Brazilian Journal of Probability and Statistics
2015, Vol. 29, No. 2, 211–281
DOI: 10.1214/14-BJPS270

From equilibrium to nonequilibrium statistical mechanics. Phase transitions and the Fourier law

Errico Presutti
INFN

Abstract. These are the notes of my lectures at EBP, August 2013. I have added some proofs which being of more technical nature have been omitted in the talks. The notes and the lectures are based on a course on atomistic and continuous descriptions of matter which I gave a few years earlier in Sperlonga, Italy. In the present notes, I have tried to underline the more probabilistic aspects of the theory.

I am afraid I have not been able to reproduce in the written notes the very lively atmosphere of the talks. Many old friends of mine were attending the lectures and helped me a lot with questions, comments and criticism, it was a pleasure for me to speak at EBP and I hope also the audience enjoined all that. In particular, I want to renew my deepest thanks to Stefan Luckhaus who is undoubtedly the best help for a speaker to have in the audience.

These notes are divided into four chapters, like in the lectures I have mostly avoided proofs trying to give qualitatively the main ideas of the theory. The only exception is in the second lecture of these notes where I have given more details on the proof of phase transitions in the canonical Ising model. This partly for completeness and partly because I have been asked by several people for details after the talk.

I am not very good with bibliography so I just quoted some of the papers I am most familiar with, the reader will forgive me, I hope, for the many omissions.

I conclude these preface by renewing my warmest thanks to the people who attended EBP and in particular to Maria Eulalia Vares for inviting me at EBP, for the nice words when she introduced my lectures but especially for the very long friendship (in the past and hopefully in the future).

Contents

Key words and phrases. Phase transitions, Ising model, Kac potentials, free energy functionals, stochastic evolutions, nonequilibrium stationary measures.
Received November 2014; accepted November 2014.

1 Lecture 1

Ideas, methods and techniques of probability theory have an important role in physics and in particular in statistical mechanics. They have also important ap-

plications in PDEs, in calculus of variation and in many other branches of mathematics, their influence is also evident in many applications to biological and social models. I will try in these lectures to underline some of these aspects discussing equilibrium and nonequilibrium problems in Statistical Mechanics.

In these lectures, I will concentrate on some very special topics and by no means the reader should regard the following notes as a text of Statistical Mechanics. In the first part devoted to equilibrium, I will focus on phase transitions, one of the most striking aspects of the theory which has been and still is object of intensive research.

Let me start with some examples from different contexts just to give the flavor of what we are going to do. The first one is a classical problem in probability. Consider a sequence of i.i.d. random variables x_i with mean zero. Call their partial sum

$$S_N = \sum_{i=1}^{N} x_i.$$

Let $\alpha > 0$, $0 < \varepsilon < \alpha$ and

$$p_{\varepsilon,\alpha,N} = P\left[\frac{S_N}{N} \in (\alpha - \varepsilon, \alpha + \varepsilon)\right].$$

Which are the typical sequences which realize the above large deviations event in the asymptotic when $N \to \infty$ and after $\varepsilon \to 0$? We may imagine two possible scenarios: in the first one most x_i contribute with a positive bias, in the second one just a few x_i are responsible for the deviation while the others behave "normally". Which one of the two scenarios appears depends on the tail of the distribution of the variables.

Similarly in an economic context, the question is whether some additional "wealth" is redistributed among all individuals or it goes to just a few of them (as it often seems to be the case). Going back to physics consider an elastic bar that we pull at the extremes trying to elongate it. The deformation looks homogeneous when strain and stress are small, but if we pull too hard anelastic effects become eventually important and the bar breaks: instead of being deformed homogeneously, there is a large deformation somewhere (a fracture) while elsewhere the bar is essentially at rest.

Consider two populations which live in an environment which is rich. They will then grow in harmony till when the resources are no longer sufficient and a conflict erupts with one of the two growing at the expenses of the other.

We are now ready for phase transitions! It is far from intuitive (and against philosophical theories of the ancient times) that gas, liquid and solid are only "accidental" attributes of matter, the same substance (same atoms and molecules) may appear in its solid liquid or vapor phase by changing the external conditions; you may change dress but you are still the same person. We know from physics

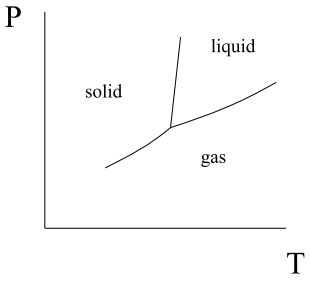

Figure 1 *The P–T (pressure–temperature) phase diagram of a simple fluid.*

(see, e.g., Chapter viii, Section 77 of the Landau and Lifschitz book on Statistical Physics (Landau and Lifshitz (1999))) that a simple, one component fluid (where we neglect the internal structure of the molecules which are then regarded as identical particles) has a phase diagram which looks like the one in Figure 1.

The thermodynamic state of the system is completely determined by the value of the pressure P and the temperature T at all points of the quadrant away from the lines shown in the diagram. Such lines separate regions where the system is in its gas, liquid and solid phases and on such lines different phases coexist, gas–solid, liquid–solid and gas–liquid. When crossing such curves moving from one phase region to the other, we see a phase transition, the curves are thus called phase transition or phase-coexistence curves.

Phase transitions may be very complex but here I restrict to the simpler case of "phase transitions of first order with order parameter the (mass) density". This simply means that there is a "forbidden interval" of densities, say (ρ', ρ''), so that if we put a mass $\rho|\Lambda|$ of fluid in the region Λ ($|\Lambda|$ the volume of Λ) with $\rho \in (\rho', \rho'')$ ("canonical constraint"), then the fluid separates into a part with density ρ' and another one with density ρ''. It does not exist an equilibrium state with homogeneous density ρ. Thus, if we "move in Λ" we go from one phase (with density ρ') to another phase (with density ρ'') and we see a "phase transition".

Existence of phase transitions is an experimentally well established fact. Much less settled is the question whether statistical mechanics is able to reproduce diagrams like the one in Figure 1 or even the existence of phase transitions as described above. Indeed a complete derivation is still an open problem, one among the most important in statistical mechanics. Does any reasonable pair interaction produce a phase transition? which are the relevant features? what is the origin of the instability which leads from an initial homogeneous state with forbidden density ρ to the final equilibrium state with coexisting phases?

1.1 The mesoscopic theory

A good feeling for what is going on can be gained by studying the problem at the mesoscopic level, intermediate between atomistic and macroscopic, we shall

discuss later how the mesoscopic theory can be derived from microscopics, but here we take it as our primitive notion. Let us then make a short detour to recall the foundations of the theory (in the particular case, we are interested in).

As in any continuum theory, points r in the physical space $\mathbb{R}^d$ are actually representative of very large "boxes" with very many particles in it. We suppose that the temperature is kept fixed to a constant value T throughout the whole body and that the state at each r is fully described by the local mass density $\rho(r) \geq 0$; we suppose the particles to be small impenetrable spheres so that for all r, $\rho(r) < \rho_c$, $\rho_c > 0$ the "close packing" density.

The states of the system are then described by functions $\rho(r)$, $r \in \mathbb{R}^d$, with values in $[0, \rho_c)$, $\rho(r)$ the local mass density at r. Let us now suppose that the fluid is contained in a region Λ so that $\rho(r) = 0$ for $r \notin \Lambda$, or, more conveniently, states are functions $\rho(r)$ with $r \in \Lambda$. To avoid problems about the interaction with the walls of Λ we take Λ a torus, equivalently we restrict to functions $\rho(r)$ on $\mathbb{R}^d$ which are periodic with periodic cell the cube Λ.

The basic quantity which specifies the nature of the system is the free energy of a state and following Ginzburg–Landau we suppose that the free energy of the state ρ is

$$F_\Lambda^{\mathrm{gl}}(\rho) = \int_\Lambda \{W_\beta(\rho(r)) + |\nabla \rho(r)|^2\}\, dr, \tag{1.1}$$

where $\beta = 1/(kT)$, k the Boltzmann constant, T the absolute temperature and

$$\begin{aligned} W_\beta(u) &= -\frac{\alpha}{2}u^2 - \frac{1}{\beta}\big(-u(\log u - 1) + u\log(1 - ua)\big), \\ 0 &< u < \rho_c = a^{-1}. \end{aligned} \tag{1.2}$$

The first term takes into account the energy among particles which are in the box represented by the point r: the minus sign is because we suppose the interaction attractive, the factor u^2 "counts" the number of pairs when the particles density is u and we suppose that each pair contributes by a factor proportional to $\alpha > 0$. The last bracket in (1.2) multiplied by the constant k is the local entropy density when the local density is u, we shall be back later on this basic formula for the entropy which goes back to the original works by Boltzmann. Thus, $W_\beta(u)$ is energy minus T times the entropy and it has the meaning of a local free energy density. The last term in (1.1) takes into account the residual interaction not taken into account by W_β, its effect is to penalize variations of $\rho(r)$.

The basic axiom of thermodynamics is that the free energy is minimal at equilibrium. Thus, the theory postulates that the thermodynamic free energy $f_\beta(u)$ of our system when its total mass density is u is given by

$$f_\beta(u) := \liminf_{\Lambda \to \mathbb{R}^d} f_{\beta,\Lambda}(u), \tag{1.3}$$

$$f_{\beta,\Lambda}(u) := \frac{1}{|\Lambda|} \inf\Big\{ F_\Lambda^{\mathrm{gl}}(\rho) \Big| \rho : \int_\Lambda \rho(r)\, dr = u|\Lambda| \Big\}. \tag{1.4}$$

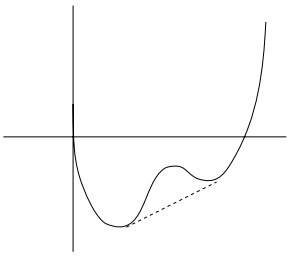

Figure 2 *The graphs of $W_\beta(u)$ and $W_\beta^{**}(u)$ (dashed line) for β large.*

The "thermodynamic limit" $\Lambda \to \mathbb{R}^d$ in (1.3) is to get rid of finite volume effects.

We have several possible ways to define the equilibrium states. Given a sequence Λ_n increasing to $\mathbb{R}^d$, it could be the sequence of minimizers of (1.4); or else any minimizing sequence $\rho^{(n)}$, namely such that

$$\begin{aligned} &\frac{1}{|\Lambda_n|}\int_{\Lambda_n} \rho^{(n)}\, dr = u \qquad \text{for all } n \quad \text{and} \\ &\lim_{n\to\infty} \frac{1}{|\Lambda|} F^{\mathrm{gl}}_{\Lambda_n}(\rho_{\Lambda_n}) = f_\beta(u) \end{aligned} \tag{1.5}$$

or maybe (in its weakest form) a particular minimizing sequence.

We shall next compute the free energy $f_\beta(u)$ and prove that for β large enough there is a phase transition.

The existence for β large enough of a phase transition will be first proved (see Theorem 1.1 below) in the thermodynamic sense by showing that the free energy density is linear in some interval (ρ', ρ'') and then (see Theorem 1.2) by proving that (ρ', ρ'') is a forbidden density interval, in the sense that the equilibrium state (see (1.5)) is in a large fraction of Λ either close to ρ' or to ρ''.

As we shall see, the origin of the phase transition comes from the loss of convexity of $W_\beta(u)$ when β is large, see Figure 2. We shall use the following features of the function $W_\beta(u)$ valid for all β large enough. Call $W_\beta^{**}(\cdot)$ the convexification of $W_\beta(\cdot)$ (i.e., the largest convex function below the graph of $W_\beta(\cdot)$). Then:

> There is an open interval (ρ', ρ'') so that $W_\beta^{**}(u) = W_\beta(u)$ for $u \notin (\rho', \rho'')$, $W_\beta^{**}(u) < W_\beta(u)$ for $u \in (\rho', \rho'')$ and $W_\beta''(u) > 0$ for $u \notin (\rho', \rho'')$.

We shall see in the next theorem that there are nonhomogeneous states with total density u such that $F^{\mathrm{gl}}_\Lambda(\rho)/|\Lambda| \approx W_\beta^{**}(u)$. Then the inf in (1.3) is achieved on nonhomogeneous states and from this the existence of a forbidden interval will follow.

Theorem 1.1. *The liminf in* (1.3) *is actually a limit and*

$$f_\beta(u) = W_\beta^{**}(u). \tag{1.6}$$

Proof. It is obtained by establishing lower and upper bounds and proving that they become equal in the limit.

Lower bound. Recalling that the convexification of a function f is the largest convex function which is below f, we have $W_\beta \geq W_\beta^{**}$ and therefore $F_\Lambda^{\rm gl}(\rho) \geq \int_\Lambda W^{**}(\rho)$, as the gradient term is nonnegative (we may not indicate the measure in the integral when clear from the context, as we have done above). Since W^{**} is convex we can use the Jensen inequality to write $\int_\Lambda W_\beta^{**}(\rho) \geq |\Lambda| W_\beta^{**}(u)$, where $\int_\Lambda \rho = u|\Lambda|$. In conclusion: $F_\Lambda^{\rm gl}(\rho) \geq |\Lambda| W_\beta^{**}(u)$.

Upper bound. When u is such that $W_\beta(u) = W_\beta^{**}(u)$ we simply write (for all Λ)

$$\inf\left\{ F_\Lambda^{\rm gl}(\rho) \Big| \int_\Lambda \rho = u|\Lambda| \right\} \leq F_\Lambda^{\rm gl}(u\mathbf{1}_\Lambda) = |\Lambda| W^{**}(u).$$

If instead $W_\beta(u) > W_\beta^{**}(u)$ then $u \in (\rho', \rho'')$ (which are the x-coordinates of the points where the dashed line starts and ends in Figure 2). ρ' and ρ'' are such that $W_\beta(\rho') = W_\beta^{**}(\rho')$, $W_\beta(\rho'') = W_\beta^{**}(\rho'')$ and W_β^{**} is linear in (ρ', ρ''). Let $p \in (0,1)$ be such that $u = p\rho' + (1-p)\rho''$, then $W_\beta^{**}(u) = pW_\beta(\rho') + (1-p)W_\beta((\rho'')$. We next construct a sequence ρ_Λ by taking a regularization $\rho(\cdot)$ (with the required mass $u|\Lambda|$) of the function $\rho'\mathbf{1}_{\Lambda'} + \rho''\mathbf{1}_{\Lambda''}$ where Λ' is (e.g.) a rectangle in Λ and Λ'' its complement (such that $|\Lambda'|/|\Lambda| = p$), see Figure 3 where we consider a linear interpolation between ρ' and ρ'' in an interval of length 1.

Then if L is the side of Λ,

$$F_\Lambda^{\rm gl}(\rho_\Lambda) \leq |\Lambda| W^{**}(u) + cL^{d-1}$$

so that $F_\Lambda^{\rm gl}(\rho_\Lambda)/|\Lambda| \to W_\beta^{**}(u)$, hence

$$\limsup_{\Lambda \to \mathbb{R}^d} \frac{1}{|\Lambda|} \inf\left\{ F_\Lambda^{\rm gl}(\rho) \Big| \int_\Lambda \rho = u|\Lambda| \right\} \leq W_\beta^{**}(u). \qquad \square$$

Next theorem studies the equilibrium states defined in (1.5) when there is a phase transition, proving phase separation.

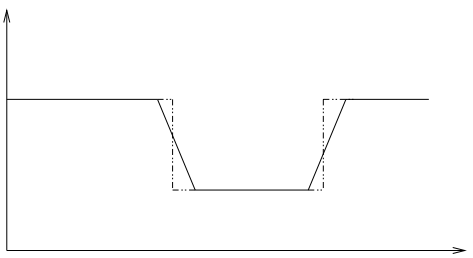

Figure 3 *The graph with continuous line is the regularization of the one with dashed line.*

Theorem 1.2. *Let $u = p\rho' + (1-p)\rho''$, $p \in (0,1)$, and let $\rho^{(\Lambda)}$ be any minimizing sequence, see (1.5), so that*

$$\lim_{\Lambda\to\mathbb{R}^d} \left| f_{\beta,\Lambda}(u) - \frac{1}{|\Lambda|} F_\Lambda\big(\rho^{(\Lambda)}\big) \right| = 0. \tag{1.7}$$

Then for any $\zeta > 0$

$$\lim_{\Lambda\to\mathbb{R}^d} \frac{1}{|\Lambda|} \big|\{r \in \Lambda : |\rho^{(\Lambda)}(r) - \rho'| > \zeta, |\rho^{(\Lambda)}(r) - \rho''| > \zeta\}\big| = 0, \tag{1.8}$$

$|\{\cdot\}|$ *the Lebesgue volume of* $\{\cdot\}$.

Proof. Define $G_\Lambda(\rho) := \int_\Lambda \{g(\rho) + (\frac{d\rho}{dx})^2\}$, where

$$g(u) := W_\beta(u) - \lambda u - c, \qquad \lambda := \frac{W_\beta(\rho'') - W_\beta(\rho')}{\rho'' - \rho'}, \qquad c := W_\beta\big(\rho'\big) - \lambda\rho'.$$

By (1.7)–(1.3)–(1.6) and writing $u = p\rho' + (1-p)\rho''$

$$\lim_{\Lambda\to\mathbb{R}^d} \frac{1}{|\Lambda|} G_\Lambda\big(\rho^{(\Lambda)}\big) = W_\beta^{**}(u) - \lambda u - c = pg\big(\rho'\big) + (1-p)g\big(\rho''\big) = 0$$

because $g(\rho') = g(\rho'') = 0$. Moreover, $g(u) > 0$ for $u \neq \rho', \rho''$ and for any $\zeta > 0$

$$m(\zeta) := \inf\big\{g(u) | u : |u - \rho'| \geq \zeta, |u - \rho''| \geq \zeta\big\} > 0.$$

Since $G_\Lambda(\rho^{(\Lambda)}) \geq |\Lambda| m(\zeta) |\{r \in \Lambda : |\rho^{(\Lambda)}(r) - \rho'| \geq \zeta, |\rho^{(\Lambda)}(r) - \rho''| \geq \zeta\}|$ and $G_\Lambda(\rho^{(\Lambda)})/|\Lambda| \to 0$ we then get (1.8). □

The proof of Theorem 1.1 suggests that the two phases present in the equilibrium state occupy regular regions with the interface between them a smooth surface. This is in fact true in general as it can be proved that the shape of the equilibrium interface is obtained by minimizing the surface tension over all possible interfaces (with the canonical constraint that the two phases occupy given fractions of the total volume, Wulff shape problems): it is such a minimization procedure which brings in the regularity properties of interface.

1.2 Zero temperature

We next move to microscopics and begin the analysis of phase transitions by studying systems at zero temperature. The problem becomes then considerably simpler and the analysis can go quite far, as we shall see. The framework is classical mechanics as we shall neglect all quantum effects: they are indeed very relevant at low temperatures and hence we are far from realistic. Reason is twofold, this is an introduction to phase transitions and we want the analysis as simple as possible.

Our systems are made of identical point particles which interact pairwise via a potential repulsive at the origin and with an attractive tail at large distances, the

prototype is the Lennard–Jones potential (see below). The basic axiom of equilibrium statistical mechanics at 0 temperature is that the equilibrium states are "ground states", namely configurations which minimize the energy (velocities are thus set equal to 0 and particle configurations will be described only by the positions of the particles). Our aim here is to prove that the ground state energy density as a function of the particle density has a phase transition with a "forbidden density interval".

1.2.1 *Inter-molecular forces.* Particle configurations, denoted by q, are countable subsets of $\mathbb{R}^d$, and even though we are ultimately interested in configurations on the whole space with infinitely many particles (thermodynamic limit) for the moment we restrict to configurations with finite cardinality, writing $|q|$ for the cardinality (or number of particles) of q. Thus, $q = (q_1, \dots, q_n)$, $|q| = n \in \mathbb{N}$, $q_i \in \mathbb{R}^d$, the order in the sequence being immaterial as the particles are identical. The energy of q is

$$H(q) = \frac{1}{2} \sum_{i \neq j} V(q_i, q_j), \tag{1.9}$$

where $V(r, r') = V(r', r)$ is the pair interaction between points at r and r' (notice that (1.9) is invariant under permutations of the sequence $(q_1, \dots, q_n)$, as all physical observables it is independent of the order in the sequence q).

Inter-molecular forces are often described by Lennard–Jones potentials, see Figure 4, where writing by an abuse of notation $V(r, r') = V(|r - r'|)$

$$V(R) = aR^{-12} - bR^{-6}, \qquad a, b > 0. \tag{1.10}$$

The positive divergence at the origin ensures stability of matter, that is, that the energy per particle is bounded from below. If on the contrary, the interaction was negative and bounded away from zero in a neighborhood of the origin then by putting all the particles of q in that set we would get $H(q) \geq -c|q|^2$ so that the energy per particle $H(q)/|q|$ would diverge to $-\infty$ and matter would not be stable (with ground states configurations having an infinite local density somewhere).

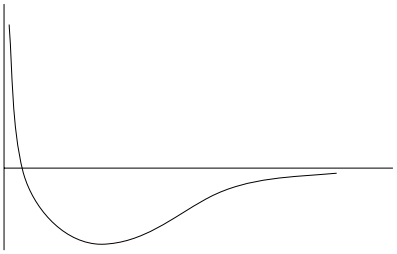

Figure 4 *Lennard–Jones potential.*

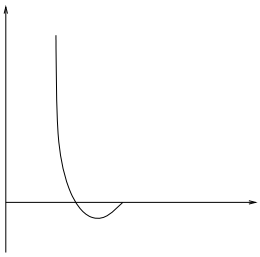

Figure 5 *Hard core plus a short range potential with attractive tail.*

The attractive tail of the Lennard–Jones potential is responsible for the occurrence of a phase transition, as we shall see.

A simpler interaction with similar features is the one in Figure 5, where V is infinite when $|r - r'| \le R_0$, $R_0 > 0$ the hard core length. It is then evident that $H(q) \ge -b|q|$, $b > 0$, and that in the ground states the local density is finite; this happens also for the Lennard–Jones interaction but it is not as easy to see.

1.2.2 *Internal energy, ground states.* By the laws of thermodynamics, the equilibrium states are those which minimize the free energy. Since we are at zero temperature, the equilibrium states are those which minimize the internal energy. Our first task is therefore to define properly the internal energy density $e(\rho)$ when the system has particle density equal to ρ.

Let Λ be a cube in $\mathbb{R}^d$ (we shall eventually restrict to $d = 2$), $|\Lambda|$ its volume, $\mathcal{X}_\Lambda$ the space of all particle configuration q with finitely many particles all in Λ. Set first

$$e_\Lambda\left(\frac{n}{|\Lambda|}\right) = \inf_{q\in\mathcal{X}_\Lambda\,:\,|q|=n} \frac{H(q)}{|\Lambda|} \tag{1.11}$$

(supposing as we do that the interaction is lower semi-continuous, the inf above is actually a minimum). We then define for any $\rho > 0$ and any increasing sequence of cubes Λ,

$$e(\rho) := \liminf_{\Lambda\to\mathbb{R}^d;\, n/|\Lambda|\to\rho} e_\Lambda\left(\frac{n}{|\Lambda|}\right). \tag{1.12}$$

We postulate that $e(\rho)$ is the internal equilibrium energy density when the particle density is ρ.

As in the mesoscopic theory, we fix an increasing sequence Λ_n of regions invading $\mathbb{R}^d$ and call equilibrium state at zero temperature either (i) the sequence of configurations where the inf in (1.11) is achieved; or else (ii) any sequence which in the limit achieves $e(\rho)$ (as defined in (1.12)); or finally, (iii), a particular sequence in the family (ii).

1.2.3 *General properties of the internal energy.* The internal energy defined in (1.12) has the good convexity properties that thermodynamics requires and which extend to the free energy at nonzero temperature.

Theorem 1.3. *The liminf in* (1.12) *is actually a limit which is a continuous convex function of* ρ *bounded from below.*

So far the analysis was pretty general, the specificity of the Lennard–Jones potentials appears next when we characterize the graph of $e(\rho)$. Since we know from Theorem 1.3 that $e(\rho)$ is convex, then to mimic the proofs for the Ginzburg–Landau functional we would need to show that the internal energy restricted to the translation invariant states is nonconvex. As we are working with atomistic rather than continuum systems, it is not a priori clear the meaning of translation invariant particle configurations. The zero temperature case is simpler as we may and shall replace translation invariance by periodicity.

Before entering into the mathematics of the problem, we discuss the physical meaning of a forbidden density interval in the present context of states at zero temperature. At zero temperature, bodies are in the solid phase with the atoms arranged in some crystalline structure. By applying stresses, we can modify the density of the body, in particular by stretching the crystal we increase the interatomic distances and the density becomes smaller. The phenomenon continues till a critical value of the density, after that the crystal becomes unstable, it does not respond anymore to our stretching and it "breaks" reducing the mesh to a value which corresponds to the critical density. Thus, densities smaller than critical are realized by a piece of crystal at a larger density (the critical one) with void around it: a fraction of the region is left empty, the other one is occupied by a crystal with the critical density, a picture which fits with the one about forbidden density intervals discussed in the Introduction with $\rho' = 0$ and ρ'' the critical density.

Purpose of our analysis is to show that this can be seen also at the mathematical level.

1.3 The Ising model

To get a feeling of what is going on, let us consider the much simpler case of a nearest neighbor "ferromagnetic" lattice gas where configurations are made of point particles whose positions are restricted to the lattice $\mathbb{Z}^2$ (for simplicity we discuss the two-dimensional case). Thus the phase space (of all particle configurations) is $\{0,1\}^{\mathbb{Z}^2}$. A particle configuration is $\eta = \{\eta(x), x \in \mathbb{Z}^2\}$ and $\eta(x) = 1$ if the configuration η has a particle at x and $\eta(x) = 0$ otherwise. The exclusion rule mimics the hard core interaction in Figure 5. The attractive part (called above ferromagnetic) is represented by a negative nearest neighbor pair interaction so that the energy of a configuration with finitely many particles is

$$H(\eta) = -\frac{J}{2} \sum_{|x-y|=1} \eta(x)\eta(y), \qquad J > 0. \tag{1.13}$$

Evidently the density ρ in the lattice gas varies in $[0, 1]$ and we have:

$$e(0) = 0, \qquad e(1) = -2J. \tag{1.14}$$

The internal energy $e(\rho)$, being convex, must then satisfy the inequality

$$e(\rho) \le -2J\rho \tag{1.15}$$

as the line $-2J\rho$ connects the energies at 0 and 1.

Equation (1.15) actually holds with equality, as it follows from showing that $e(\rho) \ge -2J\rho$. The lower bound holds even before the thermodynamic limit:

$$H(\eta) \ge -2J \sum \eta(x). \tag{1.16}$$

Equation (1.16) follows from (1.13) once we attribute an energy $-J/2$ to each one of the two particles placed in nearest neighbor sites. Then a particle has $\ge -4(J/2)$ energy (as each site has 4 nearest neighbors) hence the total energy is $\ge -2JN$ if there are N particles, hence (1.16).

