Bull. Soc. Math. France
147 (2), 2019, p. 197-219

GENERALIZED CURIE-WEISS MODEL AND QUADRATIC PRESSURE IN ERGODIC THEORY

by Renaud Leplaideur & Frédérique Watbled

Abstract. — We explain the Curie-Weiss model in statistical mechanics within an ergodic viewpoint. More precisely, we simultaneously define in $\{-1,+1\}^{\mathbb{N}}$, on the one hand a generalized Curie-Weiss model within the statistical mechanics viewpoint and on the other hand, the quadratic free energy and quadratic pressure within the ergodic theory viewpoint. We show that there are finitely many invariant measures that maximize the quadratic free energy. They are all dynamical Gibbs measures. Moreover, the probabilistic Gibbs measures for the generalized Curie-Weiss model converge to a determined combination of the (dynamical) conformal measures associated with these dynamical Gibbs measures. The standard Curie-Weiss model is a particular case of our generalized Curie-Weiss model. An ergodic viewpoint over the Curie-Weiss-Potts model is also given.

Texte reçu le 5 janvier 2017, modifié le 14 février 2018, accepté le 16 mars 2018.

Renaud Leplaideur, ISEA, Université de Nouvelle Calédonie, 145, Avenue James Cook – BP R4 98 851 – Nouméa Cedex. Nouvelle Calédonie • LMBA, UMR6205, Université de Brest • *E-mail :* `Renaud.Leplaideur@univ-nc.nc` • *Url :* `http://pagesperso.univ-brest.fr/~leplaide/`

Frédérique Watbled, LMBA, UMR 6205, Université de Bretagne Sud, Campus de Tohannic, BP 573, 56017 Vannes, France • *E-mail :* `frederique.watbled@univ-ubs.fr`

Mathematical subject classification (2010). — 37A35, 37A50, 37A60, 82B20, 82B30, 82C26.

Key words and phrases. — thermodynamic formalism, equilibrium states, Curie-Weiss model, Curie-Weiss-Potts model, Gibbs measure, phase transition.

F. Watbled thanks the IRMAR, CNRS UMR 6625, University of Rennes 1, for its hospitality. The authors thank the Centre Henri Lebesgue ANR-11-LABX-0020-01 for creating an attractive mathematical environment.

doi:10.24033/bsmf2779

Résumé (*Généralisation du modèle de Curie-Weiss et Pression quadratique en théorie ergodique*). — On explique ici un modèle généralisé de Curie-Weiss (champ moyen) en utilisant le vocabulaire de la théorie ergodique. On introduit le concept de pression quadratique en théorie ergodique et on montre que pour tout potentiel Hölder dans le sous-shift unilatère $\{-1,+1\}^{\mathbb{N}}$, il n'y a qu'un nombre fini de mesures invariantes qui maximisent la pression quadratique et, que ce sont toutes des mesures d'équilibre pour un multiple du potentiel. On montre que la limite thermodynamique des mesures de Gibbs associées à l'Hamiltonien en champ moyen convergent vers une combinaison des mesures conformes associées à chaque mesure qui maximise la pression quadratique. Le cas standard de Curie-Weiss s'obtient pour un exemple particulier de potentiel. Enfin, le modèle de Curie-Weiss-Potts est également expliqué avec le vocabulaire de la théorie ergodique.

1. Introduction

1.1. Background, main motivations and results. — The notion of Gibbs measure comes from statistical mechanics. It has been studied a lot from the probabilistic viewpoint (see [13, 6, 9, 10]). This notion was introduced in ergodic theory in the 70's by Sinai, Ruelle and Bowen (see [27, 28, 25, 24, 2]). Since that moment, the thermodynamic formalism in dynamical systems became a purely mathematical question and has somehow become isolated from the original physical questions.

It has turned out that this situation has generated sources of confusions. The first one is that while people share the same vocabulary, it is not clear that the same names precisely define the same notions in each viewpoint (ergodic vs physical). We *e.g.* refer to *phase transition, Gibbs measures, pressure*. Furthermore, the confusion is also internal to ergodic theory. Indeed, the thermodynamic formalism is presented very differently for $\mathbb{Z}$-actions (where the transfer operator plays a crucial role) or for $\mathbb{Z}^d$-actions (with $d > 1$). For this later case, the thermodynamic formalism is much closer to what people in statistical mechanics or in probability do. Several questions arising for 1-dimensional actions ergodic theory have to be exported to the higher dimensional case (see [4, 1]). Therefore, it became important to make clear similitudes and differences in the thermodynamic formalism between physical and (1-d) ergodic viewpoints.

Our first result (see Theorem 1.1) states a kind of dictionary between thermodynamic formalism in statistical mechanics and probability on the one hand, and ergodic theory on the other hand. More precisely we explain with the ergodic vocabulary the first-order phase transition arising for the Curie-Weiss model (mean field case), and make precise the link between Gibbs measures within the physical/probabilistic viewpoints and the ergodic viewpoint. We initially decided to focus on the mean field case for the following reasons. First, there is a large amount of literature dealing with this topic. Second, the mean field model is naturally represented into $\{-1,+1\}^{\mathbb{N}}$ and exhibits "physical phase transitions" that we wanted to compare with "1-d ergodic phase transitions" in $\{-1,+1\}^{\mathbb{N}}$.

From there, a subsequent task was to get a similar dictionary for the Curie-Weiss-Potts model which is a generalization of the Curie-Weiss model. This is done in Theorem 1.4.

