Cluster expansion methods in rigorous statistical mechanics

Aldo Procacci

October 2, 2016

Contents

Part I

Classical Continuous Systems

Chapter 1

Ensembles in Continuous systems

1.1 The Hamiltonian and the equations of motion

In this first part of the book we will deal with a system made by a large number N of continuous particles enclosed in a box $\Lambda \subset \mathbb{R}^d$ (we will assume Λ to be cube and denote $|\Lambda|$ its volume) performing a motion according to the laws of classical mechanics. "Large number of particles" in physics lingo means typically $N \approx 10^{23}$.
We will restrict our discussion to systems composed of identical particles with no internal structure, i.e. just "point" particles with a given mass m. The position of the i^{th} particle in the box Λ at a given time t is given by a d-component coordinate vector, denoted by $x_i = x_i(t)$ respect to some system of orthogonal axis.
The momentum of the i^{th} particle at time t is also given by a d-component vector denoted by $p_i = p_i(t)$ which is directly related to the velocity of the particles, i.e. if m is the mass of the particle, then $p_i = m\frac{dx_i}{dt}$.
In principle, the laws of mechanics permit to know the evolution of such a system in time, i.e. these laws should be able to determine which positions $x_i = x(t)$ and momenta $p_i = p_i(t)$ the particles in the system will have in the future and had in the past, provided one knows the position of the particles $x_i^0 = x(t_0)$ and the momenta of $p_i^0 = p_i(t_0)$ at a given time t_0,
As a matter of fact, the time evolution of such system is described by a real valued function $H(p_1, \ldots, p_N, x_1, \ldots, x_N)$ of particle positions and momenta (hence a function of $2dN$ real variables) called the *Hamiltonian*. In the case of isolated systems, this function H is assumed to have the form

$$H(p_1, \ldots, p_N, x_1, \ldots, x_N) \;=\; \sum_{i=1}^{N} \frac{p_i^2}{2m} + U(x_1, \ldots x_N) \tag{1.1}$$

The term $\sum_{i=1}^{N} \frac{p_i^2}{2m}$ is called the kinetic energy of the system, while the term

$U(x_1, \dots x_N)$ is the potential energy. Since we are assuming that particles has to be enclosed in a box Λ we still have to restrict $x_i \in \Lambda$ for all $i = 1, 2, \dots, N$, while no restriction is imposed on p_i, i.e. the momentum (and hence the velocity) of particles can be arbitrarily large.

Once the Hamiltonian of the system is given, one could in principle solve the system of $2dN$ differential equations

$$\begin{cases} \frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i} \end{cases} \qquad i = 1, \dots, N \tag{1.2}$$

where $\partial/\partial x_i$ and $\partial/\partial p_i$ are d-dimensional gradients.
This is a system of $2dN$ first order differential equations. The solution are $2dN$ functions $x_i(t)$, $p_i(t)$, $i = 1, \dots, N$, given the positions $x_i(t_0)$ and momenta $p_i(t_0)$ at some initial time $t = t_0$.
It is convenient to introduce a $2dN$-dimensional space $\Gamma_N(\Lambda)$, called the phase space of the system (with N particles) whose points are determined by the coordinates $(\boldsymbol{q}, \boldsymbol{p})$ with $\boldsymbol{q} = (x_1, \dots, x_N)$ and $\boldsymbol{p} = (p_1, \dots, p_N)$ ($\boldsymbol{q}$ and $\boldsymbol{p}$ are both dN vectors!) with the further condition that $x_i \in \Lambda$ for all $i = 1, 2, \dots, N$. A point $(\boldsymbol{q}, \boldsymbol{p})$ in the phase space of the system is called a *microstate* of the system. With this notations the evolution of the system during time can be interpreted as the evolution of the a point in the "plane" $\boldsymbol{q}, \boldsymbol{p}$ (actually a $2dN$ dimension space, the phase space $\Gamma_N(\Lambda)$).
We finally want to remark that, by (1.1) (isolated system), the value of the Hamiltonian H is constant during the time evolution of the system governed by the Hamilton equations (1.2). Namely $H(\boldsymbol{q}(t), \boldsymbol{p}(t)) = H(\boldsymbol{q}(0), \boldsymbol{p}(0)) = E$. To see this, just calculate the total derivative of H respect to the time using equations (1.2). This constant E of motion is called *energy* of the system. Thus the trajectory of the point $\boldsymbol{q}, \boldsymbol{p}$ in the phase space $\Gamma_N(\Lambda)$ occurs in the surface $H(\boldsymbol{p}, \boldsymbol{q}) = E$.

1.2 Gibbsian ensembles

It substantially meaningless to look for the solution of system (1.2).
First, there is a technical reason. Namely the system contains an enormous number of equation ($\approx 10^{23}$) which in general are coupled (depending on the structure of U), so that it is practically an impossible task to find the solution. However, even supposing that some very powerful entity would give us the solution, this extremely detailed description (a microscopic description) would not be useful to describe the macroscopic properties of such a system. Macroscopic properties which appear to us as the laws of thermodynamic are due presumably by some mean effects of such large systems and can in general be described in terms of very few parameters, e.g. temperature, volume, pressure, etc.. Hence we have no means and also no desire to know the microscopic state of the system at every instant (i.e. to know the functions $\boldsymbol{q}(t), \boldsymbol{p}(t)$). We thus shall adopt a statistical point of view in order to describe the system.

We know a priori some macroscopic properties of the system, e.g. an isolated system occupies the volume Λ, has N particles and has a fixed energy E.
We further know that macroscopic systems, if not perturbated from the exterior, tends to stay in a situation of macroscopic (or thermodynamic) equilibrium (i.e. a "static" situation), in which the values of some thermodynamic parameters (e.g. pressure, temperature, etc.) are well defined and fixed and do not change in time. Of course, in a system at the thermodynamic equilibrium, the situation at the microscopic level is desperately far form a static one. Particles in a gas at the equilibrium do in general complicated and crazy motions all the time, nevertheless nothing seems to happens as time goes by at the macroscopic level. So the thermodynamic equilibrium of a system must be the effect os some mean behavior at the microscopic level. This "static" mean macroscopic behavior of systems composed by a large number of particles must be produced in some way by the microscopic interactions between particle and by the law of mechanics.
Adopting the statistical mechanics point of view to describe a macroscopic system at equilibrium means that we renounce to understand how and why a system reach the thermodynamic equilibrium starting from the microscopic level, and we just assume that, at the thermodynamic equilibrium (characterized by some thermodynamic parameters), the system could be find in *any* microscopic state within a certain suitable set of microstates compatible with the fixed thermodynamic parameters. This is the so called *ergodic hypothesis*. Namely, all microstates are equivalent. We will also assume that each of this micro-state can occur with a given probability. Of course, in order to have some hope that such a point of view will work, we need to treat really "macroscopic systems". So values such N and V must be always though as very large values (i.e. close to ∞).
The statistical description of the macroscopic properties of the system at equilibrium (and in particular the laws of thermodynamic) is done in two step.

Step 1: *fixing the Gibbsian ensemble (or the space of configurations).* We choose the phase space Γ_e and we assume that the system can be found in any microstate $(\boldsymbol{q}, \boldsymbol{p}) \in \Gamma_e$. This set Γ_e has to interpreted as the set of all microstates accessible by the system and it is called the *Gibbsian ensemble* or *the space of configurations* of the system. We will see later that several choices are also possible for Γ_e. We will then think not on a single system, but in an infinite number of mental copies of the same system, one copy for each element of Γ_e.

Step 2: *fixing the Gibbs measure in the Gibbisan ensemble*
We choose a function $\rho(\boldsymbol{p}, \boldsymbol{q})$ in Γ_e which will represent the probability density in the Gibbsian ensemble, Namely, $\rho(\boldsymbol{p}, \boldsymbol{q})$ is a function such that

$$\int_{\Gamma_e} \rho(\boldsymbol{p}, \boldsymbol{q}) d\boldsymbol{p} d\boldsymbol{q} = 1$$

and $d\mu(\boldsymbol{p}, \boldsymbol{q}) = \rho(\boldsymbol{p}, \boldsymbol{q}) d\boldsymbol{p} d\boldsymbol{q}$ represents the probability to find the system in a microstate (or in the configuration) contained in an infinitesimal volume $d\boldsymbol{p} d\boldsymbol{q}$ around the point $(\boldsymbol{p}, \boldsymbol{q}) \in \Gamma_e$ where $d\boldsymbol{p}\ d\boldsymbol{q}$ is the usual Lebesgue measure in

$\mathbb{R}^{2dN}$. The measure $\mu(\boldsymbol{p},\boldsymbol{q})$ defined in Γ_e is called the *Gibbs measure* of the system.

Once a Gibbsian ensemble and a Gibbs measure are established, one can begin to do statistic in order to describe the macroscopic state of the system. When we look at a macroscopic system described via certain Gibbs ensemble we do not know in which microstate the system is at a given instant. All we know is that its microscopic state must be one of the microstates of the space configuration Γ_e with probability density given by the Gibbs measure $d\mu$.
For example, suppose that $f(\boldsymbol{p},\boldsymbol{q})$ is a measurable function respect to the Gibbs measure $d\mu$, such as the energy, the kinetic energy per particle, potential energy etc. Then we can calculate its mean value in the Gibbsian ensemble that we have chosen by the formula:

$$\langle f\rangle \;=\; \int f(\boldsymbol{p},\boldsymbol{q})\rho(\boldsymbol{p},\boldsymbol{q})d\boldsymbol{p}d\boldsymbol{q}$$

We also recall the concept of mean relative square fluctuation of f (a.k.a. standard deviation) denoted by σ_f. This quantity measures how spread is the probability distribution of $f(\boldsymbol{p},\boldsymbol{q})$ around its mean value. It is defined as

$$\sigma_f \;=\; \frac{\langle (f-\langle f\rangle)^2\rangle}{\langle f\rangle^2} \;=\; \frac{\langle f^2\rangle-\langle f\rangle^2}{\langle f\rangle^2} \tag{1.3}$$

1.3 The Micro-Canonical ensemble

There are different possible choices for the ensembles, depending on the different macroscopic situation of the system.
We start defining the micro-Canonical ensemble which is used to describe perfectly isolated systems. Hence we suppose our system totally isolated from the outside, the N particles are constrained to stay in the box Λ, and they do not exchange energy with the outside, so that the system has a given energy E, occupies a given volume $|\Lambda|$ and has a fixed number of particles N.
Thus, for such a system, we can naturally say that the space of configuration Γ_{mc} is the set of points $\boldsymbol{p},\boldsymbol{q}$ (with $\boldsymbol{p}\in\mathbb{R}^{dN}$ and $\boldsymbol{q}\in\Lambda^N$) with energy between a given value E and $E+\Delta E$ (where ΔE can be interpreted as the experimental error in the measure of the energy E). We have

$$\Gamma_{mc} \;=\; \{(\boldsymbol{p},\boldsymbol{q}) : \boldsymbol{p}\in\mathbb{R}^{dN},\quad q\in\Lambda \quad\text{and}\quad E<H(\boldsymbol{p},\boldsymbol{q})<E+\Delta E\}$$

We now choose the probability measure in such way that any microstate in the set of configurations above is equally probable, i.e. there is no reason to assign different probability to different microstate. This quite drastic hypothesis is the so-called *postulate of equal a priori distribution*, which is just a different formulation of the ergodic hypothesis. Hence

$$\rho(\boldsymbol{p},\boldsymbol{q}) \;=\; \begin{cases} [\Psi_\Lambda(E,N)]^{-1} & \text{if } (\boldsymbol{q},\boldsymbol{p})\in\Gamma_{mc} \\ \\ 0 & \text{otherwise} \end{cases}$$

where

$$\Psi_\Lambda(E,N) \;=\; \int_{E<H(\boldsymbol{p},\boldsymbol{q})<E+\Delta E} d\boldsymbol{p}\, d\boldsymbol{q} \tag{1.4}$$

is the $2dN$-dimensional "volume" in the phase space occupied by the space of configuration of the micro-canonical ensemble. $\Psi_\Lambda(E,N)$ is generally called the *partition function* of the system in the Micro-Canonical ensemble. The link between the Micro-Canonical ensemble and the thermodynamic is obtained via the definition of the thermodynamic entropy of the system by

$$S_\Lambda(E,N) \;=\; k\ln\left[\frac{1}{|\delta|}\Psi_\Lambda(E,N)\right] \tag{1.5}$$

where k is the Boltzmann's constant and $|\delta|$ is the volume of some elementary phase cell δ in the phase space, so that the pure number $|\delta|^{-1}\Psi_\Lambda(E,N)$ is the number of such cells in the configuration space.
It may seem that the value of this constant $|\delta|$ can be somewhat arbitrary, since it depends on our measure instruments, and it could be done as small as we please by improving our measure techniques. But experiments says that this constant is fixed at the value h^{dN} where h is the Plank constant . So hereafter we will assume, unless differently specificated, that $|\delta|$ is set at the value of the Plank constant. It is important to stress that the presence of this constant in the definition of the entropy has a very deep physical meaning. Actually, it is a first clue of the quantum mechanics nature of particle systems: things go as if the position $\boldsymbol{p},\boldsymbol{q}$ of a micro-state in the phase space could not be known exactly and one can just say that the micro-state is in a small cube $d\boldsymbol{p}d\boldsymbol{q}$ centered at $\boldsymbol{p},\boldsymbol{q}$ of the space phase with volume h^{dN}. This is actually the Heisenberg indetermination principle.
The definition (1.5) gives a beautiful probabilistic interpretation of the second law of thermodynamic. A macroscopic system at the equilibrium will tend to stay in a state of maximal entropy, namely, by (1.5), in the most probable macroscopic state, i.e. a macroscopic state with thermodynamics parameters fixed in such way that this state corresponds to the largest number of microstates. Namely, at equilibrium, the quantity $\Psi_\Lambda(E,N)$ should expect to reach a maximum value. Thus the entropy (which by definition is just the logarithm of the number of micro-states of a macro-state) is also expect to be maximum at equilibrium. So the second law of thermodynamics stating that the entropy of an isolated system always increases means in term of statistical mechanics that the systems tends to evolve to macrostates which are more probable, i.e. those with maximum number of microstates.
It is also interesting to check that entropy (1.5) is a so called "extensive" quantity. namely, if the macroscopic system is composed by two macroscopic subsystems whose entropies are, respectively S_1 and S_2, the entropy of the total system must be S_1+S_2.

1.4 The entropy is additive. An euristic discussion

Suppose thus to consider a system made with two subsystems, one living in a phase space Γ_1 with coordinates $\boldsymbol{p}_1,\boldsymbol{q}_1$ occupying the volume Λ_1 and described

by the Hamiltonian $H_1(\boldsymbol{p}_1, \boldsymbol{q}_1)$ and the other in a phase space Γ_2 with coordinate $\boldsymbol{p}_2, \boldsymbol{q}_2$, occupying the volume Λ_2 and described by the Hamiltonian $H_2(\boldsymbol{p}_2, \boldsymbol{q}_2)$. We also suppose that systems are isolated from each other. Consider first the micro-canonical ensemble for each subsystem taken alone. The energy of the first system will stay in a interval say $(E_1, E_1 + \Delta)$ while the second system will have an energy in $(E_2, E_2 + \Delta)$. The entropies of the subsystems will be respectively $S(E_1) = k \ln \Psi_1(E_1)$ and $S(E_2) = k \ln \Psi_2(E_2)$, where $\Psi_1(E_1)$ and $\Psi_2(E_2)$ are the volumes occupied by the two ensembles in their respective phase spaces Γ_1 and Γ_2. Consider now the micro-canonical ensemble of the total system made by the two subsystems, The composite system lives in a phase space $\Gamma_1 \times \Gamma_2$ with coordinates p_1, q_1, p_2, q_2 occupying the volume $V_1 + V_2$ and described by the Hamiltonian $H_1(p_1, q_1) + H_2(p_2, q_2)$ (system are supposed isolated one from each other). Let the total energy be in the interval say $(E, E + 2\Delta)$ ($\Delta \ll E$). This ensemble contains all the micro-states of the composite system such that:

a) N_1 particles with momenta and coordinates p_1, q_1 are in the volume V_1

b) N_2 particles with momenta and coordinates p_2, q_2 are in the volume V_2

c) The energy E_1 and E_2 of the subsystems have values satisfying the condition

$$E < E_1 + E_2 < E + 2\Delta \tag{1.6}$$

We want to calculate the partition function $\Psi(E)$ of the composite system. Clearly $\Psi_1(E_1)\Psi_2(E_2)$ is the volume in the composite phase space Γ with coordinate $(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{q}_1, \boldsymbol{q}_2)$ that corresponds to conditions a) and b) with first system at energy E_1 and second system at energy E_2 such that $E_1 + E_2 \in (E, E + 2\Delta)$. Then

$$\Psi(E) \;=\; \sum_{\substack{E_1, E_2 \\ E < E_1 + E_2 < E + 2\Delta}} \Psi_1(E_1)\Psi_2(E_2)$$

Since E_1 and E_2 are possible values of $H(\boldsymbol{p}_1, \boldsymbol{q}_1)$ and $H(\boldsymbol{p}_2, \boldsymbol{q}_2)$, suppose that $H(p_1, q_1)$ and $H(\boldsymbol{p}_2, \boldsymbol{q}_2)$ are bounded below (as it will be always the case, see later) and for simplicity let the joint lower bound be equal to 0. Hence E_1 and E_2 both varies in the interval $[0, E]$. Suppose also that E_1 and E_2 take discrete values $E_i = 0, \Delta, 2\Delta, ...$ so that in the interval $(0, E)$ there are E/Δ of such intervals. Then

$$\Psi(E) \;=\; \sum_{i=1}^{E/\Delta} \Psi_1(E_i)\Psi_2(E - E_i) \tag{1.7}$$

The entropy of the total system of $N = N_1 + N_2$ particles, volume $\Lambda = \Lambda_1 \cup \Lambda_2$ and energy E is given by

$$S_\Lambda(E, N) \;=\; k \ln \left[\sum_{i=1}^{E/\Delta} \Psi_1(E_i)\Psi_2(E - E_i) \right]$$

As subsystems are supposed macroscopic ($N_1 \to \infty$ and $N_2 \to \infty$) is easy to see that a single term in the sum (1.7) will dominate. Sum (1.7) is a sum of positive

terms, let the largest of such terms be $\Psi_1(\bar{E}_1)\Psi_2(\bar{E}_2)$ with $\bar{E}_1+\bar{E}_2 = E$. Then we have the obvious inequalities

$$\Psi_1(\bar{E}_1)\Psi_2(\bar{E}_2) \;\le\; \Psi(E) \;\le\; \frac{E}{\Delta}\Psi_1(\bar{E}_1)\Psi_2(\bar{E}_2)$$

or

$$k\ln\left[\Psi_1(\bar{E}_1)\Psi_2(\bar{E}_2)\right] \;\le\; S_\Lambda(E,N) \;\le\; k\ln\left[\Psi_1(\bar{E}_1)\Psi_2(\bar{E}_2)\right] + k\ln(E/\Delta) \quad (1.8)$$

We expect, as $N_1 \to \infty$ and $N_2 \to \infty$, that $\Psi_1 \sim C^{N_1}$ and $\Psi_2 \sim C^{N_2}$, thus $\ln\Psi_1 \propto N_1$ and $\ln Z_2 \propto N_2$. and also $E \sim N_1 + N_2$. Hence factor $\ln(E/\Delta)$ goes like $\ln N$ and can be neglected. Namely by this discussion (just a counting argument) we get

$$S_\Lambda(E,N) \;=\; S_{\Lambda_1}(\bar{E}_1,N_1) + S_{\Lambda_2}(\bar{E}_2,N_2) + O(\ln N) \quad (1.9)$$

In other words the entropy is extensive, modulo terms of order $\ln N$. Note that (1.9) also means that the two subsystems has a definite values $\bar{E}_1$ and $\bar{E}_2$ for the energy. Namely $\bar{E}_1$ and $\bar{E}_2$ are the values that maximize the number

$$\Psi_1(E_1)\Psi_2(E_2)$$

under the condition $E_1 + E_2 = E$. Using Lagrange multiplier is easy to check that

$$\left.\frac{\partial\ln\Psi_1(E_1)}{\partial E_1}\right|_{E_1\,=\,\bar{E}_1} \;=\; \left.\frac{\partial\ln\Psi_2(E_2)}{\partial E_2}\right|_{E_2\,=\,\bar{E}_2}$$

or

$$\left.\frac{\partial S(E_1)}{\partial E_1}\right|_{E_1\,=\,\bar{E}_1} \;=\; \left.\frac{\partial S(E_2)}{\partial E_2}\right|_{E_2\,=\,\bar{E}_2}$$

Since thermodynamics tells us that $\frac{\partial S(E,V)}{\partial E} = \frac{1}{T}$, we conclude that the two subsystems choose energy $\bar{E}_1$ and $\bar{E}_2$ in such way to have the same temperature. Thus the temperature in a macroscopic system can be seen as the parameter governing the equilibrium between one part of the system and the other.

1.5 Entropy of the ideal gas

If one can calculate the partition function of the Micro-canonical ensemble, then it is possible to derive the thermodynamic properties of the system. The Micro-Canonical ensemble is difficult to be treated mathematically. As a matter of fact, a direct calculation of the integral in r.h.s. of (1.4) is generally very difficult, since involves integration over complicated surfaces in high dimensions. According to elementary calculus in $\mathbb{R}^n$ we can use a volume integral to calculate the integral (1.4). Volume integrals are easier to deal with than surface integrals. Let $\omega_\Lambda(E,N)$ be the volume of the phase space surrounded by the surface $H(\boldsymbol{p},\boldsymbol{q}) = E$ (with of course the further condition that particles are constrained to stay in Λ). Then

$$\omega_\Lambda(E,N) \;=\; \int_{H(\boldsymbol{p},\boldsymbol{q})\le E} d\boldsymbol{p}\; d\boldsymbol{q} \quad (1.10)$$

Then, for small ΔE

$$\Psi_\Lambda(E,N) \;=\; \omega_\Lambda(E+\Delta E,N)-\omega_\Lambda(E,N)\approx \frac{\partial\omega_\Lambda(E,N)}{\partial E}\Delta E \tag{1.11}$$

Let for example calculate the Micro-Canonical partition function of an ideal gas, i.e. a gas of non interacting particles. with Hamiltonian

$$H \;=\; \sum_{i=1}^{N}\frac{p_i^2}{2m} \tag{1.12}$$

We thus calculate $\omega(N,\Lambda,E)$ when $H(\boldsymbol{p},\boldsymbol{q})$ is given by (1.12). In this case it is absolutely elementary to calculate the integral $\int_\Lambda dq$ which gives just $|\Lambda|$ i.e. the volume occupied by Λ, and therefore we get

$$\begin{aligned}\omega_\Lambda(E,N) \;&=\; \int_{H(\boldsymbol{p},\boldsymbol{q})\le E} d\boldsymbol{p}\; d\boldsymbol{q} \;=\; |\Lambda|^N\int_{\sum_{i=1}^N\frac{p_i^2}{2m}\le E} dp_1\dots dp_N \;=\\ &=\; |\Lambda|^N\int_{\sum_{i=1}^N p_i^2\le 2mE} dp_1\dots dp_N\end{aligned}$$

the last integral is just the volume of a dN dimensional sphere of radius $\sqrt{2mE}$. Let us thus face this geometric problem. The volume of a sphere of given radius R in a n dimensional space is the integral

$$V_n(R) \;=\; \int_{\sum_{i=1}^n x_1^2\le R^2} dx_1\dots dx_n \;=\; C_nR^n$$

In order to find C_n consider the following integral

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-(x_1^2+\dots+x_n^2)}dx_1\dots dx_n \;&=\; \int_{-\infty}^{+\infty}dx_1\dots\int_{-\infty}^{+\infty}dx_n e^{-(x_1^2+\dots+x_n^2)} \;=\\ &=\; \left(\int_{-\infty}^{+\infty}dx e^{-x^2}\right)^n \;=\; \pi^{\frac{n}{2}}\end{aligned}$$

on the other hand, noting that the integrand above depends only on $r=(x_1^2+\dots x_n^2)^{1/2}$, by a transformation to polar coordinates in n dimensions we can express the volume element $dx_1\dots dx_n$ by spherical shells $dV_n(r)=dC_nr^n=nC_nr^{n-1}dr$. Then integral above can also be calculated as

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-(x_1^2+\dots+x_n^2)}dx_1\dots dx_n \;&=\; \int_0^\infty e^{-r^2}dV_n(r) \;=\; nC_n\int_0^\infty r^{n-1}e^{-r^2}dr \;=\\ &=\; \frac{nC_n}{2}\int_0^\infty x^{n/2-1}e^{-x}dx\end{aligned} \tag{1.13}$$

Recall now the definition of the gamma function: for any $z>0$

$$\Gamma(z) \;=\; \int_0^\infty x^{z-1}e^{-x}dx$$

Among properties of the gamma function we recall

$$\Gamma(n) \;=\; (n-1)! \qquad\qquad \text{n positive integer}$$

and

$$z\Gamma(z) \;=\; \Gamma(z+1) \qquad\qquad \text{z} \in \mathbb{R}^+$$

So Gamma function is an extension of factorial in the whole positive real axis. Equation (1.13) thus becomes

$$\frac{nC_n}{2}\Gamma(n/2) \;=\; \pi^{n/2}$$

and consequently of the volume $V_n(R)$ of a sphere on radius R in n dimensions.

$$C_n \;=\; \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}, \qquad\qquad V_n(R) \;=\; \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}R^n$$

Hence we get

$$\omega_\Lambda(E,N) \;=\; \frac{\pi^{3N/2}}{\frac{3N}{2}\Gamma(\frac{3N}{2})}(2mE)^{\frac{3N}{2}}|\Lambda|^N$$

and, by (1.11)

$$\Psi_\Lambda(E,N) \;=\; \frac{\partial\omega_\Lambda(E,N)}{\partial E}\Delta E \;=\; \Delta E\;|\Lambda|^N\;\frac{\pi^{3N/2}}{\Gamma(\frac{3N}{2})}(2m)^{3N/2}E^{\frac{3N}{2}-1}$$

So, recalling definition (1.5), the entropy of an ideal gas is given by

$$S_\Lambda(E,N) \;=\; k\ln\left[\Delta E\;|\Lambda|^N\;\frac{\pi^{3N/2}}{\Gamma(\frac{3N}{2})}(2m/h^2)^{3N/2}E^{\frac{3N}{2-1}}\right]$$

Since N is a very big number, we may write, for $N \gg 1$,

$$E^{\frac{3N}{2}-1} \approx E^{\frac{3N}{2}}, \quad \ln\left[\Gamma\left(\frac{3N}{2}\right)\right] \approx \frac{3N}{2}\left(\ln(3N/2)-1\right), \quad k\ln(\Delta E) = O(1) \approx 0$$

we used the Stirling approximation for the factorial: $n! \approx \frac{n^n}{e^n}$ for large n) thus

$$S_\Lambda(E,N) \;=\; kN\left\{\frac{3}{2}+\ln\left[\;|\Lambda|\;\left(\frac{4m\pi E}{3Nh^2}\right)^{3/2}\right]\right\} \tag{1.14}$$

This equation leads to the correct equation of state for a perfect gas. In fact by definition, the inverse temperature of the system is the derivative of the entropy respect to the energy, and the pressure is the derivative of the entropy respect to the volume $V = |\Lambda|$ times the temperature, i.e.

$$\begin{aligned}\frac{1}{T} \;&=\; \frac{\partial S}{\partial E} \;=\; \frac{3}{2}\frac{Nk}{E} \qquad \text{or} \qquad E \;=\; \frac{3N}{2}kT \\ \frac{p}{T} \;&=\; \frac{\partial S}{\partial V} \;=\; \frac{Nk}{V} \qquad \text{or} \qquad PV \;=\; NkT \end{aligned} \tag{1.15}$$

Nevertheless (1.14) cannot be the correct expression for the entropy of an ideal gas. One can just observe that l.h.s. of (1.14) is not a purely extensive quantity, as the thermodynamic entropy should be. There is some deep mistake in the calculation of the entropy. This can be very well illustrated by the so called Gibbs paradox.

1.6 The Gibbs paradox

We put in ths section $V = |\Lambda|$. Consider thus the entropy of an ideal gas as a function of the temperature T, the volume V and the numer of particles N. By (1.14) and (1.15) we get

$$S(T,V,N) \;=\; kN\left\{\frac{3}{2}+\ln\left[\;V\;\left(\frac{2m\pi kT}{h^2}\right)^{3/2}\right]\right\} \tag{1.16}$$

where, using (1.15), we have posed that $E/N = \frac{3}{2}kT$. Consider now a closed system consisting initially of two adjacent volumes V_A and V_B separated by a wall. The volume A contains an ideal gas with N_A particles, and the volume V_B contains another ideal gas with N_B particles. The two subsystem are kept at the same pressure P and at the same temperature T. The entropy of such system, according to (1.16) is

$$S^i_{\text{total}} \;=\; S(T,V_A,N_A)+S(T,V_B,N_B)$$

If we now remove the wall, the two ideal gases will mix, each occupying the volume $V_A + V_B$. In the new equilibrium situation the entropy is now

$$S^f_{\text{total}} \;=\; S(T,V_A+V_B,N_A)+S(T,V_A+V_B,N_B)$$

The entropy difference between the initial state i and the final state f is, according to (1.16)

$$\Delta S \;=\; S^f_{\text{total}} - S^i_{\text{total}} \;=\; N_A k\ln(1+V_B/V_A)+N_B k\ln(1+V_A/V_B) \tag{1.17}$$

So far everything seems to be fine, since $\Delta S > 0$ as one should expect for this irreversible process (the mixing on two ideal gases). But let us now suppose that the two ideal gases are actually identical. We could repeat the argument and again we will find that the variation of the entropy is $\Delta S > 0$. However this cannot be correct since, after the removal of the wall, no macroscopic changes happens at all in the system. We could put back the wall ad we will return to the initial macroscopic situation. This paradox is clearly related to the fact that we are assuming the particles of the ideal gas as distinguishable. If particle are considered as distinguishable, then also the situation in which two identical perfect gases mixes is a irreversible process. If we put back the wall particles in the volume 1 are not the same particles of the initial situation. The paradox was resolved by Gibbs supposing that identical particles are not distinguishable. With this hypothesis the number of microstates involving N particles should be reduced by a factor $N!$, since there are exactly $N!$ ways to enumerate N identical particles. Hence the correct definition of, e,g. $\omega_\Lambda(E,N)$ should be

$$\omega_\Lambda(E,N) \;=\; \frac{1}{N!}\int_{H(\boldsymbol{p},\boldsymbol{q})\le E} dp\; dq \tag{1.18}$$

Thus, instead of (1.14), the correct entropy of a perfect gas is

$$S_\Lambda(E,N) \;=\; kN\left\{\frac{3}{2}+\ln\left[\;|\Lambda|\;\left(\frac{2m\pi kT}{h^2}\right)^{3/2}\right]\right\} - k\ln N! \tag{1.19}$$

First observe that this new definition of the entropy does not affect the equation of state of the perfect gas, since it differs from (1.14) by a term independent on E and V, so (1.19) leads to the same equations (1.15). and, for $N \gg 1$, by Stirling's formula $\ln N! \approx N \ln N - N$

$$S_\Lambda(E,N) \;=\; kN\left\{\frac{5}{2} + \ln\left[\;\frac{|\Lambda|}{N}\;\left(\frac{2m\pi kT}{h^2}\right)^{3/2}\right]\right\} \tag{1.20}$$

Thus entropy defined by (1.20) of a perfect gas is indeed a purely extensive quantity. Let us also check that (1.20) solve the Gibbs paradox.

Let us star again the argumentation which leads us to the Gibbs paradox with the new temperature dependent entropy (just using that $E = 3NKT/2$)

$$S_\Lambda(E,N) \;=\; kN\left\{\frac{5}{2} + \ln\left[\;\frac{|\Lambda|}{N}\;\left(\frac{2m\pi kT}{h^2}\right)^{3/2}\right]\right\} \tag{1.21}$$

The new variation in entropy is now

$$\Delta S \;=\; k(N_A+N_B)\left\{\frac{5}{2} + \ln\left[\;\frac{V_A+V_B}{N_A+N_B}\;\left(\frac{2m\pi kT}{h^2}\right)^{3/2}\right]\right\} -$$

$$-kN_A\left\{\frac{5}{2} + \ln\left[\;\frac{V_A}{N_A}\;\left(\frac{2m\pi kT}{h^2}\right)^{3/2}\right]\right\} - kN_B\left\{\frac{5}{2} + \ln\left[\;\frac{V_B}{N_B}\;\left(\frac{2m\pi kT}{h^2}\right)^{3/2}\right]\right\} =$$

$$= kN_A \ln\left[\frac{\frac{V_A+V_B}{N_A+N_B}}{\frac{V_A}{N_A}}\right] + kN_B \ln\left[\frac{\frac{V_A+V_B}{N_A+N_B}}{\frac{V_B}{N_B}}\right] \tag{1.22}$$

For two different gases this formula gives something similar to (1.17). But if gas A and gas B are identical, then, since in the initial state and in the final state temperature and pressure ar unchanged, we must have, by (1.15)

$$\frac{V_A}{N_A} \;=\; \frac{V_B}{N_B} \;=\; \frac{V_A+V_B}{N_A+N_B} \;=\; \frac{kP}{T}, \qquad \text{if gas A and gas B are identical} \tag{1.23}$$

(of course also for different gases $\frac{V_A}{N_A} = \frac{V_B}{N_B} = \frac{kT}{P}$ but $\frac{V_A+V_B}{N_A+N_B} \neq \frac{kT}{P}$).

Inserting formula (1.23) in (1.22) we obtain

$$\Delta S \;=\; 0 \qquad \text{if gas A and gas B are identical}$$

The necessity to divide by the factor $N!$ to escape from the Gibbs paradox is a new symptom that classical mechanics is not adequate.

1.7 The Canonical Ensemble

The Micro Canonical Ensemble is suited for isolated systems where natural macroscopic variables are the volume $|\Lambda|$, the number of particles N and the energy E. We now define a new ensemble which is appropriate to describe a system which is not isolated, but it is in thermal equilibrium with a larger

system (the heat reservoir), e.g. a gas kept in a box made by heat conducting walls which is fully immersed in a larger box containing some other gas at a fixed temperature T. Hence this system is constrained to stay in a box Λ with a fixed volume $|\Lambda|$, a fixed number of particles N, at a fixed temperature T, but its energy is no longer fixed, since system is now allowed to exchange energy with the heat reservoir through the walls.

We define the Canonical Ensemble for such a system as follows. The space of configuration of the Canonical Ensemble is

$$\Gamma_c \;=\; \Gamma_N(\Lambda)$$

The probability measure of the Canonical Ensemble is

$$d\mu_c(\boldsymbol{p},\boldsymbol{q}) \;=\; \frac{1}{Z_\Lambda(\beta,N)}\frac{1}{N!}e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}\frac{d\boldsymbol{p}d\boldsymbol{q}}{h^{3N}} \tag{1.24}$$

where $\beta \;=\; (kT)^{-1}$ is a constant proportional to the inverse temperature of the system (k is again the Boltzmann constant) and the normalization constant

$$Z_\Lambda(\beta,N) \;=\; \frac{1}{N!}\int e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}\frac{d\boldsymbol{p}d\boldsymbol{q}}{h^{3N}} \tag{1.25}$$

is the partition function of the system in the canonical ensemble.
Thermodynamics is recovered by the following definition. The thermodynamic function called free energy of the system is obtained in the canonical ensemble by the formula

$$F_\Lambda(\beta,N) \;=\; -\,kT\ln Z_\Lambda(\beta,N) \tag{1.26}$$

We now remark that in (1.25) we can perform for free the integration over momenta.

$$Z_\Lambda(\beta,N) = \frac{1}{h^{3N}N!}\int dp_1\ldots\int dp_N\int_\Lambda dx_1\ldots\int_\Lambda dx_N e^{-\beta H(p_1,\ldots,p_N,x_1,\ldots,x_N)} =$$

$$= \frac{1}{h^{3N}N!}\int dp_1\ldots\int dp_N\int_\Lambda dx_1\ldots\int_\Lambda dx_N e^{-\beta(\sum_{i=1}^N \frac{p_i^2}{2m}+U(x_1,\ldots x_N))} =$$

$$= \frac{1}{h^{3N}N!}\int e^{-\beta\frac{p_1^2}{2m}}dp_1\ldots\int e^{-\beta\frac{p_N^2}{2m}}dp_N\int_\Lambda dx_1\ldots\int_\Lambda dx_N e^{-\beta U(x_1,\ldots x_N)} =$$

$$= \frac{\left[\int_{-\infty}^{+\infty}e^{-\beta x^2/2m}dx\right]^{3N}}{h^{3N}N!}\int_\Lambda dx_1\ldots\int_\Lambda dx_N e^{-\beta U(x_1,\ldots x_N)} =$$

$$= \frac{\left[(2m\pi/\beta h^2)^{3/2}\right]^N}{N!}\int_\Lambda dx_1\ldots\int_\Lambda dx_N e^{-\beta U(x_1,\ldots x_N)}$$

The integral

$$\frac{1}{N!}\int_\Lambda dx_1\ldots\int_\Lambda dx_N e^{-\beta U(x_1,\ldots x_N)}$$

is called the "configurational" partition function of the system. Generally it is very difficult to calculate explicitly this function for a real gas. But in case of

an ideal gas the situation is again immediate. So we can now give a justification a posteriori for the definition (1.24) by considering again the case of the ideal gas. Let us thus calculate the partition function of an ideal gas (i.e. a system whose Hamiltonian is $H(\boldsymbol{p},\boldsymbol{q}) = \sum_{i=1}^{N} \frac{p_i^2}{2m}$) in the canonical ensemble. In this case calculations are much easier. In fart, by definition

$$Z_\Lambda(\beta,N) = \frac{1}{N!}\int e^{-\frac{\beta}{2m}\sum_{i=1}^{N} p_i^2}\frac{d\boldsymbol{p}d\boldsymbol{q}}{h^{3N}} = \frac{V^N}{N!}\frac{1}{h^{3N}}\left[\int_{-\infty}^{+\infty} e^{-\frac{\beta}{2m}x^2}dx\right]^{3N} =$$

$$= \frac{V^N}{N!}\left(\frac{2m\pi}{h^2\beta}\right)^{3N/2}$$

Hence, recalling that $\beta^{-1} = kT$ and using also Stirling approximation for $\ln N!$

$$F_\Lambda(\beta,N) = -kT\ln Z_\Lambda(\beta,N) = -kTN\left\{1+\ln\left[\frac{V}{N}\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\right]\right\}$$

From free energy we can calculate all thermodynamic quantities. E.g. (posing $|\Lambda| = V$)

$$P = -\frac{\partial F}{\partial V} = \frac{NkT}{V} \quad \Rightarrow \quad PV = NkT$$

$$S = -\frac{\partial F}{\partial T} = Nk\left[\frac{5}{2}+\ln\left[\frac{V}{N}\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\right]\right]$$

$$E = F+TS = \frac{3}{2}NkT$$

Results are identical to the case of Micro Canonical Ensemble!
Note also that the energy E in the canonical ensemble is not fixed and hence E has to be interpreted as mean energy. This suggest that it could be also calculate directly by the formula

$$E = \langle H(\boldsymbol{p},\boldsymbol{q})\rangle = Z_\Lambda^{-1}(\beta,N)\frac{1}{N!}\int e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}H(\boldsymbol{p},\boldsymbol{q})\frac{d\boldsymbol{p}d\boldsymbol{q}}{h^{3N}} = -\frac{\partial}{\partial\beta}\ln Z_\Lambda(\beta,N)$$

Exercise. Show that $\langle H(\boldsymbol{p},\boldsymbol{q})\rangle = \frac{3}{2}NkT$ as soon as $H(\boldsymbol{p},\boldsymbol{q}) = \sum_{i=1}^{N}\frac{p_i^2}{2m}$.

1.8 Canonical Ensemble: Energy fluctuations

A system in the Canonical Ensemble can have in principle microstates of all possible energies. This means that the energy fluctuates around its mean value $E = \langle H(\boldsymbol{p},\boldsymbol{q})\rangle$. Let us thus check the fluctuations of the energy in the Canonical Ensemble. The mean energy in the Canonical Ensemble is given by

$$E = \langle H(\boldsymbol{p},\boldsymbol{q})\rangle = \frac{\int d\boldsymbol{p}d\boldsymbol{q}H(\boldsymbol{p},\boldsymbol{q})e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}}{\int d\boldsymbol{p}d\boldsymbol{q}e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}} \tag{1.27}$$

Differentiating both side of (1.27) respect to β we get

$$\frac{\partial E}{\partial\beta} = -\frac{\int d\boldsymbol{p}d\boldsymbol{q}H^2(\boldsymbol{p},\boldsymbol{q})e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}\int d\boldsymbol{p}d\boldsymbol{q}e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}}{(\int d\boldsymbol{p}d\boldsymbol{q}e^{-\beta H(\boldsymbol{p},\boldsymbol{q})})^2} +$$

$$+ \quad \frac{\int d\boldsymbol{p}d\boldsymbol{q}H(\boldsymbol{p},\boldsymbol{q})e^{-\beta H(\boldsymbol{p},\boldsymbol{q})} \int d\boldsymbol{p}d\boldsymbol{q}H(\boldsymbol{p},\boldsymbol{q})e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}}{(\int d\boldsymbol{p}d\boldsymbol{q}e^{-\beta H(\boldsymbol{p},\boldsymbol{q})})^2} \; =$$

$$= \quad \left[\frac{\int d\boldsymbol{p}d\boldsymbol{q}H(\boldsymbol{p},\boldsymbol{q})e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}}{\int d\boldsymbol{p}d\boldsymbol{q}e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}}\right]^2 \quad - \quad \frac{\int d\boldsymbol{p}d\boldsymbol{q}H^2(\boldsymbol{p},\boldsymbol{q})e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}}{\int d\boldsymbol{p}d\boldsymbol{q}e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}}$$

Hence we get the relation, for the standard deviation of $E \; = \; \langle H(\boldsymbol{p},\boldsymbol{q})\rangle$

$$\langle H^2(\boldsymbol{p},\boldsymbol{q})\rangle - \langle H(\boldsymbol{p},\boldsymbol{q})\rangle^2 \; = \; -\frac{\partial E}{\partial \beta} \; = \; kT^2\frac{\partial E}{\partial T}$$

From thermodynamics $\frac{\partial E}{\partial T} = C_V$ where C_V is the heat capacity. In general $C_V \propto N$ as also $E \propto N$ (see e.g. the case of the perfect gas where $C_V = \frac{3}{2}Nk$ and $E = \frac{3}{2}NkT$). Hence

$$\frac{\sqrt{\langle H^2(\boldsymbol{p},\boldsymbol{q})\rangle - \langle H(\boldsymbol{p},\boldsymbol{q})\rangle^2}}{\langle H(\boldsymbol{p},\boldsymbol{q})\rangle} \; = \; \frac{\sqrt{kT^2\frac{\partial E}{\partial T}}}{E} \approx \frac{1}{\sqrt{N}} \ll 1$$

Thus the (relative) fluctuation of the energy around its mean value in the canonical ensemble are "macroscopically" small (in the sense that they are of the order of $1/\sqrt{N}$ with N being a very large number). This means that in the canonical Ensemble it is highly probable to find the system in microstates with energy equal or very close to the mean energy $E \; = \; \langle H(\boldsymbol{p},\boldsymbol{q})\rangle$. So Canonical Ensemble is "nearly" a micro-canonical Ensemble. The energy is not exactly fixed, but it can fluctuate around a fixed value with relative fluctuations of order $N^{-1/2} \approx 10^{-12}$.

1.9 The Grand Canonical Ensemble

The Micro-Canonical Ensemble applies to isolated systems with fixed N, V and E, while the Canonical Ensemble describes systems with fixed N, V and T and energy variable (e.g. systems in heat bath) The Canonical Ensemble appears more realistic than the Micro Canonical Ensemble. It is very difficult to construct a perfectly isolated system, as demanded in the Micro Canonical Ensemble. So systems whose energy is not known exactly (hence not perfectly isolated) are easier to construct experimentally.
On the other hand, in the canonical ensemble is still demanded a severe "microscopic" condition: the number of particles must fixed, i.e. system confined in Λ cannot exchange matter with the outside. This is also a very difficult situation to create experimentally. We generally deal with systems where, besides the energy, also the number of particles is not known exactly. We now thus define the Ensemble suitable to describe systems in thermodynamic equilibrium in which matter and energy can be exchanged with the exterior. The fixed thermodynamics parameters for such a system are the volume V, the temperature T and the chemical potential μ.

Hence the configuration space of the Grand canonical Ensemble is

$$\Gamma_{GC} \;=\; \bigcup_{N\geq 0} \Gamma_N(\Lambda)$$

(by convention Γ_0 represent the single micro-state in which no particle is present in the volume Λ). The (restriction to $\Gamma_N(\Lambda)$ of) probability measure of the Grand Canonical Ensemble is

$$d\mu_{GC}(\boldsymbol{p},\boldsymbol{q}) \;=\; \frac{1}{\Xi(T,\Lambda,\mu)}\frac{e^{N\beta\mu}}{N!\,h^{3N}}e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}d\boldsymbol{p}\,d\boldsymbol{q} \tag{1.28}$$

Hence $d\mu_{GC}(\boldsymbol{p},\boldsymbol{q})$ is the probability to find the system in a micro-state with exactly N particles, with momenta and positions in the small volume $d\boldsymbol{p}\,d\boldsymbol{q}$ centered at $(\boldsymbol{p},\boldsymbol{q})$ of the phase space Γ_N. By convention

$$d\mu(\Gamma_0) \;=\; \frac{1}{\Xi(T,\Lambda,\mu)}$$

is the probability to find the system in the micro-state where no particle in Λ is present.
The partition function in the grand canonical ensemble is

$$\Xi(T,\Lambda,\mu)=\sum_{N=0}^{\infty}\frac{z^N}{h^{3N}N!}\int dp_1\ldots\int dp_N\int_\Lambda dx_1\ldots\int_\Lambda dx_N e^{-\beta H(p_1,\ldots,p_N,x_1,\ldots,x_N)} \tag{1.29}$$

where $z \;=\; e^{\beta\mu}$ is called "activity" or "fugacity" of the system and $\beta=1/kT$ is the inverse temperature. The term $N=0$ is put conventionally $=1$ in the sum above while for the term $N=1$ we have $H(p_1,x_1)=p_1^2/2m$. Again remark that the integration over momenta can be done explicitly and one gets

$$\Xi(T,\Lambda,\mu) \;=\; \Xi_\Lambda(\beta,\lambda) \;=\; \sum_{N\,=\,0}^{\infty}\frac{\lambda^N}{N!}\int_\Lambda dx_1\ldots\int_\Lambda dx_N\; e^{-\beta U(x_1,\ldots,x_N)} \tag{1.30}$$

where

$$\lambda \;=\; e^{\beta\mu}\left(\frac{2\pi m}{\beta h^2}\right)^{3/2}, \qquad\qquad \beta \;=\; \frac{1}{kT} \tag{1.31}$$

The parameter λ is called configurational activity (or simply activity when it will be clear from the contest).
The connection with thermodynamic in the Grand-canonical ensemble is defined via the formula

$$\beta P_\Lambda(\beta,\lambda) \;=\; \frac{1}{V}\ln\Xi_\Lambda(\beta,\lambda) \tag{1.32}$$

and the function $P_\Lambda(\beta,\lambda)$ is identified with the thermodynamical *pressure* of the system. Another important relation is the *mean density* in the Grand Canonical ensemble. The mean density is obtained by calculating, at fixed volume $|\Lambda|$, temperature T and chemical potential μ the mean number of particle in the system.

$$\langle N\rangle \;=\; \frac{1}{\Xi_\Lambda(\beta,\lambda)}\sum_{N\,=\,0}^{\infty}\frac{N\lambda^N}{N!}\int_\Lambda dx_1\ldots\int_\Lambda dx_N\; e^{-\beta U(x_1,\ldots,x_N)} \;= \tag{1.33}$$

$$= \lambda \frac{\partial}{\partial \lambda} \ln \Xi_\Lambda(\beta, \lambda)$$

Hence calling $\rho_\Lambda(\beta, \lambda) = \frac{\langle N \rangle}{|\Lambda|}$

$$\rho_\Lambda(\beta, \lambda) = \frac{1}{|\Lambda|} \lambda \frac{\partial}{\partial \lambda} \ln \Xi_\Lambda(\beta, \lambda) \tag{1.34}$$

1.10 The ideal gas in the Grand Canonical Ensemble

To conclude this brief introduction let us consider the case of perfect gas in the Grand Canonical Ensemble. It is very easy to calculate the Grand Canonical partition function in this case, where $U(x_1, \ldots, x_N) = 0$. E.g., by (1.30) we get

$$\Xi_\Lambda^{ideal\ gas}(\beta, \lambda) = \sum_{N=0}^{\infty} \frac{\lambda^N}{N!} \int_\Lambda dx_1 \ldots \int_\Lambda dx_N = \sum_{N=0}^{\infty} \frac{(\lambda |\Lambda|)^N}{N!} = e^{\lambda |\Lambda|}$$

Hence, (1.32) and (1.34) become

$$\beta P^{ideal\ gas} = \lambda \tag{1.35}$$

$$\rho^{ideal\ gas} = \lambda \tag{1.36}$$

In particular (1.36) says that the activity λ of a perfect gas coincides with the its density $\langle N \rangle / |\Lambda|$. Putting (1.36) in (1.35) we get

$$\beta P = \rho$$

which is again the equation of state of a perfect gas.
Exercise: Calculate the fluctuation $\langle N^2 \rangle - \langle N \rangle^2$ in the case of the perfect gas and show that it is of order $\langle N \rangle$.

1.11 The Thermodynamic limit

The dependence of the density $\rho_\Lambda(\beta, \lambda)$ from the volume Λ in (1.34) must be a residual one. In fact we may think to increase the volume Λ of our system but keeping fixed the value of the chemical potential μ and the inverse temperature β. We expect in this case that the density of the system does not vary in a sensible way. Values of $|\Lambda|$ that one can take in thermodynamics are macroscopic, hence very large. We thus may think that the volume is arbitrarily large (which is the rigorous formalization of "macroscopically" large) and define

$$\beta P(\beta, \lambda) = \lim_{\Lambda \to \infty} \frac{1}{|\Lambda|} \ln \Xi_\Lambda(\beta, \lambda) \tag{1.37}$$

$$\rho(\beta, \lambda) = \lim_{\Lambda \to \infty} \frac{1}{|\Lambda|} \lambda \frac{\partial}{\partial \lambda} \ln \Xi_\Lambda(\beta, \lambda) \tag{1.38}$$

The limit $\Lambda \to \infty$ (where the way in which Λ goes to infinity has to be specified in a precise sense) is called the *thermodynamic limit* of the Grand Canonical

Ensemble. The exact thermodynamic behavior of the system is recovered at the thermodynamic limit. This limit can be understood in the physical sense as "the volume macroscopically large".
Note that, by (1.38) it is possible to express the activity of the system λ as a function of the density ρ and of the (inverse) temperature β. So the pressure of the system can be expressed in term ρ and β. This is very easy to do in the case of the ideal gas. When real gases are concerned (i.e. gases for which the potential energy U is non-zero), the formula giving the pressure of the system in term of the density is called the *virial eqution of state*.
Thermodynamic limit can also be done in the Micro Canonical and Canonical Ensemble. In this case one have to fix a give density $\rho = N/|\Lambda|$ for the system and then take the limit $\Lambda \to \infty$, $N \to \infty$ in such way that $N/|\Lambda|$ is kept constant at the value ρ.
In the Micro Canonical Ensemble, in place of (1.5) one can write

$$S(\rho, E) = \lim_{\substack{\Lambda\to\infty, N\to\infty, \\ N/|\Lambda|,\ E/|\Lambda| \text{ fixed}}} = \frac{1}{|\Lambda|} k \ln \left[\frac{1}{h^{3N}} \Psi_\Lambda(E, N) \right] \tag{1.39}$$

where $S(\rho, E)$ is the entropy per unit volume (specific entropy) which is an intensive quantity. In the Canonical Ensemble one can consider in place of (1.26)

$$F(\rho, \beta) = - \lim_{\substack{\Lambda\to\infty, N\to\infty, \\ N/|\Lambda| = \rho}} \frac{1}{|\Lambda|} kT \ln Z_\Lambda(\beta, N) \tag{1.40}$$

where $F(\rho, \beta)$ is the Gibbs free energy per unit volume.
In principle the three ensembles that we have considered are equivalent only at the thermodynamic limit, when the effect of the boundary are removed. Thus equations of the thermodynamics are exactly recovered at the thermodynamic limit.
Typical mathematical problems in statistical mechanics are thus to show the existence of limits (1.40), (1.39) and (1.38) and to show that they produce the same thermodynamic (otherwise something would be seriously wrong in the picture of the statistical mechanics).
In the following we will focus our attention mainly on the Grand Canonical Ensemble and we will investigate the existence of the limit (1.38) and the property of this limit as a function of β and z.

Chapter 2

The Grand Canonical Ensemble

2.1 Conditions on the potential energy

A system of point particles in a volume Λ in the Grand Canonical Ensemble is described by a probability measure on $\cup_N \Gamma_N(\Lambda)$ where $\Gamma_N(\Lambda) = \{(x_1, \dots x_N) \in \mathbb{R}^{dN} : x_i \in \Lambda\}$. The restriction of this probability measure to $\Gamma_N(\Lambda)$ is called the *configurational Gibbs measure* (we have already integrated over momenta)

$$d\mu(x_1, \dots, x_N) = \frac{1}{\Xi_\Lambda(\beta, \lambda)} \frac{\lambda^N}{N!} e^{-\beta U(x_1, \dots, x_N)} dx_1 \dots dx_N \tag{2.1}$$

where $\beta = (kT)^{-1}$ with T absolute temperature and k Boltzmann constant, while the activity λ is given in (1.31). $d\mu(x_1, \dots, x_N)$ is the probability to find the system in the micro-state in which exactly N particles are present and, for $i = 1, 2, \dots, N$, the i^{th} particle is in the small volume d^3x_i centered at the point $x_i \in \Lambda$. The normalization constant $\Xi(\beta, \Lambda, \lambda)$ is called the Grand Canonical partition function of the system and it is given by

$$\Xi_\Lambda(\beta, \lambda) = 1 + \sum_{N=1}^{\infty} \frac{\lambda^N}{N!} \int_\Lambda dx_1 \dots \int_\Lambda dx_N \; e^{-\beta U(x_1, \dots, x_N)} \tag{2.2}$$

The factor 1 in the sum above correspond to the micro-state in which no particle is present, which hence can occur with probability $\Xi^{-1}(\beta, \Lambda, \lambda)$. The potential energy $U(x_1, \dots, x_N)$ is assumed to be a function $U : (\mathbb{R}^d)^N \to (\mathbb{R} \cup \{+\infty\})$. We will suppose from now on that $U(x_1, \dots, x_n)$ has the following form

$$U(x_1, \dots, x_N) = \sum_{1 \le i < j \le N} V(x_i - x_j) + \sum_{i=1}^{N} \Phi_{\mathrm{e}}(x_i)$$

where $V(x)$ is a function $V : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ and $\Phi_{\mathrm{e}}(x)$ is a function $\Phi_{\mathrm{e}} : \mathbb{R}^d \to \mathbb{R}$. We will always assume that $V(x)$ is such that $V(-x) = V(x)$. We let $|x|$ to denote the Euclidean norm of x. Physically the assumption above on the potential energy means that we are restricting to the case of particles interacting

via a translational invariant *pair potential* V plus an external potential $\Phi_{\rm e}$. The interaction $\sum_{1\le i<j\le N} V(x_i - x_j)$ is "internal" in the sense that it depends only on mutual positions in space of particles (i.e. only from vectors $x_i - x_j$). The interaction $\sum_{i=1}^{N} \Phi_{\rm e}(x_i)$ depends instead on the absolute positions of particles in space and $\Phi_{\rm e}(x_i)$ is interpreted as the effect of the world "outside" the boundary of Λ on the i^{th} particle confined in Λ. For instance, suppose that outside Λ there are M particles in fixed positions $y_1, \dots y_M$, then $\Phi_{\rm e}(x_i) = \sum_{j=1}^{M} V(x_i - y_j)$.

The choice of $\Phi_{\rm e}$ is somehow arbitrary, in the sense that it will depend on conditions we are supposing outside Λ. A given choice of $\Phi_{\rm e}$ is called generally a *boundary condition*. For continuous systems the question of the effect of the boundary conditions at the thermodynamic limit is rather difficult. The pressure is expected to be independent on the choice of the boundary condition but other quantities (such as some derivatives of the pressure) may not be independent on boundary conditions. In this section we will make the mathematically simple choice $\Phi_{\rm e} = 0$ (i.e. no influence at all on particles inside Λ from the world outside) which is called *free boundary condition*. We will consider in this chapter just free boundary conditions, hence we will suppose

$$U(x_1, \dots, x_N) = \sum_{1\le i<j\le N} V(x_i - x_j) \tag{2.3}$$

Anyway, we will return later to the question of the independence of the thermodynamic limit from boundary conditions, since it is of fundamental importance in the theory of phase transitions.
By (2.3), interaction energy between particles is known once we have specified the two body potential $V(x)$. We immediately see that the function $U(x_1, \dots, x_N)$ defined in (2.3) has the following properties.

i) U is symmetric for the exchange of particles.
Let $\{\sigma(1), \sigma(2), \dots, \sigma(N)\}$ be a permutation of the set $\{1, 2, \dots, N\}$ then

$$U(x_{\sigma(1)}, \dots, x_{\sigma(N)}) = U(x_1, \dots, x_N) \tag{2.4}$$

ii) U is translational invariant.
Namely, if $(x_1, \dots, x_N)$ and $(x'_1, \dots, x'_N)$ are two configurations which differs only by a translation then

$$U(x_1, \dots, x_N) = U(x'_1, \dots, x'_N) \tag{2.5}$$

Some further conditions on the potential V must be imposed. Stability and temperedness are commonly considered as minimal conditions to guarantee a good statistical mechanics behavior of the system (see, e.g., [38] and [14]).