The bound (1.16) can also be proved using contours. Let Λ be a torus and η a configuration in Λ. Set

$$\sigma(x) := 2\eta(x) - 1, \qquad \eta(x) = \frac{1 + \sigma(x)}{2} \tag{1.17}$$

$\sigma(x) \in \{-1, 1\}$ is an "Ising spin". With such a change of variables

$$\begin{aligned} H(\eta) + 2J \sum \eta(x) &= -\frac{J}{8} \sum_{|x-y|=1} \big(\sigma(x)\sigma(y) - 1\big) \\ &= \frac{J}{4} \sum_{|x-y|=1} \mathbf{1}_{\sigma(x)\neq\sigma(y)}. \end{aligned} \tag{1.18}$$

Thus each "broken bond" (x, y) ($|x - y| = 1$, $\sigma(x) \neq \sigma(y)$) contributes to the energy, adding a finite amount $J/4$ to the minimal value $-2JN$, $N = \sum \eta(x)$.

Let Λ_n be an increasing sequence of tori which invades $\mathbb{Z}^2$, L_n their sides length. Let u_n be of the form $k/|\Lambda_n|$, k a nonnegative integer, and such that $u_n \to u$ as $n \to \infty$. Call $\eta^{(n)}$ a minimizer of the energy in the set $\mathcal{X}_n := \{\sum \eta(x) = u_n|\Lambda_n|\}$. Then

$$H\big(\eta^{(n)}\big) \le -2Ju_n|\Lambda_n| + \frac{J}{4}(4L_n). \tag{1.19}$$

Equation (1.19) is proved by bounding the left-hand side by the energy of a configuration, that we call η^* and choose in the following way. Order the points of Λ_n lexicographically, set $\eta^*(x) = 0$ in the first k points and $=1$ in the remaining points. If $k = |\Lambda_n| - u_n|\Lambda_n|$, then $\eta^* \in \mathcal{X}_n$. Since the number of broken bonds in η^* is not larger than $4L_n$, $H(\eta^*)$ is bounded as on the right-hand side of (1.19) which is therefore proved.

By (1.19), the sequence $\eta^{(n)}$ has a number of broken bonds $\leq cL_n$ and therefore if we partition the regions Λ_n into cubes of fixed side ℓ, then the number of cubes which contain a broken bond is $\leq cL_n$. Hence, the fraction of cubes with broken bonds goes like

$$\leq \frac{cL_n}{L_n^2/\ell^2} = c\frac{\ell^2}{L_n}.$$

An analogous argument applies to any minimizing sequence so that in any such sequence the fraction of cubes totally occupied or totally empty goes to 1 and the lattice gas has a phase transition with forbidden interval [0, 1].

1.4 Triangular lattice configurations

Ansatz: *the equilibrium states of identical point particles with Lennard–Jones interactions are configurations where the particles form a triangular lattice*. Such configurations, see Figure 6, will be denoted by $\mathbb{T}_R$ with R the lattice mesh (i.e., the distance between nearest neighbor points).

The validity of the statement will be proved a posteriori, there are however some considerations which make the ansatz reasonable. Physics tells us that crystals are obtained by repeating periodically a basic cell and since in our system particles are identical we may conjecture that the particles are arranged in a periodic lattice. Moreover, if there is an "optimal distance" R between nearest neighbor particles, the best lattice is the one where the lattice mesh is R and the coordination number (the number of nearest neighbors of any given point) maximal.

Such a request selects in $d = 2$ dimensions the triangular lattice. The square lattice $\mathbb{Z}^2$, for instance, has 4 nearest neighbors, while in the triangular lattice $\mathbb{T}_R$ there are 6 neighbors. To see that we cannot do better, we start from a point in the lattice, for simplicity the origin. Its nearest neighbors are all in the circle of radius R and center the origin. Thus the maximal coordination number cannot be

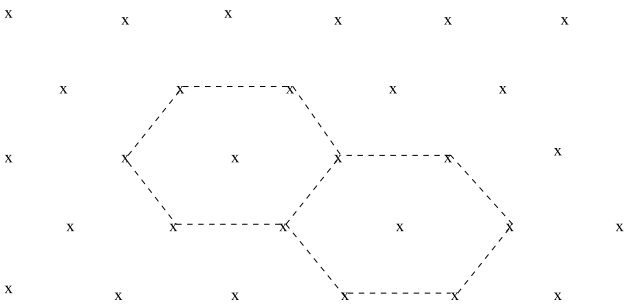

Figure 6 *A triangular configuration induces a paving of* $\mathbb{R}^2$ *into hexagons*, *the sides of the hexagon have equal length*, *equal to the distance of the vertices from the center of the hexagon*. *Each hexagon splits into* 6 *equilateral triangles*.

larger than the maximal number of points that we can put on this circle at distance $\geq R$ from each other, as the lattice must have mesh R. The angle spanned by two points on the circle at distance R from each other is $\pi/3$ (because these two points and the origin form an equilateral triangle) hence there are exactly 6 points on the circle at consecutive distances R.

According to the above ansatz the relevant configurations are triangular lattice configurations, which suggests to introduce the triangular lattice internal energy

$$e_T(\rho) := \lim_{\Lambda\to\mathbb{R}^2} \frac{H(\mathbb{T}_{R(\rho)}\cap\Lambda)}{|\Lambda|}, \qquad R(\rho): \lim_{\Lambda\to\mathbb{R}^2;} \frac{|\mathbb{T}_{R(\rho)}\cap\Lambda|}{|\Lambda|} = \rho \tag{1.20}$$

with $R(\rho) = c'\rho^{-1/2}$, $c' > 0$. In fact, the inverse $\rho(R)$ can be computed as follows. We realize $\mathbb{T}_R$ by putting particles on horizontal lines: on each line consecutive particles have distance R from each other; the particles on a line are shifted by $(R/2)$ with respect to those of the line below; the distance D between consecutive lines is $D^2 + (R/2)^2 = R^2$. This is a triangular configuration with mesh R. The density ρ of particles in $\mathbb{T}_R$ is the same as that in a horizontal strip of height D which has in the middle one of the lines with particles, hence $\rho = 1/(RD)$.

As $e_T(\rho)$ will play the role that $W_\beta(\rho)$ had in the Ginzburg–Landau phase transition discussed in the Introduction, we need to prove that $e_T(\rho)$ is nonconvex, indeed:

Proposition 1.4. *There are a' and b' both strictly positive so that*

$$e_T(\rho) = a'\rho^7 - b'\rho^4. \tag{1.21}$$

Proof. It is readily seen from (1.20) that, supposing $0 \in \mathbb{T}_R$,

$$e_T(\rho) = \frac{\rho}{2} \sum_{x\in\mathbb{T}_{R(\rho)}, x\neq 0} V\big(|x|\big). \tag{1.22}$$

In fact the right-hand side is the product of the density ρ times the energy per particle, that is, $1/2$ the interaction energy of the particle at the origin with all the others. The factor $1/2$ is to avoid counting twice a same pair of particles.

To compute the r.h.s. of (1.22), we use a scaling argument:

$$\sum_{x\in\mathbb{T}_R, x\neq 0} V\big(|x|\big) = \frac{aR^{-12}}{2} \sum_{x\in\mathbb{T}_1, x\neq 0} |x|^{-12} - \frac{bR^{-6}}{2} \sum_{x\in\mathbb{T}_1, x\neq 0} |x|^{-6} \tag{1.23}$$

hence (1.21) after recalling that $R(\rho) = c'\rho^{-1/2}$, $c' > 0$. □

Thus $e_T(\rho)$ is nonconvex, its graph and the graph of its convexification $e_T^{**}(\rho)$ are as in Figures 7 and 8.

This will prove that there is a phase transition with density interval $(0, \rho_c)$ if the energy $e_T^{**}(\rho)$ is the internal energy of the system, a statement which follows from Theil's theorem (whose proof can be found in the literature).

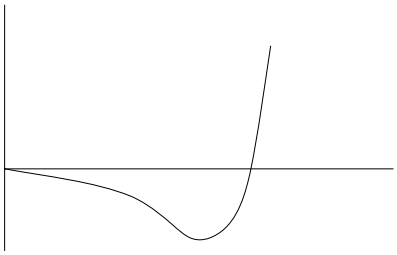

Figure 7 *$e_T(\rho)$, the energy of the triangular configuration with density ρ.*

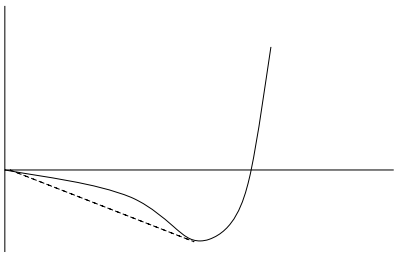

Figure 8 *The graph of $e_T^{**}(\rho)$, linear in $[0, \rho_c]$.*

1.5 Bibliographical remarks

Variational problems play a fundamental role in probability, for instance, in large deviations. In this lecture, I have underlined their role in statistical mechanics and thermodynamics. I refer to the book by Braides (2002), for mathematical foundations and to Chapter 6 of my book (Presutti (2008)) for the physical interpretation of several free energy functionals, see also the last two lectures of these notes.

Theil's theorem has been published in Theil (2006). Previous results are due to Heitmann and Radin (1980), Radin (1981), they describe the infinitely extended crystal. In bounded regions, at the phase transition, the crystal occupies only a part of the domain, the determination of its shape is the so-called Wulff problem. This has been studied by Au Yeung, Friesecke and Schmidt (2009) and by Friesecke, see http://www.acmac.uoc.gr//CKM2011/talks/Friesecke.pdf.

2 Lecture 2

The strength and beauty of equilibrium statistical mechanics is that it gives a very simple and explicit formula for the equilibrium states at positive temperatures: *the probability of a particle configuration with energy H is proportional to $e^{-\beta H}$, $\beta = 1/kT$, k the Boltzmann constant, T the absolute temperature*. Statistical mechanics gives also a formula for free energy and pressure in terms of the partition function, namely the sum, or integral, over all states of the Gibbs factor $e^{-\beta H}$. The

theory therefore constructs a bridge between microscopics (inter-molecular forces) and macroscopics (thermodynamic potentials). Details are given later after some heuristic justification of the Gibbs formulation based on the Boltzmann hypothesis in an information theory context.

2.1 The Boltzmann hypothesis and the Shannon entropy

2.1.1 *The Shannon entropy.* Consider a channel which transmits messages with a finite alphabet Ω; *we want to compute its capacity by counting how many messages can be emitted by a source with* "*parameters* $f - \phi$". By this, we mean the following: f is a real valued function on Ω, $\phi \in (\min f, \max f)$; the messages which can be sent into the channel are

$$\left\{(\omega_1, \dots, \omega_N) \in \Omega^N : A_N(\omega_1, \dots, \omega_N) := \left| \frac{1}{N} \sum_{i=1}^{N} f(\omega_i) - \phi \right| \le \delta \right\}$$

δ a positive parameter. Calling

$$K_\delta(N) = \operatorname{card}\big\{(\omega_1, \dots, \omega_N) \in \Omega^N : A_N(\omega_1, \dots, \omega_N) \le \delta\big\} \tag{2.1}$$

the capacity of the channel is then defined as:

$$\lim_{\delta \to 0} \lim_{N \to \infty} \frac{\log K_\delta(N)}{N}. \tag{2.2}$$

Instead of going into combinatorics and Stirling formulas it is more instructive for the applications to statistical mechanics to use a probabilistic approach. Let $p(\omega) > 0$ be a probability on Ω; write the identity

$$K_\delta(N) = \sum_{\omega_1, \dots, \omega_N} \mathbf{1}_{A_N \le \delta} \frac{p(\omega_1) \cdots p(\omega_N)}{p(\omega_1) \cdots p(\omega_N)}. \tag{2.3}$$

The whole trick will be to choose properly $p(\omega)$.

Calling $Z_b = \sum_{\omega \in \Omega} e^{bf(\omega)}$, we will see that the "right choice" is

$$p(\omega) = \frac{e^{bf(\omega)}}{Z_b}, \qquad b \text{ such that } \sum_{\omega \in \Omega} p(\omega) f(\omega) = \phi. \tag{2.4}$$

Existence (and uniqueness) of b follows from the fact that $\sum_{\omega \in \Omega} p(\omega) f(\omega)$ is an increasing function of b which converges to $\min f$ and $\max f$ as $b \to \mp\infty$. With the choice (2.4) for $p(\cdot)$, we get from (2.3)

$$K_\delta(N) = \sum_{\omega_1, \dots, \omega_N} \mathbf{1}_{A_N \le \delta} \big[p(\omega_1) \cdots p(\omega_N)\big] e^{-\sum (bf(\omega_i) - \log Z_b)}. \tag{2.5}$$

It is now clear why (2.4) is the good choice: the sum $\sum_{i=1}^{N} bf(\omega_i)$ in the exponent is by (2.1) approximately equal to $b\phi$ because of the condition $A_N \le \delta$, while, as we are going to see, the second equality in (2.4) ensures that the condition

$A_N \le \delta$ is satisfied with P_N probability 1 as $N \to \infty$, P_N the product probability $p(\omega_1) \cdots p(\omega_N)$. Indeed, calling $S := \log Z_b - b\phi$,

$$P_N[A_N \le \delta] e^{(S-b\delta)N} \le K_\delta(N) \le e^{(S+b\delta)N}. \tag{2.6}$$

By the law of large numbers, for any $\delta > 0$ $\lim_{N\to\infty} P_N[A_N \le \delta] = 1$ so that

$$\lim_{\delta\to 0} \lim_{N\to\infty} \frac{\log K_\delta(N)}{N} = S \tag{2.7}$$

and by an explicit computation

$$S(p) := -\sum_{\omega\in\Omega} p(\omega) \log p(\omega) = \log Z_b - b\phi = S. \tag{2.8}$$

In conclusion, the capacity of the channel equals the "Shannon information entropy" $S(p)$ and this is related to the Gibbs probability (2.4).

2.1.2 *Boltzmann hypothesis*. In analogy with the previous example, we consider a lattice gas in a cube Λ, Λ plays the role of the interval $[1, N]$ and the alphabet Ω is now $\{0, 1\}$. The messages are here particle configurations $\eta = \{\eta(x), x \in \Lambda\}$. The Hamiltonian is (extending η to the whole space with $\eta(x) = 0$ whenever $x \notin \Lambda$)

$$H(\eta) = -\frac{1}{2} \sum_{x\neq y} J(x, y)\eta(x)\eta(y), \tag{2.9}$$

where $J(x, y) = J(0, y - x)$ with $J(0, x)$ a symmetric function with compact support (finite range interactions). The analogy with the previous example comes by observing that

$$H(\eta) = \sum_x f(\tau_x \eta), \qquad f(\eta) = -\frac{1}{2} \sum_{x\neq 0} J(0, x)\eta(0)\eta(x), \tag{2.10}$$

τ_x being the shift by x: $(\tau_x\eta)(y) = \eta(y - x)$. Thus the condition that $A_N \le \delta$ becomes

$$\left\{\eta : \left| \frac{H(\eta)}{|\Lambda|} - E \right| \le \delta \right\} \tag{2.11}$$

E an energy density. Calling

$$N_{E,\Lambda,\delta} = \operatorname{card}\left\{\eta : \left| \frac{H(\eta)}{|\Lambda|} - E \right| \le \delta \right\} \tag{2.12}$$

the Boltzmann hypothesis states that the thermodynamic entropy $S(E)$ is given by:

$$S(E) := \lim_{\delta\to 0} \limsup_{n\to\infty} k \frac{\log N_{E,\Lambda_n,\delta}}{|\Lambda_n|}, \tag{2.13}$$

where Λ_n is an increasing sequence of cubes (more general regions can be considered as well but for simplicity we have formulated the hypothesis in terms of cubes) and k is the Boltzmann constant.

We are going to argue that similarly to the case of the Shannon example of the previous subsection, the computation of the entropy $S(E)$ involves Gibbs measures. We start from the trivial identity (valid for any $\beta > 0$)

$$N_{E,\Lambda,\delta} = \sum_{\eta} \mathbf{1}_{|H(\eta)-|\Lambda|E|\le\delta|\Lambda|}$$
$$= \sum_{\eta} \mathbf{1}_{|H(\eta)-|\Lambda|E|\le\delta|\Lambda|} \frac{e^{-\beta H(\eta)}}{Z_{\beta,\Lambda}} \{Z_{\beta,\Lambda} e^{\beta H(\eta)}\}.$$

We proceed by writing upper and lower bounds which hopefully coincide in the thermodynamic limit. We bound the Gibbs factor $e^{\beta H(\eta)}$ in the curly bracket by $e^{\beta(E\pm\delta)|\Lambda|}$ getting that $\frac{\log N_{E,\Lambda,\delta}}{|\Lambda|}$ is bounded by

$$\begin{aligned} &\le \frac{\log Z_{\beta,\Lambda}}{|\Lambda|} + \beta(E+\delta) \\ &\ge \frac{\log Z_{\beta,\Lambda}}{|\Lambda|} + \beta(E-\delta) + \frac{1}{|\Lambda|} \log G_{\beta,\Lambda}(\{|H_\Lambda(\sigma_\Lambda) - |\Lambda|E| \le \delta|\Lambda|\}), \end{aligned} \tag{2.14}$$

where $G_{\beta,\Lambda}(\eta) = \frac{e^{-\beta H(\eta)}}{Z_{\beta,\Lambda}}$ is the Gibbs measure (with 0 boundary conditions). By a sub-additivity argument which is omitted one can prove that the following limit exists

$$\lim_{\Lambda\to\mathbb{Z}^d} \frac{\log Z_{\beta,\Lambda}}{\beta|\Lambda|} =: P_\beta. \tag{2.15}$$

This yields

$$\lim_{\delta\to 0} \limsup_{\Lambda\to\mathbb{Z}^d} \frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} \le \inf_{\beta>0} (\beta P_\beta + \beta E). \tag{2.16}$$

The lower bound goes smoothly if there is β^* such that

$$\lim_{\delta\to 0} \lim_{\Lambda\to\mathbb{Z}^d} \frac{1}{|\Lambda|} \log G_{\beta^*,\Lambda}\left(\left\{\left|\frac{H_\Lambda(\sigma_\Lambda)}{|\Lambda|} - E\right| \le \delta\right\}\right) = 0. \tag{2.17}$$

Then

$$\lim_{\delta\to 0} \liminf_{\Lambda\to\mathbb{Z}^d} \frac{\log N_{E,\Lambda,\delta}}{|\Lambda|} \ge \beta^* P_{\beta^*} + \beta^* E \tag{2.18}$$

which together with (2.16) yields that the limit in (2.13) exists and it is equal to

$$S(E) = \beta^* P_{\beta^*} + \beta^* E = \inf_{\beta>0} (\beta P_\beta + \beta E) \tag{2.19}$$

which is the well-known thermodynamic formula for the entropy in terms of the pressure. Thus (under the assumption that (2.17) holds) we have a formula, (2.15), for the pressure in terms of the partition function associated to the Gibbs measures $G_{\beta^*,\Lambda}$.

The problem is that (2.17) is not true in general. As we will see it holds for E in a set $\mathcal{E}_{\text{erg}}$: for any $E \in \mathcal{E}_{\text{erg}}$ there is a special value of β for which (2.17) holds (we shall briefly discuss later how to proceed when $E \notin \mathcal{E}_{\text{erg}}$). The terminology suggests that the crucial estimate (2.17) is related to an ergodic theorem. Recalling in fact (2.10) the energy is the empirical average of the function f, thus if there is β^* so that (i) the (weak) limit G_{β^*} of $G_{\beta^*,\Lambda}$ as $\Lambda \to \mathbb{Z}^d$ exists, (ii) it is ergodic and (iii) the average $G_{\beta^*}[f] = E$, then it can be seen that (2.17) holds.

In $d = 1$, it is proved that for any β the limit G_β of $G_{\beta,\Lambda}$ exists and it is ergodic and the averages $G_\beta[f]$ span all the possible values of the energy E. In dimensions $d \geq 2$ the statement is no longer true due to phase transitions, but one can decompose any limit point G_β into its ergodic components and in this way it is possible to deal also with energies which do not correspond to ergodic measures. All that is beyond the purposes of this lecture, I refer here to the literature.

2.2 Gibbs measures and thermodynamic potentials

2.2.1 *Lattice gas.* Let us now give the precise definition of Gibbs measures starting from the simpler case of the lattice gas. Let $H(\eta)$ be as in (2.9) and

$$H(\eta|\bar\eta) := H(\eta + \bar\eta) - H(\bar\eta), \qquad \eta + \bar\eta \leq 1 \tag{2.20}$$

the energy of η in the field of $\bar\eta$, η and $\bar\eta$ having disjoint support and finitely many particles (by a limit procedure the definition extends to the case with $\bar\eta$ having infinitely many particles, as we are supposing the interaction to have finite range).

The Gibbs measure in the bounded region Λ with boundary conditions $\bar\eta$ ($\bar\eta$ a configuration in Λ^c), inverse temperature β and chemical potential λ is the probability on $\{0,1\}^\Lambda$

$$\mu_{\beta,\lambda;\Lambda}(\eta|\bar\eta) = Z_{\beta,\lambda;\Lambda}(\bar\eta)^{-1} e^{-\beta[H(\eta|\bar\eta)-\lambda N(\eta)]}, \tag{2.21}$$

where $\eta \in \{0,1\}^\Lambda$, $N(\eta) = \sum_{x\in\Lambda} \eta(x)$ and

$$Z_{\beta,\lambda;\Lambda}(\bar\eta) = \sum_{\eta\in\{0,1\}^\Lambda} e^{-\beta[H(\eta|\bar\eta)-\lambda N(\eta)]} \tag{2.22}$$

called the partition function. The "infinite volume" Gibbs measure also called DLR measure, at inverse temperature β and chemical potential λ is any probability μ on $\{0,1\}^{\mathbb{Z}^d}$ such that for any bounded set Λ the conditional probability of seeing η in Λ given that there is $\bar\eta$ outside Λ is

$$\mu[\eta|\mathcal{F}_{\Lambda^c}](\bar\eta) = \mu_{\beta,\lambda;\Lambda}(\eta|\bar\eta). \tag{2.23}$$

Thus the DLR measures may be regarded as a higher-dimensional extension of Markov processes.

Existence of DLR measures is proved for the lattice gas and for more general systems as well. The set of DLR measures is a weakly compact, convex subset in the space of all probabilities in $\{0,1\}^{\mathbb{Z}^d}$ and their extremal points can be obtained as weak limits as $\Lambda \to \mathbb{Z}^d$ of finite volume Gibbs measures with suitable boundary conditions.

The other basic postulate of the theory is the identification of the thermodynamic potentials in terms of the partition functions: the pressure $\pi_{\beta,\lambda}$ is in fact identified to

$$\pi_{\beta,\lambda} := \lim_{\Lambda \nearrow \mathbb{Z}^d} \frac{1}{\beta|\Lambda|} \log Z_{\beta,\lambda;\Lambda} \tag{2.24}$$

(as already mentioned existence of the limit is proved in great generality). Moreover let

$$Z^{\text{can}}_{\beta,N;\Lambda}(\bar\eta) = \sum_{\eta \in \{0,1\}^\Lambda : N(\eta)=N} e^{-\beta H(\eta|\bar\eta)} \tag{2.25}$$

the canonical partition function. Then the free energy $f_{\beta,\rho}$ at inverse temperature β and particle density ρ is:

$$f_{\beta,\rho} := -\lim_{\Lambda \nearrow \mathbb{R}^d} \lim_{N/|\Lambda| \to \rho} \frac{1}{\beta|\Lambda|} \log Z^{\text{can}}_{\beta,N;\Lambda}. \tag{2.26}$$

Again, general theorems in statistical mechanics prove the existence of the limit. Physical consistency of the above definitions is ensured by *equivalence of ensembles*:

Theorem. *$f_{\beta,\rho}$ is a convex function of ρ, $\pi(\beta,\lambda)$ a convex function of λ and they are the Legendre transform of each other*:

$$\pi_{\beta,\lambda} = \sup_\rho \{\lambda\rho - f_{\beta,\rho}\}, \qquad f_{\beta,\rho} = \sup_\lambda \{\lambda\rho - \pi_{\beta,\lambda}\}. \tag{2.27}$$

2.2.2 *Continuous particle systems*. Still considering the lattice gas, let us label the particles so that to each configuration η such that $N(\eta) = n$ we associate a sequence $(x_1, \dots, x_n)$ of distinct sites in such a way that $\eta(x_i) = 1$ for all $i = 1, \dots, n$. $(x_1, \dots, x_n)$ is a labeling of η and obviously there are $n!$ possible labelings. As the particles are undistinguishable we give probability $1/n!$ to each labeling. Thus calling $\underline{x}$ the elements of $\bigcup_{n\geq 0} \Lambda^n$, we have

$$\mu(\underline{x}) = Z^{-1} \frac{1}{|\underline{x}|!} e^{-\beta[H(\underline{x}) - \lambda|\underline{x}|]}$$

having denoted by $H(\underline{x})$ the Hamiltonian $H(\eta)$ if $\underline{x}$ is a labeling of η and $|\underline{x}|$ the number of elements in $\underline{x}$.

The Gibbs measure for continuous particle systems is simply the extension to continuum of the above expression:

$$\mu_{\beta,\lambda;\Lambda}(d\underline{x}|\bar{\underline{x}}) = Z_{\beta,\lambda;\Lambda}(\bar{\underline{x}})^{-1}\frac{1}{|\underline{x}|!}e^{-\beta[H(\underline{x}|\bar{\underline{x}})-\lambda|\underline{x}|]}\mathbf{1}_{\underline{x}\in\mathcal{X}_\Lambda}\,d\underline{x},$$

where $\mathcal{X}_\Lambda = \bigcup_{n\geq 0}\Lambda^n$, $d\underline{x} = dx_1\cdots dx_n$, if $\underline{x} = (x_1,\ldots,x_n)$, $\bar{x}$ is a configuration outside Λ and

$$Z_{\beta,\lambda;\Lambda}(\bar{\underline{x}}) = \sum_{n\geq 0}\frac{1}{n!}\int_{\Lambda^n} e^{-\beta[H(\underline{x}|\bar{\underline{x}})-\lambda n]}\,dx_1\cdots dx_n \tag{2.28}$$

is the partition function. All that requires suitable assumptions on the interaction to ensure convergence of the above sums and integrals (which are verified by the Lennard–Jones interaction, if there are hard cores and in general if the "Hamiltonian is stable").

DLR measures, pressure and free energy are defined by extending to continuum the definition given for the lattice gas. The extension is however not trivial because the variables are now unbounded and several problems arise from noncompactness. This is beyond the purpose of these lectures and I just refer to the literature.

2.3 Phase transitions at positive temperatures

There are several possible definitions of phase transitions, they are not equivalent each one captures some of the aspects of the phenomenon.

2.3.1 *Phase transitions and thermodynamic potentials*. Phase transitions of first order with order parameter the mass density are characterized in thermodynamics by the condition that the graph of the free energy density versus the mass density has a straight segment, see Figure 9 or that the derivative of the pressure with respect to the chemical potential has a discontinuity, see Figure 10,

Since the free energy is equal to the internal energy minus T times the entropy (T the absolute temperature), in the limit $T \to 0$ the above condition for phase transitions becomes what was seen in the previous chapter where the internal energy at 0 temperature is linear in an interval $[0, \rho_c]$, see Figures 7 and 8.