These two results are then the motivation for our main result (see Theorem 1.3). The key point is that the Hamiltonian for the Curie-Weiss model is almost equal to the square of a Birkhoff sum. The Birkhoff sum is a key object in dynamical systems. We thus introduce within the ergodic viewpoint the notion of *quadratic free energy*. It is equal to the entropy plus the square of an integral. We are naturally led to study a variational principle, distinguishing the invariant measures that maximize the quadratic free energy. This maximum defines the *quadratic pressure*. At the same time, we introduce a generalized Hamiltonian in the Curie-Weiss model and show the link between the associated Gibbs measures (within physical/probabilistic viewpoint) and the Gibbs measures within the ergodic viewpoint. We show how first order phase transitions for this generalized Curie-Weiss model are related to a bifurcation into the set of measures which maximize the quadratic free energy. Theorem 1.1 is thus a particular case of Theorem 1.3.

We believe that this quadratic pressure generates further possible research questions in ergodic theory. Some of them are discussed later (see Subsubsection 1.2.5). Similarly, we believe that our generalized Curie-Weiss model may have physical interest.

Finally, we show that Theorem 1.3 is not an extension of Theorem 1.4. There is no obstruction to defining and studying the quadratic pressure for a more general subshift of a finite type. Nevertheless, the Hamiltonian for the Curie-Weiss-Potts model does not write itself as a square of a Birkhoff sum, because one considers a vector-valued "potential". This is work in progress to give an extension of Theorem 1.4 with the flavour of Theorem 1.3.

1.2. Precise settings and results. —

1.2.1. *Ergodic and Dynamical setting.* — We consider a finite set Λ with a cardinality greater than or equal to 2. It is called the alphabet. We also consider the one-sided full shift $\Sigma = \Lambda^{\mathbb{N}}$ over Λ. A point x in Σ is a sequence $x_0, x_1, \ldots$ (also called an infinite word) where the x_i are in Λ. Most of the time we shall use the notation $x = x_0x_1x_2\ldots$

A $x_i \in \Lambda$ can either be called a letter, or a digit or a symbol.

The shift map σ is defined by

$$\sigma(x_0x_1x_2\ldots) = x_1x_2\ldots.$$

The distance between two points $x = x_0x_1\dots$ and $y = y_0y_1\dots$ is given by

$$d(x,y) = \frac{1}{2^{\min\{n,\ x_n\neq y_n\}}}.$$

A finite string of symbols $x_0\dots x_{n-1}$ is also called a *word*, of length n. For a word w, its length is $|w|$. A *cylinder* (of length n) is denoted by $[x_0\dots x_{n-1}]$. It is the set of points y such that $y_i = x_i$ for $i = 0,\dots n-1$. We shall also talk about n-cylinder instead of cylinder of length n.

If w is the word of finite length $w_0\dots w_{n-1}$ and x is a word, the concatenation wx is the new word $w_0w_1\dots w_{n-1}x_0x_1\dots$.

For $\psi : \Sigma \to \mathbb{R}$ continuous and $\beta > 0$, the *pressure function* is defined by

$$\mathcal{P}(\beta\psi) := \sup_{\mu}\left\{h_\mu + \beta\int_\Sigma \psi\, d\mu\right\}, \tag{1}$$

where the supremum is taken among the set $\mathcal{M}_\sigma(\Sigma)$ of σ-invariant probabilities on Σ and h_μ is the Kolmogorov-Sinaï entropy of μ. The real parameter β is assumed to be positive because it represents the inverse of the temperature in statistical mechanics. It is known that the supremum is actually a maximum and any measure for which the maximum is attained in (1) is called an *equilibrium state for $\beta\psi$*. We refer the reader to [2, 25] for basic notions on thermodynamic formalism in ergodic theory.

If ψ is Lipschitz continuous then the Ruelle theorem (see [23]) states that for every β, there is a unique equilibrium state for $\beta\psi$, which is denoted by $\widetilde{\mu}_{\beta\psi}$. It is *ergodic* and it shall be called the *dynamical Gibbs measure* (DGM for short[1]). It is the unique σ-invariant probability measure which satisfies the property that for every $x = x_0x_1\dots$ and for every n,

$$e^{-C_\beta} \le \frac{\widetilde{\mu}_{\beta\psi}([x_0\dots x_{n-1}])}{e^{\beta.S_n(\psi)(x)-n\mathcal{P}(\beta\psi)}} \le e^{C_\beta}, \tag{2}$$

where C_β is a positive real number depending only on β and ψ (but not on x or n), and $S_n(\psi)$ stands for $\psi + \psi\circ\sigma + \dots + \psi\circ\sigma^{n-1}$.

In this setting, the *$\beta\psi$-conformal measure* is the unique probability measure such that for every x and for every n,

$$\nu_{\beta\psi}([x_0\dots x_{n-1}]) = \int e^{\beta S_n(\psi)(x_0\dots x_{n-1}y)-n\mathcal{P}(\beta\psi)}\, d\nu_{\beta\psi}(y). \tag{3}$$

A precise (and more technical) definition of conformal measure is given in page 207, where the connection between conformal measures and DGM is stated. We emphasize that in our setting, conformal measures and DGM are equivalent measures and one can obtain one from the other.

If the choice of ψ is clear we shall drop the ψ and write $\widetilde{\mu}_\beta$, ν_β and $\mathcal{P}(\beta)$.

1. We prefer the adjective "dynamical" instead of "ergodic" to avoid the discussion if an ergodic Gibbs measure is ergodic or not.

1.2.2. *The Curie-Weiss model.* — We consider the case $\Lambda = \{-1, +1\}$; Σ will be denoted by Σ_2.