Definition 2.1 *A pair potential $V(x)$ is said to be stable if there exists $C \ge 0$ such that, for all $n \in \mathbb{N}$ such that $n \ge 2$ and all $(x_1, \dots, x_n) \in \mathbb{R}^{dn}$,*

$$\sum_{1\le i<j\le n} V(|x_i - x_j|) \ge -nC \tag{2.6}$$

Definition 2.2 *A pair potential $V(x)$ is said to be tempered if there exists $r_0 \geq 0$ such that*

$$\int_{|x|\geq r_0} |V(|x|)|dx < \infty \tag{2.7}$$

A stable and tempered pair potential $V(x)$ will be also called *admissible.*

Definition 2.3 *Given $V(x)$ admissible, the nonnegative number*

$$B = \sup_{\substack{n\geq 2 \\ (x_1,\dots,x_n)\in\mathbb{R}^{dn}}} -\frac{1}{n}\sum_{1\leq i<j\leq n} V(|x_i - x_j|) \tag{2.8}$$

is called the stability constant of the potential $V(x)$.

Note that temperedness of $V(x)$ implies that B is non-negative and $B = 0$ if and only if $V(|x|) \geq 0$ (i.e. "repulsive" potential). Stability and temperedness are actually deeply interconnected and the lack of one of them always produce non thermodynamic or catastrophic behaviors (see ahead).
The conditions *i*) and *ii*) are motivated by physical considerations. They originates from the observation that most of physical interaction are indeed symmetric under exchange of particles and translational invariant.
Stability and temperedness are from the physical point of view more difficult to understand. Concerning in particular temperedness, it is quite natural to assume that potential must vanish at large distances since particles far away are expected to interact in a negligible way. While is not clear why the rate of decay has to be such that (2.7) is satisfied. We will see that stability condition is somehow related to the fact that particle are not allowed to be (or pay a high price to be) at short distance from each other.
We will see below that the grand-canonical partition defined in (1.1) is a holomorphic function of λ if the potential $V(|x|)$ is stable. Moreover, under very mild additional conditions on the potential (upper-continuity) it can be proved that the converse is also true (see [38]). In other words $\Xi_\Lambda(\beta, \lambda)$ converges if and only if the potential V is stable. So, in some sense, stability is a condition *sine qua non* to construct a consistent statistical mechanics for continuous particle systems.

Proposition 2.1 *Let $U(x_1, \dots, x_N)$ be a stable interaction with stability constant B and let λ in (2.2) be allowed to vary in $\mathbb{C}$, then series in the r.h.s. of (2.2) converges absolutely for all $\lambda \in \mathbb{C}$, all $\beta \in \mathbb{R}^+$ and all Λ Lebesgue measurable set in $\mathbb{R}^d$, or in other words the function $\Xi(\beta, \Lambda, \lambda)$, as a function of λ in the complex plane, is holomorphic.*

Proof.

$$|\Xi_\Lambda(\beta,\lambda)| \leq 1 + \sum_{N=1}^{\infty} \frac{|\lambda|^N}{N!} \int_\Lambda dx_1 \dots \int_\Lambda dx_N \; e^{-\beta U(x_1,\dots,x_N)} \leq$$

$$\leq \; 1 + \sum_{N=1}^{\infty} \frac{|\lambda|^N}{N!} \int_\Lambda dx_1 \dots \int_\Lambda dx_N e^{+\beta BN} \quad =$$

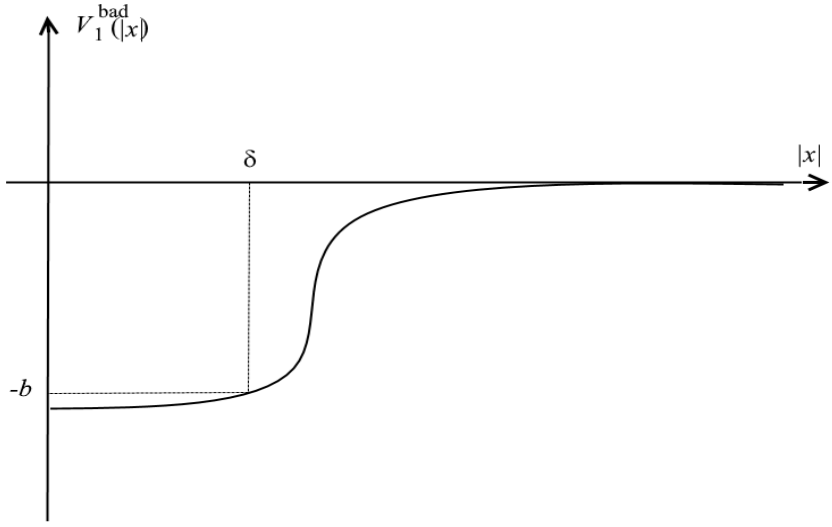


Figure 1. The potential $V_1^{\text{bad}}(|x|)$

$$= \sum_{N=0}^{\infty} \frac{(|\Lambda||\lambda|e^{\beta B})^N}{N!} = \exp\{|\Lambda||\lambda|e^{\beta B}\}$$

□

The stability condition is therefore a sufficient condition for the absolute convergence of the Grand Canonical partition function. As mentioned above, it can be also shown, that the converse also holds, i.e. stability condition is also a necessary condition (see [38]). But in that case some further conditions on the function V are needed. We rather prefer here to show how violation of stability condition and temperness produce non thermodynamic behaviors in the system.

2.2 Potentials too attractive at short distances

A simple way to violate stability is by choosing a potential $V(x)$ which is negative in the neighborhood of $x = 0$. For example, let us consider the potential $V_1^{\text{bad}}(|x|)$ as in figure 1: a continuous, not decreasing, compactly supported function of $|x|$. This potential is continuous, tempered (i.e. satisfies (2.7)), bounded (i.e. $|V_1^{\text{bad}}(|x|)| \leq \alpha$, for some $\alpha > 0$), and strictly negative around the origin $x = 0$ (i.e. $\exists \delta > 0$ and $b > 0$ such that $V_1^{\text{bad}}(|x|) \leq -b$ whenever $|x| < \delta$). This potential $V_1^{\text{bad}}(|x|)$ is not stable. Indeed, if we place N particles in positions $x_1, \ldots, x_N$ so close one to each other that $|x_i - x_j| < \delta$ (for all $i, j = 1, 2, \ldots, N$), then $U(x_1, \ldots, x_N) \leq -bN(N-1)/2$.

As we said above, it can be shown that the lack of stability destroys the convergence of the series for $\Xi(\beta, \Lambda, \lambda)$. I.e. it is possible to show that the grand

canonical partition function with the potential $V_1^{\rm bad}(|x|)$ is a divergent series. Even so, one may argue that the partition function in the canonical ensemble for the potential $V_1^{\rm bad}(|x|)$ is still well defined and calculations could then be performed in this ensemble.

Let us therefore do these calculations in the Canonical Ensemble with β and N fixed.

First we consider a catastrophic situation in which the N particles collapse in a small region inside Λ. Namely, we calculate the probability to find the system in a micro-state with the N particles being *all* located in a small sphere $S_\delta \subset \Lambda$ of radius $\delta/2$ (so that they are all at distance less than δ). By (2.1) a lower bound for such probability is given by

$$P_{\rm bad}(N) \;=\; \frac{1}{Z_\Lambda(\beta,N)} \int_{S_\delta} dx_1 \dots \int_{S_\delta} dx_N \frac{\lambda^N}{N!} e^{-\beta U(x_1,\dots,x_N)} \geq$$

$$\geq \; \frac{1}{Z_\Lambda(\beta,N)} \left[\frac{\pi\delta^3}{6}\right]^N \frac{\lambda^N}{N!} e^{+\beta b \frac{N(N-1)}{2}}$$

Now consider a configuration "macroscopically correct", i.e. a micro-state with N particles in positions $x_1, \dots x_N$ uniformly distributed in the box Λ, with density equal to the density $\rho \;=\; N/|\Lambda|$ fixed by the parameters N and $|\Lambda|$ in the canonical ensemble. If the potential $V_1^{\rm bad}$ is tempered then it is not difficult to see that for such configurations $|U(x_1,\dots,x_N)| \leq CN\rho$ where $2C \;=\; \int_{\mathbb{R}^3} dx |V(x)|$. As a matter of fact, let us consider the box Λ as the disjoint union of small cubes Δ (with volume $|\Delta|$) so that each of the particles $x_1 \dots, x_N$ belongs to one of these small cubes Δ (with volume $|\Delta|$). Hence, given a configuration $x_1, \dots, x_N$ (recall that $V_1^{\rm bad}(x)$ non-positive by assumption)

$$\sum_{\substack{j\in\{1,2,\dots,N\}\\ j\neq i}} V_1^{\rm bad}(|x_j - x_i|) \;=\; \sum_{\Delta\subset\Lambda} \sum_{\substack{j\in\{1,2,\dots,N\}\\ x_j\in\Delta,\; j\neq i}} V_1^{\rm bad}(|x_j - x_i|)| \geq$$

$$\geq \sum_{\Delta\subset\Lambda} V_1^{\rm bad}(r_\Delta) \sum_{\substack{j\in\{1,2,\dots,N\}\\ x_j\in\Delta,\; j\neq i}} 1$$

where in the last line to get the inequality we have used the assumption that $V_1^{\rm bad}$ is non decreasing and $\sum_{\Delta\subset\Lambda}$ runs over all small cubes whose disjoint union is Λ and r_Δ denotes the (minimal) distance of a point inside the cube Δ from the point x_i. Now, since we are assuming that particles are uniformly distributed in Λ and choosing the dimensions of the small cubes sufficiently large in order to still consider this cubes macroscopic, so that the particles in a small cube Δ are still uniformly distributed with density ρ or very close to ρ. Hence we may assume that there exist $\varepsilon > 0$ such that, for each $\Delta \subset \Lambda$

$$\sum_{\substack{j\in\{1,2,\dots,N\}\\ x_j\in\Delta\; j\neq i}} 1 \;\leq (1+\varepsilon)\rho|\Delta|$$

so that

$$\sum_{\substack{j\in\{1,2,\dots,N\}\\ j\neq i}} V_1^{\rm bad}(|x_j-x_i|) \;\geq\; \rho(1+\varepsilon)\sum_{\Delta} V_1^{\rm bad}(r_\Delta)\Delta \;\geq$$

$$\geq\; \rho(1+\varepsilon)\int_\Lambda V_1^{\rm bad}(x-x_i)dx \geq \rho(1+\varepsilon)\int_{\mathbb{R}^3} V_1^{\rm bad}(x)dx \qquad (2.9)$$

Hence, for such configurations we have

$$\sum_{1\leq i<j\leq N} V_1^{\rm bad}(x_i-x_j) \;=\; \frac{1}{2}\sum_{i=1}^{N}\sum_{j:\ j\neq i} V_1^{\rm bad}(x_i-x_j) \;\geq$$

$$\geq\; \frac{N\rho(1+\varepsilon)}{2}\int_{\mathbb{R}^3} V_1^{\rm bad}(x)dx \;=\; -NC\rho$$

with

$$C=\frac{(1+\varepsilon)}{2}\left[-\int_{\mathbb{R}^3} V_1^{\rm bad}(x)dx\right]$$

positive (again, recall that $V_1^{\rm bad}(x)$ non-positive by assumption). Then an upper bound for the probability for such configuration to occur is, according with (2.1)

$$P_{\rm good}(N) \;=\; \frac{1}{Z_\Lambda(\beta,N)}\int_\Lambda dx_1\dots\int_\Lambda dx_N\frac{\lambda^N}{N!}e^{-\beta U(x_1,\dots,x_N)} \;\leq$$

$$\leq\; \frac{1}{Z_\Lambda(\beta,N)}|\Lambda|^N\frac{\lambda^N}{N!}e^{+\beta\rho CN} \;=\; \frac{1}{Z_\Lambda(\beta,N)}N^N\rho^{-N}\frac{\lambda^N}{N!}e^{+\beta C\rho N}$$

Hence a lower bound for the ratio between the probability of bad configurations and good configurations is

$$\frac{P_{\rm bad}(N)}{P_{\rm good}(N)} \;\geq\; \frac{\left[\frac{\pi}{6}\rho\delta^3\right]^N\frac{\lambda^N}{N!}e^{+\beta b\frac{N(N-1)}{2}}}{N^N\frac{\lambda^N}{N!}e^{+\beta C\rho N}} \;=\; \left[\frac{\frac{\pi}{6}\rho\delta^3}{e^{\beta(\frac{b}{2}+C\rho)}}\right]^N\frac{e^{+\beta\frac{b}{2}N^2}}{N^N}$$

and, no matter how small δ and/or b are we have that

$$\lim_{N\to\infty}\frac{P_{\rm bad}(N)}{P_{\rm good}(N)} \;=\; +\infty$$

This means that is far more probable to find the system in a micro-state in which all particles are all contained within a small sphere of diameter δ in some place of Λ rather than in a micro-state "macroscopically correct", i.e. a configuration with particles uniformly distributed in Λ with a constant (approximately) density $\rho=N/|\Lambda|$.

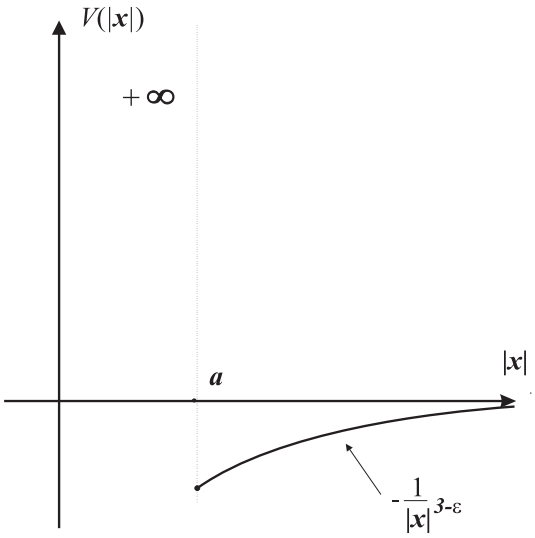


Figure 2. The potential $V_2^{\text{bad}}(|x|)$

2.3 Infrared catastrophe

We now show that the lack of stability of the pair potential yields to non-thermodynamic situations. In the previous example $V_1^{\text{bad}}(|x|)$ was a non-stable (but tempered) potential not preventing particles to accumulate in arbitrarily small regions of the space. So we consider a second case of "bad" potential which this time does not allow particles to come together at arbitrarily short distances but it is too *attractive* at large distances. Let the space dimension be set at $d = 3$ and let $a > 0$, $3 > \varepsilon > 0$ and define,

$$V_2^{\text{bad}}(|x|) \;=\; \begin{cases} +\infty & \text{if } |x| \le a \\ \\ -|x|^{-3+\varepsilon} & \text{otherwise} \end{cases} \tag{2.10}$$

It can be proved that this potential, illustrated in figure 2, is neither stable nor tempered. This is a first example of a so called *hard-core type potential* (where the hard core condition is $V_2^{\text{bad}} = +\infty$ if $|x| \le a$). It describes a system of interacting hard spheres of radius a. In fact, since $V_2^{\text{bad}}(|x|)$ is $+\infty$ whenever $|x| \le a$, then $U(x_1, \ldots, x_N) = +\infty$ whenever $|x_i - x_j| \le a$ for some i, j, thus the Gibbs factor for such configuration (i.e where some particles are at distances less or equal a) is $\exp\{-\beta U(x_1, \ldots, x_N)\} = 0$ and hence it has zero probability to occur.

This means that such system cannot take densities greater than a certain density ρ_{cp} called the close-packing density, where particle are as near as possible one to each other compatiblely with the hard core condition.

Let us again consider the system in the Canonical Ensemble at fixed inverse temperature β, fixed volume $|\Lambda|$ and fixed number of particles N, hence at fixed $\rho = N/V$. We choose N and $|\Lambda|$ in such way that ρ is much smaller that the close-packing density ρ_{cp}, i.e. $\rho/\rho_{\text{cp}} \ll 1$.

We now compare the probability to find the system in a micro-state near the close-packing situation, e.g. the close-packing configuration slightly dilated of a factor near to one. Namely, we assume that each particle can move in a small sphere S_δ of radius δ without violating the close-packing configuration. In this set of configurations the density can vary form the maximum $\rho_{\rm cp} = \frac{Const}{a^3}$ to a minimum $\tilde{\rho}_{\rm cp} = \frac{Const}{(a+2\delta)^3} = (1+\frac{2\delta}{a})^{-3}\rho_{\rm cp}$. In this case the system does not fill uniformly all the available volume $|\Lambda|$, it rather occupies a fraction $|\Lambda_{\rm cp}| = |\Lambda|\frac{\rho}{\rho_{\rm cp}}$ of the available volume and leaves a region (with volume $|\Lambda|(\rho_{\rm cp}-\rho)/\rho_{\rm cp}$) empty inside Λ. Of course such configurations are non thermodynamics.
In these configurations the potential energy U is strongly negative. An upper bound for the value of U for such type of configurations is, by a reasoning similar to that took us to the (2.9).
Indeed, suppose that particles are arranged in a configuration near the close packing (in the sense specified above). Then

$$\sum_{\substack{j\in\{1,2,\dots,N\}\\ j\neq i}} V_2^{\rm bad}(|x_j-x_i|) = \sum_{\Delta\subset\Lambda}\sum_{\substack{j\in\{1,2,\dots,N\}\\ j\neq i,\, x_j\in\Delta}} V_2^{\rm bad}(|x_j-x_i|)| \leq$$

$$\leq \sum_{\Delta\subset\Lambda} V_2^{\rm bad}(r_\Delta^{\max}) \sum_{\substack{j\in\{1,2,\dots,n\}\\ j\neq i,\, x_j\in\Delta}} 1$$

where this time $r_\Delta^{\max}$ represent the maximal distance of a point in the small cube Δ from x_i. Now we have that

$$\sum_{\substack{j\in\{1,2,\dots,n\}\\ j\neq i,\, x_j\in\Delta}} 1 \geq \tilde{\rho}_{\rm cp}|\Delta|$$

Moreover there exists surely an ϵ (depending on Δ) such that

$$\sum_{\Delta\subset\Lambda} V_2^{\rm bad}(r_\Delta^{\max})|\Delta| \leq -(1-\epsilon)B_\varepsilon(\Lambda_{\rm cp})$$

where $\Lambda_{\rm cp}$ denotes the region inside Λ where the N close-packed particles are situated and

$$B_\varepsilon(\Lambda_{\rm cp}) = \int_{x\in\Lambda_{\rm cp},\ |x|>a} |x|^{-3+\varepsilon}dx$$

and hence

$$\sum_{\substack{j\in\{1,2,\dots,N\}\\ j\neq i}} V_2^{\rm bad}(|x_j-x_i|) \leq -(1-\epsilon)B_\varepsilon(\Lambda_{\rm cp})\tilde{\rho}_{\rm cp}$$

Finally, note that, for some constant C we have that $B_\varepsilon(\Lambda_{\rm cp}) = C|\Lambda_{\rm cp}|^{\frac{\varepsilon}{3}}$ and thus calling $2C' = (1-\epsilon)C$ we get

$$\sum_{\substack{j\in\{1,2,\dots,N\}\\ j\neq i}} V_2^{\rm bad}(|x_j-x_i|) \leq -2C'\tilde{\rho}_{\rm cp}|\Lambda_{\rm cp}|^{\frac{\varepsilon}{3}}$$

and therefore

$$U(x_1,\dots,x_N) \le -NC'(1+\frac{2\delta}{a})^{-3}\rho_{\rm cp}|\Lambda_{\rm cp}|^{\frac{\varepsilon}{3}}$$

This allows to bound from below the probability $P_{\rm bad}(N)$ to find the system (in the canonical ensemble) in such bad configurations near the close-packing as follows.

$$P_{\rm bad}(N) \;=\; \frac{1}{Z_\Lambda(\beta,N)}\int_{S_\delta} dx_1 \dots \int_{S_\delta} dx_N \frac{\lambda^N}{N!} e^{-\beta U(x_1,\dots,x_N)} \ge$$

$$\ge \frac{1}{Z_\Lambda(\beta,N)}\left[\frac{4}{3}\pi\delta^3\right]^N \frac{\lambda^N}{N!} e^{+\beta N(1+2\delta/a)^{-3}\rho_{\rm cp}C'|\Lambda_{\rm cp}|^{\frac{\varepsilon}{3}}}$$

As far as good configurations (i.e. those with uniform density $\rho = N/|\Lambda|$) are concerned, proceeding similarly one can bound

$$U(x_1,\dots,x_N) \ge -NC'\rho|\Lambda|^{\frac{\varepsilon}{3}}$$

and thus an upper bound for the "good" configurations is

$$P_{\rm good}(N) \;=\; \frac{1}{Z_\Lambda(\beta,N)}\int_{\Lambda} dx_1 \dots \int_{\Lambda} dx_N \frac{\lambda^N}{N!} e^{-\beta U(x_1,\dots,x_N)} \le$$

$$\le \frac{1}{Z_\Lambda(\beta,N)}|\Lambda|^N \frac{\lambda^N}{N!} e^{+\beta NC'\rho|\Lambda|^{\frac{\varepsilon}{3}}}$$

Hence the ratio between the probability of bad and good configurations is

$$\frac{P_{\rm bad}(N)}{P_{\rm good}(N)} \ge \frac{\left[\frac{4}{3}\pi\delta^3\right]^N e^{+\beta N(1+2\delta/a)^{-3}\rho_{\rm cp}C'|\Lambda_{\rm cp}|^{\frac{\varepsilon}{3}}}}{V^N e^{+\beta NC'\rho|\Lambda|^{\frac{\varepsilon}{3}}}}$$

recalling that

$$|\Lambda| \;=\; N/\rho \qquad |\Lambda_{\rm cp}| \;=\; N/\rho_{\rm cp}$$

we get

$$\frac{P_{\rm bad}(N)}{P_{\rm good}(N)} \ge \left[\frac{4}{3}\pi\delta^3\rho\right]^N \frac{e^{+\beta C' N^{1+\frac{\varepsilon}{3}}\left[(1+2\delta/a)^{-3}\rho_{\rm cp}^{1-\frac{\varepsilon}{3}}-\rho^{1-\frac{\varepsilon}{3}}\right]}}{N^N}$$

Now observe that factor $\left[(1+2\delta)^{-3}\rho_{\rm cp}^{1-\frac{\varepsilon}{3}} - \rho^{1-\frac{\varepsilon}{3}}\right]$ in the exponential is positive if ρ is suffciently smaller than $\rho_{\rm cp}$ and δ sufficiently small. Hence, calling

$$C_1 \;=\; e^{+\beta C\left[(1+2\delta)^{-3}\rho_{\rm cp}^{1-\frac{\varepsilon}{3}}-\rho^{1-\frac{\varepsilon}{3}}\right]}, \qquad C_2 \;=\; \left[\frac{4}{3}\pi\delta^3\rho\right]^{-1}$$

and noting that $C_1 > 1$ we get

$$\frac{P_{\rm bad}(N)}{P_{\rm good}(N)} \ge \frac{C_1^{N^{1+\frac{\varepsilon}{3}}}}{C_2^N N^N}$$

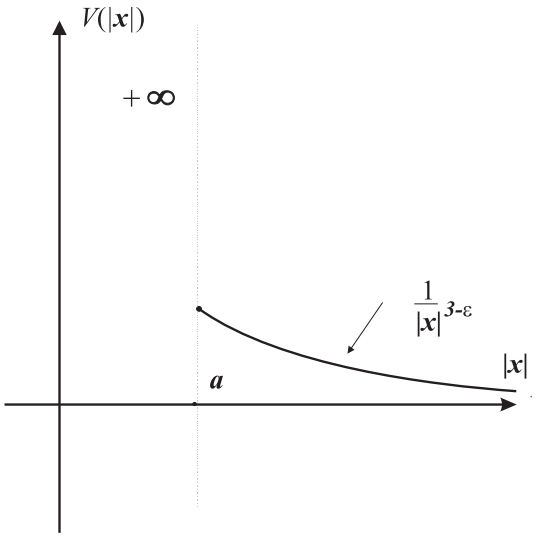


Figure 3. The potential $V_3^{\text{bad}}(|x|)$

It is just a simple exercise to show that this ratio, if $C_1 > 1$ goes to infinity as $N \to \infty$ (write e.g. $N^N = e^{N \ln N}$).
It is interesting to stress that gravitational interaction behaves at large distances exactly as $\sim |x|^{-3+\varepsilon}$ with $\varepsilon = 2$. Hence we can expect that the matter in the universe do not obey the laws of thermodynamics and in particular it is not distributed as a homogeneous low density gas (an indeed it is really the case!).

Consider now a similar case where the pair potential as the same decay as in (2.10), but now is purely repulsive, i.e. suppose

$$V_3^{\text{bad}}(x) = \begin{cases} +\infty & \text{if } |x| \le a \\ \frac{1}{|x|^{3-\varepsilon}} & \text{otherwise} \end{cases} \tag{2.11}$$

This is indeed a stable potential (since it is postive!) but not tempered. It will produce with high probability bad configurations in which particles tend to accumulate at the boundary of Λ in a close packed configuration hence forming a layer. By as argument identical to the one of case 2, supposing $\rho << \rho_{\text{cp}}$, one again shows that such configurations are far more probable than "thermodynamic configurations" (with particles uniformly distributed in Λ). Therefore particles interacting via a potential of type (2.11) will tend to leave the center of the container and accumulate in a layer at the boundary of the container.
To simplify the calculations let us suppose that the volume Λ enclosing the system is a sphere of radius L and let us estimate the probability of a bad configuration in which the N particles are in configurations $x_1, \dots, x_N$ which are nearly close-packed (i.e. they can move in little spheres of radius δ without violating the hard-core condition) with minimal density $\tilde{\rho}_{\text{cp}} = (1 + \frac{2\delta}{a})^{-3}\rho_{\text{cp}}$ occupying a region (not smaller than) Λ_{cp} which is a layer stick at the boundary of the box Λ with thickness ΔL. Since we are assuming that Λ is a sphere

of radius L we have that the volume of the occupied region $\Lambda_{\rm cp}$ is $|\Lambda_{\rm cp}| = \frac{4}{3}\pi L^3 - (L-\Delta L)^3 \approx 4\pi L^2 \Delta L$ for $\rho \ll \rho_{\rm cp}$. Now, since $|\Lambda_{\rm cp}| = N/r_{\rm cp}$ and $|\Lambda| = N/\rho$, we get $|\Lambda_{\rm cp}|\rho_{\rm cp} = |\Lambda|\rho$ and thus $4\pi L^2 \Delta L \rho_{\rm cp} = \frac{4}{3}\pi L^3 \rho$ and finally

$$\Delta L = \frac{\rho}{3\rho_{\rm cp}} L \tag{2.12}$$

Now, by the same argument seen above we have that for such configurations (recall that now the pair potential is everywhere positive) we have

$$\sum_{\substack{j\in\{1,2,\dots,N\}\\ j\neq i}} V_2^{\rm bad}(|x_j - x_i|) \;\leq\; B_\varepsilon(\Lambda_{\rm cp})\tilde{\rho}_{\rm cp}$$

where now (recall that Λ is now supposed to be a sphere of radius L)

$$B_\varepsilon(\Lambda_{\rm cp}) = \int_{\Lambda_{\rm cp}} \frac{1}{x^{3-\varepsilon}} dx = 4\pi \int_{L-\Delta L}^{L} \frac{r^2}{r^{3-\varepsilon}} dx = \frac{4\pi}{\varepsilon}\left[L^\varepsilon - (L-\Delta L)^\varepsilon\right] \approx 4\pi L^{\varepsilon-1}\Delta L$$

and thus, recalling also (2.12)

$$U(x_1,\dots,x_N) \leq \frac{N}{2}(1+\frac{2\delta}{a})^{-3} 4\pi\rho_{\rm cp} L^{\varepsilon-1}\Delta L \leq \frac{N}{2}\frac{4\pi}{3}(1+\frac{2\delta}{a})^{-3}\rho L^\varepsilon$$

and hence

$$P_{\rm bad}(N) \;=\; \frac{1}{Z_\Lambda(\beta,N)} \int_{S_\delta} dx_1 \dots \int_{S_\delta} dx_N \frac{\lambda^N}{N!} e^{-\beta U(x_1,\dots,x_N)} \geq$$

$$\geq \frac{1}{Z_\Lambda(\beta,N)} \left[\frac{4}{3}\pi\delta^3\right]^N \frac{\lambda^N}{N!} e^{-\beta\frac{N}{2}\frac{4\pi}{3}(1+\frac{2\delta}{a})^{-3}\rho L^\varepsilon}$$

On the other hand, for "good" configurations $x_1,\dots,x_N$ in which the particles are uniformly distributed in Λ with density $\rho = N/|\Lambda|$ we have

$$\sum_{\substack{j\in\{1,2,\dots,N\}\\ j\neq i}} V_2^{\rm bad}(|x_j - x_i|) \;\geq\; B_\varepsilon(\Lambda)\rho$$

where now

$$B_\varepsilon(\Lambda) = 4\pi \int_a^L \frac{r^2}{r^{3-\varepsilon}} dr = \frac{4\pi}{\varepsilon}\left[L^\varepsilon - a^\varepsilon\right]$$

and thus

$$U(x_1,\dots,x_N) \geq \frac{N}{2}\frac{4\pi}{\varepsilon}\rho\left[L^\varepsilon - a^\varepsilon\right]$$

and so

$$P_{\rm good}(N) \;=\; \frac{1}{Z_\Lambda(\beta,N)} \int_{\Lambda} dx_1 \dots \int_{\Lambda} dx_N \frac{\lambda^N}{N!} e^{-\beta U(x_1,\dots,x_N)} \leq$$

$$\leq \frac{1}{Z_\Lambda(\beta,N)} |\Lambda|^N \frac{\lambda^N}{N!} e^{-\beta\frac{N}{2}\frac{4\pi}{\varepsilon}\rho[L^\varepsilon - a^\varepsilon]}$$

Therefore the ratio between the probability of bad and good configurations is now

$$\frac{P_{\rm bad}(N)}{P_{\rm good}(N)} \geq \frac{\left[\frac{4}{3}\pi\delta^3\right]^N e^{-\beta\frac{N}{2}\frac{4\pi}{3}(1+\frac{2\delta}{a})^{-3}\rho L^\varepsilon}}{|\Lambda|^N e^{-\beta\frac{N}{2}\frac{4\pi}{\varepsilon}\rho[L^\varepsilon - a^\varepsilon]}} = \left[\frac{4}{3}\pi\frac{\delta^3}{|\Lambda|}\right]^N e^{\beta\frac{N}{2}4\pi L^\varepsilon\rho\left[\frac{1}{\varepsilon}(1-\frac{a^\varepsilon}{L^\varepsilon})-\frac{1}{3}(1+\frac{2\delta}{a})^{-3}\right]}$$

and recalling that $L^\varepsilon = (\frac{3}{4\pi}|\Lambda|)^{\frac{\varepsilon}{3}} = (\frac{3}{4\pi})^{\frac{\varepsilon}{3}}\rho^{-\frac{\varepsilon}{3}}N^{\frac{\varepsilon}{3}}$ we get

$$\frac{P_{\rm bad}(N)}{P_{\rm good}(N)} \geq \left[\frac{4\rho\pi}{3}\frac{\delta^3}{N}\right]^N e^{\beta N^{1+\frac{\varepsilon}{3}}2\pi(\frac{3}{4\pi})^{\frac{\varepsilon}{3}}\rho^{1-\frac{\varepsilon}{3}}\left[\frac{1}{\varepsilon}(1-\frac{a^\varepsilon}{L^\varepsilon})-\frac{1}{3}(1+\frac{2\delta}{a})^{-3}\right]}$$

which diverges as $N\to\infty$ as soon as

$$\frac{1}{\varepsilon}(1-\frac{a^\varepsilon}{L^\varepsilon})-\frac{1}{3}(1+\frac{2\delta}{a})^{-3}$$

is positive, which surely occurs for L sufficiently large and $\varepsilon < 3$.
Again physics gives us an example of potential such as (2.11). That is, the purely repulsive Coulomb potential (for which $\varepsilon = 2$) between charged particles *with the same charge*. Indeed electrons in excess inside a conductor tend to accumulate at the boundary of the conductor forming layers.

2.4 The Ruelle example

In Example 1, $V_1^{\rm bad}$ was tempered but not stable. It was a potential strictly negative at the origin. Therefore a necessary condition for a pair potential $V(x)$ to be stable is $V(0)\geq 0$. However the stability condition is in fact more subtle. Failure of stability can occur also for tempered potential strictly positive in the neighborhood of the origin. We will now consider a very interesting and surprising example of a potential in $d=3$ dimensions which is strongly positive near the origin, tempered, that nevertheless is a non stable potential. This subtle example, originally due to Ruelle, illustrates very well the intuitive fact that stability condition is there to avoid the collapse of many particles into a bounded region of $\mathbb{R}^d$ and it also shows the key role played by the continuum where we have always the possibility to put an arbitrary number of particles in a small region of $\mathbb{R}^d$.
Let $R>0$ and let $\delta>0$

$$V_4^{\rm bad}(x) \;=\; \begin{cases} 11 & \text{if } |x|<R-\delta \\ -1 & \text{if } R-\delta\leq|x|\leq R+\delta \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

This is clearly a tempered potential (actually it is a *finite range* potential: particles at distances greater than $R+\delta$ do not interact at all). But we will show that this is a non stable potential by proving that the grand canonical partition function diverges for such a potential. Fix an integer n and let $\tilde{x}_1,\dots,\tilde{x}_n$ be sites of a face-centered cubic lattice in three dimension with nearest neighborhoods

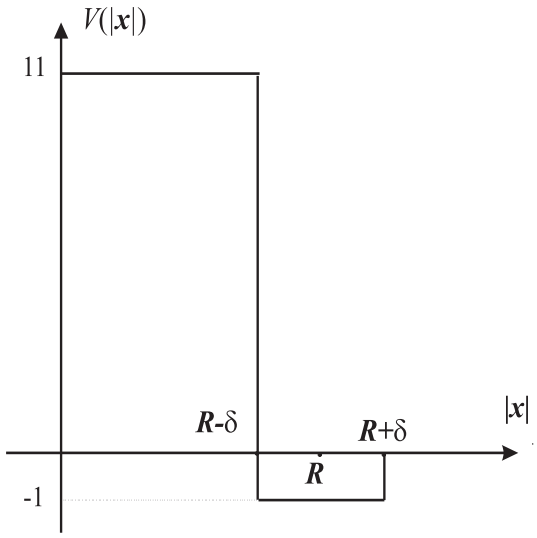


Figure 4. A catastrophic potential. The Ruelle potential

at distance R. We recall that a face-centered cubic lattice is a lattice whose unit cells are cubes, with lattice points at the center of each face of the cube, as well as at the vertices.

Let $B(\tilde{x}_1, \ldots, \tilde{x}_n) = \{\{i, j\} : |\tilde{x}_i - \tilde{x}_j| = R\}$ the set of nearest neighborhood bonds in $\tilde{x}_1, \ldots, \tilde{x}_n$ and let $B_n = |B(\tilde{x}_1, \ldots, \tilde{x}_n)|$ the cardinality of this set. Suppose that $\tilde{x}_1, \ldots, \tilde{x}_n$ are arranged in such way to maximize B_n, hence in a "close-packed" configuration. We remind that in the face centered cubic lattice every site has 12 nearest neighborhoods. If we take $\tilde{x}_1, \ldots, \tilde{x}_n$ to be close-packed, then if n is sufficiently large, the number of nearest neighborhoods bond are of the order $B_n \sim 6n$. In fact each site is the vertex of 12 nearest neighborhoods bonds, each nearest neighborhood bond is shared between two sites. If the fixed integer n is chosen sufficiently large then it is surely possible to find a close-packed configuration $\tilde{x}_1, \ldots, \tilde{x}_n$ in such way that

$$B_n > \frac{11}{2} n + \varepsilon \tag{2.14}$$

for some $\varepsilon > 0$. Suppose thus that n is chosen so large such that the close-packed face centered configuration $\tilde{x}_1, \ldots, \tilde{x}_n$ is such that (2.14) is satisfied. Then, by (2.13)

$$\sum_{i=1}^{n} \sum_{j=1}^{n} V_4^{\text{bad}}(|\tilde{x}_i - \tilde{x}_j|) \;=\; n V_4^{\text{bad}}(0) + 2 B_n V_4^{\text{bad}}(R) \;=\; 11n - 2B_n < -2\varepsilon < 0$$

Consider now the function

$$\Phi : \mathbb{R}^{6n} \to \mathbb{R} : (y_1, \ldots, y_n, z_1, \ldots, z_n) \mapsto \sum_{i=1}^{n} \sum_{j\,=\,1}^{n} V_4^{\text{bad}}(|y_i - z_j|)$$

We have that $\Phi(\tilde{x}_1,\dots,\tilde{x}_n,\tilde{x}_1,\dots,\tilde{x}_n) < -2\varepsilon$. Moreover, if

$$|y_1-\tilde{x}_1|<\frac{\delta}{2},\quad \dots,\quad |y_n-\tilde{x}_n|<\frac{\delta}{2},\quad |z_1-\tilde{x}_1|<\frac{\delta}{2},\quad \dots\quad |z_n-\tilde{x}_n|<\frac{\delta}{2}$$

with δ being the constant appearing in (2.13), then also

$$\Phi(y_1,\dots,y_n,z_1,\dots,z_n)=\sum_{i=1}^{n}\sum_{j\,=\,1}^{n}V_4^{\text{bad}}(|y_i-z_j|)<-2\varepsilon \tag{2.15}$$

Let now $S_\delta^i = \{x\in\mathbb{R}^3 : |x-\tilde{x}_i|<\delta/2\}$ be the open sphere in $\mathbb{R}^3$ with radius $\delta/2$ and center in $\tilde{x}_i$ and define $\Lambda_\delta = \cup_{i\,=\,1}^{n} S_\delta^i$ (Λ_δ is of course a subset of $\mathbb{R}^3$).
Let s be a positive integer and define M_s as the following subset of $\mathbb{R}^{3sn}$

$$M_s = \{(x_1,\dots,x_{sn})\in\mathbb{R}^{3sn} : |x_{(i-1)s+p}-\tilde{x}_i|<\delta/2,\ \ p=1,...,s\ \ i=1,...,n\}$$

Namely, $(x_1,\dots,x_{sn})\in M_s$ means that the sn-uple $(x_1,\dots,x_{sn})$ is such that the first s variables $x_1,\dots,x_s$ of the sn-uple are all inside the sphere S_δ^1, the variables $x_{s+1},\dots,x_{2s}$ are all inside the sphere S_δ^2, the variables $x_{2s+1},\dots,x_{3s}$ are all inside the sphere S_δ^3, and so on until the last s variables of the sn-uple, which are $x_{(n-1)s+1},\dots,x_{sn}$, and are all inside the sphere S_δ^n.
Now, if $(x_1,\dots,x_{sn})\in M_s$, then,

$$U(x_1,...,x_{sn})=\sum_{1\le i<j\le sn}V_4^{\text{bad}}(|x_i-x_j|)=\frac{1}{2}\left[\sum_{i=1}^{sn}\sum_{j\,=\,1}^{sn}V_4^{\text{bad}}(|x_i-x_j|)-snV_4^{\text{bad}}(0)\right]=$$

$$=\ \frac{1}{2}\left[\sum_{p\,=\,1}^{s}\sum_{p'\,=\,1}^{s}\left(\sum_{i=1}^{n}\sum_{j\,=\,1}^{n}V_4^{\text{bad}}(|x_{(i-1)s+p}-x_{(j-1)s+p'}|)\right)-snV_4^{\text{bad}}(0)\right]$$

for fixed p and p' call $x_{(i-1)s+p} = y_i$ and $x_{(i-1)s+p'} = z_i$. By definition of M_s we have that $|y_i-\tilde{x}_i|<\delta/2$ and $|z_i-\tilde{x}_i|<\delta/2$. Hence by (2.15)

$$\sum_{i=1}^{n}\sum_{j\,=\,1}^{n}V_4^{\text{bad}}(|x_{(i-1)s+p}-x_{(j-1)s+p'}|)\ =\ \sum_{i=1}^{n}\sum_{j\,=\,1}^{n}V_4^{\text{bad}}(|y_i-z_j|)\ <\ -2\varepsilon$$

hence we conclude that

$$U(x_1,\dots,x_{sn})<-\left(s^2\varepsilon+\frac{11}{2}sn\right)\qquad\text{whenever}\ \ (x_1,\dots,x_{sn})\in M_s$$

Therefore, if V_δ denote the volume of the sphere of radius $\delta/2$ in $\mathbb{R}^3$ we have

$$\Xi(\beta,\Lambda,\lambda)\ =\ 1+\sum_{N=1}^{\infty}\frac{\lambda^N}{N!}\int_\Lambda dx_1\dots\int_\Lambda dx_N\ e^{-\beta U(x_1,\dots,x_N)}\ge$$

$$\ge 1+\sum_{s=1}^{\infty}\frac{\lambda^{sn}}{(sn)!}\int_\Lambda dx_1\dots\int_\Lambda dx_{sn}\ e^{-\beta U(x_1,\dots,x_{sn})}\ge$$

$$\geq 1 + \sum_{s=1}^{\infty} \frac{\lambda^{sn}}{(sn)!} \int_{M_s} dx_1 \dots dx_{sn} \; e^{-\beta U(x_1,\dots,x_{sn})} \; \geq$$

$$\geq \sum_{s=1}^{\infty} \frac{\lambda^{sn}}{(sn)!} V_\delta^{sn} \; e^{\beta(s^2\varepsilon + \frac{11}{2} sn)} \; = \; \sum_{s=1}^{\infty} \frac{\left[\lambda V_\delta e^{\frac{11\beta}{2}}\right]^{sn}}{(sn)!} \left(e^{\beta\varepsilon}\right)^{s^2} \; = \; \sum_{s=1}^{\infty} a_s$$

This last series is a series with positive terms whose term s^{th} term is given by

$$a_s \; = \; \frac{\left[\lambda V_\delta e^{\frac{11\beta}{2}}\right]^{sn}}{(sn)!} \left(e^{\beta\varepsilon}\right)^{s^2}$$

In this formula recall that n is a fixed value while s is the variable integer. It is now easy to show that $\sum_{s=1}^{\infty} a_s$ diverges. As a matter of fact, by the ratio test for positive term series, we have that

$$\lim_{s\to\infty} \frac{a_{s+1}}{a_s} \; = \; \lim_{s\to\infty} \frac{(sn)!}{[(s+1)n)]!} \left[\lambda V_\delta e^{\frac{11\beta}{2}}\right]^n \left(e^{\beta\varepsilon}\right)^{2s+1} \geq$$

$$\geq \lim_{s\to\infty} \left[\frac{\lambda V_\delta e^{\frac{11\beta}{2}}}{(s+1)n}\right]^n \left(e^{2\beta\varepsilon}\right)^s \; = \; \left[\frac{\lambda V_\delta e^{\frac{11\beta}{2}}}{n}\right]^n \lim_{s\to\infty} \frac{\left(e^{2\beta\varepsilon}\right)^s}{(s+1)^n} \; = \; +\infty$$

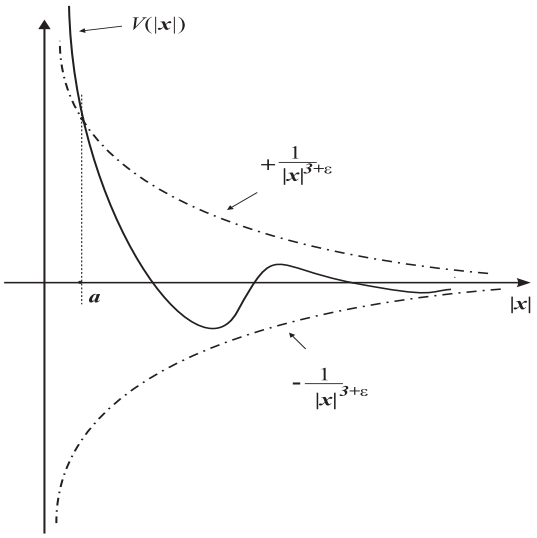


Figure 5. A Lennard-Jones potential

2.5 Admissible potentials

Let us start by discussing first three classes of potentials which are both tempered and stable.

1 - "Repulsive" temperate potentials

$$V \geq 0 \quad \text{and} \quad \int_{|x|\geq r_0} V(x)dx < +\infty \qquad \text{for some } r_0 \geq 0$$

2 - Positive type potentials: absolutely integrable potentials which are the Fourier Transform of a positive function

$$V(x) \;=\; \int \tilde{V}(\mathbf{k}) e^{i\mathbf{k}\cdot x} d^3\mathbf{k}, \qquad \tilde{V}(\mathbf{k}) \geq 0$$

3 - Lennard-Jones Potentials:

$$V(x) \geq \frac{C}{|x|^{3+\varepsilon}} \;\text{ for }\; x \leq a, \qquad |V(x)| \leq \frac{C}{|x|^{3+\varepsilon}} \;\text{ for }\; x > a$$

The potentials in the class 1, which are automatically stable are called "repulsive" in a somewhat improper manner: a positive pair potential should be monotonically decreasing in order to be really purely repulsive. We remark that tempered potentials with hard core ($V(x) \;=\; +\infty$ if $|x| \leq a$) can be included in case 3 or in case 1.
It is immediate to see that a potential of type 1 is admissible. It is tempered (i.e. it satisfies (2.7)) by definition and it is stable, since from $V(x) \geq 0$ we get

$$U(x_1, \ldots, x_n) \geq 0 \qquad \forall n, x_1, \ldots, x_n$$

hence (2.6) is verified. Moreover, since $\inf_{n,x_1,\dots,x_n} U(x_1,\dots,x_n) = 0$, we have that $B_V = 0$.
Let us show now that a potential of type 2 is admissible. A positive type potential V is stable because

$$U(x_1,\dots,x_n) = \sum_{1\le i<j\le n} V(x_i-x_j) = \frac{1}{2}\sum_{i\neq j} V(x_i-x_j) =$$

$$= \frac{1}{2}\sum_{i,j} V(x_i-x_j) - \frac{n}{2}V(0) = \frac{1}{2}\sum_{i,j}\int e^{\mathbf{i}\mathbf{k}\cdot(x_i-x_j)}\tilde{V}(\mathbf{k})d^3\mathbf{k} - \frac{n}{2}V(0) =$$

$$= \frac{1}{2}\int\left[\sum_{i,j} e^{\mathbf{i}\mathbf{k}\cdot(x_i-x_j)}\right]\tilde{V}(\mathbf{k})d^3\mathbf{k} - \frac{n}{2}V(0) =$$

$$= \frac{1}{2}\int\left|\sum_{i=1}^{n} e^{\mathbf{i}\mathbf{k}\cdot x_i}\right|^2\tilde{V}(\mathbf{k})d^3\mathbf{k} - \frac{n}{2}V(0) \ge -\frac{n}{2}V(0)$$

where last inequality follows from the assumption $\tilde{V}(\mathbf{k}) \ge 0$.
Hence the stability condition (2.6) is satisfied by choosing $B = \frac{1}{2}V(0)$. Potential 2 is tempered since it is supposed to admit Fourier transform, hence it needs to be absolutely integrable.

The example 3, the Lennard Jones type potential, is of major interest in applications, since it is the most popular and used by physicists and chemists to model the interactions between molecules in real gases.
Let us show that Lennard-Jones type potential is admissible. Such a potential is indeed tempered by definition, thus we have just to show that it is also stable. In a given configuration $x_1,\dots,x_n$, denote with the number $r_{\min} = \min_{i\neq j}|x_i - x_j|$ the minimum distance between particles.
We distinguish two cases: 1) $r_{\min} < a/2$; and 2) $r_{\min} \ge a/2$.

Case 1) $r_{\min} < a/2$. Suppose without loss of generality that $|x_1 - x_2| = r_{\min}$ and all other distances are at distances greater or equal than $r_{\min}$. We have

$$U(x_1,\dots,x_n) = V(x_1-x_2) + \sum_{j=3}^{n} V(x_1-x_j) + U(x_2,\dots,x_n) \ge$$

$$\ge \frac{C_1}{r_{\min}^{3+\varepsilon}} + \sum_{\substack{j\in\{3,\dots,n\}\\ |x_1-x_j|\ge a}} V(x_1-x_j) + U(x_2,\dots,x_n)$$

Note that the second sum of last inequality above is over those particle with index $j \in \{2,3,\dots,n\}$ at distance greater or equal to a. The inequality follows from the fact that $V(x_1-x_j) \ge 0$ if $|x_1-x_j| < a$.
We will now get a lower bound for the term $\sum_{j\in\{3,\dots,n\}:\ |x_1-x_j|\ge a} V(x_1-x_j)$. First note that, since in this sum $|x_1-x_j| \ge a$ for all j, we can bound

$$\sum_{\substack{j\in\{3,\dots,n\}:\\ |x_1-x_j|\ge a}} V(x_1-x_j) \ge -\sum_{\substack{j\in\{3,\dots,n\}:\\ |x_1-x_j|\ge a}} \frac{C_2}{|x_1-x_j|^{3+\varepsilon}}$$

We then proceed as follows. Draw around each x_j a cube Q_j with side $r_{\min}/\sqrt{12}$ (in such way that its maximal diagonal is $r_{\min}/2$) with x_j being the vertex farthest away from x_1. Since any two points among $x_3, \dots, x_n$ are at mutual distances $\geq r_{\min}$ the cubes so constructed do not overlap. Moreover if we consider the open sphere $S_1 = \{x \in \mathbb{R}^3 : |x - x_1| < \frac{a}{2}\}$ and denote by S_1^c its complementar in $\mathbb{R}^3$, then by construction (recall that $r_{\min} < a/2$) all cubes $Q_j \subset S_1^c$, i.e. they lay outside the open sphere with center x_i and radius $a/2$. Furthermore we have

$$\frac{1}{|x_1 - x_j|^{3+\varepsilon}} \leq \frac{(\sqrt{12})^3}{r_{\min}^3} \int_{Q_j} \frac{1}{|x_1 - x|^{3+\varepsilon}} dx$$

recall in fact that the cube Q_j is chosen in such way that $|x - x_1| \leq |x_1 - x_j|$ for all $x \in Q_j$. Therefore

$$\sum_{\substack{j \in \{3,\dots,n\}: \\ |x_1 - x_j| \geq a}}^{n} V(x_1 - x_j) \geq -C_2 \frac{(\sqrt{12})^3}{r_{\min}^3} \sum_{\substack{j \in \{3,\dots,n\}: \\ |x_1 - x_j| \geq a}}^{n} \int_{Q_j} \frac{1}{|x_1 - x|^{3+\varepsilon}} d^3x =$$

$$= -C_2 \frac{(\sqrt{12})^3}{r_{\min}^3} \int_{\cup_j Q_j} \frac{1}{|x_1 - x|^{3+\varepsilon}} d^3x \geq -C_2 \frac{(\sqrt{12})^3}{r_{\min}^3} \int_{|x - x_1| \geq a/2} \frac{1}{|x - x_1|^{3+\varepsilon}} d^3x =$$

$$= -C_2 \frac{(\sqrt{12})^3}{r_{\min}^3} \int_{|y| \geq a/2} \frac{1}{|y|^{3+\varepsilon}} d^3y = -\frac{C_2 K_a(\varepsilon)}{r_{\min}^3}$$

where

$$K_a(\varepsilon) = (\sqrt{12})^3 \frac{4\pi 2^\varepsilon}{\varepsilon a^\varepsilon}$$

So

$$U(x_1, \dots, x_n) \geq \frac{C_1}{r_{\min}^{3+\varepsilon}} - \frac{C_2 K_a(\varepsilon)}{r_{\min}^3} + U(x_2, \dots, x_n)$$

I.e.

$$U(x_1, \dots, x_n) \geq -C + U(x_2, \dots, x_n)$$

for a suitable C_ε positive, e.g. $-C_\varepsilon$ is the minimum of the function $F(x) = \frac{C_1}{x^{3+\varepsilon}} - \frac{K_a(\varepsilon)}{x^3}$ for $x > 0$. Therefore, iterating this formula we get

$$U(x_1, \dots, x_n) \geq -C_\varepsilon n$$

Case 2) $r_{\min} \geq a/2$. In this case we write

$$U(x_1, \dots, x_n) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j \in \{1,2,\dots,n\}:\ j \neq i} V(x_i - x_j)$$

and we will get an estimate $\sum_{j \in \{1,2,\dots,n\}, j \neq i} V(x_i - x_j)$. First note that, since all particles are at distances $\geq a/2$ we can bound

$$\sum_{\substack{j \in \{1,2,\dots,n\} \\ j \neq i}} V(x_i - x_j) \geq - \sum_{\substack{j \in \{1,2,\dots,n\} \\ j \neq i}} \frac{C_2}{|x_i - x_j|^{3+\varepsilon}}$$

We then proceed analogously as before. This time, for fixed i, we can draw around each x_j a cube Q_j with side $a/\sqrt{48}$ (such that the maximal diagonal of Q_j is $a/4$) with x_j being the vertex farthest away from x_i. Since any two points among $x_1, \ldots, x_n$ are at mutual distances $\geq a/2$ the cubes so constructed do not overlap. Moreover if we consider the open sphere $S_i = \{x \in \mathbb{R}^3 : |x - x_i| < \frac{a}{4}\}$ and denote by S_i^c its complementar in $\mathbb{R}^3$, then by construction $\cup_{j \neq i} Q_j \subset S_i^c$. In other words all cubes Q_j lay outside the open sphere with center x_i and radius $a/4$. Furthermore we have

$$\frac{1}{|x_i - x_j|^{3+\varepsilon}} \leq \frac{(\sqrt{48})^3}{a^3} \int_{Q_j} \frac{1}{|x_i - x|^{3+\varepsilon}} d^3x$$

recall in fact that the cube Q_j is chosen in such way that $|x - x_i| \leq |x_i - x_j|$ for all $x \in Q_j$.
Therefore

$$\sum_{\substack{j \in \{1,2,\ldots,n\} \\ j \neq i}} V(x_i - x_j) \;\geq\; -C_2 \frac{(\sqrt{48})^3}{a^3} \sum_{\substack{j \in \{1,2,\ldots,n\} \\ j \neq i}} \int_{Q_j} \frac{1}{|x_i - x|^{3+\varepsilon}} d^3x \;=$$

$$= \; -C_2 \frac{(\sqrt{48})^3}{a^3} \int_{\cup_{j \neq i} Q_j} \frac{1}{|x_i - x|^{3+\varepsilon}} d^3x \;\geq$$

$$\geq \; -C_2 \frac{(\sqrt{48})^3}{a^3} \int_{|x - x_i| \geq a/4} \frac{1}{|x_i - x|^{3+\varepsilon}} d^3x \;\geq\; -\bar{K}_a(\varepsilon)$$

where

$$\bar{K}_a(\varepsilon) \;=\; C_2 \frac{(\sqrt{48})^3}{a^3} \int_{|x - x_i| \geq a/4} \frac{1}{|x_i - x|^{3+\varepsilon}} d^3x$$

Hence

$$\sum_{\substack{j \in \{1,2,\ldots,n\} \\ j \neq i}} V(x_i - x_j) \;\geq\; -\bar{K}_a(\varepsilon) \tag{2.16}$$

and

$$U(x_1, \ldots, x_n) \;\geq\; -\frac{\bar{K}_a(\varepsilon)}{2} n$$

2.5.1 Basuev Criteria

We now present two very efficient criteria for stability of tempered potentials first proposed by Basuev [1].