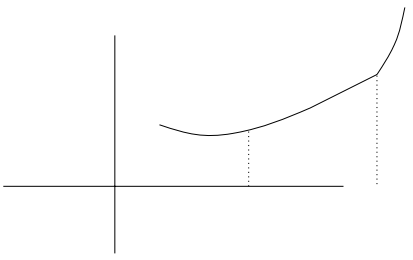

Figure 9 *Free energy versus density, the graph has a straight segment.*

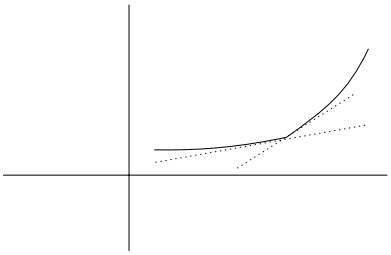

Figure 10 *Pressure versus chemical potential, the derivative jumps at a critical value.*

2.3.2 *Phase transitions and DLR measures*. In this context, phase transitions are characterized in two ways: (i) there are more than one DLR measures; (ii) there are more than one translation invariant DLR measures.

The formulation of phase transitions as outlined in the first lecture requires some more care.

Definition 2.1. Denote by $\mathcal{D}^{(\ell)}$, $\ell = 2^k, k \in \mathbb{Z}$, partitions of $\mathbb{R}^d$ into the cubes of side ℓ

$$\begin{aligned} C^{(\ell)}_{\underline{n}\ell} &= \big\{(r_1, \dots, r_d) \in \mathbb{R}^d : n_i\ell \le r_i < (n_i + 1)\ell\big\}, \\ \underline{n} &= (n_1, \dots, n_d) \in \mathbb{Z}^d. \end{aligned} \tag{2.29}$$

Write $C^{(\ell)}_r$, $r \in \mathbb{R}^d$ for the cube in $\mathcal{D}^{(\ell)}$ which contains r.

Notice that $\mathcal{D}^{(\ell)}$ refines $\mathcal{D}^{(2\ell)}$ by splitting each atom of the latter into 2^d atoms of the former. We then say that $\mathcal{D}^{(\ell')}$ is finer than $\mathcal{D}^{(\ell)}$ if $\ell' < \ell$ and that $\mathcal{D}^{(\ell)}$ is coarser than $\mathcal{D}^{(\ell')}$.

We are now ready to define (empirical) density profiles of particle configurations:

Definition 2.2. Given a particle configuration $\underline{x}$ we define its (empirical) density profile on the scale ℓ as

$$\rho_\ell(r; \underline{x}) := \frac{|\underline{x} \cap C^{(\ell)}_r|}{\ell^d}, \tag{2.30}$$

$|\underline{x} \cap C^{(\ell)}_r|$ being the number of particles of the configuration $\underline{x}$ which are in $C^{(\ell)}_r$.

Therefore, the empirical density $\rho_\ell(r; \underline{x})$ is a nonnegative function constant in r on the atoms of $\mathcal{D}^{(\ell)}$.

We distinguish between two very different scenarios: when ρ is not in a forbidden interval, that is, away from phase transitions, then with large probability $\rho_\ell(r; \underline{x})$ is close to ρ in a large fraction of the volume $|\Lambda|$. When instead ρ is inside the forbidden interval $(\rho^-_\beta, \rho^+_\beta)$, then with large probability $\rho_\ell(r; \underline{x})$ differs

from ρ in a large fraction of the volume $|\Lambda|$, taking values close either to ρ_β^- or to ρ_β^+.

We need to quantify the above considerations by specifying what we mean by "large probability" and "being close to", we also need to give conditions on the values of the mesh ℓ and on the region Λ.

First of all, notice that the restriction to "large probability events" is necessary: we are dealing with probabilities and stochastic effects must be taken into account. Indeed the canonical Gibbs measure gives positive probability to all open sets and therefore for any given Λ with positive probability we may see very atypical events. But even in the typical cases we have deviations from the expected behavior. Thus, given an accuracy parameter $\varepsilon > 0$ we define "the set of typical configurations" (in the first scenario with no phase transitions)

$$G^{\text{no PT}}_{\rho,\Lambda,\ell,\varepsilon} := \{\underline{x} : |\{r \in \Lambda : |\rho_\ell(r;\underline{x}) - \rho| \le \varepsilon\}| \ge (1-\varepsilon)|\Lambda|\} \tag{2.31}$$

while in the second scenario (with phase transitions)

$$\begin{aligned} G^{\text{PT}}_{\rho,\Lambda,\ell,\varepsilon} := \{\underline{x} : |\{r \in \Lambda : |\rho_\ell(r;\underline{x}) - \rho_\beta^-| > \varepsilon, \\ |\rho_\ell(r;\underline{x}) - \rho_\beta^+| > \varepsilon\}| \le \varepsilon|\Lambda|\}. \end{aligned} \tag{2.32}$$

Denote below by $\mu^{\text{can}}_{\beta,[\rho|\Lambda|];\Lambda}$ the conditional Gibbs measure given that in Λ there are $[\rho|\Lambda|]$ particles ($[a]$ the integer part of a). Notice that this conditional probability does not depend on the chemical potential.

Definition 2.3. ρ is "not forbidden" if for any increasing sequence Λ of cubes invading $\mathbb{R}^d$:

$$\lim_{\varepsilon\to 0}\lim_{\ell\to\infty}\lim_{\Lambda\nearrow\mathbb{R}^d} \mu^{\text{can}}_{\beta,[\rho|\Lambda|];\Lambda}\big[G^{\text{no PT}}_{\rho,\Lambda,\ell,\varepsilon}\big] = 1 \tag{2.33}$$

$(\rho_\beta^-, \rho_\beta^+)$ is a forbidden density interval if for any $\rho \in (\rho_\beta^-, \rho_\beta^+)$ and any increasing sequence of cubes Λ invading $\mathbb{R}^d$:

$$\lim_{\varepsilon\to 0}\lim_{\ell\to\infty}\lim_{\Lambda\nearrow\mathbb{R}^d} \mu^{\text{can}}_{\beta,[\rho|\Lambda|];\Lambda}\big[G^{\text{PT}}_{\rho,\Lambda,\ell,\varepsilon}\big] = 1. \tag{2.34}$$

Equation (2.34) can be formulated by saying that given any $\varepsilon > 0$ there is ℓ_ε and for any $\ell \ge \ell_\varepsilon$ there is Λ_ℓ so that for all $\Lambda \supseteq \Lambda_\ell$

$$\mu^{\text{can}}_{\beta,[\rho|\Lambda|];\Lambda}\big[G^{\text{PT}}_{\rho,\Lambda,\ell,\varepsilon}\big] \ge 1 - \varepsilon \tag{2.35}$$

with an analogous statement in the case of absence of phase transitions.

A clear separation of the two phases appears when the mesh ℓ is large, but by (2.35) ℓ large requires that the region Λ should be large enough, hence the question on how large can we take ℓ given a (very large) region Λ: notice that larger ℓ means less fragmentation between the regions where $\rho_\ell(r;\underline{x})$ is close to

ρ_β^- and ρ_β^+. Suppose Λ is a cube $C^{(\ell^*)}$ of $\mathcal{D}^{(\ell^*)}$, then if $\ell = \ell^*$, $\rho_\ell(r; \underline{x}) = \rho$ for all $r \in \Lambda$, while if ℓ is large but much smaller than ℓ^* then $\rho_\ell(r; \underline{x})$ is close to ρ_β^- and ρ_β^+ in a large fraction of Λ (since the support of the canonical measure is $\mathcal{X}_{\Lambda,[\rho|\Lambda|]}$ the fraction of good cubes where $\rho_\ell(r; \underline{x})$ is close to ρ_β^- and ρ_β^+ must balance in such a way that the total density is approximatively ρ). There is therefore a transition from this regime and the final one where $\rho_{\ell^*}(r; q) = \rho$ and therefore it is no longer close to ρ_β^- and ρ_β^+: we expect such a transition only when ℓ becomes macroscopic, that is, of the order of the side L of Λ. All that is related to the Wulff shape problem for which we refer to the literature.

When ρ is in the forbidden interval $(\rho_\beta^-, \rho_\beta^+)$ the canonical constraint forces the two phases to coexist in each "typical configuration". This is not what happens in the gran canonical ensemble. Let $\mu_{\beta,\lambda,\Lambda}$ be the gran canonical measure on the torus Λ and take $\lambda = \lambda_\beta$, λ_β the value of the chemical potential for which there is phase transition. Then the typical configurations of $\mu_{\beta,\lambda,\Lambda}$ have either a density close to ρ_β^- or to ρ_β^+ and therefore the two phases do not coexist in a same configuration $\underline{x}$ but only in a statistical average. On the other hand this happens only for the special value $\lambda = \lambda_\beta$ of the chemical potential, in contrast with the canonical picture where for any value $\rho \in (\rho_\beta^-, \rho_\beta^+)$ there is coexistence, in agreement with the predictions of thermodynamics, see Figures 10 and 9.

2.3.3 *Phase transitions, results.* There are general theorems on the absence of phase transition at low densities:

Theorem 2.4 (Absence of phase transitions). *For any $\beta > 0$, there is $\rho'_\beta > 0$ so that (2.33) holds for $\rho \le \rho'_\beta$. Moreover in $d = 1$, dimensions there is no phase transition if the pair interaction V satisfies the condition $\int |V(|r|)|r|\, dr < \infty$.*

Existence of phase transitions at small temperatures is only a conjecture (but largely believed to be correct):

Conjecture. *In $d \ge 2$ dimensions for any $\beta > 0$ large enough there are $0 < \rho_\beta^- < \rho_\beta^+$ so that (2.34) holds true. Moreover $(\rho_\beta^-, \rho_\beta^+)$ converges as $\beta \to \infty$ to the 0 temperature forbidden density interval.*

For particle systems in the continuum the conjecture is still among the most important open problems in statistical mechanics, but, for the lattice gas it is proved, as sketched in the sequel. We suppose $d = 2$ and nearest neighbor attractive interactions, such restrictions can be at least partially relaxed using the Pirogov–Sinai theory (Sinai (1982)).

2.4 The Ising model

We shall prove in this section that a lattice gas in $\mathbb{Z}^d$, $d \geq 2$, with nearest neighbor attractive interactions has a phase transition at low temperatures. As already seen in the first lecture, a lattice gas becomes an Ising ± 1 spin system by setting

$$\sigma(x) = 2\eta(x) - 1, \qquad \eta(x) = \frac{\sigma(x)+1}{2}. \tag{2.36}$$

The particle density in the lattice gas is then replaced by the magnetization density (sum of spins over volume) in the Ising language.

We shall restrict in the sequel to the Ising system in $d = 2$ with nearest neighbor, ferromagnetic interactions. Let Λ be a bounded region $\sigma_\Lambda \in \{-1, 1\}^\Lambda$ and $\bar{\sigma}_{\Lambda^c} \in \{-1, 1\}^{\Lambda^c}$, then the Ising energy of σ_Λ when interacting with $\bar{\sigma}_{\Lambda^c}$ is

$$\begin{aligned} H_\Lambda(\sigma_\Lambda|\bar{\sigma}_{\Lambda^c}) = &-\frac{J}{2}\sum_{|x-y|=1}\sigma_\Lambda(x)\sigma_\Lambda(y) - h\sum_x \sigma_\Lambda(x) \\ &- J\sum_{|x-y|=1, x\in\Lambda, y\in\Lambda^c}\sigma_\Lambda(x)\bar{\sigma}_{\Lambda^c}(y), \qquad J>0, h\in\mathbb{R} \end{aligned} \tag{2.37}$$

h in (2.37) plays the role that the chemical potential has in the lattice gas; physically h is an external magnetic field. The spin–spin interaction is ferromagnetic, because it is minimal when the two spins are aligned, the magnetic field h instead wants the spins to have its same sign (i.e., to be aligned to h).

From what was proved in the first lecture at $T = 0$, there is phase transition at $T = 0$ with forbidden magnetization interval $[-1, 1]$ (in the lattice gas this translates into the interval $[0, 1]$). We shall first prove a phase transition by showing that at $h = 0$ for β large enough there are at least two DLR measures.

Define

$$\begin{aligned} \mu_{\beta,\Lambda,\bar{\sigma}_{\Lambda^c}}(\sigma_\Lambda) &= \frac{1}{Z_{\beta,\Lambda,\bar{\sigma}_{\Lambda^c}}} e^{-\beta H(\sigma_\Lambda|\bar{\sigma}_{\Lambda^c})}, \\ Z_{\beta,\Lambda,\bar{\sigma}_{\Lambda^c}} &= \sum_{\sigma_\Lambda\in\{-1,1\}^\Lambda} e^{-\beta H^{\pm}(\sigma_\Lambda|\bar{\sigma}_{\Lambda^c})} \end{aligned} \tag{2.38}$$

writing $\mu^{\pm}_{\beta,\Lambda}$ when $\bar{\sigma}_{\Lambda^c} = \pm\mathbf{1}_{\Lambda^c}$.

Theorem 2.5. *For β large enough, there are two probability measures $\mu^{\pm}_\beta$ on $\{-1, 1\}^{\mathbb{Z}^d}$ such that*

$$\text{(weak)} \lim_{\Lambda\to\mathbb{Z}^d} \mu^{\pm}_{\beta,\Lambda} = \mu^{\pm}_\beta \tag{2.39}$$

for any increasing sequence Λ of bounded regions which invades $\mathbb{Z}^d$. $\mu^{\pm}_\beta$ are translation invariant and there is $m_\beta > 0$, $\lim_{\beta\to\infty} m_\beta = 1$, so that

$$E_{\mu^{\pm}_\beta}[\sigma(x)] = \pm m_\beta \qquad \textit{for any } x \in \mathbb{Z}^d.$$

We just mention (referring to the literature for a proof) that for β small enough and for any β but $h \neq 0$ there is a unique DLR measure.

Weak limit means convergence of the expectation of any cylindrical function f, f being cylindrical if it depends only on finitely many spins. The notion is equivalent to demand that for any bounded set Δ

$$\lim_{\Lambda \to \mathbb{Z}^d} E_{\mu^{\pm}_{\beta,\Lambda}}\Big[\prod_{x\in\Delta} \sigma(x)\Big] = E_{\mu^{\pm}_{\beta}}\Big[\prod_{x\in\Delta} \sigma(x)\Big].$$

The theorem proves a persistent dependence of the Gibbs measures on the boundary conditions $\pm\mathbf{1}_{\Lambda^c}$ even when they are sent away to infinity ($\Lambda \to \mathbb{Z}^d$) which is the starting point of our proofs.

The existence of two distinct DLR mesures claimed in Theorem 2.5 follows from that following theorem.

Theorem 2.6. *For any β large enough, there is $c = c(\beta) < \frac{1}{2}$ so that for any bounded Λ*

$$\sup_{x\in\Lambda} \mu^{+}_{\beta,\Lambda}\big[\sigma(x) = -1\big] \leq c. \tag{2.40}$$

In fact by Theorem 2.6 and the spin flip symmetry,

$$\sup_{x\in\Lambda} \mu^{-}_{\beta,\Lambda}\big[\sigma(x) = 1\big] \leq c < \frac{1}{2}. \tag{2.41}$$

Hence, $\mu^{-}_{\beta,\Lambda}[\sigma(x) = 1] \leq c < 1 - c \leq \mu^{+}_{\beta,\Lambda}[\sigma(x) = 1]$ and therefore the weak limits as $\Lambda \nearrow \mathbb{Z}^d$ of $\mu^{\pm}_{\beta,\Lambda}$ are different from each other (their existence (by subsequences) follows from compactness, as the space of all probabilities on a compact space, as $\{-1, 1\}^{\mathbb{Z}^d}$, is weakly compact, see Parthasarathy (1967)). As a consequence, there are two Gibbs probabilities $\mu^{\pm}_{\beta}$ (weak limits of $\mu^{\pm}_{\beta,\Lambda}$) which are not equal to each other (but much more is still needed for Theorem 2.5).

In the next section, we shall prove Theorem 2.6 and in the complements to this section we shall prove Theorem 2.5 and give a sketch of the proof of phase transitions in the canonical ensemble.

2.5 The plus Gibbs measures

We shall prove in this section that (2.40) holds for β large enough. The proof is based on the Peierls argument which is based on the notion of contours.

2.5.1 *Contours and the Peierls bounds*. We shall argue that the "contours" are the "microscopic interfaces" which separate the plus and minus "microscopic phases".

Definition 2.7. The dual of $\mathbb{Z}^2$ is the set of straight lines in $\mathbb{R}^2$:

$$\mathbb{Z}^2_* = \big\{(r_1, r_2) \in \mathbb{R}^2 : r_1 - \tfrac{1}{2} \in \mathbb{Z}\big\} \cup \big\{(r_1, r_2) \in \mathbb{R}^2 : r_2 - \tfrac{1}{2} \in \mathbb{Z}\big\}.$$

A "node" is a point where two straight lines of $\mathbb{Z}^2_*$ intersect, namely $(r_1, r_2) \in \mathbb{Z}^2_* : (r_1 - \frac{1}{2}, r_2 - \frac{1}{2}) \in \mathbb{Z}^2$; an "edge" is the unit segment which joins two neighboring nodes and a "bond" is the unit segment joining two neighboring sites of the original lattice $\mathbb{Z}^2$. The bond and the edge which crosses it are "dual" to each other.

Thus $\mathbb{Z}^2_*$ is the union of all vertical lines passing through the semi-integers of the x-axis and all the horizontal lines through the semi-integers of the y-axis.

Definition 2.8. The complement of $\mathbb{Z}^2_*$ is the union of all the unit open squares with center in $\mathbb{Z}^2$, they are denoted by $C_x, x \in \mathbb{Z}^2$. To visualize a spin configuration σ, we paint in red all C_x with $\sigma(x) = 1$ and in white those where $\sigma(x) = -1$.

Definition 2.9. Given a configuration $\sigma \in \{-1, 1\}^{\mathbb{Z}^2}$ call $\mathcal{R}(\sigma)$ the union of all the edges in between a red and a white square, namely such that the dual bond joins two sites with opposite spins; $|\mathcal{R}(\sigma)|$ denotes the number of edges in $\mathcal{R}(\sigma)$.

Contours are then defined as:

Definition 2.10. The contours of a configuration σ are the maximal connected components of $\mathcal{R}(\sigma)$, they are denoted by $\underline{\gamma} = (\gamma_1, \gamma_2, \ldots)$. An abstract contour is a set in $\mathbb{Z}^2_*$ which is a contour for some $\sigma \in \{-1, 1\}^{\mathbb{Z}^2}$.

Let Λ be a finite set in $\mathbb{Z}^2$. We call E_Λ the union of all the edges which are sides of C_x for some $x \in \Lambda$ and denote by $\{\mathcal{R}\}_\Lambda$ the space of all possible $\mathcal{R}(\sigma)$ with σ constant outside Λ. We state in the next theorem some basic properties of $\{\mathcal{R}\}_\Lambda$, referring to the literature for their proof.

Theorem 2.11. *Let $\mathcal{R} \in \{\mathcal{R}\}_\Lambda$, writing $\gamma = \mathcal{R}$ if $\mathcal{R}$ is connected* (*and thus a contour*). *$\mathcal{R} \subset E_\Lambda$ and*:

(1) *Let σ and σ' be such that $\mathcal{R}(\sigma) = \mathcal{R}(\sigma') = \mathcal{R}$. Then σ is constant outside Λ and either $\sigma = \sigma'$ or $\sigma = -\sigma'$.*

(2) *From each node of $\mathbb{Z}^2_*$ it departs an even number of edges belonging to $\mathcal{R}$* (*i.e., either* 0, *or* 2 *or* 4).

(3) *Viceversa, if $A \subset E_\Lambda$ is union of edges such that from each node of $\mathbb{Z}^2_*$ it departs an even number of edges belonging to A, then $A \in \{\mathcal{R}\}_\Lambda$.*

(4) *If $\mathcal{R}$ is connected* (*so that it is a contour*) *then $\gamma := \mathcal{R}$ is a closed polygonal path*: *namely there exists a continuous curve $r(t), t \in [0, 1]$, with range γ, which starts and ends at the same point ($r(0) = r(1) \in \gamma$) and which can only self intersect at the nodes of γ* (*i.e., if $r \in \gamma$ is not a node then there is a unique t so that $r = r(t)$*).

(5) *The complement of γ is union of maximal connected components*: *one of them*, *called* $\mathrm{ext}(\gamma)$, *is unbounded*, *the others*, *called* $\mathrm{int}_i(\gamma)$, *are bounded open sets*. *We call* $\mathrm{int}(\gamma)$ *the union of all the sets* $\mathrm{int}_i(\gamma)$.

(6) *Call* $\delta_{\mathrm{ext}}(\gamma)$ *the union of all* $x \in \mathbb{Z}^2$ *such that* $C_x \subset \mathrm{ext}(\gamma)$ *and the closure of* C_x *intersects* γ (*recall that* C_x *is the unit open square with center* x). *Analogously* $\delta_{\mathrm{int}_i}(\gamma)$ *is the union of all* $x \in \mathbb{Z}^2$ *such that* $C_x \subset \mathrm{int}_i(\gamma)$ *and the closure of* C_x *intersects* γ. *Then if* γ *is a contour of* σ, $\sigma(x)$ *is constant on* $\delta_{\mathrm{ext}}(\gamma)$ *and on each* $\delta_{\mathrm{int}_i}(\gamma)$. γ *is called a plus* (*minus*) *contour if* $\sigma(x) = 1$ $(\sigma(x) = -1)$ *on* $\delta_{\mathrm{ext}}(\gamma)$. $\mathrm{int}_i(\gamma)$ *is called a minus* (*plus*) *interior if* $\sigma(x) = -1$ $(\sigma(x) = 1)$ *on* $\delta_{\mathrm{int}_i(\gamma)}$.

We are now ready to prove the Peierls bounds. Recall that $h = 0$ and $\mu^+_{\beta,\Lambda}$ is the Gibbs measure on $\mathcal{X}_\Lambda := \{-1, 1\}^\Lambda$ with boundary condition $\mathbf{1}_{\Lambda^c}$ equal to 1 outside Λ. $\mathcal{X}_\Lambda$ is in one to one correspondence with $\{\mathcal{R}\}_\Lambda$ by setting $\sigma_\Lambda \to \mathcal{R}(\sigma_\Lambda, \mathbf{1}_{\Lambda^c})$ and by an abuse of notation we write $\mu^+_{\beta,\Lambda}(\mathcal{R})$ for $\mu^+_{\beta,\Lambda}(\sigma_\Lambda)$. By (2.37)

$$\begin{aligned} H_\Lambda(\sigma_\Lambda|\mathbf{1}_{\Lambda^c}) &= -\frac{J}{2}\sum_{|x-y|=1}\big(\sigma_\Lambda(x)\sigma_\Lambda(y) - 1 + 1\big) \\ &= -\frac{J}{2}|E_\Lambda| + J\sum_{|x-y|=1}\mathbf{1}_{\sigma(x)\neq\sigma(y)} \end{aligned}$$

so that the energy, modulo an additive constant, is equal to $2J|\mathcal{R}(\sigma)|$, hence

$$\mu^+_{\beta,\Lambda}(\mathcal{R}) = \frac{1}{Z^+_{\beta,\Lambda}}e^{-\beta 2J|\mathcal{R}|}, \qquad Z^+_{\beta,\Lambda} = \sum_{\mathcal{R}\in\{\mathcal{R}\}_\Lambda} e^{-\beta 2J|\mathcal{R}|}. \tag{2.42}$$

Lemma 2.12. *Let* $\gamma \in \{\mathcal{R}_\Lambda\}$ *be a contour*; *denote by* $\{\sigma \Rightarrow \gamma\}$ *the set of all* σ_Λ *such that* γ *is a contour of* $(\sigma_\Lambda, \mathbf{1}_{\Lambda^c})$. *Then*

$$\mu^+_{\beta,\Lambda}\big[\{\sigma \Rightarrow \gamma\}\big] \le e^{-\beta 2J|\gamma|}. \tag{2.43}$$

Proof. If $\sigma_\Lambda \in \{\sigma \Rightarrow \gamma\}$ then γ is a maximal connected component of $\mathcal{R}(\sigma_\Lambda, \mathbf{1}_{\Lambda^c})$ and, by property (3) of Theorem 2.11, $\mathcal{R} := \mathcal{R}(\sigma_\Lambda, \mathbf{1}_{\Lambda^c}) \setminus \gamma$ is also an element of $\{\mathcal{R}\}_\Lambda$ and we write $\mathcal{R} \sim \gamma$. Thus, the sum of $\mu^+_{\beta,\Lambda}(\sigma_\Lambda)$ over all $\sigma_\Lambda \in \{\sigma \Rightarrow \gamma\}$ can be written as

$$\begin{aligned} \mu^+_{\beta,\Lambda}\big[\{\sigma \Rightarrow \gamma\}\big] &= \frac{1}{Z^+_{\beta,\Lambda}}\sum_{\mathcal{R}\in\{\mathcal{R}\}_\Lambda\,:\,\mathcal{R}\sim\gamma} e^{-\beta 2J(|\mathcal{R}|+|\gamma|)} \\ &\le e^{-\beta 2J|\gamma|}\frac{1}{Z^+_{\beta,\Lambda}}\sum_{\mathcal{R}\in\{\mathcal{R}\}_\Lambda} e^{-\beta 2J|\mathcal{R}|} \end{aligned}$$

which by (2.42) is equal to $e^{-\beta 2J|\gamma|}$. □

Let $(\gamma_1, \dots, \gamma_k)$ be pairwise disjoint contours and denote by $\{\sigma \Rightarrow (\gamma_1, \dots, \gamma_k)\}$ the set of all σ such that $\gamma_1, \dots, \gamma_k$ are among the contours of σ, then

$$\mu^+_{\beta,\Lambda}[\{\sigma \Rightarrow (\gamma_1, \dots, \gamma_k)\}] \le e^{-\beta 2J(|\gamma_1|+\cdots+|\gamma_k|)}. \tag{2.44}$$

The proof of (2.44) is similar to that of (2.43) and omitted.

Proof of Theorem 2.6. To prove Theorem 2.6, we need to bound the probability that $\{\sigma_\Lambda(x) = -1\}$. Recalling that we have plus boundary conditions if $\sigma_\Lambda(x) = -1$, $x \in \Lambda$, there must be a contour $\gamma \in \mathcal{R}_\Lambda$ such that $x \in \mathrm{int}(\gamma)$. Thus,

$$\mu^+_{\beta,\Lambda}[\sigma_\Lambda(x) = -1] \le \sum_{\gamma\,:\,\mathrm{int}(\gamma)\ni x} e^{-2\beta J|\gamma|}. \tag{2.45}$$

The last sum is bounded by

$$\mu^+_{\beta,\Lambda}[\sigma_\Lambda(x) = -1] \le \sum_{\gamma\,:\,x+(1/2)e_1\in\gamma} |\gamma| e^{-2\beta J|\gamma|} \le \sum_{\gamma\,:\,x+(1/2)e_1\in\gamma} e^{-(2\beta J-1)|\gamma|}$$

because any $\gamma' : \mathrm{int}(\gamma') \ni x$ can be found by first choosing $\gamma : x + \frac{1}{2}e_1 \in \gamma$ (e_1 the unit vector in the horizontal direction) and then by considering all horizontal translates of γ which have $x \in \mathrm{int}(\gamma)$: their number is bounded by $|\gamma|$. In the second inequality, we have written $|\gamma| \le e^{|\gamma|}$.