If $\omega_0 \dots \omega_{n-1}$ is a finite word, we set

$$H_n(\omega) := -\frac{1}{2n} \sum_{i,j=0}^{n-1} \omega_j \omega_i. \tag{4}$$

It is called the *Curie-Weiss Hamiltonian.* The *empirical magnetization* for ω is $m_n(\omega) := \frac{1}{n} \sum_{j=0}^{n-1} \omega_j$. Then we have

$$H_n(\omega) = -\frac{n}{2}(m_n(\omega))^2. \tag{5}$$

We denote by $\mathbb{P} := \rho^{\otimes \mathbb{N}}$ the product measure on Σ_2, where ρ is the uniform measure on $\{-1, 1\}$, i.e. $\rho(\{1\}) = \rho(\{-1\}) = \frac{1}{2}$, and we define the *probabilistic Gibbs measure* (PGM for short) $\mu_{n,\beta}$ on Σ_2 by

$$\mu_{n,\beta}(d\omega) := \frac{e^{-\beta H_n(\omega)}}{Z_{n,\beta}} \mathbb{P}(d\omega), \tag{6}$$

where $Z_{n,\beta}$ is the normalization factor

$$Z_{n,\beta} = \frac{1}{2^n} \sum_{\omega',\ |\omega'|=n} e^{-\beta H_n(\omega')}.$$

Note that $\mu_{n,\beta}$ can also be viewed as a probability defined on Λ^n.

The measure $\mathbb{P}$ is a Bernoulli measure and is σ-invariant. In ergodic theory it is usually called the Parry-measure (see [21]) and turns out to be the unique measure with maximal entropy. With our previous notations it corresponds to the DGM $\widetilde{\mu}_0$.

If P_n, P are probability measures on the Borel sets of a metric space S, we say that P_n converges weakly to P if $\int_S f\, dP_n \to \int_S f\, dP$ for each f in the class $C_b(S)$ of bounded, continuous real functions f on S. In this case we write $P_n \underset{n\to+\infty}{\overset{w}{\longrightarrow}} P$.

Our first result concerns the weak convergence of the measures $\mu_{n,\beta}$.

Theorem 1.1 (Weak convergence for the CW model). — *Let ξ_β be the unique point in $[0,1]$ which realizes the maximum for*

$$\varphi_I(x) := \log(\cosh(\beta x)) - \frac{\beta}{2} x^2.$$

Let $\widetilde{\mu}_b$ be the dynamical Gibbs measure for $b(\mathbb{1}_{[+1]} - \mathbb{1}_{[-1]})$. Then

$$\mu_{n,\beta} \underset{n\to+\infty}{\overset{w}{\longrightarrow}} \begin{cases} \widetilde{\mu}_0 & \text{if } \beta \leq 1, \\ \frac{1}{2}\left[\widetilde{\mu}_{\beta\xi_\beta} + \widetilde{\mu}_{-\beta\xi_\beta}\right] & \text{if } \beta > 1. \end{cases} \tag{7}$$

REMARK 1.2. — Actually $\mu_{n,\beta}$ converges towards $\frac{1}{2}\left[\widetilde{\mu}_{\beta\xi_\beta}+\widetilde{\mu}_{-\beta\xi_\beta}\right]$ for every $\beta>0$ since we shall see that for $\beta\leq 1$ we have $\xi_\beta=0$, and it is clear that $\widetilde{\mu}_0=\rho^{\otimes\mathbb{N}}$.

We refer to [8], sections IV.4 and V.9, for a discussion of the Curie-Weiss model and historical references (see also [22], section 3.4). We also mention that the weak convergence of $\mu_{n,\beta}$ was already proved by Orey ([20], corollary 1.2) by a nice simple probabilistic argument. We recall that our motivation is the dictionary aspect and not the convergence.

We emphasize the equality

$$m_n(\omega):=\frac{1}{n}S_n(\mathbb{1}_{[+1]}-\mathbb{1}_{[-1]})(\omega) \tag{8}$$

which shows that m_n can be written as a Birkhoff mean of a continuous function.

A consequence of (8) is that (5) can be rewritten under the form

$$H_n(\omega)=-\frac{n}{2}\left(\frac{1}{n}S_n(\psi)(\omega)\right)^2,$$

where $\psi:=\mathbb{1}_{[+1]}-\mathbb{1}_{[-1]}$.

1.2.3. *The generalized Curie-Weiss model.* — If ψ is a Hölder continuous function on Σ_2, we define the *generalized Curie-Weiss Hamiltonian* H_n^ψ associated to ψ by setting

$$H_n^\psi(\omega)=-\frac{n}{2}\left(\frac{1}{n}S_n(\psi)(\omega)\right)^2.$$

Then $\mu_{n,\beta}^\psi$ is the PGM defined by

$$d\mu_{n,\beta}^\psi(d\omega):=\frac{e^{-\beta H_n^\psi(\omega)}}{Z_{n,\beta}^\psi}d\mathbb{P}(\omega),\ \text{with}\ Z_{n,\beta}^\psi=\int_{\Sigma_2}e^{-\beta H_n^\psi}\,d\mathbb{P}. \tag{9}$$

If μ is an invariant measure on Σ_2, we define its *quadratic free energy* by

$$h_\mu+\frac{\beta}{2}\left(\int_{\Sigma_2}\psi\,d\mu\right)^2.$$

Then we define the *quadratic pressure function (associated to* Ψ*)* by

$$\mathcal{P}_2(\beta\psi):=\sup_\mu\left\{h_\mu+\frac{\beta}{2}\left(\int_{\Sigma_2}\psi\,d\mu\right)^2\right\}. \tag{10}$$

Upper semi-continuity for the entropy immediately shows that the supremum is a maximum. The function $\beta\mapsto\mathcal{P}_2(\beta\psi)$ is obviously convex (thus continuous).

