{\mathbf{x}})$, we should consider a theory that satisfies symmetry $U(1)$ locally! The transformations are

$$\begin{cases} \Phi(\vec{\mathbf{x}}) \to e^{i\alpha(\vec{\mathbf{x}})}\Phi(\vec{\mathbf{x}}) \\ \Phi^*(\vec{\mathbf{x}}) \to e^{-i\alpha(\vec{\mathbf{x}})}\Phi(\vec{\mathbf{x}}) \end{cases} \tag{14.37}$$

hence, exists $A_\mu(\vec{\mathbf{x}})$ interacting with $\Phi(\vec{\mathbf{x}})$.

The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}(\vec{\mathbf{x}})F^{\mu\nu}(\vec{\mathbf{x}}) + (D_\mu\Phi(\vec{\mathbf{x}}))^*(D_\mu\Phi(\vec{\mathbf{x}})) - V(\Phi, \Phi^*) \tag{14.38}$$

where

$$D_\mu \Phi = (\partial_\mu + iqA_\mu)\Phi \qquad \text{Gauge-covariant derivative} \tag{14.39a}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \qquad \text{field strength tensor} \tag{14.39b}$$

$$V(\Phi, \Phi^*) = \frac{m^2}{2}\Phi\Phi^* + \frac{\lambda}{4}(\Phi\Phi^*)^2 \tag{14.40}$$

- Case $m^2 > 0$: the minimum is in $\Phi = 0$.
- Case $m^2 < 0$: the minimum is in $\Phi = \sqrt{-\frac{m^2}{\lambda}} \equiv v$ (circle of radius $|\Phi| = v$).

Let us choose the state

$$\bar{\Phi}_1 = v, \quad \bar{\Phi}_2 = 0 \tag{14.41}$$

and consider

$$\Phi(x) = (v + \delta\Phi_1) + i\delta\Phi_2 \tag{14.42}$$

Inserting in the Lagrangian and keeping in mind $-m^2 = v^2\lambda$

$$\begin{aligned}\mathcal{L} = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(\partial_\mu\delta\Phi_1)^2 + \frac{1}{2}(\partial_\mu\delta\Phi_2)^2 \\ &- \lambda v^2\delta\Phi_1^2 + q^2v^2A_\mu A^\mu - qvA^\mu\partial_\mu\delta\Phi_2 + \text{higher order terms}\end{aligned} \tag{14.43}$$

1. Term $\lambda v^2\delta\Phi_1^2$: the field $\delta\Phi_1$ is massive with mass $m = v\sqrt{2}$ (Higgs boson).
2. Term $q^2v^2A_\mu A^\mu$: this terms means that the Gause boson A_μ, the photon, has got a mass

   $$M_A = 2qv \tag{14.44}$$

   Remark. Since now A_μ is massive, it has three indipendent polarization states.
3. Term $qvA^\mu\partial_\mu\delta\Phi_2$: the field $\delta\Phi_2$ is not massive (no term $\propto \delta\Phi_2^2$) and is mixed with A_μ. Dynamically this means that a propagating photon can transform itself into a field $\delta\Phi_2$ (photon r Goldstone boson).

Since $\delta\Phi_2$ does not seem to be a physical field it should be eliminated by a Gauge transformation.

Indeed a gauge transformation is also characterized by the transformation

$$A_\mu(\vec{\mathbf{x}}) \to A_\mu(\vec{\mathbf{x}}) + \frac{1}{qv}\partial_\mu\delta\Phi_2(x) \tag{14.45}$$

Inserting (14.45) in the Lagrangian, we eliminate the mixed term $qvA^\mu\partial_\mu\delta\Phi_2$ and $\frac{1}{2}(\partial_\mu\delta\Phi_2)^2$.

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(\partial_\mu\delta\Phi_1)^2 - \lambda v^2\delta\Phi_1 + q^2v^2A_\mu A^\mu + \text{higher order terms} \tag{14.46}$$

Remark. Among the higher order terms there are

- $\propto \delta\Phi_1 A_\mu A^\mu$.
- $\propto \delta\Phi_1^2 A_\mu A^\mu$.

The new Lagrangian contains two fields: one is a massive photon with spin 1 and the second field $\delta\Phi_1$ is massive too, but has spin 0 (scalar).

The Goldstone boson has been "laten" by the Gauge boson that now is massive!