To sum over all γ, we use property (4) of Theorem 2.11 so that contours are polygonal paths. If we fix the length ℓ of the contours, there are $\le 3^{\ell-1}$ contours with length ℓ which contain the edge which goes through $x + \frac{1}{2}e_1$ because each successive edge can be chosen at most in 3 ways. Thus for β large enough,

$$\mu^+_{\beta,\Lambda}[\sigma_\Lambda(x) = -1] \le \sum_{\ell\ge 4} 3^{\ell-1} e^{-(2\beta J-1)\ell} \le \frac{(3e^{-(2\beta J-1)})^4}{1 - 3e^{-(2\beta J-1)}} \tag{2.46}$$

which proves Theorem 2.6. □

2.6 Complements

The following is more technical and it can be skipped to a first reading.

2.6.1 *A percolation argument.* We shall prove weak convergence of the plus Gibbs measures in the thermodynamic limit, namely that μ^+_{β,Λ_n} has a weak limit μ^+_β for any increasing sequence of finite regions invading $\mathbb{Z}^2$ with the limit μ^+_β independent of the sequence. We follow the proof of an analogous statement in Presutti (2008) which applies to a nonferromagnetic system.

A simpler proof could be given in our case using ferromagnetic inequalities, ours is solely based on the Peierls bounds and therefore, pending their validity, it extends to more general systems. The key ingredient is Theorem 2.13 below which will ensure exponential decay of correlations, see Theorem 2.18.

We use the following notation: Δ is a cube, Λ' and Λ'' two bounded regions containing Δ and μ', μ'' the corresponding plus Gibbs measures (at inverse temperature β). $\Delta_0 \subset \Delta$ is a cube which has distance $\geq R$ from Δ^c, $Y(\Delta_0)$ is the external boundary of Δ_0 made of all sites in the complement of Δ_0 which are connected to Δ_0. The result in the next theorem is meaningful if R grows at least as $c \log L$ (L the side of Δ) with c large enough.

Theorem 2.13. *For any β large enough* (*and with the above notation*)

$$\sup_{f \in \mathcal{C}_{\Delta_0}, \|f\| \leq 1} |E_{\mu'}[f] - E_{\mu''}[f]| \leq 2|Y(\Delta_0)|e^{-\beta J R}, \tag{2.47}$$

where $\|f\|$ is the L^∞ norm, $\mathcal{C}_{\Delta_0}$ is the set of all functions $f(\sigma)$ which depend only on $\{\sigma(x), x \in \Delta_0\}$, they are called "cylindrical functions in Δ".

The theorem can be stated in an equivalent form. Call μ'_{Δ_0} and μ''_{Δ_0} the marginal laws of μ' and μ'' on $\{-1, 1\}^{\Delta_0}$, that is,

$$\mu'_{\Delta_0}(\sigma_{\Delta_0}) = \sum_{\sigma'_{\Lambda'} : \sigma'_{\Lambda'}(x) = \sigma_{\Delta_0}(x), x \in \Delta_0} \mu'(\sigma'_{\Lambda'})$$

with the analogous expression for μ''_{Δ_0}. The total variation distance of μ'_{Δ_0} and μ''_{Δ_0} is

$$\|\mu'_{\Delta_0} - \mu''_{\Delta_0}\| = \sum_{\sigma_{\Delta_0}} |\mu'_{\Delta_0}(\sigma_{\Delta_0}) - \mu''_{\Delta_0}(\sigma_{\Delta_0})|. \tag{2.48}$$

It can be checked that

$$\|\mu'_{\Delta_0} - \mu''_{\Delta_0}\| \leq \sup_{f \in \mathcal{C}_{\Delta_0}, \|f\| \leq 1} |\mu'[f] - \mu''[f]| \tag{2.49}$$

while, obviously,

$$\sup_{f \in \mathcal{C}_{\Delta_0}, \|f\| \leq 1} |\mu'[f] - \mu''[f]| \leq \|\mu'_{\Delta_0} - \mu''_{\Delta_0}\|. \tag{2.50}$$

The proof of Theorem 2.13 is given after some preliminary lemmas. It is based on the method of duplicated variables in the frame of the van den Berg and Maes "disagreement percolation technique". Let $\Omega = \{\sigma = (\sigma', \sigma'') \in \{-1, 1\}^{\Lambda'} \times \{-1, 1\}^{\Lambda''}\}$, $\mu = \mu' \times \mu''$, shorthand $f' = f(\sigma')$, $f'' = f(\sigma'')$, then

$$E_{\mu'}[f] - E_{\mu''}[f] = E_{\mu}[f' - f'']. \tag{2.51}$$

Definition. The sites x, y in $\mathbb{Z}^2$ are connected if they are joined by a bond; a finite set $A \subset \mathbb{Z}^2$ is connected if any x, y in A are connected by a path in A, namely they are first and last site in a sequence (path) $x_1, \ldots, x_n$ of sites in A such that x_i, x_{i+1}, $i = 1, \ldots, n-1$ are connected. We also call $Y(A)$ the sites in A^c connected to A. We then define $\mathcal{A}_\sigma$, $\sigma \in \Omega$, as the collection of all connected sets $A \subset \mathbb{Z}^2$ such that $\Delta_0 \subset A$, $A \cup Y(A) \subset \Delta$ and $\sigma'(x) = \sigma''(x) = 1$ for all $x \in Y(A)$.

Lemma 2.14. *If $\mathcal{A}_\sigma = \varnothing$,there is a connected path $\{x_i\}$ in $\Delta \setminus \Delta_0$ which starts from a site connected to Δ_0 and ends at a site connected to Δ^c, moreover $\sigma'(x) + \sigma''(x) < 2$ all along the path. If $\mathcal{A}_\sigma \neq \varnothing$, it has a maximal element C_σ, namely $C_\sigma \supseteq A$ for all $A \in \mathcal{A}_\sigma$; C_σ is one-connected (meaning that the complement has only one maximal connected component) and measurable on the complement of C_σ (meaning that if σ^* is any configuration equal to σ in the complement of C_σ then $C_\sigma = C_{\sigma^*}$).*

Proof.

Suppose first $\mathcal{A}_\sigma = \varnothing$.

Let B_i be the maximal connected components of the set $\{x : \sigma'(x) + \sigma''(x) < 2\}$ and let B be the union of Δ_0 and all B_i connected to Δ_0. B is a connected set which is connected to Δ^c: otherwise $Y(B)$ would be a subset of Δ and hence $B \in \mathcal{A}_\sigma$, against the assumption that $\mathcal{A}_\sigma$ is empty. Thus B is connected to Δ^c, hence it contains a site y either in Δ^c or connected to Δ^c and a site $x \in \Delta_0$. Then there is a connected path $\{z_i\}$ in B which starts from x and ends at y, such a path must eventually leave Δ_0 and get to a site in Δ connected to Δ^c. This portion of $\{z_i\}$ verifies the property required in the first statement of the lemma.

Suppose next $\mathcal{A}_\sigma \neq \varnothing$.

$\mathcal{A}_\sigma$ is partially ordered (in the sense of inclusion), because if A_1 and A_2 are in $\mathcal{A}_\sigma$ then also $A_1 \cup A_2 \in \mathcal{A}_\sigma$. Thus $\mathcal{A}_\sigma$ has a maximal element C_σ. Given a finite set A call $\mathrm{int}(A)$ the set of all points $x \in A^c$ which cannot be connected to infinity in within A^c and define $c(A) := A \cup \mathrm{int}(A)$. Since $Y(c(A)) \subset Y(A)$, if $A \in \mathcal{A}_\sigma$ then also $c(A) \in \mathcal{A}_\sigma$, as a consequence C_σ, being maximal, is one-connected. If $A \in \mathcal{A}_\sigma$ and σ^* is any configuration obtained from σ by changing spins only in A, then $A \in \mathcal{A}_{\sigma^*}$, hence the last statement in the lemma. □

The proof of Theorem 2.13 is based on two facts: (i) once conditioned on $\{\mathcal{A}_\sigma \neq \varnothing\}$ the marginal laws of σ' and σ'' restricted to Δ_0 are equal; (ii) if instead $\mathcal{A}_\sigma = \varnothing$, then the set $\{x : \sigma'(x) + \sigma''(x) < 2\}$ percolates from Δ_0 to Δ and such an event has very small probability because of the Peierls bounds.

To implement the above ideas, let $\mathcal{K}$ be the collection of all one-connected sets C such that $C \supset \Delta_0$ and $C \cup Y(C) \subset \Delta$ (see Lemma 2.14 and the preceding definition for notation), let

$$G_C = \{\sigma \in \Omega : C_\sigma = C\}, \qquad C \in \mathcal{K} \tag{2.52}$$

then

$$E_\mu[f' - f''] = \sum_{C \in \mathcal{K}} E_\mu[f' - f''; G_C] + E_\mu[f' - f''; \{\sigma : \mathcal{A}_\sigma = \varnothing\}], \tag{2.53}$$

where $E[f; A]$ means the integral of f over the set A.

Lemma 2.15.

$$E_\mu[f' - f''; G_C] = 0 \qquad \textit{for all } C \in \mathcal{K}. \tag{2.54}$$

Proof. By the last statement in Lemma 2.14, there is a set G_C^* in $\{-1,1\}^{\Lambda'\setminus C}\times\{-1,1\}^{\Lambda''\setminus C}$ so that $(\sigma'_{\Lambda'},\sigma''_{\Lambda''})\in G_C$ iff $(\sigma'_{\Lambda'\setminus C},\sigma''_{\Lambda''\setminus C})\in G_C^*$. Moreover if $(\sigma'_{\Lambda'},\sigma''_{\Lambda''})\in G_C$, then $H^+_{\Lambda'}(\sigma'_{\Lambda'})=H^+_C(\sigma'_C)+H_{\Lambda'\setminus C}(\sigma'_{\Lambda'\setminus C}|\mathbf{1}_{(\Lambda')^c})$ with the analogous expression for $H^+_{\Lambda''}(\sigma''_{\Lambda''})$. Using such properties, we have

$$
\begin{aligned}
&E_\mu[f';G_C]\\
&\quad=\frac{1}{Z^+_{\beta,\Lambda'}Z^+_{\beta,\Lambda''}}\sum_{\sigma'_{\Lambda'},\sigma''_{\Lambda''}}\mathbf{1}_{(\sigma'_{\Lambda'},\sigma''_{\Lambda''})\in G_C}e^{-\beta[H^+_{\Lambda'}(\sigma'_{\Lambda'})+H^+_{\Lambda''}(\sigma'_{\Lambda''})]}f(\sigma'_{\Lambda'})\\
&\quad=\frac{1}{Z^+_{\beta,\Lambda'}Z^+_{\beta,\Lambda''}}\\
&\qquad\times\sum_{\sigma'_C,\sigma'_{\Lambda'\setminus C},\sigma''_C,\sigma''_{\Lambda''\setminus C}}\mathbf{1}_{(\sigma'_{\Lambda'\setminus C},\sigma''_{\Lambda''\setminus C})\in G_C^*}e^{-\beta[H^+_{\Lambda'}(\sigma'_{\Lambda'})+H^+_{\Lambda''}(\sigma'_{\Lambda''})]}f(\sigma'_{\Lambda'})\\
&\quad=\frac{(Z^+_{\beta,C})^2}{Z^+_{\beta,\Lambda'}Z^+_{\beta,\Lambda''}}E_{\mu^+_{\beta,C}}[f]\\
&\qquad\times\sum_{(\sigma'_{\Lambda'\setminus C},\sigma''_{\Lambda''\setminus C})}\mathbf{1}_{(\sigma'_{\Lambda'\setminus C},\sigma''_{\Lambda''\setminus C})\in G_C^*}e^{-\beta[H_{\Lambda'\setminus C}(\sigma'_{\Lambda'\setminus C}|\mathbf{1}_{(\Lambda')^c})+H_{\Lambda''\setminus C}(\sigma''_{\Lambda''\setminus C}|\mathbf{1}_{(\Lambda')^c})]}.
\end{aligned}
$$

Same expression is obtained for $E_\mu[f'';G_C]$ and (2.54) is proved. □

Proof of Theorem 2.13. We bound the last term in (2.53) as

$$
\big|E_\mu[f'-f'';\{\sigma:\mathcal{A}_\sigma=\varnothing\}]\big|\le 2\mu[\{\sigma:\mathcal{A}_\sigma=\varnothing\}].
$$

We shall bound $\mu[\{\sigma:\mathcal{A}_\sigma=\varnothing\}]$ using the following property (whose proof is postponed):

Claim. *If $\mathcal{A}_\sigma=\varnothing$ then there are two sequences $(\gamma'_1,\dots,\gamma'_{m'})$, $(\gamma''_1,\dots,\gamma''_{m''})$ of contours for σ' and respectively, σ'' such that* (i) $(\mathrm{int}(\gamma'_1)\cup\cdots\cup\mathrm{int}(\gamma''_{m''}))\cap Y(\Delta_0)\neq\varnothing$; (ii) $S:=\gamma'_1\cup\cdots\cup\gamma''_{m''}$ *is connected and* (iii) $|S|\ge R$.

Call $\mathcal{D}_x$, $x\in Y(\Delta_0)$, the set of all $(\underline{\gamma}',\underline{\gamma}'')=(\gamma'_1,\dots,\gamma'_{m'}),(\gamma''_1,\dots,\gamma''_{m''})$ as above and such that $x\in(\mathrm{int}(\gamma'_1)\cup\cdots\cup\mathrm{int}(\gamma''_{m''}))$. Using the Claim, we then have

$$
\begin{aligned}
\mu[\{\sigma:\mathcal{A}_\sigma=\varnothing\}]&\le\sum_{x\in Y(\Delta_0)}\sum_{(\underline{\gamma}',\underline{\gamma}'')\in\mathcal{D}_x}\mu'[\sigma'\Rightarrow\underline{\gamma}']\mu''[\sigma''\Rightarrow\underline{\gamma}'']\\
&\le\sum_{x\in Y(\Delta_0)}\sum_{(\underline{\gamma}',\underline{\gamma}'')\in\mathcal{D}_x}e^{-2\beta J(|\gamma'_1|+\cdots+|\gamma''_{m''}|)}.
\end{aligned}
$$

By the property (iii) of the Claim $|\gamma_1'| + \cdots + |\gamma_{m''}''| \geq R$ so that

$$\mu\big[\{\sigma : \mathcal{A}_\sigma = \varnothing\}\big] \leq e^{-\beta J R} \sum_{x \in Y(\Delta_0)} \sum_{(\underline{\gamma}', \underline{\gamma}'') \in \mathcal{D}_x} e^{-\beta J(|\gamma_1'| + \cdots + |\gamma_{m''}''|)}. \tag{2.55}$$

If $(\underline{\gamma}', \underline{\gamma}'') \in \mathcal{D}_x$, then $S := \gamma_1' \cup \cdots \cup \gamma_{m''}''$ is connected and there is a shift be_1, $b \in \mathbb{Z}$, $b \leq |S|$, of S which contains $x + e_1/2$, $|S|$ the number of edges in S. Since there are at most $3^{|S|}$ sequences $(\underline{\gamma}', \underline{\gamma}'')$ which give rise to a same S, we then get

$$\mu\big[\{\sigma : \mathcal{A}_\sigma = \varnothing\}\big] \leq e^{-\beta J R} \sum_{x \in Y(\Delta_0)} \sum_{S \in \mathcal{S}_x} |S| 3^{|S|} e^{-\beta J |S|},$$

where $\mathcal{S}_x$ denotes the collection of all finite, connected union of edges which contain $x + e_1/2$. We then bound (for β large enough)

$$\begin{aligned} \mu\big[\{\sigma : \mathcal{A}_\sigma = \varnothing\}\big] &\leq e^{-\beta J R} \sum_{x \in Y(\Delta_0)} \sum_{S \in \mathcal{S}_x} e^{-(\beta J - 4)|S|} \\ &\leq e^{-\beta J R} |Y(\Delta_0)|. \end{aligned} \tag{2.56}$$

The sum over S has been bounded using the following combinatorial lemma (Lemma 3.1.2.4 in Presutti (2008)), whose proof is reported below for the readers' convenience.

Lemma 2.16. *Let $b > 0$ be such that*

$$e^{-b} 2^4 < 1. \tag{2.57}$$

Then

$$\sum_{S \,:\, s_x \in S} e^{-b|S|} < 1, \tag{2.58}$$

where the sum is over connected sets $S \in \mathbb{Z}_*^2$ which contain the vertical edge s_x.

Proof. The lemma is proved by reducing to a sum over trees. We thus introduce a graph which starts from a vertex called root, whose label is 0. The root is connected to 4 new vertices, called individuals of the first generation, with label $01, \ldots, 04$. Each one of them is in its turn connected to 4 new elements, the collection of all such new elements are the individuals of the second generation, labeled by $0i_1i_2$, $i_j \in \{1, \ldots, 4\}$; from each of them spring 4 new elements and so on, the structure repeating indefinitely. A tree T is a connected subset of this graph which contains the root.

We can now associate to each S in (2.58) a tree T by the following rule. We order in some arbitrary but translational invariant fashion the 4 edges connected to a given node. We then associate to the highest vertical node ξ in s_x the root of the tree. The elements of the first generation are the nodes which are reached from ξ

using edges in S with the pre-assigned order. The elements of the next generations are defined with same rule but excluding edges already considered. In this way, we establish a one to one correspondence between sets S in (2.58) and finite trees T. Calling $|T|$ the number of individuals in T,

$$\sum_{S\,:\,S\ni s_x} e^{-b|S|} \le \sum_{T\,:\,0<|T|<\infty} e^{-b|T|}.$$

Call $n(T)$ the number of generations in T, then

$$\sum_{T\,:\,n(T)=1} e^{-b|T|} = \left[1+e^{-b}\right]^4 - 1 = \sum_{k=1}^{4} \frac{4!}{k!(4-k)!} e^{-bk} \le e^{-b}2^4 < 1.$$

By induction, suppose $z_N := \sum_{T\,:\,0<n(T)\le N} e^{-b|T|} < 1$ then

$$z_{N+1} = \left[1+e^{-b}z_N\right]^4 - 1 \le e^{-b}2^4$$

where the 1 in the square bracket is when the ith element of the first generation is absent, the second factor is when it is present, in such a case it may be seen as the root of a tree with $\le N$ generations. □

Proof of the Claim. We first state some elementary properties of contours. Let γ_1 and γ_2 be two distinct contours of a spin configuration, then $\gamma_1 \cap \gamma_2 = \varnothing$ and either: (i) $\gamma_2 \subset \mathrm{int}(\gamma_1)$, (ii) $\gamma_1 \subset \mathrm{int}(\gamma_2)$, (iii) both $\gamma_2 \subset \mathrm{ext}(\gamma_1)$ and $\gamma_1 \subset \mathrm{ext}(\gamma_2)$. If (i) we say $\gamma_2 < \gamma_1$, if (ii) $\gamma_1 < \gamma_2$. Then given any x either no contour γ has x in its interior or there is a maximal contour γ among those having x in their interior.

Let $\mathcal{A}_\sigma = \varnothing$ then there is a connected path $(x_1, \dots, x_n)$ in $\Delta \setminus \Delta_0$ with x_1 connected to Δ_0 and x_n to Δ^c and moreover $\sigma'(x_i) + \sigma''(x_i) < 2$ for all $i = 1, \dots, n$. Call γ_i' the maximal contour among those for σ' which has x_i in its interior (if no such contour exists we put $\gamma_i' = \varnothing$). γ_i'' is defined in the same way but referring to σ''. Since at least one among $\sigma'(x_i)$ and $\sigma''(x_i)$ is equal to -1 then $\gamma_i' \cup \gamma_i'' \neq \varnothing$. We delete from the collection $(\gamma_1', \dots, \gamma_n'')$ all elements which are smaller (in the above sense) of other elements in the sequence; the sequence $(\underline{\gamma}', \underline{\gamma}'')$ which is left verifies the properties of the Claim. By construction in fact the union of all the interiors of $(\underline{\gamma}', \underline{\gamma}'')$ contains the sequence $x_1, \dots, x_n$ so that properties (i) and (iii) of the Claim hold true. The proof of (ii) is obtained as follows. Suppose by contradiction that S is not connected, then the union of all the interiors of $(\underline{\gamma}', \underline{\gamma}'')$ is not connected as well which contradicts the fact that it contains the connected path $x_1, \dots, x_n$. Thus the Claim is proved and by (2.56) the proof of Theorem 2.13 is complete. □

□

A corollary of Theorem 2.13. *The following theorem is a corollary of the proof of Theorem 2.13. We keep the same notation and call $\mu_\Delta^{\tilde\sigma}$ the Gibbs measure on $\{-1,1\}^\Delta$ with boundary condition $\bar\sigma, \bar\sigma$ a configuration on a set in the complement*

of Δ which contains $Y(\Delta)$ (due to the nearest neighbor assumption on the interaction the values of $\bar\sigma$ on $Y(\Delta)$ are the only relevant ones). We then write $\mu^{\bar\sigma}_{\Delta;\Delta_0}$ for the marginal of $\mu^{\bar\sigma}_{\Delta}$ on $\{-1,1\}^{\Delta_0}$.

Theorem 2.17. *In the same context as in Theorem 2.13,*

$$\begin{aligned}\mu[(\sigma'_{\Lambda'\setminus\Delta},\sigma''_{\Lambda''\setminus\Delta}) : \|\mu^{\sigma'_{\Lambda'\setminus\Delta}}_{\Delta;\Delta_0} - \mu^{\sigma''_{\Lambda''\setminus\Delta}}_{\Delta;\Delta_0}\| \geq e^{-(\beta J/2)R}] \\ \leq 2|Y(\Delta_0)|e^{-(\beta J/2)R}.\end{aligned} \tag{2.59}$$

Proof. The left-hand side of (2.59) is bounded by the expectation

$$I := e^{(\beta J/2)R}\mu[\|\mu^{\sigma'_{\Lambda'\setminus\Delta}}_{\Delta;\Delta_0} - \mu^{\sigma''_{\Lambda''\setminus\Delta}}_{\Delta;\Delta_0}\|]. \tag{2.60}$$

Let $\chi_{\sigma_{\Delta_0}}$ be the function on $\{-1,1\}^{\Delta}$ equal to 1 if σ'_Δ agrees with σ_{Δ_0} on Δ_0 and 0 otherwise. Then

$$\|\mu^{\sigma'_{\Lambda'\setminus\Delta}}_{\Delta;\Delta_0} - \mu^{\sigma''_{\Lambda''\setminus\Delta}}_{\Delta;\Delta_0}\| = \sum_{\sigma_{\Delta_0}} |\mu^{\sigma'_{\Lambda'\setminus\Delta}}_{\Delta}[\chi_{\sigma_{\Delta_0}}] - \mu^{\sigma''_{\Lambda''\setminus\Delta}}_{\Delta}[\chi_{\sigma_{\Delta_0}}]|.$$

The variation norm can be bounded using the set $\mathcal{A}_\sigma$ as in the proof of Theorem 2.13. Recall that $\sigma = (\sigma'_{\Lambda'}, \sigma''_{\Lambda''})$ but by its definition it only depends on their restriction to $\Delta\setminus\Delta_0$, we shall thus write $\mathcal{A}_{\sigma'_{\Delta\setminus\Delta_0},\sigma''_{\Delta\setminus\Delta_0}}$. Call $\mu^{\sigma'_{\Lambda'\setminus\Delta},\sigma''_{\Lambda''\setminus\Delta}}_{\Delta}$ the product of $\mu^{\sigma'_{\Lambda'\setminus\Delta}}_{\Delta}$ and $\mu^{\sigma''_{\Lambda''\setminus\Delta}}_{\Delta}$, then, as in the proof of Theorem 2.13, the quantity $|\mu^{\sigma'_{\Lambda'\setminus\Delta}}_{\Delta}[\chi_{\sigma_{\Delta_0}}] - \mu^{\sigma''_{\Lambda''\setminus\Delta}}_{\Delta}[\chi_{\sigma_{\Delta_0}}]|$ is bounded by

$$\mu^{\sigma'_{\Lambda'\setminus\Delta},\sigma''_{\Lambda''\setminus\Delta}}_{\Delta}[(\chi'_{\sigma_{\Delta_0}} + \chi''_{\sigma_{\Delta_0}}); \mathcal{A}_{\sigma'_{\Delta\setminus\Delta_0},\sigma''_{\Delta\setminus\Delta_0}} \neq \varnothing].$$

Hence,

$$\|\mu^{\sigma'_{\Lambda'\setminus\Delta}}_{\Delta;\Delta_0} - \mu^{\sigma''_{\Lambda''\setminus\Delta}}_{\Delta;\Delta_0}\| \leq 2\mu^{\sigma'_{\Lambda'\setminus\Delta},\sigma''_{\Lambda''\setminus\Delta}}_{\Delta}[\mathcal{A}_{\sigma'_{\Delta\setminus\Delta_0},\sigma''_{\Delta\setminus\Delta_0}} \neq \varnothing].$$

Going back to (2.60) and using (2.56), we get

$$I \leq e^{(\beta J/2)R}2\mu[\mathcal{A}_{\sigma'_{\Delta\setminus\Delta_0},\sigma''_{\Delta\setminus\Delta_0}} \neq \varnothing] \leq e^{(\beta J/2)R}2|Y(\Delta_0)|e^{-\beta JR} \tag{2.61}$$

which concludes the proof of the theorem. □

2.6.2 *Proof of Theorem 2.5.* Fix arbitrarily an increasing sequence Λ_n which invades $\mathbb{Z}^d$. Let Δ_0 be a cube, $f \in \mathcal{C}_{\Delta_0}$, $\|f\| \leq 1$, Δ a cube of side L with same center as Δ_0, suppose that Λ_n and $\Lambda_{n'}$ contain Δ. Then by Theorem 2.13,

$$|\mu^+_{\beta,\Lambda_n}[f] - \mu^+_{\beta,\Lambda_{n'}}[f]| \leq 2|Y(\Delta_0)|e^{-\beta JR}. \tag{2.62}$$

The r.h.s. vanishes as $\Delta \to \mathbb{Z}^d$, therefore $E_{\mu^+_{\beta,\Lambda_n}}[f]$ is a bounded Cauchy sequence, call a_f its limit. We have thus proved that for any cylindrical function f (i.e., which depends only on finitely many spins)

$$\lim_{n\to\infty} \mu^+_{\beta,\Lambda_n}[f] = a_f.$$

a_f is a bounded, linear nonnegative functional on the space of cylindrical functions (a property inherited from μ^+_{β,Λ_n}), which by the Caratheodory theorem extends uniquely to a probability μ^+_β such that $\mu^+_\beta[f] = a_f$ for all cylindrical function f. As the cylindrical functions are dense in the set of continuous functions we then conclude that μ^+_{β,Λ_n} converges weakly to μ^+_β. A priori however μ^+_β may depend on the sequence Λ_n used in its definition. However by Theorem 2.13 if $\Lambda \supset \Delta$, $\Lambda_n \supset \Delta$

$$|\mu^+_{\beta,\Lambda}[f] - \mu^+_{\beta,\Lambda_n}[f]| \le 2|Y(\Delta_0)|e^{-\beta J R}$$

and taking $n \to \infty$

$$|\mu^+_{\beta,\Lambda}[f] - \mu^+_\beta[f]| \le 2|Y(\Delta_0)|e^{-\beta J R} \tag{2.63}$$

which proves that μ^+_β is actually independent of the sequence Λ_n used in its definition.