THEOREM 1.3 (Weak convergence for the generalized CW model). — *Let ψ be a Hölder continuous function on Σ_2, let β be a positive real number.*

1. *There are finitely many invariant probabilities* $m_1, \cdots, m_J$ *(with* $J = J(\beta)$*) whose quadratic free energy (for* β*) is maximal and thus equal to the quadratic pressure* $\mathcal{P}_2(\beta\psi)$*.*
2. *Each* m_i *is the unique equilibrium state* $\widetilde{\mu}_{\beta t_i \psi}$ *for the potential* $\beta t_i \psi$*.*
3. *The numbers* $t_1, \cdots, t_J$ *are the maxima of the auxiliary function*

$$\varphi_{OS}(t) := \mathcal{P}(\beta t\psi) - \frac{\beta}{2}t^2.$$

4. *As* n *goes to* $+\infty$*,* $\mu^{\psi}_{n,\beta}$ *converges weakly to a convex combination of the conformal measures* $\nu_{\beta t_j}$*'s associated to* $\beta t_j \psi$*:*

$$\mu_{n,\beta} \xrightarrow[n\to+\infty]{w} \sum_{j=1}^{J} c_j \nu_{\beta t_j}.$$

The c_j*'s are well identified (see formulas* (27) *and* (28)*).*

We emphasize that Theorem 1.1 is a particular case of Theorem 1.3 with $\psi = \mathbb{1}_{[+]} - \mathbb{1}_{[-]}$. In that case the pressure is easy to compute and is equal to

$$\mathcal{P}(\beta\psi) = \log 2 + \log(\cosh \beta),$$

and we get $\varphi_{OS}(x) = \log 2 + \varphi_I(x)$. Note that for this particular case, the DGM is also the conformal measure.

1.2.4. *Comparison of definitions of phase transition.* — Nowadays, a phase transition in ergodic theory means the lack of analyticity for the pressure function (see *e.g.* [5, 26, 19]). It is known that the loss of analyticity is transversal to the number of equilibrium states: one can have a loss of analyticity with only one equilibrium state (see the Manneville-Pomeau example with good parameters, [30]) or analyticity with several equilibrium states (see [18]). This means that the two notions of phase transition are transversal.

For the quadratic pressure, things may be different. We recall that $z \mapsto \mathcal{P}(z\psi)$ is analytic (for Hölder continuous ψ). Each t_i is a maximum for φ_{OS} and then satisfies $\mathcal{P}'(\beta t_i) = t_i$. It is thus highly probable that $t_i(\beta)$ is locally analytic (and surely locally $\mathcal{C}^\infty$). Then, the quadratic pressure satisfies

$$\mathcal{P}_2(\beta) = h_{\widetilde{\mu}_{\beta t_i}\psi} + \frac{\beta}{2}\left(\int_{\Sigma_2} \psi \, d\widetilde{\mu}_{\beta t_i \psi}\right)^2 = \mathcal{P}(\beta t_i \psi) - \beta t_i + \frac{\beta}{2} t_i^2.$$

It is thus reasonable to expect $\mathcal{P}_2(\beta)$ to be at least piecewise $\mathcal{C}^\infty$ and even probably piecewise $\mathcal{C}^\omega$. Moreover, we expect the borders of intervals of analyticity to be exactly where there is a change in the number of t_i's.

It is therefore very likely that the loss of analyticity for the quadratic pressure is equivalent to a bifurcation in the number of "quadratic" equilibrium states. Actually, this is corroborated by Theorem 1.1, where the quadratic pressure is piecewise analytic (and not analytic) and the number of quadratic equilibrium states changes with respect to β exactly where analyticity fails.

1.2.5. *Some consequences of Theorem 1.3.* — Several questions naturally arise from Theorem 1.3. At this stage, we do not have more precise conjectures or ideas for answers.

- As we said in the introduction, there is no obstruction to define and study the quadratic pressure for more general subshifts of a finite type. Nevertheless, the connection with the Curie-Weiss-Potts model, which is the extension of the mean-field model to alphabets with higher cardinality is not clear at all. Indeed, the Hamiltonian for the Curie-Weiss-Potts model is vector-valued and is not immediately written as the square of a Birkhoff sum.
- For more geometric dynamical systems, one usually considers or studies the special class of *physical or/and SRB-measures.* These measures are usually considered as the most natural ones with the measures of maximal entropy. It is clear that measures of maximal entropy also maximize $h_\mu + \left(\int_{\Sigma_2} \psi\, d\mu\right)^2$ for $\psi \equiv 0$.

  A natural question is that for a system admitting one SRB-measure, does there exists some potential ψ such that the SRB measure maximizes the quadratic free energy $h_\mu + \left(\int_{\Sigma_2} \psi\, d\mu\right)^2$.
- More generally, one can ask how big is the set of measures that maximizes the quadratic pressure for some potential ψ ? It is for instance known that any ergodic measure is an equilibrium state for some continuous potential (see [25, Cor. 3.17]). Does this still hold for quadratic pressure ?
- Ergodic optimization studies what happens to DGM $\widetilde{\mu}_{\beta\psi}$ as β goes to $+\infty$. It is known that any accumulation point maximizes the integral of ψ among invariant measures. The goal is to find out if there is convergence and how is the limit selected among the simplex of *ψ-maximizing measures.* The same kind of questions may be studied with the quadratic pressure. Note that non-linearity may introduce new and different phenomena compared to the "usual pressure".

1.2.6. *The Curie-Weiss-Potts model. Probabilistic settings and result.* — The *Curie-Weiss-Potts model* will be for $\Lambda = \{\theta^1, \dots, \theta^q\}$ with $q > 2$. In that case we shall write Σ_q instead of Σ.

The Curie-Weiss-Potts Hamiltonian is defined for a finite word $\omega = \omega_0 \cdots \omega_{n-1}$ by

$$H_n(\omega) := -\frac{1}{2n} \sum_{i,j=0}^{n-1} 1\!\!1_{\omega_j = \omega_i}. \tag{11}$$

We define the vector $L_n(\omega) = (L_{n,1}(\omega), \cdots, L_{n,q}(\omega))$ where

$$L_{n,k}(\omega) = \sum_{i=0}^{n-1} 1\!\!1_{\omega_i=\theta^k}$$

is the number of digits of ω which take the value θ^k, so that we can write

$$\sum_{i,j=0}^{n-1} 1\!\!1_{\omega_j=\omega_i} = \sum_{k=1}^{q} \left(\sum_{i=0}^{n-1} \mathbf{1}_{\omega_i=\theta^k} \right)^2 = \|L_n(\omega)\|^2,$$

where $\|\cdot\|$ stands for the Euclidean norm on $\mathbb{R}^q$.