The mechanics trough which the gauge boson becomes massive is the so called *Higgs mechanism.*

In summary: (according to the degrees of freedom)

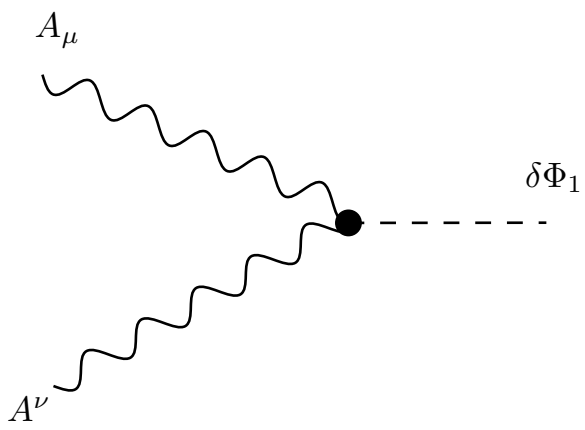


Figure 14.5: Description.

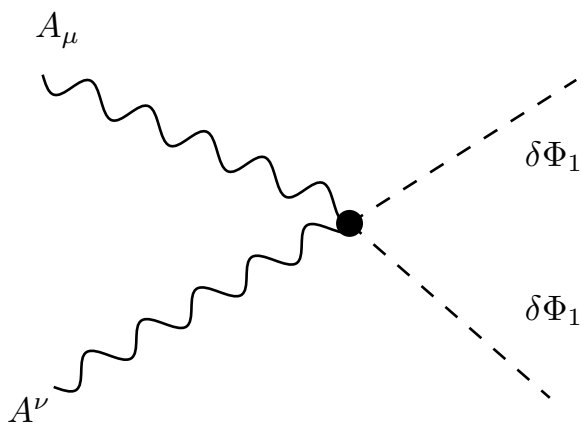


Figure 14.6: Description.

- Global $U(1) \Rightarrow$ 2 scalar fields with mass 1+1 $\overset{\text{symmetry breaking}}{\Longrightarrow}$ 1 scalar field massive and 1 scalar field not massive 1+1.

- Local gauge $U(1) \Rightarrow$ 2 massive scalar fields + 1 photon (2+2) two polarizations $\overset{\text{symmetry breaking}}{\Longrightarrow}$ 1 massive scalar field, 1 massive photon 1 + 3 polarization.

Remark. The presence of the massive photon $m_A^2 = q^2 v^2$, $q = 2l$ in superconductivity, gives rise to the experimental drop

$$B(x) = B(0) \exp\Big(-\frac{x}{l}\Big) \tag{14.47}$$

inside the system.

- We cannot introduce by hand a massive photon i.e. a term like $\frac{1}{2} m_A^2 A_\mu A^\mu$ in the Lagrangian because we would violate explicitely the gauge symmetry!

- The Lagrangian is gauge invariant.

- Symmetry breaking occurs at the level of the vacuum state.

- A gauge symmetry that is explicitely broken is not renormalizable.

14.5.1 Non abelian gauge theories

Example 33. Electro-weak interactions theory (Glashow-Weinberg-Solam) (theory of leptons). Lagrangian has $\underbrace{SU(2)}_{\substack{\text{weak}\\ \text{interactions}}} \times \underbrace{U(1)}_{\text{electromagnetian}}$

Example 34 (Quantum chromodynamic (quarks+gluons))**.**

In this case, one has a term that is $SU(3)$ invariant + the GWS lagrangian with symmetry $SU(2) \times U(1)$, implies

$$SU(3) \times SU(2) \times U(1) \tag{14.48}$$

Because of the groups $SU(2)$ and $SU(3)$ the symmetries above ar not abelian. (For example in $SU(2)$ two matrices $U(\alpha)$ and $U(\beta)$ do not commute in general).

14.6 Extension of Higgs mechanism to non abelian theories

14.6.1 GWS model

Complex field $SU(2)$

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix} = \begin{pmatrix} \Phi_a(\vec{\mathbf{x}}) \\ \Phi_b(\vec{\mathbf{x}}) \end{pmatrix} \tag{14.49}$$

where Φ_a, Φ_b are complex fields.