Theorem 2.1 *Let $V(|x|)$ be a pair potential, and let $a > 0$ such that $V(a) > 0$, $V(|x|) \geq V(a)$ for all $|x| \leq a$ and*

$$V(a) \geq \mu(a) \tag{2.17}$$

where

$$\mu(a) = \sup_{\substack{n \in \mathbb{N}, \\ (x_1, \ldots, x_n) \in \mathbb{R}^{dn} \\ |x_i - x_j| > a}} \sum_{i=1}^{n} V^-(|x_i|) \tag{2.18}$$

with

$$V^-(|x|) = \max\{-V(|x|), 0\}.$$

Then the potential $V(|x|)$ is stable with stability constant B_V such that

$$B_V \leq \mu(a)$$

Proof. Let us first consider the case in all particles are mutually at distance greater than a. We have that

$$\sup_{\substack{n\in\mathbb{N},\ (x_1,\dots,x_n)\in\mathbb{R}^{dn}\\ |x_i-x_j|>a}} -\frac{1}{n}U(x_1,\dots,x_n) = \sup_{\substack{n\in\mathbb{N},\ (x_1,\dots,x_n)\in\mathbb{R}^{dn}\\ |x_i-x_j|>a}} -\frac{1}{n}\sum_{1\leq k<s\leq n} V(|x_k-x_s|) =$$

$$\sup_{\substack{n\in\mathbb{N},\ (x_1,\dots,x_n)\in\mathbb{R}^{dn}\\ |x_i-x_j|>a}} -\frac{1}{2n}\sum_{k=1}^{n}\sum_{s\neq k} V(|x_k-x_s|) \leq \frac{1}{2n}\sum_{k=1}^{n} \sup_{\substack{s\in\mathbb{N}\\ (x_1,\dots,x_s)\in\mathbb{R}^{ds}\\ |x_i-x_j|>a}} \sum_{j=1}^{s} V^-(|x_j|) = \frac{\mu(a)}{2}$$

We can therefore limit ourselves to configurations in which there are particles at distance a or smaller than a. Consider thus a configuration $(x_1,\dots,x_n)$ such that there exists $\{i,j\}\subset[n]$ such that $|x_i - x_j| \leq a$. Thus there is a particle, (which, without loss of generality, we can assume to be the particle indexed by 1 at position $x_1 = 0$), which has the maximum number of particles among $x_2,\dots,x_n$ at distance less than or equal a. We have to estimate

$$U(x_1,\dots,x_n) = E_1 + U(x_2,\dots,x_n)$$

where

$$E_1 = \sum_{j=2}^{n} V(|x_j|)$$

Say that the number of this particles close to x_1 less or equal to a is l (clearly $l \geq 1$ by assumption). The energy E_1 of the particle at position x_1 is thus

$$E_1 \geq lV(a) - \sum_{\substack{k\in[n]\\ |x_k|>a}} V^-(|x_k|)$$

To control the sum $\sum_k V^-(|x_k|)$ observe that we are supposing that each particle has at most l other particles at distance less or equal than a. Thus take the k such that $V^-(|x_k|)$ is maximum. Again, without loss of generality we can suppose $k = 2$. In the sphere with center x_2 and radius a there are at most $l+1$ particles (the particle at position x_2, for which the value $V_a^-(|x_2|)$ is maximum plus at most l other particles) Hence

$$\sum_{\substack{k\in[n]\\ |x_k|>a}} V^-(|x_k|) \leq (l+1)V^-(|x_2|) + \sum_{\substack{k\in[n]\\ |x_k|>a, |x_k-x_2|>a}} V^-(|x_k|)$$

Iterating we get, recalling definition (2.18)

$$\sum_{\substack{k\in[n]\\ |x_k|>a}} V^-(|x_k|) \leq (l+1) \sum_{\substack{k\in[n]\\ |x_i-x_j|>a}} V^-(|x_k|) \leq (l+1)\mu(a)$$

Therefore we have that

$$E_1 \geq lV(a) - (l+1)\mu(a) = l(V(a) - \mu(a)) - \mu(a) \geq -\mu(a)$$

So we have obtained

$$U(x_1, \ldots, x_n) \geq -\mu(a) + U(x_2, \ldots, x_n)$$

Iterating we get

$$U(x_1, \ldots, x_n) \geq -n\mu(a)$$

□

Example. Let us consider the potential

$$V(x) = \begin{cases} A & \text{if } |x| \leq R \\ -1 & \text{if } R < |x| \leq R + \delta \\ 0 & \text{otherwise} \end{cases} \tag{2.19}$$

Let us prove that $V(x)$ is stable if $A \geq 12$ and δ sufficiently small.
It is known that the maximum number of points that can be put on the surface of the sphere $S(R)$ of radius R in such way that the distance between any pair of such points is greater or equal R is 12. Let us now show that if δ is sufficiently small 12 is also the maximum number of points that can be fitted in the layer $S_\delta(R) = \{x \in \mathbb{R}^3 : R \leq |x| \leq R + \delta\}$. Indeed, suppose by absurd that for any $\delta > 0$ one can find 13 points on S_δ such that the distance between any pair of these points is greater or equal than R. Then by choosing $\delta = \frac{1}{n}$, for $n \in \mathbb{N}$ one can construct 13 sequences $x_n^1, \ldots, x_n^{13}$ all contained in S_δ such that, for any n and for any $i, j \in \{1, 2, \ldots, 13\}$ it holds that $|x_n^i - x_n^j| \geq R$ since S_δ is compact and $\lim_{n\to\infty} S_\delta(R) = S(R)$, for each of the sequences x_n^i there exists a subsequence convergent to some point $y^i \in S(R)$ such that for any pair $|y^i - y^j| \geq R$. So we have found 13 points of the sphere $S(R)$ of radius R all at distance from one to another greater that R, contradicting the statement that the maximal number of such points is 12. This means that the maximum number of points we can put on the layer S_δ in such way that the distance between any pair of such points is greater or equal R is 12 and therefore the potential of Figure 4 with 11 is replaced by any number $A \geq 12$ and δ is sufficiently small satisfies the hypothesis of Theorem 2.1 and hence is stable.

Theorem 2.2 *Let $V(|x|)$ be a tempered potential, such that there exists $a > 0$ such that*

$$V(|x|) \geq V(a) > 0 \qquad \text{for all} \quad |x| \leq a \tag{2.20}$$

and

$$V(a) > 2\mu(a) \tag{2.21}$$

with $\mu(a)$ defined in (2.18). Then the potential $V(|x|)$ is stable with stability constant $B_V \leq \frac{1}{2}\mu(a)$. Moreover the representation

$$V(|x|) = V_a(|x|) + K_a(|x|)$$

with

$$V_a(|x|) = \begin{cases} V(a) & \text{if } |x| \leq a \\ V(|x|) & \text{if } |x| > a \end{cases} \tag{2.22}$$

and

$$K_a(|x|) = \begin{cases} V(|x|) - V(a) & \text{if } |x| \leq a \\ 0 & \text{if } |x| > a \end{cases} \tag{2.23}$$

is such that the potential $V_a(|x|)$ defined in (2.22) is also stable and it has the same stability constant B_V of the full potential $V(|x|)$ and the potential $K_a(|x|)$ is positive and supported in $[0, a]$.

Proof. The thesis is trivial if $V^-(|x|) = 0$ (i.e. if $V(|x|)$ is purely repulsive). So we may assume that $V^-(|x|) \neq 0$. For any $(x_1, \dots, x_n) \in \mathbb{R}^{nd}$ and any $i \in [n]$, let

$$E_i(x_1, \dots, x_n) = \sum_{j \in [n]:\, j \neq i} V(|x_i - x_j|)$$

so that $U(x_1, \dots, x_n) = \frac{1}{2}\sum_{i \in [n]} E_i(x_1, \dots, x_n)$. Let now $(x_1, \dots, x_n) \in \mathbb{R}^{nd}$ be a configuration such that there is a particle in position say x_1 (without loss of generality) such that $E_1(x_1, \dots, x_n) \geq 0$. Then

$$U(x_1, \dots, x_n) = E_1(x_1, \dots, x_n) + U(x_2, \dots, x_n)$$

and, since we are assuming $E_1(x_1, \dots, x_n) \geq 0$, we have

$$U(x_1, \dots, x_n) \geq U(x_2, \dots, x_n)$$

i.e.

$$-U(x_1, \dots, x_n) \leq -U(x_2, \dots, x_n)$$

i.e., since $\frac{1}{n} < \frac{1}{n-1}$,

$$-\frac{1}{n}U(x_1, \dots, x_n) < -\frac{1}{n-1}U(x_2, \dots, x_n)$$

Thus the configuration $(x_2, \dots, x_n)$ produces a value $-\frac{1}{n-1}U(x_2, \dots, x_n)$ which is nearer to B_V than $-U(x_1, \dots, x_n)/n$. Whence we can look for minimal energy configurations $(x_1, \dots, x_n)$ limiting ourselves to those configurations in which the energy per particle $E_i(x_1, \dots, x_n)$ is negative for all $i \in [n]$.
Now let us consider the system of particles interacting via the pair potential $V_a(|x|)$ defined in (2.22) and let us assume that conditions (2.20) and (2.21) holds. Note first that, due to condition (2.20), $V_a^-(|x|) = \max\{0, -V_a(|x|)\} = V^-(|x|)$. Consider then a configuration $(x_1, \dots, x_n)$ such that there exists $\{i, j\} \subset [n]$ such that $|x_i - x_j| \leq a$ thus there is at least a particle, (which,

without loss of generality, we can assume to be the particle indexed by 1 at position x_1), which has the maximum number of particles among $x_2, \ldots, x_n$ at distance less than or equal than a. Say that the number of these particles close to x_1 less or equal to a is l (clearly $l \geq 1$ by assumption). The energy E_1 of the particle at position x_1 can thus be estimated as follows.

$$E_1(x_1, \ldots, x_n) \geq lV(a) - \sum_{\substack{k \in [n] \\ |x_k - x_1| > a}} V^-(|x_k|)$$

To control the sum

$$- \sum_{\substack{k \in [n] \\ |x_k - x_1| > a}} V^-(|x_k|))$$

observe that we are supposing that each particle has at most l other particles at distance less or equal than a. Thus take the $k \in [n]$ such that $V^-(|x_k|)$ is maximum. Again, without loss of generality we can suppose $k = 2$. In the sphere with center x_2 and radius a there are at most $l+1$ particles (the particle at position x_2, for which the value $V^-(|x_2|)$ is maximum plus at most l other particles). Hence

$$\sum_{\substack{k \in [n] \\ |x_k - x_1| > a}} V^-(|x_k|) \leq (l+1)V^-(|x_2|) + \sum_{\substack{k \in [n] \\ |x_k - x_1| > a, |x_k - x_2| > a}} V^-(|x_k|)$$

Now we have to control the sum

$$\sum_{\substack{k \in [n] \\ |x_k - x_1| > a, |x_k - x_2| > a}} V^-(|x_k|)$$

Note that in this sum all particles are at distance greater than a for the particle in x_1 and also from particle at position x_2 moreover each of the particle in the sum has at most l particles at distance less or equal than a. Suppose without loss of generality that the particle at position x_3 is such that

$$V^-(x_3) = \max_{\substack{k \in [n] \\ |x_k - x_1| > a, |x_k - x_2| > a}} V^-(x_k)$$

and this particle at x_3 has at most l particles at distance less or equal than a. Therefore

$$\sum_{\substack{k \in [n] \\ |x_k - x_1| > a, |x_k - x_2| > a}} V^-(|x_k|) \leq (l+1)V^-(|x_3|) + \sum_{\substack{k \in [n] \\ |x_k - x_1| > a, |x_k - x_2| > a, |x_k - x_3| > a}} V^-(|x_k|)$$

Iterating we get

$$\sum_{\substack{k \in [n] \\ |x_k - x_1| > a}} V^-(|x_k|) \leq (l+1) \sum_{\substack{k \in [n] \\ |x_i - x_j| > a}} V^-(|x_k|) \tag{2.24}$$

where now in the sum in the r.h.s of (2.24) all pairs of particles are at distance greater than a to each other. Therefore, recalling definition (2.18) we have that

$$\sum_{\substack{k\in[n]\\ |x_i-x_j|>a}} V^-(|x_k|) \leq \mu(a)$$

and hence

$$E_1(x_1,\dots,x_n) \geq lV(a) - (l+1)\mu(a)$$

I.e. we have that $E_1 > 0$ whenever

$$V(a) > \frac{l+1}{l}\mu(a)$$

Using now assumption (2.21) and since $\frac{l+1}{l} < 2$ we get

$$E_1(x_1,\dots,x_n) > 0$$

The conclusion is that if a configuration $(x_1,\dots,x_n)$ is such that some particles are at distance less or equal than a then there is at least a particle whose energy is positive. Hence the minimal energy configurations for $V_a(|x|)$ must be searched among those configurations in which all particles are at distance greater than a from each other. But for these configurations $V_a(|x|) = V(|x|)$ which implies that V_a and V, if stable, have the same stability constant B (and hence also the same $\bar{B}$). It is now easy to see that $V_a(|x|)$ is stable. Indeed observe that, for any configuration $(x_1,\dots,x_n)$ for which particles are at distance greater than a from each other we have

$$U_a(x_1,\dots,x_n) = \frac{1}{2}\sum_{i=1}^{n}\sum_{\substack{j\in[n]\\ j\neq i}} V(|x_i-x_j|) \geq -\frac{1}{2}n\mu(a)$$

which implies that $V_a(|x|)$ is stable with stability constant $B \leq \frac{\mu(a)}{2}$. □

We have shown above that a Lennard-Jones type potential $V(|x|)$ is stable. Here below we prove that a Lennard-Jones type potential is also Basuev.

Theorem 2.3 *Let $V(|x|)$ be a pair potential on $\mathbb{R}^d$ such that there exist constants $w, r_1, r_2 > 0$, with $r_1 \leq r_2$, and non-negative monotonic decreasing functions $\xi(|x|)$, $\eta(|x|)$ such that*

$$V(|x|)\begin{cases} \geq \xi(|x|) & \text{if } |x| \leq r_1 \\ \geq -w & \text{if } r_1 < |x| < r_2 \\ \geq -\eta(|x|) & \text{if } |x| \geq r_2 \end{cases} \tag{2.25}$$

with

$$\lim_{a\to 0}\xi(a)a^d = +\infty \tag{2.26}$$

and

$$\int_{|x|\geq r_2}\eta(|x|)dx < +\infty \tag{2.27}$$

Then $V(|x|)$ satisfies Theorem 2.2.

Proof. By Theorem 2.2 we just need to show that there exists a such that (2.20) and (2.21) are satisfied. Fix $a \in (0, r_1)$, let $\bar{w} = \max\{w, \eta(r_2)\}$

$$\bar{\eta}(|x|) = \begin{cases} \eta(|x|) & \text{if } |x| > r_2 \\ \\ \bar{w} & \text{if } |x| \leq r_2 \end{cases}$$

Then, by construction $\bar{\eta}(|x|)$ is monotonic decreasing and such that $\int_{\mathbb{R}^d} \bar{\eta}(|x|) \leq \infty$. Moreover by conditions (2.25) we have that

$$V^-(|x|) \leq \bar{\eta}(|x|)$$

Hence, recalling (2.18) and considering also that, since we took $a \in (0, r_1)$, by hypothesis $V^-(|x|) = 0$ for all $|x| \leq a$, we have

$$\mu(a) \leq \sup_{\substack{n \in \mathbb{N},\ (x_1,\ldots,x_n) \in \mathbb{R}^{dn} \\ |x_i - x_j| > a,\ |x_i| > a}} \sum_{i=1}^{n} \bar{\eta}(|x_i|)$$

To estimate from above $\sum_{i=1}^n \bar{\eta}(|x_i|)$, having in mind that all particles are at mutual distances greater than a and are at distance greater than a from the origin, we proceed as follows. We draw for each x_j a hypercube Q_j with side $a/2\sqrt{d}$ (such that the maximal diagonal of Q_j is $a/2$) in such a way that x_j is a vertex of the cube Q_j and at the same time is the point $x \in Q_j$ which is the farthest away from the origin 0. Since any two points among $x_1, \ldots, x_n$ are at mutual distances $\geq a$ the cubes so constructed do not overlap. Furthermore, using the fact that $\bar{\eta}(|x|)$ is monotonic decreasing, we have

$$\bar{\eta}(|x_j|) \leq \frac{(4d)^{\frac{d}{2}}}{a^d} \int_{Q_j} \bar{\eta}(|x|) dx$$

recall in fact that the cube Q_j is chosen in such way that $|x| \leq |x_j|$ for all $x \in Q_j$. Therefore

$$\sum_{i=1}^{n} \bar{\eta}(|x_i|) \ \leq \ \frac{(4d)^{\frac{d}{2}}}{a^d} \sum_{i=1}^{n} \int_{Q_i} \bar{\eta}(|x|) dx \ = \ \frac{(4d)^{\frac{d}{2}}}{a^d} \int_{\cup_i Q_i} \bar{\eta}(|x|) dx \ \leq$$

$$\leq \ \frac{(4d)^{\frac{d}{2}}}{a^d} \int_{\mathbb{R}^d} \bar{\eta}(|x|) dx \ = \ \frac{C_d}{a^d}$$

where $C_d = \ (4d)^{\frac{d}{2}} \int_{\mathbb{R}^d} \bar{\eta}(|x|) dx$. Hence we get

$$\mu(a) \leq \frac{C_d}{a^d}$$

Now, in view of condition (2.26), we can always choose a such that $\xi(a) a^d > 2C_d$. Thus we get

$$\xi(a) a^d \ > 2C_d \quad \Longrightarrow \quad \xi(a) > 2\frac{C_d}{a^d} \quad \Longrightarrow \quad V(a) > 2\frac{C_d}{a^d} \quad \Longrightarrow \quad V(a) > 2\mu(a)$$

□

Exercise: prove that a Lennard-Jones pair potential satisfies Theorem 2.3 and hence Theorem 2.2 (thus a Lennard-Jones potential can be written as the sum of a positive potential plus an absolutely integrable and stable potential with stability constant B_V equal to the one of the full potential).

2.6 The infinite volume limit

We will now start to consider the problem of the existence of the thermodynamic limit for the pressure of a system of particles in the grand canonical ensemble interacting via a pair potential stable and tempered. The mathematical problem is to show the existence of the infinite volume pressure defined as

$$\beta p(\beta, \lambda) \;=\; \lim_{\Lambda \to \infty} \frac{1}{|\Lambda|} \ln \Xi_\Lambda(\beta, \lambda) \tag{2.28}$$

where we recall that $|\Lambda|$ denotes the volume of Λ.
First of all we need to give a mathematical meaning to the notation $\Lambda \to \infty$ in (2.28). We know that Λ is a finite region of $\mathbb{R}^3$ and it can tend to infinity (namely its volume tends to infinity) in various way. For instance Λ could be a cylinder of fixed base A and height h and we could let $h \to \infty$. I.e. like a cigar with increasing length. It is obvious that such a system (particularly if A is very small) is not expected to have a thermodynamic behavior even if $h \to \infty$. Thus $\Lambda \to \infty$ cannot simply be $|\Lambda| \to \infty$, since we want to exclude cases like "the cigar" which is not a scandalous if they do not yield thermodynamic behavior. We need thus that $\Lambda \to \infty$ roughly in such way that Λ is big in every direction, e.g. a sphere of increasing radius, a cube of increasing size etc. We will review the following definitions

Definition 2.4 *Λ is said to go to infinity in the sense of Van Hove if the following occurs.*
Cover Λ with small cubes of size a and let $N_+(\Lambda, a)$ the number of cubes with non void intersection with Λ and $N_-(\Lambda, a)$ the number of cubes strictly included in Λ. Then $\Lambda \to \infty$ in the sense of Van Hove if

$$N_-(\Lambda, a) \to \infty, \qquad \text{and} \qquad \frac{N_-(\Lambda, a)}{N_+(\Lambda, a)} \to 1 \qquad \forall a$$

See figure 7.

As an example, it is easy to show that if $\Lambda \to \infty$ as in the case of the "cigar" seen before, then Λ is not tending to infinity in the sense of Van Hove. Consider thus, for sake of simplicity in the plane, a rectangle R with sizes l fixed and t variable and going to infinity. Let consider squares of size a to cover the rectangle. Then

$$N_-(R, a) \;=\; \frac{l}{a}\frac{t}{a}, \qquad N_+(R, a) \;=\; \frac{l}{a}\frac{t}{a} + 2\frac{t}{a} + 2\frac{l}{a}$$

Figure 6. A box Λ going to infinity in a "non-thermodynamic" way

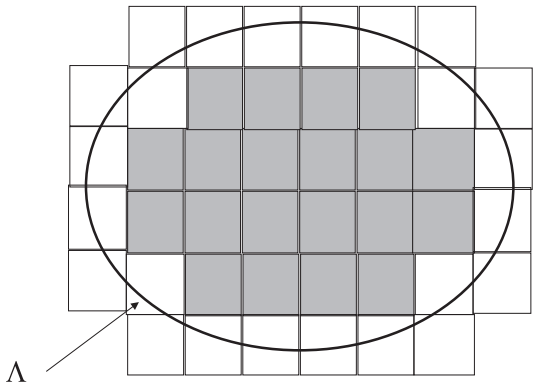


Figure 7. A set Λ with $N_+ = 44$ and $N_- = 20$

thus

$$\lim_{t\to\infty} \frac{N_-(R,a)}{N_+(R,a)} = \frac{l}{l+2a} \neq 1$$

On the other hand, let us check that a rectangle of sizes $(\sqrt{t}, t)$ and tends to infinity, as $t \to \infty$, in the sense of Van Hove. As a matter of fact

$$N_-(R,a) = \frac{\sqrt{t}}{a}\frac{t}{a}, \qquad N_+(R,a) = \frac{\sqrt{t}}{a}\frac{t}{a} + 2\frac{t}{a} + 2\frac{\sqrt{t}}{a}$$

thus

$$\lim_{t\to\infty} \frac{N_-(R,a)}{N_+(R,a)} = \lim_{t\to\infty} \frac{t\sqrt{t}}{t\sqrt{t} + at + a\sqrt{t}} = 1$$

We now give a second and more stringent definition of $\Lambda \to \infty$. For a given Λ let Λ_h denotes the set of points at distance less or equal to h from the boundary of Λ and let $|\Lambda_h|$ denotes its volume. Let finally $d(\Lambda)$ denote the diameter of Λ (i.e. $d(\Lambda) = \sup_{x,y\in\Lambda}\{|x-y|\}$) .

Definition 2.5 *We say that Λ tends to infinity in the sense of Fischer if $|\Lambda| \to \infty$ and it exists a positive function $\pi(\alpha)$ such that $\lim_{\alpha\to 0} \pi(\alpha) = 0$ and for α sufficiently small and for all Λ*

$$\frac{|\Lambda_{\alpha d(\Lambda)}|}{|\Lambda|} \leq \pi(\alpha)$$

A rectangle R of sizes $f(t), t$ such that $\lim_{t\to\infty} \frac{f(t)}{t} = 0$ does not go to infinity in the sense of Fischer. As a matter of fact

$$d(R) = [t^2 + f^2(t)]^{1/2}$$

$$
\begin{aligned}
|\Lambda_{\alpha d(R)}| &= 2\alpha[t^2+f^2(t)]^{1/2}[t-2\alpha[t^2+f^2(t)]^{1/2}+f(t)] \\
\frac{|\Lambda_{\alpha d(R)}|}{|R|} &= \frac{2\alpha[t^2+f^2(t)]^{1/2}[t-2\alpha[t^2+f^2(t)]^{1/2}+f(t)]}{tf(t)} \\
&= \frac{2\alpha t^2\sqrt{1+\frac{f^2(t)}{t^2}}[1-2\alpha\sqrt{1+\frac{f^2(t)}{t^2}}+\frac{f(t)}{t}]}{tf(t)}
\end{aligned}
$$

For any fixed α the quantity above can be make large at will, as $t \to \infty$. Thus it is not possible to find any $\pi(\alpha)$ such that $V_{\alpha d(R)}(R)/V(R) \leq \pi(\alpha)$ for all R (i.e. for all t).
On the other hand, the square S of size (t,t) goes to infinity in the sense of Fischer

$$
\begin{aligned}
d(S) &= \sqrt{2}t, \qquad\qquad |\Lambda_{\alpha d(S)}| = 4\alpha\sqrt{2}t[t-\alpha\sqrt{2}t] \\
\frac{|\Lambda_{\alpha d(S)}|}{|S|} &= \frac{4\alpha\sqrt{2}\,t[t-\alpha\sqrt{2}t]}{t^2} = 4\alpha\sqrt{2}[1-\alpha\sqrt{2}] \leq 4\alpha\sqrt{2}
\end{aligned}
$$

Hence one can choose $\pi(\alpha) = 4\alpha\sqrt{2}$.

2.7 Example: finite range potentials

We will prove in this section the existence of the thermodynamic limit for the function $\beta p(\beta,\lambda)$ defined in (2.28), but we will not treat the general case, namely particles interacting via a pair potential stable and tempered and Λ going to infinity as e.g. Van Hove. We will rather give our proof in a simpler case. Namely, we will put ourselves in $d = 3$ dimensions and we will assume that particles interact through a stable and tempered pair potential $V(x)$ with the further property that it exists $\bar{r} > 0$ such that $V(x) \leq 0$ if $|x| \geq \bar{r}$, i.e. the potential is negative for distances greater than $\bar{r}$.
In order to make things even simpler we will also suppose that Λ is a cube of size L and $\Lambda \to \infty \Leftrightarrow L \to \infty$. We now choose two particular sequences $\Lambda_1, \Lambda_2, \ldots, \Lambda_n, \ldots$ and $\tilde{\Lambda}_1, \tilde{\Lambda}_2, \ldots, \tilde{\Lambda}_n, \ldots$ of cubes. Let Λ_1 be a cube of size L_1 with volume V_1 while $\tilde{\Lambda}_1$ is a new cube of size $\tilde{L}_1 = L_1 + \bar{r}$ which consists of Λ_1 plus a frame of thickness $\frac{\bar{r}}{2}$. We denote $\tilde{V}_1$ it volume (of course $\tilde{V}_1 > V_1$) Λ_2 is a cube of size $L_2 = 2L_1 + \bar{r}$, thus in Λ_2 we can arrange $2^3 = 8$ cubes Λ_1 with frames of thickness $\bar{r}/2$ in such way that any point in a given cube Λ_1 inside Λ_2 is at distance greater that $\bar{r}$ from any point in any other cube Λ_1 inside Λ_2. Of course $\tilde{\Lambda}_2$ is a cube of size $2\tilde{L}_1$, i.e. is the cube Λ_2 plus a frame of thickness $\bar{r}/2$. See Figure 8.
In general Λ_{n+1} is a cube of size $L_{n+1} = 2L_n + \bar{r}$ and $\tilde{\Lambda}_{n+1}$ is a cube of size $\tilde{L}_{n+1} = 2\tilde{L}_n = 2L_n + 2\bar{r}$. Note that $|\Lambda_{n+1}| > 8|\Lambda_n|$ and $|\tilde{\Lambda}_{n+1}| = 8|\tilde{\Lambda}_n|$ and $\lim_{n\to\infty} |\tilde{\Lambda}_n|/|\Lambda_n| = 1$.
We will now show that the sequence of functions

$$
P_n(\beta,\lambda) = \frac{1}{|\Lambda_n|}\ln \Xi_{\Lambda_n}(\beta,\lambda) \tag{2.29}
$$

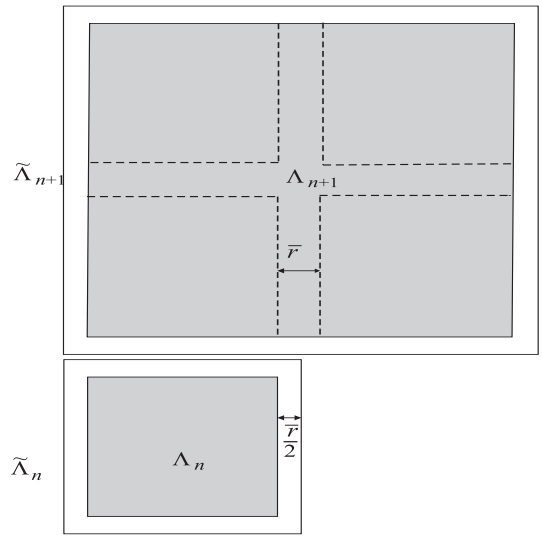


Figure 8. The cubes Λ_{n+1}, $\tilde{\Lambda}_{n+1}$ and the cubes Λ_n, $\tilde{\Lambda}_n$

tends to a limit. Define the sequence

$$\tilde{P}_n \;=\; \frac{1}{|\tilde{\Lambda}_n|}\ln \Xi_{\Lambda_n}(\beta,\lambda) \tag{2.30}$$

We will first show that this sequence converges to a limit. Consider the sequence $\Xi_n \;=\; \Xi_{\Lambda_n}(\beta,\lambda)$. We have

$$\Xi_{n+1} \;=\; \sum_{N=0}^{\infty}\frac{\lambda^N}{N!}\int_{\Lambda_{n+1}^N} dx_1\dots dx_N e^{-\beta U(x_1,\dots,x_N)}$$

We now think Λ_{n+1} as the union of 8 cubes Λ_n^j $(j=1,2,\dots,8)$ plus the (internal) frames. Obviously if we calculate Ξ_{n+1} eliminating the configurations in which particles can stay in the frames we are underestimating Ξ_{n+1}. I.e.

$$\Xi_{n+1} > \sum_{N\,=\,0}^{\infty}\frac{\lambda^N}{N!}\int_{\cup_j\Lambda_n^j} dx_1\dots dx_N e^{-\beta U(x_1,\dots,x_N)}$$

Thus the N particles are arranged in such way that N_1 are in the box Λ_n^1, N_2 are in the box Λ_n^2..., and N_8 are in the box Λ_n^8. Thus we can rename coordinates $x_1,\dots,x_N$ as $x_1^1,\dots x_{N_1}^1,\dots,x_1^8,\dots x_{N_8}^8$. The interaction between particles in different boxes is surely non positive (since they are at distances greater or equal to $\bar{r}$) then

$$\begin{aligned} U(x_1,\dots,x_N) \;&=\; U(x_1^1,\dots x_{N_1}^1,\dots,x_1^8,\dots x_{N_8}^8)\le \\ &\le U(x_1^1,\dots x_{N_1}^1) \;+\; \dots \;+\; U(x_1^8\dots x_{N_8}^8) \end{aligned}$$

and hence

$$e^{-\beta U(x_1,\dots,x_N)} \ge e^{-\beta\left[U(x_1^1,\dots x_{N_1}^1)\;+\;\dots\;+\;U(x_1^8\dots x_{N_8}^8)\right]}$$

The number of ways in which such arrangement can occur is

$$\binom{N}{N_1}\binom{N-N_1}{N_2}\cdots\binom{N-N_1-N_2-\ldots-N_7}{N_8} \;=\; \frac{N!}{N_1!\ldots N_8!}$$

Hence

$$\Xi_{n+1} > \sum_{N_1,\ldots,N_8} \frac{\lambda^{N_1+\ldots+N_8}}{N!}\frac{N!}{N_1!\ldots N_8!}\int_{\Lambda_n^1} dx_1^1\ldots\int_{\Lambda_n^1} dx_{N_1}^1\cdots$$

$$\cdots\int_{\Lambda_n^8} dx_1^8\ldots\int_{\Lambda_n^8} dx_{N_8}^8 e^{-\beta U(x_1^1,\ldots,x_{N_1}^1)}e^{-\beta U(x_1^2,\ldots,x_{N_2}^2)}\ldots e^{-\beta U(x_1^8,\ldots,x_{N_8}^8)} \;=$$
$$=\; (\Xi_n)^8$$

So we have shown that

$$\Xi_{n+1} \;>\; (\Xi_n)^8$$

Hence, since $f(x) = \ln x$ is a monotonic increasing function, we also get

$$\ln\Xi_{n+1} > 8\ln(\Xi_n)$$

and

$$\frac{1}{\tilde{V}(\Lambda_{n+1})}\ln\Xi_{n+1} > \frac{8}{\tilde{V}(\Lambda_{n+1})}\ln(\Xi_n) \;=\; \frac{1}{\tilde{V}(\Lambda_n)}\ln\Xi_n$$

So the sequence $\tilde{P}_n$ is monotonic increasing with n. On the other hand by stability we have that.

$$\Xi_n \;=\; \sum_{N\,=\,0}^{\infty}\frac{\lambda^N}{N!}\int_{\Lambda_n^N} d^3x_1\ldots d^3x_N e^{-\beta U(x_1,\ldots,x_N)} \;\leq$$

$$\leq \;\sum_{N\,=\,0}^{\infty}\frac{\lambda^N}{N!}\int_{\Lambda_n^N} d^3x_1\ldots d^3x_N e^{+\beta BN} \;=\; \exp\{\lambda V_n e^{\beta B}\}$$

therefore,

$$\ln\Xi_n \;\leq\; \lambda V_n e^{\beta B}$$

and

$$\tilde{P}_n \;=\; \frac{1}{\tilde{V}_n}\ln\Xi_n \leq \frac{1}{\tilde{V}_n}\lambda V_n e^{\beta B} \leq \lambda e^{\beta B}$$

This means that the sequence $\tilde{P}_n$ is monotonic increasing and bounded above. Hence $\lim_{n\to\infty}\tilde{P}_n \;=\; P$ exists. But now

$$P_n \;=\; \frac{1}{V_n}\ln\Xi_n \;=\; \frac{\tilde{V}_n}{V_n}\tilde{P}_n$$

whence

$$\lim_{n\to\infty} P_n \;=\; \lim_{n\to\infty}\frac{\tilde{V}_n}{V_n}\tilde{P}_n \;=\; \lim_{n\to\infty}\frac{\tilde{V}_n}{V_n}\lim_{n\to\infty}\tilde{P}_n \;=\; P$$

Therefore we show the existence of the thermodynamic limit for a class of systems interacting via a potential which, beyond being tempered and stable,

has the further property to be non positive for $|x| \geq \bar{r}$, when Λ goes to infinity along the sequence of cubes Λ_n. It is not difficult from here to show that the existence of such limit implies also that the limits exists if Λ is a cube which goes at infinity in sense homothetic (i.e. the size $\Lambda \to \infty$).
Actually, the existence of limits can be proved for much more general cases, see e.g. Theorem 3.3.12 in [38]

2.8 Properties the pressure

Let us now show some general properties of the limit for the pressure $\beta p(\beta, \lambda)$ in (2.28). The pressure $p(\beta, \lambda)$ is a function of two variables λ and β which are the two independent thermodynamic parameters describing the macroscopic equilibrium state in the Grand canonical ensamble.
We are interested to study the function $\beta p(\beta, \lambda)$ only for the "physically" admissible values of λ and β. These physical values are: λ real number in the interval $(0, +\infty)$ and β real number in the interval $(0, +\infty)$ (recall definition (1.31)).
The Grand canonical partition function $\Xi(\beta, \Lambda, \lambda)$ defined in (2.2) where we are supposing of course that $U(x_1, \ldots, x_N)$ is derived from a stable and tempered pair potential has the following structure

$$\begin{aligned} \Xi_\Lambda(\beta, \lambda) &= 1 + Z_1(\Lambda, \beta)\lambda + Z_2(\Lambda, \beta)\lambda^2 + Z_3(\Lambda, \beta)\lambda^3 + \ldots = \\ &= \sum_{n=0}^{\infty} Z_n(\Lambda, \beta)\lambda^n \end{aligned}$$

I.e. is a power series in λ with convergence radius $R = \infty$ (this is true for any Λ such that $V(\Lambda) < \infty$), i.e., $\Xi_\Lambda(\beta, \lambda)$ is analytic as a function of λ in the whole complex plane. Hence *a fortiori* $\Xi_\Lambda(\beta, \lambda)$ is analytic for all $\lambda \in (0, +\infty)$. This is true for all Λ such that $V(\Lambda) < \infty$.
The coefficients $Z_n(\Lambda, \beta)$ are explicitly given by

$$Z_n(\Lambda, \beta) = \frac{1}{n!} \int_\Lambda dx_1 \ldots \int_\Lambda dx_n \; e^{-\beta U(x_1, \ldots, x_n)} \tag{2.31}$$

They are clearly all positive numbers and due to stability (recall proposition 1) they admit the upper bound $Z_n(\Lambda, \beta) \leq [V(\Lambda)]^n \, e^{nB\beta}/n!$. Moreover as functions of β the coefficients $Z_n(\Lambda, \beta)$ are analytic in β in the whole complex plane and hence *a fortiori* for all $\beta \in (0, \infty)$.
So, $\Xi_\Lambda(\beta, \lambda)$, for all Λ such that $|\Lambda| < \infty$, and for all $\lambda \in (0, +\infty)$ is also analytic as a function of β in the whole complex plane. Hence *a fortiori* $\Xi_\Lambda(\beta, \lambda)$ is analytic for all $\beta \in (0, +\infty)$.
Now the function $\ln \Xi_\Lambda(\beta, \lambda)$ has no reason to continue analytic in λ and β in the whole complex plane, but it is indeed analytic in λ for any $\lambda \in (0, +\infty)$ and it is analytic in β for all $\beta \in (0, +\infty)$. This is due to the fact that coefficients $Z_n(\Lambda, \beta)$ are positive numbers. Hence $\Xi_\Lambda(\beta, \lambda)$ has no zeroes in the intervals $\lambda \in (0, +\infty)$ and $\beta \in (0, +\infty)$. This means that its logarithm is analytic for such intervals.

In conclusion we can state that the function

$$\beta p_\Lambda(\beta,\lambda) \;=\; \frac{1}{|\Lambda|}\ln \Xi_\Lambda(\beta,\lambda) \;=\; f_\Lambda(\beta,\lambda) \tag{2.32}$$

is analytic in λ for all $\lambda \in (0,+\infty)$ and it is also analytic in β for all $\beta \in (0,+\infty)$, whenever on choose a box Λ such that $|\lambda| < \infty$.
This fact of course does not imply that also in the limit $\Lambda \to \infty$ the function $\beta p(\beta,\lambda)$ will stay analytic in the whole physical intervals $\lambda \in (0,+\infty)$ and $\beta \in (0,+\infty)$.
Let us now list some properties of the function $f_\Lambda(\beta,\lambda)$ defined by (2.32).

Property 0a. *$f_\Lambda(\beta,\lambda)$ defined in (2.32) is continuous as a function of λ and all its derivatives are continue as functions of λ in the whole interval $\lambda \in (0,+\infty)$ and for all $\beta \in (0,+\infty)$ and for all Λ such that $V(\Lambda) < \infty$.*
Property 0b. *$f_\Lambda(\beta,\lambda)$ defined in (2.32) is continuous as a function of β and all its derivatives are continue as functions of β in the whole interval $\beta \in (0,+\infty)$ and for all $\lambda \in (0,+\infty)$ and for all Λ such that $V(\Lambda) < \infty$.*
This properties follow trivially from the fact that $\beta p_\Lambda(\beta,\lambda)$ is analytic in λ and β when they vary in the interval $(0,+\infty)$.
Property 1. *$f_\Lambda(\beta,\lambda)$ defined in (2.32) is monotonic increasing as a function of λ in the interval $\lambda \in (0,+\infty)$, for all $\beta \in (0,+\infty)$ and for all Λ such that $V(\Lambda) < \infty$*

In order to show the property 1 it is sufficient to show that $\frac{\partial f_\Lambda(\beta,\lambda)}{\partial \lambda} \geq 0$. But

$$\frac{\partial f_\Lambda(\beta,\lambda)}{\partial \lambda} \;=\; \frac{1}{V(\Lambda)}\frac{\partial \Xi_\Lambda(\beta,\lambda)/\partial\lambda}{\Xi_\Lambda(\beta,\lambda)} \;=\; \frac{\rho_\Lambda(\beta,\lambda)}{\lambda} \tag{2.33}$$

where $\rho_\Lambda(\beta,\lambda) \;=\; \langle N\rangle/V(\Lambda)$ is the density and $\langle N\rangle_\Lambda(\beta,\lambda)$ is the mean number of particles in the grand canonical ensemble at fixed values of λ, β and Λ. Explicitly $\langle N\rangle_\Lambda(\beta,\lambda)$ is given by

$$\langle N\rangle_\Lambda(\beta,\lambda) \;=\; \frac{\sum_{N=0}^{\infty}\lambda^N N Z_N(\Lambda,\beta)}{\sum_{N=0}^{\infty}\lambda^N Z_N(\Lambda,\beta)}$$

Hence, since $\langle N\rangle_\Lambda$ is surely a positive number for $\lambda > 0$, we get

$$\frac{\partial f_\Lambda(\beta,\lambda)}{\partial \lambda} > 0$$

Property 2. *$f_\Lambda(\beta,\lambda)$ defined in (2.32) is convex as a function of $\ln\lambda$ in the interval $\lambda \in (0,+\infty)$, for all $\beta \in (0,+\infty)$ and for all Λ such that $V(\Lambda) < \infty$. Moreover the finite volume density $\rho_\Lambda(\beta,\lambda) \;=\; \frac{\partial}{\partial(\ln\lambda)} f_\Lambda(\beta,\lambda)$ is a monotonic increasing function of $\ln\lambda$.*
As a matter of fact

$$\frac{\partial}{\partial(\ln\lambda)} f_\Lambda(\beta,\lambda) \;=\; \lambda\frac{\partial}{\partial\lambda} f_\Lambda(\beta,\lambda) \;=\; \rho_\Lambda(\beta,\lambda)$$

Last line follows by (2.33). Moreover

$$\frac{\partial^2}{\partial(\ln\lambda)^2} f_\Lambda(\beta,\lambda) \;=\; \lambda\frac{\partial}{\partial\lambda}\rho_\Lambda(\beta,\lambda) \;=\; \frac{\lambda}{|\Lambda|}\frac{\partial}{\partial\lambda}\langle N\rangle_\Lambda \;=$$

$$= \;\frac{\lambda}{|\Lambda|}\frac{\partial}{\partial\lambda}\left[\frac{\sum_{N=0}^{\infty}\lambda^N N Z_N(\Lambda,\beta)}{\sum_{N\,=\,0}^{\infty}\lambda^N Z_N(\Lambda,\beta)}\right] \;=\; \frac{1}{|\Lambda|}\left(\langle N^2\rangle_\Lambda - \langle N\rangle_\Lambda^2\right) \;=$$

$$= \;\frac{\langle (N-\langle N\rangle_\Lambda)^2\rangle_\Lambda}{|\Lambda)|} \tag{2.34}$$

where

$$\langle N^2\rangle_\Lambda \;=\; \frac{\sum_{N=0}^{\infty}\lambda^N N^2 Z_N(\Lambda,\beta)}{\sum_{N=0}^{\infty}\lambda^N Z_N(\Lambda,\beta)}$$

thus, since the factor $\langle (N-\langle N\rangle_\Lambda)^2\rangle_\Lambda$ is always positive we get

$$\frac{\partial^2}{\partial(\ln\lambda)^2} f_\Lambda(\beta,\lambda) \;=\; \frac{\partial}{\partial(\ln\lambda)}\rho_\Lambda(\beta,\lambda) > 0$$

This prove that $f_\Lambda(\beta,\lambda)$ is convex in the variable $\ln\lambda$ and that $\rho_\Lambda(\beta,\lambda)$ is monotonic increasing in $\ln\lambda$.

Property 3. *$f_\Lambda(\beta,\lambda)$ defined in (2.32) is convex as a function of β in the interval $\beta\in(0,+\infty)$, for all $\lambda\in(0,+\infty)$ and for all Λ such that $V(\Lambda)<\infty$.* As a matter of fact

$$\frac{\partial}{\partial\beta} f_\Lambda(\beta,\lambda) \;=\; \frac{1}{V(\Lambda)}\frac{\partial}{\partial\beta}\ln\Xi(\Lambda,\beta,\lambda) \;=\; \frac{\langle -U\rangle}{V(\Lambda)}$$

where

$$\langle -U\rangle \;=\; \frac{1}{\Xi(\Lambda,\beta,\lambda)}\sum_{N\,=\,0}^{\infty}\frac{\lambda^N}{N!}\int_\Lambda dx_1\ldots\int_\Lambda dx_N[-U(x_1,\ldots,x_N)]e^{-\beta U(x_1,\ldots,x_N)}$$

Deriving one more time respect to β

$$\frac{\partial^2}{\partial\beta^2} f_\Lambda(\beta,\lambda) \;=\; \frac{\langle U^2\rangle - \langle U\rangle^2}{V(\Lambda)}$$

where

$$\langle U^2\rangle \;=\; \frac{1}{\Xi(\Lambda,\beta,\lambda)}\sum_{N\,=\,0}^{\infty}\frac{\lambda^N}{N!}\int_\Lambda dx_1\ldots\int_\Lambda dx_N[U(x_1,\ldots,x_N)]^2\, e^{-\beta U(x_1,\ldots,x_N)}$$

and since $\langle U^2\rangle - \langle U\rangle^2 \;=\; \langle (U-\langle U\rangle)^2\rangle \geq 0$ we obtain

$$\frac{\partial^2}{\partial\beta^2} f_\Lambda(\beta,\lambda) \geq 0$$

and $f_\Lambda(\beta,\lambda)$ is a convex function of β for all Λ such that $V(\Lambda)$ is finite.

Property 4. $\beta p(\beta,\lambda) = f(\beta,\lambda) = \lim_{\Lambda\to\infty}\beta p_\Lambda(\beta,\lambda) = \lim_{\Lambda\to\infty} f_\Lambda(\beta,\lambda)$ *is convex and hence continuous as a function of* β *and* $\ln\lambda$ *in the interval* $\beta,\lambda\in(0,+\infty)$.
This property which is a very important property about the pressure in the thermodynamic limit follows trivially by the fact that the limit of a converging sequence of convex functions is also a convex function and that any convex function defined in an open set is continuous in the same open set.

Property 5. *It is possible to express, for any* Λ *finite, the pressure* p_Λ *as a function of* ρ_Λ *and* β*, i.e.*

$$p_\Lambda = g_\Lambda(\rho_\Lambda,\beta)$$

moreover the function $g_\Lambda(\rho_\Lambda,\beta)$ *is monotonic increasing as a function of* ρ_Λ.
The finite volume density is

$$\rho_\Lambda = \rho_\Lambda(\beta,\lambda) = \rho_\Lambda(e^{\ln\lambda},\beta) = F_\Lambda(\ln\lambda,\beta)$$

Since, by property 3 the function $F_\Lambda(x,\beta)$ is strictly increasing as a function of x for any Λ finite and any $\beta\in(0,+\infty)$, then it admit, as function of x, an inverse, say $x = G_\Lambda(\rho_\Lambda,\beta)$. Hence

$$\ln\lambda = G_\Lambda(\rho_\Lambda,\beta), \qquad \lambda = e^{G_\Lambda(\rho_\Lambda,\beta)}$$

Thus the function $g_\Lambda(\rho_\Lambda,\beta)$ can be indeed constructed and is given explicitly by

$$p_\Lambda = g_\Lambda(\rho_\Lambda,\beta) = \frac{1}{\beta} f_\Lambda(e^{G_\Lambda(\rho_\Lambda,\beta)},\beta)$$

It is now easy to check that this function is monotonic increasing. As a matter of fact

$$\frac{\partial}{\partial\rho_\Lambda}p_\Lambda = \frac{1}{\beta}\frac{\partial}{\partial\rho_\Lambda}(\beta p_\Lambda) = \frac{1}{\beta}\frac{\partial(\ln\lambda)}{\partial\rho_\Lambda}\frac{\partial}{\partial(\ln\lambda)}(\beta p_\Lambda) =$$

$$= \frac{1}{\beta}\rho_\Lambda\left[\frac{\partial\rho_\Lambda}{\partial(\ln\lambda)}\right]^{-1} = \frac{1}{\beta}\rho_\Lambda\left[\lambda\frac{\partial\rho_\Lambda}{\partial\lambda}\right]^{-1}$$

recalling now that $\rho_\Lambda = \langle N\rangle/V(\Lambda)$ and (2.34) we obtain

$$\frac{\partial}{\partial\rho_\Lambda}p_\Lambda = \frac{1}{\beta}\rho_\Lambda\left[\lambda\frac{\partial\rho_\Lambda}{\partial\lambda}\right]^{-1} = \frac{1}{\beta}\frac{\langle N\rangle}{\langle N^2\rangle-\langle N\rangle^2} \tag{2.35}$$

Formula (2.35) shows that $(\partial/\partial\rho_\Lambda)p_\Lambda$ is always positive. Actually (2.35) tells us also that the value of $(\partial/\partial\rho_\Lambda)p_\Lambda$ is $\langle N\rangle/\beta(\langle N^2\rangle-\langle N\rangle^2)$. Thus if we are able to prove that $\langle N\rangle(\langle N^2\rangle-\langle N\rangle^2)^{-1}$ stay bounded away from $+\infty$ for any Λ we can conclude that $p(\rho,\beta) = \lim_{\Lambda\to\infty} p_\Lambda(\rho_\Lambda,\beta)$ is continuous a function of the density $\rho = \lim_{\Lambda\to\infty}\rho_\Lambda$.

Property 6. $p = g(\rho,\beta)$ *is monotonic increasing as a function of* ρ *(hence monotonic decreasing as a function of* ρ^{-1}*)*

The monotonicity follows trivially from that fact that $p_\Lambda(\rho_\Lambda, \beta)$ is monotonic increasing for any Λ.
Note now that

$$\frac{\partial}{\partial \rho_\Lambda} p_\Lambda \;=\; \frac{1}{\beta} \frac{\langle N \rangle}{\langle N^2 \rangle - \langle N \rangle^2} \;=\; C_\Lambda \geq 0 \tag{2.36}$$

C_Λ is a constant in general depending on Λ.

2.9 Continuity of the pressure

Experimentally the thermodynamic pressure (i.e. the infinite volume limit of the finite volume pressure p_Λ) is not only increasing as a function of the density ρ but it appears furthermore to have no (jump) discontinuities. A general proof of this fact is still lacking. It has been proven that under suitable conditions on the potential (super-stability) the pressure is indeed continuous as a function of the density (see e.g. Ruelle Comm. Math. Phys. vol. 18, 127-159 (1970)). We prove here this fact in a much simpler case, namely we assume that the pair potential is either hard core, or non negative (i.e. purely repulsive). Our strategy will consist in proving that the constant C_Λ in equation (2.36) is bounded uniformly in Λ. This will allow us to conclude that $p_\Lambda(\rho_\Lambda, \beta)$ has a bounded derivative in ρ_Λ uniformly in Λ, so the limit $p(\rho, \beta) \;=\; \lim_{\Lambda \to \infty} p_\Lambda(\rho_\Lambda, \beta)$, cannot have vertical jumps as a function of ρ, i.e. $p(\rho, \beta)$ is continuous as a function of the density.
Let us thus prove that

$$\frac{\langle N \rangle}{\langle N^2 \rangle - \langle N \rangle^2} \leq C \tag{2.37}$$

assuming that the pair potential between particles is hard core or purely repulsive.

Theorem 2.4 *Let V be a tempered and stable pair potential. If V is either positive ($V \geq 0$) or such that $\exists\, a :\; V(x) \;=\; +\infty$ whenever $|x| \leq a$ (i.e. hard core), then*

$$\frac{\langle N \rangle}{\langle N^2 \rangle - \langle N \rangle^2} \leq (1 + D\lambda)$$

where D is uniform in Λ.

Proof. We will use the following short notations

$$X_N \;=\; x_1, \ldots, x_N, \qquad\qquad dX_N \;=\; dx_1 \ldots dx_N \;,$$

$$U(X_N) \;=\; U(x_1, \ldots, x_N), \qquad\qquad W(x, X_N) \;=\; \sum_{j=1}^{N} V(x_j - x)$$

$$Z_N \;=\; \int_\Lambda dx_1 \ldots \int_\Lambda dx_N \, e^{-\beta U(x_1, \ldots, x_N)} \;=\; \int_{\Lambda^N} dX_N \, e^{-\beta U(X_N)}$$

The partition function $\Xi(\lambda, \Lambda, \beta)$ (denoted shortly by Ξ) is thus rewritten

$$\Xi \;=\; \sum_{N\,=\,0}^{\infty} \frac{\lambda^N}{N!} Z_N \tag{2.38}$$

With these definitions we have,

$$Z_{N+1}^2 \;=\; \left[\int_{\Lambda^N} dX_N\, e^{-\beta U(X_N)} \int_{\Lambda} dx e^{-\beta W(x,X_N)}\right]^2 \;=$$

$$=\; \left[\int_{\Lambda^N} dX_N\, e^{-\frac{\beta}{2}U(X_N)} \int_{\Lambda} dx e^{-\frac{\beta}{2}U(X_N)}\, e^{-\beta W(x,X_N)}\right]^2 \;=$$

$$=\; \left[\int_{\Lambda^N} dX_N\, F(X_N) G(X_N)\right]^2$$

where

$$F(X_N) \;=\; e^{-\frac{\beta}{2}U(X_N)} \qquad\qquad G(X_N) \;=\; \int_{\Lambda} dx e^{-\frac{\beta}{2}U(X_N)}\, e^{-\beta W(x,X_N)}$$

Using Schwartz inequality we get

$$\left[\int_{\Lambda^N} dX_N\, F(X_N) G(X_N)\right]^2 \le \int_{\Lambda^N} dX_N\, F^2(X_N) \int_{\Lambda^N} dX_N\, G^2(X_N)$$

thus

$$Z_{N+1}^2 \le \int_{\Lambda^N} dX_N\, e^{-\beta U(X_N)} \int_{\Lambda^N} dX_N \int_{\Lambda^2} dx dy e^{-\beta U(X_N)}\, e^{-\beta W(x,X_N)} e^{-\beta W(y,X_N)} =$$

$$=\; Z_N \int_{\Lambda^N} dX_N \int_{\Lambda} dx \int_{\Lambda} dy e^{-\beta U(X_N,x,y)} \left(e^{+\beta V(x-y)} - 1 + 1\right) \;=$$

$$=\; Z_N\, Z_{N+2} + Z_N \int_{\Lambda^N} dX_N \int_{\Lambda} dx \int_{\Lambda} dy e^{-\beta U(X_N,x,y)} \left(e^{+\beta V(x-y)} - 1\right) \;=$$

$$=\; Z_N\, Z_{N+2} \;+\; Z_N \int_{\Lambda^N} dX_N \int_{\Lambda} dx e^{-\beta U(X_N,x)} \int_{\Lambda} dy e^{-\beta W(X_N,y)} \left(1 - e^{-\beta V(x-y)}\right)$$

Now, since we are assuming $V(x)$ hard core or positive we have

$$e^{-\beta W(X_N,y)} \le K(\beta) \;=\; \begin{cases} 1 & \text{if } V \ge 0 \\ \\ e^{\beta B} & \text{if } V \text{ has an hard core } a \end{cases}$$

where B is a constant $\propto a^{-3}$ (see equation (2.16)).
Thus

$$Z_{N+1}^2 \le Z_N Z_{N+2} \;+\; K(\beta) Z_N \int_{\Lambda^N} dX_N \int_{\Lambda} dx e^{-\beta U(X_N,x)} \int_{\Lambda} dy \left(1 - e^{-\beta V(x-y)}\right) \le$$

$$\le Z_N\, Z_{N+2} + K(\beta) Z_N Z_{N+1} \int_{\Lambda} dy \left|1 - e^{-\beta V(x-y)}\right|$$

Note now that in the case in which V is positive or have hard core, and also recalling that V must also be tempered, we have

$$\int_\Lambda dy \left|1 - e^{-\beta V(x-y)}\right| \leq \int_{\mathbb{R}^3} dx |1 - e^{-\beta V(x)}| \leq C(\beta) < +\infty$$

and calling $D(\beta) = K(\beta)C(\beta)$ we get finally

$$Z_{N+1}^2 \leq Z_N \, Z_{N+2} + D(\beta) Z_N Z_{N+1} \iff \frac{Z_{N+1}^2}{Z_N} \leq Z_{N+2} + Z_{N+1} D(\beta) \quad (2.39)$$

Now consider

$$(\langle N \rangle [1 + \lambda D(\beta)])^2 = \left(\sum_{N=0}^{\infty} \frac{\lambda^N}{N!} N \frac{Z_N}{\Xi} [1 + \lambda D(\beta)] \right)^2 =$$

$$= \left(\sum_{N=0}^{\infty} \frac{\lambda^N}{N!} \frac{Z_N}{\Xi} \left[\frac{\lambda^{N+1} Z_{N+1}}{\lambda^N Z_N} + N\lambda D(\beta) \right] \right)^2 =$$

$$= \left(\sum_{N=0}^{\infty} \frac{\lambda^N}{N!} \frac{Z_N}{\Xi} \left[\frac{\lambda \, Z_{N+1}}{Z_N} + N\lambda D(\beta) \right] \right)^2$$

where in the second line we have used

$$\sum_{N=0}^{\infty} \frac{\lambda^N}{N!} N Z_N = \sum_{N=0}^{\infty} \frac{\lambda^{N+1}}{N!} Z_{N+1}$$

and by definition (2.38)

$$\sum_{N=0}^{\infty} \frac{\lambda^N}{N!} \frac{Z_N}{\Xi} = 1$$

Now

$$\left(\sum_{N=0}^{\infty} \frac{\lambda^N}{N!} \frac{Z_N}{\Xi} \left[\frac{\lambda Z_{N+1}}{Z_N} + N\lambda D(\beta) \right] \right)^2 = \left(\sum_{N=0}^{\infty} F_N \cdot G_N \right)^2$$

where

$$F_N = \left(\frac{\lambda^N}{N!} \frac{Z_N}{\Xi} \right)^{1/2} \qquad G_N = \left(\frac{\lambda^N}{N!} \frac{Z_N}{\Xi} \right)^{1/2} \left[\frac{\lambda Z_{N+1}}{Z_N} + N\lambda D(\beta) \right]$$

using thus again Schwartz inequality we get

$$\left(\sum_{N=0}^{\infty} F_N \cdot G_N \right)^2 \leq (\sum_{N=0}^{\infty} F_N^2)(\sum_{N=0}^{\infty} G_N^2) =$$

$$= 1 \times \sum_{N=0}^{\infty} \left(\frac{\lambda^N}{N!} \frac{Z_N}{\Xi} \right) \left[\lambda \frac{Z_{N+1}}{Z_N} + N\lambda D(\beta) \right]^2 =$$

$$= \sum_{N=0}^{\infty} \left(\frac{\lambda^N}{N!} \frac{1}{\Xi} \right) \left[\lambda^2 \frac{Z_{N+1}^2}{Z_N} + 2N\lambda^2 D(\beta) Z_{N+1} + Z_N N^2 \lambda^2 D^2(\beta) \right]$$

hence using also (2.39)

$$(\langle N\rangle[1+\lambda D(\beta)])^2 \leq \sum_{N\,=\,0}^{\infty}\left(\frac{\lambda^N}{N!}\frac{1}{\Xi}\right)\lambda^2\Big[Z_{N+2}\ +$$

$$+\ Z_{N+1}D(\beta)+2ND(\beta)Z_{N+1}+Z_N N^2D^2(\beta)\Big]$$

Now observe that

$$\sum_{N=0}^{\infty}\left(\frac{\lambda^N}{N!}\frac{1}{\Xi}\right)\lambda^2 Z_{N+2}\ =\ \langle N(N-1)\rangle$$

$$\sum_{N=0}^{\infty}\left(\frac{\lambda^N}{N!}\frac{1}{\Xi}\right)\lambda^2 Z_{N+1}D(\beta)\ =\ \lambda D(\beta)\langle N\rangle$$

$$\sum_{N=0}^{\infty}\left(\frac{\lambda^N}{N!}\frac{1}{\Xi}\right)2N\lambda^2 D(\beta)Z_{N+1}\ =\ 2D(\beta)\lambda\langle N(N-1)\rangle$$

$$\sum_{N=0}^{\infty}\left(\frac{\lambda^N}{N!}\frac{1}{\Xi}\right)Z_N N^2\lambda^2D^2(\beta)\ =\ D^2(\beta)\lambda^2\langle N^2\rangle$$

thus

$$(\langle N\rangle[1+\lambda D(\beta)])^2 \leq$$

$$\leq\ \langle N(N-1)\rangle+\lambda D(\beta)\langle N\rangle\ +\ 2D(\beta)\lambda\langle N(N-1)\rangle\ +\ D^2(\beta)\lambda^2\langle N^2\rangle\ =$$

$$=\ \langle N^2\rangle-\langle N\rangle+\lambda D(\beta)\langle N\rangle\ +\ 2D(\beta)\lambda(\langle N^2\rangle-\langle N\rangle)\ +\ D^2(\beta)\lambda^2\langle N^2\rangle\ =$$

$$=\ \langle N^2\rangle(1+D(\beta)\lambda)^2-\langle N\rangle(1+\lambda D(\beta))$$

Thus we are arrived at the inequality

$$\langle N\rangle^2[1+\lambda D(\beta)]^2 \leq\ \langle N^2\rangle(1+D(\beta)\lambda)^2-\langle N\rangle(1+\lambda D(\beta))$$

i.e.

$$\langle N\rangle\ \ \leq\ (\langle N^2\rangle-\langle N\rangle^2)(1+\lambda D(\beta))$$

which is as to say

$$\frac{\langle N\rangle}{\langle N^2\rangle-\langle N\rangle^2}\leq 1+\lambda D(\beta)$$

and the proof is completed.□

2.10 Analiticity of the pressure

We have seen by property 0a that the pressure at finite volume $p_\Lambda(\beta, \lambda)$ is analytic as a function of its parameters β and λ in the whole physical domain $\lambda > 0$, $\beta > 0$. We can now ask if the infinite volume pressure $p(\beta, \lambda)$ is also analytic in its parameters. If this were the case then we would be really in trouble and we should conclude that statistical mechanics is not sufficient to describe the macroscopic behaviour of a system with a large number of particles. As a matter of fact experiments tells us that the physical pressure can indeed be non analytic. For example the graphic of the pressure versus the density at constant temperature (if the temperature is not too high, i.e. not above the critical point) for a real gas is drawn below.
When ρ reaches the values ρ_0 the gas starts to condensate to its liquid phase and during the whole interval $[\rho_0, \rho_1]$ the gas performs a phase transitions. i.e. from its gas phase to its liquid phase, at the same pressure p_0. Above ρ_1 the system is totally in its liquid phase. The change occurs abruptly and is usually characterized by singular behaviour in thermodynamic functions.