To prove translation invariance, namely that for any f as above

$$\mu^+_\beta[f] = \mu^+_\beta[\tau f], \qquad \tau f \text{ a shift of } f \tag{2.64}$$

we consider a sequence Λ_n as above. Then, since the interaction is translation invariant,

$$\mu^+_\beta[\tau f] = \lim_{n\to\infty} \mu^+_{\beta,\Lambda_n}[\tau f] = \lim_{n\to\infty} \mu^+_{\beta,\tau^{-1}\Lambda_n}[f] = \mu^+_\beta[f]$$

having used in the last equality that the limit does not depend on the sequence.

By the spin flip symmetry, we deduce that also $\mu^-_{\Lambda_n}$ converges to μ^-_β (the spin flip of μ^+_β) independently of the sequence Λ_n. By using (2.41), we then complete the proof of Theorem 2.5.

The above implies an exponential decay of the correlations for the measure μ^+_β (by spin flip the result applies as well to μ^-_β). We keep the same notation used so far denoting by g any function $\|g\| \le 1$ which depends on finitely many sites all outside Δ.

Theorem 2.18. *With the above notation and in the same context of Theorem* 2.13

$$|\mu^+_\beta[fg] - \mu^+_\beta[f]\mu^+_\beta[g]| \le e^{-(\beta J/2)R}\big(1 + 6|Y(\Delta_0)|\big). \tag{2.65}$$

Proof. Let Λ be a cube large enough, eventually $\Lambda \to \mathbb{Z}^2$. By the DLR property

$$\mu^+_{\beta,\Lambda}[fg] = \mu^+_{\beta,\Lambda}[g\mu^{\sigma_{\Lambda\setminus\Delta}}_{\Delta}[f]],$$

where $\mu^{\sigma_{\Lambda\setminus\Delta}}_{\Delta}$ is the Gibbs measure on $\{-1,1\}^{\Delta}$ with boundary condition $\sigma_{\Lambda\setminus\Delta}$. By Theorem 2.17 with $\Lambda' = \Lambda$ and $\Lambda'' = \Delta$, we have

$$\mu^+_{\beta,\Lambda}[\mathcal{B}] \geq 1 - 2|Y(\Delta_0)|e^{-(\beta J/2)R} \tag{2.66}$$

with

$$\mathcal{B} = \{\sigma_{\Lambda\setminus\Delta} : \|\mu^{\sigma_{\Lambda\setminus\Delta}}_{\Delta;\Delta_0} - \mu^+_{\Delta;\Delta_0}\| \leq e^{-(\beta J/2)R}\}.$$

We then get

$$|\mu^+_{\beta,\Lambda}[fg] - \mu^+_{\beta,\Lambda}[g;\mathcal{B}]\mu^+_{\Delta}[f]| \leq e^{-(\beta J/2)R} + 2|Y(\Delta_0)|e^{-(\beta J/2)R}.$$

Using again (2.66)

$$|\mu^+_{\beta,\Lambda}[fg] - \mu^+_{\beta,\Lambda}[g]\mu^+_{\Delta}[f]| \leq e^{-(\beta J/2)R} + 4|Y(\Delta_0)|e^{-(\beta J/2)R}.$$

Letting $\Lambda \to \mathbb{Z}^2$ we get

$$|\mu^+_{\beta}[fg] - \mu^+_{\beta}[g]\mu^+_{\Delta}[f]| \leq e^{-(\beta J/2)R} + 4|Y(\Delta_0)|e^{-(\beta J/2)R}$$

and use (2.63) to get (2.65). □

2.6.3 *Phase separation*. We suppose hereafter that β is as in Theorem 2.5 and drop it from the notation. Let $\Lambda = \{(x_1, x_2) \in \mathbb{Z}^2 : |x_i| \leq L\}$, L a positive integer, and

$$M_\Lambda := \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sigma_\Lambda(x) \tag{2.67}$$

the empirical magnetization in Λ. For any $\zeta > 0$, by the Chebyshev inequality

$$\mu^+[|M_\Lambda - m_\beta| \geq \zeta] \leq \frac{1}{(\zeta|\Lambda|)^2}\sum_{x,y\in\Lambda} E_{\mu^+}[(\sigma(x) - m_\beta)(\sigma(y) - m_\beta)].$$

It then follows from Theorem 2.18 that:

Theorem 2.19. *For any* $\zeta > 0$

$$\mu^+[|M_\Lambda - m_\beta| \geq \zeta] \leq \frac{c}{\zeta^2|\Lambda|} \tag{2.68}$$

c a positive constant.

The theorem shows that the plus measure μ^+ concentrates on configurations whose empirical magnetization is m_β. The phenomenon of phase separation concerns instead what happens to the empirical magnetization when we consider the canonical Gibbs measure with magnetization $m \in (-m_\beta, m_\beta)$. Let μ_Λ, Λ as above, be the Gibbs measure at inverse temperature β with 0 boundary condi-

tions: that is, the Gibbs factor is $e^{-\beta H(\sigma_\Lambda)}$, with no interaction with spins outside Λ. Fix hereafter $m \in (-m_\beta, m_\beta)$ and let $[m|\Lambda|]$ be the largest number in $\{-|\Lambda|, -|\Lambda|+2, \dots, |\Lambda|\}$ which is $\le m|\Lambda|$. Then *the canonical Gibbs measure with magnetization m* is

$$\mu_{m;\Lambda}(\sigma_\Lambda) = \frac{\mu_\Lambda(\sigma_\Lambda)\mathbf{1}_{M_\Lambda|\Lambda|=[m|\Lambda|]}}{\mu_\Lambda[M_\Lambda|\Lambda| = [m|\Lambda|]]}. \tag{2.69}$$

We partition $\mathbb{Z}^2$ into squares of side ℓ and denote by $\Delta_{i;\ell}$, $i = 1, \dots, N$, those which are entirely contained in Λ, thus $N \approx L^2/\ell^2$.

Theorem 2.20. *Call* $I_\zeta := \{x \in [-1,1] : |x - m_\beta| > \zeta, |x + m_\beta| > \zeta\}$. *Then*

$$\lim_{\zeta\to 0}\lim_{\ell\to\infty}\lim_{L\to\infty} \mu_{m,\Lambda}\left[\sum_{i=1}^{N} \mathbf{1}_{M_{\Delta_{i,\ell}}\in I_\zeta} \ge \zeta N\right] = 0. \tag{2.70}$$

It would be too long to prove the theorem in all its details and I will only give some hints which the more experienced readers can use to get a full proof. Surprisingly, we do not need more than what done so far even though here we deal with "wrong" boundary conditions (zero instead of plus–minus boundary conditions), and moreover the canonical constraint is in conflict with the Peierls bounds (as their proof is based on a comparison of events obtained one from the other by spin flip).

However to prove (2.70), it suffices to get upper bounds which vanish in the limit and an upper bound for the numerator of the fraction defining the canonical measure in (2.69) is simply obtained by dropping the characteristic function of the canonical constraint. Moreover, since the interaction of the spins at the boundary of Λ with spins external to Λ is bounded from below by $-cL$ we can bound $\mu_\Lambda(\sigma_\Lambda) \le \mu^+_\Lambda(\sigma_\Lambda)e^{\beta cL}$. The extra exponential can be handled if we have the super-exponential bound

$$\mu^+_\Lambda\left[\sum_{i=1}^{N} \mathbf{1}_{M_{\Delta_{i,\ell}}\in I_\zeta} \ge \zeta N\right] \le e^{-b(\ell)\zeta N} \tag{2.71}$$

with $b(\ell) > 0$. Recall that $N \approx L^2/\ell^2$ so that for ζ and ℓ fixed the bound in (2.71) is exponential in $-L^2$ and it thus wins against $e^{\beta cL}$. We shall prove that the denominator in (2.69) is bounded from below by

$$\mu_\Lambda\big[M_\Lambda|\Lambda| = \big[m|\Lambda|\big]\big] \ge e^{-aL}, \qquad a > 0 \tag{2.72}$$

so that the bound (2.71) covers all these exponential bounds.

The event in (2.71) requires the existence of at least ζN squares $\Delta_{i,\ell}$ which are "bad". We shall easily see that bad squares have small probability but the crucial point is to derive a bound for the simultaneous appearance of bad squares in terms of a product of their probabilities. Notice however that we also need to control an entropic factor due to the number of ways the bad squares may appear.

Factorization is the crucial point. In fact, the probability of finding deviations from m_β alone:

$$\mu_\Lambda^+\left[\sum_{i=1}^N \mathbf{1}_{|M_{\Delta_{i,\ell}}-m_\beta|\geq\zeta} \geq \zeta N\right]$$

is only exponentially and not super-exponentially small as in (2.71). This can be seen by bounding $\mu_\Lambda^+(\sigma_\Lambda) \leq \mu_\Lambda^-(\sigma_\Lambda)e^{-\beta cL}$ and then using Theorem 2.19 for μ_Λ^- to get an upper bound of the form $c'e^{-\beta cL}$.

It is thus essential for the bound (2.71) to deal with deviations of the magnetization both from m_β and $-m_\beta$. We shall in fact see that if the empirical magnetization in a square deviates both from m_β and $-m_\beta$ then in that square there is most likely a long piece of a contour.

Notation. Let $R = \sqrt{\ell}$ and $\Delta_{i;\ell}^0$ the cube with same center as $\Delta_{i;\ell}$ and distance (1 plus the integer part of) R from the complement of $\Delta_{i;\ell}$. Let $\mathcal{A}_i(\sigma)$ be the family of connected sets A such that $A \cup Y(A) \subset \Delta_{i;\ell}$ and $\Delta_{i;\ell}^0 \subset A$ and either $\sigma(x) \equiv 1$ on $Y(A)$ or $\sigma(x) \equiv -1$ on $Y(A)$; as before there is a maximal element in $\mathcal{A}_i(\sigma)$ and this is a one-connected set. Calling $\mathcal{K}_i$ the family of one-connected sets C such that $C \cup Y(C) \subset \Delta_{i;\ell}$ and $\Delta_{i;\ell}^0 \subset C$ we denote by $G_{C,i}$ the set of configurations σ such that C is the maximal element in $\mathcal{A}_i(\sigma)$, calling $G_{C,i}^\pm$ the set of spin configurations such that $\sigma_\Lambda(x) \equiv 1$, respectively, $\sigma_\Lambda(x) \equiv -1$, for all $x \in Y(C)$.

"Good and bad cubes". *The cube* $\Delta_{i;\ell}$ *is* $[\zeta > 0]$ *good if it is neither* 1*-bad nor* 2*-bad*:

- $\Delta_{i;\ell}$ is 1-bad if $\mathcal{A}_i(\sigma) = \varnothing$,
- $\Delta_{i;\ell}$ is 2-bad if $\mathcal{A}_i(\sigma) \neq \varnothing$ but $M_{\Delta_{i,\ell}} \in I_\zeta$.

Since the bad cubes include those where $M_{\Delta_{i,\ell}} \in I_\zeta$, it suffices to prove that

$$\mu_\Lambda^+\left[\sum_{i=1}^N \mathbf{1}_{\Delta_{i,\ell} \text{ is bad}} \geq \zeta N\right] \leq e^{-b(\ell)\zeta N}. \tag{2.73}$$

We obviously have

$$\begin{aligned}\mu_\Lambda^+&\left[\sum_{i=1}^N \mathbf{1}_{\Delta_{i,\ell} \text{ is bad}} \geq \zeta N\right] \\ &\leq \sum_{n\geq(\zeta/2)N} \sum_{i_1,\dots,i_n} \left\{\mu_\Lambda^+\left[\bigcap_{j=1}^n \Delta_{i_j,\ell} \text{ is 1-bad}\right]\right. \\ &\qquad \left. + \mu_\Lambda^+\left[\bigcap_{j=1}^n \Delta_{i_j,\ell} \text{ is 2-bad}\right]\right\}.\end{aligned} \tag{2.74}$$

We start from the 2-bad squares. Let $\Delta \equiv \Delta_{i_j,\ell}$ be one of the 2-bad squares, fix $C \in \mathcal{K}_{i_j}$ and consider the set G^+_{C,i_j} (the argument for G^-_{C,i_j} is similar). Since G^+_{C,i_j} is measurable on the complement of C the conditional probability of σ_C given the values of all the spins outside C is $\mu^+_C(\sigma_C)$. Since Δ is 2-bad,

$$\frac{1}{|\Delta|}\Big|\sum_{x\in\Delta}\big(\sigma_\Delta(x)-m_\beta\big)\Big| \ge \zeta$$

so that for ℓ large enough

$$\frac{1}{|\Delta_0|}\Big|\sum_{x\in\Delta_0}\big(\sigma_{\Delta_0}(x)-m_\beta\big)\Big| \ge \frac{\zeta}{2}.$$

By (2.68), there is a constant $c>0$ so that

$$\mu^+_C\Big[\frac{1}{|\Delta_0|}\Big|\sum_{x\in\Delta_0}\big(\sigma_{\Delta_0}(x)-m_\beta\big)\Big| \ge \frac{\zeta}{2}\Big] \le \frac{c}{\zeta\ell}$$

and as a consequence

$$\mu^+_\Lambda\Big[\bigcap_{j=1}^{n}\Delta_{i_j,\ell} \text{ is 2-bad}\Big] \le \Big(\frac{c}{\zeta\ell}\Big)^n. \tag{2.75}$$

It then follows that

$$\begin{aligned}\sum_{n\ge(\zeta/2)N}\sum_{i_1,\dots,i_n}\mu^+_\Lambda\Big[\bigcap_{j=1}^{n}\Delta_{i_j,\ell} \text{ is 2-bad}\Big] &\le \sum_{n\ge(\zeta/2)N}\binom{N}{n}\Big(\frac{c}{\zeta\ell}\Big)^n \\ &\le e^{-b\zeta N}, \qquad b>0\end{aligned} \tag{2.76}$$

if ℓ is large enough.

The bound of the 1-bad squares is more involved. If $\mathcal{A}_i(\sigma_\Lambda)=\varnothing$ it means that there is neither a plus nor a minus circuit in $\Delta\setminus\Delta_0$ ($\Delta=\Delta_{i,\ell}$) which contains Δ_0 in its interior. This implies that there are two paths which connect Δ_0 and Δ^c along which the spins are constant, in one of the paths the spins being identically plus and in the other identically minus.

It remains to prove (2.72). We split $\Lambda=\Lambda_1\cup\Lambda_2\cup\Lambda_3$ into three disjoint rectangles of vertical side $2L+1$ and horizontal side L_i, $i=1,2,3$. We set $L_2=k$, k a large integer which will be specified later which is fixed independently of L. L_1 and L_3 are chosen in such a way that

$$m_\beta\big(|\Lambda_1|-|\Lambda_3|\big) \qquad \text{is as close as possible to } m\big(|\Lambda_1|+|\Lambda_3|\big).$$

Calling $\mu^\pm$ the infinite volume plus–minus Gibbs measures and let $\nu^+_{\Lambda_1}(\sigma_{\Lambda_1})$ be the marginal probability of μ^+ over $\{-1,1\}^{\Lambda_1}$, $\nu^+_{\Lambda_3}(\sigma_{\Lambda_3})$ that of μ^- over $\{-1,1\}^{\Lambda_3}$.

Call $\mu_{\Lambda_i}^{+,\sigma_{\Lambda_2}}(\sigma_{\Lambda_i})$, $i = 1, 3$, the μ_{Λ}^{+} conditional probabilities of σ_{Λ_i} given σ_{Λ_2}. Then there exists a constant c so that for any σ_{Λ_2}

$$\mu_{\Lambda_1}^{+,\sigma_{\Lambda_2}}(\sigma_{\Lambda_1}) \geq e^{-\beta cL}\nu_{\Lambda_1}^{+}(\sigma_{\Lambda_1}), \qquad \mu_{\Lambda_3}^{+,\sigma_{\Lambda_2}}(\sigma_{\Lambda_3}) \geq e^{-\beta cL}\nu_{\Lambda_3}^{-}(\sigma_{\Lambda_3}).$$

Analogously to Theorem 2.19, one can prove that there is a constant $a > 0$ so that

$$\nu_{\Lambda_1}^{+}\big(|M_{\Lambda_1} - m_\beta||\Lambda_1| \leq aL\big) \geq \tfrac{1}{2}, \qquad \nu_{\Lambda_3}^{-}\big(|M_{\Lambda_3} + m_\beta||\Lambda_3| \leq aL\big) \geq \tfrac{1}{2}.$$

Suppose that σ_{Λ_i}, $i = 1, 3$ are such that $|M_{\Lambda_1} - m_\beta||\Lambda_1| \leq aL$ and $|M_{\Lambda_3} - m_\beta||\Lambda_3| \leq aL$, then if $L_2 = k$ is large enough there is σ_{Λ_2} so that

$$M_{\Lambda_1}|\Lambda_1| + M_{\Lambda_3}|\Lambda_3| + M_{\Lambda_3}|\Lambda_3| = \big[|\Lambda|m\big].$$

Hence,

$$\mu_{\Lambda}^{+}\big[M_\Lambda = \big[m|\Lambda|\big]\big] \geq e^{-2\beta cL}\tfrac{1}{4}\big(2^{-|\Lambda_2|}e^{-2\beta J|\Lambda_2|}\big)$$

the last factor bounding the probability of σ_{Λ_2}. In this way, one can prove (2.72), details are omitted.

2.7 Bibliographical remarks

The classical reference for Gibbs measures and equilibrium statistical mechanics is Ruelle's book (Ruelle (1969)), and his paper (Ruelle (1970)) on superstable interactions, for the existence of the thermodynamic limit for the pressure.

The Peierls argument can be found in Georgii's book (Georgii (1988)). The extension to more general interaction is the content of the Pirogov–Sinai theory, see Sinai (1982). The analysis of the Ising phase transition in the canonical ensemble goes back to the work by Minlos and Sinai (1968), and later by Dobrushin, Kotecký and Shlosman on their fundamental paper on the Wulff construction (Dobrushin, Kotecký and Shlosman (1992)).

3 Lecture 3

The extension to positive temperatures of the zero temperature phase transitions for particles systems in $\mathbb{R}^d$ is still a completely open problem, one of the most important in equilibrium statistical mechanics. As seen in Lecture 2, the occurrence of a phase transition is proved for lattice gases at low temperatures. The main features used in the proof are:

- A symmetry between phases, that is, between empty and occupied sites (which in the Ising language is the spin flip symmetry).
- The energy excitations have a positive gap, that is, any perturbation of the ground state carries an energy bounded away from 0.

- The excitations can be classified in terms of contours, contours carry an excess energy proportional to their length and are simple geometrical objects (connected sets in $\mathbb{Z}^2_*$).

For particle systems in the continuum, the above properties fail. We do not have nor expect a symmetry between the solid and the vapour phases, but this is maybe not really essential. There is a very robust theory, the Pirogov–Sinai theory (Sinai (1982)), Zahradnik (1998), which covers cases where the symmetry is broken.

In the continuum, perturbations of the ground state can be infinitesimal and therefore the energy excitations cannot be bounded away from zero. Also this however has been overcome as there are examples of phase transitions in continuum spin systems.

The real difficulty when dealing with continuum particle systems seems to be the very complex nature of their excitations. We have to deal with phonons, that is, compression and rarefaction waves, dislocations, fractures and all the other phenomena which arise in elasticity. The excess energy is no longer well localized on contours as in the lattice case and the complexity of their structure has so far defeated all efforts to extend the analysis to the continuum.

In conclusion, the analysis of the gas–solid transition curve at low temperatures seems to require new ideas and techniques and the existence of a gas–solid phase transition in continuum systems is among the most important open problems in statistical mechanics. In this lecture, I shall discuss the liquid–vapour branch of the phase diagram. Both the physics and the mathematics of the liquid–vapour transition are quite different than the gas–solid transition at small temperatures.

I start by recalling the van der Waals equation of state and the mean field theory. We begin with the perfect gas and then analyze the corrections which appear when the interaction among molecules is taken into account.

3.1 The ideal and the imperfect gas

The equation of state of an ideal or perfect gas is $PV = NkT$, maybe one of the most famous laws in physics. Let us derive it from the postulates of statistical mechanics by computing the gran canonical pressure of the ideal gas, that is, a system of noninteracting particles. Recalling (2.28)

$$Z_{\beta,\lambda;\Lambda} = \sum_{n\geq 0} \frac{1}{n!} \int_{\Lambda^n} e^{\beta\lambda n}\, dx_1 \cdots dx_n = \exp\{|\Lambda| e^{\beta\lambda}\}.$$

The average number of particles computed using the corresponding Gibbs measure is

$$Z^{-1}_{\beta,\lambda;\Lambda} \sum_{n\geq 0} \frac{n}{n!} \int_{\Lambda^n} e^{\beta\lambda n}\, dx_1 \cdots dx_n = |\Lambda| e^{\beta\lambda}$$

so that the density is $\rho = e^{\beta\lambda}$ and the pressure according to (2.24) is

$$\pi_{\beta,\lambda} = \frac{\log Z_{\beta,\lambda;\Lambda}}{\beta|\Lambda|} = \frac{e^{\beta\lambda}}{\beta} = \beta^{-1}\rho \tag{3.1}$$

which is indeed the equation of state $PV = NkT$ of a perfect gas.

When the interaction among particles is not disregarded the gas is called "imperfect". The equation of state of an imperfect gas must thus take into account the influence of the long attractive tail of the molecules interaction as well as its short range repulsive part, as described, for instance, by the Lennard–Jones potential, see (1.10). The van der Waals equation of state is:

$$\left(P + \alpha\frac{\rho^2}{2}\right)(V - aN) = NkT, \qquad \alpha, a > 0. \tag{3.2}$$

The heuristics behind (3.2) is the following. In an ideal gas, the pressure P is the force (per unit surface) that the walls exert to keep the system in the volume V, in an imperfect gas the attractive forces between molecules add to the pressure hence the term $\alpha\rho^2/2$ proportional to the number of interacting pairs of molecules. The repulsive part of the interaction is schematized in the van der Waals equation by regarding the particles as hard, impenetrable, small spheres. Thus if there are N particles in a volume V, the free volume available to each particle is not V but rather $V - Na$, a the volume occupied by each particle.

Rewrite (3.2) as

$$P = -\alpha\frac{\rho^2}{2} + \beta^{-1}\frac{\rho}{1 - a\rho}. \tag{3.3}$$

According to the laws of thermodynamics P should be a nondecreasing function of ρ (as $P = \rho f'(\rho) - f(\rho)$, $P' = \rho f''$ and f, the free energy, is a convex function of ρ, $f''(\rho) \geq 0$) while for β large P' has a graph like in Figure 11.

The van der Waals equation of state must therefore be modified and Maxwell proposed to do that by cutting the graph of $P'(\rho)$ along the dashed line as in

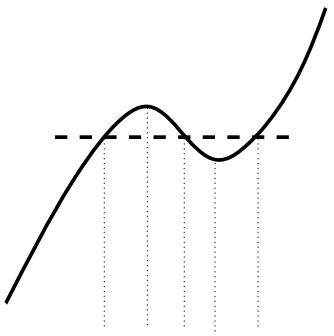

Figure 11 *$P' = dP/d\rho$ as a function of ρ for large β.*

Figure 11 chosen so that the parts above and below the dashed line have the same area. This is the famous "Maxwell equal area rule". In this way, $P'(\rho) = 0$ in the interval where the graph has been cut and this corresponds in thermodynamics to a forbidden interval and a phase transition. Van der Waals interpreted this as a liquid–vapour phase transition.

3.2 The mean field theory

The van der Waals argument is certainly clever and appealing in its simplicity but definitely heuristic, hence the many attempts to justify it. I shall present below the one based on the mean field approximation.

The basic assumption of the mean field theory is that (i) any particle interacts with any other in the same way (independently of their positions); (ii) the total interaction of a particle with all the others is uniformly bounded. Notice that by (i) and (ii) the interaction must depend on the region where the system is confined (which is the really weak point of the theory), indeed the mean field Hamiltonian is:

$$H_\Lambda^{\mathrm{mf}}(\underline{x}) = -\frac{1}{2}\sum_{i\neq j}\frac{\alpha}{|\Lambda|}\mathbf{1}_{x_i,x_j\in\Lambda} = -\left(\alpha\frac{|\underline{x}|(|\underline{x}|-1)}{2|\Lambda|^2}\right)|\Lambda| \tag{3.4}$$

so that if the number of particles $|\underline{x}|$ grows like $\rho|\Lambda|$, then $H_\Lambda^{\mathrm{mf}}(\underline{x}) \approx -\alpha\frac{\rho^2}{2}|\Lambda|$ and the energy per particles is uniformly bounded.

The theory with the Hamiltonian (3.4) is however ill-posed, as the gran canonical partition function (2.28) diverges (because the Gibbs factor grows as e^{cn^2} and it thus dominates the convergent factor $1/n!$). The traditional way to make the sum convergent is consistent with the van der Waals argument as it is based on identifying a particle at x with a "hard ball" of radius $R > 0$ and center x with the condition that different balls cannot intersect. This is the same as adding to the Hamiltonian a hard core pairwise interaction which is infinite if $|x_i - x_j| \le 2R$ and $=0$ otherwise. Having added a local interaction, we are no longer in the simple context of the mean field theory, the composite Hamiltonian is far from simple as the hard core system alone is very complex. This is true but in $d = 1$ where the canonical partition function $Z \equiv Z_{\beta,N,\Lambda}$ can be computed exactly. Calling $a = 2R$ the volume occupied by a particle (recall $d = 1$) we have

$$\begin{aligned} Z &= e^{\beta[\alpha N(N-1)/(2|\Lambda|^2)]|\Lambda|}\int \mathbf{1}_{x_{i+1}>x_i+a,i=1,\dots,N-1;0\le x_1;|\Lambda|\ge x_N}\,dx_1\cdots dx_N \\ &= e^{\beta[\alpha N(N-1)/(2|\Lambda|^2)]|\Lambda|}\frac{1}{N!}\int_{[0,|\Lambda|-(N-1)a]^N} dx_1\cdots dx_N. \end{aligned} \tag{3.5}$$

The first equality is obtained by ordering the position of the particles so that the $N!$ term (in the definition of Z) disappears. The second equality is obtained by the change of variables $x_i \to x_i + (i-1)a$ and then dropping the ordering of the particles so that the $N!$ term reappears.