Gauge transformation $SU(2) \times U(1)$:

$$\begin{pmatrix} \Phi_a(\vec{\mathbf{x}}) \\ \Phi_b(\vec{\mathbf{x}}) \end{pmatrix} \to e^{\frac{i}{2}\alpha_0(\vec{\mathbf{x}})} e^{\frac{i}{2}\vec{\tau}\cdot\vec{\alpha}(\vec{\mathbf{x}})} \begin{pmatrix} \Phi_a(\vec{\mathbf{x}}) \\ \Phi_b(\vec{\mathbf{x}}) \end{pmatrix} \tag{14.50}$$

where $\vec{\tau}$ are Pauli matrices, $\alpha_0, \alpha_1, \alpha_2, \alpha 3$ are four real functions (4 vectorial mesons).

$$\vec{\alpha}(\vec{\mathbf{x}}) \to W_\mu^a(\vec{\mathbf{x}}) = \left(W_\mu^{(1)}(\vec{\mathbf{x}}), W_\mu^{(2)}(\vec{\mathbf{x}}), W_\mu^{(3)}(\vec{\mathbf{x}})\right) \tag{14.51}$$

The scalar gauge field is

$$\alpha_0(\vec{\mathbf{x}}) \to B_\mu(\vec{\mathbf{x}}) \tag{14.52}$$

with B_μ is a linear combination of A_μ and $W_\mu^{(3)}$.

Lagrangian:

$$\mathcal{L} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - \mu^2\Phi^*\Phi - \lambda(\Phi^*\Phi)^2 - \frac{1}{4}b^{\mu\nu}b_{\mu\nu} - \frac{1}{4}f_a^{\mu\nu}f_{\mu\nu}^a \tag{14.53}$$

$$D_\mu \to \partial_\mu \frac{1}{2} ig\tau^a W_\mu^a - \frac{i}{2}g' B_\mu \tag{14.54}$$

$$f_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g\varepsilon^{abc} W_\mu^b W_\nu^a \tag{14.55}$$

$$b_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{14.56}$$

$$W_\mu^a \to W_\mu^a - \varepsilon^{abc}\alpha_b(\vec{\mathbf{x}})W_\mu^c(\vec{\mathbf{x}}) + \frac{1}{g}\partial_\mu\alpha^a(\vec{\mathbf{x}}) \tag{14.57a}$$

$$B_\mu \to B_\mu + \frac{1}{g'}\frac{\partial\alpha_0}{\partial x_\mu} \tag{14.57b}$$

$$\nu \sim \Phi_1^2 + \Phi_2^2 + \Phi_3^2 + \Phi_4^2 = v^2 \tag{14.58}$$

Choosing the direction on the sphere in $\mathbb{R}^4$, 3 symmetries are broken a 3 Goldstone bosons.

14.6.2 Higgs mechanism

Higgs scalar field

$$\delta\Phi = \begin{pmatrix} \Phi^+ \\ \Phi_0 \end{pmatrix} \tag{14.59}$$

such that

$$\langle 0| \Phi |0\rangle = \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{14.60}$$

$$\Rightarrow \mathcal{L}_{Higgs} = \frac{1}{2}(g\nu)^2 W_\mu^+ W^{-\mu} + \frac{1}{2}v^2\Big(gW_\mu^{(3)} - g'B_\mu\Big)^2 \tag{14.61}$$

where

$$W_\mu^{(1)} = \frac{1}{\sqrt{2}}(W_\mu^+ + W_\mu^-) \tag{14.62a}$$

$$W_\mu^{(2)} = \frac{1}{\sqrt{2}}(W_\mu^+ - W_\mu^-) \tag{14.62b}$$

Mass of the W^+ particle and its antiparticle

$$M_W^2 = \frac{1}{2}(gv)^2 \tag{14.63}$$

The 2nd term is a linear combination of W_μ^3 and B_μ which corresponds to Z^0, the field for a third weak gauge boson.

To make Z_μ^0 and A_μ orthogonal we should consider

$$A_\mu = (\cos\theta_W)B_\mu + (\sin\theta_W)W_\mu^3 \tag{14.64a}$$

$$Z_\mu^0 = (-\sin\theta_W)B_\mu + (\cos\theta_W)W_\mu^3 \tag{14.64b}$$

where θ_W is the Weiberg angle:

$$\tan\theta_W = \frac{g'}{g} \tag{14.65}$$

$$M_{Z^0}^2 = \frac{1}{2}\left(\frac{vg}{\cos\theta_W}\right)^2 = \frac{Mw^2}{\cos^2\theta_W} \tag{14.66}$$

Conclusions

Bibliography

[1] Luca Peliti *Statistical Mechanics in a Nutshell.*

[2] Wikipedia `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).*