We denote by $\mathbb{P} := \rho^{\otimes \mathbb{N}}$ the product measure on Σ_q, where ρ is the uniform measure on Λ, i.e. $\rho = \frac{1}{q}\sum_{k=1}^{q} \delta_{\theta^k}$, and we define the probabilistic Gibbs measure $\mu_{n,\beta}$ on Σ_q by

$$\mu_{n,\beta}(d\omega) := \frac{e^{-\beta H_n(\omega)}}{Z_{n,\beta}} \mathbb{P}(d\omega) = \frac{e^{\frac{\beta}{2n}\|L_n(\omega)\|^2}}{Z_{n,\beta}} \mathbb{P}(d\omega), \tag{12}$$

where $Z_{n,\beta}$ is the normalization factor

$$Z_{n,\beta} = \frac{1}{q^n} \sum_{\omega',\ |\omega'|=n} e^{\frac{\beta}{2n}\|L_n(\omega')\|^2}.$$

Now we can state the analog of Theorem 1.1.

Theorem 1.4 (Weak convergence for the CWP model). — *For $1 \le k \le q$, $b \in \mathbb{R}$, let $\widetilde{\mu}_b^k$ be the dynamical Gibbs measure for $b1\!\!1_{[\theta^k]}$. Let $\beta_c = \frac{2(q-1)\log(q-1)}{q-2}$. For $0 < \beta < \beta_c$ set $s_\beta = 0$ and for $\beta \ge \beta_c$ let s_β be the largest solution of the equation*

$$s = \frac{e^{\beta s} - 1}{e^{\beta s} + q - 1}. \tag{13}$$

Then,

$$\mu_{n,\beta} \xrightarrow[n\to+\infty]{w} \begin{cases} \rho^{\otimes\mathbb{N}} & \text{if } 0 < \beta < \beta_c, \\ \frac{1}{q}\sum_{k=1}^{q} \widetilde{\mu}_{\beta s_\beta}^k & \text{if } \beta > \beta_c, \\ \frac{A\,\widetilde{\mu}_0^1 + B\sum_{k=1}^{q} \widetilde{\mu}_{\beta_c s_{\beta_c}}^k}{A+qB} & \text{if } \beta = \beta_c, \end{cases} \tag{14}$$

with $A = \left(1 - \frac{\beta_c}{q(q-1)}\right)^{\frac{q-2}{2}}$ and $B = \left(1 - \frac{\beta_c}{q}\right)^{\frac{q-2}{2}}$.

Remark 1.5. — Actually $\mu_{n,\beta}$ converges towards $\frac{1}{q}\sum_{k=1}^{q} \widetilde{\mu}_{\beta s_\beta}^k$ for every $\beta \neq \beta_c$ since $s_\beta = 0$ for $\beta < \beta_c$, and it is clear that $\widetilde{\mu}_0^k = \rho^{\otimes\mathbb{N}}$ for each $1 \le k \le q$.

We refer to [11] for a discussion of the Curie-Weiss-Potts model and historical references. Orey ([20], Theorem 4.4) mentions the weak convergence of $\mu_{n,\beta}$ towards an explicit atomic measure, but he makes a mistake concerning the case $\beta = \beta_c$, as pointed out in [11].

1.3. Composition of the paper. — The paper is composed as follows.

In Section 2 we prove Theorem 1.3, in Section 3 we prove Theorem 1.4. Both proofs are based on the convergence of $\mu_{n,\beta}(C)$ where C is a cylinder in Σ.

We note that in Theorem 1.3 the proofs of the parts (3)–(4) and of the parts (1)–(2) are independent.

Theorem 1.1 is a simple consequence of Theorem 1.3 as mentioned above.

2. Proof of Theorem 1.3

2.1. Convergence of $\boldsymbol{\mu^{\psi}_{n,\beta}}$. — To lighten the notations we drop the ψ in H^{ψ}_n, $\mu^{\psi}_{n,\beta}$, $Z^{\psi}_{n,\beta}$. To prove the weak convergence of $\mu_{n,\beta}$ towards a measure μ, it is enough to show that for every cylinder C,

$$\lim_{n\to\infty} \mu_{n,\beta}(C) = \mu(C). \tag{15}$$

Let $\omega = \omega_0 \dots \omega_{p-1}$ be a finite word of length p, let $n > p$. We use the equality

$$e^{a^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}+\sqrt{2}ax} dx,$$

sometimes called the Hubbard-Stratonovich transformation ([15, 29]), to compute the following.

$$\begin{aligned}
Z_{n,\beta}\mu_{n,\beta}([\omega]) &= \int_{\Sigma_2} e^{\frac{\beta}{2n}(S_n(\psi)(\alpha))^2} \mathbb{1}_{[\omega]}(\alpha) d\mathbb{P}(\alpha) \\
&= \int_{\Sigma_2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} e^{\sqrt{\frac{\beta}{n}} x S_n(\psi)(\alpha)} \mathbb{1}_{[\omega]}(\alpha)\, dx\, d\mathbb{P}(\alpha), \\
&= \sqrt{\frac{\beta n}{2\pi}} \int_{-\infty}^{+\infty} e^{-n\frac{\beta}{2}z^2} \int_{\Sigma_2} e^{\beta z S_n(\psi)(\alpha)} \mathbb{1}_{[\omega]}(\alpha) d\mathbb{P}(\alpha) dz,
\end{aligned} \tag{16}$$

where we have made the change of variable $\beta z = \sqrt{\frac{\beta}{n}} x$.