We can now perform the thermodynamic limit getting

$$\begin{aligned} f_{\beta,\rho} &= \lim_{|\Lambda|\to\infty, N/|\Lambda|\to\rho} \frac{-\log Z_{\beta,N,\Lambda}}{\beta|\Lambda|} \\ &= -\alpha\frac{\rho^2}{2} + \frac{1}{\beta}\{\rho(\log\rho - 1) - \rho\log(1-\rho a)\}. \end{aligned} \tag{3.6}$$

Let $\lambda = df_{\beta,\rho}/d\rho$ and $\pi_{\beta,\rho} = \lambda\rho - f_{\beta,\rho}$ then one can check that $\pi_{\beta,\rho} = P$ as given in (3.3) and that the equal area Maxwell rule is equivalent to taking the convexification $f^{**}_{\beta,\rho}$ of $f_{\beta,\rho}$.

On the other hand, if we work directly in the gran canonical ensemble and compute the thermodynamical limit of the gran ganonical pressure we find it equal to the convexification $f^{**}_{\beta,\rho}$ of $f_{\beta,\rho}$, thus the grand canonical ensemble proves the equal area Maxwell rule.

With the van der Waals theory and its statistical mechanics interpretation in terms of mean field, we have thus recovered phase transitions. There are however three main drawbacks in the derivation:

- The Hamiltonian depends on the volume and it is not well defined in the thermodynamic limit $|\Lambda| \to \infty$.
- The computation of the thermodynamic potentials is explicit only in $d = 1$, in $d > 1$ one can still prove the existence of a phase transition but not as explicitly.
- The Maxwell equal area rule or the convexification of the free energy $f_{\beta,\rho}$ is indeed derived working within the gran canonical ensemble, but in mean field the equivalence of ensembles is not valid and there is no reason to prefer gran canonical to canonical.

3.2.1 *Variants of the mean field model.* The addition of hard cores is not needed in the lattice gas as the variables are bounded. The Hamiltonian of the corresponding Ising model is (in any dimension)

$$H_\Lambda(\sigma) = -\frac{\alpha}{2|\Lambda|}\sum_{x\neq y}\sigma(x)\sigma(y).$$

The mean field free energy density is then

$$f_{\beta,m} = -\alpha\frac{m^2}{2} - \frac{S(m)}{\beta}, \tag{3.7}$$

$$S(m) = \frac{-1}{\beta}\left\{\frac{1-m}{2}\log\frac{1-m}{2} + \frac{1+m}{2}\log\frac{1+m}{2}\right\}. \tag{3.8}$$

The mean field pressure (obtained by working in the gran canonical ensemble) is the Legendre transform of $f_{\beta,m}$.

In the continuum we may avoid hard cores by adding mean field repulsive forces. The mean field LMP Hamiltonian is

$$H_\Lambda(\underline{x}) = |\Lambda|\Big(-\frac{\rho^2}{2} + \frac{\rho^4}{4!}\Big), \qquad \rho := \frac{|\underline{x}|}{|\Lambda|}, \tag{3.9}$$

where the quadratic term comes from pair interactions and the quartic from four body interactions. All interactions are supposed independent of the positions of the particles (with self interactions included in (3.9)).

The LMP mean field free energy density is then

$$f_{\beta,\rho} = -\frac{\rho^2}{2} + \frac{\rho^4}{4!} + \frac{1}{\beta}\rho(\log\rho - 1). \tag{3.10}$$

The LMP mean field pressure (obtained by working in the gran canonical ensemble) is the Legendre transform of $f_{\beta,m}$.

Both the Ising and the LMP model (in their mean field version) have a phase transition of the van der Waals type.

3.3 Kac potentials

The mean field free energy may be a reasonable approximation of the true one in regions large enough to be considered thermodynamically relevant yet still small relatively to the range of the attractive forces. If such conditions are met, one may see a double well free energy as the one appearing in the mean field theory. If however we look at the behavior of the same system on much larger regions, we will then see a "thermodynamic correct" behavior of the free energy which will rather look like the convexification of the mean field shape. Such a *two scales behavior* will be the center of our discussion in the remaining part of this lecture.

Following Kac, we quantify the above ideas by considering an attractive pair interaction constant in a ball of radius γ^{-1} and equal to 0 outside, thus getting

$$H(\underline{x}) = -\frac{1}{2}\sum_{i\neq j}\gamma^d \mathbf{1}_{|x_i - x_j|\leq\gamma^{-1}}$$

(γ^d the inverse of the volume of the interaction ball, to be compared to the mean field Hamiltonian). The above gives indeed a mean field behavior in regions of diameter suitably smaller than γ^{-1}. To remove the discontinuity, we will replace the above Hamiltonian by

$$H(\underline{x}) = -\frac{1}{2}\sum_{i\neq j} V_\gamma(x_i, x_j), \tag{3.11}$$

where

$$V_\gamma(r, r') = \gamma^d V(\gamma r, \gamma r') \tag{3.12}$$

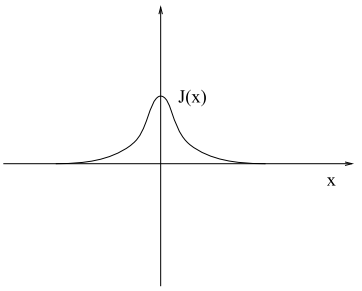

Figure 12 *$J(x)$ denotes the function $J(0,r)$ when r varies along the x-axis.*

with $V(r,r')$ a symmetric, translation invariant ($V(r,r') = J(0,r'-r)$) smooth probability kernel (see Figure 12 where $J(0,r)$ stands for $V(0,r)$) which we suppose for simplicity to vanish for $|r-r'| \geq 1$, $\gamma > 0$ is "the Kac scaling parameter".

As in mean field, we need to add to the Kac potentials repulsive forces to ensure stability. In the Ising model, this is not necessary because the variables are bounded:

3.3.1 *Ising model with Kac potentials.* The Ising Hamiltonian with Kac potentials in a bounded region $\Lambda \subset \mathbb{Z}^d$ is

$$H_\Lambda(\sigma) = -\frac{1}{2}\sum_{x\neq y} J_\gamma(x,y)\sigma(x)\sigma(y), \qquad \sigma \in \{-1,1\}^\Lambda \tag{3.13}$$

with J_γ having the same properties stated before for V_γ. Notice that

$$\int J_\gamma(r,r')\,dr' = \int J(r,r')\,dr' = 1. \tag{3.14}$$

The following theorem holds true:

Theorem 3.1. *In $d \geq 2$ there is $\beta_{c,\gamma} > 0$ and for any $\beta > \beta_{c,\gamma}$ there is $m_{\beta,\gamma} > 0$ so that $(-m_{\beta,\gamma}, m_{\beta,\gamma})$ is the only forbidden magnetization interval. Moreover*

$$\lim_{\gamma\to 0} \beta_{c,\gamma} = 1, \qquad \lim_{\gamma\to 0} m_{\beta,\gamma} = m_\beta, \tag{3.15}$$

where

$$m_\beta = \tanh\{\beta m_\beta\} > 0, \qquad \beta > 1 \tag{3.16}$$

and $(-m_\beta, m_\beta)$ is the forbidden interval for the mean field model whose critical temperature is 1.

The theorem accomplishes the Kac program to derive the van der Waals or the mean field thermodynamics in a statistical mechanics context. Indeed for any

$\gamma > 0$ the Ising Hamiltonian (3.13) is a legitimate finite range Hamiltonian and the system exhibits a phase transition like its mean field analogue.

A weaker statement, namely that the thermodynamic pressure $\pi_{\beta,h;\gamma}$ (h an external magnetic field) converges to the mean field pressure $\pi^{\mathrm{mf}}_{\beta,h}$ is valid in much more generality (the Kac and Lebowitz–Penrose theory), but it does not ensure that the phase transition in the limit is also present before the limit, as a sequence of strictly convex functions may converge to a function which is linear in a segment. In fact, this is what happens in $d = 1$ where by Theorem 2.4 there is no phase transition.

We shortly discuss the ideas of the proofs after stating the analogous result for the LMP model.

3.3.2 *The LMP model.* As mentioned earlier, a way to "stabilize" the Kac Hamiltonian (3.11) is to add a hard core interaction. This is what done originally by Kac but for such a system we only have a phase transition after taking the limit $\gamma \to 0$ and miss a theorem like Theorem 3.1. In the LMP context instead, the whole approach works fine. A key notion in the definition of Kac potential for LMP is the following: given a particles configuration $\underline{x}$ introduce its *local particles density* $\rho_\gamma(r;\underline{x})$ as

$$\rho_\gamma(r;\underline{x}) = J_\gamma * \underline{x}(r) := \sum_i J_\gamma(r, x_i), \qquad r \in \mathbb{R}^d, \tag{3.17}$$

where J_γ is as above. If $J_\gamma(r,r') = \frac{\mathbf{1}_{|r-r'|\le\gamma^{-1}}}{|B_{\gamma^{-1}}(r)|}$, namely the indicator of the ball $B_{\gamma^{-1}}(r)$ of radius γ^{-1} divided by the volume $|B_{\gamma^{-1}}(r)| = \gamma^{-d}|B_1|$ of the ball ($|B_1|$ the volume of the unit ball), then

$$J_\gamma * \underline{x}(r) := \frac{|\underline{x} \cap B_{\gamma^{-1}}(r)|}{|B_{\gamma^{-1}}(r)|}$$

would be the "empirical" particles density in the ball $B_{\gamma^{-1}}(r)$.

The basic assumption in the LMP model is that the Hamiltonian $H_\gamma(\underline{x})$ is a function of the local particles density $\rho_\gamma(r;\underline{x})$

$$H_\gamma(\underline{x}) = \int e_\lambda\big(\rho_\gamma(r;\underline{x})\big)\,dr, \tag{3.18}$$

where $e_\lambda(\rho)$ is an assigned energy density (which includes the chemical potential contribution). The specific choice in LMP is

$$e(\rho) = -\frac{\rho^2}{2} + \frac{\rho^4}{4!} - \lambda\rho. \tag{3.19}$$

As already mentioned, the LMP Hamiltonian may be written as a sum of two and four body interactions whose range is $2\gamma^{-1}$.

Theorem 3.2. *In $d \geq 2$, there are $0 < \beta_c < \beta_0$ and for any $\beta \in (\beta_c, \beta_0)$ for all γ small enough there are $\rho_{\beta,\gamma,-} < \rho_{\beta,\gamma,+}$ so that the interval $(\rho_{\beta,\gamma,-}, \rho_{\beta,\gamma,+})$ is a forbidden density interval. Moreover, as $\gamma \to 0$ the forbidden interval converges to the mean field forbidden interval.*

3.4 Free energy functionals

The proof of Theorems 3.1 and 3.2 uses renormalization group ideas to reduce to a Ising like system at small temperatures where the approach described in Lecture 2 can be reproduced. The role of the 0 temperature ground states is played by the minimizers of a free energy functional. In this section, we shall describe these functionals.

3.4.1 *The mesoscopic limit.* As discussed in Lecture 1, the states in a continuum theory are order parameter valued functions on the domain occupied by the system, in our case states will be mass-density profiles $\rho(r)$ or magnetization profiles $m(r)$. The continuum theory is then characterized by a free energy functional which specifies the free energy of any state (i.e., any density or magnetization profile).

Continuum theories are derived from microscopic particle models in a regime where the size of the spatial domain (where the system is confined) is much larger than the inter-molecular distances. We shall distinguish between mesoscopic and macroscopic continuum limits starting here from the former.

The mesoscopic regime describes systems where the ratio between the range of the (Kac) interaction and the inter-particles distance diverges while the ratio between the size of the domain and the range of the interaction stays bounded. Thus in mesoscopic spatial units the domain is a bounded region, for simplicity the torus $\mathcal{T}_L$. In microscopic units, the same domain becomes the torus $\mathcal{T}_{\gamma^{-1}L}$. Mesoscopic states are functions $\rho(r), r \in \mathcal{T}_L$, microscopic states are configurations $\underline{x}$ of particles positions in $\mathcal{T}_{\gamma^{-1}L}$. They can be related as follows. Let $\mathcal{D}^{(\ell)}$ be a partition of $\mathbb{R}^d$ into cubes of side ℓ, $C_r^{(\ell)}$ denoting the cube of the partition which contains r. Then the empirical density profile (with grain ℓ) associated to a particles configurations $\underline{x}$ is the function

$$A^{(\ell)}(r|\underline{x}) := \frac{1}{|C_r^{(\ell)}|} \sum_i \mathbf{1}_{x_i \in C_r^{(\ell)}} \tag{3.20}$$

$r \in \mathbb{R}^d$ which is constant on each atom of $\mathcal{D}^{(\ell)}$. We may suppose for simplicity that $\mathcal{T}_{\gamma^{-1}L}$ is $\mathcal{D}^{(\ell)}$ measurable so that $A^{(\ell)}(r|\underline{x}) = 0$ outside $\mathcal{T}_{\gamma^{-1}L}$.

The choice of ℓ is quite arbitrary, we only demand that it is so large that fluctuations are negligible but vanishingly smaller when compared to the range γ^{-1} of the interaction (so that in interaction range units $C_r^{(\ell)}$ looks like a point). For the sake of definiteness, we take $\ell = \gamma^{-1/2}$. The empirical density profile defined in

this way becomes a mesoscopic state if we rescale space: $A^{(\ell)}(r|\underline{x}) \to A^{(\ell)}(\gamma r|\underline{x})$. Then the statistical weight of a mesoscopic state $\rho(\cdot)$ is defined as the constrained partition function

$$\begin{aligned} Z_{L,\gamma,\rho;\zeta} &:= \sum_{n\geq 0} \frac{1}{n!} \int_{(\gamma^{-1}\mathcal{T}_L)^n} e^{-\beta H_\gamma(\underline{x})} \mathbf{1}_{\|A^{(\ell)}(\gamma r|\underline{x})-\rho(r)\|_1<\zeta}\, d\underline{x}, \\ \ell &= \gamma^{-1/2}, \end{aligned} \tag{3.21}$$

where H_γ is the LMP Hamiltonian with $\lambda = 0$. We can then define the mesoscopic free energy of the profile ρ as

$$F_{\mathcal{T}_L}(\rho) := -\lim_{\zeta\to 0}\lim_{\gamma\to 0} \frac{\gamma^d}{\beta} \log Z_{L,\gamma,\rho;\zeta}. \tag{3.22}$$

The Lebowitz–Penrose technique allows to show that

$$F_{\mathcal{T}_L}(\rho) = \int_{\mathcal{T}_L} \{e_0\big(J*\rho(r)\big) + \frac{1}{\beta}\rho(r)\big[\log\rho(r)-1\big]\, dr, \tag{3.23}$$

where $e_\lambda(\cdot)$ is defined in (3.19) (analogous expressions have been derived by Lebowitz–Penrose for a large class of systems).

An analogous procedure in the Ising case leads to the functional

$$F_{\mathcal{T}_L}(m) = \int_{\mathcal{T}_L} \left\{-\frac{1}{2}m(r)[J*m](r) - \frac{S(m(r))}{\beta}\right\} dr, \tag{3.24}$$

where $m(r)$ is a magnetization profile (with values in $[-1,1]$), the entropy $S(m)$ is defined in (3.7). Equation (3.24) can be written in a more significant way as:

$$\begin{aligned} F_{\mathcal{T}_L}(m) = \int_{\mathcal{T}_L} \left\{-\frac{m(r)^2}{2} - \frac{S(m(r))}{\beta}\right\} dr \\ + \frac{1}{4}\int\int J(r,r')\big[m(r)-m(r')\big]^2\, dr\, dr' \end{aligned} \tag{3.25}$$

which reminds of the Ginzburg–Landau functional (1.1). Indeed the latter is obtained by approximating the last term by a gradient squared.

The connection with large deviations goes as follows. Call μ_γ the LMP Gibbs measure (with $\lambda = 0$) on the torus $\gamma^{-1}\mathcal{T}_L$. Then

$$\begin{aligned} &\lim_{\zeta\to 0}\lim_{\gamma\to 0} \gamma^d \log\mu_\gamma\big[\{\|A^{(\ell)}(\gamma r|\underline{x}) - \rho(r)\|_1 < \zeta\}\big] \\ &\qquad = -\beta\big\{F_{\mathcal{T}_L}(\rho) + \pi_{\beta,0}\big\}, \end{aligned} \tag{3.26}$$

where

$$\pi_{\beta,\lambda} = -\inf_{\rho\in\mathbb{R}_+}\left\{e_\lambda(\rho) + \frac{1}{\beta}\big(\rho[\log\rho - 1]\big)\right\}. \tag{3.27}$$

The right-hand side of (3.26) is the large deviation rate function as $\gamma\to 0$ for the probability of observing the profile ρ on the mesoscopic scale (where space is increased by a factor γ^{-1}). An analogous expression holds for the Ising case.

3.4.2 *The macroscopic limit*. The macroscopic regime describes systems where the ratio between the size of the domain and the range of the interaction diverges. For the systems with Kac potentials that we are studying, we also demand that the ratio between the range of the Kac interaction and the inter-particles distances diverges.

We suppose again that in macroscopic units the domain is the torus $\mathcal{T}_L$, in microscopic units the same domain becomes the torus $\mathcal{T}_{\varepsilon^{-1}L}$, $0 < \varepsilon \ll \gamma \ll 1$, for the sake of definiteness we take $\varepsilon = \gamma^{100}$. As in the mesoscopic regime, we associate to a configuration $\underline{x}$ of particles in $\varepsilon^{-1}\mathcal{T}_L$ the density profile

$$A^{(\ell)}(r|\underline{x}) := \frac{1}{|C_r^{(\ell)}|}\sum_i \mathbf{1}_{x_i \in C_r^{(\ell)}}. \tag{3.28}$$

For instance, $\ell = \gamma^{-2}$. Analogously to (3.21), we set

$$Z_{L,\gamma,\rho;\zeta} := \sum_{n\geq 0}\frac{1}{n!}\int_{(\varepsilon^{-1}\mathcal{T}_L)^n} e^{-\beta H_\gamma(\underline{x})]}\mathbf{1}_{\|A^{(\ell)}(\varepsilon r|\underline{x})-\rho(r)\|_1<\zeta}\, d\underline{x} \tag{3.29}$$

and define analogously to (3.22)

$$F^{\text{macro}}_{\mathcal{T}_L}(\rho) := -\lim_{\zeta\to 0}\lim_{\gamma\to 0}\frac{\varepsilon^d}{\beta}\log Z_{L,\gamma,\rho;\zeta}. \tag{3.30}$$

We then have

$$F^{\text{macro}}_{\mathcal{T}_L}(\rho) = \int_{\mathcal{T}_L}\Big(e_0(\rho) + \frac{1}{\beta}\big(\rho[\log\rho - 1]\big)\Big)^{**} dr, \tag{3.31}$$

where g^{**} is the convexification of g.

The relation with the mesoscopic functional (3.23) is not what one may have naively guessed by doing a local approximation of the latter, we refer to Lecture 1 where such questions have been addressed in the context of the Ginzburg–Landau functional.

3.5 Contours and Peierls estimates in LMP

We sketch here the proof of Theorems 3.2 and 3.1. As mentioned, they are based on an extension of the Peierls estimates for the low temperatures Ising model proved in Lecture 2. The extension is far from simple and we only give some semi-heuristic ideas.

Recall that the free energy functional for the LMP model with chemical potential λ is

$$F_{\mathcal{T}_L}(\rho) = \int_{\mathcal{T}_L}\{e_\lambda\big(J*\rho(r)\big) + \frac{1}{\beta}\rho(r)\big[\log\rho(r) - 1\big]\, dr. \tag{3.32}$$

It is easy to prove (see Presutti (2008)) that the minimizers of the functional are spatially constant profiles and that if $\beta > \beta_c$ there exist λ_β, $\rho_{\beta,-} < \rho_{\beta,+}$ so that

the minimum of $F_{\mathcal{T}_L}(\rho)$ is achieved at $\rho_+(r) = \rho_{\beta,+}$ and $\rho_-(r) = \rho_{\beta,-}$ and these are the only minimizers. Recalling from (3.20), the relation between continuum density profiles and particle configurations goes via the empirical density profile $A^{(\ell)}(r|\underline{x})$ and the analysis stands on the ansatz that $A^{(\ell)}(r|\underline{x})$ is close to the equilibrium values $\rho_{\beta,\pm}$. In the nearest neighbor Ising model any spin flip from the constant configuration (all spins equal to 1, or to -1) gives rise to a contour, here we cannot be so strict. Indeed $A^{(\ell)}(r|\underline{x})$ is the average particles density in the cube $C_r^{(\ell)}$, and if $\underline{x}$ is distributed with a nice extremal Gibbs measure then we expect fluctuations of the order $\ell^{-d/2}$ so that we must allow for small deviations before saying that there is a contour.

To be quantitative, we take $\ell = \gamma^{-(1-\alpha)} =: \ell_{-,\gamma}$ with $0 < \alpha \ll 1$ and set $\eta^{(\zeta,\ell_{-,\gamma})}(\underline{x}; r) = \pm 1$ if $|A^{(\ell_{-,\gamma})}(r|\underline{x}) - \rho_{\beta,\pm}| \le \zeta := \gamma^a$, $0 < a \ll \alpha$, and $\eta^{(\zeta,\ell_{-,\gamma})}(\underline{x}; r) = 0$ otherwise. We also introduce the scale $\ell_{+,\gamma} = \gamma^{-(1+\alpha)}$, suppose that for all r, $C_r^{(\ell_{-,\gamma})} \subset C_r^{(\ell_{+,\gamma})}$, and say that there is a deviation from equilibrium at r if $\eta^{(\zeta,\ell_{-,\gamma})}(\underline{x}; r') = 0$ for some $r' \in I_r$, I_r the union of $C_r^{(\ell_{+,\gamma})}$ and all the cubes $C_{r'}^{(\ell_{+,\gamma})}$ connected to $C_r^{(\ell_{+,\gamma})}$. We summarize all that by introducing the phase indicator $\Theta^{(\zeta,\ell_{-,\gamma},\ell_{+,\gamma})}(\underline{x}; r)$ which is set equal to 0 if there is a deviation from equilibrium in the above sense; otherwise we set $\Theta^{(\zeta,\ell_{-,\gamma},\ell_{+,\gamma})}(\underline{x}; r) = 1$ if $\eta^{(\zeta,\ell_{-,\gamma})}(\underline{x}; r') = 1$ for all $r' \in I_r$ and $\Theta^{(\zeta,\ell_{-,\gamma},\ell_{+,\gamma})}(\underline{x}; r) = -1$ if $\eta^{(\zeta,\ell_{-,\gamma})}(\underline{x}; r') = -1$ for all $r' \in I_r$.

Contours. A contour Γ of a configuration $\underline{x}$ consists of a bounded set $\mathrm{sp}(\Gamma)$, called the spatial support of Γ, which is one of the maximal connected components of the region $\{r : \Theta^{(\zeta,\ell_{-,\gamma},\ell_{+,\gamma})}(\underline{x}; r) = 0\}$, and the restriction η_Γ of $\eta^{(\zeta,\ell_{-,\gamma})}(\underline{x}; r)$ to $\mathrm{sp}(\Gamma)$. Thus, a contour is the pair $\Gamma = (\mathrm{sp}(\Gamma), \eta_\Gamma)$. We shall hereafter simply write η and Θ without indicating the parameters which enter in their definition.

Geometry of contours. As in the Ising case, we shall restrict to boundary conditions such that Θ is identically 1 (or -1) outside of a compact. Thus $\mathrm{sp}(\Gamma)$ is a bounded, $\mathcal{D}^{(\ell_{+,\gamma})}$-measurable connected set. We call $\mathrm{ext}(\Gamma)$ and $\mathrm{int}_i(\Gamma)$ the maximal connected components of its complement, $\mathrm{ext}(\Gamma)$ being the unbounded one. We denote by D_{ext} the union of all $\mathcal{D}^{(\ell_{+,\gamma})}$ in $\mathrm{ext}(\Gamma)$ which are connected to $\mathrm{sp}(\Gamma)$ and by D_i the union of those in $\mathrm{int}_i(\Gamma)$ connected to $\mathrm{sp}(\Gamma)$. It is proved that all such sets are connected. Moreover, $\Theta \neq 0$ on any such sets and $\Theta(\underline{x}; r) = \eta(\underline{x}, r')$ if $C_{r'}^{(\ell_{+,\gamma})} \subset \mathrm{sp}(\Gamma)$ is connected to $C_r^{(\ell_{+,\gamma})}$. Thus, Θ is constant on D_{ext} and on each D_i and its value is determined by η_Γ. We may add a $\pm$ superscript to the sets D to indicate the sign of Θ.

Weight of contours. Γ is a plus (minus) contour if $\Theta = 1$ $(= -1)$ on D_{ext}. C denotes the union of $\mathrm{sp}(\Gamma)$ and all $\mathrm{int}_i(\Gamma)$ and $Z_{C,\underline{x}_{D_{\mathrm{ext}}}}(A)$ stands for the partition function in C with boundary conditions $\underline{x}_{D_{\mathrm{ext}}}$ and constraint A, A a set of configurations in C (we are dropping here the parameters β, λ, γ). Then if Γ is a

plus contour for any $\underline{x}_{D_{\rm ext}}$ such that $\eta(\underline{x}_{D_{\rm ext}}, r) = 1$ on $D_{\rm ext}$ we define

$$W(\Gamma, \underline{x}_{D_{\rm ext}}) := \frac{Z_{C,\underline{x}_{D_{\rm ext}}}[\eta(\underline{x}; r) = \eta_\Gamma(r), r \in \mathrm{sp}(\Gamma), \Theta(\underline{x}, r) = \pm 1, r \in D_i^\pm]}{Z_{C,\underline{x}_{D_{\rm ext}}}[\eta(\underline{x}; r) = 1, r \in \mathrm{sp}(\Gamma), \Theta(\underline{x}, r) = 1, r \in D_i^\pm]}. \tag{3.33}$$

Thus the numerator is the weight of seeing the contour Γ conditioned to the outside of $\mathrm{sp}(\Gamma)$ while the denominator is the weight with the same conditioning that the contour Γ is absent and replaced by the plus configurations. Analogous definition is given for the minus contours.