Hence, in spite of the fact that pressure is continuous even in the limit, its derivatives may be not continuous in the infinite volume limit. Note that this fact substantially justifies the necessity to take the thermodynamic limit. Until we stay at finite volume, all thermodynamic functions are analytic, hence in order to describe a phenomenon like phase transition we are forced to consider the infinite volume limit. We can thus give the following mathematical definition for phase transition

Definition 2.6 *Any non analytic point of the grand canonical pressure defined in (2.28) occurring for real positive β or λ is called a phase transition point.*

People believe in general that the pressure $p(\beta, \lambda)$ is a piecewise analytic function of its parameters in the physical interval $\lambda > 0$ $\beta > 0$.
Hence it is very important to see which are the values of parameter λ and β for which the pressure is analytic. Guided by the physical intuition, for low values of λ and β one expects that the pressure is indeed analytic. In fact λ low means that the system is at low density, while β low means that the system is at high temperature (e.g. above the critical point thus so high that the system is always a gas and never condensates). For such values the system is indeed a gas and in general very near to a perfect gas. I.e. for temperature sufficiently high and/or density sufficiently low the system is in the gas phase and no phase transition occurs.
Hence it should exist a theorem stating that the pressure $p(\beta, \lambda)$ is analytic for β and/or λ sufficiently small. We would see later the such a theorem can effectively be proved.
We may ask the following question. We know that $p(\beta, \rho)$ is continuous while e.g. $\frac{\partial p}{\partial \rho}$ may be not. But if $\frac{\partial p}{\partial \rho}$ is not continuous in some point it means that at that point ρ_0 the function can take two values. So what is the thermodynamic limit when $\rho = \rho_0$? Or, in other words which of the possible values of $\frac{\partial p}{\partial \rho}$ the system chooses at the thermodynamic limit?

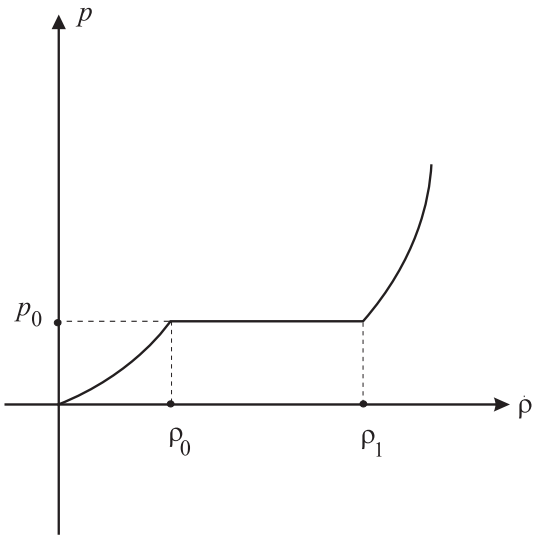


Figure 9. Pressure versus density for a physical gas. The gas-liquid phase transition

The answer to this question resides in the concept of boundary conditions. Up to now boundary conditions where "open", i.e. we were studying a system of particles enclosed in a box Λ supposing that outside Λ there was nothing, just empty space.
But of course we could also have done things in a different way, or in more proper words we could have put a different boundary condition . For instance we can put n particles outside the box Λ at fixed point $y_1, \ldots, y_n$.
In this case the grand canonical partition function looks as

$$\Xi^y_\Lambda(\beta,\lambda) \;=\; 1 \;+\; \sum_{N=1}^{\infty} \frac{\lambda^N}{N!} \int_\Lambda dx_1 \ldots \int_\Lambda dx_N \; e^{-\beta U(x_1,\ldots,x_N)} e^{-\beta W(x_1,\ldots,x_N,y_1,\ldots,y_n)} \tag{2.40}$$

where

$$W(x_1,\ldots,x_N,y_1,\ldots,y_n) \;=\; \sum_{i=1}^{N}\sum_{j=1}^{n} V(x_i - y_j)$$

Hence the function $\Xi_\Lambda(\beta,\lambda)$ may depend also from boundary conditions. This means that also p_Λ the pressure at finite volume depends on boundary conditions.
Does the infinite volume pressure depend on boundary conditions?
The answer to this question must be no (always guided by physical intuition), at least for not too strange systems and/or not too strange boundary conditions. It is possible to show this quite easily for a finite range (with range $\bar{r}$) hard core (with hard core a) potential. I.e., particles must stay at distances a or higher and they do not interact if the distance is greater than $\bar{r}$.
In this case particles outside Λ than give contribution to the partition function (2.40) are just in a frame of radius $\bar{r}$ outside Λ. Supposing that Λ is a cube of

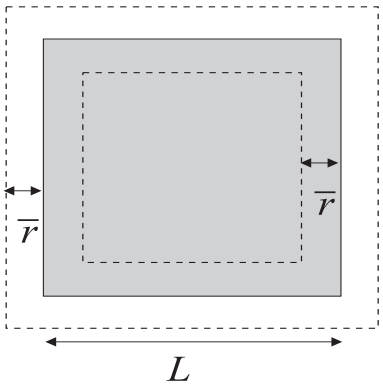


Figure 10.

size L, the volume of this frame is of the order $L^2\bar{r}$. Any particle inside Λ can interact just with particles inside a sphere of radius $\bar{r}$ centers at the particles position, and due to the hard core condition in this sphere one can arrange at most of the order of $(\bar{r}/a)^3$. This means that a particle inside can interact at most with $(\bar{r}/a)^3$ particles outside Λ.
On the other hand particles inside Λ that can interact with particles outside Λ are also contained in a (internal) frame of size $\bar{r}$, and the maximum number of particles inside Λ(hence contributing to W in (2.40)) is of the order

$$\frac{L^2\bar{r}}{a^3}$$

With this observations it is not difficult to see that

$$|W(x_1,\dots,x_N,y_1,\dots,y_n)| \leq \text{Const. } \frac{L^2\bar{r}}{a^3}\left(\frac{\bar{r}}{a}\right)^3$$

Thus

$$\frac{1}{L^3}\ln \Xi_y(\Lambda,\beta,\lambda) \;=\; \frac{1}{L^3}\ln \Xi_{\text{open}}(\beta,\lambda,\lambda) \pm \frac{1}{L^3}\text{Const. } \frac{L^2\bar{r}}{a^3}\left(\frac{\bar{r}}{a}\right)^3$$

and the factor

$$\frac{1}{L^3}\text{Const. } \frac{L^2\bar{r}}{a^3}\left(\frac{\bar{r}}{a}\right)^3$$

goes to zero as $L \to \infty$. I.e. the pressure does not depend on boundary conditions.
But now in general the derivative of the pressure may depend on boundary condition. This happens precisely when the derivative are not continuous. In such points different boundary condition may force different values of derivatives. Changing boundary condition one can thus change the value on a derivative in a discontinuity point. This can be interpreted as an alternative definition of phase transition.

Definition 2.7 *A phase transition point is a point in which the value of some derivative of the infinite volume pressure depends on boundary conditions even at the thermodynamic limit.*

Thus when the system is sensible to change of boundary conditions even in the infinite volume limit we say that there is a phase transition.
By this rough discussion we see that lack of analiticity of the pressure or sensitivity of the system to change in boundary conditions are two ways to characterize a phase transition.

Chapter 3

High temperature low density expansion

3.1 The Mayer series

3.1.1 Some Notations about graphs

We first give some definitions about abstract graphs which will be useful below. Let A be any finite set, we denote by $|A|$ the number of elements of A. We denote by $P(A)$ the power set of A (i.e. the set of all subsets of A). We denote by $P^k(A) = \{U \subset A : |U| = k\}$ (i.e. the set of all subsets of A with cardinality k). If $A = \{1, 2, \dots, n\}$, we shortly put $P(\{1, 2, \dots, n\}) \equiv P_n$ and $P^k(\{1, 2, \dots, n\}) \equiv P_n^k$.

Definition 3.1 *A graph g is a pair $g = (V_g, E_g)$ where V_g is a countable set and $E_g \subset P^2(V_g)$. The set V_g is called the vertex set of g and the elements of V_g are called vertices of the graph g, while E_g is called the edge set of g and the elements of E_g are called called edges of the graph g.*

Given two graphs $g = (V_g, E_g)$ and $f = (V_f, E_f)$ we say that $f \subset g$ if $V_f \subset V_g$ and $E_f \subset E_g$.

Definition 3.2 *A graph $g = (V_g, E_g)$ is said to be connected if for any pair B, C of subsets of V_g such that $B \cup C = V_g$ and $B \cap C = \emptyset$, there is a $e \in E_g$ such that $e \cap B \neq \emptyset$ and $e \cap C \neq \emptyset$.*

We denote $\mathcal{G}_V$ the set of all graphs with vertex set V and by G_V the set of all connected graphs with vertex set V. We will shortly denote $\mathcal{G}_n$ the set $\mathcal{G}_{\{1,2,\dots,n\}}$ of all graphs with vertex set $\{1, 2, \dots, n\}$ and by G_n the set $G_{\{1,2,\dots,n\}}$ of all connected graphs with vertex set $\{1, 2, \dots, n\}$. We will also denote shortly

$$[n] = \{1, 2, \dots, n\} \tag{3.1}$$

$$\mathrm{E}_n = \{\{i, j\} \subset [n]\} \tag{3.2}$$

In particular E_n is the set of all unordered pairs $\{i, j\}$ in the set $[n]$.

Definition 3.3 *A graph $\tau \in G_V$ such that $|E_\tau| = |V| - 1$ is called a tree graph.*

We denote T_V the set of all tree graphs with vertex set V. Note that $T_V \subset G_V$. The set of all the tree graph over $[n]$ will be denoted by T_n.

Definition 3.4 *Let $g = (V_g, E_g)$ be a graph and let $x \in V_g$. Then the degree (or number of incidence) d_x of the vertex $x \in V_g$ in g is the number of edges $e \in E_g$ such that $x \in e$.*

Observe that in a tree $\tau \in T_n$ the numbers of incidence $d_1, \ldots, d_n$ at vertices $1, \ldots, n$ satisfies the identity

$$d_1 + d_2 + \ldots + d_n = 2n - 2 \tag{3.3}$$

This identity is because any tree τ has $n-1$ edges and each edge has two vertices. Moreover we have clearly the following bound for the number of incidence d_i in a vertex of tree τ.

$$1 \leq d_i \leq n - 1 \tag{3.4}$$

This is again because any tree τ has $n - 1$ edges.
The number of trees in T_n is explicitly computable using the Cayley formulas. These formulas can be presented by the the following lemma.

Lemma 3.1 *The number of trees in T_n with numbers of incidence in vertices $1, 2, \ldots, n$ fixed at the values $d_1, d_2, \ldots, d_n$ is given by*

$$\sum_{\substack{\tau \in T_n \\ d_1, d_2, \ldots, d_n \text{ fixed}}} 1 = \frac{(n-2)!}{\prod_{i=1}^{n} (d_i - 1)!} \tag{3.5}$$

Moreover

$$|T_n| = \sum_{\tau \in T_n} 1 = n^{n-2} \tag{3.6}$$

Proof. We first show (3.5) by induction. Formula (3.5) is trivially true if $n = 2$. We suppose now that (3.5) is true for all tree with n vertices and with any incidence numbers $d_1, \ldots, d_n$ and we will prove that (3.5) holds also for all tree with $n + 1$ vertices and with any incidence numbers $d_1, \ldots, d_n, d_{n+1}$.
Take a tree τ with $n+1$ vertices and numbers of incidence in vertices fixed at the values $d_1, \ldots, d_{n+1}$. Such tree has at least two vertices with number of incidence equal to 1. Without lost in generality let us suppose that $d_{n+1} = 1$ (otherwise we can always rename the vertices of τ). Now if $d_{n+1} = 1$, then there is an edge in τ which links the vertex $n + 1$ with some vertex $j \in \{1, \ldots, n\}$. This vertex j has surely $d_j \geq 2$, i.e. there is at least another edge staring from j which ends in some other vertex $k \in \{1, \ldots, n\}$, since τ is connected. So we can count trees with $n + 1$ vertices and fixed number of incidence proceeding as follows. For each time we fix the edge joining $n + 1$ to some vertex $j \in \{1, \ldots, n\}$,

we count all trees in $\{1,\dots,n\}$ with number of incidence fixed at the values $d_1,\dots,d_{j-1},d_j-1,d_{j+1},\dots d_n$, by using the induction hypothesis. Namely,

$$\sum_{\substack{\tau\in T_{n+1}\\ d_1,d_2,\dots,d_{n+1}\ \text{fixed}}} 1 \;=$$

$$= \sum_{\substack{j\in\{1,2,\dots,n\}:\\ d_j\geq 2}} \frac{(n-2)!}{(d_1-1)!\cdots(d_{j-1}-1)!(d_j-2)!(d_{j+1}-1)!\cdots(d_n-1)!} \;=$$

$$= \sum_{\substack{j\in\{1,2,\dots,n\}:\\ d_j\geq 2}} \frac{(n-2)!(d_j-1)}{(d_1-1)!\cdots(d_j-1)!\cdots(d_n-1)!(d_{n+1}-1)!} \;=$$

$$= \frac{(n-2)!}{\prod_{i\,=\,1}^{n+1}(d_i-1)!}\sum_{j\,=\,1}^{n}(d_j-1)$$

where in the second line we just rewrite

$$\frac{1}{(d_j-2)!} \;=\; \frac{(d_j-1)}{(d_j-1)!(d_{n+1}-1)!}$$

using the assumption that $d_{n+1} \;=\; 1$ and hence $(d_{n+1}-1)! \;=\; 0! \;=\; 1$. Now using again that $d_{n+1}-1=0$ and by (3.3) with $n+1$ in place of n, we get

$$\sum_{j\,=\,1}^{n}(d_j-1)=\sum_{j\,=\,1}^{n+1}(d_j-1) \;=\; \sum_{j\,=\,1}^{n+1}d_j-(n+1) \;=\; 2(n+1)-2-(n+1) \;=\; n-1$$

hence

$$\sum_{\substack{\tau\in T_{n+1}\\ d_1,d_2,\dots,d_{n+1}\ \text{fixed}}} 1 \;=\; \frac{(n-2)!}{\prod_{i=1}^{n+1}(d_i-1)!}(n-1) \;=\; \frac{(n-1)!}{\prod_{i=1}^{n+1}(d_i-1)!}$$

and (3.5) is proved. Now (3.6) is an easy consequence of (3.5). As a matter of fact, just observe, using (3.3), (3.4), and (3.5), that

$$\sum_{\tau\in T_n}1 \;=\; \sum_{\substack{d_1,\dots,d_n:\ 1\leq d_i\leq n-1\\ d_1+\dots+d_n=2n-2}}\frac{(n-2)!}{\prod_{i=1}^{n}(d_i-1)!} \;=\; \sum_{\substack{s_1,\dots,s_n:\ 0\leq s_i\leq n-2\\ s_1+\dots+s_n\ =\ n-2}}\frac{(n-2)!}{\prod_{i=1}^{n}s_i!} \;=\; n^{n-2}$$

□

3.1.2 Mayer Series: definition

We now come back to the affirmation that a system of particles, interacting via a reasonable potential energy (e.g. defined via a stable and tempered pair potential) should look as a gas at sufficiently low density and/or sufficiently high temperature, hence the pressure of such system should be analytic in the thermodynamic parameters in this region. Since we are considering just the

Grand Canonical ensemble, the region of high temperature and low density will be in the case the region of λ small and β small.
Let us thus consider again the grand Canonical partition function

$$\Xi(\beta,\Lambda,\lambda) \;=\; 1+\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\int_\Lambda dx_1\ldots\int_\Lambda dx_n\; e^{-\beta U(x_1,\ldots,x_n)} \tag{3.7}$$

with

$$U(x_1,\ldots,x_n) \;=\; \sum_{1\le i<j\le n} V(x_i-x_j) \tag{3.8}$$

where $V(x)$ is a stable and tempered pair potential. We denote shortly $x_{[n]}$ the set of coordinates $x_1,\ldots,x_n$ and if $I\subset[n]$ then x_I will denote the set of coordinates x_i with $i\in I$. Hence for $I\subset[n]$

$$U(x_I)=\sum_{\{i,j\}\subset I} V(x_i-x_j)$$

We now look for an expansion of the log of such function valid for λ small. This is clearly possible, due to the structure of $\Xi(\beta,\Lambda,\lambda)$. In fact, as a power series of λ the partition function has the form

$$\Xi(\beta,\Lambda,\lambda) \;=\; 1+c_1\lambda+c_2\lambda^2+c_3\lambda^3+\ldots \;=\; 1+O(\lambda)$$

and $\ln(1+O(\lambda))$ can indeed be expanded also in power series of λ. Let us show that there exists a formal expansion of $\ln\Xi$ in powers of λ called the Mayer series of the pressure. We thus prove now the following theorem.

Theorem 3.1 *Let $\Xi(\beta,\Lambda,\lambda)$ be defined as in (3.7) and (3.8). Then*

$$\frac{1}{|\Lambda|}\ln\Xi(\beta,\Lambda,\lambda) \;=\; \sum_{n=1}^{\infty} C_n(\beta,\Lambda)\lambda^n \tag{3.9}$$

where

$$C_n(\beta,\Lambda) \;=\; \frac{1}{n!}\frac{1}{|\Lambda|}\int_\Lambda dx_1\ldots\int_\Lambda dx_n \Phi^T(x_1,\ldots,x_n) \tag{3.10}$$

with $\Phi^T(x_1,\ldots,x_n)$ given by the following expressions:

A)

$$\Phi^T(x_1,\ldots,x_n) \;=\; \begin{cases} \sum\limits_{g\in G_n}\;\prod\limits_{\{i,j\}\in E_g}\left[e^{-\beta V(x_i-x_j)}-1\right] & \text{if } n\ge 2 \\ 1 & \text{if } n \;=\; 1 \end{cases} \tag{3.11}$$

where recall that G_n is the set of all connected graphs in $\{1,2,\ldots,n\}$ and if $g\in G_n$ then its edge set is denoted by E_g.

B)

$$\Phi^T(x_1,\ldots,x_n) \;=\; \sum_{k=1}^{n}(-1)^{k-1}(k-1)!\sum_{\{I_1,I_2,\ldots,I_k\}\in\Pi_n} e^{-\beta\sum_{\alpha=1}^{k}U(x_{I_\alpha})} \tag{3.12}$$

where Π_n is the set of all partitions of $[n]$, $I=\{i_1,\ldots,i_k\}\subset[n]$ and $x_I=(x_{i_1},\ldots,x_{i_k})$ with $U(x_I)=U(x_{i_1},\ldots,x_{i_k})$.

The series in the r.h.s. of (3.9) is called the Mayer series and the coefficient $C_n(\beta,\Lambda)$ $(n \geq 1)$ is called Mayer (or Ursell) coeffcient of order n. For the moment this series is merely a formal series, since theorem does not say if the series converges or not.
Proof of A.

$$\Xi(\beta,\Lambda,\lambda) \;=\; 1+\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\int_\Lambda dx_1\ldots\int_\Lambda dx_n\; e^{-\beta\sum_{1\leq i<j\leq n}V(x_i-x_j)} \;=$$

$$=\; 1+\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\int_\Lambda dx_1\ldots\int_\Lambda dx_n\;\prod_{1\leq i<j\leq n} e^{-\beta V(x_i-x_j)} \;=$$

$$=\; 1+\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\int_\Lambda dx_1\ldots\int_\Lambda dx_n\;\prod_{1\leq i<j\leq n}\Big[\Big(e^{-\beta V(x_i-x_j)}-1\Big)+1\Big] \;=$$

Develop now the product in the factor

$$\prod_{1\leq i<j\leq n}\Big[\Big(e^{-\beta V(x_i-x_j)}-1\Big)+1\Big]$$

Then it is not difficult to see that it can be rewritten as

$$\prod_{1\leq i<j\leq n}\Big[\Big(e^{-\beta V(x_i-x_j)}-1\Big)+1\Big] \;=\; \sum_{g\in\mathcal{G}_n}\prod_{\{i,j\}\in E_g}\Big[e^{-\beta V(x_i-x_j)}-1\Big]$$

where recall that $\mathcal{G}_n$ is the set of all graphs (connected or not connected) in the set $\{1,2\ldots,n\}$. In $\mathcal{G}_n$ we also include the empty graph, i.e the graph g such that $E_g \;=\; \emptyset$, and its contribution is the factor 1 in the development of the product.
We now reorganize the sum over graphs in $\mathcal{G}_n$.
For $1\leq k\leq n$, let $\{I_1,I_2,\ldots,I_k\}$ denote a partition of the set $[n]$. Namely, for any $i,j \;=\; 1,2\ldots,k$ we have that $I_i\neq\emptyset$, $I_i\cap I_j \;=\; \emptyset$ and $\cup_{j\;=\;1}^{k} I_j \;=\; [n]$. We denote by x_{I_j} the set of coordinates x_i with $i\in I_j$. We also denote with Π_n the set of all partitions of the set $[n]$.
Then it is not difficult to see that

$$\sum_{g\in\mathcal{G}_n}\prod_{\{i,j\}\in E_g}\Big[e^{-\beta V(x_i-x_j)}-1\Big] \;=\; \sum_{k=1}^{n}\sum_{\{I_1,I_2,\ldots,I_k\}\in\Pi_n}\prod_{j=1}^{k}\Phi^T(x_{I_j}) \qquad (3.13)$$

where

$$\Phi^T(x_{I_j}) \;=\; \begin{cases}\displaystyle\sum_{g\in G_{I_j}}\prod_{\{l,s\}\in E_g}\big[e^{-\beta V(x_l-x_s)}-1\big] & \text{if } |I_j|\geq 2\\ 1 & \text{if } |I_j| \;=\; 1\end{cases}$$

Note now that $\sum_{g\in G_{I_j}}$ runs over all *connected* graphs in the set I_j.

Hence

$$\Xi(\beta,\Lambda,\lambda) \;=\; 1+\sum_{n\,=\,1}^{\infty}\frac{\lambda^n}{n!}\int_\Lambda dx_1\ldots\int_\Lambda dx_n\sum_{k\,=\,1}^{n}\sum_{\{I_1,I_2,\ldots,I_k\}\in\Pi_n}\prod_{j\,=\,1}^{k}\Phi^T(x_{I_j})$$

Now observe that each $\Phi^T(x_{I_j})$ depends only on the set of coordinates x_i with $i\in I_j$ and since $I_1,\ldots I_k$ is a partition of $\{1,\ldots,n\}$ we can write

$$\int_\Lambda dx_1\ldots\int_\Lambda dx_n \;=\; \int_{\Lambda^{|I_1|}} dx_{I_1}\ldots\int_{\Lambda^{|I_k|}} dx_{I_k}$$

where of course $dx_{I_i} \;=\; \prod_{i\in I_j} dx_i$. Hence

$$\Xi(\beta,\Lambda,\lambda) \;=\; 1+\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\sum_{k=1}^{n}\sum_{\{I_1,I_2,\ldots,I_k\}\in\Pi_n}\prod_{j\,=\,1}^{k}\int_{\Lambda^{|I_j|}} dx_{I_j}\Phi^T(x_{I_j})$$

Observe now that the factor

$$\int_{\Lambda^{|I_j|}} dx_{I_j}\Phi^T(x_{I_j})$$

depends only on $|I_j|$ and not anymore from x_i (since space coordinates are integrated) and not even on I_j since $i\in I_j$ is just an index attached to a mute variable. I.e.

$$\int_{\Lambda^{|I_j|}} dx_{I_j}\Phi^T(x_{I_j}) \;=\; \int_\Lambda dx_1\ldots\int_\Lambda dx_{|I_j|}\Phi^T(x_1,\ldots x_{|I_j|}) \;=\; \phi(|I_j|)$$

Note that the numbers $|I_j|$ are positive integers subjected to the condition $\sum_{j=1}^{k}|I_j|=n$. Hence

$$\Xi_\Lambda(\beta,\lambda) \;=\; 1+\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\sum_{k=1}^{n}\sum_{\{I_1,I_2,\ldots,I_k\}\in\Pi_n}\prod_{j=1}^{k}\phi(|I_j|)$$

Now observe that

$$\sum_{\{I_1,I_2,\ldots,I_k\}\in\Pi_n}\prod_{j=1}^{k}\phi(|I_j|) \;=\; \frac{1}{k!}\sum_{\substack{m_1,\ldots,m_k:\;\; m_i\geq 1\\ m_i+\ldots+m_k=n}}\;\sum_{\substack{I_1,\ldots,I_k:\\ |I_1|=m_1,\ldots,|I_k|=m_k}}\prod_{j=1}^{k}\phi(m_j) \;=$$

$$=\; \frac{1}{k!}\sum_{\substack{m_1,\ldots,m_k:\;\; m_i\geq 1\\ m_1+\ldots+m_k\;=\;n}}\prod_{j=1}^{k}\phi(m_j)\sum_{\substack{I_1,\ldots,I_k:\\ |I_1|=m_1,\ldots,|I_k|=m_k}} 1$$

I.e.

$$\sum_{\{I_1,I_2,\ldots,I_k\}\in\Pi_n}\prod_{j=1}^{k}\phi(|I_j|)=\frac{1}{k!}\sum_{\substack{m_1,\ldots,m_k:\;\; m_i\geq 1\\ m_1+\ldots m_k=n}}\prod_{j=1}^{k}\phi(m_j)\frac{n!}{m_1!\ldots m_k!} \tag{3.14}$$

Indeed, $n!/(m_1!\dots m_k!)$ is the number of partitions of $[n]$ in k subsets $I_1, ..., I_k$ such that the numbers $|I_1|, ..., |I_k|$ are fixed at the values $m_1, \dots, m_k$ respectively. We have also to divide by $k!$ because the same partition $I_1, \dots I_k$ of $[n]$ appears exactly $k!$ times in the sum

$$\sum_{\substack{m_1,\dots,m_k:\, m_i\geq 1\\ m_1+\dots m_k=n}} \quad \sum_{\substack{I_1,\dots,I_k:\\ |I_1|=m_1,\dots,|I_k|=m_k}}$$

Hence we can write

$$\Xi_\Lambda(\beta,\lambda) \;=\; 1+\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\sum_{k=1}^{n}\frac{1}{k!}\sum_{\substack{m_1,\dots,m_k:\, m_i\geq 1\\ m_1+\dots+m_k=n}}\frac{n!}{m_1!\dots m_k!}\prod_{j=1}^{k}\phi(m_j) \;=$$

$$=\; 1+\sum_{n=1}^{\infty}\sum_{k=1}^{n}\frac{1}{k!}\sum_{\substack{m_1,\dots,m_k:\, m_i\geq 1\\ m_1+\dots+m_k=n}}\frac{\lambda^{m_1}\cdots\lambda^{m_k}}{m_1!\dots m_k!}\prod_{j=1}^{k}\phi(m_j) \;=$$

$$=\; 1+\sum_{k=1}^{\infty}\frac{1}{k!}\sum_{\substack{m_1,\dots,m_k\\ m_i\geq 1}}\frac{\lambda^{m_1}\cdots\lambda^{m_k}}{m_1!\dots m_k!}\prod_{j=1}^{k}\phi(m_j) \;=\; 1+\sum_{k=1}^{\infty}\frac{1}{k!}\left[\sum_{m=1}^{\infty}\frac{\lambda^m}{m!}\phi(m)\right]^k$$

i.e we have found

$$\Xi_\Lambda(\beta,\lambda) \;=\; 1+\sum_{k=1}^{\infty}\frac{1}{k!}\left[\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\phi(n)\right]^k \;=\; \exp\left\{\left[\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\phi(n)\right]\right\}$$

Hence

$$\ln\Xi_\Lambda(\beta,\lambda) \;=\; \sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\phi(n) \;=\; \sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\int_\Lambda dx_1\dots\int_\Lambda dx_n\Phi^T(x_1,\dots,x_n) \;=$$

$$=\; \sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\int_\Lambda dx_1\dots\int_\Lambda dx_n\sum_{g\in G_n}\prod_{\{i,j\}\in E_g}\left[e^{-\beta V(x_i-x_j)}-1\right]$$

which concludes the proof of part A) of the theorem.□

Proof of B). We start by writing

$$Z_n=\int_\Lambda dx_1\dots\int_\Lambda dx_n\; e^{-\beta U(x_{[n]})}$$

and in general, if $I\subset[n]$,

$$Z_I=\int_{\Lambda^{|I|}} dx_I\; e^{-\beta U(x_I)}=Z_{|I|}$$

so that

$$\Xi_\Lambda(\beta,\lambda) \;=\; 1+\sum_{n=1}^{\infty}\frac{\lambda^n}{n!}Z_n$$

Hence, formally

$$\ln \Xi_\Lambda(\beta,\lambda) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left[\sum_{n=1}^{\infty} \frac{\lambda^n}{n!} Z_n\right]^k = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{\substack{m_1,\dots,m_k \\ m_i \geq 1}} \prod_{j=1}^{k} \frac{\lambda^{m_j} Z_{m_j}}{m_j!} =$$

$$= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{n=k}^{\infty} \frac{\lambda^n}{n!} \sum_{\substack{m_1,\dots,m_k:\ m_i \geq 1 \\ m_1+\dots+m_k=n}} (\prod_{j=1}^{k} Z_{m_j}) \frac{n!}{m_1! \cdots m_k!}$$

Now, by equation (3.14) we have that

$$\sum_{\substack{m_1,\dots,m_k:\ m_i \geq 1 \\ m_1+\dots+m_k=n}} (\prod_{j=1}^{k} Z_{m_j}) \frac{n!}{m_1! \cdots m_k!} = k! \sum_{\{I_1,I_2,\dots,I_k\} \in \Pi_n} \prod_{j=1}^{k} Z_{(|I_j|)} =$$

$$= k! \int_\Lambda dx_1 \dots \int_\Lambda dx_n \sum_{\{I_1,I_2,\dots,I_k\} \in \Pi_n} e^{-\sum_{\alpha=1}^{k} \beta U(x_{I_\alpha})}$$

Thus we get

$$\ln \Xi_\Lambda(\beta,\lambda) = \sum_{k=1}^{\infty} (-1)^{k-1}(k-1)! \sum_{n=k}^{\infty} \frac{\lambda^n}{n!} \int_\Lambda dx_1 \dots \int_\Lambda dx_n \sum_{\{I_1,I_2,\dots,I_k\} \in \Pi_n} e^{-\beta \sum_{\alpha=1}^{k} U(x_{I_\alpha})}$$

Finally, exchanging the sum over k with the sum over n we get

$$\ln \Xi_\Lambda(\beta,\lambda) = \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} \int_\Lambda dx_1 \dots \int_\Lambda dx_n \sum_{k=1}^{n} (-1)^{k-1}(k-1)! \sum_{\{I_1,\dots,I_k\} \in \Pi_n} e^{-\beta \sum_{\alpha=1}^{k} U(x_{I_\alpha})}$$

whence (3.12) follows. □

3.1.3 The combinatorial problem

The very structure of the Mayer series for $\ln \Xi$ hides a very hard combinatorial problem. To understand this problem let us find a bound uniform in Λ for the coefficient of order n of the Mayer series, as given by formula (3.10), i.e.

$$C_n(\beta,\Lambda) = \frac{1}{n!} \frac{1}{|\Lambda|} \int_\Lambda dx_1 \dots \int_\Lambda dx_n \Phi^T(x_1,\dots,x_n) =$$

$$= \frac{1}{n!} \frac{1}{|\Lambda|} \int_\Lambda dx_1 \dots \int_\Lambda dx_n \sum_{g \in G_n} \prod_{\{i,j\} \in E_g} \left[e^{-\beta V(x_i - x_j)} - 1\right] \tag{3.15}$$

Note that $C_n(\beta,\Lambda)$ is a function of β and Λ. The simplest way to bound $|C_n(\beta,\Lambda)|$ is as follows

$$|C_n(\beta,\Lambda)| \leq \frac{1}{n!} \frac{1}{|\Lambda|} \int_\Lambda dx_1 \dots \int_\Lambda dx_n \left| \sum_{g \in G_n} \prod_{\{i,j\} \in E_g} \left[e^{-\beta V(x_i - x_j)} - 1\right] \right| \leq$$

$$\leq \frac{1}{n!}\frac{1}{|\Lambda|}\sum_{g\in G_n}\int_\Lambda dx_1\ldots\int_\Lambda dx_n\prod_{\{i,j\}\in E_g}\left|e^{-\beta V(x_i-x_j)}-1\right|$$

Now observe that, if $V(x)$ is stable then necessarily it is bounded below by $2B$, i.e.

$$V(x_i-x_j)\geq -2B$$

which is simply the stability condition $U(x_1,\ldots,x_n)\geq -Bn$ for the case in which $n=2$. Thus factor

$$|e^{-\beta V(x_i-x_j)}-1|\leq \max\{1,e^{2\beta B}-1\}\leq 1 \tag{3.16}$$

for β sufficiently small.
Now observe that any connected graph $g\in G_n$ contains at least one tree $\tau\in T_n$. Let us thus choose for each g a tree τ such that $\tau\subset g$. It is allowed to choose the same tree for different g's.
So, by (3.16) we have

$$\prod_{\{i,j\}\in E_g}\left|e^{-\beta V(x_i-x_j)}-1\right|\leq\prod_{\{i,j\}\in E_\tau}\left|e^{-\beta V(x_i-x_j)}-1\right|$$

We now prove the following

Proposition 3.1 *For any $\tau\in T_n$ it holds the following inequality*

$$\int_\Lambda dx_1\ldots\int_\Lambda dx_n\prod_{\{i,j\}\in E_\tau}\left|e^{-\beta V(x_i-x_j)}-1\right| \;\leq\; |\Lambda|\left[\int_{\mathbb{R}^3}|e^{-\beta V(x)}-1|dx\right]^{n-1} \tag{3.17}$$

Proof. Let to fix ideas $E_\tau=\{i_1,j_1\},\ldots\{i_{n-1},j_{n-1}\}$, so that

$$\int_\Lambda dx_1\ldots\int_\Lambda dx_n\prod_{\{i,j\}\in E_\tau}\left|e^{-\beta V(x_i-x_j)}-1\right| \;=$$

$$=\;\int_\Lambda dx_1\ldots\int_\Lambda dx_n\prod_{k=1}^{n-1}\left|e^{-\beta V(x_{i_k}-x_{j_k})}-1\right|$$

Define a change of variables in the integral as follows

$$y_k \;=\; x_{i_k}-x_{j_k},\qquad \forall k\;=\;2\ldots,n$$

$$y_1\;=\;x_1$$

The Jacobian of this transformation is clearly 1, then

$$\int_\Lambda dx_1\ldots\int_\Lambda dx_n\prod_{k\,=\,1}^{n-1}\left|e^{-\beta V(x_{i_k}-x_{j_k})}-1\right| \;\leq$$

$$\leq\;\int_\Lambda dx_1\int_{\mathbb{R}^3}dy_2\ldots\int_{\mathbb{R}^3}dy_n\prod_{j\,=\,2}^{n}\left|e^{-\beta V(y_j)}-1\right| \;=$$

$$= |\Lambda| \left[\int_{\mathbb{R}^3} |e^{-\beta V(x)} - 1| dx \right]^{n-1}$$

□

It is now a simple exercise to see that stability and temperness imply that

$$\int_{\mathbb{R}^3} |e^{-\beta V(x)} - 1| dx \ = \ C(\beta) < \infty$$

Actually, it is not even neccessary for the potential to be stable. A sufficient condition for the finiteness of $C(\beta)$ is, if $S(a)$ is a sphere of radius a and center in $x = 0$, $V(x)$ integrable in $\mathbb{R}^3 \backslash S(a)$ and non negative for $|x| \leq a$.
So we get

$$|C_n(\beta, \Lambda)| \leq \frac{[C(\beta)]^{n-1}}{n!} \sum_{g \in G_n} 1 \ = \ \frac{[C(\beta)]^{n-1}}{n!} B_n$$

where B_n is the number of connected graphs in the set $\{1, 2, \ldots, n\}$. Hence $|C_n(\beta, \Lambda)|$ has as upper bound a quantity uniform in Λ, which is good, hence the coefficents for the Mayer series of the pressure indeed admit a bound uniform in Λ!
Thus looking for the absolute convergence of the Mayer series for the finite volume pressure we get

$$|\beta p_\Lambda(\beta, \lambda)| \leq \sum_{n=1}^{\infty} |C_n((\beta, \Lambda))\lambda^n| \leq \sum_{n=1}^{\infty} \frac{[C(\beta)]^{n-1}}{n!} B_n \lambda^n$$

If we could now get a good bound for B_n, the number of connected graphs in a finite set of n elements, e.g. a bound at worst such as $n!B^n$, we would have the proof that the infinite volume pressure is analytic in the region

$$|\lambda| C(\beta) B < 1$$

with the further restriction (3.16) on β. But it is not the case since it is easy to show that for all $n \geq 2$

$$B_n \geq 2^{\frac{(n-1)(n-2)}{2}} \tag{3.18}$$

Indeed, first observe that the number of graphs, denoted with $\mathcal{B}_n$, (connected or not connected) in the set $\{1, 2, \ldots, n\}$ is

$$\mathcal{B}_n \ = \ \sum_{g \in \mathcal{G}_n} 1 \ = \ 2^{\frac{1}{2} n(n-1)} \tag{3.19}$$

As a matter of fact, we can construct a one-to-one correspondence between the set of graphs in $\mathcal{G}_n$ and the set Ω_n of sequences $\{\sigma_{\{i,j\}}\}_{\{i,j\} \subset \{1,\ldots,n\}}$ with $\sigma_{\{i,j\}} = 0, 1$, with the rule that, $\sigma_{\{i,j\}} = 1$ if $\{i, j\} \in E_g$ and $\sigma_{\{i,j\}} = 0$ if $\{i, j\} \notin E_g$. Hence a graph $g \in \mathcal{G}_n$ can be viewed as a ordered sequence of $n(n-1)/2$ numbers which can be either 0 or 1. Thus the total number of graphs equals the total number of such sequences which is clearly $2^{n(n-1)/2}$.

It is now simple to get a bound for $B_n = |G_n|$. Consider the subset of $\tilde{\Omega}_n \subset \Omega_n$ formed by sequences $\{\sigma_{\{i,j\}}\}_{\{i,j\}\subset\{1,\dots,n\}}$ such that $\sigma_{\{1,2\}} = 1, \sigma_{\{2,3\}} = 1, \dots, \sigma_{\{n-1,n\}} = 1$, hence $n-1$ links are fixed while $n(n-1)/2 - (n-1) = (n-1)(n-2)/2$ are arbitrary. Clearly any graphs corresponding to a sequence in $\tilde{\Omega}_n$ is connected by construction, since it contains the tree $\tau = \{\{1,2\},\{2,3\}...,\{n-1,n\}\}$ and $|\tilde{\Omega}_n| = 2^{(n-1)(n-2)/2}$. Thus

$$B_n \geq 2^{(n-1)(n-2)/2}$$

Such a bound shows that we have no hope to control the (absolute) convergence of the Mayer series if we don't exploit cancellations hidden in the factor

$$\left| \sum_{g\in G_n} \prod_{\{i,j\}\in E_g} \left[e^{-\beta V(x_i - x_j)} - 1 \right] \right|$$

One may think to use the alternative expression (3.12) for the Ursell coefficients $\Phi^T(x_1,\dots,x_n)$ which looks more well behaved as a function of n and also looks more suitable to take advantage of stability. The problem is that the expression (3.12), if bounded naively, behaves badly in the volume $|\Lambda|$. Using naively (3.12) we get the following bound for the absolute value of $C_n(\beta,\Lambda)$.

$$|C_n(\beta,\Lambda)| \leq \frac{1}{n!}\frac{1}{|\Lambda|}\int_\Lambda dx_1 \dots \int_\Lambda dx_n |\Phi^T(x_1,\dots,x_n)| \ \leq$$

$$\leq \ \frac{1}{|\Lambda|}\sum_{k=1}^{n}(k-1)! \sum_{\{I_1,I_2,\dots,I_k\}\in\Pi_n} \frac{1}{n!}\int_\Lambda dx_1 \dots \int_\Lambda dx_n e^{-\beta\sum_{\alpha=1}^{k} U(x_{I_\alpha})}$$

and using then stability we get

$$|C_n(\beta,\Lambda)| \leq \frac{1}{|\Lambda|}\sum_{k=1}^{n}(k-1)! \sum_{\{I_1,I_2,\dots,I_k\}\in\Pi_n} \frac{1}{n!}\int_\Lambda dx_1 \dots \int_\Lambda dx_n e^{+\beta\sum_{\alpha=1}^{k} B|I_\alpha|} \leq$$

$$\leq |\Lambda|^{n-1} e^{\beta B n} \sum_{k=1}^{n} \frac{k!}{n!} \sum_{\{I_1,I_2,\dots,I_k\}\in\Pi_n} 1 = e^{+\beta B}(|\Lambda| e^{\beta B})^{n-1} \sum_{k=1}^{n} \frac{k!}{n!} S(n,k)$$

where the terms $S(n,k)$ are the Stirling numbers of the second kind[1]. It holds

$$S(n,k) \leq \binom{n}{k} k^{n-k}$$

Therefore we obtain

$$|C_n(\beta,\Lambda)| \leq e^{\beta B}(|\Lambda| e^{\beta B})^{n-1} \sum_{k=1}^{n} \frac{k!}{n!}\binom{n}{k} k^{n-k} \ = \ e^{\beta B}(|\Lambda| e^{\beta B})^{n-1} \sum_{k=1}^{n} \frac{k^{n-k}}{(n-k)!} \ =$$

[1] $S(n,k)$ is the number of ways to partition a set of n objects into k non-empty subsets.

$$= e^{\beta B}(|\Lambda|e^{\beta B})^{n-1}\sum_{s=0}^{n-1}\frac{(n-s)^s}{s!} \leq e^{\beta B}(|\Lambda|e^{\beta B})^{n-1}\sum_{s=0}^{n-1}\frac{n^s}{s!} \leq$$

$$\leq e^{\beta B}(|\Lambda|e^{\beta B})^{n-1}\sum_{s=0}^{\infty}\frac{n^s}{s!} = |\Lambda|^{n-1}(e^{+\beta B+1})^n$$

This bound $|C_n(\beta,\Lambda)|$ for has a well combinatioral behavior in n, but it grows as $|\Lambda|^n$ in the volume Λ and hence the (lower) bound on the convergence radius of the Mayer series obtained from this bound shrinks to zero as $\Lambda \to \infty$.

3.2 Convergence of the Mayer series

The best rigorous upper bound on $|C_n(\beta,\Lambda)|$ until recently (and hence the best lower bound on the convergence radius of the Mayer series) for stable and tempered pair potentials was that obtained by Penrose and Ruelle in 1963 [28, 39].

Theorem 3.2 (Penrose-Ruelle) *Let V be a stable and tempered pair potential with stability constant B. Then the n-order Mayer coefficient $C_n(\beta,\Lambda)$ defined in (3.15) is bounded by*

$$|C_n(\beta,\Lambda)| \leq e^{2\beta B(n-2)}n^{n-2}\frac{[C(\beta)]^{n-1}}{n!} \tag{3.20}$$

where

$$C(\beta) = \int_{\mathbb{R}^d} dx\ |e^{-\beta V(x)} - 1| \tag{3.21}$$

Therefore the Mayer series (3.9) converges absolutely, uniformly in Λ, for any complex λ inside the disk

$$|\lambda| < \frac{1}{e^{2\beta B+1}C(\beta)} \tag{3.22}$$

I.e. the convergence radius of the Mayer series (3.9) admits the following lower bound

$$R_V \geq \frac{1}{e^{2\beta B+1}C(\beta)} \tag{3.23}$$

The estimate (3.20) leading to the lower bound (3.23) was obtained by attacking the problem with a rather indirect approach. Namely Penrose and Ruelle looked at an infinite class of functions of the systems, the so called correlation functions, and showed that they can be expressed via an absolute convergent expansion. The next section is entirely devoted to the proof of Theorem 3.2.

3.2.1 Kirkwood-Salsburg equations: the original proof of Theorem 3.2 (according to Penrose)

We are considering a system of particles enclosed in a box Λ and interacting via a pair potential $V(x-y)$ stable and tempered. We thus define the correlation

functions of the system as follows. For any $n \geq 1$

$$\rho_n(x_1, \dots x_n; \lambda) = \frac{1}{\Xi_\Lambda(\beta, \lambda)} \sum_{m=0}^{\infty} \frac{\lambda^{n+m}}{m!} \int_\Lambda dy_1 \dots \int_\Lambda dy_m \, e^{-\beta U(x_1,\dots,x_n,y_1,\dots,y_m)} \tag{3.24}$$

with the convention that

$$U(x_1) = 0 \tag{3.25}$$

and for $n = 0$

$$\rho_0(\emptyset; \lambda) = 1 \tag{3.26}$$

The numbers $\rho_n(x_1, \dots x_n; \lambda)$ represent the probability density of finding in the system n particles at positions $x_1, \dots x_n$ irrespective of where the other particle are. It is important to observe that the functions $\rho_n(x_1, \dots x_n; \lambda)$ are by construction symmetric under permutation of positions $x_1, \dots x_n$. I.e. if $\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\} : i \mapsto \sigma(i)$ is one-to-one (i.e. a permutation) then

$$\rho_n(x_1, \dots, x_n; \lambda) = \rho_n(x_{\sigma(1)}, \dots, x_{\sigma(n)}; \lambda) \tag{3.27}$$

Observe also that, due to stability, the numerator and the denominator of (3.24) are holomorphic functions of λ (see proposition 2.1), hence $\rho_n(x_1, \dots x_n; \lambda)$ is a meromorphic function of λ. The series expansion $\rho(x_1, \dots x_n; \lambda)$ in powers of λ around $\lambda = 0$ has convergence radius at least equal to the convergence radius of $\log \Xi_\Lambda(\beta, \lambda)$ (i.e. where in the region where the denominator $\Xi_\Lambda(\beta, \lambda)$ in l.h.s of (3.24) is free of zeros). Let us write this series expansion as

$$\rho_n(x_1, \dots, x_n; \lambda) = \sum_{\ell=0} \rho_{n,\ell}\,(x_1, \dots, x_n) \lambda^{n+\ell} \tag{3.28}$$

Observe that if we consider the one-point correlation function $\rho_1(x_1)$, it is easy to see that the coefficients $\rho_{1,\ell}(x_1)$ of its series expansion in power of λ, integrated over the volume Λ, are related with the Ursell coefficient of the Mayer expansion for $\log \Xi_\Lambda(\beta, \lambda)$. As a matter of fact, since

$$\rho_1(x_1; \lambda) = \frac{1}{\Xi_\Lambda(\beta, \lambda)} \sum_{m=0}^{\infty} \frac{\lambda^{1+m}}{m!} \int_\Lambda dy_1 \dots \int_\Lambda dy_m \, e^{-\beta U(x_1,y_1,\dots,y_m)} \tag{3.29}$$

it is easy to check that

$$\int_\Lambda \rho_1(x_1; \lambda) dx_1 = \lambda \, \frac{d(\log \Xi_\Lambda(\beta, \lambda))}{d\lambda}$$

Thus

$$\frac{1}{|\Lambda|} \int_\Lambda \rho_{1,\ell-1}\,(x_1) dx_1 = \frac{1}{(\ell-1)!} \frac{1}{|\Lambda|} \int_\Lambda dx_1 \dots \int_\Lambda dx_\ell \, \Phi^T(x_1, \dots, x_\ell) \; = \; \ell C_l(\beta, \Lambda) \tag{3.30}$$

One can now calculate easily the first term in the expansion (3.28) by division in (3.24). As a matter of fact, since $\Xi_\Lambda(\beta, \lambda) = 1 + O(\lambda)$, then also $\Xi_\Lambda^{-1}(\beta, \lambda) = 1 + O(\lambda)$ and the first term of (3.24) is thus the first term of the numerator. Therefore

$$\rho_{n,0}\,(x_1, \dots x_n) = e^{-\beta U(x_1 \dots, x_n)} \tag{3.31}$$

Moreover observe that, by the convention (3.26), we also have

$$\rho_{0,\ell}(\emptyset) = \delta_{0,\ell} \tag{3.32}$$

where δ_{ij} is the Kronecker symbol, i.e $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. Let us now find recurrence relations for the higher coefficients $\rho_{n,\ell}(x_1, \dots x_n)$ for $n \geq 1$. It is easy to see that the functions $\rho_n(x_1, \dots x_n; \lambda)$ satisfies a set of integral functions named the Kirkwood-Salsburg equations. These equations are obtained first by decomposing the energy $U(x_1, \dots, x_n, y_1, \dots, y_m)$ appearing in l.h.s. of (3.24) as

$$U(x_1, \dots, x_n, y_1, \dots, y_m) = W(x_1; x_2, \dots, x_n) + \\ + \sum_{j=1}^{m} V(x_1 - y_j) + U(x_2, \dots, x_n, y_1, \dots, y_m) \tag{3.33}$$

where

$$W(x_1; x_2, \dots, x_n) = \sum_{i=2}^{n} V(x_1 - x_i)$$

Putting (3.33) into (3.24) one gets

$$\rho_n(x_1, \dots x_n; \lambda) = \frac{\lambda e^{-\beta W(x_1; x_2, \dots, x_n)}}{\Xi_\Lambda(\beta, \lambda)} \sum_{m=0}^{\infty} \frac{\lambda^{n-1+m}}{m!} \int_\Lambda dy_1 \dots \int_\Lambda dy_m$$

$$e^{-\beta(\sum_{j=1}^{m} V(x_1 - y_j) + U(x_2, \dots, x_n, y_1, \dots, y_m))} \quad =$$

$$= \frac{\lambda e^{-\beta W(x_1; x_2, \dots, x_n)}}{\Xi_\Lambda(\beta, \lambda)} \sum_{m=0}^{\infty} \frac{\lambda^{n-1+m}}{m!} \int_\Lambda dy_1 \dots \int_\Lambda dy_m$$

$$\prod_{j=1}^{m} \left[\left(e^{-\beta V(x_1 - y_j)} - 1 \right) + 1 \right] e^{-\beta U(x_2, \dots, x_n, y_1, \dots, y_m)}$$

$$= \frac{\lambda e^{-\beta W(x_1; x_2, \dots, x_n)}}{\Xi_\Lambda(\beta, \lambda)} \sum_{m=0}^{\infty} \frac{\lambda^{n-1+m}}{m!} \int_\Lambda dy_1 \dots \int_\Lambda dy_m$$

$$\sum_{s=0}^{m} \sum_{\substack{j_1, \dots j_s \\ 1 \leq j_1 < j_2 < \dots < j_s \leq m}} \prod_{k=1}^{s} (e^{-\beta V(x_1 - y_{j_k})} - 1) e^{-\beta U(x_2, \dots, x_n, y_1, \dots, y_m)}$$

Here above the term with $s = 0$ correspond to 1. Since the variables y are dummy indices, for any term of the sum over indices $j_1, \dots j_s$ we can relabel the variables $y_1, \dots y_m$ in such way that $y_{j_1} = y_1, \dots, y_{j_s} = y_s$ and keeping in account that

$$\sum_{\substack{j_1, \dots j_s \\ 1 \leq j_1 < j_2 < \dots < j_s \leq m}} 1 = \binom{m}{s} = \frac{m!}{s!(m-s)!}$$

and the fact that $U(z_1, \dots z_p)$ is symmetric under permutations of its variables, we obtain

$$\rho_n(x_1, \dots x_n; \lambda) = \frac{\lambda e^{-\beta W(x_1; x_2, \dots, x_n)}}{\Xi_\Lambda(\beta, \lambda)} \sum_{m=0}^{\infty} \lambda^{n-1+m} \int_\Lambda dy_1 \dots \int_\Lambda dy_m$$

$$\sum_{s=0}^{m} \frac{1}{s!(m-s)!} \prod_{k=1}^{s} (e^{-\beta V(x_1 - y_k)} - 1) e^{-\beta U(x_2, \dots, x_n, y_1, \dots, y_m)}$$

We can interchange the sum over m and s by observing that

$$\sum_{m=0}^{\infty} \sum_{s=0}^{m} \dots = \sum_{s=0}^{\infty} \sum_{m=s}^{\infty} \dots$$

and calling $t = m - s$. Hence we get, after a suitable renomination of the dummy indices y

$$\rho_n(x_1, \dots x_n; \lambda) = \frac{\lambda e^{-\beta W(x_1; x_2, \dots, x_n)}}{\Xi_\Lambda(\beta, \lambda)} \sum_{s=0}^{\infty} \frac{1}{s!} \int_\Lambda dy_1 \dots \int_\Lambda dy_s \prod_{k=1}^{s} (e^{-\beta V(x_1 - y_k)} - 1)$$

$$\sum_{t=0}^{\infty} \frac{\lambda^{n-1+s+t}}{t!} \int_\Lambda dy'_1 \dots \int_\Lambda dy'_t e^{-\beta U(x_2, \dots, x_n, y_1, \dots, y_s, y'_1, \dots, y'_t)}$$

i.e.

$$\rho_n(x_1, \dots x_n; \lambda) = \lambda e^{-\beta W(x_1; x_2, \dots, x_n)} \sum_{s=0}^{\infty} \frac{1}{s!} \int_\Lambda dy_1 \dots \int_\Lambda dy_s \prod_{k=1}^{s} (e^{-\beta V(x_1 - y_k)} - 1)$$

$$\rho_{n-1+s}(x_2, \dots x_n, y_1, \dots, y_s; \lambda) \tag{3.34}$$

The latter are the Kirkwood-Salsburg equations. Substituting now (3.28) into (3.34) we get

$$\sum_{k=0}^{\infty} \rho_{n,k}(x_1, \dots x_n) \lambda^k = e^{-\beta W(x_1; x_2, \dots, x_n)} \sum_{s=0}^{\infty} \frac{1}{s!} \int_\Lambda dy_1 \dots \int_\Lambda dy_s \prod_{k=1}^{s} (e^{-\beta V(x_1 - y_k)} - 1)$$

$$\sum_{r=0}^{\infty} \rho_{n-1+s,r}(x_2, \dots x_n, y_1, \dots, y_s) \lambda^{s+r}$$

and equating coefficients with the same power, say k, of λ

$$\rho_{n,k}(x_1, \dots x_n) = e^{-\beta W(x_1; x_2, \dots, x_n)} \sum_{s=0}^{k} \frac{1}{s!} \int_\Lambda dy_1 \dots \int_\Lambda dy_s \prod_{k=1}^{s} (e^{-\beta V(x_1 - y_k)} - 1)$$

$$\rho_{n-1+s,k-s}(x_2, \dots x_n, y_1, \dots, y_s) \tag{3.35}$$

Recalling (3.32), this equation holds for all integers $n \geq 1$ and $k \geq 0$. Note that for $n = 1$ the formula (3.35) degenerates into

$$\rho_{1,k}(x_1) = \sum_{s=0}^{k} \frac{1}{s!} \int_\Lambda dy_1 \dots \int_\Lambda dy_s \prod_{k=1}^{s} (e^{-\beta V(x_1 - y_k)} - 1) \rho_{s,k-s}(y_1, \dots, y_s) \tag{3.36}$$

In order to estimate the convergence radius of the correlation functions $\rho_n(x_1, \dots x_n; \lambda)$ and hence of the $\log \Xi_\Lambda(\beta, \lambda)$, we have to calculate efficient upper bounds on the coefficients $\rho_{n,k}(x_1, \dots x_n)$ of the expansion of the correlations.
Looking to the structure of the equation (3.35) one can observe the following. Given a pair of indices (n, k), let $M = n + k$, then (3.35) says that $\rho_{n,k}$ is a function of the $\rho_{i,j}$ such that $i + j = M - 1$. Therefore it is absolutely natural to look for a bound on $\rho_{n,k}$ by induction on $n + k$. Namely we look for

$$|\rho_{n,M-n}(x_1, \dots x_n)| \le K_{n,M-n} \qquad (n = 1, 2, \dots, M) \tag{3.37}$$

and make the induction on the integer M. By (3.31) and (3.25)

$$K_{1,0} = 1 \tag{3.38}$$

Therefore inequality (3.37) is satisfied for $M = 1$ with $|\rho_{1,0}(x_1)| \le K_{1,0} = 1$. Now, for $M > 1$, we proceed by induction on M. Assume that (3.37) is true $M - 1$ and for all $n = 1, \dots, M - 1$, then, by (3.35)

$$|\rho_{n,M-n}(x_1, \dots x_n)| \le e^{-\beta W(x_1; x_2, \dots, x_n)} \sum_{s=0}^{M-n} \frac{1}{s!} [C(\beta)]^s K_{n-1+s,M-n-s} \tag{3.39}$$

where

$$C(\beta) = \int_{\mathbb{R}^2} \left| e^{-\beta V(x)} - 1 \right| dx \tag{3.40}$$

In order to bound the factor $\exp\{-\beta W(x_1; x_2, \dots, x_n)\}$ in (3.39), we explicitly make use of the symmetry of the function $\rho_{n,M-n}(x_1, \dots x_n)$ under permutation of $x_1, \dots, x_n$.