Let us define the transfer operator $\mathcal{L}_\xi$, depending on a real or complex parameter ξ, by

$$\mathcal{L}_\xi(T)(x) := \sum_{y,\ \sigma(y)=x} e^{\xi\psi(y)} T(y).$$

Then for every $n \in \mathbb{N}$,

$$\mathcal{L}^n_\xi(T)(x) = \sum_{y,\ \sigma^n(y)=x} e^{\xi S_n\psi(y)} T(y). \tag{17}$$

We recall the following properties for $\mathcal{L}_\xi$ (see [2, 21]):

Proposition 2.1. — *The operator $\mathcal{L}_\xi$ acts on continuous and Hölder continuous functions. For Hölder continuous functions, its spectral radius λ_ξ is a simple dominating eigenvalue. We denote by H_ξ a positive associated eigenfunction. The rest of the spectrum of $\mathcal{L}_\xi$ is included into the disk of radius $\lambda_\xi e^{-\varepsilon(\xi)}$ for some positive $\varepsilon(\xi)$. The dual operator (for continuous functions) acts on measures. There exists a unique probability measure ν_ξ which is the eigen-measure for the eigen-value λ_ξ. The measure ν_ξ is the conformal measure. The DGM is then equal to*

$$d\widetilde{\mu}_\xi = H_\xi d\nu_\xi,$$

where H_ξ is normalized such that $\widetilde{\mu}_\xi$ is a probability measure. The pressure is $\log \lambda_\xi$.

We recall that the Ionescu-Tulcea & Marinescu theorem (see *e.g.* [3, 16]) holds for the transfer operator. This allows us to write

$$\mathcal{L}_\xi^n(\mathbb{1}_{[\omega]})(x) = \lambda_\xi^n \nu_\xi([\omega]) H_\xi(x) + \lambda_\xi^n e^{-n\varepsilon(\xi)} T(n,\xi)(x) \tag{18}$$

with $||T(n,\xi)||_\infty \leq 1$.[2]

For $\alpha \in \Sigma_2$ one writes $\alpha = \bar{\alpha}\theta$ with $\bar{\alpha}$ equal to the suffix of length n of α and θ in Σ_2. Using (17) we can rewrite (16) as

(19)

$$\begin{aligned} Z_{n,\beta}\mu_{n,\beta}([\omega]) &= \frac{1}{2^n}\sqrt{\frac{\beta n}{2\pi}} \int_{-\infty}^{+\infty} e^{-n\frac{\beta}{2}z^2} \int_{\Sigma_2} \sum_{\bar{\alpha}} e^{\beta z S_n(\psi)(\bar{\alpha}\theta)} \mathbb{1}_{[\omega]}(\bar{\alpha}) d\mathbb{P}(\theta)\, dz, \\ &= \frac{1}{2^n}\sqrt{\frac{\beta n}{2\pi}} \int_{-\infty}^{+\infty} e^{-n\frac{\beta}{2}z^2} \int_{\Sigma_2} \mathcal{L}_{\beta z}^n(\mathbb{1}_{[\omega]})(\theta) d\mathbb{P}(\theta)\, dz. \end{aligned}$$

The normalization factor $Z_{n,\beta}$ can be computed by replacing $[\omega]$ by Σ_2, and we get

$$\mu_{n,\beta}([\omega]) = \frac{\int_{-\infty}^{+\infty} e^{-n\frac{\beta}{2}z^2} \int_{\Sigma_2} \mathcal{L}_{\beta z}^n(\mathbb{1}_{[\omega]})(\theta) d\mathbb{P}(\theta) dz}{\int_{-\infty}^{+\infty} e^{-n\frac{\beta}{2}z^2} \int_{\Sigma_2} \mathcal{L}_{\beta z}^n(\mathbb{1})(\theta) d\mathbb{P}(\theta)\, dz} =: \frac{N_{n,\beta}}{D_{n,\beta}}. \tag{20}$$

Using (18) we get

$$\begin{aligned} N_{n,\beta} = \int_{-\infty}^{+\infty} & e^{-n\frac{\beta}{2}z^2 + n\log\lambda_{\beta z}} \\ & \left[\int_{\Sigma_2} \left(\nu_{\beta z}([\omega]) H_{\beta z}(\theta) + e^{-n\varepsilon(\beta z)} T(n,\beta z)(\theta)\right) d\mathbb{P}(\theta)\right] dz. \end{aligned} \tag{21}$$

2. Actually, this holds for the stronger Hölder norm, thus for the weaker norm on continuous functions. We will just need this weaker inequality.

We want to use the Laplace method here, but the last term in the inner integral depends on n. This term converges to zero as n goes to infinity but the speed of convergence depends on z and $|z|$ may go to infinity. Setting $A := \|\psi\|_\infty$, we deduce from (17) that for every n, every ξ and every T continuous

$$||\mathcal{L}_\xi^n(T)||_\infty \leq 2^n e^{n\xi A}||T||_\infty. \tag{22}$$

Therefore the term in the integral defining the numerator $N_{n,\beta}$ in (20) is bounded from above by $e^{-n\frac{\beta}{2}z^2+n\log 2+n\beta zA}$. Furthermore, $Z(\beta)$ exists such that for $|z| > Z(\beta)$

$$-\frac{\beta}{2}z^2 + \log 2 + \beta zA \leq -\frac{\beta}{4}z^2 \tag{23}$$

holds, from which we deduce for every $n > p = |\omega|$,

$$\int_{|z|\geq Z(\beta)} e^{-n\frac{\beta}{2}z^2}\int_{\Sigma_2}\mathcal{L}_{\beta z}^n(\mathbb{1}_{[\omega]})(\theta)d\mathbb{P}(\theta)\,dz \leq \frac{4}{n\beta Z(\beta)}e^{-\frac{n\beta}{4}Z^2(\beta)}. \tag{24}$$

From this we claim that the computation of the integral in (21) can be done for z in the compact set $[-Z(\beta), Z(\beta)]$ instead of $\mathbb{R}$. We shall *a posteriori* check that claim, when we will get an estimate for the integral in the compact set $[-Z(\beta), Z(\beta)]$.