As in the Ising model of Lecture 2, everything works fine if we have good bounds on $W(\Gamma, \underline{x}_{D_{\rm ext}})$. To get a feeling of what we could get, let us approximate the partition functions in (3.33) by the free energy functional, which, in the present case, is a conditional free energy functional. Let A be a bounded region, $\bar\rho$ the boundary condition, supported outside A. We then define

$$\int_{\mathbb{R}^d} \Big\{ e_\lambda\big(J * [\rho_A + \bar\rho]\big) - e_\lambda(J * \bar\rho) + \frac{1}{\beta}\rho_A(r)\big[\log \rho_A(r) - 1\big] \Big\}\, dr.$$

In our case $\lambda = \lambda_\beta$, $A = \gamma C$, $\bar\rho(r) = A^{(\ell)}(\gamma r | \underline{x}_{D_{\rm ext}})$. Call $\mathcal{X}$ and $\mathcal{X}^+$ the set of all the configurations $\underline{x}$ in the region C which contribute to the numerator and, respectively, the denominator of (3.33). Let $\mathcal{R}$ and $\mathcal{R}^+$ be, respectively, the sets of $\rho_{\gamma C}(r)$, $r \in \gamma C$, such that

$$\rho_{\gamma C}(r) = A^{(\ell)}(\gamma r | \underline{x}), \underline{x} \in \mathcal{X}; \qquad \rho_{\gamma C}(r) = A^{(\ell)}(\gamma r | \underline{x}), \underline{x} \in \mathcal{X}^+. \tag{3.34}$$

Then the numerator in (3.33) is approximated by

$$\exp\Big\{ -\beta \gamma^{-d} \inf_{\mathcal{R}} F_{\gamma C}(\rho_{\gamma C} | \bar\rho) \Big\} \tag{3.35}$$

while the denominator is approximated by

$$\exp\Big\{ -\beta \gamma^{-d} \inf_{\mathcal{R}^+} F_{\gamma C}(\rho_{\gamma C} | \bar\rho) \Big\}. \tag{3.36}$$

The ratio of (3.35) and (3.36) can be proved to be bounded as the right-hand side of (3.37) thus if (3.35) and (3.36) were not an approximation we would get

$$W(\Gamma, \underline{x}_{D_{\rm ext}}) \le e^{-c\gamma^{-d(1-2\alpha-2a)} N_\Gamma}, \qquad N_\Gamma \ell_{+,\gamma}^d = |\mathrm{sp}(\Gamma)|. \tag{3.37}$$

With this bound, we could actually repeat the proofs of the Ising case as in Lecture 2. However the bound (3.35) and (3.36) become true only in the limit $\gamma \to 0$ which is not good enough as we want to keep γ fixed, maybe very small but fixed. To make (3.37) correct, we should increase the right-hand side by a factor $\exp\{b\gamma^{1/2}|C|\}$, $b > 0$, so that the bound becomes useless if the interior of C is much larger than $\mathrm{sp}(\Gamma)$.

We can avoid this difficulty by localizing the estimate to $\mathrm{sp}(\Gamma)$ so that the error is now $\exp\{b\gamma^{1/2}|\mathrm{sp}(\Gamma)|\}$ which does not spoil the bound on the right-hand side of (3.37). Referring to Chapter 11 in Presutti (2008) for the details, we just say here that using a Lebowitz–Penrose argument we get

$$W(\Gamma, \underline{x}_{D_{\mathrm{ext}}}) \le e^{-\beta(c_1\zeta^2 - c'\gamma^{1/2-2\alpha d})\ell^d_{-,\gamma} N_\Gamma} \frac{e^{\beta I^-(\mathrm{int}^-(\Gamma))}\hat{Z}^-_{\mathrm{int}^-(\Gamma),\chi^-}}{e^{\beta I^+(\mathrm{int}^-(\Gamma))}\hat{Z}^+_{\mathrm{int}^-(\Gamma),\chi^+}}, \tag{3.38}$$

where $\chi^\pm(r) \equiv \rho_{\beta,\pm}$, $\hat{Z}^\pm_{\mathrm{int}^-(\Gamma)}$ stands for the partition function in $\mathrm{int}^-(\Gamma)$ with the constraint that $\Theta = \pm 1$ identically on the cubes of $\mathcal{D}^{\ell_+,\gamma}$ connected to $\mathrm{sp}(\Gamma)$ and

$$\begin{aligned} I^\pm(\Lambda) = &\int_{\Lambda^c} \{e_{\lambda(\beta)}(\rho_{\beta,\pm}) - e_{\lambda(\beta)}(J_\gamma * \rho_{\beta,\pm}\mathbf{1}_{\Lambda^c})\} \\ &- \int_\Lambda e_{\lambda(\beta)}(J_\gamma * \rho_{\beta,\pm}\mathbf{1}_{\Lambda^c}). \end{aligned} \tag{3.39}$$

An analogous expression is derived for the Ising model with Kac potential and the analogous of the bound (3.38) would conclude the proof of the Peierls bounds because the ratio of the partition functions in (3.38) is equal to 1 by the spin flip symmetry and for the same reason $I^+(\Lambda) = I^-(\Lambda)$.

In the LMP model instead, there is no symmetry between the plus and minus states and indeed the ratio of the plus and minus partition functions in (3.38) will in general be of the order of the surface which would then compete with the first factor (the gain term) on the right-hand side of (3.38). This is exactly the framework of the Pirogov–Sinai theory, the crucial point is the fact that for a special value of the chemical potential the ratio of the partition functions becomes of the order of the surface times a small factor which for γ small is negligible with respect to the gain term.

The idea is grosso modo the following. We can control the dependence on the boundaries of the partition functions if we have good estimates on the decay of the correlation functions; if we have a good dependence on the boundaries the Peierls bounds hold true and then by the arguments used in Lecture 2 we derive the desired decay of correlations. The argument is clearly circular and Pirogov–Sinai could "rectify it" by using an induction on the volume argument. This require to adjust the value of the chemical potential at each step of the induction and in the end it leads to the special choice $\lambda = \lambda_{\beta,ga}$ of Theorem 3.2.

3.6 Additional bibliography

As already mentioned the approach used in LMP is more general and applies to other systems with Kac potentials, in particular it has been used to study a version of the Potts model in the continuum with Kac potentials, see De Masi et al. (2008, 2009). In the last reference (as well as in the original LMP paper), there is

a computation of the effective Hamiltonian which is proved to be a Hamiltonian with many body interactions.

The first results where the Kac program has been carried through keeping γ fixed refer to the Ising model with ferromagnetic Kac interactions, see Cassandro and Presutti (1996), Bovier and Zahradnik (1997) and also Chapter 9 of Presutti (2008).

The statement about the absence of phase transitions in $d = 1$ models is not entirely correct, at least in the generality in which it is stated. The simple spin system made of a chain of harmonic oscillators with harmonic nearest neighbor interactions gives a counter-example: in Benfatto, Presutti and Pulvirenti (1978) it is shown that there is an infinity of extremal Gibbs measures, but only one translational invariant (Gibbs measures being defined as weak limits of finite volume Gibbs measures).

Absence of phase transition is true for compact spin systems with interaction having second moment finite. Problems arise in the noncompact case: one may take extremely large boundary spins so that they drive the neighboring ones to high values and the effect may persist in the thermodynamic limit, as in the counter-example quoted above. This does not apply though to the LMP model because of the repulsive nature of the interaction at high densities. So that very high density boundary conditions do not drive toward high densities, but the opposite happens.

3.7 Bibliographical remarks

In the 60s, Kac proposed a scheme to derive the van der Waals theory from statistical mechanics. The original papers are by Kac, Uhlenbeck and Hemmer (1963a, 1963b, 1964), by Kac and Thompson (1966) and by Kac and Helfand (1963). The extension to more general systems with the introduction of coarse-graining techniques and renormalization group ideas is due to Lebowitz and Penrose (1966), Penrose and Lebowitz (1971). An analysis of the nonlocal free energy functionals which arise in the theory can be found in the papers by Gates and Penrose (1969, 1970a, 1970b).

The simplest model for the Lebowitz–Penrose coarse graining is maybe the Ising model with Kac potentials, the detailed analysis of this system can be found in Section 4.2 of my book (Presutti (2008)). The first proof of phase transitions in the Ising model with a fixed positive value of the Kac scaling parameter γ is due to Cassandro and Presutti (1996), and Bovier and Zahradnik (1997). The LMP model has been introduced by Lebowitz, Mazel and Presutti who proved the existence of a phase transition for positive values of γ (Lebowitz, Mazel and Presutti (1999)). The validity of the Gibbs phase rule for the LMP model has been established (Bovier et al. (2004)). A detailed analysis of the structure of the pure phases in the LMP model can be found in my book (Presutti (2008), Chapters 10–12).

4 Lecture 4

In this last lecture, I will discuss nonequilibrium. The theory is not as developed as in equilibrium and many fundamental problems are still open. For the sake of definiteness I will focus on a single subject: the analysis of states which are stationary but not in thermal equilibrium. The example to have in mind is a metal bar whose extremes are kept at two different temperatures $T_- > T_+$ by adding and removing heat. After a while, the system equilibrates and a steady heat flow goes from the left where the temperature is T_- to the right where the temperature is T_+. The system is then in a stationary state but not in thermal equilibrium as the temperature is not constant throughout the body.

As any probability (in a finite state space) can be written as a Gibbs measure with a suitable Hamiltonian also the stationary nonequilibrium state describing the above situation can be written as a Gibbs state with some effective Hamiltonian and the dream is to discover a rule for finding such a Hamiltonian. Most likely, this is just a dream but, as we shall see, something can be done in this direction.

4.1 The symmetric simple exclusion process

The easiest system we can possibly conceive is certainly the free lattice gas, namely a system of particles in $\mathbb{Z}^d$ where the only interaction is hard-core which enforces the exclusion rule that no two particles can occupy the same site. Thus particle configurations are elements $\eta \in \{0,1\}^{\mathbb{Z}^d}$. Gibbs equilibrium states are then product measures which are therefore specified by the probability that $\eta(x) = 1$:

$$\mu_\lambda\big[\eta(x) = 1\big] = \frac{e^\lambda}{1+e^\lambda}, \qquad x \in \mathbb{Z}^d \tag{4.1}$$

with $\lambda \in \mathbb{R}$ the chemical potential (temperature does not play any role here as there is no interaction among particles except exclusion, we have thus set $\beta = 1$).

We introduce a dynamics by requiring "locality", namely that particles can only move to nearest neighbor sites and that all μ_λ should be invariant. This still leaves open a lot of choices, we take maybe the simplest one by requiring that dynamics is given by the Markov process on $\{0,1\}^{\mathbb{Z}^d}$ defined by the generator L^{ssep} below which acts on the cylindrical functions f as

$$\begin{aligned} L^{\text{ssep}} f(\eta) &= \sum_x \sum_{y\,:\,|y-x|=1} \frac{1}{2}\big(\eta(x)\big[1-\eta(y)\big] + \eta(y)\big[1-\eta(x)\big]\big) \\ &\qquad\qquad \times \big(f\big(\eta^{(x,y)}\big) - f(\eta)\big) \\ &= \sum_{x,y\,:\,|y-x|=1} \frac{1}{2}\big(f\big(\eta^{(x,y)}\big) - f(\eta)\big), \end{aligned} \tag{4.2}$$

where $\eta^{(x,y)}$ is obtained from η by exchanging its occupation numbers at x and y.

L is the symmetric simple exclusion process (SSEP), where each particle independently tries to jump to a nearest neighbor site after an exponential time of mean $1/2$: the jump then takes place if the chosen site is empty. This is the same as the generator of the stirring process where independently and at rate $1/2$ the occupation numbers of nearest neighbor sites exchange their values, as described in the last expression in (4.2). It is obvious that the measures μ_λ are invariant. The proof that they are the only extremal, invariant measures follows from the De Finetti theorem and the fact that they are exchangeable, see, for instance (Liggett (1985)).

Let us now move to the stationary nonequilibrium states for our lattice gas. We represent the metal bar as an interval $[-N, N]$ on $\mathbb{Z}$, N a positive integer. We do not have temperatures here, the role of temperature is played by the chemical potential and correspondingly we have a mass rather than a heat flow. The evolution is described by the SSEP but we have to add something else which in the physical example of the metal bar corresponds to fixing two different temperatures at the extremes of the bar. This can be done by adding the following two new updating mechanisms, one at $-N$ and the other at N. At rate $1/2$ we choose $-N$ and independently of its previous value we set $\eta(-N) = 1$ with probability $\mu_{\lambda_-}[\eta(-N) = 1] := \rho_-$ and $\eta(-N) = 0$ with probability $\mu_{\lambda_-}[\eta(-N) = 0]$; analogous mechanism acts at N with λ_+ instead of λ_- and ρ_+ instead of ρ_-. We also suppose $\lambda_- > \lambda_+$ and thus $\rho_- > \rho_+$. The generator of the process is then $L = L_- + L_+ + L^{\text{ssep}}$ where

$$\begin{aligned} L_+ f(\eta) = {} & \tfrac{1}{2}\big(\eta(N)(1-\rho_+)\big[f(\eta - \mathbf{1}_N) - f(\eta)\big] \\ & + \big(1-\eta(N)\big)\rho_+\big[f(\eta + \mathbf{1}_N) - f(\eta)\big]\big) \end{aligned} \tag{4.3}$$

and $\mathbf{1}_N$ is the configuration with all 0 except at N. L_- is defined analogously.

The one-body correlation function is

$$\rho(x,t) = E_\mu\big[\eta(x,t)\big], \tag{4.4}$$

where E_μ denotes the expectation of the process with generator L starting from some initial measure μ. The time derivative of $\rho(x,t)$ is obtained by applying the generator to the function $f(\eta) = \eta(x)$:

$$\begin{aligned} \frac{d\rho(x,t)}{dt} &= \frac{1}{2}\big(\rho(x+1,t) + \rho(x-1,t) - 2\rho(x,t)\big), \qquad |x| < N, \\ \frac{d\rho(-N,t)}{dt} &= \frac{1}{2}\big(\rho(-N+1,t) - \rho(-N,t)\big) + \frac{1}{2}\big(\rho_- - \rho(-N,t)\big), \\ \frac{d\rho(N,t)}{dt} &= \frac{1}{2}\big(\rho(N-1,t) - \rho(N,t)\big) + \frac{1}{2}\big(\rho_+ - \rho(N,t)\big). \end{aligned}$$

It is certainly atypical that the one-body correlation functions satisfy a closed equation, in general their evolution depends on the two-body correlations which in turns

evolve depending on the three-body correlations and so on so that there is a whole hierarchy of equations that one has to solve to determine the evolution of the one-body correlations. In our case instead the equations are closed, this is related to the self-duality of the SSEP, an issue which will be discussed later on.

The above evolution equations can be solved in a rather explicit fashion. Indeed let $x(t)$ be a simple symmetric random walk on $\mathbb{Z}$ (i.e., after an exponential time of mean 1 it jumps with equal probability on its nearest neighbor sites). We stop $x(t)$ once outside the interval $\{|x| \le N\}$, as we are interested to the case when $|x(0)| \le N$ this means that $x(t)$ is the simple random walk till the random time τ when $|x(\tau)| = N+1$, after which it does not move anymore. Let $\rho_0(x) = \rho(x,0)$ for $|x| \le N$ and $\rho_0(x) = \rho_+$ for $x > N$ and $= \rho_-$ for $x < -N$. Then

$$\rho(x,t) = \mathcal{E}_x\big[\rho_0\big(x(t)\big)\big], \qquad |x| \le N \tag{4.5}$$

$\mathcal{E}_x$ denoting the expectation of the stopped random walk starting from x. It is also easy to compute the limit $\rho^{\mathrm{st};N}(x)$ of $\rho(x,t)$ as $t \to \infty$:

$$\rho^{\mathrm{st};N}(x) = \rho_- + \frac{x-(-N-1)}{N+1-(-N-1)}(\rho_+ - \rho_-), \qquad |x| \le N \tag{4.6}$$

which extends to a linear profile connecting ρ_- to ρ_+ as x varies from $-N-1$ to $N+1$.

4.2 The Fick's law

The local current $j(x,t)$ is defined as the expected number of particles crossing the bond $(x, x+1)$ at time t from left to right minus that of particles crossing the bond in the reverse direction. In the SSEP, this is

$$j(x,t) = -\tfrac{1}{2}\big(\rho(x+1,t) - \rho(x,t)\big). \tag{4.7}$$

Thus, the local current is proportional to minus the density gradient, the proportional factor being equal to $1/2$. Fourier's law states that the heat current is equal to minus the temperature gradient times the diffusivity coefficient; Fick's law, which is the analogue of the Fourier's law for mass transport, states that the mass current is equal to minus the density gradient times the mass-diffusivity coefficient. In the SSEP therefore the diffusivity coefficient is equal to $1/2$.

In the stationary state, the current is

$$j^{\mathrm{st};N}(x) = -\frac{1}{2}\frac{1}{2(N+1)}(\rho_+ - \rho_-) \tag{4.8}$$

which vanishes as $N \to \infty$ like the inverse of the size of the system. Indeed if we introduce spatial macroscopic units, $x \to r = \varepsilon x$, $\varepsilon = 1/N$, the stationary profile $\rho^{\mathrm{st};N}(x)$ expressed in macroscopic unit converges as $N \to \infty$ to

$$\lim_{N\to\infty} \lim_{Nx\to r} \rho^{\mathrm{st};N}(x) = \rho^{\mathrm{st}}(r) = \rho_- + \frac{r-(-1)}{2}(\rho_+ - \rho_-), \qquad |r| \le 1. \tag{4.9}$$

The time macroscopic units are $t \to \varepsilon^2 t$ and from (4.5) (adding a superscript N to $\rho(r,t)$)

$$\lim_{N\to\infty} \lim_{Nx\to r} \rho^{(N)}\big(x, N^2 t\big) = \rho(r,t), \qquad |r| \le 1, \tag{4.10}$$

where the limit $\rho(r,t)$ is the solution of the Dirichlet equation in $[-1,1]$

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\frac{\partial^2 \rho}{\partial t^2} \tag{4.11}$$

with boundary conditions:

$$\rho(r,0) = \rho_0(r), \qquad \rho(\pm 1, t) = \rho_\pm \tag{4.12}$$

having supposed that the initial datum $\rho_0^{(N)}(x)$ converges to $\rho_0(r)$ in the sense of (4.10).

Equations (4.11)–(4.12) are compatible with (4.9) in the sense that $\rho^{\mathrm{st}}(r)$ is the stationary solution of (4.11)–(4.12). This means that the limits $t \to \infty$ and $N \to \infty$ can be exchanged: this is so in the SSEP but it is not at all a general fact.

Let us now briefly recall the Fick's law. This is about a system where mass is transported in a conservative way, so that the evolution equation is the conservation law

$$\frac{\partial \rho}{\partial t} = -\frac{\partial I}{\partial r} \tag{4.13}$$

with boundary conditions (4.12). To close the equation, we need a so-called constitutive law which expresses the flux I as a function of ρ. The Fick's law says that this is given by

$$I = -D(\rho)\frac{\partial \rho}{\partial r}$$

$D(\rho) > 0$ the diffusion coefficient. Comparing with (4.11), we see that for the SSEP $D(\rho) = 1/2$ so that by (4.9) the stationary current I^{st} is

$$I^{\mathrm{st}} = -\frac{\rho_+ - \rho_-}{4}. \tag{4.14}$$

If we express (4.8) in macro unit, we get

$$j^{\mathrm{st};N} \to \left(\frac{\varepsilon^2}{\varepsilon}\right) j^{\mathrm{st};N}, \tag{4.15}$$

where ε^2 comes from the time rescaling (current is the mass flow per unit time) and ε from the space, the bond $(x, x+1)$ has macro-length ε. It is readily seen that the right-hand side of (4.15) converges to the right-hand side of (4.14).

In conclusion, there is perfect agreement with the predictions of the macrosocpic theory and the behavior of the SSEP.

4.3 The invariant measure

Let us now move one of the main questions posed in the beginning of this lecture about the structure of the nonequilibrium stationary states. First of all, observe that for any N there exists a unique invariant measure $\mu^{(N)}$, as it follows from general theorems on Markov processes with finite state space under the condition that any state can be reached from any other in a finite time. This is a strong Döblin condition which also ensures exponentially fast convergence from any initial state.

We are interested in the limit $N \to \infty$ which in equilibrium was used to derive the thermodynamics of the system and was therefore called the thermodynamic limit. In our case, is more proper to call it the macroscopic limit. We thus want to understand the behavior of $\mu^{(N)}$ in the macroscopic limit $N \to \infty$. First issue is of course the choice of the topology, we start from the weak topology used so far in equilibrium, a choice in many ways inadequate as we shall see. We say that $\mu^{(N)} \to \mu$ weakly as $N \to \infty$ if for any cylindrical function f:

$$\lim_{N\to\infty} \mu^{(N)}[f] = \mu[f]. \tag{4.16}$$

Recall that a function f is cylindrical if it depends only on finitely many $\eta(x)$, so that the cylindrical functions can be obtained by taking linear combinations of functions of the form $\prod_{i=1}^{N} \eta(x_i)$. The expectations

$$\mu\left[\prod_{i=1}^{N} \eta(x_i)\right] =: \rho(x_1, \dots, x_n)$$

are called the n-body correlation functions of the measure μ and μ is uniquely determined by the totality of its correlation functions. We shall use below that by weak-compactness $\mu^{(N)}$ converges weakly by subsequences.

Recall that $\mu^{(N)}$ is invariant if for all functions f on $\{-1, 1\}^{[-N,N]}$ we have

$$\mu^{(N)}\big[L^{(N)} f\big] = 0, \tag{4.17}$$

where $L^{(N)} = L_-^{(N)} + L_+^{(N)} + L^{\mathrm{ssep},N}$ (we have added the super-script N to the previous notation). Recalling the definition of $L_\pm^{(N)}$ we have $L_\pm^{(N)} f = 0$ if f depends only on $\eta(x)$, $|x| < N$, thus for such functions

$$\mu^{(N)}\big[L^{\mathrm{ssep}} f\big] = 0, \qquad L^{\mathrm{ssep},N} f = L^{\mathrm{ssep}} f \tag{4.18}$$

and therefore if μ is the weak limit along a convergent subsequence N_k for any cylindrical function f

$$\lim_{N_k\to\infty} \mu^{(N_k)}\big[L^{\mathrm{ssep}} f\big] = \mu\big[L^{\mathrm{ssep}} f\big] = 0. \tag{4.19}$$

Thus, μ is SSEP-invariant and it is therefore a superposition of Bernoulli measures ν_ρ, $\rho \in [0, 1]$, where $\nu_\rho[\eta(x) = 1] = \rho$.

The result gives a negative answer to the question about the structure of stationary nonequilibrium states: in the macroscopic limit the stationary nonequilibrium

states become thermal equilibrium states. Thus, the macroscopic limit is not useful to understand the stationary nonequilibrium states, but before discussing the issue we state and prove the following theorem which improves (4.19).

Theorem 4.1. *$\mu^{(N)}$ converges weakly to the Bernoulli measure ν_ρ, $\rho = \frac{\rho_- - \rho_+}{2}$.*

Proof. We already know from (4.9) that for any x

$$\lim_{N\to\infty} \mu^{(N)}\big[\eta(x) = 1\big] = \frac{\rho_- - \rho_+}{2} \tag{4.20}$$

so that if μ is the weak limit of $\mu^{(N)}$ along a convergent subsequence then $\mu[\eta(x) = 1] = \frac{\rho_- - \rho_+}{2}$. Since μ is a superposition of Bernoulli measures ν_ρ the theorem will follow once we prove that for any $x \neq y$

$$\mu\big[\eta(x) = 1, \eta(y) = 1\big] = \left(\frac{\rho_- - \rho_+}{2}\right)^2. \tag{4.21}$$

Let us go back to the time dependent case, denote by $\mathcal{E}^{(N)}_{\eta_0}$ the expectation for the process starting from η_0, define $\rho^{(N)}(z,t) = \mathcal{E}^{(N)}_{\eta_0}[\eta(z,t)]$ and

$$v^{(N)}(x,y;t) := \mathcal{E}^{(N)}_{\eta_0}\big[\big(\eta(x,t) - \rho^{(N)}(x,t)\big)\big(\eta(y,t) - \rho^{(N)}(y,t)\big)\big]. \tag{4.22}$$

We shall prove (4.21) by showing that

$$\lim_{N\to\infty}\lim_{t\to\infty} v^{(N)}(x,y;t) = 0. \tag{4.23}$$

Recall in fact that for each N the law of the process at time t converges as $t \to \infty$ to $\mu^{(N)}$ independently of the initial state so that

$$\lim_{t\to\infty} v^{(N)}(x,y;t) = \mu^{(N)}\big[\eta(x) = 1, \eta(y) = 1\big] - \mu^{(N)}\big[\eta(x) = 1\big]\mu^{(N)}\eta(y) = 1].$$

Denote by L^{st} the stirring generator which exchanges independently at rate $1/2$ the occupation numbers at $(x, x+1)$ and by $L^{\mathrm{st},N}$ the one where $(x, x+1)$ are both in $[-N, N]$. It can be checked that for x and $y \neq x$ in $[-N, N]$

$$\begin{aligned}\frac{d}{dt}v^{(N)}(x,y;t) &= L^{\mathrm{st},N}v^{(N)}(x,y;t)\\ &\quad - \frac{1}{2}\big(\rho^{(N)}(y,t) - \rho^{(N)}(x,t)\big)^2\mathbf{1}_{|x-y|=1}\\ &\quad - \frac{1}{2}v^{(N)}(x,y;t)(\mathbf{1}_{|x|=N} + \mathbf{1}_{|y|=N}),\end{aligned} \tag{4.24}$$

where $v^{(N)}(x,y;t)$ is thought of as a function of the positions x and y of two stirring particles and $L^{\mathrm{st},N}v^{(N)}(x,y;t)$ is defined accordingly. Equation (4.24) can

be rewritten as an equation for $v^{(N)}(x,y;t)$, $x \neq y$, $|x| \leq N+1$, $|y| \leq N+1$,

$$\begin{aligned}\frac{d}{dt}v^{(N)}(x,y;t) &= L^{\mathrm{st}}v^{(N)}(x,y;t)\\ &\quad -\frac{1}{2}\big(\rho^{(N)}(y,t)-\rho^{(N)}(x,t)\big)^2\mathbf{1}_{|x-y|=1}, \qquad (4.25)\\ v^{(N)}(x,y;t) &= 0 \qquad \text{in } \big\{|x|=N+1\big\}\cup\big\{|y|=N+1\big\} \text{ and when } t=0\end{aligned}$$

which is solved by

$$\begin{aligned}v^{(N)}(x,y;t) = -\frac{1}{2}\int_0^t E^{\mathrm{st}}_{x,y}\big[\big(\rho^{(N)}\big(y(s),t-s\big)-\rho^{(N)}\big(x(s),t-s\big)\big)^2\\ \times \mathbf{1}_{|x(s)-y(s)|=1}\mathbf{1}_{\tau_x>s,\tau_y>s}\big],\end{aligned} \qquad (4.26)$$

where $E^{\mathrm{st}}_{x,y}$ denotes expectation with respect to the process of two stirring particles which start from x and y in $[-N,N]$ and τ_x, τ_y are the first time when particle x, respectively, particle y, reaches $\pm(N+1)$. There is a constant $b>0$ so that

$$E^{\mathrm{st}}_{x,y}[\mathbf{1}_{\tau_x>s}] \leq e^{-b\varepsilon^2 s}, \qquad \varepsilon = \frac{1}{N}$$

as the left-hand side is the probability that a simple random walk which starts from x remains inside $[-N,N]$ till time s. With this bound we can take the limit as $t\to\infty$ in (4.26) and get, recalling (4.6),

$$\begin{aligned}&\mu^{(N)}\big(\eta(x)=1,\eta(y)=1\big)-\rho^{\mathrm{st};N}(x)\rho^{\mathrm{st};N}(y)\\ &\quad=-\frac{1}{2}\Big(\frac{1}{2N+1}\Big)^2\int_0^\infty E^{\mathrm{st}}_{x,y}[\mathbf{1}_{|x(s)-y(s)|=1}\mathbf{1}_{\tau_x>s,\tau_y>s}].\end{aligned} \qquad (4.27)$$

By Cauchy–Schwarz

$$\begin{aligned}&\big|\mu^{(N)}\big(\eta(x)=1,\eta(y)=1\big)-\rho^{\mathrm{st};N}(x)\rho^{\mathrm{st};N}(y)\big|\\ &\quad\leq\frac{\varepsilon^2}{2}\int_0^\infty e^{-(b/2)\varepsilon^2 s}\Big(\frac{c}{\sqrt{s}}\Big)^{1/2}\leq c'\varepsilon^{1/2}.\end{aligned} \qquad (4.28)$$

The bound $E^{\mathrm{st}}_{x,y}[\mathbf{1}_{|x(s)-y(s)|=1}]\leq\frac{c}{\sqrt{s}}$ is obtained from classical estimates on simple random walks after coupling the two stirring particles, $x(s)$, $y(s)$, with two simple independent random walks, $x^0(s)$, $y^0(s)$, in such a way that if $x^0(0)=x(0)$, $y^0(0)=y(0)$ then at all times

$$\big|x^0(s)-y^0(s)\big|\leq\big|x(s)-y(s)\big|\leq\big|x^0(s)-y^0(s)\big|+1.$$

The coupling is defined by matching the displacements of the rightmost and leftmost particles. This cannot be achieved when $|x(s)-y(s)|=1$ and that is the reason for the inequality.