Let $i \in \{1, 2, \dots, n\}$ and let $\sigma_{1 \leftrightarrows i}$ be a permutation in $\{1, 2, \dots, n\}$ such that $\sigma_{1 \leftrightarrows i}(1) = i$, $\sigma_{1 \leftrightarrows i}(i) = 1$ and $\sigma_{1 \leftrightarrows i}(k) = k$ for all $k \neq 1, i$ (i.e. $\sigma_{1 \leftrightarrows i}$ exchange only i with 1). Then, by stability for at least one $j \in \{1, 2, \dots, n\}$, it holds

$$W(x_{\sigma_{1 \leftrightarrows j}(1)}; x_{\sigma_{1 \leftrightarrows j}(2)}, \dots, x_{\sigma_{1 \leftrightarrows j}(n)}) \ge -2B \tag{3.41}$$

Indeed, by stability we have that

$$\sum_{i=1}^{n} W(x_{\sigma_{1 \leftrightarrows i}(1)}; x_{\sigma_{1 \leftrightarrows i}(2)}, \dots, x_{\sigma_{1 \leftrightarrows i}(n)}) = 2U(x_1, \dots, x_n) \ge -2nB$$

and this immediately implies (3.41) for at least one j. Thus, choosing $j \in [n]$ such that the permutation $\sigma_{1 \leftrightarrows j}$ satisfies (3.41), we have

$$|\rho_{n,M-n}(x_1, \dots x_n)| = |\rho_{n,M-n}(x_{\sigma_{1 \leftrightarrows j}(1)}, \dots x_{\sigma_{1 \leftrightarrows j}(n)})| \le$$

$$\le \; e^{2\beta B} \sum_{s=0}^{M-n} \frac{1}{s!} [C(\beta)]^s K_{n-1+s,M-n-s} \tag{3.42}$$

Hence (3.37) holds also for M provided that, for $n = 1, \dots, M$

$$K_{n,M-n} \ge \; e^{2\beta B} \sum_{s=0}^{M-n} \frac{1}{s!} [C(\beta)]^s K_{n-1+s,M-n-s} \tag{3.43}$$

In conclusion we have proved by induction that (3.37) holds for all $M \geq 1$ if (3.38) holds and (3.43) for $M > 1$. To find the best value for $K_{n,M-n}$ let us solve the set of equations

$$K_{n,M-n} = e^{2\beta B} \sum_{s=0}^{M-n} \frac{1}{s!} [C(\beta)]^s \, K_{n-1+s,M-n-s} \tag{3.44}$$

This equations are recursive. I.e., for a fixed integer M all coefficients $K_{n,M-n}$ with $n = 1, \dots, M$ are functions of coefficients $K_{n,M-1-n}$ with $n = 1, \dots M-1$. So the initial condition (3.38) plus the equations (3.44) determines uniquely all coefficients $K_{n,M-n}$. It is worth to check that its solution is

$$K_{n,\ell} = e^{2\beta B(n+\ell-1)} n(n+\ell)^{\ell-1} \frac{[C(\beta)]^\ell}{\ell!} \tag{3.45}$$

In conclusion we have that the coefficients of the power series in λ given by (3.28) can be bounded as

$$|\rho_{n,\ell}(x_1, \dots x_n)| \leq e^{2\beta B(n+\ell-1)} n(n+\ell)^{\ell-1} \frac{[C(\beta)]^\ell}{\ell!} \tag{3.46}$$

This implies that (3.28) has convergence radius at least

$$R \geq \frac{1}{e^{2\beta B+1} C(\beta)}$$

As a matter of fact, the n-point correalation $|\rho_n(x_1, \dots x_n)|$, see (3.28) and (3.46) is less or equal to

$$|\rho_n(x_1, \dots, x_n; \lambda)| \leq \sum_{\ell=0}^{\infty} e^{2\beta B(n+\ell-1)} n(n+\ell)^{\ell-1} \frac{[C(\beta)]^\ell}{\ell!} \lambda^{n+\ell} \leq$$

$$\leq \sum_{\ell=0}^{\infty} e^{2\beta B(n+\ell-1)} (n+\ell)^{\ell} \, \frac{[C(\beta)]^\ell}{\ell!} \lambda^{n+\ell} \; =$$

$$= \; e^{2\beta B(n-1)} \lambda^n \sum_{\ell=0}^{\infty} \left(1 + \frac{n}{\ell}\right)^{\ell} \, \frac{\ell^\ell}{\ell!} \, \left[e^{2\beta B} C(\beta) \lambda\right]^\ell \leq$$

$$\leq e^{2\beta B(n-1)} \lambda^n \sum_{\ell=0}^{\infty} e^n \, e^\ell \left[e^{2\beta B} C(\beta) \lambda\right]^\ell \leq$$

$$\leq e^{(2\beta B+1)n} \lambda^n \sum_{\ell=0}^{\infty} \left[e^{2\beta B+1} C(\beta) \lambda\right]^\ell$$

It is also interesting to calculate the the bound we obtain for the Ursell coefficients of the Mayer series. Observe first that in place of (3.39) we have, recalling the special case $n = 1$ (3.36),

$$|\rho_{1,M-1}(x_1)| \leq \sum_{s=0}^{M-1} \frac{1}{s!} [C(\beta)]^s \, K_{s,M-1-s} \tag{3.47}$$

(i.e.the factor $e^{-\beta W(x_1;x_2,\ldots,x_n)}$ is not present in the $n=1$ case). So, using (3.44) and (3.45) we get

$$|\rho_{1,\ell}(x_1)| \leq e^{-2\beta B} K_{1,\ell} = e^{2\beta B(\ell-1)}(1+\ell)^{\ell-1}\frac{[C(\beta)]^\ell}{\ell!}$$

Hence, recalling the definitions (3.9) and (3.10) and using (3.30), we get

$$n|C_n(\beta,\Lambda)| \;\leq\; \left| n\frac{\lambda^n}{n!}\frac{1}{|\Lambda|}\int_\Lambda dx_1 \ldots \int_\Lambda dx_n \; \Phi^T(x_1,\ldots,x_n)\right| \;=$$

$$= \left|\frac{1}{|\Lambda|}\int_\Lambda \rho_{1,n-1}(x_1)dx_1\right| \;\leq\; e^{2\beta B(n-2)}(n)^{n-2}\frac{[C(\beta)]^{n-1}}{(n-1)!}$$

I.e. we obtain the bound (3.20)

$$|C_n(\beta,\Lambda)| \leq e^{2\beta B(n-2)} n^{n-2}\frac{[C(\beta)]^{n-1}}{n!} \tag{3.48}$$

which completes the proof of Theorem 3.2.
We stress once again that this bound is valid for any stable pair potential $V(x)$ such that $C(\beta)$ defined in (3.40) is finite. The bound is very efficient, but it is indeed obtained in a rather involved and indirect way. We will now show an alternative way to get the same bound (3.48) in a much more direct way, i.e. obtaining directly a bound for the absolute values of the Ursell coefficients $|C_n(\beta,\Lambda)|$ starting from their explicit expression (3.10).

3.3 The Penrose Tree graph Identity

The tree-graph identity that we present was proposed by Penrose [29] in 1967 and it was based on the existence of a map from the set T_n of the trees with vertex set $[n]$ to the set G_n of the connected graphs with vertex set $[n]$ inducing a so-called *partition scheme* in G_n. This tree graph identity permits to rewrite the Ursell coefficient defined in (3.11), whose expression we recall

$$\Phi^T(x_1,\ldots,x_n) \;=\; \sum_{g\in G_n}\;\prod_{\{i,j\}\in E_g}\left[e^{-\beta V(x_i-x_j)}-1\right] \tag{3.49}$$

in terms of a sum over trees rather than over connected graphs.

Definition 3.5 *A map $\mathfrak{M}: T_n \to G_n$ is called a partition scheme in the set of the connected graphs G_n if, for all $\tau \in T_n$, $\tau \subset \mathfrak{M}(\tau)$ and $G_n = \biguplus_{\tau\in\mathcal{T}}[\tau,\mathfrak{M}(\tau)]$ where $\biguplus$ means disjoint union and $[\tau,\mathfrak{M}(\tau)] = \{g \in G_n : \tau \subset g \subset \mathfrak{M}(\tau)\}$ is a boolean interval (with respect to the set-inclusion).*

Once a partition scheme in G_n has been given, we have the following identity

Theorem 3.3 (General Penrose identity) *Let $V(x)$ be a pair potential. Let $n \geq 2$. Let $\mathfrak{M} : T_n \to G_n$ be a partition scheme in G_n. Then, for any $(x_1, \ldots, x_n) \in \mathbb{R}^{dn}$ the following identity holds*

$$\Phi^T(x_1, \ldots, x_n) \;=\; \sum_{\tau \in T_n} e^{-\beta \sum\limits_{\{i,j\} \in E_{\mathfrak{M}(\tau)} \setminus E_\tau} V(x_i - x_j)} \prod_{\{i,j\} \in E_\tau} \Big(e^{-\beta V(x_i - x_j)} - 1 \Big) \tag{3.50}$$

Proof. Let us pose shortly $V_{ij} = \beta V(x_i - x_j)$. Since G_n is the disjoint union $G_n = \biguplus_{\tau \in T_n} [\tau, \mathfrak{M}(\tau)]$ we can write

$$\sum_{g \in G_n} \prod_{\{i,j\} \in E_g} \big(e^{-V_{ij}} - 1 \big) \;=\; \sum_{\tau \in T_n} \sum_{g \in [\tau, \mathfrak{M}(\tau)]} \prod_{\{i,j\} \in E_g} \big(e^{-V_{ij}} - 1 \big) \;=$$

$$= \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} \big(e^{-V_{ij}} - 1 \big) \sum_{g \in [\tau, \mathfrak{M}(\tau)]} \prod_{\{i,j\} \in E_g \setminus E_\tau} \big(e^{-V_{ij}} - 1 \big) =$$

$$= \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} \big(e^{-V_{ij}} - 1 \big) \sum_{\substack{g \in G_n \\ E_\tau \subset E_g \subset E_{\mathfrak{M}(\tau)}}} \prod_{\{i,j\} \in E_g \setminus E_\tau} \big(e^{-V_{ij}} - 1 \big) \;=$$

$$= \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} \big(e^{-V_{ij}} - 1 \big) \sum_{E \subset E_{\mathfrak{M}(\tau)} \setminus E_\tau} \prod_{\{i,j\} \in E} \big(e^{-V_{ij}} - 1 \big) \;=$$

$$= \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} \big(e^{-V_{ij}} - 1 \big) \prod_{\{i,j\} \in E_{\mathfrak{M}(\tau)} \setminus E_\tau} \big[\big(e^{-V_{ij}} - 1 \big) + 1 \big] \;=$$

$$= \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} \big(e^{-V_{ij}} - 1 \big) \prod_{\{i,j\} \in E_{\mathfrak{M}(\tau)} \setminus E_\tau} e^{-V_{ij}}$$

which concludes the proof. □

In general it is not so simple to check whether a given map $\mathfrak{M} : T_n \to G_n$ is a partition scheme. The proposition below can be useful.

Proposition 3.2 *Assume we have two maps*

$$G_n \underset{\mathfrak{M}}{\overset{\mathfrak{T}}{\rightleftarrows}} T_n$$

such that $\mathfrak{T}^{-1}(\tau) = \{g \in G_n : \tau \subset g \subset \mathfrak{M}(\tau)\}$ for every $\tau \in T_n$. Then $\mathfrak{M}$ is a partition scheme in G_n.

Proof. Since $g \in \mathfrak{T}^{-1}(\mathfrak{T}(g))$, we have $\mathfrak{T}(g) \subset g$ for all $g \in G_n$. In particular, for every tree τ we have $\mathfrak{T}(\tau) \subset \tau$ which implies $\mathfrak{T}(\tau) = \tau$ because both are trees. The relation $\tau \in \mathfrak{T}^{-1}(\tau)$ implies $\tau \subset \mathfrak{M}(\tau)$, i.e., $\mathfrak{T}$ is a surjective and thus the boolean intervals $\mathfrak{T}^{-1}(\tau)$ are nonempty and G_n is the disjoint union of the boolean interval $G_n = \bigcup_{\tau \in T_n} \mathfrak{T}^{-1}(\tau) = \bigcup_{\tau \in T_n} [\tau, \mathfrak{M}(\tau)]$. Hence in view of definition 3.5 we conclude that $\mathfrak{M}$ is a partition scheme in G_n. □

3.3.1 The original Penrose map

The original partition scheme proposed by Penrose was exclusively based on labels $1, \dots, n$ (rather than $x_1, \dots, x_n$) and it involves two explicit maps, say $\mathbf{t} : G_n \to T_n$ and $\mathbf{m} : T_n \to G_n$ satisfying Proposition 3.2. Let us first construct the map $\mathbf{t} : G_n \to T_n$. To define this map we have first of all to choose a root among vertices $1, 2, \dots, n$. So we identify for example a vertex among $\{1, 2, \dots, n\}$ as the root, e.g., to fix the ideas, let the root be the vertex 1 (as in the original paper of Penrose). Once the root 1 has been chosen, let us denote, for any $g \in G_n$, by $d_g(i)$ the graph distance of the vertex i from the root 1 in g. Given thus $g \in G_n$, we construct the tree $\mathbf{t}(g)$ as follows.

1) We first delete all edges $\{i, j\}$ in E_g with $d_g(i) = d_g(j)$.

After this operation we are left with a connected graph g' such that $d_{g'}(i) = d_g(i)$ for all vertices $i = 1, \dots, n$. Moreover each edge $\{i, j\}$ of g' is such that $|d_{g'}(i) - d_{g'}(j)| = 1$.

2) For any $i \neq 1$ let now delete from the graph g' all edges $\{i, j\}$ in $E_{g'}$ such that $d_{g'}(j) = d_{g'}(i) - 1$ except the one with j minimal.

The resulting graph $g'' \doteq \mathbf{t}(g)$ is by construction a connected graph in G_n, i.e. $\mathbf{t}(g) \in G_n$, which is a subgraph of g, i.e. $\mathbf{t}(g) \subset g$, and which has no cycles, i.e. $\mathbf{t}(g) \in T_n$. Observe that the map $\mathbf{t}$ is a surjection from G_n to T_n (because $\mathbf{t}(\tau) = \tau$ for all $\tau \in T_n$).

We now define the map $\mathbf{m} : T_n \to G_n$.
First, observe that if $\tau \in T_n$ is thought as rooted in 1, then vertices of $\tau \in T_n$ may be thought as *partially ordered* so that each vertex i of τ has a unique parent, which we denote by i', and s_i children, denoted by $i^1, \dots, i^{s_i}$. The number s_i is also called the branching factor of i. Of course, the root has no parent. For any $\tau \in T_n$ we also denote by $d_\tau(i)$ the tree distance of the vertex i from the root 1 ($d_\tau(i)$ is also called the *generation number* of the vertex i in τ). If i is a vertex of τ such that $s_i = 0$ (i.e. i has no children) then i is called a *leaf* (or end-point) of τ.

Definition 3.6 (Penrose partition scheme) *The Penrose partition scheme is the map $\mathbf{m} : T_n \to G_n$ such that to each tree $\tau \in T_n$ associates the graph $\mathbf{m}(\tau) \in G_n$ formed by adding to τ all edges $\{i, j\}$ such that either:*

(p1) $d_\tau(i) = d_\tau(j)$ (edges between vertices of the same generation), or

(p2) $d_\tau(j) = d_\tau(i) - 1$ and $j > i'$ (edges between vertices with generations differing by one).

It is clear, by construction, that, for any $\tau \in T_n$, we have $\mathbf{t}(\mathbf{m}(\tau)) = \tau$ and if $g \in G_n$ such that $\mathbf{t}(g) = \tau$ then $g \in [\tau, \mathbf{m}(\tau)]$. In other words the maps $\mathbf{m}$ and $\mathbf{t}$ satisfy Proposition 3.2 and the map $\mathbf{m}$ is a partition scheme.

The original *Penrose identity [29]*, involving the explicit map $\mathbf{m}$, is the following

Theorem 3.4 *Let V be a pair potential and let $\mathbf{m}$ be the map described in Definition 3.6. Then the following identity holds.*

$$\Phi^T(x_1,\dots,x_n) \;=\; \sum_{\tau\in T_n} e^{-\beta\sum_{\{i,j\}\in E_{\mathbf{m}(\tau)}\setminus E_\tau} V(x_i-x_j)} \prod_{\{i,j\}\in E_\tau}\Big(e^{-\beta V(x_i-x_j)}-1\Big) \tag{3.51}$$

The problem with this identity is that, supposing stability of the potential $V(x)$, it is in general very hard (if not impossible) to efficiently estimate the factor $-\sum_{\{i,j\}\in E_{\mathbf{m}(\tau)}\setminus E_\tau} V(x_i-x_j)$ using stability. In other words, let us consider the following conjecture

Conjecture 3.1 *Let $\mathbf{m}$ be the map described in Definition 3.6. Let $V(x)$ be a stable pair potential with stability constant B. Then there exists a constant $\tilde B$ such that for any $n\geq 2$, for any $\tau\in T_n$ and for any $(x_1,\dots,x_n)\in\mathbb{R}^{dn}$ it holds*

$$\sum_{\{i,j\}\in E_{\mathbf{m}(\tau)}\setminus E_\tau} V(x_i-x_j)\geq -\tilde B n$$

If the conjecture above is true, then we would immediately get from (3.51) the inequality

$$|\Phi^T(x_1,\dots,x_n)| \;\leq\; e^{\beta\tilde B n}\sum_{\tau\in T_n}\prod_{\{i,j\}\in\tau}\Big|e^{-\beta V(x_i-x_j)}-1\Big|$$

and thus, recalling (3.15), the Mayer coefficients would be bounded as follows

$$C_n(\beta,\Lambda)\leq\frac{1}{|\Lambda|}\frac{e^{\beta\tilde B n}}{n!}\sum_{\tau\in T_n}\int_\Lambda dx_1\dots\int_\Lambda dx_n\prod_{\{i,j\}\in E_\tau}\Big|e^{-\beta V(x_i-x_j)}-1\Big|\leq$$

$$\leq\frac{e^{\beta\tilde B n}}{n!}\left[\int_{\mathbb{R}^d}dx\,\Big|e^{-\beta V(x)}-1\Big|\right]^{n-1}\sum_{\tau\in T_n}1\;\leq\;\frac{e^{\beta\tilde B n}}{n!}C(\beta)^{n-1}n^{n-2}$$

The latter estimates yield a lower bound for the convergence radius R_V of the Mayer series as

$$R_V\geq\frac{1}{e^{\beta\tilde B+1}C(\beta)}$$

which would have been better than the Penrose-Ruelle bound (3.23) provided $\tilde B\leq 2B$.

Unfortunately, Conjecture 3.1, as far as the map $\mathbf{m}$ of Definition 3.6 is concerned, has never been proven to be true. We will see later that using a partition scheme different from $\mathbf{m}$, it is possible to prove the Conjecture 3.1.

Let us conclude this section by showing a first consequence of the tree graph identity (3.51). We show that when the potential is positive the Mayer series of the pressure has the remarkable property to be an alternate series. Using the tree graph identity (3.51), it is very easy to see the following.

Proposition 3.3 *If $V(x) \geq 0$ then*

$$\Phi^T(x_1, \ldots, x_n) = (-1)^{n-1} |\Phi^T(x_1, \ldots, x_n)|$$

and therefore the Mayer series of the finite volume pressure is, for any Λ, an alternate series, i.e.

$$(-1)^{n-1} C_n(\beta, \Lambda) \geq 0$$

Proof. We have

$$\Phi^T(x_1, \ldots, x_n) \;=\; \sum_{g \in G_n} \prod_{\{i,j\} \in E_g} \left[e^{-\beta V(x_i - x_j)} - 1 \right]$$

and, using the tree graph identity (3.51), with, for any fixed $x_1, \ldots, x_n$, we can write

$$\Phi^T(x_1, \ldots, x_n) \;=\; \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} \left(e^{-\beta V(x_i - x_j)} - 1 \right) e^{-\beta \sum_{\{i,j\} \in E_{\mathbf{m}(\tau)} \setminus E_\tau} V(x_i - x_j)} =$$

$$= \; (-1)^{n-1} \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} \left(1 - e^{-\beta V(x_i - x_j)} \right) e^{-\beta \sum_{\{i,j\} \in E_{\mathbf{m}(\tau)} \setminus E_\tau} V(x_i - x_j)}$$

$$= \; (-1)^{n-1} \left| \Phi^T(x_1, \ldots, x_n) \right|$$

□

An interesting consequence of such proposition is that for positive potential the Mayer series with convergence radius R has surely a singularity at the point $\lambda = -R$. This fact is quite unpleasant, since the singularity occurs in a non physical region.

3.4 The hard-sphere gas via Penrose identity

The original Penrose identity (3.51), although it has not been proven useful to deal with general stable potentials, has been proven to be extremely powerful for purely hard-core pair potentials. Namely for those potentials of the form

$$V^a_{\text{h.c.}}(x) = \begin{cases} +\infty & \text{if } |x| \leq a \\ \\ 0 & \text{otherwise} \end{cases} \tag{3.52}$$

where $a > 0$. We remind that particles interacting via the pair potential $V^a_{\text{h.c.}}(|x|)$ are in fact a system of free hard spheres of diameter $a > 0$.
Observe that the potential $V^a_{\text{h.c.}}$ is stable with stability constant $B = 0$ (since $V^a_{\text{h.c.}}$ is non-negative) and tempered (since $\int_{x>a} V^a_{\text{h.c.}}(|x|)dx = 0$). We can therefore apply Theorem 3.2 which gives us the following lower bound the convergence radius of the Mayer series for such a system

$$R_{V^a_{\text{h.c.}}} \geq \frac{1}{e S_d(a)} \tag{3.53}$$

where

$$S_d(a) = \int_{\mathbb{R}^d} |e^{-\beta V^a_{\text{h.c.}}(|x|)} - 1| dx = \int_{|x|\le a} 1 dx$$

is the volume of the d-dimensional sphere of radius a.
Let us now use the Penrose tree graph identity (3.51) to bound directly the Mayer coefficients of a system of free hard spheres of diameter a (i.e particles interacting via the potential (3.52)).
Given $(x_1, \dots, x_n) \in \mathbb{R}^{dn}$, let us use the short notation $i \sim j$ when $|x_i - x_j| > a$ and $i \nsim j$ when $|x_i - x_j| \le a$.

Definition 3.7 *Let $V^a_{\text{h.c.}}(x)$ be a purely hard core pair potential as in formula (3.52). Then, for any $n \ge 2$ and any $(x_1, \dots, x_n) \in \mathbb{R}^{dn}$, we define the set of Penrose trees $P_a(x_1, \dots, x_n) \subset T_n$ as follows. A tree $\tau \in T_n$ belongs to $P_a(x_1, \dots, x_n)$ if the following conditions are satisfied*

(t0) if $\{i, j\} \in E_\tau$ then $i \nsim j$ (i.e. $|x_i - x_j| \le a$)

(t1) if two vertices i and j are cousins in τ (i.e. such that $d_\tau(i) = d_\tau(j)$), then $i \sim j$ (i.e. $|x_i - x_j| > a$);

(t2) if two vertices i and j are such that $d_\tau(j) = d_\tau(i) - 1$ and $j > i'$, then $i \sim j$ (i.e. $|x_i - x_j| > a$);

Then we have the following

Theorem 3.5 *For any purely hard core pair potential $V^a_{\text{h.c.}}$, for any $n \ge 2$, for any $(x_1, \dots, x_n) \in \mathbb{R}^{dn}$ and for any $\beta \in (0, +\infty)$ it holds the identity*

$$\sum_{g \in G_n} \prod_{\{i,j\} \in E_g} \left(e^{-\beta V^a_{\text{h.c.}}(x_i - x_j)} - 1 \right) \;=\; (-1)^{n-1} \sum_{\tau \in T_n} \mathbb{1}_{\{P_a(x_1, \dots, x_n)\}}(\tau) \qquad (3.54)$$

where

$$\mathbb{1}_{\{P_a(x_1, \dots, x_n)\}}(\tau) = \begin{cases} 1 & \text{if } \tau \in P_a(x_1, \dots, x_n) \\ \\ 0 & \text{otherwise} \end{cases}$$

Proof. Observe that, by definition (3.52) we have that $\beta V^a_{\text{h.c.}}(x) = V^a_{\text{h.c.}}(x)$, for any $\beta \in (0, +\infty)$ and any $x \in \mathbb{R}^d$. Using now (3.51) we have

$$\sum_{g \in G_n} \prod_{\{i,j\} \in E_g} \left(e^{-\beta V^a_{\text{h.c.}}(x_i - x_j)} - 1 \right) \;=\; \sum_{g \in G_n} \prod_{\{i,j\} \in E_g} \left(e^{- V^a_{\text{h.c.}}(x_i - x_j)} - 1 \right) \;=\; \sum_{\tau \in T_n} w_\tau(x_1, \dots, x_n)$$

where, for any

$$w_\tau(x_1, \dots, x_n) = e^{- \sum\limits_{\{i,j\} \in E_{\mathbf{m}(\tau)} \setminus E_\tau} V^a_{\text{h.c.}}(x_i - x_j)} \prod_{\{i,j\} \in E_\tau} \left(e^{- V^a_{\text{h.c.}}(x_i - x_j)} - 1 \right)$$

We have

$$\prod_{\{i,j\} \in E_\tau} \left(e^{- V^a_{\text{h.c.}}(x_i - x_j)} - 1 \right) = \begin{cases} (-1)^{n-1} & \text{if } |x_i - x_j| \le a \text{ for all } \{i, j\} \in E_\tau \\ \\ 0 & \text{otherwise} \end{cases}$$

and

$$e^{-\sum\limits_{\{i,j\}\in E_{\mathbf{m}(\tau)}\setminus E_\tau} V^a_{\text{h.c.}}(x_i-x_j)} = \begin{cases} 1 & \text{if } |x_i-x_j|>a \text{ for all } \{i,j\}\in E_{\mathbf{m}(\tau)}\setminus E_\tau \\ 0 & \text{otherwise} \end{cases}$$

Therefore

$$w_\tau(x_1,\dots,x_n) = \begin{cases} (-1)^{n-1} & \text{if } \begin{smallmatrix} |x_i-x_j|\le a \text{ for all } \{i,j\}\in E_\tau \\ |x_i-x_j|>a \text{ for all } \{i,j\}\in E_{\mathbf{m}(\tau)}\setminus E_\tau \end{smallmatrix} \\ 0 & \text{otherwise} \end{cases}$$

Now, recalling the Definition 3.6 of the map $\mathbf{m}$, we have

$$\{i,j\}\in E_{\mathbf{m}(\tau)}\setminus E_\tau \quad\Longrightarrow\quad \text{either } d_\tau(i)=d_\tau(j) \text{ or } d_\tau(j)=d_\tau(i)-1 \text{ and } j>i'$$

whence

$$w_\tau(x_1,\dots,x_n) = (-1)^{n-1}\begin{cases} 1 & \text{if } \begin{smallmatrix} |x_i-x_j|\le a \text{ for all } \{i,j\}\in E_\tau \\ |x_i-x_j|>a \text{ for all } \{i,j\} \text{ s.t. } d_\tau(i)=d_\tau(j) \\ \text{and for all } \{i,j\} \text{ s.t. } d_\tau(j)=d_\tau(i)-1 \text{ and} j>i' \end{smallmatrix} \\ 0 & \text{otherwise} \end{cases}$$

I.e., recalling Definition 3.7,

$$w_\tau(x_1,\dots,x_n) = (-1)^{n-1}\mathbb{1}_{\{P_a(x_1,\dots,x_n)\}}(\tau) \tag{3.55}$$

□

We now derive a useful inequality from (3.54). To do this we need one more definition.

Definition 3.8 *Let $V^a_{\text{h.c.}}(x)$ be a purely hard core pair potential as in formula (3.52). Then, for any $n\ge 2$ and any $(x_1,\dots,x_n)\in\mathbb{R}^{dn}$, we define the set of weakly Penrose trees $P^*_a(x_1,\dots,x_n)\subset T_n$ as follows. A tree $\tau\in T_n$ belongs to $P^*_a(x_1,\dots,x_n)$ if the following conditions are satisfied*

(t0) if $\{i,j\}\in E_\tau$ then $i\nsim j$ (i.e. $|x_i-x_j|\le a$);

(t1) if two vertices i and j are siblings (i.e. $d_\tau(i)=d_\tau(j)$ and moreover they have the same parent $i'=j'$), then $i\sim j$ (i.e. $|x_i-x_j|>a$);

Then we have the following

Theorem 3.6 *For any purely hard core pair potential $V^a_{\text{h.c.}}$ it holds the inequality*

$$\Big|\sum_{g\in G_n}\prod_{\{i,j\}\in E_g}\Big(e^{-V^a_{\text{h.c.}}(x_i-x_j)}-1\Big)\Big| \;\le\; \sum_{\tau\in T_n} w^*_\tau(x_1,\dots,x_n) \tag{3.56}$$

with

$$w^*_\tau(x_1,\dots,x_n) = \mathbb{1}_{\{P^*_a(x_1,\dots,x_n)\}}(\tau)$$

Proof. The inequality follows immediately from Theorem 3.5 by noting that $P_a^*(x_1,\dots,x_n) \supset P_a(x_1,\dots,x_n)$. □

Observe now that

$$w_\tau^*(x_1,\dots,x_n) = \begin{cases} 1 & \text{if} \quad \begin{array}{l} |x_i-x_j|\le a \text{ for all } \{i,j\}\in E_\tau \\ |x_i-x_j|>a \text{ for all } i,j \text{ siblings in } \tau \end{array} \\ 0 & \text{otherwise} \end{cases} \tag{3.57}$$

Therefore n-order Mayer coefficient for a system interacting via the potential $V_{\text{h.c.}}^a$ is bounded by

$$|C_n(\beta,\Lambda)| \le \frac{1}{n!}\sum_{\tau\in T_n} S_\Lambda(\tau) \tag{3.58}$$

with

$$S_\Lambda(\tau) = \frac{1}{|\Lambda|}\int_\Lambda dx_1 \dots \int_\Lambda dx_n w_\tau^*(x_1,\dots,x_n)$$

By (3.57) we have

$$S_\Lambda(\tau) \;\le\; g_d(d_1)\prod_{i=2}^n g_d(d_i-1) \tag{3.59}$$

where d_i is the degree of the vertex i in τ,

$$g_d(k) \;=\; \int_{\substack{|x_i|\le a \\ |x_i-x_j|>a}} dx_1\dots dx_k \;=\; a^{dk}\int_{\substack{|y_i|\le 1 \\ |y_i-y_j|>1}} dy_1\dots dy_k$$

for k positive integer, and $g_d(0)=1$ by definition. Recalling that $S_d(a)$ denotes the volume of the d-dimensional sphere of radius a, it is convenient to write

$$g_d(k) \;=\; [S_d(a)]^k\,\widetilde{g}_d(k) \tag{3.60}$$

with

$$\widetilde{g}_d(k) \;=\; \frac{1}{[S_d(1)]^k}\int_{\substack{|y_i|\le 1 \\ |y_i-y_j|>1}} dy_1\dots dy_k \tag{3.61}$$

for k positive integer and $\widetilde{g}_d(0)=1$. We observe that $\widetilde{g}_d(k)\le 1$ for all values of k. From (3.59)–(3.61) we conclude that

$$\begin{aligned} S_\Lambda(\tau) \;&\le\; [S_d(a)]^{d_1}\,\widetilde{g}_d(d_1)\prod_{i=2}^n [S_d(a)]^{d_i-1}\,\widetilde{g}_d(d_i-1) \\ &=\; [S_d(a)]^{n-1}\,\widetilde{g}_d(d_1)\prod_{i=2}^n \widetilde{g}_d(d_i-1)\;. \end{aligned}$$

The last identity follows from the fact that for every tree of n vertices, $d_1+\dots+d_n = 2n-2$. The τ-dependence of this last bound is only through the degree of the vertices, hence it leads, upon insertion in (3.58), to the inequality

$$|C_n(\beta,\Lambda)| \;\le\; \frac{[S_d(a)]^{n-1}}{n!}\sum_{\substack{d_1,\dots,d_n:\, d_i\ge 1 \\ d_1+\dots+d_n=2n-2}} \widetilde{g}_d(d_1)\prod_{i=2}^n \widetilde{g}_d(d_i-1)\,\frac{(n-2)!}{\prod_{i=1}^n (d_i-1)!} \;\le$$

$$\leq \frac{[S_d(a)]^{n-1}}{n} \sum_{\substack{d_1,\dots,d_n:\, d_i\geq 1 \\ d_1+\dots+d_n=2n-2}} \frac{\widetilde{g}_d(d_1)}{d_1!} \prod_{i=2}^{n} \frac{\widetilde{g}_d(d_i-1)}{(d_i-1)!}$$

where we recall that the quantity $(n-2)!/[\prod_{i=1}^{n}(d_i-1)!]$ in the first line of inequality above is precisely the number of trees with n vertices and fixed degrees $d_1,\dots,d_n$, according to Cayley formula and in the second line we have used the bound $d_1 \leq n-1$.
At this point, following [33], we multiply and divide by μ^{n-1} where $\mu > 0$ is a parameter to be chosen in an optimal way. This leads us to the inequality

$$\begin{aligned} |C_n(\beta,\Lambda)| &\leq \frac{1}{n}\left[\frac{[S_d(a)]}{\mu}\right]^{n-1} \sum_{\substack{d_1,\dots,d_n:\, d_i\geq 1 \\ d_1+\dots+d_n=2n-2}} \frac{\widetilde{g}_d(d_1)\,\mu^{d_1}}{d_1!} \prod_{i=2}^{n} \frac{\widetilde{g}_d(d_i-1)\,\mu^{d_i-1}}{(d_i-1)!} \\ &\leq \frac{1}{n}\left[\frac{[S_d(a)]}{\mu}\right]^{n-1} \left(\sum_{s\geq 0} \frac{\widetilde{g}_d(s)\mu^s}{s!}\right)^n \end{aligned}$$

Therefore the convergence radius of the Mayer series of the of the gas of free hard spheres of diameter a admit the following lower bound

$$R_{V^a_{\text{h.c.}}} \geq \frac{1}{S_d(a)} \max_{\mu>0} \frac{\mu}{C_d(\mu)}$$

where

$$C_d(\mu) \;=\; \sum_{s\geq 0} \frac{\widetilde{g}_d(s)}{s!}\,\mu^s$$

(attention! This is a polynomial in μ). Let us show that for $d=2$ the quantitative improvement given by this condition with respect to the classical bound (3.53) can be substantial. In the $d=2$ case (i. e. the two-dimensional hard sphere gas)

$$C_2(\mu) \;=\; \sum_{s=0}^{5} \frac{\widetilde{g}_2(s)}{s!}\mu^s$$

where, by definition, $\widetilde{g}_2(0) = \widetilde{g}_2(1) = 1$. The factor $\widetilde{g}_2(2)$ can be explicitly evaluated in terms of straightforward integrals and we get

$$\widetilde{g}_2(2) = \frac{1}{\pi^2} \int_{|x|\leq 1} d^2x \int_{|x'|\leq 1} d^2x' \Theta(|x-x'|>1)$$

where $\Theta(|x-x'|>1) = 1$ if $|x-x'|>1$ and zero otherwise. Using polar coordinates

$$\widetilde{g}_2(2) = \frac{2\pi}{\pi^2} \int_0^1 A(\rho)\rho d\rho$$

where $A(\rho)$ is the area of the region $S_0 \backslash S_\rho$ with $S_0 = \{(x,y)\in\mathbb{R}^2 : x^2+y^2\leq 1\}$ and $S_\rho = \{(x,y)\in\mathbb{R}^2 : (x-\rho)^2+y^2\leq 1\}$. We get

$$A(\rho) = 2\left[\int_{-1}^{\rho/2} \sqrt{1-x^2}dx - \int_{-1+\rho}^{\rho/2} \sqrt{1-(x-\rho)^2}\,dx\right] = 4\int_0^{\rho/2} \sqrt{1-x^2}dx =$$

$$= 2\left[\arcsin(\rho/2) + \frac{\rho}{2}\sqrt{1-\frac{\rho^2}{4}}\right]$$

Hence

$$\widetilde{g}_2(2) = \frac{2}{\pi}\int_0^1 A(\rho)\rho d\rho = \frac{4}{\pi}\int_0^1 \left[\arcsin(\rho/2) + \frac{\rho}{2}\sqrt{1-\frac{\rho^2}{4}}\right]\rho d\rho =$$

$$= \frac{16}{\pi}\int_0^{1/2}\left[u\arcsin(u) + u^2\sqrt{1-u^2}\right]du$$

But, since

$$\int u\arcsin(u)du = \frac{1}{4}\left[2u^2\arcsin(u) - \arcsin(u) + u\sqrt{1-u^2}\right]$$

and

$$\int u^2\sqrt{1-u^2} = \frac{1}{8}\left[\arcsin(u) + u\sqrt{1-u^2}(1-2u^2)\right]$$

We obtain

$$\widetilde{g}_2(2) = \frac{3\sqrt{3}}{4\pi}$$

The other terms of the sum can be numerically evaluated (e.g. using Mathematica), obtaining

$$\widetilde{g}_2(3) = 0,0589 \quad \widetilde{g}_2(4) = 0,0013 \quad \widetilde{g}_2(5) \leq 0,0001$$

Choosing $\mu = \left[\frac{8\pi}{3\sqrt{3}}\right]^{1/2}$ (a value for which $\frac{\mu}{C_d(\mu)}$ is close to its maximum) we get

$$R_{V^a_{\text{h.c.}}} \geq \frac{1}{S_d(a)}\,0.5107$$

This should be compared with the bound $R_{V^a_{\text{h.c.}}} \geq \frac{1/e}{S_d(a)}$ obtained through the classical condition (3.53).

3.5 Stable and tempered potentials

We explain in this section how to get an improvement on the lower bound (3.23) for the convergence radius of the Mayer series of a system of particles interacting via stable and tempered potential using a modified version of the Penrose identity. Such a new bound was given in [36].
Given a pair interaction V in $\mathbb{R}^d$ and given $(x_1,\ldots,x_n) \in \mathbb{R}^{dn}$, for every connected graph $g \in G_n$ there is at least a tree $\tau \subset g$, among the trees $\tau' \subset g$, which minimizes the value of $\sum_{\{i,j\}\in E_{\tau'}} V(x_i - x_j)$. We call such a tree a minimum spanning tree. If V and $(x_1,\ldots,x_n) \in \mathbb{R}^{dn}$ are such that for any $g \in G_n$ this minimum spanning tree is unique then we have a map $\boldsymbol{T} : G_n \to T_n$ and it is possible to show that this map is a partition scheme in G_n. Of course the problem is that in general a pair potential V does not have such property for any fixed $(x_1,\ldots,x_n) \in \mathbb{R}^{dn}$ and for some particular $(x'_1,\ldots,x'_n) \in \mathbb{R}^{dn}$

multiple minimal spanning trees are possible for a given graph g. In order to avoid such (possible) multiple minima, our strategy will be to add coordinates associated to the labels of the edges.
We recall that a totally ordered abelian monoid is a structure $(\mathbb{K}, +, 0, \geq)$ such that $(\mathbb{K}, +, 0)$ is an abelian (i.e. commutative) monoid, $(\mathbb{K}, \geq)$ is a totally ordered set (a chain) and for all $x, y, z \in \mathbb{K}$ we have that $x \geq y$ implies $x + z \geq y + z$. We also recall that E_n denotes the set of all unordered pairs $\{i, j\}$ in $[n]$.

Definition 3.9 *Let $f : E_n \to \mathbb{K}$ where $\mathbb{K}$ is a totally ordered abelian monoid. We say that f is admissible if $\sum_{e \in E_\tau} f(e)$ is different for different trees $\tau \in T_n$. Then, for every $g \in G_n$ there is a unique spanning tree $\tau \subset g$ for which $\sum_{e \in E_\tau} f(e)$ is minimum. We define the map $\boldsymbol{T}_f : G_n \to T_n$ such that $\boldsymbol{T}_f(g)$ is this unique minimum spanning tree of g.*

Definition 3.10 *Let $\mathbb{K}$ be a totally ordered abelian monoid. Let $f : E_n \to \mathbb{K}$ be an admissible function and let $\boldsymbol{T}_f : G_n \to T_n$ the minimum spanning tree map associated to f. We define the map $\boldsymbol{M}_f : T_n \to G_n$ such that $\boldsymbol{M}_f(\tau)$ is the graph with vertex-set $[n]$ and whose edges are the $\{i, j\}$ such that $f(\{i, j\}) \geq f(e)$ for every edge $e \in E_\tau$ belonging to the path from i to j through τ.*

Thus we have constructed

$$G_n \underset{\boldsymbol{M}_f}{\overset{\boldsymbol{T}_f}{\rightleftarrows}} T_n$$

Observe that $\tau \subset \boldsymbol{M}_f(\tau)$ and $\boldsymbol{T}_f(g) \subset g$. The following lemma shows that the maps $\boldsymbol{M}_f$ and $\boldsymbol{T}_f$ satisfy the hypothesis of the Proposition 3.2.

Lemma 3.2 *Let $f : E_n \to \mathbb{K}$ be an admissible function and let $\tau \in T_n$. Then*

$$\boldsymbol{T}_f^{-1}(\tau) = \{g \in G_n : \tau \subset g \subset \boldsymbol{M}_f(\tau)\}$$

Proof. Let $g \in \boldsymbol{T}_f^{-1}(\tau)$. We have $\tau = \boldsymbol{T}_f(g) \subset g$. Now take $\{i, j\} \in E_g$, and let $d \in E_\tau$ be any edge belonging to the path from i to j in τ. Consider τ' the graph obtained from τ after replacing d by $\{i, j\}$. Clearly τ' is connected and has $n-1$ edges, so it is a tree. By minimality of τ we must have $f(d) \leq f(\{i, j\})$, whence $\{i, j\} \in E_{\boldsymbol{M}_f(\tau)}$. Therefore $g \subset \boldsymbol{M}_f(\tau)$.
Conversely, let $\tau \subset g \subset \boldsymbol{M}_f(\tau)$. We must show $\boldsymbol{T}_f(g) = \tau$. By cardinality, it suffices to show $\boldsymbol{T}_f(g) \subset \tau$. Proceeding by contradiction, take $\{i, j\} \in E_{\boldsymbol{T}_f(g)} \setminus E_\tau$. Consider the path in τ joining i with j. Since $\boldsymbol{T}_f(g) \subset \boldsymbol{M}_f(\tau)$, $f(\{i, j\})$ is greater than the corresponding value for any edge in the path. If we remove $\{i, j\}$ from $\boldsymbol{T}_f(g)$, the tree splits into two trees. Necessarily, at least one of the edges in the path joins a vertex of one tree with a vertex of the other. Thus, by adding this edge we obtain a connected graph with $n - 1$ edges, a new tree, which contradicts the minimality of $\boldsymbol{T}_f(g)$. $\square$.

Remark 3.1 *Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ denote the set of whole numbers and let $m \in \mathbb{N}$ a positive integer. Then the set $\mathbb{R} \times \mathbb{N}_0^m$ has a canonical structure of abelian monoid. This set is also endowed with a natural total order. Indeed, $\mathbb{N}_0^m$ is totally ordered lexicographically, according to the usual order of $\mathbb{N}_0$ and then we consider the lexicographical order on $\mathbb{R} \times \mathbb{N}_0^m$, prioritizing the first coordinate.*

We will regard E_n as a totally ordered set according to some previously chosen order. E.g., we can choose the following order in E_n. Given a pair $e = \{i, j\}$ and $e' = \{i', j'\}$ we say that $e < e'$ if $i < i'$ or $i = i'$ and $j < j'$ so that $\{1, 2\}$ is the first edge, $\{1, 3\}$ is the second edge and so on until $\{n - 1, n\}$ which is the last edge. We also denote $\mathbb{N}_0^{\mathrm{E}_n}$ the set $\mathbb{N}_0^{|\mathrm{E}_n|}$ such that the m^{th} entry (with $1 \leq m \leq |\mathrm{E}_n|$) of an element $x \in \mathbb{N}_0^{|\mathrm{E}_n|}$ corresponds to the m^{th} edge in the order chosen in E_n. Of course $\mathbb{R} \times \mathbb{N}_0^{\mathrm{E}_n}$ is a totally ordered abelian monoid.

Definition 3.11 *Given a pair potential $V : \mathbb{R}^d \to \mathbb{R}$ and given $(x_1, \ldots, x_n) \in \mathbb{R}^{dn}$, we define the function $f_V : \mathrm{E}_n \to \mathbb{R} \times \mathbb{N}_0^{\mathrm{E}_n} : \{i, j\} \mapsto V(x_i - x_j) \times \mathbf{1}_{\{i,j\}}$ where $\mathbf{1}_{\{i,j\}}$ denotes the element of $\mathbb{N}_0^{|\mathrm{E}_n|}$ with all entries zero except the one at the position corresponding to the edge $\{i, j\}$ which is equal to one. Observe that f_V is, for every pair potential V, admissible according to Definition 3.9.*

In the following the symbol E_τ^+ denotes the set of edges of the tree τ with positive energy. That is,

$$E_\tau^+ = \{\{i, j\} \in E_\tau : \ V(x_i - x_j) \geq 0\}$$

We sometimes refer to E_τ^+ and $E_{\boldsymbol{M}_f(\tau)}$ simply by τ^+ and $\boldsymbol{M}_f(\tau)$ respectively to ease the notation.

Lemma 3.3 *Let $V : \mathbb{R}^d \to \mathbb{R}$ be a stable pair potential with stability constant B. Then, for every $\tau \in T_n$ and for any $(x_1, \ldots, x_n) \in \mathbb{R}^{dn}$*

$$\sum_{\{i,j\} \in \boldsymbol{M}_{f_V}(\tau) \setminus \tau^+} V(x_i - x_j) \geq -Bn \tag{3.62}$$

Proof. In the following, V_{ij} denotes $V(x_i - x_j)$, and we will make implicit use of the following trivial fact:

$$(x, \sigma) \geq (x', \sigma') \Rightarrow x \geq x' \quad \text{for } (x, \sigma), (x', \sigma') \in \mathbb{R} \times \mathbb{N}_0^{E_n}$$

Now we proceed to show that the inequality (3.62) holds true. The set of edges $E_\tau \setminus E_{\tau^+}$ is the forest $\{\tau_1, ..., \tau_k\}$. Let us denote $\mathcal{V}_{\tau_s}$ the vertex set of the tree τ_s of the forest. Assume $i \in \mathcal{V}_{\tau_a}$, $j \in \mathcal{V}_{\tau_b}$. If $a \neq b$, the path from i to j through τ involves an edge e in τ^+. Thus, if in addition $\{i, j\} \in \boldsymbol{M}_{f_V}(\tau)$, we have $V_{ij} \geq V_e \geq 0$. If $a = b$, the path from i to j through τ is contained in τ_a. Thus, if in addition $\{i, j\} \notin \boldsymbol{M}_{f_V}(\tau)$, we must have $V_{ij} \leq V_e \leq 0$ for some edge e in that path. This allows to bound:

$$\sum_{\{i,j\} \in \boldsymbol{M}_{f_V}(\tau) \setminus \tau^+} V_{ij} \geq \sum_{s=1}^{k} \sum_{\{i,j\} \subset \mathcal{V}_{\tau_s}} V_{ij} \geq \sum_{s=1}^{k} -|\mathcal{V}_{\tau_s}| B \geq -nB$$

□

Theorem 3.7 *Let V be a stable pair potential with stability constant B. Then for any $n \in \mathbb{N}$ such that $n \geq 2$ and any $(x_1, \dots, x_n) \in \mathbb{R}^{dn}$ the following inequality holds.*

$$\left| \sum_{g \in G_n} \prod_{\{i,j\} \in E_g} (e^{-\beta V(x_i - x_j)} - 1) \right| \leq e^{\beta B n} \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} (1 - e^{-\beta |V(x_i - x_j|)}) \quad (3.63)$$

Proof. Again using the short notation V_{ij} for $\beta V(x_i - x_j)$, by Lemma 3.2 we have that

$$\sum_{g \in G_n} \prod_{\{i,j\} \in E_g} (e^{-V_{ij}} - 1) = \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} (e^{-V_{ij}} - 1) \prod_{\{i,j\} \in E_{M_{f_V}(\tau)} \setminus E_\tau} e^{-V_{ij}}$$

where f_V is the admissible function given in Definition 3.11 and $\boldsymbol{M}_{f_V}$ is the map given in Definition 3.10. Using now the trick proposed in [32] we observe that, for any $\tau \in T_n$

$$\prod_{\{i,j\} \in E_\tau} |e^{-V_{ij}} - 1| = \Big[\prod_{\{i,j\} \in E_\tau} (1 - e^{-|V_{ij}|}) \Big] \, e^{-\sum_{\{i,j\} \in E_\tau \setminus E_\tau^+} V_{ij}}$$

so that

$$\begin{aligned} \Big| \sum_{g \in G_n} \prod_{\{i,j\} \in E_g} (e^{-V_{ij}} - 1) \Big| &\leq \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} |e^{-V_{ij}} - 1| \prod_{\{i,j\} \in E_{\boldsymbol{M}_{f_V}(\tau)} \setminus E_\tau} e^{-V_{ij}} = \\ &= \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} (1 - e^{-|V_{ij}|}) \prod_{\{i,j\} \in E_{\boldsymbol{M}_{f_V}(\tau)} \setminus E_\tau^+} e^{-V_{ij}} \leq \\ &\leq e^{\beta B n} \sum_{\tau \in T_n} \prod_{\{i,j\} \in E_\tau} (1 - e^{-|V_{ij}|}) \end{aligned}$$

where in the last line we have used Lemma 3.3. □

From Theorem 3.7 we have immediately the following Theorem.

Theorem 3.8 *Let V be a stable and tempered pair potential with stability constant B. Then the n-order Mayer coefficient $C_n(\beta, \Lambda)$ defined in (3.15) is bounded by*

$$|C_n(\beta, \Lambda)| \leq e^{\beta B n} n^{n-2} \frac{[\tilde{C}(\beta)]^{n-1}}{n!} \quad (3.64)$$

where

$$\tilde{C}(\beta) = \int_{\mathbb{R}^d} dx \; |e^{-\beta |V(x)|} - 1| \quad (3.65)$$

Therefore the Mayer series (3.9) converges absolutely, uniformly in Λ, for any complex λ inside the disk

$$|\lambda| < \frac{1}{e^{\beta B + 1} \tilde{C}(\beta)} \quad (3.66)$$

I.e. the convergence radius of the Mayer series (3.9) admits the following lower bound

$$R_V \geq \frac{1}{e^{\beta B + 1} \tilde{C}(\beta)} \quad (3.67)$$

Proof. From Theorem 3.7 we have that

$$|C_n(\beta,\Lambda)| \leq \frac{1}{n!}\frac{1}{|\Lambda|}\int_\Lambda dx_1 \dots \int_\Lambda dx_n |\Phi^T(x_1,\dots,x_n)| \;\leq$$

$$\leq\; e^{\beta B n}\frac{1}{n!}\left[\int_{\mathbb{R}^d}\left[1-e^{-\beta|V(x)|}\right]dx\right]^{n-1}\sum_{\tau\in T_n}1 =$$

$$=\; e^{\beta B n}\frac{n^{n-2}}{n!}\left[\tilde{C}(\beta)\right]^{n-1}$$

This concludes the proof of Theorem 3.8. □.

The improvement on the lower bound of the convergence radius of the Mayer series for stable and tempered potentials given by Theorem 3.7 with respect to Theorem 3.2 is twofold. First, the factor $e^{\beta B+1}$ in (3.67) replaces the factor $e^{2\beta B+1}$ in (3.23). Second, the factor $\tilde{C}(\beta)$ in (3.67) replaces the factor $C(\beta)$ in (3.23) and clearly, recalling their definitions (3.65) and (3.21) respectively, we have $\tilde{C}(\beta) \leq C(\beta)$ where the equality only holds if V is non-negative (purely repulsive). Moreover observe that while $\tilde{C}(\beta)$ grows at most linearly in β, the factor $C(\beta)$ grows exponentially with β. So, if we denote by $R_{PR} = [e^{2\beta B+1}C(\beta)]^{-1}$ the Penrose-Ruelle lower bound for the convergence radius given in Theorem 3.2 and by $R^* = [e^{\beta B+1}\tilde{C}(\beta)]^{-1}$ the lower bound given by Theorem 3.8 for the same convergence radius we get that $R^*/R_{PR} = e^{\beta B}[C(\beta)/\tilde{C}(\beta)]$. This ratio, always greater than one, is the product of two factors, $e^{\beta B}$ and $[\tilde{C}(\beta)/C(\beta)]$, both growing exponentially fast with β. To give an idea, for a gas of particles interacting via the classical Lennard-Jones potential

$$V(|x|) = \frac{1}{|x|^{12}} - \frac{2}{|x|^6}$$

at inverse temperature $\beta = 1$, using the value $B_{\mathrm{LJ}} = 8.61$ for its stability constant (see [20]), the lower bound (3.67) is at least 8.5×10^4 larger than the Penrose-Ruelle lower bound, while for $\beta = 10$ is at least 7.26×10^{43} larger than the Penrose-Ruelle lower bound.

Part II

Discrete systems

Chapter 4

The Abstract Polymer gas

4.1 Setting

The abstract polymer gas is a discrete model which plays a very important role in many physical situations, such as spin systems on the lattice at low or high temperature, or continuous or discrete field theories. The polymer gas was first introduced by Gruber and Kunz [18] in 1970. In the original Gruber and Kunz paper polymer were finite subsets of $\mathbb{Z}^d$ (the unit cubic lattice). Later, Kotecyý and Preiss proposed the abstract model in which polymers were abstract objects belonging to some set $\mathcal{P}$, the polymer set, whose unique structure was given by mean of a symmetric and reflexive relation in $\mathcal{P}$, that they called the incompatibility relation. As we will see, this is equivalent to assume that the interaction between polymers occurs via a hard core pair potential. In the present chapter we revisit the abstract polymer gas.

We first specify a countable set $\mathcal{P}$ whose elements are all possible polymers (i.e. $\mathcal{P}$ is the *single particle state space*). We then associate to each polymer $\gamma \in \mathcal{P}$ a complex number z_γ (a positive number in physical situations) which is interpreted as the *activity* of the polymer γ. We will denote $\boldsymbol{z} = \{z_\gamma\}_{\gamma\in\mathcal{P}}$ and for any $\Lambda \subset \mathcal{P}$, $\boldsymbol{z}_\Lambda = \{z_\gamma\}_{\gamma\in\Lambda}$.