As the spectral gap $\xi \mapsto \varepsilon(\xi)$ is lower semi-continuous (see [14]), the map $z \mapsto \varepsilon(\beta z)$ attains its infimum on $[-Z(\beta), Z(\beta)]$ so that $\int_{\Sigma_2} e^{-n\varepsilon(\beta z)}T(n,\beta z)(\theta)d\mathbb{P}(\theta)$ converges uniformly to zero on $[-Z(\beta), Z(\beta)]$. This yields that one can use the Laplace method for the convergence in (21), as we now explain.

The Laplace method shows that if $\varphi : I \to \mathbb{R}$ is a twice continuously differentiable function, if φ' vanishes on a single point ξ in the interior of the interval I, with $\varphi''(\xi) < 0$, and if $f : I \to \mathbb{R}$ is continuous with $f(\xi) \neq 0$, then

$$\int_I e^{n\varphi(y)}f(y)dy \underset{n\to\infty}{\sim} \frac{\sqrt{2\pi}}{\sqrt{|\varphi''(\xi)|}}e^{n\varphi(\xi)}f(\xi)n^{-1/2}, \tag{25}$$

where by $u_n \underset{n\to\infty}{\sim} v_n$ we mean that $u_n = v_n(1+\epsilon(n))$ with $\lim_{n\to+\infty}\epsilon(n) = 0$. We refer to [12] for a report about the Laplace method, and for a generalization to the case where the least integer k is such that $\varphi^{(k)}(\xi) \neq 0$ is greater than two. Of course when φ has a finite number of maxima we may break up the integral into a finite number of integrals so that in each integral φ reaches its maximum at only one interior point.

In our case we claim that the function $\varphi_{\mathrm{OS}} : z \mapsto -\dfrac{\beta}{2}z^2 + \log\lambda_{\beta z}$ admits only finitely many maxima. Indeed, for $z \notin [-Z(\beta), Z(\beta)]$, $\varphi(z) < -\frac{\beta}{4}z^2 < 0$ (this is a consequence of (22) and (23)) and $\varphi_{\mathrm{OS}}(0) = \log 2 > 0$. Therefore, the maxima for φ_{OS} must be in the compact interval $[-Z(\beta), Z(\beta)]$. If there are infinitely many, there must be some accumulation point. As φ_{OS} is analytic in some complex neighborhood of $[-Z(\beta), Z(\beta)]$, it must be equal to the constant function, which is clearly not the case. Let $t_1, \cdots, t_J$ be the points where φ_{OS} attains its maximum. We write the integral (21) over the segment $[-Z(\beta), Z(\beta)]$ as a finite sum of integrals over segments $[a_j, b_j]$ where each segment $[a_j, b_j]$ contains exactly one of the points t_j, $1 \leq j \leq J$.

We state the following lemma, which is an immediate adaptation of the Laplace method.

Lemma 2.2. — *Let $\varphi : [a, b] \to \mathbb{R}$ a function of class C^2, with φ' vanishing on a single point c in $]a, b[$ and $\varphi''(c) < 0$. Let $(f_n)_{n\geq 1}$, f be some continuous functions from $[a, b]$ to $\mathbb{R}$ such that f_n converges to f uniformly on $[a, b]$, and $f(c) \neq 0$. Then as $n \to \infty$*

$$\int_a^b e^{n\varphi(x)} f_n(x)\, dx \sim \sqrt{\frac{\pi}{2|\varphi''(c)|}} e^{n\varphi(c)} f(c) n^{-1/2}.$$

We apply this lemma on every $[a_j, b_j]$ to the functions f_n defined by

$$f_n(z) = \int_{\Sigma_2} \nu_{\beta z}([\omega]) H_{\beta z}(\theta) + e^{-n\varepsilon(\beta z)} T(n, \beta z)(\theta) d\mathbb{P}(\theta).$$

The functions f_n converge uniformly on $[a_j, b_j]$ to f defined by

$$f(z) = \left(\int_{\Sigma_2} H_{\beta z}(\theta) d\mathbb{P}(\theta) \right) \nu_{\beta z}([\omega]).$$

Putting together (24) and the result of Lemma 2.2 applied to every $[a_j, b_j]$, assuming for the moment that $\varphi''_{\mathrm{OS}}(t_j) < 0$ for every $j = 1, \cdots, J$, we obtain that $N^r_{n,\beta}$ is equivalent, when n goes to infinity, to

$$\sqrt{\frac{\pi}{2n}} e^{n\varphi_{\mathrm{OS}}(t_1)} \sum_{j=1}^{J} \frac{\left(\int_{\Sigma_2} H_{\beta t_j}(\theta) d\mathbb{P}(\theta) \right) \nu_{\beta t_j}([\omega])}{\sqrt{|\varphi''_{\mathrm{OS}}(t_j)|}}$$

and $D^r_{n,\beta}$ is equivalent to

$$\sqrt{\frac{\pi}{2n}} e^{n\varphi_{\mathrm{OS}}(t_1)} \sum_{j=1}^{J} \frac{\left(\int_{\Sigma_2} H_{\beta t_j}(\theta) d\mathbb{P}(\theta) \right)}{\sqrt{|\varphi''_{\mathrm{OS}}(t_j)|}},$$

where $N^r_{n,\beta}$ and $D^r_{n,\beta}$ stand for the same integrals as $N_{n,\beta}$, $D_{n,\beta}$, but restricted to $[-Z(\beta), Z(\beta)]$.

We can now check the claim we made above. We have just seen that the restrictions of integrals involved in (21) or (20) on the compact set $[-Z(\beta), Z(\beta)]$ are equivalent to constant $\times\, e^{n\varphi_{\mathrm{OS}}(t_i)} n^{-\frac{1}{2}}$, with $\varphi_{\mathrm{OS}}(t_i) > \log 2 > 0$, whereas the integrals outside the compact set are in $O(e^{-\frac{n\beta}{4} Z^2(\beta)}/n\beta Z(\beta))$. Therefore, integrals outside the compact set are negligible with respect to the integrals in the compact set and the claim is correct.