By (4.28) $|\mu^{(N)}(\eta(x)=1,\eta(y)=1)-\rho^{\mathrm{st};N}(x)\rho^{\mathrm{st};N}(y)|\to 0$ as $N\to\infty$ at least as fast as $N^{-1/2}$ and the theorem is proved. □

The use of the Cauchy–Schwarz inequality at the end of the proof could be avoided. Define for $r \neq r'$, both in $(-1, 1)$,

$$C^{(N)}(r, r') = N\big(\mu^{(N)}\big(\eta(x) = 1, \eta(y) = 1\big) - \rho^{\mathrm{st};N}(x)\rho^{\mathrm{st};N}(y)\big), \tag{4.29}$$

where x and y are the integer parts of rN and $r'N$. It can be proved that:

Theorem 4.2. *Denote by $E_{r,r'}$ expectation with respect to the law of two independent Brownian motions, $B(s)$ and $B'(s)$ starting from r and $r' \neq r$ both in $(-1, 1)$. Let τ and τ' be the hitting times at ± 1 of $B(s)$ and $B'(s)$, then*

$$\begin{aligned} &\lim_{N\to\infty} C^{(N)}(r, r') = C(r, r'), \\ &C(r, r') = -\frac{1}{2}\left(\frac{\rho_- - \rho_+}{2}\right)^2 \int_0^\infty E_{r,r'}\big[\delta_{B(s)}(B'(s))\mathbf{1}_{\tau>s,\tau'>s}\big] \end{aligned} \tag{4.30}$$

$\delta_b(x)$ the Dirac delta at b.

The theorem is proved by showing that (4.27) converges in the limit to (4.30), this is based on a comparison between stirring and independent random walks and convergence of the latter to Brownian motions. Equation (4.30) has a physical meaning and will play an important role in the analysis of the stationary nonequilibrium states as we shall see later on.

We conclude the section with an obvious extension of Theorem 4.1. Denote by $\mu^{(N;r)}$, $r \in (-1, 1)$, the measure on $\{0, 1\}^{[-N+X,N+X]}$, X the integer part of Nr, obtained from $\mu^{(N)}$ by translation.

Theorem 4.3. *For any $r \in (-1, 1)$ $\mu^{(N;r)}$ converges weakly to the Bernoulli measure $\nu_{\rho^{\mathrm{st}}(r)}$ where $\rho^{\mathrm{st}}(r)$ is defined in (4.9):*

$$\rho^{\mathrm{st}}(r) = \rho_- + \frac{r - (-1)}{2}(\rho_+ - \rho_-).$$

Theorem 4.3 is in agreement with the picture we have of continuum systems: at each macroscopic point r the system is described by an equilibrium state with order parameter the local density at r. We have thus proved that the stationary nonequilibrium state is in the macroscopic limit described by the density profile $\rho^{\mathrm{st}}(r)$. However, to have a complete characterization of the continuum theory we would need to derive a free energy functional associated to all density profiles $\rho(r)$ which should then reach its minimum at $\rho^{\mathrm{st}}(r)$. We shall discuss later this issue.

4.4 Density reservoirs

We have seen that the local structure of the invariant measure "degenerates" in the macroscopic limit to thermal equilibrium so that in the infinite volume we do not

see anymore the nonequilibrium effects. In particular, the current is proportional to $\varepsilon = 1/N$ and thus vanishes in the macroscopic limit $N \to \infty$.

On the contrary, we would like to find a system which has a stationary state which on the right at infinity has density ρ_+, on the left at infinity has density $\rho_- > \rho_+$ and throughout the system there is a nonzero current. This may seem impossible as the existence of a positive current means that we are transporting mass at constant rate to the right which seems to imply that the density on the left should decrease while that on the right increase.

The question is related to the way we fix in real systems the density (or the temperature) at the extremes of the system. This is done by putting the system in contact with density (or thermal) reservoirs. A reservoir at density ρ (let us stick to the case of mass transport for definiteness) is a system in thermal equilibrium at density ρ which exchanges mass (particles) with the system to which it is put in contact. The main property of the reservoir is its very high capacity so that despite the fact that it may loose (or gain) mass yet its density remains essentially unchanged. Thus the system at the contact with the reservoir has essentially the same density as the reservoir.

These features are behind the definitions of the boundary operators $L_\pm$ introduced in (4.4). In particular the action of L_+ can be thought of as a stirring exchange between the sites $(N, N+1)$ if $\eta(N+1)$ is averaged with mean ρ_+. This mimics the fact that particles enter or exit from the system to the reservoir. The fact that successive events have same rates independently of the past corresponds to the assumption that in the reservoir there is an instantaneous thermalization.

The natural question is whether there exist particle models for reservoirs, as we shall see a positive answer would enable to construct examples of infinite systems with the desired properties of having a nonzero current and asymptotically distinct densities. The infinite system being composed by our original system plus the two (infinite) reservoirs to its right and left.

Since in a reservoir loss or gain of mass should not change its total density a reservoir must be an infinite system. With this in mind, we go back to our SSEP model in $[-N, N]$ and add two infinite systems one in $[-\infty, -N-1]$ and the other in $[N+1, \infty]$ and suppose that initially the particles are distributed independently in $[-\infty, -N-1]$ with density ρ_- and in $[N+1, \infty]$ with density ρ_+. The dynamics is then the SSEP process on the whole $\mathbb{Z}$. This however has not the desired properties because the stationary measure for the SSEP are superposition of Bernoulli, thus asymptotically in time the current vanishes and the density equilibrates on the whole $\mathbb{Z}$: the systems in $[-\infty, -N-1]$ and in $[N+1, \infty]$ do not behave as reservoirs.

A simple way out is to replace the sets $[-\infty, -N-1]$ and $[N+1, \infty]$ by higher-dimensional half spaces, so that the stirring process acts on the space

$$\begin{aligned} S_{d,N} = \{x \in \mathbb{Z}^d : x_1 < -N\} \cup \{x \in \mathbb{Z} : x \in [-N, N]\} \\ \cup \{x \in \mathbb{Z}^d : x_1 > N\} \end{aligned} \tag{4.31}$$

(here x_1 denotes the first component of x). Calling $S^{\text{left}}_{d,N}$, $S^{\text{center}}_{d,N}$ and $S^{\text{right}}_{d,N}$ the three sets on the right-hand side of (4.31), we suppose as before that the initial measure μ_0 is a product measure with constant averages ρ_- on $S^{\text{left}}_{d,N}$ and ρ_+ on $S^{\text{right}}_{d,N}$. Dynamics is the SSEP on $S_{d,N}$ with jumps on nearest neighbor sites; call μ_t its law at time t (starting from μ_0).

Theorem 4.4. *With the above notation, if $d \geq 3$, then μ_t converges weakly as $t \to \infty$ to a measure μ such that*

$$\lim_{x_1 \to \pm\infty} \mu\big[\eta(x) = 1\big] = \rho_{\pm}. \tag{4.32}$$

Instead if $d \leq 2$ for any x

$$\mu\big[\eta(x) = 1\big] = \frac{\rho_- - \rho_+}{2}. \tag{4.33}$$

Proof. Calling again $\rho(x,t) = \mu_t[\eta(x) = 1]$,

$$\frac{d}{dt}\rho(x,t) = \frac{1}{2} \sum_{y\,:\,|y-x|=1} \big(\rho(y,t) - \rho(x,t)\big) \tag{4.34}$$

and therefore

$$\rho(x,t) = E_x\big[\rho\big(x(t),0\big)\big], \tag{4.35}$$

where E_x is the expectation of a simple random walk on $S_{d,N}$, $x(t)$ its position at time t. In $\mathbb{Z}^d$, $d \geq 3$, the random walk is transient therefore

$$\lim_{T\to\infty} P_x\big[x(t) \in S^{\text{center}}_{d,N} \text{ for some } t \geq T\big] = 0. \tag{4.36}$$

By (4.36) are well defined the probabilities $P^{\pm}_x$ that as $t \to \infty$ the walk is in $S^{\text{right}}_{d,N}$ and, respectively, $S^{\text{left}}_{d,N}$, moreover $P^+_x + P^-_x = 1$. Then

$$\lim_{t\to\infty} \rho(x,t) = \rho_- P^-_x + \rho_+ P^+_x. \tag{4.37}$$

In an analogous way, one can prove the analogue of (4.36) in the case of n stirring particles starting from any $x_1, \ldots, x_n \in S_{d,N}$ and since

$$\mu_t\big[\eta(x_1) = \cdots = \eta(x_n) = 1\big] = E_{x_1,\ldots,x_n}\big[\rho\big(x_1(t),0\big)\cdots\rho\big(x_n(t),0\big)\big] \tag{4.38}$$

we deduce that the left-hand side converges as $t \to \infty$, details are omitted as well as the analysis of the case $d \leq 2$. □

Thus, the stationary measure μ on $\{0,1\}^{S_{d,N}}$, $d \geq 3$, is a good candidate for studying the characteristic features of stationary nonequilibrium states: it has a nonzero current and the asymptotic densities are $\rho_\pm$ as desired. The analysis seems doable and interesting, but it has not yet been carried out. Also the study of the

behavior of the system when N is large is still missing, other interesting open question is a comparison of the stationary measure restricted to $\{0,1\}^{S^{\text{center}}_{d,N}}$ and the stationary measure $\mu^{(N)}$ with the birth–death processes at $\pm N$ in the regime of large N.

Of course, all the above refers to a very simple model, the SSEP process, and the conclusions one may learn from this analysis may be influenced by the peculiarity of the model. Thus, the main question is whether the above can be extended to more general systems. The most ambitious plan is to study Hamiltonian particle systems looking for the stationary solutions of the BBGKY hierarchy in a spatial geometry like the one described above. This is a completely open question even though there are some results in the case of a hard core gas at very low density, but to my knowledge nothing written has yet appeared.

A more doable problem which is still open as far as I know is the same one but in the context of the Boltzmann equation. Well studied instead is the derivation of the Fourier law in the Boltzmann equation which is carried out by using the Hilbert expansion method in bounded domains with "thermal boundaries" which simulate the reservoirs.

4.5 The $1/N$ corrections

Let us go back to the model in $[-N, N]$ with the boundary operators $L_\pm$. We have seen that in the macroscopic limit the invariant measure $\mu^{(N)}$ converges weakly to the Bernoulli equilibrium measure so that the information about the nonequilibrium effects have been lost. To recover them, at least partially, we may then look at the small deviations as $N \to \infty$. To first order the occupation numbers $\eta(x)$ are independent and have mean $\rho^{\mathrm{st};N}(x)$, see (4.6). Thus to first order

$$\mu^{(N)}(\eta) = \prod_x \frac{e^{\lambda_N(x)\eta(x)}}{1+e^{\lambda_N(x)}} =: \nu^{(N)}(\eta), \tag{4.39}$$

where $\lambda_N(x)$ is such that

$$\rho^{\mathrm{st};N}(x) = \frac{e^{\lambda_N(x)}}{1+e^{\lambda_N(x)}}.$$

By (4.29) and Theorem 4.2 to order $1/N$ the $\eta(x)$ are no longer independent. To catch the effect we replace $\nu^{(N)}$ by the measure

$$\nu^{(N)}(\eta)\Big(1+\frac{1}{N}\sum_{x\neq y}\frac{1}{\chi_N(x)\chi_N(y)}C^{(N)}\Big(\frac{x}{N},\frac{y}{N}\Big) \times \big[\eta(x)-\rho^{\mathrm{st};N}(x)\big]\big[\eta(y)-\rho^{\mathrm{st};N}(y)\big]\Big), \tag{4.40}$$

where $\chi_N(z) := \rho^{\mathrm{st};N}(z) - \rho^{\mathrm{st};N}(z)^2$. Calling $\nu_1^{(N)}$ the measure in (4.40) and using (4.29) we then get

$$\nu_1^{(N)}\big[\eta(x)\eta(y)\big] = \mu^{(N)}\big[\eta(x)\eta(y)\big].$$

It is certainly tempting to write

$$\nu_1^{(N)} \approx Z^{-1} e^{-H}. \tag{4.41}$$

with

$$\begin{aligned} H(\eta) = -\frac{1}{N} \sum_{x \neq y} \frac{1}{\chi_N(x)\chi_N(y)} C^{(N)}\left(\frac{x}{N}, \frac{y}{N}\right)\left[\eta(x) - \rho^{\mathrm{st};N}(x)\right] \\ \times \left[\eta(y) - \rho^{\mathrm{st};N}(y)\right] \end{aligned} \tag{4.42}$$

which suggests that if $\mu^{(N)}$ is a Gibbs measure its effective Hamiltonian is most likely given by a Kac potential with range comparable to the size of the system. Nonequilibrium seems to bring in long range correlations and hence long range effective interactions if the measure is written as Gibbsian.

The correction term $C^{(N)}(r, r')$ is related to the "small" fluctuations of the density fields as discussed in the next section.

4.6 Density fields and nonequilibrium thermodynamics

The density fields are random variables of the form

$$X_t^{(N)}(\phi) := \varepsilon \sum_x \phi(\varepsilon x)\eta\big(x, \varepsilon^{-2}t\big), \qquad \varepsilon = \frac{1}{N}, \tag{4.43}$$

where ϕ is a test function defined on $[-1, 1]$. Let $X^{(N)}(\phi)$ be the density field of (4.43) with $t = 0$. By Theorem 4.3, the distribution of $X^{(N)}(\phi)$ under $\mu^{(N)}$ which, converges as $N \to \infty$ to a Dirac delta on $\int \phi(r)\rho^{\mathrm{st}}(r)\,dr$ with $\rho^{\mathrm{st}}(r)$ defined in (4.9).

This is another way to say that the stationary particle density profile becomes in the macroscopic limit a deterministic linear profile. What can we say about the other, nonstationary density profiles? What stated above tells us that their probability vanishes in the macroscopic limit so that to see something we must renormalize. The density fluctuation fields (in the stationary case) are

$$Y^{(N)}(\phi) := \sqrt{\varepsilon} \sum_x \phi(\varepsilon x)\big[\eta(x) - \rho^{\mathrm{st},N}(x)\big]. \tag{4.44}$$

Using Theorem 4.2, it has been shown that the distribution of the density fluctuation fields becomes in the macroscopic limit Gaussian with variance

$$E\big[Y(\phi)Y(\psi)\big] = \int \chi(r)\psi(r)\phi(r) + \int\int \psi(r)\phi\big(r'\big)C\big(r, r'\big) \tag{4.45}$$

with $C(r, r')$ as in (4.30) and $\chi := \rho^{\mathrm{st}} - (\rho^{\mathrm{st}})^2$. This confirms the appearance of long range correlations, again in the $1/N$ regime as in the previous section.

To explore the density fields away from the linear stationary profile, we need to study the large deviations. As in equilibrium, this requires to take log of probabilities and normalize by dividing by N: we are in fact looking at exponentially small

probabilities. The analysis is far from trivial and the knowledge we have acquired so far on the stationary measure $\mu^{(N)}$ does not help much. Since the analysis is very technical we just refer to the literature.

Theorem 4.5. *There exists a functional $F(\rho)$ on the space $\mathcal{M}$ of continuous densities $\rho(\cdot)$ on $[-1,1]$ with values $\rho_\pm$ at ± 1 such that the following holds. For any $\rho \in \mathcal{M}$*

$$\lim_{N\to\infty} \frac{-1}{N} \log \mu^{(N)}\left[\left|X^{(N)}(\phi_i) - \int \rho\phi_i\right| > \zeta, i = 1,\dots,n\right] = F(\rho) \tag{4.46}$$

for any n, any $\phi_1,\dots,\phi_n$ and any $\zeta > 0$.

The theorem has been proved by Derrida and Lebowitz who have an explicit expression for the functional $F(\rho)$. By its very construction $F(\rho)$ is nonnegative and 0 at ρ^{st}. The result has been framed in the context of temporal large deviations by Bertini et al. in within a general theory for the analysis of the large deviations in stationary nonequilibrium states. In the Bertini version, $F(\rho)$ can be written as

$$F(\rho) = \int \left\{ \rho\log\rho + (1-\rho)\log(1-\rho) + (1-\rho)\phi - \log\bigl(1+e^{\phi}\bigr) + \log\frac{\nabla\phi}{\nabla\rho^{\mathrm{st}}} \right\},$$

where ϕ is the solution of the equation

$$\frac{\Delta\phi}{|\nabla\phi|^2} + \frac{1}{1+e^{\phi}} = \rho$$

with boundary conditions $\phi(\pm 1) = \log\frac{\rho_\pm}{1-\rho_\pm}$. One can check that if $\rho = \rho^{\mathrm{st}}$ then $\phi = \log\frac{\rho}{1-\rho}$ and that $F(\rho) = 0$ for $\rho = \rho^{\mathrm{st}}$.

In equilibrium, we have used density fields and a formula analogous to (4.46) to define the free energy of profiles ρ and for this reason Jona-Lasinio used $F(\rho)$ to define the theory of nonequilibrium thermodynamics interpreting $F(\rho)$ as a nonequilibrium free energy (or rather minus entropy) functional. This is at first sight quite confusing as we have already from the equilibrium theory an expression for the free energy of a profile ρ given by

$$F^{\mathrm{eq}}(\rho) := \int f_\beta(\rho(r))\,dr, \tag{4.47}$$

where f_β is the thermodynamic free energy. In the case of the simple exclusion,

$$F^{\mathrm{eq}}(\rho) = \int \{\rho\log\rho + (1-\rho)\log(1-\rho)\}\,dr. \tag{4.48}$$

But we cannot have two different expressions for the free energy of a same profile ρ! Indeed the full system in the context of the Fick's law must include the two reservoirs: namely, to specify the state we should specify not only the profile ρ but also whether the system is in contact with reservoirs. This is more clear in the

representation of the reservoirs as infinite systems given in the previous section: then the state is specified by a density profile which extends to the whole space and not only to the interval $[-N, N]$ we are interested in.

The large deviations for the system of the previous section has not yet been done and it is not clear whether it leads to a functional analogous to $F(\rho)$. There are as you can see a lot of question marks and very many interesting questions in the study of stationary nonequilibrium states, too much for this short survey. I hope the reader is at this point intrigued by all that and it is therefore a good point to end here this introduction to the subject addressing the reader to the research papers in this field which is still very active.

4.7 Bibliographical remarks

The first paper (as far as I know) with hydrodynamic limit in the title is by Galves et al. (1981), where it is proved that the Fourier law is valid for the SSEP. It was then extended to a stochastic system of harmonic oscaillators by Kipnis, Marchioro and Presutti (1982). Lebowitz proposed to investigate the nature of the stationary measure by studying its large deviations. This was accomplished using the Derrida's matrix approach by Derrida, Lebowitz and Speer (2002). The analysis of the large deviations for the current was inititaed by Bodineau and Derrida (2007). A general theory of large deviations in nonequilibrium stationary measures and their relation with nonequilibrium thermodynamics has been developed in the last ten years by Bertini, De Sole, Gabrielli, Jona-Lasinio and Landim, see their most recent survey paper (Bertini et al. (2014)).

References

Au Yeung, Y., Friesecke, G. and Schmidt, B. (2009). Minimizing atomic configurations of short range pair potentials in two dimensions: Crystallization in the Wulff shape. *Calc. Var. Partial Differential Equations* **44** 81–100. MR2898772

Benfatto, G., Presutti, E. and Pulvirenti, M. (1978). DLR measures for one dimensional harmonic systems. *Z. Wahr. verw. Gebiete* **41**, 305–312. MR0467969

Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G. and Landim, C. (2014). Macroscopic fluctuation theory. Available at arXiv:1404.6466 [cond-mat.stat-mech].

Bodineau, T. and Derrida, B. (2007). Cumulants and large deviations of the current through nonequilibrium steady states. *C. R. Physique* **8**, 540–555.

Bovier, A., Merola, I., Presutti, E. and Zahradnik, M. (2004). On the Gibbs phase rule in the Pirogov–Sinai regime. *J. Stat. Phys.* **114**, 1235–1267. MR2039478

Bovier, A. and Zahradnik, M. (1997). The low-temperature phase of Kac–Ising models. *J. Stat. Phys.* **87**, 311–332. MR1453742

Braides, A. (2002). *Gamma-Convergence for Beginners*. *Oxford Lecture Series in Mathematics and Its Applications* **22**. Oxford: Oxford Univ. Press MR1968440

Cassandro, M. and Presutti, E. (1996). Phase transitions in Ising systems with long but finite range interactions. *Markov Process. Related Fields* **2**, 241–262. MR1414119

De Masi, A., Merola, I., Presutti, E. and Vignaud, Y. (2008). Potts models in the continuum. Uniqueness and exponential decay in the restricted ensembles. *J. Stat. Phys.* **133**, 281–345. MR2448535

De Masi, A., Merola, I., Presutti, E. and Vignaud, Y. (2009). Coexistence of ordered and disordered phases in Potts models in the continuum. *J. Stat. Phys.* **134**, 243–306. MR2485716

Derrida, B., Lebowitz, J. L. and Speer, E. (2002). Large deviations of the density profile in the steady state of the open symmetric simple exclusion process. *J. Stat. Phys.* **107**, 599–634. MR1898851

Dobrushin, R. L., Kotecký, R. and Shlosman, S. (1992). *The Wulff Construction: A Global Shape for Local Interactions*. Providence: Amer. Math. Soc. MR1181197

Galves, A., Kipnis, C., Marchioro, C. and Presutti, E. (1981). Non equilibrium measures which exhibit a temperature gradient: Study of a model. *Comm. Math. Phys.* **81**, 124–147. MR0630334

Gates, D. J. and Penrose, O. (1969). The van der Waals limit for classical systems. I. A variational principle. *Comm. Math. Phys.* **15**, 255–276. MR0260283

Gates, D. J. and Penrose, O. (1970a). The van der Waals limit for classical systems. II. Existence and continuity of the canonical pressure. *Comm. Math. Phys.* **16**, 231–237. MR1552568

Gates, D. J. and Penrose, O. (1970b). The van der Waals limit for classical systems. III. Deviation from the van der Waals–Maxwell theory. *Comm. Math. Phys.* **17**, 194–209. MR1552570

Georgii, H. O. (1988). *Gibbs Measures and Phase Transitions. de Gruyter Studies in Mathematics* **9**. Berlin: Walter de Gruyter. MR0956646

Heitmann, R. C. and Radin, C. (1980). The ground state for sticky disks. *J. Stat. Phys.* **22**, 281–287. MR0570369

Kac, M. and Helfand, E. (1963). Study of several lattice systems with long range forces. *J. Math. Phys.* **4**, 1078–1088. MR0153303

Kac, M. and Thompson, C. J. (1966). On the mathematical mechanism of phase transition. *Proc. Natl. Acad. Sci. USA* **55**, 676–683. MR0197136

Kac, M., Uhlenbeck, G. and Hemmer, P. C. (1963a). On the van der Waals theory of vapor–liquid equilibrium. I. Discussion of a one dimensional model. *J. Math. Phys.* **4**, 216–228. MR0148416

Kac, M., Uhlenbeck, G. and Hemmer, P. C. (1963b). On the van der Waals theory of vapor–liquid equilibrium. II. Discussion of the distribution functions. *J. Math. Phys.* **4**, 229–247. MR0148417

Kac, M., Uhlenbeck, G. and Hemmer, P. C. (1964). On the van der Waals theory of vapor–liquid equilibrium. III. Discussion of the critical region. *J. Math. Phys.* **5**, 60–74. MR0157692

Kipnis, C., Marchioro, C. and Presutti, E. (1982). An exactly solvable model for heat conduction. *J. Stat. Phys.* **27**, 65–74. MR0656869

Landau, L. D. and Lifshitz, E. M. (1999). *Statistical Physics*. Elmsford: Pergamon Press. MR0586944

Lebowitz, J. L., Mazel, A. and Presutti, E. (1999). Liquid–vapor phase transitions for systems with finite-range interactions. *J. Stat. Phys.* **94**, 955–1025. MR1694123

Lebowitz, J. L. and Penrose, O. (1966). Rigorous treatment of the van der Waals–Maxwell theory of the liquid–vapour transition. *J. Math. Phys.* **7**, 98–113. MR0187835

Liggett, T. M. (1985). *Interacting Particle Systems*. Berlin: Springer. MR0776231

Parthasarathy, K. R. (1967). *Probability Measures on Metric Spaces*. San Diego: Academic Press. MR0226684

Penrose, O. and Lebowitz, J. L. (1971). Rigorous treatment of metastable states in the van der Waals–Maxwell theory. *J. Stat. Phys.* **3**, 211–236. MR0293957

Presutti, E. (2008). *Scaling Limits in Statistical Mechanics and Microstructures in Continuum Mechanics*. Heidelberg: Springer. MR2460018

Radin, C. (1981). The ground state for soft disks. *J. Stat. Phys.* **26**, 365–373. MR0643714

Ruelle, D. (1969). *Statistical Mechanics: Rigorous Results*. Singapore: World Scientific, Imperial College Press. MR1747792

Ruelle, D. (1970). Superstable interactions in classical statistical mechanics. *Comm. Math. Phys.* **12**, 127–159. MR0266565

Sinai, Ya. G. (1982). *Theory of Phase Transitions: Rigorous Results*. Budapest: Akademiai Kiadó. MR0691854

Theil, F. (2006). A proof of crystallization in two dimensions. *Comm. Math. Phys.* **262**, 209–236. MR2200888

Zahradnik, M. (1998). A short course on the Pirogov–Sinai theory. *Rend. Mat. Appl.* **18**, 411–486. MR1686840

Gran Sasso Science Institute, Center for advanced studies
INFN
Viale Francesco Crispi 7
67100 L'Aquila
Italy
E-mail: errico.presutti@gmail.com