In the general situation polymers interact through a pair potential. Namely, the energy E of a configuration $\gamma_1, \ldots, \gamma_n$ of n polymers is given by

$$E(\gamma_1, \ldots, \gamma_n) = \sum_{1\le i<j\le n} V(\gamma_i, \gamma_j) \tag{4.1}$$

where pair potential $V(\gamma, \gamma')$ is a symmetric function in $\mathcal{P} \times \mathcal{P}$ taking values in $\mathbb{R} \cup \{+\infty\}$. We will make the assumption that the pair interaction $V(\gamma, \gamma')$ is purely hard-core. Namely $V(\gamma, \gamma')$ takes values in the set $\{0, +\infty\}$. Observe that an hard-core pair potential $V(\gamma, \gamma')$ in $\mathcal{P}$ induces a relation $\mathcal{R}_V$ in $\mathcal{P} \times \mathcal{P}$ (i.e. $\mathcal{R}_V$ is a subset of $\mathcal{P} \times \mathcal{P}$). Namely, we say that a pair (γ, γ') belongs to $\mathcal{R}_V$ if and only if $V(\gamma, \gamma') = +\infty$. Clearly $\mathcal{R}_V$ is symmetric because by assumption V is symmetric. When $(\gamma, \gamma') \in \mathcal{R}_V$ (i.e. $V(\gamma, \gamma') = +\infty$), we write $\gamma \not\sim \gamma'$ and say that γ and γ' are *incompatible*. Conversely, if $(\gamma, \gamma') \notin \mathcal{R}_V$ we say that the polymers γ and γ' are *compatible* and we write $\gamma \sim \gamma'$. Note that if V is such that $\mathcal{R}_V$ is reflexive, then $\gamma \not\sim \gamma$ for all $\gamma \in \mathcal{P}$. The relation $\mathcal{R}_V$ induced by the pair potential V is called *the incompatibility relation*.

Fix now a finite set $\Lambda \subset \mathcal{P}$ (the "volume" of the gas). Then, for $z_\gamma \geq 0$, the probability to see the configuration $(\gamma_1, \ldots, \gamma_n) \in \Lambda^n$ is given by

$$Prob(\gamma_1, \ldots, \gamma_n) = \frac{1}{\Xi_\Lambda} \frac{1}{n!} z_{\gamma_1} z_{\gamma_2} \ldots z_{\gamma_n} e^{-\sum_{1 \leq i < j \leq n} V(\gamma_i, \gamma_j)}$$

where the normalization constant Ξ_Λ is the grand-canonical partition function in the volume Λ and is given by

$$\Xi_\Lambda(\boldsymbol{z}_\Lambda) = 1 + \sum_{n \geq 1} \frac{1}{n!} \sum_{(\gamma_1, \ldots, \gamma_n) \in \Lambda^n} z_{\gamma_1} z_{\gamma_2} \ldots z_{\gamma_n} e^{-\sum_{1 \leq i < j \leq n} V(\gamma_i, \gamma_j)} \tag{4.2}$$

Note that configurations $\gamma_1, \ldots, \gamma_n$ for which there exists some incompatible pair γ_i, γ_j, i.e. such that $V(\gamma_i, \gamma_j) = +\infty$ have zero probability to occur, i.e. are forbidden.

Remark. The fact that $V(\gamma, \gamma') \geq 0$ immediately implies that Ξ_Λ is convergent and

$$|\Xi_\Lambda(\boldsymbol{z}_\Lambda)| \leq 1 + \sum_{n \geq 1} \frac{1}{n!} \left[\sum_{\gamma \subset \Lambda} |z_\gamma| \right]^n \leq \exp\left\{ \sum_{\gamma \in \Lambda} |z_\gamma| \right\} \leq \max_{\gamma \in \Lambda} e^{|z_\gamma||\Lambda|}$$

Actually, $V(\gamma\, \gamma') \geq 0$ implies that $\Xi_\Lambda(z)$ is analytic in the whole $\mathbb{C}^{|\Lambda|}$ ($|\Lambda|$ is the cardinality of Λ).

The "pressure" of this gas is defined via the formula

$$P_\Lambda(\boldsymbol{z}_\Lambda) = \frac{1}{|\Lambda|} \log \Xi_\Lambda(\boldsymbol{z}_\Lambda) \tag{4.3}$$

While the partition function $\Xi_\Lambda(\boldsymbol{z}_\Lambda)$ diverges as $\Lambda \to \mathcal{P}$, the pressure $P_\Lambda(\boldsymbol{z}_\Lambda)$ is supposed to have a finite limit at least when z varies in some finite polydisc $|z_\gamma| \leq \rho_\gamma$ with $\boldsymbol{\rho} = \{\rho_\gamma\}_{\gamma \in \mathcal{P}}$ being some positive function $\boldsymbol{\rho} : \mathcal{P} \to \mathbb{R}^+ : \gamma \mapsto \rho_\gamma$ defined on $\mathcal{P}$ independent on Λ. So that in principle it should be possible to give an upper bound for $|P_\Lambda(\boldsymbol{z})|$ which is uniform in Λ.

The pressure (4.3) can be written as a formal series using the Mayer trick in the partition function (4.2) by writing the Gibbs factor as

$$e^{-\sum_{1 \leq i < j \leq n} V(\gamma_i, \gamma_j)} = \prod_{1 \leq i < j \leq n} \left[(e^{-V(\gamma_i, \gamma_j)} - 1) + 1 \right]$$

Proceeding exactly as in section 3.1.2 with the only difference if that now $V(\gamma_i, \gamma_j)$ is in place of $V(x_i - x_j)$ and $\sum_{\gamma_i \in \Lambda}$ is in place of $\int_\Lambda dx_i$ we obtain

$$\log \Xi_\Lambda(\boldsymbol{z}_\Lambda) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{(\gamma_1, \ldots, \gamma_n) \in \Lambda^n} \phi^T(\gamma_1, \ldots, \gamma_n)\, z_{\gamma_1} \ldots z_{\gamma_n} \tag{4.4}$$

with

$$\phi^T(\gamma_1, \ldots, \gamma_n) = \begin{cases} 1 & \text{if } n = 1 \\ \displaystyle\sum_{g \in G_n} \prod_{\{i,j\} \in E_g} (e^{-V(\gamma_i, \gamma_j)} - 1) & \text{if } n \geq 2 \end{cases} \tag{4.5}$$

where $\sum_{g\in G_n}$ is the sum over all connected graphs between $[n]$.
Observe that, since $V(\gamma,\gamma')\geq 0$, by Proposition 3.3 it holds

$$\phi^T(\gamma_1,\dots,\gamma_n) = (-1)^{n-1}|\phi^T(\gamma_1,\dots,\gamma_n)| \tag{4.6}$$

The equation (4.4) makes sense only for those $\boldsymbol{z}\in\mathbb{C}^{\mathcal{P}}$ such that the formal series in the r.h.s. of (4.4) converge absolutely. To study absolute convergence, we will consider, for any Λ, the positive term series

$$|\log\Xi|_\Lambda(\boldsymbol{\rho}_\Lambda) = \sum_{n=1}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\dots,\gamma_n)\in\Lambda^n}|\phi^T(\gamma_1,\dots,\gamma_n)|\rho_{\gamma_1}\cdots\rho_{\gamma_n} \tag{4.7}$$

for $\boldsymbol{\rho}\in(0,\infty)^{\mathcal{P}}$. Note that

$$|\log\Xi_\Lambda(\boldsymbol{z}_\Lambda)| \leq |\log\Xi|_\Lambda(\boldsymbol{\rho}_\Lambda)$$

for all $\boldsymbol{z}\in\mathbb{C}^{|\Lambda|}$ in the poly-disc $\{|z_\gamma|\leq\rho_\gamma\}_{\gamma\in\Lambda}$, so that if we are able to prove that the series (4.7) converges, for all Λ, at some value $\boldsymbol{\rho}=\{\rho_\gamma\}_{\gamma\in\mathcal{P}}\in(0,\infty)^{\mathcal{P}}$, then we have also proved that the series (4.4) converges absolutely, for all Λ, whenever $\boldsymbol{z}$ is in the poly-disk $\{|z_\gamma|\leq\rho_\gamma\}_{\gamma\in\mathcal{P}}$. We also observe that, since $V(\gamma,\gamma')\geq 0$, by (4.6)

$$|\log\Xi|_\Lambda(\boldsymbol{\rho}_\Lambda) = -\log\Xi_\Lambda(-\boldsymbol{\rho}_\Lambda) \tag{4.8}$$

To study the convergence of the pressure (4.3) it is convenient to consider the quantity

$$|\Pi|_{\gamma_0}(\boldsymbol{\rho}) = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in\mathcal{P}^n}|\phi^T(\gamma_0,\gamma_1,\dots,\gamma_n)|\ \rho_{\gamma_1}\cdots\rho_{\gamma_n} \tag{4.9}$$

Note that in (4.9) the sum over each polymer γ_i is not anymore restricted to the finite "volume" Λ and it runs over the full polymer space $\mathcal{P}$.
If we are able to show that $|\Pi|_{\gamma_0}(\boldsymbol{\rho})$ converges for some (bounded) positive function $\boldsymbol{\rho}\in[0,\infty)^{\mathcal{P}}$, then also the pressure $P_\Lambda(\boldsymbol{z}_\Lambda)$ converges absolutely, for all Λ, whenever z is in the poly-disk $\{|z_\gamma|\leq\rho_\gamma\}_{\gamma\in\mathcal{P}}$ and in this poly-disk it uniformly in Λ. Indeed in the poly-disk $\{|z_\gamma|\leq\rho_\gamma\}_{\gamma\in\mathcal{P}}$ we have

$$|P_\Lambda(\boldsymbol{z}_\Lambda)| = \frac{1}{|\Lambda|}|\log\Xi_\Lambda(\boldsymbol{z}_\Lambda)| \leq \frac{1}{|\Lambda|}|\log\Xi|_\Lambda(\boldsymbol{\rho}_\Lambda)\ =$$

$$= \frac{1}{|\Lambda|}\sum_{n=1}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\dots,\gamma_n)\in\Lambda^n}|\phi^T(\gamma_1,\dots,\gamma_n)|\rho_{\gamma_1}\cdots\rho_{\gamma_n} =$$

$$= \frac{1}{|\Lambda|}\sum_{\gamma_0\in\Lambda}\rho_{\gamma_0}\sum_{n=1}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\dots,\gamma_{n-1})\in\Lambda^{n-1}}|\phi^T(\gamma_0,\gamma_1,\dots,\gamma_{n-1})|\rho_{\gamma_1}\cdots\rho_{\gamma_{n-1}}\ =$$

$$\leq\ \frac{1}{|\Lambda|}\sum_{\gamma_0\in\Lambda}\rho_{\gamma_0}\sum_{m=0}^{\infty}\frac{1}{(m+1)!}\sum_{(\gamma_1,\dots,\gamma_n)\in\Lambda^m}|\phi^T(\gamma_0,\gamma_1\dots,\gamma_m)|\rho_{\gamma_1}\cdots\rho_{\gamma_m}\ \leq$$

$$\frac{1}{|\Lambda|}\sum_{\gamma_0\in\Lambda}\rho_{\gamma_0}\sum_{m=0}^{\infty}\frac{1}{m!}\sum_{(\gamma_1,\dots,\gamma_n)\in\Lambda^m}|\phi^T(\gamma_0,\gamma_1\dots,\gamma_m)|\rho_{\gamma_1}\cdots\rho_{\gamma_m} \;=$$

$$=\;\frac{1}{|\Lambda|}\sum_{\gamma_0\in\Lambda}\rho_{\gamma_0}|\Pi|_{\gamma_0}(\boldsymbol{\rho})\;\le\sup_{\gamma_0\in\Lambda}\;\rho_{\gamma_0}\,|\Pi|_{\gamma_0}(\boldsymbol{\rho})\;\le\;\sup_{\gamma_0\in\mathcal{P}}\;\rho_{\gamma_0}\,|\Pi|_{\gamma_0}(\boldsymbol{\rho})$$

In short, for all $\boldsymbol{z}$ in the polydisk $\{|z_\gamma|\le\rho_\gamma\}_{\gamma\in\mathcal{P}}$, it holds

$$|P_\Lambda(\boldsymbol{z}_\Lambda)|\;\le\;\sup_{\gamma_0\in\mathcal{P}}\;\rho_{\gamma_0}\,|\Pi|_{\gamma_0}(\boldsymbol{\rho}) \tag{4.10}$$

So in the next sections we will focus our attention on the formal series $|\Pi|_{\gamma_0}(\boldsymbol{\rho})$ defined in equation (4.9). Indeed, when one is able to prove that the series (4.9) then, by (4.10) he has also proved the absolute convergence of the pressure of the polymer gas uniformly in the volume Λ.

4.2 Convergence of the abstract polymer gas

As explained in the previous section, to study absolute convergence of the pressure we will just need to prove that the positive term series defined in (4.9)

$$|\Pi|_{\gamma_0}(\boldsymbol{\rho})=\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in\mathcal{P}^n}|\phi^T(\gamma_0,\gamma_1,\dots,\gamma_n)|\;\rho_{\gamma_1}\cdots\rho_{\gamma_n}$$

is convergent for some $\boldsymbol{\rho}\in(0,+\infty)^{\mathcal{P}}$.
Since the interaction $V(\gamma,\gamma')$ is purely hard core, we can use the original Penrose identity to bound the factor Ursell factor $|\phi^T(\gamma_0,\gamma_1,\dots,\gamma_n)|$. We define the map $\boldsymbol{m}:T_n^0\to G_n^0$ as in Definition 3.6. By rooting trees in T_n^0 in 0, given $\tau\in T_n^0$, $d(i)$ denotes the depth of the vertex i (i.e. its edge distance from 0) and that i' denotes the parent of i. The map $\mathbf{m}:T_n^0\to G_n^0$ is therefore the map such that to each tree $\tau\in T_n^0$ associates the graph $\mathbf{m}(\tau)\in G_n^0$ formed by adding to τ all edges $\{i,j\}$ such that either:

(p1) $d_\tau(i)=d_\tau(j)$ (edges between vertices of the same generation), or

(p2) $d_\tau(j)=d_\tau(i)-1$ and $j>i'$ (edges between vertices with generations differing by one).

As shown previously the map $\mathbf{m}$ is a partition scheme and therefore by Theorem 3.4 we have that

$$\phi^T(\gamma_0,\gamma_1,\dots,\gamma_n)=\sum_{g\in G_n^0}\prod_{\{i,j\}\in E_g}(e^{-V(\gamma_i,\gamma_j)}-1)=\sum_{\tau\in T_n^0}w_\tau(\gamma_0\,\gamma_1,\dots,\gamma_n)$$

with

$$w_\tau(\gamma_0\,\gamma_1,\dots,\gamma_n)=e^{-\sum_{\{i,j\}\in E_{\mathbf{m}(\tau)}\setminus E_\tau}V(\gamma_i,\gamma_j)}\prod_{\{i,j\}\in\tau}\Big(e^{-V(\gamma_i,\gamma_j)}-1\Big)$$

As shown in Section 3.4 (see there Theorem 3.5 and formula (3.55)) we have

$$w_\tau(\gamma_0\,\gamma_1,\dots,\gamma_n) = (-1)^n \mathbb{1}_{P(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau)$$

where $P(\gamma_0,\gamma_1,\dots,\gamma_n)$ is the set of Penrose trees and $\mathbb{1}_{\tau\in P(\gamma_0,\gamma_1,\dots,\gamma_n)}$ is the characteristic function of the set $P(\gamma_0,\gamma_1,\dots,\gamma_n)$ in T_n^0, i.e.

$$\mathbb{1}_{P(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau) \;=\; \begin{cases} 1 & \text{if } \tau \in P(\gamma_0,\gamma_1,\dots,\gamma_n) \\ \\ 0 & \text{otherwise} \end{cases}$$

We remind its definition below.

Definition 4.1 *Given* $(\gamma_0,\gamma_1,\dots,\gamma_n)\in\mathcal{P}^{n+1}$, *the set of* Penrose trees $P(\gamma_0,\gamma_1,\dots,\gamma_n)$, *is formed by the trees* $\tau\in T_n^0$ *such that*

(t0) *if* $\{i,j\}\in E_\tau$ *then* $\gamma_i \nsim \gamma_j$

(t1) *if two vertices* i *and* j *siblings then* $\gamma_i\sim\gamma_j$;

(t2) *if two vertices* i *and* j *are not siblings but* $d_\tau(i)=d_\tau(j)$, *then* $\gamma_i\sim\gamma_j$;

(t3) *if two vertices* i *and* j *are s. t.* $d_\tau(j)=d_\tau(i)-1$ *and* $j>i'$, *then* $\gamma_i\sim\gamma_j$.

Therefore we get

$$\phi^T(\gamma_0,\gamma_1,\dots,\gamma_n) \;=\; (-1)^n \sum_{\tau\in T_n^0} \mathbb{1}_{P(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau) \tag{4.11}$$

Using (4.11) we can now rewrite the formal series (4.9) as

$$\begin{aligned} |\Pi|_{\gamma_0}(\boldsymbol{\rho}) &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in\mathcal{P}^n}\sum_{\tau\in T_n^0}\mathbb{1}_{P_G(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau)\;\rho_{\gamma_1}\dots\rho_{\gamma_n} \;= \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{\tau\in T_n^0}\sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in\mathcal{P}^n}\mathbb{1}_{P_G(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau)\;\rho_{\gamma_1}\dots\rho_{\gamma_n} \;= \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{\tau\in T_n^0}\phi_{\gamma_0}(\tau) \end{aligned} \tag{4.12}$$

where

$$\phi_{\gamma_0}(\tau) = \sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in\mathcal{P}^n}\mathbb{1}_{P_G(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau)\;\rho_{\gamma_1}\dots\rho_{\gamma_n} \tag{4.13}$$

The structure of (4.12) is crucial. This equation shows that the formal series $|\Pi|_{\gamma_0}(\rho)$ can be reaorganized as a sum over terms associated to labelled trees. We stress that here the factor $\phi_{\gamma_0}(\tau)$ depends on the *labelled* tree τ because of the Penrose condition (t3) which indeed depends on the labelling of the tree. We will see below that efficient bounds on the factor $\phi_{\gamma_0}(\tau)$ defined in (4.13) above can be obtained by choosing a family of trees $\tilde{P}(\gamma_0,\gamma_1,\dots,\gamma_n)$ such that $P(\gamma_0,\gamma_1,\dots,\gamma_n)\subset\tilde{P}(\gamma_0,\gamma_1,\dots,\gamma_n)$. Below we will consider three possible choice of $\tilde{P}(\gamma_0,\gamma_1,\dots,\gamma_n)$ which are the choices which yield the three known criteria for the convergence of cluster expansion of the abstract polymer gas.

Definition 4.2 *Given $(\gamma_0, \gamma_1, \dots, \gamma_n) \in \mathcal{P}^{n+1}$, the set of weakly the Penrose trees, which we denote by is defined as $P^*(\gamma_0, \gamma_1, \dots, \gamma_n)$, is formed by all trees τ with vertex set $\{0, 1, \dots, n\}$ and edge set E_τ such that*

(t0) *if $\{i, j\} \in E_\tau$ then $\gamma_i \not\sim \gamma_j$*

(t1)* *if i and j are siblings then $\gamma_i \sim \gamma_j$;*

Definition 4.3 *Given $(\gamma_0, \gamma_1, \dots, \gamma_n) \in \mathcal{P}^{n+1}$, the set of the Dobrushin trees, which we denote by is defined as $P^{\mathrm{Dob}}(\gamma_0, \gamma_1, \dots, \gamma_n)$, is the set the trees τ with vertex set $\{0, 1, \dots, n\}$ and edge set E_τ such that*

(t0) *if $\{i, j\} \in E_\tau$ then $\gamma_i \not\sim \gamma_j$*

(t1)$^{\mathrm{D}}$ *if i and j are siblings then $\gamma_i \neq \gamma_j$;*

Definition 4.4 *Given $(\gamma_0, \gamma_1, \dots, \gamma_n) \in \mathcal{P}^{n+1}$, the set of Kotecký-Preiss trees, which we denote by $P^{\mathrm{KP}}(\gamma_0, \gamma_1, \dots, \gamma_n)$ is formed by all trees τ with vertex set $\{0, 1, \dots, n\}$ and edge set E_τ such that*

(t0) *if $\{i, j\} \in E_\tau$ then $\gamma_i \not\sim \gamma_j$*

Note that we have, by definition, that

$$P(\gamma_0, \gamma_1, \dots, \gamma_n) \subset P^*(\gamma_0, \gamma_1, \dots, \gamma_n) \subset P^{\mathrm{Dob}}(\gamma_0, \gamma_1, \dots, \gamma_n) \subset P^{\mathrm{KP}}(\gamma_0, \gamma_1, \dots, \gamma_n) \tag{4.14}$$

So that, by (3.51), we have the bounds

$$|\phi^T(\gamma_0, \gamma_1, \dots, \gamma_n)| \leq |P^*(\gamma_0, \gamma_1, \dots, \gamma_n)| = \sum_{\tau \in T_n^0} \mathbb{1}_{P^*(\gamma_0, \gamma_1, \dots, \gamma_n)}(\tau) \leq \tag{4.15}$$

$$\leq |P^{\mathrm{Dob}}(\gamma_0, \gamma_1, \dots, \gamma_n)| = \sum_{\tau \in T_n^0} \mathbb{1}_{P^{\mathrm{Dob}}(\gamma_0, \gamma_1, \dots, \gamma_n)}(\tau) \leq \tag{4.16}$$

$$\leq |P^{\mathrm{KP}}(\gamma_0, \gamma_1, \dots, \gamma_n)| = \sum_{\tau \in T_n^0} \mathbb{1}_{P^{\mathrm{KP}}(\gamma_0, \gamma_1, \dots, \gamma_n)}(\tau) \tag{4.17}$$

The latter (4.17), i.e. the worst among the three bound proposed, is known as the Rota bound. This was the bound which was used by Cammarota [8], Brydges [6] and Simon [43] to obtain a direct proof of the absolute convergence of the pressure of a Polymer gas by directly estimating the Ursell coefficent.
Hence, we can bound the positive term series $|\Pi|_{\gamma_0}(\boldsymbol{\rho})$ defined in (4.9), using of course the estimate (4.15) which is the best among the three (4.15)-(4.17) proposed, as

$$|\Pi|_{\gamma_0}(\boldsymbol{\rho}) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(\gamma_1, \gamma_2, \dots, \gamma_n) \in \mathcal{P}^n} |\phi^T(\gamma_0, \gamma_1, \dots, \gamma_n)| \rho_{\gamma_1} \dots \rho_{\gamma_n} \leq$$

$$\leq \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(\gamma_1, \gamma_2, \dots, \gamma_n) \in \mathcal{P}^n} \sum_{\tau \in T_n^0} \mathbb{1}_{P^*(\gamma_0, \gamma_1, \dots, \gamma_n)}(\tau) \rho_{\gamma_1} \dots \rho_{\gamma_n} =$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\tau \in T_n^0} \sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in \mathcal{P}^n} \mathbb{1}_{P^*(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau) \rho_{\gamma_1} \dots \rho_{\gamma_n}$$

So we get

$$|\Pi|_{\gamma_0}(\boldsymbol{\rho}) \;\le\; \Pi^*_{\gamma_0}(\boldsymbol{\rho}) \tag{4.18}$$

where

$$\Pi^*_{\gamma_0}(\boldsymbol{\rho}) \;=\; \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\tau \in T_n^0} \phi^*_{\gamma_0}(\tau, \boldsymbol{\rho}) \tag{4.19}$$

with

$$\phi^*_{\gamma_0}(\tau, \boldsymbol{\rho}) = \sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in \mathcal{P}^n} \mathbb{1}_{P^*(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau)\; \rho_{\gamma_1} \dots \rho_{\gamma_n} \tag{4.20}$$

Analogously, using the bounds (4.16) and (4.17) we can define two more series which also majorize the series $|\Pi|_{\gamma_0}(\rho)$. Namely

$$\Pi^{\mathrm{Dob}}_{\gamma_0}(\boldsymbol{\rho}) \;=\; \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\tau \in T_n^0} \phi^{\mathrm{Dob}}_{\gamma_0}(\tau, \rho) \tag{4.21}$$

and

$$\Pi^{\mathrm{KP}}_{\gamma_0}(\boldsymbol{\rho}) \;=\; \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\tau \in T_n^0} \phi^{\mathrm{KP}}_{\gamma_0}(\tau, \rho) \tag{4.22}$$

with

$$\phi^{\mathrm{Dob}}_{\gamma_0}(\tau, \boldsymbol{\rho}) = \sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in \mathcal{P}^n} \mathbb{1}_{P^{\mathrm{Dob}}(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau)\; \rho_{\gamma_1} \dots \rho_{\gamma_n} \tag{4.23}$$

and

$$\phi^{\mathrm{KP}}_{\gamma_0}(\tau, \boldsymbol{\rho}) = \sum_{(\gamma_1,\gamma_2,\dots,\gamma_n)\in \mathcal{P}^n} \mathbb{1}_{P^{\mathrm{KP}}(\gamma_0,\gamma_1,\dots,\gamma_n)}(\tau)\; \rho_{\gamma_1} \dots \rho_{\gamma_n} \tag{4.24}$$

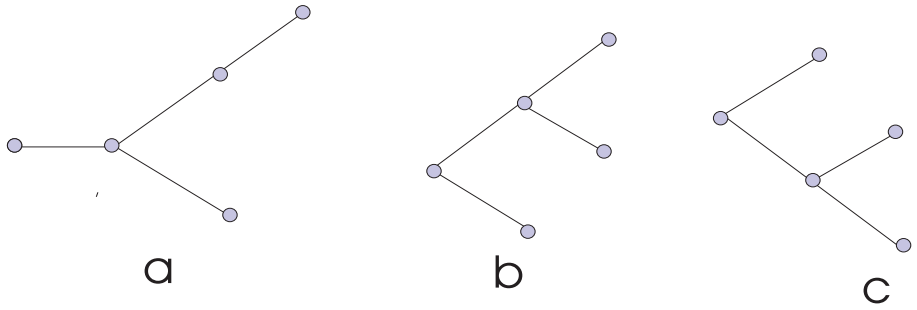
Here it is important to stress that, differently form the factor $\phi_{\gamma_0}(\tau)$ defined in (4.13), the three factors $\phi^{*,\,\mathrm{Dob},\,\mathrm{KP}}_{\gamma_0}(\tau)$ do not depend on the labels of the tree τ, but only on its topological structure. This means that the terms in the series $\Pi^*_{\gamma_0}(\rho)$ can further be grouped together in terms of non unlabelled rooted trees. We will make this concept precise in the next section. We conclude this section by remarking that inequalities (4.15)-(4.17) immediately imply

$$|\Pi|_{\gamma_0}(\boldsymbol{\rho}) \;\le\; \Pi^*_{\gamma_0}(\boldsymbol{\rho}) \;\le\; \Pi^{\mathrm{Dob}}_{\gamma_0}(\boldsymbol{\rho}) \;\le\; \Pi^{\mathrm{KP}}_{\gamma_0}(\boldsymbol{\rho}) \tag{4.25}$$

4.2.1 Reorganization of the series $\Pi^*_{\gamma_0}(\boldsymbol{\rho})$

We now reorganize the sum over rooted labelled trees appearing in formula (4.19) in terms of the plane rooted trees. Such reorganization is motivated by the observation that the factor (4.20) does not depend on the labels assigned to the vertices of τ but only its topological structure. As a matter of fact, to each labelled ordered rooted tree $\tau \in T_n^0$ we can associate a drawning in the

plane known as the "plane rooted tree" associated to τ. The drawing of τ is obtained by putting parents at the left of their children which are ordered in the top-to-bottom order consistently with the order of their labels. For example the plane rooted trees with $n+1=5$ vertices associated to the trees **a** with edge set $\{0,3\},\{1,3\},\{2,3\},\{1,4\}$, **b** with edge set $\{0,2\},\{0,3\},\{1,2\},\{2,4\}$ and **c** with edge set $\{0,2\},\{0,4\},\{4,3\},\{1,4\}$ are drawn below



Figure

Observe that b and c are *different* plane rooted tree (because of the rule of the ordering of the children from top-to-bottom).

In this way we have defined a map $m : \tau \mapsto m(\tau)$ which associate to each labelled tree $\tau \in T_n^0$ a unique drawing $t = m(\tau)$ in the plane, called the *planar rooted tree* associated to τ. We denote by $\mathcal{T}_n^0 =$ the set of all planar rooted trees with $n+1$ vertices and by $\mathcal{T}^{0,k}$ the set of planar rooted trees with maximal generation number k; let also $\mathcal{T}^0 = \cup_{n\geq 0}\mathcal{T}_n^0 = \cup_{k\geq 0}\mathcal{T}^{0,k}$ be the set of all planar rooted trees. An element $t \in \mathcal{T}_n^0$ can also be viewed as an equivalence class of elements $\tau \in T_{n+1}$ with the equivalence relation being that two elements τ and τ' are equivalent if the produce the same planar rooted tree. So when we write $\tau \in t$ with $t \in \mathcal{T}_n^0$ we mean that τ is an element of the set of all labelled trees in T_n^0 that produce the same plane ordered rooted tree.
We will use the following notations. Given a vertex $v \neq 0$ in a rooted tree (with root 0), we denote by v' is parent, we denote by s_v the number of its children and we denoted by $v^1, \dots, v^{s_v}$ the children of v. If $s_v = 0$ we say that v is an end-point or a *leaf* of τ.
Note that the set of vertices of a labeled rooted tree $\tau \in T_n^0$ (a plane rooted tree $t \in \mathcal{T}_n^0$) can be endowed with a total order $\prec$ in a natural way in such way that for any v the father v' of v is such that $v' \prec v$ and children of any vertex v of a labeled tree $\tau \in T_n^0$ (a plane rooted tree $t \in \mathcal{T}_n^0$) are ordered following the order of their labels (from high to low); namely, the ordering of $v^1, \dots, v^{s_v}$ is such that $v^1 \prec v^2 \prec \cdots \prec v^{s_v}$.

Clearly the map $\tau \mapsto m(\tau) = t$ is many-to-one and the cardinality of the pre-image of a planar rooted tree t (=number of ways of labelling the n non-root vertices of the tree with n distinct labels consistently with the rule "from high to low") is given by

$$\left|\{\tau \in T_n^0 : m(\tau) = t\}\right| \;=\; \frac{n!}{\prod_{v\succeq 0} s_v!} \tag{4.26}$$

As a matter of fact, it is very easy to count how many labeled trees $\tau \in T_n^0$ belong to the same equivalent class t, i.e. are associated to the same plane rooted tree. One have just to count all permutations σ of $\{0, 1, 2, \dots, n\}$ which leaves the root unchanged and which respect the order of the children in any vertex. Let $\tau \in T_n^0$ and let $t = [\tau]$ the plane root tree associated to τ characterized by the sequence $\{s_v\}_{v \succeq 0}$ then we have

$$|[\tau]| = \frac{n!}{\prod_{v \succeq 0} s_v!}$$

Indeed, $n!$ is the number of all permutations in the set $\{1, 2, \dots, n\}$, while for each vertex v, $s_v!$ are the permutation of the children. So $n!/\prod_{v \succeq 0} s_v!$ is the number of permutations of the vertices of τ different from the root which do not change the order of the children in every vertex.
Now observe that the factor (4.20) depend actually only on the plane rooted tree associated to τ. We have indeed

$$\phi^*_{\gamma_0}(\tau, \boldsymbol{\rho}) = \phi^*_{\gamma_0}([\tau], \boldsymbol{\rho}) = \phi^*_{\gamma_0}(t, \boldsymbol{\rho}) = \prod_{v \succeq 0} \left[\sum_{\substack{(\gamma_{v1}, \dots, \gamma_{vs_v}) \in \mathcal{P}^{s_v} \\ \gamma_{vi} \not\sim \gamma_v,\ \gamma_{vi} \sim \gamma_{vj}}} \rho_{\gamma_{v1}} \dots \rho_{\gamma_{vs_v}} \right] \quad (4.27)$$

Note that, since in each vertex v the sum over polymers $\gamma_{v_1}, \dots, \gamma_{v_{s_v}}$ associated to children of v depends on the polymer γ_v associated to v, in the expression above the order of the product is relevant and it is organized in such way that products corresponding to ancestors are at the left of products corresponding to descendants. In (4.27) it also adopted the convention that the product in brackets is equal to 1 for a vertex v such that $s_v = 0$.
We now ready to reorganize the sum in the the r.h.s. of (4.19)

$$\Pi^*_{\gamma_0}(\boldsymbol{\rho}) = \sum_{n \geq 0} \frac{1}{n!} \sum_{\tau \in T_n^0} \phi^*_{\gamma_0}(\tau, \boldsymbol{\rho}) \ = \sum_{n \geq 0} \frac{1}{n!} \sum_{t \in \mathcal{T}_n^0} \sum_{\tau \in t} \phi^*_{\gamma_0}(t, \rho) =$$

$$= \sum_{n \geq 0} \frac{1}{n!} \sum_{t \in \mathcal{T}_n^0} \phi^*_{\gamma_0}(t, \rho) \sum_{\tau \in t} 1 = \sum_{n \geq 0} \frac{1}{n!} \sum_{t \in \mathcal{T}_n^0} \phi^*_{\gamma_0}(t, \rho) |t| =$$

$$= \sum_{n \geq 0} \sum_{t \in \mathcal{T}_n^0} \Big[\prod_{v \succeq 0} \frac{1}{s_v!} \Big] \phi^*_{\gamma_0}(t, \rho) = \sum_{n \geq 0} \sum_{t \in \mathcal{T}_n^0} \prod_{v \succeq 0} \left[\frac{1}{s_v!} \sum_{\substack{(\gamma_{v1}, \dots, \gamma_{vs_v}) \in \mathcal{P}^{s_v} \\ \gamma_{vi} \not\sim \gamma_v,\ \gamma_{vi} \sim \gamma_{vj}}} \rho_{\gamma_{v1}} \dots \rho_{\gamma_{vs_v}} \right]$$

In conclusion we have obtained

$$\Pi^*_{\gamma_0}(\boldsymbol{\rho}) \ =$$

$$= \sum_{t \in \mathcal{T}^0} \prod_{v \succeq 0} \left\{ \frac{1}{s_v!} \sum_{(\gamma_{v1}, \dots, \gamma_{vs_v}) \in \mathcal{P}^{s_v}} \prod_{i=1}^{s_v} \mathbb{1}_{\{\gamma_{vi} \nsim \gamma_v\}} \prod_{1 \leq i < j \leq s_v} \mathbb{1}_{\{\gamma_{vi} \sim \gamma_{vj}\}} \rho_{\gamma_{v1}} \dots \rho_{\gamma_{vs_v}} \right\} \quad (4.28)$$

In a completely analogous way we also can obtain

$$\Pi^{\mathrm{Dob}}_{\gamma_0}(\boldsymbol{\rho}) \ =$$

$$= \sum_{t\in\mathcal{T}^0}\prod_{v\succeq 0}\left\{\frac{1}{s_v!}\sum_{(\gamma_{v1},\dots,\gamma_{vs_v})\in\mathcal{P}^{s_v}}\prod_{i=1}^{s_v}\mathbb{1}_{\{\gamma_{vi}\nsim\gamma_v\}}\prod_{1\le i<j\le s_v}\mathbb{1}_{\{\gamma_{vi}\neq\gamma_{vj}\}}\rho_{\gamma_{v1}}\dots\rho_{\gamma_{vs_v}}\right\} \tag{4.29}$$

and

$$\Pi^{\mathrm{KP}}_{\gamma_0}(\boldsymbol{\rho}) \;=\; \sum_{t\in\mathcal{T}^0}\prod_{v\succeq 0}\left\{\frac{1}{s_v!}\sum_{(\gamma_{v1},\dots,\gamma_{vs_v})\in\mathcal{P}^{s_v}}\prod_{i=1}^{s_v}\mathbb{1}_{\{\gamma_{vi}\nsim\gamma_v\}}\rho_{\gamma_{v1}}\dots\rho_{\gamma_{vs_v}}\right\} \tag{4.30}$$

4.2.2 Trees and convergence

We start by defining a proper domain in the space of function $(0,\infty)^{\mathcal{P}}$, i.e. the space of the functions $\boldsymbol{\mu}:\mathcal{P}\to(0,\infty):\gamma\mapsto\mu_\gamma$. We are agree that given two functions $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ in $(0,\infty)^{\mathcal{P}}$, we say that $\boldsymbol{\mu}<\boldsymbol{\nu}$ if and only if $\mu_\gamma<\nu_\gamma$ for all $\gamma\in\mathcal{P}$.

Let us consider, for all $n\in\mathbb{N}$ and for all $(\gamma_0,\gamma_1,\dots,\gamma_n)\in\mathcal{P}^{n+1}$, numbers $b_n(\gamma_0;\gamma_1,\dots,\gamma_n)$ such that $b_n(\gamma_0;\gamma_1,\dots,\gamma_n)\ge 0$. Once numbers $b_n(\gamma_0;\gamma_1,\dots,\gamma_n)$ are given, we can define a function

$$\varphi^b:(0,\infty)^{\mathcal{P}}\to(0,\infty]^{\mathcal{P}}:\boldsymbol{u}\mapsto\varphi^b(\boldsymbol{u})$$

with entries

$$[\varphi^b(\boldsymbol{u})]_\gamma\doteq\varphi^b_\gamma(\boldsymbol{u}) \;=\; 1+\sum_{n\ge1}\sum_{(\gamma_1,\dots,\gamma_n)\in\mathcal{P}^n} b_n(\gamma;\gamma_1,\dots,\gamma_n)\,u_{\gamma_1}\dots u_{\gamma_n} \tag{4.31}$$

We define the set

$$\mathcal{D}^b=\Big\{\boldsymbol{u}\in(0,\infty)^{\mathcal{P}}:\;\;\varphi^b_\gamma(\boldsymbol{u})<+\infty,\;\forall\gamma\in\mathcal{P}\Big\}$$

Then the restriction of φ^b to $\mathcal{D}^b$ is a function in $(0,\infty)^{\mathcal{P}}$, i.e.

$$\varphi^b_\gamma(\boldsymbol{u})\;<\;\infty,\;\forall\gamma\in\mathcal{P},\qquad\text{whenever }\boldsymbol{u}\in\mathcal{D}^b$$

Note also that, if $\boldsymbol{u}\in\mathcal{D}^b$ and $\boldsymbol{u}'<\boldsymbol{u}$ then also $\boldsymbol{u}'\in\mathcal{D}^b$.
Let now $\boldsymbol{\mu}:\mathcal{P}\to(0,\infty)$ be a function in the set $\mathcal{D}^b\subset(0,\infty)^{\mathcal{P}}$ and let $\boldsymbol{r}\in(0,\infty)^{\mathcal{P}}$ be defined such that its entries r_γ, as γ varies in $\mathcal{P}$, are given by

$$r_\gamma=\frac{\mu_\gamma}{\varphi^b_\gamma(\boldsymbol{\mu})} \tag{4.32}$$

Note that $\boldsymbol{r}\in\mathcal{D}^b$ because $\boldsymbol{r}\le\boldsymbol{\mu}$ by construction, since $\varphi^b_\gamma(\boldsymbol{\mu})\ge1$ for all $\gamma\in\mathcal{P}$. Morevover the assumption $\boldsymbol{\mu}\in(0,\infty)^{\mathcal{P}}$ is equivalent to say

$$\mu_\gamma>0\qquad\text{for all }\gamma_0\in\mathcal{P} \tag{4.33}$$

while assumption $\boldsymbol{\mu}\in\mathcal{D}^b$ means that

$$\varphi^b_\gamma(\boldsymbol{\mu})<+\infty\qquad\text{for all }\gamma\in\mathcal{P} \tag{4.34}$$

Assumptions (4.33) and (4.34) guarantee that $\boldsymbol{r} \in (0,\infty)^{\mathcal{P}}$, i.e.

$$r_\gamma > 0 \qquad \text{for all } \gamma \in \mathcal{P}$$

Let us now consider, for any $\boldsymbol{\rho} \in \mathcal{D}^p$, the map $T^{\boldsymbol{\rho}} = \boldsymbol{\rho}\varphi^b$. $T^{\boldsymbol{\rho}}$ is the map

$$T^{\boldsymbol{\rho}} : (0,\infty)^{\mathcal{P}} \to (0,\infty]^{\mathcal{P}} : \boldsymbol{u} \mapsto T^{\boldsymbol{\rho}}(\boldsymbol{u})$$

with entries

$$[T^{\boldsymbol{\rho}}(\boldsymbol{u})]_\gamma \doteq T^{\boldsymbol{\rho}}_\gamma(\boldsymbol{u}) = \rho_\gamma \varphi^b_\gamma(\boldsymbol{u}) \qquad\qquad \gamma \in \mathcal{P}$$

From (4.32) we get

$$\boldsymbol{\mu} = T^{\boldsymbol{r}}(\boldsymbol{\mu}) \tag{4.35}$$

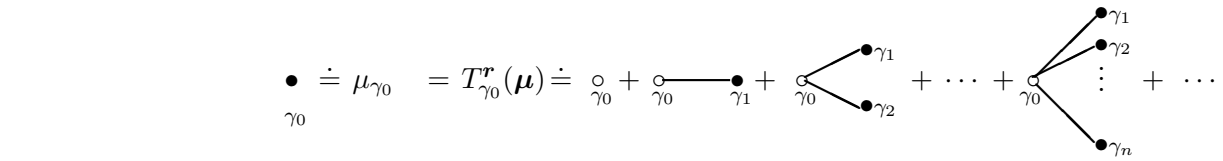
I.e. $\boldsymbol{\mu}$ is fixed point for the map $T^{\boldsymbol{r}}$. So, by (4.35) we have that, for all $\gamma_0 \in \mathcal{P}$

$$\mu_{\gamma_0} = T^{\boldsymbol{r}}_{\gamma_0}(\boldsymbol{\mu}) =$$

$$= r_{\gamma_0} \; + \; r_{\gamma_0} \sum_{\gamma_1 \in \mathcal{P}} b_1(\gamma_0;\gamma_1)\, \mu_{\gamma_1} + r_{\gamma_0} \sum_{(\gamma_1,\gamma_2)\in\mathcal{P}^2} b_2(\gamma_0;\gamma_1,\gamma_2)\, \mu_{\gamma_1}\mu_{\gamma_2} + \; \dots$$

$$\dots \; + \; r_{\gamma_0} \sum_{(\gamma_1,\dots,\gamma_n)\in\mathcal{P}^n} b_n(\gamma_0;\gamma_1,\dots,\gamma_n)\, \mu_{\gamma_1}\dots\mu_{\gamma_n} \; + \dots \tag{4.36}$$

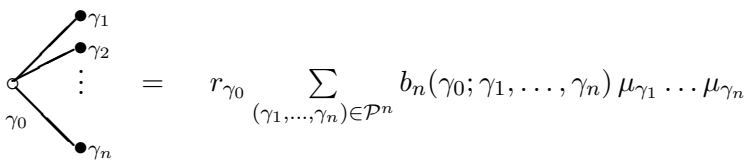
Equation (4.36), recalling the definition (4.31) of $\varphi^b_\gamma(\boldsymbol{\mu})$ can be visualized in the diagrammatic form



where

$$\circ_{\gamma_0} = r_{\gamma_0} \qquad\qquad \bullet_{\gamma_i} = \mu_{\gamma_i}$$

and, for any $n \geq 1$



The iteration $[T^r]^2(\mu) = T^r(T^r(\mu))$ corresponds to replacing each of the bullets by each one of the diagrams of the expansion for T^r.
This leads to plane rooted trees of up to two generations, with open circles at first-generation vertices and bullets at second-generation ones. The k-th iteration of T involves all possible *plane rooted trees* previously seen with generation

up to k. In each tree of the expansion, vertices of the last generation are occupied by bullets and all the others by open circles. Let us recall that $\mathcal{T}^{0,k}$ denotes the set of trees with maximal depth (or maximal generation number) equal to k. A straightforward inductive argument shows that

$$[T^{\boldsymbol{r}}]^k_{\gamma_0}(\boldsymbol{\mu}) \;=\; r_{\gamma_0}\Big[\sum_{\ell=0}^{k-1}\Phi^{(\ell)}_{\gamma_0}(\boldsymbol{r}) + R^{(k)}_{\gamma_0}(\boldsymbol{r},\boldsymbol{\mu})\Big] \tag{4.37}$$

with

$$\Phi^{(\ell)}_{\gamma_0}(\boldsymbol{r}) \;=\; \sum_{t\in\mathcal{T}^{0,\ell}}\prod_{v\succeq 0}\Big\{\sum_{(\gamma_{v^1},\dots,\gamma_{v^{s_v}})\in\mathcal{P}^{s_v}} b_{s_v}(\gamma_v;\gamma_{v^1},\dots,\gamma_{v^{s_v}})\, r_{\gamma_{v^1}}\dots r_{\gamma_{v^{s_v}}}\Big\} \tag{4.38}$$

while

$$R^{(k)}_{\gamma_0}(\boldsymbol{r},\boldsymbol{\mu}) = \sum_{t\in\mathcal{T}^{0,k}}\prod_{v\succeq 0}\Big\{\sum_{(\gamma_{v^1},\dots,\gamma_{v^{s_v}})\in\mathcal{P}^{s_v}} b_{s_v}(\gamma_v;\gamma_{v^1},\dots,\gamma_{v^{s_v}})\, \chi^t_{\gamma_{v^1}}\dots\chi^t_{\gamma_{v^{s_v}}}\Big\} \tag{4.39}$$

where

$$\chi^t_{\gamma_v} = \begin{cases} r_{\gamma_v} & \text{if} \quad d_t(v) < k \\ \\ \mu_{\gamma_v} & \text{if} \quad d_t(v) = k \end{cases} \tag{4.40}$$

with, we recall, $d_t(v)$ indicating the depth (distance from the root) of v in t. In other words $R^{(k)}_{\gamma_0}(\boldsymbol{r},\boldsymbol{\mu})$ has an expression similar to $\Phi^{(k)}_{\gamma_0}(\boldsymbol{r})$ but with the activities of the vertex of the k-th generation weighted by $\boldsymbol{\mu}$. Here we agree that if v is such that $s_v = 0$ then $b_0(\gamma_v) \equiv 1$. Now, by (4.36) we have

$$[T^{\boldsymbol{r}}]^k_{\gamma_0}(\boldsymbol{\mu}) \;=\; \mu_{\gamma_0}$$

which implies immediately, via (4.37),

$$r_{\gamma_0}\sum_{\ell=0}^{k-1}\Phi^{(\ell)}_{\gamma_0}(\boldsymbol{r}) \;\le\; \mu_{\gamma_0} \qquad \text{for all } k\in\mathbb{N} \tag{4.41}$$

Equation (4.41) immediately implies the following proposition.

Proposition 4.1 *Let $\boldsymbol{\mu}$ be a function $\boldsymbol{\mu}\in(0,\infty)^{\mathcal{P}}$ and, for any $\gamma\in\mathcal{P}$ let $\varphi^b_\gamma(\boldsymbol{\mu})$ be function defined in (4.31) supposed to satisfy (4.34). Let $\boldsymbol{r}\in(0,\infty)^{\mathcal{P}}$ be defined by (4.32). Then, for all $\boldsymbol{\rho}\le\boldsymbol{r}$*

i) The series

$$\Phi^b_{\gamma_0}(\boldsymbol{\rho}) \;:=\; \sum_{t\in\mathcal{T}^0}\prod_{v\succeq 0}\Big\{\sum_{(\gamma_{v^1},\dots,\gamma_{v^{s_v}})\in\mathcal{P}^{s_v}} b_{s_v}(\gamma_v;\gamma_{v^1},\dots,\gamma_{v^{s_v}})\,\rho_{\gamma_{v^1}}\dots\rho_{\gamma_{v^{s_v}}}\Big\} \tag{4.42}$$

converges for each $\gamma_0\in\mathcal{P}$ and admits, for each $\gamma_0\in\mathcal{P}$, the bound

$$\Phi^b_{\gamma_0}(\boldsymbol{\rho}) \le \Phi^b_{\gamma_0}(\boldsymbol{r}) \le \varphi^b_{\gamma_0}(\boldsymbol{\mu}) \tag{4.43}$$

ii)

$$\rho_{\gamma_0}\Phi^b_{\gamma_0}(\boldsymbol{\rho}) = \lim_{n\to\infty}[T^{\boldsymbol{\rho}}]^n_{\gamma_0}(\boldsymbol{\rho}) \tag{4.44}$$

and $\rho_{\gamma_0}\Phi^b_{\gamma_0}(\boldsymbol{\rho})$ is solution of the equation (4.36), i.e. is fixed point of the map $T^{\boldsymbol{\rho}}$, i.e.

$$\rho_{\gamma_0}\Phi^b_{\gamma_0}(\boldsymbol{\rho}) = T^{\boldsymbol{\rho}}_{\gamma_0}\Big(\rho_\gamma\Phi^b_\gamma(\boldsymbol{\rho})\Big) \tag{4.45}$$

This proposition can be viewed as a generalization of the Lagrange inversion formula for series depending on infinite (countable) variables.

Proof.
i). By (4.37) we get

$$r_{\gamma_0}\sum_{\ell=0}^{n}\Phi^{(\ell)}_{\gamma_0}(\boldsymbol{r}) \;\le\; [T^{\boldsymbol{r}}]^{n+1}_{\gamma_0}(\boldsymbol{\mu}) \qquad \forall n\in\mathbb{N}$$

but, by definition (4.36) we have, for any $k\in\mathbb{N}$ that $[T^{\boldsymbol{r}}]^k_{\gamma_0}(\boldsymbol{\mu}) = \mu_{\gamma_0}$. So we obtain

$$r_{\gamma_0}\sum_{\ell=0}^{n}\Phi^{(\ell)}_{\gamma_0}(\boldsymbol{r}) \;\le\; \mu_{\gamma_0} \qquad \text{for all } n$$

i.e., by (4.32),

$$\sum_{\ell=0}^{n}\Phi^{(\ell)}_{\gamma_0}(\boldsymbol{r}) \;\le\; \varphi^b_{\gamma_0}(\boldsymbol{\mu}) \qquad \text{for all } n$$

which implies

$$\Phi^b_{\gamma_0}(\boldsymbol{r}) \le \varphi^b_{\gamma_0}(\boldsymbol{\mu})$$

Therefore, by monotonicity, for any $\boldsymbol{\rho}\le\boldsymbol{r}$

$$\Phi^b_{\gamma_0}(\boldsymbol{\rho}) \le \Phi^b_{\gamma_0}(\boldsymbol{r}) \le \varphi^b_{\gamma_0}(\boldsymbol{\mu})$$

ii) By (4.37) we have that

$$[T^{\boldsymbol{\rho}}]^k_{\gamma_0}(\boldsymbol{\rho}) \;=\; \rho_{\gamma_0}\Big[\sum_{\ell=0}^{k-1}\Phi^{(\ell)}_{\gamma_0}(\boldsymbol{\rho}) + R^{(k)}_{\gamma_0}(\boldsymbol{\rho},\boldsymbol{\mu})|_{\boldsymbol{\mu}=\boldsymbol{\rho}}\Big] \tag{4.46}$$

But, recalling definition (4.40)

$$R^{(k)}_{\gamma_0}(\boldsymbol{\rho},\boldsymbol{\mu})|_{\boldsymbol{\mu}=\boldsymbol{\rho}} \;=$$

$$= \sum_{t\in\mathcal{T}^{0,k}}\prod_{v\succeq 0}\left\{\sum_{(\gamma_{v^1},\dots,\gamma_{v^{s_v}})\in\mathcal{P}^{s_v}} b_{s_v}(\gamma_v;\gamma_{v^1},\dots,\gamma_{v^{s_v}})\,\rho_{\gamma_{v^1}}\dots\rho_{\gamma_{v^{s_v}}}\right\} \;=\; \Phi^{(k)}_{\gamma_0}(\boldsymbol{\rho})$$

So

$$[T^{\boldsymbol{\rho}}]^k_{\gamma_0}(\boldsymbol{\rho}) \;=\; \rho_{\gamma_0}\sum_{\ell=0}^{k}\Phi^{(\ell)}_{\gamma_0}(\boldsymbol{\rho})$$

Hence, for any $\boldsymbol{\rho} \leq \boldsymbol{r}$

$$\lim_{k\to\infty} [T^{\boldsymbol{\rho}}]^k_{\gamma_0}(\boldsymbol{\rho}) \;=\; \lim_{k\to\infty} \rho_{\gamma_0} \sum_{\ell=0}^{k} \Phi^{(\ell)}_{\gamma_0}(\boldsymbol{\rho}) \;=\; \rho_{\gamma_0} \Phi^b_{\gamma_0}(\boldsymbol{\rho})$$

Finally, for any $\boldsymbol{\rho} \leq \boldsymbol{r}$

$$\rho_{\gamma_0} \Phi^b_{\gamma_0}(\boldsymbol{\rho}) = \lim_{n\to\infty} [T^{\boldsymbol{\rho}}]^{n+1}_{\gamma_0}(\boldsymbol{\rho}) \;=\; \lim_{n\to\infty} T^{\boldsymbol{\rho}}_{\gamma_0}\Big([T^{\boldsymbol{\rho}}]^n_{\gamma_0}(\boldsymbol{\rho})\Big) \;=\; \rho_{\gamma_0} \lim_{n\to\infty} \varphi^b_{\gamma_0}\Big([T^{\boldsymbol{\rho}}]^n_{\gamma_0}(\boldsymbol{\rho})\Big) =$$

$$= \rho_{\gamma_0} \varphi^b_{\gamma}\Big(\lim_{n\to\infty} [T^{\boldsymbol{\rho}}]^n_{\gamma_0}(\boldsymbol{\rho})\Big) \;=\; \rho_{\gamma_0} \varphi^b_{\gamma_0}\Big(\rho_{\gamma_0} \Phi^b_{\gamma_0}(\boldsymbol{\rho})\Big) \;=\; T^{\boldsymbol{\rho}}_{\gamma_0}\Big(\rho_{\gamma_0} \Phi^b_{\gamma_0}(\boldsymbol{\rho})\Big)$$

□

4.2.3 Convergence criteria

Fernández-Procacci criterion.
Let us now choose

$$b_n(\gamma_0;\gamma_1,\dots,\gamma_n) = b^*_n(\gamma_0;\gamma_1,\dots,\gamma_n) \doteq \frac{1}{n!} \prod_{i=1}^{n} \mathbb{1}_{\{\gamma_i \not\sim \gamma_0\}} \prod_{1\leq i<j\leq n} \mathbb{1}_{\{\gamma_i \sim \gamma_j\}} \tag{4.47}$$

and thus

$$\varphi^{b^*}_{\gamma_0}(\boldsymbol{\mu}) \;=\; 1 + \sum_{n\geq 1} \frac{1}{n!} \sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{P}^n \\ \gamma_i \not\sim \gamma_0,\ \gamma_i\sim\gamma_j}} \mu_{\gamma_1} \dots \mu_{\gamma_n} \;=\; \Xi_{\mathcal{P}_{\gamma_0}}(\boldsymbol{\mu}) \tag{4.48}$$

Then, proposition 4.1 tells us that the series

$$\Phi^{b^*}_{\gamma_0}(\boldsymbol{\rho}) \;=\; \sum_{t\in\mathcal{T}^0} \prod_{v\succeq 0} \left\{ \sum_{(\gamma_{v1},\dots,\gamma_{vs_v})\in\mathcal{P}^{s_v}} \frac{1}{s_v!} \prod_{i=1}^{s_v} \mathbb{1}_{\{\gamma_{vi} \not\sim \gamma_v\}} \prod_{1\leq i<j\leq s_v} \mathbb{1}_{\{\gamma_{vi}\sim\gamma_{vj}\}}\, \rho_{\gamma_{v1}} \dots \rho_{\gamma_{vs_v}} \right\} \tag{4.49}$$

converges as soon as $\boldsymbol{\rho} \leq \boldsymbol{r}^*$ with

$$r^*_\gamma = \frac{\mu_\gamma}{\Xi_{\mathcal{P}_\gamma}(\boldsymbol{\mu})} \tag{4.50}$$

comparing (4.28) with (4.49) we immediately see that

$$\Pi^*_{\gamma_0}(\boldsymbol{\rho}) = \Phi^{b^*}_{\gamma_0}(\boldsymbol{\rho})$$

So we immediately get, by Proposition 4.1, the following criterion for the convergence of cluster expansions.

Theorem 4.1 *Choose $\boldsymbol{\mu} \in \mathcal{D}^{b^*} \subset (0,\infty)^{\mathcal{P}}$ and let $\boldsymbol{r}^* \in (0,\infty)^{\mathcal{P}}$ s.t.*

$$r^*_{\gamma_0} = \frac{\mu_{\gamma_0}}{\Xi_{\mathcal{P}_{\gamma_0}}(\boldsymbol{\mu})} \tag{4.51}$$

Let $\boldsymbol{\rho}$ such that

$$\rho_\gamma \;\leq\; r^*_\gamma \qquad \forall \gamma \in \mathcal{P} \tag{4.52}$$

Then the series $|\Pi|_{\gamma_0}(\boldsymbol{\rho})$ *defined in (4.9) is finite for each* $\gamma_0 \in \mathcal{P}$ *and*

$$|\Pi|_{\gamma_0}(\boldsymbol{\rho}) \leq \Xi_{\mathcal{P}_{\gamma_0}}(\boldsymbol{\mu}) \tag{4.53}$$

and hence

$$\rho_{\gamma_0}|\Pi|_{\gamma_0}(\boldsymbol{\rho}) \leq \mu_{\gamma_0} \tag{4.54}$$

for each $\gamma_0 \in \mathcal{P}$.

Proof. By proposition 4.1 we have immediately that the series $\Pi^*_{\gamma_0}(\boldsymbol{\rho})$ defined in (4.19) is finite for each $\gamma_0 \in \mathcal{P}$ and for all $\boldsymbol{\rho}$ such that $\rho_\gamma \leq r^*_\gamma$ where r^*_γ is defined in (4.52). Moreover

$$\Pi^*_{\gamma_0}(\boldsymbol{\rho}) \leq \Xi_{\mathcal{P}_{\gamma_0}}(\boldsymbol{\mu})$$

for each $\gamma_0 \in \mathcal{P}$. Now recalling (4.18) we obtain that the same is true also for the series $|\Pi|_{\gamma_0}(\boldsymbol{\rho})$. □

<u>*Dobrushin criterion*</u>.
If we now choose

$$b_n(\gamma_0;\gamma_1,\dots,\gamma_n) = b^{\mathrm{Dob}}_n(\gamma_0;\gamma_1,\dots,\gamma_n) \doteq \frac{1}{n!}\prod_{i=1}^{n} \mathbb{1}_{\{\gamma_i \not\sim \gamma_0\}} \prod_{1\leq i<j\leq n} \mathbb{1}_{\{\gamma_i \neq \gamma_j\}} \tag{4.55}$$

and thus

$$\varphi^{b^{\mathrm{Dob}}}_{\gamma_0}(\boldsymbol{\mu}) \;=\; 1 + \sum_{n\geq 1}\frac{1}{n!}\sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{P}^n \\ \gamma_0\nsim\gamma_i\,,\,\gamma_i\neq\gamma_j}} \mu_{\gamma_1}\dots\mu_{\gamma_n} \;=\; \prod_{\gamma\nsim\gamma_0}[1+\mu_\gamma]\,, \tag{4.56}$$

Then again proposition 4.1 tells us that the series

$$\Phi^{b^{\mathrm{Dob}}}_{\gamma_0}(\boldsymbol{\rho}) \;=\; \sum_{t\in\mathcal{T}^0}\prod_{v\succeq 0}\left\{\sum_{(\gamma_{v1},\dots,\gamma_{vs_v})\in\mathcal{P}^{s_v}} \frac{1}{s_v!}\prod_{i=1}^{s_v}\mathbb{1}_{\{\gamma_{vi}\not\sim\gamma_v\}} \prod_{1\leq i<j\leq s_v}\mathbb{1}_{\{\gamma_{vi}\neq\gamma_{vj}\}}\,\rho_{\gamma_{v1}}\dots\rho_{\gamma_{vs_v}}\right\} \tag{4.57}$$

converges as soon as $\boldsymbol{\rho} \leq \boldsymbol{r}^{\mathrm{Dob}}$ with

$$r^{\mathrm{Dob}}_\gamma = \frac{\mu_\gamma}{\prod_{\gamma\nsim\gamma_0}[1+\mu_\gamma]} \tag{4.58}$$

Comparing (4.29) with (4.57) we get

$$\Pi^{\mathrm{Dob}}_{\gamma_0}(\boldsymbol{\rho}) = \Phi^{b^{\mathrm{Dob}}}_{\gamma_0}(\boldsymbol{\rho})$$

So Proposition 4.1 also yields the following (weaker) criterion for the convergence of cluster expansions.

Corollary 4.1 (Dobrushin) *Choose* $\boldsymbol{\mu} \in \mathcal{D}^{b^{\mathrm{Dob}}}$ *and let* $\boldsymbol{r}^{\mathrm{Dob}} \in (0,\infty)^{\mathcal{P}}$ *s.t.*

$$r^{\mathrm{Dob}}_\gamma = \frac{\mu_\gamma}{\prod_{\tilde\gamma\nsim\gamma}[1+\mu_{\tilde\gamma}]} \tag{4.59}$$

Let $\boldsymbol{\rho} \in [0,\infty)^{\mathcal{P}}$ *such that*

$$\rho_\gamma \;\le\; r_\gamma^{\mathrm{Dob}} = \frac{\mu_\gamma}{\prod_{\tilde\gamma \nsim \gamma}[1+\mu_{\tilde\gamma}]} \qquad \forall \gamma \in \mathcal{P} \tag{4.60}$$

Then the series $|\Pi|_{\gamma_0}(\boldsymbol{\rho})$ *defined in (4.9) is finite for each* $\gamma \in \mathcal{P}$ *and*

$$|\Pi|_\gamma(\boldsymbol{\rho}) \le \prod_{\tilde\gamma \nsim \gamma}[1+\mu_{\tilde\gamma}] \tag{4.61}$$

or

$$\rho_\gamma |\Pi|_\gamma(\boldsymbol{\rho}) \le \mu_\gamma \tag{4.62}$$

for each $\gamma \in \mathcal{P}$.

For the benefit of the readers we stress that in the literature the Dobrushin condition is generally written in a different (but equivalent) form. In particular, in the Dobrushin paper [10] and also in [26] the condition (4.60) is written as follows.

$$\rho_\gamma \;\le\; r_\gamma^{\mathrm{Dob}} = (e^{\tilde\mu_\gamma}-1)e^{-\sum_{\tilde\gamma\in\mathcal{P}:\,\tilde\gamma\nsim\gamma}\tilde\mu_{\tilde\gamma}} \qquad \forall \gamma \in \mathcal{P} \tag{4.63}$$

This is clearly the same condition (4.60) by defining

$$\tilde\mu_\gamma = \log[1+\mu_\gamma] \tag{4.64}$$

Kopteckýˊ-Preiss criterion.
Finally, if we now choose

$$b_n(\gamma_0;\gamma_1,\dots,\gamma_n) = b_n^{\mathrm{KP}}(\gamma_0;\gamma_1,\dots,\gamma_n) \doteq \frac{1}{n!}\prod_{i=1}^{n} \mathbb{1}_{\{\gamma_i \nsim \gamma_0\}}$$

and

$$\varphi_{\gamma_0}^{\mathrm{KP}}(\boldsymbol{\mu}) \;=\; 1+\sum_{n\ge1}\frac{1}{n!}\sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{P}^n\\ \gamma_0\nsim\gamma_i\,,\,1\le i\le n}} \mu_{\gamma_1}\dots\mu_{\gamma_n} \;=\; \exp\Big[\sum_{\gamma\nsim\gamma_0}\mu_\gamma\Big] \tag{4.65}$$

Then once again proposition 4.1 tells us that the series

$$\Phi_{\gamma_0}^{b^{\mathrm{KP}}}(\boldsymbol{\rho}) \;=\; \sum_{t\in\mathcal{T}^0}\prod_{v\succeq 0}\left\{\sum_{(\gamma_{v^1},\dots,\gamma_{v^{s_v}})\in\mathcal{P}^{s_v}} \frac{1}{s_v!}\prod_{i=1}^{s_v}\mathbb{1}_{\gamma_{v^i}\nsim\gamma_v}\;\; \rho_{\gamma_{v^1}}\dots\rho_{\gamma_{v^{s_v}}}\right\} \tag{4.66}$$

converges as soon as $\boldsymbol{\rho} \le \boldsymbol{r}^{\mathrm{KP}}$ with

$$r_\gamma^{\mathrm{KP}} = \frac{\mu_\gamma}{\exp\Big[\sum_{\tilde\gamma\nsim\gamma}\mu_{\tilde\gamma}\Big]} \tag{4.67}$$

Comparing (4.30) with (4.66) we get

$$\Pi_{\gamma_0}^{\mathrm{KP}}(\boldsymbol{\rho}) = \Phi_{\gamma_0}^{b^{\mathrm{KP}}}(\boldsymbol{\rho})$$

So we get the criterion of Kotecký and Preiss. Namely,

Corollary 4.2 (Kotecký-Preiss) *Choose* $\boldsymbol{\mu} \in \mathcal{D}^{b^{\mathrm{KP}}}$ *and let* $\boldsymbol{r}^{\mathrm{KP}} \in (0,\infty)^{\mathcal{P}}$ *s.t.*

$$r^{\mathrm{KP}}_{\gamma_0} = \frac{\mu_{\gamma_0}}{\exp\Big[\sum_{\gamma \nsim \gamma_0} \mu_\gamma\Big]} \tag{4.68}$$

Let $\boldsymbol{\rho} \in [0,\infty)^{\mathcal{P}}$ *such that*

$$\rho_\gamma \;\le\; r^{\mathrm{KP}}_\gamma \;=\; \frac{\mu_\gamma}{\exp\Big[\sum_{\tilde\gamma \nsim \gamma} \mu_{\tilde\gamma}\Big]}, \qquad \forall \gamma \in \mathcal{P} \tag{4.69}$$

Then the series $|\Pi|_{\gamma_0}(\boldsymbol{\rho})$ *defined in (4.9) is finite for each* $\gamma_0 \in \mathcal{P}$ *and*

$$|\Pi|_{\gamma_0}(\boldsymbol{\rho}) \le \exp\Big[\sum_{\gamma \nsim \gamma_0} \mu_\gamma\Big] \tag{4.70}$$

or

$$\rho_{\gamma_0} |\Pi|_{\gamma_0}(\boldsymbol{\rho}) \le \mu_{\gamma_0}$$

for each $\gamma_0 \in \mathcal{P}$.

Remark. Again for the benefit of the readers we stress that in the literature the Kotecky-Preiss condition is generally written in a different (but equivalent) form. Namely in the original Kotecky-Preiss paper convergence is guaranteed by choosing $\boldsymbol{\rho}$ such that there is a function a_γ such that

$$\sum_{\tilde\gamma \nsim \gamma} \rho_{\tilde\gamma} e^{a_{\tilde\gamma}} \le a_\gamma \qquad\qquad \forall \gamma \in \mathcal{P} \tag{4.71}$$

This is clearly the same condition (4.68) by setting

$$\mu_\gamma = \rho_\gamma e^{a_\gamma} \tag{4.72}$$

Summing up, available convergence conditions are of the form

$$\rho_\gamma \;\le\; r_\gamma = \frac{\mu_\gamma}{\varphi_\gamma(\mu)} \tag{4.73}$$

with

$$\varphi_\gamma(\boldsymbol{\mu}) \;=\; \begin{cases} \exp\Big[\sum_{\tilde\gamma \nsim \gamma} \mu_{\tilde\gamma}\Big] & \text{(Kotecký-Preiss)} \\ \prod_{\tilde\gamma \nsim \gamma}(1+\mu_{\tilde\gamma}) & \text{(Dobrushin)} \\ \Xi_{\mathcal{P}_\gamma}(\boldsymbol{\mu}) & \text{(Fernández-Procacci)} \end{cases} \tag{4.74}$$

Each condition is strictly weaker than the preceding one. Namely, since, for fixed $\boldsymbol{\mu} \in (0,\infty)^{\mathcal{P}}$,

$$\Xi_{\mathcal{P}_\gamma}(\boldsymbol{\mu}) \le \prod_{\tilde\gamma \nsim \gamma} [1+\mu_{\tilde\gamma}] \le \exp\Big[\sum_{\gamma \nsim \gamma_0} \mu_{\tilde\gamma}\Big]$$

we get

$$\boldsymbol{r}_\gamma^* \geq \boldsymbol{r}_\gamma^{\mathrm{Dob}} \geq \boldsymbol{r}_\gamma^{\mathrm{KP}}$$

So the criterion given by the Corollary 4.2 (i.e. Kotecky-Preiss condition) yields the worst estimate for convergence radius for the cluster expansion; the Dobrishin Criterion of Corollary 4.1 gives an estimate which is better (i.e. larger) than that given by the Kotecký-Preiss criterion for the same radius and finally the criterion 4.1 give the best estimate for convergence radius for the cluster expansion among the three proposed.

4.2.4 Elementary Examples

In this section we give some elementary in order to illustrate how the criterion (4.52) represents a sensible improvement on previous criteria in applications.

Example 1. The Domino model on $\mathbb{Z}^2$.

This model has also been considered by Dobrushin in [10]. The elements of the polymer space $\mathcal{P}$ are in this case nearest neighbor bonds of the bidimensional cubic lattice. For any $\gamma \in \mathcal{P}$ we put $\rho_\gamma = \varepsilon$, where $\varepsilon > 0$ (all polymers have the same activity). Two polymers are incompatible if and only if they have non empty intersection. We can choose by symmetry that the function μ_γ appearing in the Kotecky-Preiss, Dobrushin and Fernández-Procacci criteria are constant at the value μ.

The Kotecky-Preiss criterion (4.69) for the domino model then reads as

$$\rho_\gamma \leq \mu_\gamma e^{-\sum_{\tilde{\gamma} \not\sim \gamma} \mu_{\tilde{\gamma}}} \iff \varepsilon \leq \mu e^{-7\mu}$$

which yelds at best

$$\varepsilon \leq \frac{1}{7e} \approx 0.0525$$

On the other hand the Dobrushin condition (4.60) reads

$$\rho_\gamma \leq \frac{\mu_\gamma}{\prod_{\tilde{\gamma} \not\sim \gamma}[1+\mu_{\tilde{\gamma}}]} \iff \varepsilon \leq \frac{\mu}{(1+\mu)^7}$$

which yelds at best

$$\varepsilon \leq \frac{\frac{1}{6}}{(1+\frac{1}{6})^7} \approx 0.0566$$

Finally, the condition (4.52) gives

$$\rho_\gamma \leq \frac{\mu_\gamma}{\Xi_{\mathcal{P}_\gamma}(\boldsymbol{\mu})} \iff \varepsilon \leq \frac{\mu}{1+7\mu+9\mu^2}$$

which yields at best

$$\varepsilon \leq \frac{1}{13} \approx 0.0769$$

Example 2. The lattice gas on a bounded degree graph $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ with hard core self repulsion and hard core pair interaction and the triangular lattice on then plane

Let $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ be a bounded degree infinite graph with vertex set $\mathbb{V}$ and edge set $\mathbb{E}$, and maximum degree Δ. A polymer system is obtained by choosing $\mathcal{P} = \mathbb{V}$ and by defininig the incompatibility relation $\nsim$ by saying that two polymers γ and γ' (i.e. two vertices of $\mathbb{G}$) are incompatible if and only if either $\gamma = \gamma'$ or $\{\gamma, \gamma'\} \in \mathbb{E}$. This polymer gas realization is called the self repulsive hard core lattice gas on $\mathbb{G}$. In this case the polymers are the vertices of $\mathbb{G}$ and two polymers $\{x, y\} \subset \mathbb{V}$ are incompatible if either $y = x$ (self repulsion) or $\{x, y\} \in \mathbb{E}$ (hard core pair interaction). In general, since polymers have no structure (they are just vertices in a graph) one can suppose that the activity of a polymer $x \in \mathcal{P}$ is a constant, i.e. $\rho_x = \rho$ for all $x \in \mathbb{V}$. Of course expect that the convergence radius depends strongly on the topological structure of $\mathbb{G}$. We first consider the worst case i.e. when the graph $\mathbb{G}$ is such that the nearest neighbors of any vertex are pairwise compatible. This happens e.g. if $\mathbb{G}$ is a tree or if it is the cubic lattice $\mathbb{Z}^d$. The Kotecky-Preiss condition for this model then reads as

$$\rho_x \leq \mu_x e^{-\sum_{y \nsim x} \mu_y} \iff \rho \leq \mu e^{-(\Delta+1)\mu}$$

which yields at best

$$\rho \leq \frac{1}{(\Delta+1)e} \tag{4.75}$$

On the other hand the Dobrushin condition reads

$$\rho_x \leq \frac{\mu_x}{\prod_{y \nsim x}[1+\mu_y]} \iff \rho \leq \frac{\mu}{(1+\mu)^{\Delta+1}}$$

which yields at best

$$\rho \leq \frac{\frac{1}{\Delta}}{(1+\frac{1}{\Delta})^{\Delta+1}} = \frac{\Delta^\Delta}{(\Delta+1)^{\Delta+1}} = \frac{1}{\Delta+1}\frac{1}{(1+\frac{1}{\Delta})^\Delta} \tag{4.76}$$

Finally, the condition (4.52) gives

$$\rho_x \leq \frac{\mu_x}{\Xi_{\mathcal{P}_x}(\boldsymbol{\mu})} \iff \rho \leq \frac{\mu}{1+(\Delta+1)\mu+\sum_{k=2}^{\Delta}\binom{\Delta}{k}\mu^k} = \frac{\mu}{\mu+(1+\mu)^\Delta}$$

which yields at best

$$\rho \leq \frac{\frac{1}{\Delta-1}}{\frac{1}{\Delta-1}+\left(1+\frac{1}{\Delta-1}\right)^\Delta} = \frac{1}{1+\frac{\Delta^\Delta}{(\Delta-1)^{\Delta-1}}} = \frac{1}{\Delta(1+\frac{1}{\Delta-1})^{\Delta-1}+1} \tag{4.77}$$

To illustrate that Theorem 4.1 permits to improve this last bound (4.77) if we know more about the topological structure of $\mathbb{G}$, we now consider a case of the triangular lattice in $d = 2$ (a regular graph with degree $\Delta = 6$, where our

bound turns to be more efficient than the Dobrushin bound and the Shearer-Sokal bound (4.75). For the triangular lattice the tree bound (4.77) gives

$$\varepsilon < \frac{5^5}{6^6} \approx 0,067$$

while our bound gives

$$\varepsilon \leq \frac{c}{\Xi_\gamma(c)} = \frac{c}{1 + 7c + 9c^2 + 2c^3}$$

The maximum occurs when $4c^3 + 9c^2 - 1 = 0$, which is somewhere between $1/3$ and $3/10$. For example choosing $c = 1/3$ (which is not the best choice) we obtain

$$\varepsilon \leq \frac{c}{\Xi_\gamma(c)} = \frac{\frac{1}{3}}{1 + \frac{7}{3} + 1 + \frac{2}{27}} \approx 0,075$$

4.3 Gas of non overlapping finite subsets

In this section we will study a particular realization of the polymer gas which appears in the most part of the examples in statistical mechanics.
We will suppose that it is given an infinite countable set $\mathbb{V}$, and we define the space of polymers as

$$\mathcal{P}_\mathbb{V} = \{R \subset \mathbb{V} : |R| < \infty\}$$

and the incompatibility relation in $\mathcal{P}_\mathbb{V}$ is defined as

$$\gamma \not\sim \tilde{\gamma} \iff \gamma \cap \tilde{\gamma} \neq \emptyset$$

Note that now polymers have a cardinality, so that we can speak about big polymers and small polymers. Of course, as before, to a polymer γ is associated an activity $\zeta(\gamma)$. We assume in general that $\zeta(\gamma) \in \mathbb{C}$ as far as $\gamma \in \mathcal{P}_\mathbb{V}$ and we set

$$|\zeta(\gamma)| = \rho(\gamma) \tag{4.78}$$

Note that here we allow the value $\zeta(\gamma) = 0$ for some γ in order to stay more general. For example in the polymer expansion of high temperature spin systems, the polymer space is always $\mathcal{P}_\mathbb{V}$ for some suitable $\mathbb{V}$ but it happens that $\zeta(\gamma) = 0$ whenever $|\gamma| = 1$.

In most of the physics realizations $\mathbb{V}$ is the vertex set of an infinite graph $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ with edge set $\mathbb{E}$. For example $\mathbb{V} = \mathbb{Z}^d$ and $\mathbb{E}$ is the set of nearest neighbor in $\mathbb{Z}^d$. When $\mathbb{V}$ is the vertex set of a graph $\mathbb{G}$ then$\mathbb{V}$ has a natural metric structure induced by the graph distance in $\mathbb{G}$. This metric structure on $\mathbb{V}$ allow us to talk about how spread is a polymers (a polymer is spread if its points are far apart) and we can say now if two polymers γ and γ' are close or far apart. So, from the abstract context we can pass to more concrete realizations which have richer structures. Namely, if one suppose that polymers are finite subsets of an underlying countable set $\mathbb{V}$ with $\not\sim = \cap$, the any polymer has an activity and a cardinality, so we an distinguish between big and small polymers. If we further suppose that the underlying set $\mathbb{V}$ is the vertex set of

some graph $\mathbb{G}$ we also can talk about distance between polymers and spread polymers. In any case, in the whole section below we will not suppose any graph structure for the set $\mathbb{V}$. Our abstract polymer space is just the set of all finite subsets of a countable set with the incompatibility relation being the the non void intersection.

Let now Λ be a finite set of $\mathbb{V}$. A configuration of polymer gas in Λ is given once we specify the set of polymers which are present in Λ. Of course this polymers must be pairwise compatible, i.e. a configuration in Λ is an unordered n-ple $\{\gamma_1, \ldots, \gamma_n\}$ such that $\gamma_i \cap \gamma_j = \emptyset$ for all $i, j = 1, \ldots, n$. The "probability"[1] to see the configuration $\{\gamma_1, \ldots, \gamma_n\}$ in the box Λ is defined as

$$\mathrm{Prob}_{\boldsymbol{\zeta}}(\gamma_1, \ldots, \gamma_n) = \Xi_\Lambda^{-1} \prod_{i=1}^{n} \zeta(\gamma_i) \tag{4.79}$$

where Ξ_Λ is the partition function defined as

$$\Xi_\Lambda(\boldsymbol{\zeta}) = 1 + \sum_{n \geq 1} \sum_{\substack{\{\gamma_1, \ldots, \gamma_n\}:\ \gamma_i \subset \Lambda \\ \gamma_i \cap \gamma_j = \emptyset}} \zeta(\gamma_1) \ldots \zeta(\gamma_n) \tag{4.80}$$

4.3.1 Convergence via the abstract polymer criteria

We compare the three conditions for this model. Starting with the Kotecky-Preiss condition, choosing $\mu(\gamma) = \rho(\gamma) e^{a|\gamma|}$ (recall: $|\zeta(\gamma)| = \rho(\gamma)$), the condition (4.69) becomes the well known inequality

$$\sum_{\substack{\tilde{\gamma} \in \mathcal{P}_\mathbb{V} \\ \tilde{\gamma} \not\sim \gamma}} \rho(\tilde{\gamma})\, e^{a|\tilde{\gamma}|} \ \leq\ a|\gamma|, \qquad \forall \gamma \in \mathcal{P} \tag{4.81}$$

Now using that $\tilde{\gamma} \not\sim \gamma$ means for the present model $\tilde{\gamma} \cap \gamma \neq \emptyset$ we have that

$$\sum_{\tilde{\gamma} \not\sim \gamma} \rho(\tilde{\gamma}) e^{a|\tilde{\gamma}|} \leq |\gamma| \sup_{x \in \mathbb{V}} \sum_{\tilde{\gamma} \ni x} \rho(\tilde{\gamma}) e^{a|\tilde{\gamma}|}$$

Hence (4.81) becomes the well known condition

$$\sup_{x \in \mathbb{V}} \sum_{\substack{\gamma \in \mathcal{P}_\mathbb{V} \\ \gamma \ni x}} \rho(\gamma)\, e^{a|\gamma|} \ \leq\ a \tag{4.82}$$

On the other hand the Dobrushin condition (4.60) can be written as

$$\rho(\gamma) \leq \frac{c(\gamma)}{\prod_{\tilde{\gamma} \in \mathcal{P}_\mathbb{V}:\, \tilde{\gamma} \not\sim \gamma} [1 + c(\tilde{\gamma})]}$$

choosing again $c(\gamma) = |\rho(\gamma)| e^{a|\gamma|}$ the condition above becomes

$$\prod_{\substack{\tilde{\gamma} \in \mathcal{P}_\mathbb{V} \\ \tilde{\gamma} \not\sim \gamma}} (1 + \rho(\tilde{\gamma}) e^{a|\tilde{\gamma}|}) \leq e^{a|\gamma|}, \qquad \forall \gamma \in \mathcal{P}$$

[1] (4.79) is a real probability only if $\zeta(\gamma) \in [0, +\infty)$

i.e.

$$\sum_{\substack{\tilde\gamma\in\mathcal{P}_{\mathbb{V}}\\ \tilde\gamma\not\sim\gamma}} \log(1+\rho(\tilde\gamma)e^{a|\tilde\gamma|}) \le a|\gamma| \qquad \forall\gamma\in\mathcal{P} \tag{4.83}$$

i.e.

$$\sup_{x\in\mathbb{V}} \sum_{\substack{\gamma\in\mathcal{P}_{\mathbb{V}}\\ \gamma\ni x}} \log(1+\rho(\gamma)e^{a|\gamma|}) \le a \tag{4.84}$$

which is slightly better than (4.82).
Finally, the condition (4.52), putting again $c(\gamma)=\rho(\gamma)e^{a|\gamma|}$, becomes

$$\Xi_{\mathcal{P}_\gamma}(c) \le e^{a|\gamma|} \tag{4.85}$$

where $\mathcal{P}_\gamma=\{\gamma'\in\mathcal{P}_{\mathbb{V}} : \gamma'\cap\gamma\neq\emptyset\}$ and

$$\Xi_{\mathcal{P}_\gamma}(c) = 1+\sum_{n=1}^{|\gamma|}\frac{1}{n!}\sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{P}^n_{\mathbb{V}}\\ \gamma_i\not\sim\gamma,\gamma_i\sim\gamma_j}}\prod_{i=1}^{n}\rho(\gamma_i)e^{a|\gamma_i|}$$

We again use the fact that $\gamma_i\not\sim\gamma_j \iff \gamma_i\cap\gamma_j\neq\emptyset$ and $\gamma_i\sim\gamma_j\iff\gamma_i\cap\gamma_j=\emptyset$ to estimate the factor

$$\sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{P}^n_{\mathbb{V}}\\ \gamma_i\not\sim\gamma,\gamma_i\sim\gamma_j}}\prod_{i=1}^{n}\rho(\gamma_i)e^{a|\gamma_i|}$$

Note that this factor is zero whenever $n>|\gamma|$, since there is no way to choose n subsets γ_i of $\mathbb{V}$ such that they are all pairwise compatible (i.e. non intersecting) and all incompatible (i.e. intersecting) with a fixed subset γ of $\mathbb{V}$ with a number of elements equal to $|\gamma|$. On the other hand, when the sum above is not zero, i.e. for $n\le|\gamma|$, it can be bounded at least by (a very rough bound)

$$\sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{P}^n\\ \gamma_i\not\sim\gamma,\gamma_i\sim\gamma_j}}\prod_{i=1}^{n}\rho(\gamma_i)e^{a|\gamma_i|} \le |\gamma|(|\gamma|-1)\cdots(|\gamma|-n+1)\left[\sup_{x\in\mathbb{V}}\sum_{\substack{\gamma\in\mathcal{P}_{\mathbb{V}}\\ x\in\gamma}}\rho(\gamma)e^{a|\gamma|}\right]^n =$$

$$= \binom{|\gamma|}{n}n!\left[\sup_{x\in\mathbb{V}}\sum_{\substack{\gamma\in\mathcal{P}_{\mathbb{V}}\\ x\in\gamma}}\rho(\gamma)e^{a|\gamma|}\right]^n$$

Thus

$$\Xi^\gamma_{\mathcal{P}_{\mathbb{V}}}(c) \le 1+\sum_{n=1}^{|\gamma|}\binom{|\gamma|}{n}\left[\sup_{x\in\mathbb{V}}\sum_{\substack{\gamma\in\mathcal{P}_{\mathbb{V}}\\ x\in\gamma}}\rho(\gamma)e^{a|\gamma|}\right]^n = \left[1+\sup_{x\in\mathbb{V}}\sum_{\substack{\gamma\in\mathcal{P}_{\mathbb{V}}\\ x\in\gamma}}\rho(\gamma)e^{a|\gamma|}\right]^{|\gamma|}$$

Thus (4.85) can be written as

$$\left[1+\sup_{x\in\mathbb{V}}\sum_{\substack{\gamma\in\mathcal{P}_{\mathbb{V}}\\ x\in\gamma}}\rho(\gamma)e^{a|\gamma|}\right]^{|\gamma|} \le e^{a|\gamma|}$$

i.e.

$$\sup_{x\in\mathbb{V}} \sum_{\substack{\gamma\in\mathcal{P}_{\mathbb{V}}\\ x\in\gamma}} \rho(\gamma)e^{a|\gamma|} \leq e^a - 1 \tag{4.86}$$

Note that by (4.53) we also get the upper bound

$$\Pi_\gamma(\rho) \leq e^{a|\gamma|}$$

We have therefore proved the following theorem

Theorem 4.2 *Let $\mathbb{V}$ be a countable set and let $\mathcal{P}_{\mathbb{V}} = \{\gamma \subset \mathbb{V} : |\gamma| < \infty\}$ be the polymer set with incompatibility relation: $\gamma \not\sim \gamma' \Leftrightarrow \gamma \cap \gamma' \neq \emptyset$ and (complex) activity $\zeta(\gamma)$. Assume that there is a positive number $a > 0$, such that, for all $x \in \mathbb{V}$*

$$\sum_{\substack{\gamma\in\mathcal{P}_{\mathbb{V}}\\ x\in\gamma}} |\zeta(\gamma)|\ e^{a|\gamma|} \leq e^a - 1 \tag{4.87}$$

then, for all finite $\Lambda \subset \mathbb{V}$, we have that $\frac{1}{|\Lambda|}\ln \Xi_\Lambda(\boldsymbol{\zeta})$, where $\Xi_\Lambda(\boldsymbol{\zeta})$ is the partition function define in (4.80), can be written as an absolutely convergent series uniformly in Λ.

Chapter 5

Two systems in the cubic lattice

To show the utility of the polymer expansion in the study os discrete systems, we will consider in this chapter two systems on the d-dimensional cubic lattice $\mathbb{Z}^d$.

5.1 Self repuslive Lattice gas

A lattice gas in $\mathbb{Z}^d$ is a simple case of a polymer system. Namely, a lattice is a polymers gas in which the polymer space $\mathcal{P}$ are "particles" which can occupy the vertices x of the unit cubic lattice $\mathbb{Z}^d$. In a lattice gas the activity of a particle occupying the site x is assumed to be constant λ, i.e. $z_x = \lambda$ for all $x \in \mathbb{Z}^d$. Finally, the pair potential of a lattice gas is assumed to be symmetric $V(x, y) = V(y, x)$ and self repulsive, i.e. only one particle can occupy a site $x \in \mathbb{Z}^d$. Namely:

$$V(x, x) = +\infty \qquad \text{for all } x \in \mathbb{Z}^d \tag{5.1}$$

We will also assume that the pair potential is summable in the following sense

$$\sup_{x \in \mathbb{Z}^d} \sum_{\substack{y \in \mathbb{Z}^d \\ y \neq x}} |V(x, y)| = C < \infty \tag{5.2}$$

The assumptions (5.1) and (5.2) automatically guarantee that the pair potential is stable, i.e. for any $n \in \mathbb{N}$ and $(x_1, \ldots, x_n) \in (\mathbb{Z}^d)^n$,

$$\sum_{1 \leq i < j \leq n} V(x_i, x_j) \geq -nC, \tag{5.3}$$

5.1.1 Covergence by direct Mayer expansion

We can obtain the convergent condition for the high tempretaure/low activity phase of such lattice gas by directly perform a Mayer expansion of the grand canonical partition function. Let us sketch how. The Grand Canonical partition

function is

$$Z_\Lambda(\beta,\lambda) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \sum_{(x_1,\dots,x_n)\in\Lambda^n} e^{-\beta\sum_{1\le i<j\le n} V(x_i,x_j)}$$

where of course we are assuming in that $V(x,x) = +\infty$. By Mayer expansion on the factor $e^{-\beta\sum_{i<j} V(x_i,x_j)}$ we find, as usual

$$\ln Z_\Lambda(\beta,\lambda) \;=\; \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} \sum_{(x_1,\dots,x_n)\in\Lambda^n} \sum_{g\in G_n} \prod_{\{i,j\}\in E_g} (e^{-\beta V(x_i,x_j)} - 1)$$

Then, by Theorem 3.7 and tree graph inequality (3.63), we have, using stability also condition (5.3),

$$|\sum_{g\in G_n} \prod_{\{i,j\}\in E_g} (e^{-\beta V(x_i,x_j)} - 1)| \le e^{\beta n C} \sum_{\tau\in T_n} \prod_{\{i,j\}\in E_\tau} (1 - e^{-\beta|V(x_i,x_j)|})$$

hence

$$|\ln Z_\Lambda(\beta,\lambda)| \le \sum_{n\,=\,1}^{\infty} \frac{|\lambda|^n}{n!} \sum_{(x_1,\dots,x_n)\in\Lambda^n} e^{n\beta C} \sum_{\tau\in T_n} \prod_{\{i,j\}\in E_\tau} (1 - e^{-\beta|V(x_i,x_j)|}) \le$$

$$\le \sum_{n\,=\,1}^{\infty} \frac{|\lambda|^n}{n!}\, e^{n\beta C} \sum_{\tau\in T_n} \sum_{(x_1,\dots,x_n)\in\Lambda^n} \prod_{\{i,j\}\in E_\tau} (1 - e^{-\beta|V(x_i,x_j)|})$$

but similarly to what we saw in Proposition 3.1, formula (3.17) it is not difficult to show that

$$\sum_{(x_1,\dots,x_n)\in\Lambda^n} \prod_{\{i,j\}\in E_\tau} (1 - e^{-\beta|V(x_i,x_j)|}) \le |\Lambda| \left[\sup_{x\in\mathbb{Z}^d} \sum_{y\in\mathbb{Z}^d} (1 - e^{-\beta|V(x,y)|})\right]^{n-1}$$

note that the sum over y after the sup includes also $y = x$ where $V(x,x) \;=\; \infty$. Therefore defining

$$C(\beta) \;=\; \sup_{x\in\mathbb{Z}^d} \sum_{\substack{y\in\mathbb{Z}^d \\ y\neq x}} (1 - e^{-\beta|V(x,y)|})$$

we get

$$\sum_{(x_1,\dots,x_n)\in\Lambda^n} \prod_{\{i,j\}\in E_\tau} (1 - e^{-\beta|V(x_i,x_j)|}) \le |\Lambda|\,[1 + C(\beta)]^{n-1}$$

and therefore

$$|\ln Z_\Lambda(\beta,\lambda)| \le \sum_{n\,=\,1}^{\infty} \frac{|\lambda|^n}{n!}\, e^{n\beta C} |\Lambda| [1 + C(\beta)]^{n-1} n^{n-2}$$

thus in this case the condition for the convergence is

$$|\lambda| e^{\beta C+1} [1 + C(\beta)] < 1 \tag{5.4}$$

This condition is quite unsatisfactory since it says that, for any temperature, even very high (i.e even for β very small when $C(\beta)$ is very small), one needs to set the activity smaller than $1/e$ to ensure convergence!

5.1.2 Convergence via polymer expansion

We now obtain the convergence condition for the same $\ln Z_\Lambda(\beta,\lambda)$ performing a first example of high temperature polymer expansion. We will get a much satisfactory bound and this will illustrate quite well how convenient can be to perform polymer expansion in discrete systems when possible!
The *Grand canonical partition function* of the same lattice gas enclosed in $\Lambda \subset \mathbb{Z}^d$, with activity λ, inverse temperature β, interacting via a pair potential $V(x,y)$ (x and y sites in Λ) such that $V(x,x) = \infty$ (a site x can be occupied can be written as follows

$$Z_\Lambda(\beta,\lambda) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \sum_{(x_1,\dots,x_n)\in\Lambda^n} e^{-\beta\sum_{i<j} V(x_i,x_j)} =$$

$$= \sum_{n=0}^{|\Lambda|} \frac{\lambda^n}{n!} \sum_{\substack{(x_1,\dots,x_n)\in\Lambda^n \\ x_i\neq x_j}} e^{-\beta\sum_{i<j} V(x_i,x_j)} =$$

$$= \sum_{n=0}^{|\Lambda|} \lambda^n \sum_{\{x_1,\dots,x_n\}\subset\Lambda} e^{-\beta\sum_{i<j} V(x_i,x_j)} =$$

$$= \sum_{S\subset\Lambda} \lambda^{|S|}\, e^{-\beta\sum_{\{x,y\}\subset S} V(x,y)}$$

Define then for $x\in\Lambda$ the variable n_x taking values in the set $\{0,1\}$ (n_x can be interpreted as the occupation number of the site x: $n_x = 0$ means that the site is empty and $n_x = 1$ means that the site is occupied). We denote by n_Λ a possible configuration of occupied numbers in Λ and by N_Λ the set of all possible configurations n_Λ (and in general, if $R\subset\mathbb{Z}^d$, we denote by n_R a possible configuration of occupied numbers in R and by N_R the set of all possible configurations n_R). Clearly there is a one to one correspondence between $n_\Lambda\in N_\Lambda$ and $S\subset\Lambda$ by defining $S(n_\Lambda) = \{x\in\Lambda : n_x = 1\}$. Therefore we can write

$$Z_\Lambda(\beta,\lambda) = \sum_{S\subset\Lambda} \lambda^{|S|}\, e^{-\beta\sum_{\{x,y\}\subset S} V(x,y)} =$$

$$= \sum_{n_\Lambda\in N_\Lambda} \lambda^{\sum_{x\in\Lambda} n_x} e^{-\beta\sum_{\{x,y\}\subset\Lambda} n_x n_y V(x,y)} \tag{5.5}$$

Expand now the exponential in (5.5)

$$e^{-\beta\sum_{\{x,y\}\subset\Lambda} n_x n_y V(x,y)} = \prod_{\{x,y\}\subset\Lambda} [e^{-\beta n_x n_y V(x,y)} - 1 + 1] =$$

$$= \sum_{s=1}^{|\Lambda|} \sum_{\{R_1,\dots,R_s\}\in\pi(\Lambda)} \rho(R_1)\cdots\rho(R_r)$$

where $\pi(\Lambda)$ = set of all partitions of Λ (so that $\{R_1, \ldots, R_s\}$ is a partition of Λ), and

$$\rho(R) = \begin{cases} 1 & \text{if } |R| = 1 \\ \sum_{g \in G_R} \prod_{\{x,y\} \in E_g} [e^{-\beta n_x n_y V(x,y)} - 1] & \text{if } |R| \geq 2 \end{cases}$$

where G_R is the set of connected graph with vertex set R.
Thus (5.5) can be written as,

$$Z_\Lambda(\beta, \lambda) = \sum_{n_\Lambda \in N_\Lambda} \lambda^{\sum_{x \in \Lambda} n_x} \sum_{s=1}^{|\Lambda|} \sum_{R_1, \ldots, R_s \in \pi(\Lambda)} \rho(R_1) \cdots \rho(R_s) =$$

$$= \sum_{s=1}^{|\Lambda|} \sum_{R_1, \ldots, R_s \in \pi(\Lambda)} \left[\sum_{n_{R_1} \in N_{R_1}} \rho(R_1) \lambda^{\sum_{x \in R_1} n_x} \right] \cdots \left[\sum_{n_{R_s} \in N_{R_s}} \rho(R_s) \lambda^{\sum_{x \in R_s} n_x} \right]$$

Define now

$$\tilde{\rho}(R) = \sum_{n_R \in N_R} \rho(R) \, \lambda^{\sum_{x \in R} n_x}$$

Since, for any $g \in G_R$,

$$\prod_{\{x,y\} \in E_g} [e^{-\beta n_x n_y V(x,y)} - 1] \neq 0$$

if $n_x = 1$ for all $x \in R$, then

$$\tilde{\rho}(R) = \begin{cases} 1 + \lambda & \text{if } |R| = 1 \\ \lambda^{|R|} \sum_{g \in G_R} \prod_{\{x,y\} \in E_g} [e^{-\beta V(x,y)} - 1] & \text{if } |R| \geq 2 \end{cases}$$

Define now

$$\zeta(R) = \begin{cases} 1 & \text{if } |R| = 1 \\ \frac{\lambda^{|R|}}{(1+\lambda)^{|R|}} \sum_{g \in G_R} \prod_{\{x,y\} \in E_g} [e^{-\beta V(x,y)} - 1] & \text{if } |R| \geq 2 \end{cases} \tag{5.6}$$

and obtain

$$Z_\Lambda(\beta, \lambda) = (1+\lambda)^{|\Lambda|} \sum_{s=1}^{|\Lambda|} \sum_{R_1, \ldots, R_s \in \pi(\Lambda)} \zeta(R_1) \cdots \zeta(R_s)$$

$$= (1+\lambda)^{|\Lambda|} \sum_{s \geq 0} \sum_{\substack{\{R_1, \ldots, R_s\} \subset \Lambda \\ |R_i| \geq 2, \ R_i \cap R_j = \emptyset}} \zeta(R_1) \cdots \zeta(R_s)$$

where the term $s = 0$ in the last sum is equal to 1 and corresponds to the partition of Λ in $|\Lambda|$ subsets each of cardinality 1. Clearly

$$\Xi_\Lambda(\beta,\lambda) \doteq \sum_{s\geq 0} \sum_{\substack{\{R_1,\dots,R_s\}\subset\Lambda \\ |R_i|\geq 2,\ R_i\cap R_j=\emptyset}} \zeta(R_1)\cdots\zeta(R_s) =$$

$$= \ 1+\sum_{n\geq 1}\frac{1}{n!} \sum_{\substack{R_1,\dots,R_n \\ |R_i|\geq 2,\, R_i\cap R_j \,=\, \emptyset}} \zeta(R_1)\cdots\zeta(R_n)$$

I.e. $\Xi_\Lambda(\beta,\lambda)$ is the grand canonical partition function of a polymer gas in which polymers are finite subsets of $R\subset\mathbb{Z}^d$ such that $|R|\geq 2$ and with activity $\zeta(R_1)$. We finally have

$$Z_\Lambda(\beta,\lambda) = (1+\lambda)^{|\Lambda|}\Xi_\Lambda(\beta,\lambda)$$

and hence

$$\ln Z_\Lambda(\beta,\lambda) = |\Lambda|\ln(1+\lambda) + \ln\Xi_\Lambda(\beta,\lambda)$$

Hence the pressure of the lattice gas $\frac{1}{|\Lambda|}\log Z_\Lambda(\beta,\lambda)$ converges absolutely uniformly in Λ if $\frac{1}{|\Lambda|}\log\Xi_\Lambda(\beta,\lambda)$ does.
I.e., by (4.86) conditions if,

$$\sum_{n\geq 2} e^{an}\sup_{x\in\mathbb{Z}^d}\sum_{\substack{R\subset\mathbb{Z}^d: \\ x\in R,\ |R|=n}}|\zeta(R)|\leq e^a-1$$

or a little bit "roughly" (i.e choosing $a=\ln 2$ which is not an optimal choice)

$$\sum_{n\geq 2} 2^n \sup_{x\in\mathbb{Z}^d}\sum_{\substack{R\subset\mathbb{Z}^d: \\ x\in R,\ |R|=n}}|\zeta(R)| \ \leq \ 1 \tag{5.7}$$

Recalling definition (5.6) of the activity $\zeta(R)$ and setting $\tilde\lambda = \frac{\lambda}{(1+\lambda)}$ we have

$$\sup_{x\in Z^d}\sum_{\substack{R\subset\mathbb{Z}^d: \\ x\in R,\ |R|=n}}|\zeta(R)| \ = \ |\tilde\lambda|^n \sup_{x\in Z^d}\sum_{\substack{R\subset\mathbb{Z}^d: \\ x\in R,\ |R|=n}}\Big|\sum_{g\in G_R}\prod_{\{x,y\}\in E_g}[e^{-\beta V(x,y)}-1]\Big|$$

Now note that

$$\sum_{\substack{R\subset\mathbb{Z}^d: \\ x\in R,\ |R|=n}} \ = \ \frac{1}{(n-1)!}\sum_{\substack{(x_1,\dots,x_n)\in\mathbb{Z}^{dn} \\ x_1=x,\ x_i\neq x_j \text{ for } i\neq j}}$$

Hence

$$\sup_{x\in\mathbb{Z}^d}\sum_{\substack{R\subset\mathbb{Z}^d: \\ x\in R,\ |R|=n}}|\zeta(R)| \ = \ \frac{|\tilde\lambda|^n}{(n-1)!}\sup_{x\in\mathbb{Z}^d}\sum_{\substack{(x_1,\dots,x_n)\in\mathbb{Z}^{dn} \\ x_1=x,\ x_i\neq x_j \text{ for } i\neq j}}\Big|\sum_{g\in G_n}\prod_{\{i,j\}\in E_g}[e^{-\beta V(x_i,x_j)}-1]\Big| \tag{5.8}$$

Once again, by Theorem 3.7, tree graph inequality (3.63) and using stability condition (5.3) we have the estimate

$$\Big| \sum_{g\in G_n} \prod_{\{i,j\}\in E_g} [e^{-\beta V(x_i,x_j)} - 1] \Big| \le e^{\beta n C} \sum_{\tau\in T_n} \prod_{\{i,j\}\in E_\tau} (1 - e^{-\beta|V(x_i,x_j)|})$$

Now observe that, for any $\tau \in T_n$

$$\sum_{\substack{(x_1,\dots,x_n)\in\mathbb{Z}^{dn} \\ x_1=x,\ x_i\neq x_j \text{ for } i\neq j}} \prod_{\{i,j\}\in E_\tau} (1-e^{-\beta|V(x_i,x_j)|}) \ \le \ \left[\sup_{x\in\mathbb{Z}^d} \sum_{\substack{y\in\mathbb{Z}^d \\ y\neq x}} (1 - e^{-\beta|V(x,y)|}) \right]^{n-1} =$$

$$= \ [C(\beta)]^{n-1}$$

Therefore, recalling also Cayley formula (3.6), we get

$$\sup_{x\in Z^d} \sum_{\substack{R\subset\mathbb{Z}^d: \\ x\in R,\ |R|=n}} |\zeta(R)| \ \le \ \frac{n^{n-2}}{(n-1)!}[C(\beta)]^{n-1}(|\tilde\lambda|e^{\beta C})^n \ \le \ |\tilde\lambda|e^{\beta C}\Big[|\tilde\lambda|e^{\beta C+1}\, C(\beta)\Big]^{n-1}$$

Hence the condition for convergence is by (5.7)

$$2|\tilde\lambda|e^{\beta C}\sum_{n\ge 2}\Big[2|\tilde\lambda|e^{\beta C+1}\, C(\beta)\Big]^{n-1} \ \le \ 1 \tag{5.9}$$

I.e., after some calculus

$$|\tilde\lambda|e^{\beta C+1} \le \frac{1}{C(\beta)}\frac{1}{1+\sqrt{1+\frac{4}{eC(\beta)}}} \tag{5.10}$$

This is a much better condition than (5.4). Indeed, recalling that $\tilde\lambda = \lambda/(1+\lambda)$, observe that for $\beta < \beta_c$ where β_c is the solution of the equation

$$e^{\beta C+1} \ = \ \frac{1}{C(\beta)}\frac{1}{1+\sqrt{1+\frac{4}{eC(\beta)}}}$$

the pressure of the lattice gas is an absolute convergent expansion for all λ real. The temperature β_c is a first example of critical temperature: below β_c the lattice gas is in a pure phase for any activity $\lambda > 0$.

5.2 Ising model

The Ising model is a lattice system enclosed in a box $\Lambda \subset \mathbb{Z}^d$. The box Λ is a finite set, generally a square of size L which contains $|\Lambda| \ = \ L^d$ sites of the lattice $\mathbb{Z}^d$. In each site $x \in \Lambda$ there is a random variable σ_x which can take one of the two possible value $\sigma_x \ = \ \pm 1$. A configuration σ_Λ of the system is given when one declares the value of the spin σ_x for each site $x \in \Lambda$. Hence σ_Λ is a

set $\sigma_\Lambda = (\sigma_{x_1}, \ldots, \sigma_{x_{|\Lambda|}})$ of $|\Lambda|$ numbers ± 1. Equivalently one can say that a configuration σ_Λ of the system is a function $\sigma_\Lambda : \Lambda \to \{+1, -1\} : x \mapsto \sigma_x$. Observe that the total number of configurations of the system in Λ is $2^{|\Lambda|}$. We denote by Ω_Λ the set of all possible spin configurations in Λ.

Given a configurations σ_Λ of the system in Λ, the energy of such configuration is

$$H_\Lambda(\sigma_\Lambda) = -J \sum_{\substack{|x-y|=1 \\ \{x,y\}\subset\Lambda}} \sigma_x\sigma_y - h\sum_{x\in\Lambda} \sigma_x + \mathcal{B}(\sigma_\Lambda) \tag{5.11}$$

where h is an "external magnetic field", $J > 0$ is a positive constant and $\mathcal{B}(\sigma_\Lambda)$ represent the interaction of the spins inside the box Λ with the world outside. Of course $\mathcal{B}(\sigma_\Lambda)$ is rather arbitrary. We will list in a moment some typical boundary conditions. Let us just remark here that $\mathcal{B}(\sigma_\Lambda)$ has to be in any case a "surface term", i.e.

$$\lim_{\Lambda\to\infty} \frac{\max_{\sigma_\Lambda} |\mathcal{B}(\sigma_\Lambda)|}{|\Lambda|} = 0 \tag{5.12}$$

The statistical mechanics is obtained by assigning to any configuration σ_Λ a probability to occur $P(\sigma_\Lambda)$. This probability is given by

$$P(\sigma_\Lambda) = \frac{e^{-\beta H_\Lambda(\sigma_\Lambda)}}{Z_\Lambda(\beta, h)} \tag{5.13}$$

where

$$Z_\Lambda(\beta, h) = \sum_{\sigma_\Lambda\in\Omega_\Lambda} e^{-\beta H_\Lambda(\sigma_\Lambda)}$$

is the partition function in the grand canonical ensemble. If $f(\sigma_\Lambda)$ is a function depending on the configuration of the system, its mean value in the grand canonical ensemble is

$$\langle f(\sigma_\Lambda)\rangle_\Lambda = \sum_{\sigma_\Lambda\in\Omega_\Lambda} P(\sigma_\Lambda) f(\sigma_\Lambda) = \sum_{\sigma_\Lambda\in\Omega_\Lambda} \frac{e^{-\beta H_\Lambda(\sigma_\Lambda)} f(\sigma_\Lambda)}{Z_\Lambda(\beta, h)}$$

The function

$$f_\Lambda(\beta, h) = \frac{1}{|\Lambda|} \ln Z_\Lambda(\beta, h) \tag{5.14}$$

is called finite volume free energy. Note that this function is analytic in β for all $\beta > 0$ and in h for all $h \in (-\infty, +\infty)$.
Thermodynamic is recovered evaluating the following limit

$$f(\beta, h) = \lim_{\Lambda\to\infty} \frac{1}{|\Lambda|} \ln Z_\Lambda(\beta, h) \tag{5.15}$$

The function $f(\beta, h)$ is called the free energy of the system. It is easy to show that the limit exists, but in general is not expected to be analytic in the whole physical region. It is also worth to stress that (5.12) ensures that the limit above in independent on the boundary conditions $\mathcal{B}$. As a matter of fact, let $\mathcal{B}(\sigma_\Lambda)$ and $\mathcal{B}'(\sigma_\Lambda)$ two different boundary conditions and let us denote

$$H_\Lambda^{\text{bulk}}(\sigma_\Lambda) = -J \sum_{\substack{|x-y|=1 \\ \{x,y\}\subset\Lambda}} \sigma_x\sigma_y - h\sum_{x\in\Lambda} \sigma_x$$

then we have

$$\frac{Z_{\Lambda,\mathcal{B}}(\beta,h)}{Z_{\Lambda,\mathcal{B}'}(\beta,h)} \;=\; \frac{\sum_{\sigma_\Lambda} e^{-\beta H_\Lambda^{\rm bulk}(\sigma_\Lambda)} e^{-\beta \mathcal{B}(\sigma_\Lambda)}}{\sum_{\sigma_\Lambda} e^{-\beta H_\Lambda^{\rm bulk}(\sigma_\Lambda)} e^{-\beta \mathcal{B}'(\sigma_\Lambda)}} \leq \max_{\sigma_\Lambda} e^{\beta|\mathcal{B}(\sigma_\Lambda)|+\beta|\mathcal{B}'(\sigma_\Lambda)|}$$

i.e

$$\frac{Z_{\Lambda,\mathcal{B}}(\beta,h)}{Z_{\Lambda,\mathcal{B}'}(\beta,h)} \leq e^{\beta \max_{\sigma_\Lambda}(|\mathcal{B}(\sigma_\Lambda)|+|\mathcal{B}'(\sigma_\Lambda)|)}$$

analogously one can get

$$e^{-\beta \max_{\sigma_\Lambda}(|\mathcal{B}(\sigma_\Lambda)|+|\mathcal{B}'(\sigma_\Lambda)|)} \leq \frac{Z_{\Lambda,\mathcal{B}}(\beta,h)}{Z_{\Lambda,\mathcal{B}'}(\beta,h)}$$

whence

$$\frac{1}{|\Lambda|}|\ln Z_{\Lambda,\mathcal{B}}(\beta,h) - \ln Z_{\Lambda,\mathcal{B}'}(\beta,h)| \leq \frac{1}{|\Lambda|}\beta \max_{\sigma_\Lambda}(|\mathcal{B}(\sigma_\Lambda)| + |\mathcal{B}'(\sigma_\Lambda)|) \qquad (5.16)$$

and taking the limit $\Lambda \to \infty$ and using (5.12) we get the result.
We list below some typical boundary conditions.
1) *Open (or free) boundary conditions.* This is the case

$$\mathcal{B}_0(\sigma_\Lambda) \;=\; 0$$

2) *Periodic boundary conditions.* This is the case in which Λ is a Torus, i.e. spins of opposite faces interact via the constant $-J$, as if they were nearest neighbors. Clearly this can be obtained by a suitable choice of $\mathcal{B}_p(\sigma_\Lambda)$

3) $+$ *boundary conditions.* This is the following case. Let $\partial\Lambda = \{x \in \mathbb{Z}^2 \backslash \Lambda : |x-y| = 1 \text{ for some } y \in \Lambda\}$ be the "external boundary" of Λ and put $\sigma_x \;=\; +1$ for all $x \in \partial\Lambda$. Then

$$B_+(\sigma_\Lambda) \;=\; -J \sum_{x\in\partial\Lambda} \sum_{\substack{y\in\Lambda \\ |x-y|=1}} \sigma_y$$

Physically this means to fix the spins outside Λ (those who can interact with spin inside!) to the value $\sigma_x = +1$.

The fact that the free energy is independent from boundary conditions does not mean that the system is stable respect to boundary conditions. Precisely the instability of the system respect to boundary conditions is an evidence for phase transition. Even if the free energy is independent of boundary conditions it still can occur a discontinuity of some derivative of the free energy at some point in the (β, h) region of the physical parameters. We will show below that this indeed happens.

We will consider ahead the magnetization of the system, which measures if spin are mostly oriented up or down, is the partial derivative of the free energy respect to the magnetic field.

$$M_\Lambda^{\mathcal{B}}(\beta,h) \;=\; \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\sigma_x\rangle_{\Lambda,\mathcal{B}} \;=\; \beta^{-1}\frac{\partial}{\partial h} f_\Lambda(\beta,h)$$

We will see that the magnetization in the thermodynamic limit can take different values according to different boundary conditions.

5.2.1 High temperature expansion

We now develop a (high temperature) polymer expansion for the partition function of the two dimensional Ising model with zero magnetic field (i.e. we set $h = 0$ in (5.11)). We suppose from now on that Λ is a square of size L which thus contains L^2 lattice sites. We also choose free boundary conditions (i.e. we set $\mathcal{B}(\sigma_\Lambda) = 0$ in (5.11)). The Hamiltonian of the zero field, free boundary conditions Ising model is thus

$$H_\Lambda(\sigma_\Lambda) \;=\; -J \sum_{\substack{|x-y|=1 \\ \{x,y\}\subset\Lambda}} \sigma_x\sigma_y \;=\; -J \sum_{b\in B(\Lambda)} \tilde\sigma_b$$

where $B(\Lambda)$ is the set of all pairs $b = \{x, y\} \subset \Lambda$ such that $|x-y| \;=\; 1$ (nearest neighbor pairs) and for $b = \{x, y\}$ we put $\tilde\sigma_b = \sigma_x\sigma_y$. Recall that, since Λ is a $L \times L$ box in $\mathbb{Z}^2$, then $|B(\Lambda)| \;=\; 2L(L-1)$
The partition function of the Ising model at zero magnetic field is thus

$$Z_\Lambda(\beta) \;=\; \sum_{\sigma_\Lambda\in\Omega_\Lambda} \prod_{b\in B(\Lambda)} e^{+\beta J\tilde\sigma_b}$$

We want to prove that the free energy

$$f_\Lambda(\beta) = \frac{1}{|\Lambda|} \ln Z_\Lambda(\beta) \tag{5.17}$$

is an analytic function of β is β is sufficiently small (i.e. in the high temperature regime). Observe that $\tilde\sigma_b \;=\; \pm 1$, hence

$$e^{+\beta J\tilde\sigma_b} \;=\; \cosh(\beta J\tilde\sigma_b) + \sinh(\beta J\tilde\sigma_b) \;=\; \cosh(\beta J) + \tilde\sigma_b \sinh(\beta J) \;=$$

$$=\; \cosh(\beta J)[1 + \tilde\sigma_b \tanh(\beta J)]$$

hence

$$Z_\Lambda(\beta) \;=\; [\cosh(\beta J)]^{2L(L-1)} \sum_{\sigma_\Lambda\in\Omega_\Lambda} \prod_{b\in B(\Lambda)} [1 + \tilde\sigma_b \tanh(\beta J)]$$

Developing the product $\prod_{b\in B(\Lambda)}[1 + \tilde\sigma_b \tanh(\beta J)]$ we get terms of the type

$$[\tanh(\beta J)]^k \tilde\sigma_{b_1} \dots \tilde\sigma_{b_k}$$

which has a clear geometric interpretation. The set of bonds $b_1, \dots, b_k$ form a graph (connected or not) in Λ whose links are nearest neighbors.
When one perform the sum over σ_Λ we get that

$$\sum_{\sigma_\Lambda} \tilde\sigma_{b_1} \dots \tilde\sigma_{b_k}$$

is zero whenever there is a not paired spin. See Figures 11 and 12.
The only graphs which yield a non vanishing contribution to the partition function are those whose vertices have incidence number two or four (Figure 11),

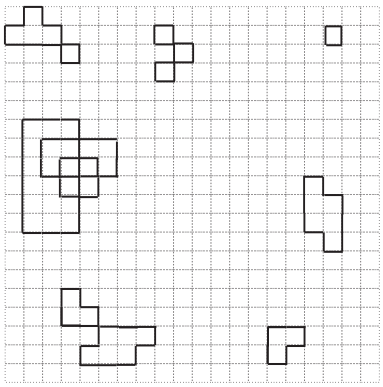


Figure 11. A non vanishing graph with seven connected components

while all other graphs are zero once the sum over configurations σ_Λ has been bone (Figure 12). If the graph $\tilde{\sigma}_{b_1} \dots \tilde{\sigma}_{b_k}$ is non vanishing then

$$\sum_{\sigma_\Lambda} \tilde{\sigma}_{b_1} \dots \tilde{\sigma}_{b_k} \;=\; 2^{L^2}$$

We can naturally split a non vanishing graph in non intersecting connected components which we will call *lattice animals*. For example in figure 11 it is drawn a non vanishing graph formed by seven non intersecting lattice animals. A lattice animal γ is thus nothing but a graph g with edge set $E_\gamma = \{b_1, \dots, b_k\} \subset B(\Lambda)$ formed by nearest neighbor links $b = \{x, y\}$ and with vertex set $V_g = \cup_{i=1}^k b_i \subset \Lambda$ which is a connected graph in V_γ (in usual sense). The allowed lattice animals are only those γ with incidence number at the vertices equal to two or four. Let us denote by $\mathcal{L}$ the set of possible lattice animals in $\mathbb{Z}^d$ and by $\mathcal{L}_\Lambda$ the set of all possible lattice animals in Λ.

Two lattice animals γ and γ' are non overlapping (i.e. compatible), and we write $\gamma \sim \gamma'$ *if and only if* $V_\gamma \cap V'_{\gamma'} = \emptyset$. We will denote shortly $|\gamma| = |E_\gamma|$ (i.e. $|\gamma|$ is the number of nearest neighbor bonds which constitute γ, i.e. if $\gamma = \{b_1, \dots, b_k\}$ then $|\gamma| = k$. Note that only such lattice animals (i.e. just those with incidence number at the vertices equal to 2 or to 4) survive because we are using free boundary conditions. Note also that lattice animal $c \in \mathcal{L}$ with incidence number equal to 2 in anyone of its vertices is a simple cycle. For a cycle c we have that $|V_c| = |c|$, while for a lattice lattice animal γ which is not a cycle we have $|V_\gamma| < |\gamma|$.

In conclusion we can write

$$Z_\Lambda(\beta) \;=\; [\cosh(\beta J)]^{2L(L-1)} 2^{L^2}\, \Xi_\Lambda(\beta) \tag{5.18}$$

where

$$\Xi_\Lambda(\beta) \;=\; 1 + \sum_{n \geq 1} \sum_{\substack{\{\gamma_1, \dots \gamma_n\} \subset \mathcal{L}_\Lambda \\ |\gamma_i| \geq 4,\ \gamma_i \sim \gamma_j}} \xi(\gamma_1) \dots \xi(\gamma_n) \tag{5.19}$$

where γ denote a allowed lattice animal (whence the condition $|\gamma| \geq 4$) with activity

$$\xi(\gamma) \;=\; [\tanh(\beta J)]^{|\gamma|} \tag{5.20}$$

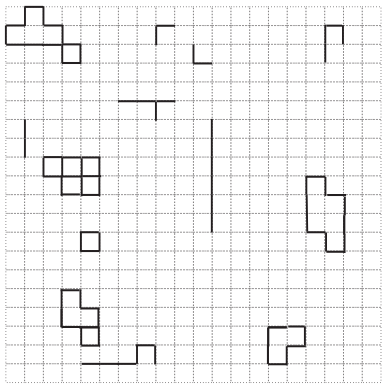


Figure 12. A vanishing graph.

Thus the partition function of the Ising model (5.18) can be written, modulo a constant (the factor $[\cosh(\beta J)]^{2L(L-1)}2^{L^2}$), as the partition function of a hard core polymer gas, i.e the term (5.19). In this case polymers are lattice animals, i.e. elements of $\mathcal{L}$ with the incompatibility relation $\gamma \nsim \gamma'$ if and only if $V_\gamma \cap V_{\gamma'} \neq \emptyset$.
Let us apply to this polymer gas the convergence criterion (4.52) of Theorem 4.1 to the polymers system with partition function $\Xi_\Lambda(\beta)$ given by (5.19) above. Namely we need to find numbers $\mu(\gamma) \in (0, +\infty)$ such that

$$\xi(\gamma) \leq \frac{\mu(\gamma)}{\Xi_{\mathcal{L}_\gamma}(\boldsymbol{\mu})} \tag{5.21}$$

with $\mathcal{L}_\gamma = \{\gamma' \in \mathcal{L} : \gamma' \nsim \gamma\}$.
As we did previoulsy we set $\mu(\gamma) = \xi(\gamma)e^{a|\gamma|}$ so that condition becomes

$$\Xi_{\mathcal{L}_\gamma}(\boldsymbol{\mu}) \leq e^{a|\gamma|} \tag{5.22}$$

where

$$\Xi_{\mathcal{L}_\gamma}(\boldsymbol{\mu}) = 1 + \sum_{n=1}^{|\gamma|} \frac{1}{n!} \sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{L}_\gamma^n \\ \gamma_i\nsim\gamma_j}} \prod_{i=1}^{n} |\xi(\gamma_i)|e^{a|\gamma_i|}$$

Consider now the factor

$$\sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{L}_\gamma^n \\ \gamma_i\nsim\gamma_j}} \prod_{i=1}^{n} |\xi(\gamma_i)|e^{a|\gamma_i|} \tag{5.23}$$

We have thus to choose n lattice animals $\gamma_1, \dots \gamma_n$ all incompatible with a given lattice animal γ and all pairwise compatible. We recall that two lattice animals are incompatible of they share a vertex os $\mathbb{Z}^2$. The factor (5.23) is zero whenever $n > |V_\gamma|$, $\gamma_i \nsim \gamma$ since γ has $|V_\gamma|$ vertices and thus we can arrange at most $|V_\gamma|$ lattice animals pairwise compatible each one sharing a different vertex of V_γ.

Therefore, when the factor above is not zero, i.e. for $n \leq |V_\gamma|$, it can be bounded at least by (again a very rough bound)

$$\sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{L}_\gamma^n \\ \gamma_i\sim\gamma_j}} \prod_{i=1}^{n} |\xi(\gamma_i)|e^{a|\gamma_i|} \leq |V_\gamma|(|V_\gamma|-1)\cdots(|V_\gamma|-n+1)\left[\sup_{x\in\mathbb{Z}^2}\sum_{\substack{\gamma\in\mathcal{L}\\ x\in\gamma}}|\xi(\gamma)|e^{a|\gamma|}\right]^n =$$

$$= \binom{|V_\gamma|}{n} n!\left[\sup_{x\in\mathbb{Z}^2}\sum_{\substack{\gamma\in\mathcal{L}\\ x\in\gamma}}|\xi(\gamma)|e^{a|\gamma|}\right]^n$$

Thus

$$\Xi_{\mathcal{L}_\gamma}(\boldsymbol{\mu}) \leq 1+\sum_{n=1}^{|V_\gamma|}\binom{|V_\gamma|}{n}\left[\sup_{x\in\mathbb{Z}^2}\sum_{\substack{\gamma\in\mathcal{L}\\ x\in\gamma}}|\rho(\gamma)|e^{a|\gamma|}\right]^n = \left[1+\sup_{x\in\mathbb{Z}^2}\sum_{\substack{\gamma\in\mathcal{L}\\ x\in\gamma}}|\xi(\gamma)|e^{a|\gamma|}\right]^{|V_\gamma|}$$

Thus (5.22) can be written as

$$\left[1+\sup_{x\in\mathbb{Z}^2}\sum_{\substack{\gamma\in\mathcal{L}\\ x\in\gamma}}|\xi(\gamma)|e^{a|\gamma|}\right]^{|V_\gamma|} \leq e^{a|\gamma|} \tag{5.24}$$

and since, $|V_\gamma| \leq |\gamma|$ for any $\gamma \in \mathcal{L}$ (the equality holding only if γ is a cycle) we have that (5.24) is surely satisfied if

$$\sup_{x\in\mathbb{Z}^2}\sum_{\substack{\gamma\in\mathcal{L}\\ x\in\gamma}}|\xi(\gamma)|e^{a|\gamma|} \leq e^a - 1 \tag{5.25}$$

Observe finally that, due to the symmetry of the problem the function

$$f(x) = \sum_{\substack{\gamma\in\mathcal{L}\\ x\in\gamma}}|\xi(\gamma)|e^{a|\gamma|}$$

is constant as x varies in $\mathbb{Z}^2$. Therefore (5.25) is equivalent to the condition

$$\sum_{\substack{\gamma\in\mathcal{L}\\ 0\in\gamma}}|\xi(\gamma)|e^{a|\gamma|} \leq e^a - 1 \tag{5.26}$$

where 0 is the origin in $\mathbb{Z}^2$.
The condition (5.26) is a (high temperature) convergence condition for the analyticity of the free energy of the Ising model at zero magnetic field and free boundary conditions. Now recalling (5.20) and due to the symmetry of the problem

$$\sum_{\substack{\gamma\in\mathcal{L}\\ 0\in\gamma}}[\tanh(\beta J)]^{|\gamma|}e^{a|\gamma|} = \sum_{n\geq 4}[\tanh(\beta J)]^n e^{an}\sum_{\substack{\gamma\in\mathcal{L}\\ 0\in\gamma,\ |\gamma|=n}} 1 = \sum_{n\geq 4}[\tanh(\beta J)]^n e^{an}C_n$$

with

$$C_n = \sum_{\substack{\gamma\in\mathcal{L}\\ 0\in\gamma,\ |\gamma|=n}} 1$$

being the number of lattice animals in $\mathcal{L}$ made by n nearest neighbor bonds containing the origin. I.e. we need to count all lattice animals γ with a given cardinality $|\gamma| = n$ that pass through the origin. To do this just observe that the nearest neighbor bonds of a lattice animal form a graph with degree 2 or 4, i.e. a graph with even degree vertices. It is long known that any graph with even degree at its vertices admits an Eulerian cycle (i.e. a graph cycle that crosses each edge exactly once). Therefore all lattice animals containing the origin can be formed by performing a cycle starting at the origin. To form a cycle starting at the origin in $\mathbb{Z}^2$ we can take each time 3 directions (also at the beginning since the cycle can be traveled in two direction) This immediately implies that

$$C_n = \sum_{\substack{\gamma\in\mathcal{L}\\ 0\in\gamma,\ |\gamma|=n}} 1 \leq 3^n$$

Therefore condition (5.26) is surely satisfied if

$$\sum_{n\geq 4}[3\tanh(\beta J)]^n e^{an} \leq e^a - 1$$

i.e. posing $x = |3\tanh(\beta J)|$ as soon as

$$e^{4a}[3\tanh(\beta J)]^4 + (e^{2a} - e^a)[3\tan(\beta J)] - (e^a - 1) \leq 0$$

which yields (for $a = 0.21$)

$$3\tanh(\beta J) \leq 0.4577$$

Hence convergence of the free energy, uniformly in the volume Λ occurs as soon as

$$\beta \leq \beta_0 \doteq \frac{1}{3J}\tanh^{-1}(0.1525) \approx \frac{0.151}{J} \tag{5.27}$$

5.2.2 Low temperature expansion

We consider now the zero magnetic field Ising model with + boundary conditions in a box Λ which is assumed to be a $L \times L$ square with L^2 sites. Thus, given a configuration σ_Λ of the spins inside Λ, the Hamiltonian is

$$H^+_\Lambda(\sigma_\Lambda) \;=\; -J \sum_{\substack{\{x,y\}\subset\Lambda\\ |x-y|\,=\,1}} \sigma_x\sigma_y - J \sum_{y\in\partial\Lambda} \sum_{\substack{x\in\Lambda,\\ |x-y|\,=\,1}} \sigma_x$$

Another way to write the Hamiltonian of the zero field + boundary condition Ising model is as follows is

$$H^+_\Lambda(\sigma_\Lambda) \;=\; -J \sum_{\substack{\{x,y\}\cap\Lambda\neq\emptyset\\ |x-y|\,=\,1}} \sigma_x\sigma_y$$

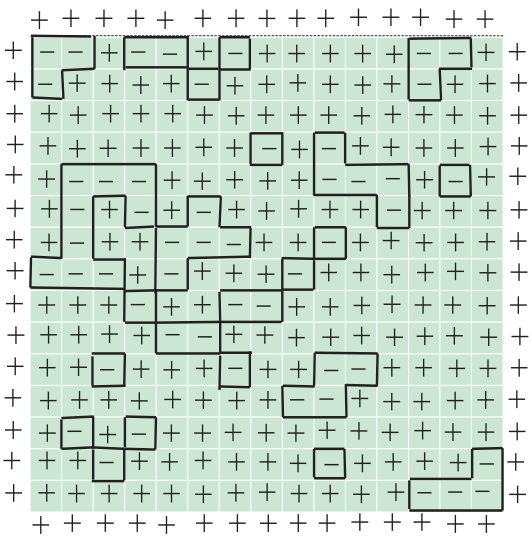


Figure 13. The grey square is Λ^*. The outer spins are those fixed by the boundary conditions.

recalling that $\sigma_y = +1$ whenever $y \in \partial\Lambda$.

We now rewrite the partition function $Z^+_\Lambda(\beta)$ via a contour gas in the following way. For a fixed configuration σ_Λ draw a unit segment perpendicular to the center of each bond b of nearest neighbors having opposite spins at its extremes (in three dimensions this segment becomes a unite square surface). Such unit segments have extremities which are in a lattice shifted respect to the original lattice by a factor $1/2$ along both x and y axis. This lattice is called the dual lattice and denoted by $\mathbb{Z}^{*2}$. Hence this set of unit segments forms a graph of nearest neighbor links in $\Lambda^* \in \mathbb{Z}^{*2}$ where Λ^* is a $(L+1)\times(L+1)$ square in the dual lattice. With the particular choice of $+$ boundary conditions such graphs are exactly the same of Figure 11, with the only difference that now they live in Λ^*, i.e. in a square with $(L+1)\times(L+1)$ sites in the dual lattice. See Figure 13.

To any configuration σ_Λ in Λ we can thus associate a graph Γ in the dual lattice made by nearest neighbor links with incidence number equal either 2 or 4. The graph Γ splits into its connected components in the usual manner. Hence $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_k\}$ with $\gamma_i \cap \gamma_j = \emptyset$. In this case (low temperature expansion), such connected components γ are called *contours* and $|\gamma|$ is now called perimeter of the contour. Note that now a contour configuration correspond to a spin configuration. I.e., given the spin configuration σ_Λ in Λ we can recover the contour configuration $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_k\}$ and conversely, given the contour configuration $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_k\}$ we can recover the spin configuration σ_Λ which originated it.

I.e. there is a one to one correspondence $\sigma_\Lambda \to \{\gamma_1, \ldots, \gamma_k\}$. We can express the Hamiltonian, which is a function of σ_Λ, as a function of the contour configuration $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_k\}$ uniquely determined by σ_Λ. This is done very

easily, just observing that, for a fixed configuration σ_Λ we have

$$H^+_\Lambda(\sigma_\Lambda) \;=\; -J \sum_{\substack{\{x,y\}\cap\Lambda\neq\emptyset \\ |x-y| \,=\, 1}} \sigma_x\sigma_y \;=\; -J\left[B^+ - B^-\right]$$

where

$$B^+ \;=\; \sharp \text{ of nearest neighbor pairs } \sigma_x\sigma_y \text{ with } \sigma_x \;=\; \sigma_y$$

$$B^- \;=\; \sharp \text{ of nearest neighbor pairs } \sigma_x\sigma_y \text{ with } \sigma_x \;=\; -\,\sigma_y$$

If now $\{\gamma_1,\dots,\gamma_k\}$ is the contour configurations corresponding to σ_Λ we have that the number of nearest neighbor pairs with opposite signal is by construction equal to $\sum_{i=1}^k |\gamma_i|$, where recall that $|\gamma_i|$ is the perimeter of γ

$$B^- \;=\; \sum_{i=1}^k |\gamma_i|$$

Denoting with

$$\tilde{B}_\Lambda \;=\; 2L(L-1) + 4L \;=\; 2L(L+1)$$

the number of nearest neighbor pairs with non void intersection with Λ, we also have

$$B^+ \;=\; \tilde{B}_\Lambda - B^-$$

hence

$$H^+_\Lambda(\sigma_\Lambda) \;=\; H_{\Lambda^*}(\gamma_1,\dots,\gamma_k) \;=\; -J\tilde{B}_\Lambda + 2J\sum_{i=1}^k |\gamma_i|$$

Hence the partition function of the Ising model with $h \;=\; 0$ and with +boundary conditions can be written as

$$Z^+_\Lambda(\beta) \;=\; e^{+\beta J\tilde{B}_\Lambda}\left[1 + \sum_{n\geq 1}\sum_{\substack{\{\gamma_1,\dots,\gamma_n\}\subset\mathcal{L}_{\Lambda^*} \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} e^{-2\beta J\sum_{i=1}^n |\gamma_i|}\right] \tag{5.28}$$

the factor 1 inside the bracket correspond to the situation in which no contour is present or equivalently when all spin are +1. As usual $\gamma_i \sim \gamma_j$ means that $V_{\gamma_i} \cap V_{\gamma_j} \;=\; \emptyset$ where recall that V_{γ_i} and V_{γ_j} are subsets of Λ^*.
The expression in square bracket in (5.28) it is the grand canonical partition function of a hard core lattice polymer gas in which the polymers are as before contours γ_i but now with activity $\exp[-2\beta J|\gamma_i|]$, i.e.

$$\Xi_{\Lambda^*}(\beta) \;=\; 1 + \sum_{n\geq 1}\sum_{\substack{\{\gamma_1,\dots,\gamma_n\}\subset\mathcal{L}_{\Lambda^*} \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} e^{-2\beta J\sum_{i=1}^n |\gamma_i|} \;=$$

$$1 + \sum_{n\geq 1}\frac{1}{n!}\sum_{\substack{(\gamma_1,\dots,\gamma_n)\in\mathcal{L}^n_{\Lambda^*} \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} \rho(\gamma_1)\dots\rho(\gamma_n)$$

where now the activity of contours is

$$\rho(\gamma) \;=\; e^{-2\beta J|\gamma|} \tag{5.29}$$

Note that the activity of a contour goes to zero if the perimeter $|\gamma|$ of the contour increases, so that big contours are depressed, but also the activity of a contour goes to zero when β goes to infinity, i.e all contours tend to be depressed when the temperature is lower and lower. Hence at very low temperature one is expected to see a sea of + with very small and very few contours which surround islands of -. This means that the in Ising model at zero magnetic field and very low temperature spins are predominantly oriented +, and this fact occurs independently on the size Λ where the system is confined. This give a quite clear a detailed picture of the low temperature phase of the Ising model with + boundary condition and $h \;=\; 0$.
We can now formulate a condition for the analyticity of the free energy of the Ising model at $h \;=\; 0$ with +boundary conditions in a completely analogous way as we did for the case of high temperature. The only thing that changes is that now the activity of contours (which in the high temperature case where called lattice animals, but they are the same objects) is defined by (5.29). Again we have to check a formula identical to (5.26) where now in place of $\xi(\gamma)$ we have to put $\rho(\gamma)$ defined in (5.29). Hence we reduce ourselves to check for which β the following inequality holds

$$\sum_{n\geq 4}[3e^a\, e^{-2\beta J}]^n \leq e^a - 1$$

which is satisfied, for $a = 0.21$, if

$$3e^{-2\beta J} \leq 0.4577$$

i.e.

$$\beta \geq \beta_1 \;\doteq\; \frac{1}{2J}\ln\left[\frac{3}{0.4577}\right] \approx \frac{0.94}{J} \tag{5.30}$$

we have that the finite volume free energy $f_\Lambda(\beta)$ is analytic in β for all $\beta \geq \beta_1 = \frac{0.94}{J}$ uniformly in Λ. Note that in this case $f_\Lambda(\beta)$ is calculated with + boundary condition, but the limit $f(\beta) \;=\; \lim_{\Lambda\to\infty} f_\Lambda(\beta)$ which is also analytic in β for all $\beta > \beta_1$, does not depend on boundary conditions.
In conclusion we have proved the following theorem

Theorem. The free energy of the Ising model at zero magnetic field

$$f(\beta, h=0) \;=\; \lim_{\Lambda\to\infty}\frac{1}{|\Lambda|}\ln Z_\Lambda(\beta, h=0)$$

is analytic in β for all $\beta \in (0,\beta_0]\cup[\beta_1,+\infty)$ where β_0 is defined in (5.27) and β_1 is defined in (5.30). We now show that there is at least a point in the interval (β_0,β_1) where $f(\beta, h=0)$ is non analytic.

5.2.3 Existence of phase transitions

First recall the definition of the magnetization of the system as

$$M_\Lambda(\beta) \;=\; \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\sigma_x\rangle_\Lambda \tag{5.31}$$

The quantity $M_\Lambda(\beta)$ measures the mean orientation of spins of the finite system confined in Λ. If $M_\Lambda(\beta) > 0$ then it means that spin + are predominant while if $M_\Lambda(\beta) < 0$ this means that spin - are predominant. It is just a simple calculus to show the identity

$$M_\Lambda(\beta) \;=\; \beta^{-1}\,\frac{\partial f_\Lambda(\beta,h)}{\partial h}\Big|_{h\,=\,0}$$

which tells us that the (finite volume) magnetization is, modulo a factor β^{-1}, the first derivative respect to h of the finite volume free energy of the system.

By a Theorem called "Lee Yang theorem" is possible to show that $f(\beta,h)$ is analytic for all $\beta > 0$ when $h \neq 0$ hence the only singularity of this function can occur when $h = 0$.

Hence in order to show existence of transition phase, our strategy will be to show that, depending on boundary condition the infinite volume magnetization $M(\beta) \;=\; \lim_{\Lambda\to\infty} M_\Lambda(\beta)$ is not stable for variations of the boundary conditions.

Thus we will show first that the magnetization $M^+_\Lambda(\beta)$ of the Ising model with + boundary conditions (the superscript "+" remember us that we are using + boundary conditions) is a number arbitrarily near 1 for β sufficiently big. The result will also immediately imply that the magnetization $M^-_\Lambda(\beta)$ of the Ising model with - boundary conditions is a number arbitrarily near -1 for β sufficiently big.

Let us denote $\langle\cdot\rangle^+_\Lambda$ the mean values when the Hamiltonian is taken with $h \;=\; 0$ and with + boundary condition. It is easy to see that, for any $x \in \Lambda$

$$\langle\sigma_x\rangle^+_\Lambda \;=\; 1 - 2\Big[\mathrm{Prob}^+_\Lambda(\sigma_x = -1)\Big] \tag{5.32}$$

where $\mathrm{Prob}^+_\Lambda(\sigma_x = -1)$ denotes the probability that the spin σ_x at the site x is equal to -1 for the Ising model with zero magnetic field and + boundary conditions. As a matter of fact, by definition

$$\langle\sigma_x\rangle^+_\Lambda \;=\; (+1)\,\Big[\mathrm{Prob}^+_\Lambda(\sigma_x = +1)\Big] \;+\; (-1)\,\Big[\mathrm{Prob}^+_\Lambda(\sigma_x = -1)\Big] \;=$$

$$=\; 1 \;-\; 2\Big[\mathrm{Prob}^+_\Lambda\sigma_x = -1)\Big]$$

Thus in order to evaluate $\langle\sigma_x\rangle^+_\Lambda$ it is sufficient to evaluate $\mathrm{Prob}^+_\Lambda(\sigma_x = -1)$. We have seen before that the Ising model with + boundary condition and zero magnetic field can be mapped in a contour gas, and that there is a one to one correspondence between spin configurations σ_Λ and contour configurations $\gamma_1, \ldots \gamma_k$. Now a σ_Λ such that the spin at x is -1 is such that the site x is the

"interior" of at least one contour associated to σ_Λ, i.e. x is surrounded at least by a contour (actually by an odd number of contours). Hence, if we denote $\mathrm{Prob}^+_\Lambda(\exists\gamma \odot x)$ the probability that at least one contour surrounds x, we have surely

$$\mathrm{Prob}^+_\Lambda(\sigma_x \;=\; -1) \leq \mathrm{Prob}^+_\Lambda(\exists\gamma \odot x)$$

But now

$$\mathrm{Prob}^+_\Lambda(\exists\gamma \odot x) \;=\; \frac{\sum_{\substack{\{\gamma_1,\ldots,\gamma_n\}:\exists\gamma_i\odot x \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} e^{-2\beta J\sum_{i=1}^n|\gamma_i|}}{\sum_{\substack{\{\gamma_1,\ldots,\gamma_n\} \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} e^{-2\beta J\sum_{i=1}^n|\gamma_i|}}$$

where the denominator is the grand canonical partition function of the contour gas and in the sum in the denominator is also included the empty graph which contribute with the factor 1. Now we can write

$$\mathrm{Prob}^+_\Lambda(\exists\gamma \odot x) \;=\; \frac{\sum_{\substack{\{\gamma_1,\ldots,\gamma_n\}:\exists\gamma_i\odot x \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} e^{-2\beta J\sum_{i=1}^n|\gamma_i|}}{\sum_{\substack{\{\gamma_1,\ldots,\gamma_n\} \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} e^{-2\beta J\sum_{i=1}^n|\gamma_i|}} \;=$$

$$= \frac{\sum_{\gamma\odot x} e^{-2\beta J|\gamma|}\sum_{\substack{\{\gamma_1,\ldots,\gamma_n\}:\gamma_i\sim\gamma \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} e^{-2\beta J\sum_{i=1}^n|\gamma_i|}}{\sum_{\substack{\{\gamma_1,\ldots,\gamma_n\} \\ |\gamma_i|\geq 4,\ \gamma_i\sim\gamma_j}} e^{-2\beta J\sum_{i=1}^n|\gamma_i|}} \leq$$

$$\leq \sum_{\gamma\odot x} e^{-2\beta J|\gamma|}$$

I.e. in conclusion we get

$$\mathrm{Prob}^+_\Lambda(\exists\gamma \odot x) \leq \sum_{\gamma\odot x} e^{-2\beta J|\gamma|} \leq \sum_{n\geq 4} e^{-2\beta Jn}\sum_{\gamma\odot x} 1$$

It is now easy to bound $\sum_{\gamma\odot x}$ with $n3^n$. As a matter of fact, let us denote with x' the point of intersection of a contour γ (such that $|\gamma| = n$)which surrounds x with the horizontal axis passing through x. Let us ask ourselves how many possible x' we can get. Obviously not more than n. Now the possible contours with fixed perimeter $|\gamma| = n$ which pass through a fixed x' are at most 3^n, so $\sum_{\gamma\odot x} 1 \leq n3^n$. Hence

$$\mathrm{Prob}^+_\Lambda(\exists\gamma \odot x) \leq \sum_{n\geq 4} e^{-2\beta Jn} n3^n \;\doteq\; C(\beta)$$

Note that the sum in l.h.s. converges to if $3e^{-2\beta J} < 1$ i.e. if

$$\beta > \beta_1' \doteq \frac{1}{2J}\ln 3 \;\approx\; \frac{0.55}{J}$$

Note that $C(\beta)$ defined above goes to zero as $\beta \to \infty$. Thus when $\beta > \frac{1}{2J}\ln 3$ we have that

$$\langle\sigma_x\rangle^+_\Lambda \;\geq\; 1 - 2C(\beta)$$

and recalling (5.31)

$$M_\Lambda(\beta, h=0) \geq 1-2C(\beta) \implies M(\beta, h=0) = \lim_{\Lambda\to\infty} M_\Lambda(\beta, h=0) \geq 1-2C(\beta)$$

Note that $\langle\sigma_x\rangle_\Lambda^+$ tends to 1 as $\beta \to \infty$ and this estimate is uniform in Λ.
On the other side, if we bound $\langle\sigma_x\rangle_\Lambda^-$ (i.e. the mean value of the spin at a site x with - boundary conditions) we have

$$\langle\sigma_x\rangle_\Lambda^- \;=\; (+1)\,\Big[\mathrm{Prob}_\Lambda^-(\sigma_x=+1)\Big] \;+\; (-1)\,\Big[\mathrm{Prob}_\Lambda^-(\sigma_x=-1)\Big] \;=$$

$$=\; 2\,\Big[\mathrm{Prob}_\Lambda^-(\sigma_x=+1)\Big] \;-\; 1$$

but now we have obviously that $\mathrm{Prob}_\Lambda^-(\sigma_x = +1) \;=\; \mathrm{Prob}_\Lambda^+(\sigma_x = -1)$ and hence

$$\langle\sigma_x\rangle_\Lambda^- \;\leq\; 2C(\beta) - 1$$

In this case $\langle\sigma_x\rangle_\Lambda^-$ tends to -1 as $\beta \to \infty$.
In conclusion the system show spontaneous magnetization for β sufficiently high uniformly in the volume Λ. I.e. we have shown that e.g.

$$\lim_{\Lambda\to\infty} \langle\sigma_x\rangle_\Lambda^+ \neq 0 \qquad \text{if} \;\; \beta > \frac{1}{2J}\ln 3$$

In other words the system is not stable even at the infinite volume limit to boundary conditions.
As a last computation we show that when β is small there is no such instability. As a matter of fact, we can express $\langle\sigma_x\rangle_\Lambda^+$ in term of high temperature lattice animals. Recall that the Hamiltonian on the Ising model with + boundary conditions and zero magnetic field is

$$H_\Lambda^+(\sigma_\Lambda) \;=\; -\,J \sum_{\substack{\{x,y\}\cap\Lambda\neq\emptyset \\ |x-y| \,=\, 1}} \sigma_x\sigma_y$$

with $\sigma_y \;=\; +1$ whenever $y \in \partial\Lambda$. And the partition function is

$$Z_\Lambda^+(\beta) \;=\; \sum_{\sigma_\Lambda} e^{+\beta J \sum_{\substack{\{x,y\}\cap\Lambda\neq\emptyset \\ |x-y| \,=\, 1}} \sigma_x\sigma_y} \;=\; \sum_{\sigma_\Lambda} \prod_{\substack{\{x,y\}\cap\Lambda\neq\emptyset \\ |x-y| \,=\, 1}} e^{+\beta J\sigma_x\sigma_y} \;=$$

$$=\; \sum_{\sigma_\Lambda} \prod_{\substack{b \\ b\cap\Lambda\neq\emptyset}} e^{+\beta J\tilde\sigma_b}$$

where now the nearest neighbor pair b are those strictly contained in Λ plus the nearest neighbor pairs $\{x, y\}$ for which $x \in \Lambda$ and $y \in \partial\Lambda$. In this last case, since we are using + boundary conditions, $\tilde\sigma_b = \sigma_x\sigma_y = +\sigma_x$. As before we can write

$$Z_\Lambda^+(\beta) \;=\; \sum_{\sigma_\Lambda} \prod_{\substack{b \\ b\cap\Lambda\neq\emptyset}} \cosh\beta J\,[1 + \tilde\sigma_b \tanh\beta J]$$

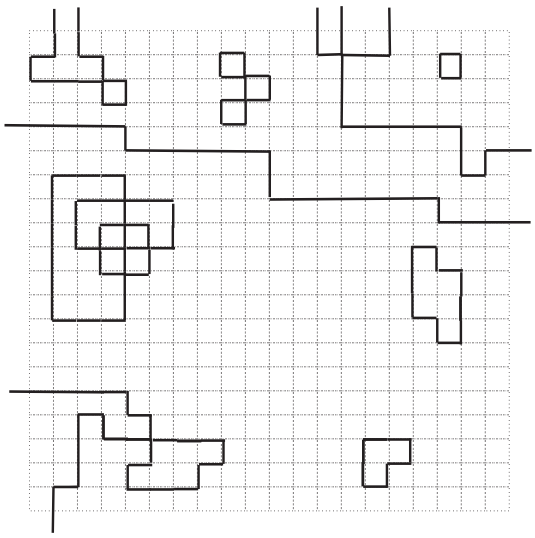


Figure 14. A non vanishing configuration of lattice animals in $Z^+_\Lambda(\beta)$.

and supposing as usual that Λ is a square of size L we have that the number of b such that $b \cap \Lambda \neq \emptyset$ is $2L(L-1) + 4L \;=\; 2L(L+1)$, hence

$$Z^+_\Lambda(\beta) \;=\; [\cosh \beta J]^{2L(L+1)} \sum_{\sigma_\Lambda} \prod_{\substack{b \\ b \cap \Lambda \neq \emptyset}} [1 + \tilde{\sigma}_b \tanh \beta J]$$

As before the development of the product

$$\prod_{\substack{b \\ b \cap \Lambda \neq \emptyset}} [1 + \tilde{\sigma}_b \tanh \beta J] \tag{5.33}$$

gives rise to terms of the form

$$(\tanh \beta J)^k \tilde{\sigma}_1 \dots \tilde{\sigma}_k$$

which can be associated to graphs in $\Lambda \cup \partial\Lambda$. It is important to stress that when $b \subset \Lambda$ then we call b internal bond, i.e. $b \;=\; \{x, y\}$ with $x \in \Lambda$ and $y \in \Lambda$ and $\tilde{\sigma}_b \;=\; \sigma_x \sigma_y$ where *both* σ_x *and* σ_y are summed in the summation $\sum_{\sigma_\Lambda}$ over spin configurations. On the other hand when $b \cap \partial\Lambda \neq \emptyset$, then we call b boundary bond i.e. $b \;=\; \{x, y\}$ with $x \in \Lambda$ and $y \in \partial\Lambda$ and $\tilde{\sigma}_b \;=\; \sigma_x \sigma_y$ where only σ_x is summed in the summation $\sum_{\sigma_\Lambda}$, while $\sigma_y \;=\; 1$

So this time graphs which will not vanish after the sum over spin configuration are not only closed polygons type lattice animals as before. There are also new type of lattice animals. New non vanishing lattice animals are also those which start in a site $y \in \partial\Lambda$ and end in another site $y' \in \partial\Lambda$. See Figure 14 below. Any non vanish factor $(\tanh \beta)^k J \tilde{\sigma}_1 \dots \tilde{\sigma}_k$, independently of the nature of the bonds b (i.e. internal bonds or boundary bonds) becomes $(\tanh \beta)^k J 2^{L^2}$. Note that all non vanishing factors $(\tanh \beta)^k J \tilde{\sigma}_1 \dots \tilde{\sigma}_k$ are all positive even if we were imposing $-$ boundary conditions (so that now for a boundary bond

$b = \{x, y\}$ with $y \in \partial\Lambda$ we should have $\sigma_y = -1$), because any new type (connected) lattice animals has always an even number of paths ending at the boundary so that it has only an *even* number of such points where $\sigma_x = 1$. So that negative values of spins at the boundary cancels. In conclusion we have again

$$Z^+_\Lambda(\beta) = [\cosh(\beta J)]^{2L(L-1)} 2^{L^2} \, \Xi^+_\Lambda(\beta) \tag{5.34}$$

where

$$\Xi^+_\Lambda(\beta) = 1 + \sum_{n \geq 1} \sum_{\substack{\{\gamma_1, \dots \gamma_n\}: \, |\gamma_i| \geq 4 \\ \gamma_i \sim \gamma_j}} \xi(\gamma_1) \dots \xi(\gamma_n) \tag{5.35}$$

where again γ denote a allowed lattice animal (just recall that there are in this case more new lattice animals allowed respect to the case of free boundary conditions) with activity

$$\xi(\gamma) = [\tanh(\beta J)]^{|\gamma|}$$

Let us now express the mean value $\langle \sigma_x \rangle^+_\Lambda$ (where $x \in \Lambda$) in terms of lattice animals. By definition we have

$$\langle \sigma_x \rangle^+_\Lambda = \frac{\sum_{\sigma_\Lambda} \sigma_x e^{+\beta J \sum_{\substack{\{x,y\} \cap \Lambda \neq \emptyset \\ |x-y| = 1}} \sigma_x \sigma_y}}{Z^+_\Lambda(\beta)} \tag{5.36}$$

The numerator of the expression above can be easily rewritten in term of lattice animals as

$$\sum_{\sigma_\Lambda} \sigma_x e^{+\beta J \sum_{\substack{\{x,y\} \cap \Lambda \neq \emptyset \\ |x-y| = 1}} \sigma_x \sigma_y} =$$

$$= [\cosh(\beta J)]^{2L(L-1)} 2^{L^2} \sum_{n \geq 1} \sum_{\substack{\{\gamma_1, \dots \gamma_n\}: \\ \gamma_i \sim \gamma_j, \, \exists \gamma_x}} \xi(\gamma_1) \dots \xi(\gamma_n) \tag{5.37}$$

where now in $\sum$ the notation $\exists \gamma_x$ means that among lattice animals $\gamma_1, \dots \gamma_n$ at least one of them starts at x and end at a boundary site, see Figure 15 for such kind of lattice animals (actually are also allowed lattice animals which have incidence number 3 in x). Let as indicate with $\mathcal{L}_x$ the set os all this lattice animals

Note also that the factor 1 (i.e that corresponding to the empty lattice animal) is no more present in the numerator (5.37).

Thus we can write

$$\langle \sigma_x \rangle^+_\Lambda = \frac{\sum_{\substack{\{\gamma_1, \dots \gamma_n\}: \\ \gamma_i \sim \gamma_j, \, \exists \gamma_x}} \xi(\gamma_1) \dots \xi(\gamma_n)}{1 + \sum_{\substack{\{\gamma_1, \dots \gamma_n\}: \, |\gamma_i| \geq 4 \\ \gamma_i \sim \gamma_j}} \xi(\gamma_1) \dots \xi(\gamma_n)} \leq$$

$$\leq \frac{\sum\limits_{\substack{\gamma \in \mathcal{L}_x \\ V_\gamma \cap \partial\Lambda \neq \emptyset}} \xi(\gamma) \sum_{\substack{\{\gamma_1, \dots \gamma_n\}: \, |\gamma_i| \geq 4 \\ \gamma_i \sim \gamma_j, \, \gamma_i \sim \gamma}} \xi(\gamma_1) \dots \xi(\gamma_n)}{1 + \sum_{\substack{\{\gamma_1, \dots \gamma_n\}: \, |\gamma_i| \geq 4 \\ \gamma_i \sim \gamma_j}} \xi(\gamma_1) \dots \xi(\gamma_n)} \leq \sum_{\substack{\gamma \in \mathcal{L}_x \\ V_\gamma \cap \partial\Lambda \neq \emptyset}} \xi(\gamma)$$

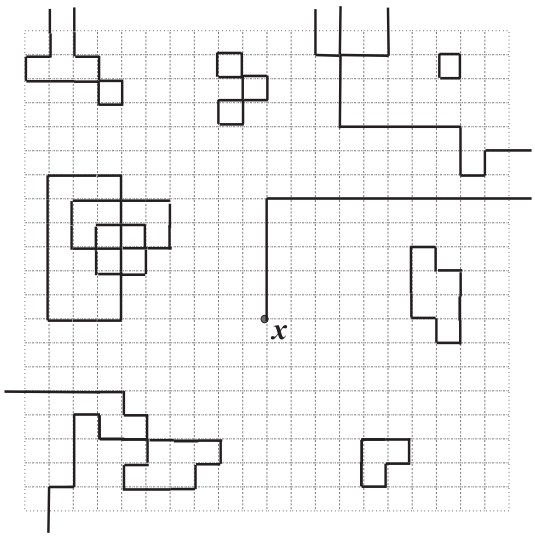


Figure 15. A non vanishing configuration of lattice animals in the numerator of $\langle \sigma_x \rangle^+_\Lambda$.

In conclusion we have that

$$\langle \sigma_x \rangle^+_\Lambda \;\leq\; \sum_{\substack{\gamma \in \mathcal{L}_x \\ V_\gamma \cap \partial\Lambda \neq \emptyset}} \xi(\gamma) \tag{5.38}$$

Observe now that if $\mathcal{L}^1_x$ denotes the set of paths starting at x (i.e. lattice animals having degree 1 at x) we can reorganize the sum above as follows

$$\sum_{\substack{\gamma \in \mathcal{L}_x \\ V_\gamma \cap \partial\Lambda \neq \emptyset}} \xi(\gamma) \leq \sum_{\substack{\gamma \in \mathcal{L}^1_x \\ V_\gamma \cap \partial\Lambda \neq \emptyset}} \xi(\gamma) \left[1 + \left(\sum_{\substack{\tilde\gamma \in \mathcal{L}^1_x \\ V_{\tilde\gamma} \cap \partial\Lambda \neq \emptyset}} \xi(\tilde\gamma) \right)^2 \right]$$

If we denote by $d(x, \partial\Lambda)$ the minimum distance (in nearest neighbor bonds) between x and the boundary $\partial\Lambda$, we have clearly that a γ in the sum must have at leat $|\gamma| \;=\; d(x, \partial\Lambda)$ bonds. So

$$\sum_{\substack{\gamma \in \mathcal{L}^1_x \\ V_\gamma \cap \partial\Lambda \neq \emptyset}} \xi(\gamma) \;\leq\; \sum_{n \geq d(x,\partial\Lambda)} (\tanh \beta J)^n \sum_{\substack{\gamma \in \mathcal{L}^1_x \colon |\gamma| = n \\ V_\gamma \cap \partial\Lambda \neq \emptyset}} 1 \;\leq\; \frac{4}{3} \sum_{n \geq d(x,\partial\Lambda)} (3 \tanh \beta J)^n$$

where we have estimated

$$\sum_{\substack{\gamma \in \mathcal{L}^1_x \colon |\gamma| = n \\ V_\gamma \cap \partial\Lambda \neq \emptyset}} \leq 4 \cdot 3^{n-1}$$

The series $\sum_{n \geq d(x,\partial\Lambda)} (3 \tanh \beta J)^n$ converges for $3 \tanh \beta J < 1$ i.e for

$$\beta < \beta'_0 \doteq \frac{1}{J} \tanh^{-1}(1/3) \approx \frac{0.34}{J} \tag{5.39}$$

and in conclusion

$$\sum_{\substack{\gamma\in\mathcal{L}^1_x\\ V_\gamma\cap\partial\Lambda\neq\emptyset}} \xi(\gamma) \;\leq\; \frac{4}{3}\frac{(3\tanh\beta J)^{d(x,\partial\Lambda)}}{1-3\tanh\beta J}$$

Therfore we get finally, for β as small as in condition (5.39)

$$\langle\sigma_x\rangle^+_\Lambda \;\leq\; \frac{4}{3}\frac{(3\tanh\beta J)^{d(x,\partial\Lambda)}}{1-3\tanh\beta J}\left[1+\left(\frac{4}{3}\frac{(3\tanh\beta J)^{d(x,\partial\Lambda)}}{1-3\tanh\beta J}\right)^2\right] \;\leq$$

$$\leq\; \frac{100}{3}\frac{(3\tanh\beta J)^{d(x,\partial\Lambda)}}{(1-3\tanh\beta J)^3}$$

where in the las line we have used that $3\tanh\beta J<1$. Hence the magnetization at zero magnetic field and at finite volume is bounded by

$$M^+_\Lambda(\beta,h=0) \;=\; \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\sigma_x\rangle^+_\Lambda \;\leq\; \frac{100}{3(1-3\tanh\beta J)^3}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}[3\tanh\beta J]^{d(x,\partial\Lambda)} \leq$$

$$\leq \frac{100}{3(1-3\tanh\beta J)^3}\frac{1}{|\Lambda|}\sum_{n=1}^{L/2}[3\tanh\beta J]^n \sum_{\substack{x\in\Lambda\\ d(x,\partial\Lambda)=n}} 1 \leq \frac{300\tanh\beta J}{3(1-3\tanh\beta J)^4}\frac{|\partial\Lambda|}{|\Lambda|}$$

hence we get, for $\beta<\beta_0'$

$$\lim_{\Lambda\to\infty} M^+_\Lambda(\beta,h=0) \;=\; 0$$

Note also, for any Λ

$$\lim_{\beta\to 0} M^+_\Lambda(\beta,h=0) \;=\; 0$$

Thus when the temperature is sufficiently high the magnetization of the system, even with + boundary conditions, tends to be zero in the thermodynamic limit. This in contrast with the result that we have shown for low temperature, where $M^+_\Lambda(\beta,h=0)$ is definitely away from zero and is near to one.
This result say to us that the Ising model does not present the phenomenon of the spontaneous magnetization when the temperature is sufficiently high, or in other words the bulk system is not sensible to change of boundary conditions in the thermodynamic limit. On the contrary, when the temperature is very low, the system shows indeed spontaneous magnetization and its bulk becomes sensible to boundary conditions even in the thermodynamic limit.
The interpretation of this fact is that the Ising model has a phase transition at zero magnetic field and at some critical value β_c of the inverse temperature. Below β_c the system behaves like in the high temperature regime and above β_c the system behave like the low temperature regime.
Note also that $\lim_{\beta\to 0} M^+_\Lambda(\beta,h=0)$ and $\lim_{\beta\to 0} M^-_\Lambda(\beta,h=0)$ has different values, the first near 1 and the second near -1, and therefore the thermodynamic limit $M(\beta,h=0)$ of the magnetization (i.e. the derivative of the free energy $f(\beta,h)$) does not have a definite value. This is an evidence that $f(\beta,h)$ has

discontinous derivative respect to h at $h = 0$ when the inverse temperature is greater than $\beta > \beta_1$. On the other hand we have seen that $\lim_{\beta\to 0} M^+_\Lambda(\beta, h = 0) = \lim_{\beta\to 0} M^-_\Lambda(\beta, h = 0) = 0$ when $\beta < \beta'_0$, therefore there is a value β_c such that for $\beta < \beta_c$ the infinite volume magnetization is zero and for $\beta > \beta_c$ the magnetization is different from zero. It can be shown that the free energy is non analytic in $\beta = \beta_c$ as a function of β.

5.2.4 The critical temperature

As a very last exercise we show that if there is a unique non analytic point of the free energy $f(\beta)$ at $h = 0$ in the interval $(0, \infty)$ then this point can stay only on a well defined β_c.

Consider the low temperature expansion of the partition function of the zero magnetic field, + boundary conditions Ising model in a box Λ which is a square of size $L - 1$ where L is an integer. Then contours live in the square Λ^* of size L and the partition function is

$$Z^+_{(L-1)\times(L-1)}(\beta) \;=\; e^{+\beta J 2L(L-1)} \sum_{\substack{\gamma_1,\dots,\gamma_n:\ \gamma_i\sim\gamma_j \\ \gamma_i\in L\times L}} \prod_{i=1}^{n} e^{-2\beta J|\gamma_i|} \tag{5.40}$$

On the other hand, via high temperature expansion, the partition function of the zero magnetic field, free boundary conditions Ising model in a box of size L can be written as

$$Z^{\text{open}}_{L\times L}(\beta) \;=\; \cosh(\beta J)^{2L(L-1)} 2^{L^2} \sum_{\substack{\gamma_1,\dots,\gamma_n:\ \gamma_i\sim\gamma_j \\ \gamma\in L\times L}} \prod_{i=1}^{n} [\tanh \beta J]^{|\gamma_i|} \tag{5.41}$$

Compare now (5.40) and (5.41) and note that the sums on closed polymers in both equations are identical.

Let thus pose $\beta^* = \beta^*(\beta)$ the function of β defined by the equation

$$e^{-2\beta^* J} \;=\; \tanh \beta J$$

i.e.

$$\beta^* \;=\; -\,\frac{\ln[\tanh \beta J]}{2J} \tag{5.42}$$

Then we can write, by (5.40) and (5.41)

$$\frac{Z^+_{(L-1)\times(L-1)}(\beta^*)}{e^{+2\beta^* J(L-1)(L-2)}} \;=\; \frac{Z^{\text{open}}_{L\times L}(\beta)}{\cosh(\beta J)^{2L(L-1)} 2^{L^2}}$$

and taking the logarithm on both sides

$$\ln Z^+_{(L-1)\times(L-1)}(\beta^*) - 2\beta^* J(L-1)(L-2) \;=$$

$$=\; \ln Z^{\text{open}}_{L\times L}(\beta) - 2L(L-1)\ln\cosh(\beta J) - L^2 \ln 2$$

Dividing by L^2 and taking the limit $L \to \infty$

$$\lim_{L\to\infty} \left\{ \frac{(L-1)(L-1)}{L^2} \frac{1}{(L-1)(L-1)} \ln Z^{+}_{(L-1)\times(L-1)}(\beta^*) \; - \right.$$

$$\left. -2\beta^* J \frac{(L-1)(L-2)}{L^2} \right\} \; =$$

$$= \lim_{L\to\infty} \left\{ \frac{1}{L^2} \ln Z^{\text{open}}_{L\times L}(\beta) - \frac{2L(L-1)}{L^2} \ln\cosh(\beta J) - \ln 2 \right\}$$

i.e., since free energy $f(\beta) = \lim_{\Lambda\to\infty} \ln Z^{\tau}_{\Lambda}(\beta)$ does not depend on boundary conditions

$$f(\beta^*) - 2\beta^* J = f(\beta) - \ln[2\cosh^2(\beta J)]$$

hence we get

$$f(\beta) = f(\beta^*) + \ln[2\cosh^2(\beta J)] - 2\beta^* J$$

i.e.

$$f(\beta) = f(\beta^*) + \ln[2\cosh^2(\beta J)] + \ln[\tanh \beta J] \tag{5.43}$$

We now suppose that $f(\beta)$ can have at most one singularity, and we conclude via (5.43) that this singularity, if exists, can occur only for $\beta = \beta^*$.
Note that $\beta^* = \beta^*(\beta)$ (see (5.42)) is analytic as a function of β for all $\beta \in (0,\infty)$. Now, by (5.43), $f(\beta) - f(\beta^*(\beta))$ is analytic for all $\beta \in (0,\infty)$. Let now $\beta' \in (0,\infty)$ such that $\beta' \neq \beta^*$, then by (5.43) $f(\beta)$ is analytic in $\beta = \beta'$. In fact suppose, by absurd, that $f(\beta)$ is non analytic in $\beta = \beta'$, then $f(\beta^*(\beta'))$ must also be non analytic, since only in this way $f(\beta') - f(\beta^*(\beta')) = \ln 2\cosh(\beta' J) + 2\ln[\tanh \beta' J]$ can be non singular. But if $\beta' \neq \beta^*$ then $f(\beta^*(\beta')) = f(\beta'')$ with $\beta'' \neq \beta'$. Hence $f(\beta)$ would have two singularities which is in contradiction with the hypothesis. This equation is saying that if $f(\beta)$ has a singularity point in the domain $\beta \in (0,\infty)$, this point can occur only when $\beta = \beta^*$, i.e. when

$$e^{-2\beta J} = \tanh \beta J$$

which is the correct value calculated via the Onsager solution.

Bibliography

[1] A. G. Basuev (1978) : *A theorem on minimal specific energy for classical systems*. Teoret. Mat. Fiz. **37**, no. 1, 130–134.

[2] A. G. Basuev (1979): *Representation for the Ursell functions, and cluster estimates*. Teoret. Mat. Fiz. **39**, no. 1, 94-105.

[3] Battle, G. A.; Federbush, P.: A note on cluster expansion, tree graph identity, extra $1/N!$ factors!!! *Lett. Math. Phys.* **8** (1984), no. 1, 55-57.

[4] Borgs, C. : *Absence of Zeros for the Chromatic Polynomial on Bounded Degree Graphs* Preprint (2005), to appear in Combinatorics, Probability and Computing.

[5] D. Brydges: *A Short Course on Cluster Expansion*, Les Houches 1984, K. Osterwalder, R. Stora eds., North Holland Press (1986)

[6] Brydges, D.: *A short course on cluster expansion*. Les Houches 1984, K. Osterwalder, R. Stora eds., North Holland Press

[7] J. C. Butcher: *The Numerical Analysis of Ordinary Differential Equations. Runge-Kutta and General Linear Methods*, Wiley-Interscience Publication (1987)

[8] Cammarota, C.: *Decay of Correlations for Infinite range Interactions in unbounded Spin Systems*, *Comm. Math Phys.* **85**, 517-528 (1982) (1986).

[9] A. Cayley: *A Theorem on trees*. Quarterly Journal of Pure and Applied Mathematics, **23**, 376–378 (1889).

[10] Dobrushin, R.L. Estimates of semi-invariants for the Ising model at low temperatures. In Topics in Statistical and Theoretical Physics, American Mathematical Society Translations, Ser. 2, 177 59-81 (1996).

[11] R. Fernández and A. Procacci: *Cluster expansion for abstract polymer models.New bounds from an old approach*, Commun. Math Phys, **274**, 123–140 (2007).

[12] R. Fernández; A. Procacci; B. Scoppola: *The analyticity region of the hard sphere gas. Improved bounds*. J. Stat. Phys., **128**, n.5, 1139–1143 (2007).

[13] M. E. Fisher and D. Ruelle: *The Stability of Many-Particle Systems*, J. Math. Phys. **7**, 260–270 (1966).

[14] Gallavotti G.: Statistical Mechanics. A short treatise, Springer Verlag, Berlin (1999)

[15] Gill Williamson, S. *Combinatorics for computer science* , computer science press, Rockville, MD. (1985).

[16] Glimm, J. and Jaffe, A.: Quantum Physics, A functional integral point of view, Spinger, New York (1987)

[17] J. Groeneveld: *Two theorems on classical many-particle systems.* Phys. Lett., **3**, 50–51 (1962).

[18] C. Gruber and H. Kunz, *General properties of polymer systems*, *Comm. Math Phys.*, **22**, 133-161 (1971).

[19] K. Huang: *Statisitical Mechanics*, John Wiley and Sons, New York (1987).

[20] J. E. Jones; A. E. Ingham (1925): *On the calculation of certain crystal potential constants, and on the cubic crystal of least potential energy.* Proc. Roy. Soc. Lond. A **107**, 636–653.

[21] Kuna, Tobias; Kondratiev, Yuri G.; Da Silva, Jose L.: Marked Gibbs measures via cluster expansion. *Methods Funct. Anal. Topology 4* , no. 4, 50–81 (1998).

[22] R. Kotecký and D. Preiss, Cluster expansion for abstract Polymer models, *Comm. Math Phys.*, **103**, 491-498 (1986).

[23] J. E. Mayer: *The Statistical Mechanics of Condensing Systems. I*, J. Chem. Phys. **5**, 67–73 (1937).

[24] J. E. Mayer : *Contribution to Statistical Mechanics*, J. Chem. Phys., **10**, 629–643 (1942).

[25] J. E. Mayer and M. G. Mayer: *Statistical Mechanics*, John Wiley & Sons, Inc. London: Chapman & Hall, Limited (1940).

[26] S. Miracle-Solé, *On the convergence of cluster expansions*, Physica A **279**, 244-9 (2000).

[27] T. Morais, A. Procacci and B. Scoppola: *On Lennard-Jones type potentials and hard-core potentials with an attractive tail*, J. Stat. Phys., **157** , p. 17-39 (2014).

[28] O. Penrose (1963): *Convergence of Fugacity Expansions for Fluids and Lattice Gases*, J. Math. Phys. **4**, 1312 (9 pages).

[29] O. Penrose, *Convergence of fugacity expansions for classical systems.* In "Statistical mechanics: foundations and applications", A. Bak (ed.), Benjamin, New York (1967).

[30] D. Ya. Petrina; V. I. Gerasimenk; P. V. Malyshev: *Mathematical foundations of classical statistical mechanics. Continuous systems.* Translated from the Russian by P. V. Malyshev and D. V. Malyshev. Second edition. Advanced Studies in Contemporary Mathematics, 8. Taylor & Francis, London (2002).

[31] Ch.E. Pfister, *Large deviation and phase separation in the two-dimensional Ising model*, Helv. Phys. Acta **64**, 953-1054 (1991).

[32] A. Procacci: *Erratum and Addendum: "Abstract Polymer Models with General Pair Interactions"*, Journal of Statistical Physics **135**, Issue 4, 779-786 (2009)

[33] A. Procacci and B. Scoppola: *Polymer gas approach to N-body lattice systems.* J. Statist. Phys. **96**, 49–68 (1999)

[34] Procacci, A.; de Lima, B. N. B.; Scoppola, B.: A Remark on high temperature polymer expansion for lattice systems with infinite range pair interactions. *Lett. Math. Phys.* **45** (1998), no. 4, 303-322.

[35] Procacci, A, Scoppola, B. and Gerasimov, V.: *Potts model on innite graphs and the limit of chromatic polynomials.* Commun. Math. Phys. **235** (2003) 215-231.

[36] A. Procacci, S. A. Yuhjtman: *Convergence of Mayer and Virial expansions and the Penrose tree-graph identity*, Preprint arxiv.org/abs/1508.07379v1.

[37] G. Rota (1964): *On the foundations of combinatorial theory. I. Theory of Möbius functions.* Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, **2**, 340–368.

[38] Ruelle, D.: Statistical Mechanics, Rigorous results, W.A. Benjamin, Inc., New York (1969).

[39] D. Ruelle (1963): *Cluster Property of the Correlation Functions of Classical Gases*, Rev. Mod. Phys., **36**, 580–584.

[40] Scott, A.; Sokal, A.: *The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma.* J. Stat. Phys. **118** (2005), no. 5-6, 1151–1261.

[41] Sokal, A.: *Bounds on the complex zeros of (di)chromatic polynomials and Potts-model partition functions.* Combin. Probab. Comput. **10** (2001) 41-77

[42] J. B. Shearer, *On a problem of Spencer*, Combinatorica **5** (1985), 241245.

[43] Simon B.: The Statistical Mechanics of Lattice Gases, Volume 1, Princeton University Press (1993)

[44] S. G. Williamson (1985): *Combinatorics for computer science.* Computers and Math Series. Computer Science Press, Rockville, MD, 1985.

[45] S. A. Yuhjtman (2015): *A sensible estimate for the stability constant of the Lennard-Jones potential*, J. Stat. Phys. **160**, no. 6, 1684–1695.

Index