We can now finish the proof of Theorem 1.3. Recalling (20) we get that $\mu_{n,\beta}([\omega])$ converges to

$$\sum_{j=1}^{J} c_j \nu_{\beta t_j}([\omega]), \tag{26}$$

where

$$c_j := \frac{\dfrac{\displaystyle\int_{\Sigma_2} H_{\beta t_j}\, d\mathbb{P}}{\sqrt{|\varphi''_{\mathrm{OS}}(t_j)|}}}{\displaystyle\sum_{i=1}^{J} \dfrac{\displaystyle\int_{\Sigma_2} H_{\beta t_i}\, d\mathbb{P}}{\sqrt{|\varphi''_{\mathrm{OS}}(t_i)|}}}. \tag{27}$$

If $\varphi''_{\mathrm{OS}}(t_i) = 0$ then the contribution of the integral over $[a_i, b_i]$ is of order $e^{n\varphi_{\mathrm{OS}}(t_i)} n^{-1/k_i}$ where k_i is the least integer such that $\varphi_{\mathrm{OS}}^{(k_i)}(t_i) < 0$. Note that all $\varphi_{\mathrm{OS}}(t_i)$ are equal and the k_i's are all even numbers because φ_{OS} reaches its maximum at each t_i.

Let $K := \max k_i$ and let $\mathcal{I}$ be the set of indexes i's such that $k_i = K$. Then we still get the convergence of $\mu_{n,\beta}([\omega])$ to a convex combination (26) of measures $\nu_{\beta t_j}$'s, but with $c_j = 0$ whenever $j \notin \mathcal{I}$ and

$$c_j := \frac{\dfrac{\displaystyle\int_{\Sigma_2} H_{\beta t_j}\, d\mathbb{P}}{|\varphi_{\mathrm{OS}}^{(K)}(t_j)|^{1/K}}}{\displaystyle\sum_{i\in\mathcal{I}} \dfrac{\displaystyle\int_{\Sigma_2} H_{\beta t_i}\, d\mathbb{P}}{|\varphi_{\mathrm{OS}}^{(K)}(t_i)|^{1/K}}} \tag{28}$$

for $j \in \mathcal{I}$. This finishes the proof of part (4) of Theorem 1.3.

2.2. Measures maximizing the quadratic pressure.

— We want to determine the invariant measures m which maximize

$$h_m + \frac{\beta}{2}\left(\int_{\Sigma_2} \psi\, dm\right)^2.$$

We know that the space $\mathcal{M}_\sigma(\Sigma)$ of σ-invariant probabilities is a non-empty convex set which is metric compact in the weak* topology. Moreover in our setting the map $\nu \mapsto h_\nu$ is affine and upper semi-continuous (see for instance [31], Theorems 8.1 and 8.2). We set $\overline{A}$ (resp. $\underline{A}$) for $\displaystyle\max_{m\ \sigma-inv} \int_{\Sigma_2} \psi\, dm$ (resp.

$\min_{m\ \sigma-inv}\int_{\Sigma_2}\psi\,dm)$. For $z\in\mathbb{R}$, we set

$$\overline{H}(z) := \begin{cases} \max_{m\ \sigma-inv}\left\{h_m,\ \int_{\Sigma_2}\psi\,dm = z\right\} & \text{if } z\in[\underline{A},\overline{A}],\\ -\infty & \text{if not.}\end{cases}$$

Let us show that $\overline{H}$ is concave and upper semi-continuous. Let z_1, z_2 be in $[\underline{A},\overline{A}]$, λ be in $[0,1]$. There exists μ_i such that $\overline{H}(z_i) = h_{\mu_i}$ with $\int_{\Sigma_2}\psi\,d\mu_i = z_i$, for $i = 1, 2$. Set $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$. Then $\int_{\Sigma_2}\psi\,d\mu = \lambda z_1 + (1-\lambda)z_2$, hence $h_\mu \le \overline{H}(\lambda z_1 + (1-\lambda)z_2)$. As $h_\mu = \lambda h_{\mu_1} + (1-\lambda)h_{\mu_2}$ we deduce that

$$\lambda\overline{H}(z_1) + (1-\lambda)\overline{H}(z_2) \le \overline{H}\left(\lambda z_1 + (1-\lambda)z_2\right),$$

which proves the concavity of $\overline{H}$. To prove its upper semi-continuity we fix c in $\mathbb{R}$ and we show that the inverse image of $[c,+\infty]$ by $\overline{H}$ is closed. Let $(z_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ converging to z with $\overline{H}(z_n)\ge c$. Then z_n, z are in $[\underline{A},\overline{A}]$ and there exists μ, μ_n in $\mathcal{M}_\sigma(\Sigma)$ such that $\overline{H}(z) = h_\mu$, $\overline{H}(z_n) = h_{\mu_n}$, with $\int_{\Sigma_2}\psi\,d\mu = z$, $\int_{\Sigma_2}\psi\,d\mu_n = z_n$. Let $(\mu_{n_k})_{k\in\mathbb{N}}$ be a converging subsequence with limit m in $\mathcal{M}_\sigma(\Sigma)$. Then z_{n_k} converges to $\int_{\Sigma_2}\psi\,dm = z$, hence by definition $h_m\le\overline{H}(z)$. But we know that $h_m\ge c$ because $h_{\mu_{n_k}}\ge c$ and the map $\nu\mapsto h_\nu$ is upper semi-continuous. Therefore $\overline{H}(z)\ge c$, which concludes the proof of the upper semi-continuity of $\overline{H}$.

We note the equality

$$\mathcal{P}_2(\beta\psi) := \max_m\left\{h_m + \frac{\beta}{2}\left(\int_{\Sigma_2}\psi\,dm\right)^2\right\} = \max_{z\in[\underline{A},\overline{A}]}\left\{\overline{H}(z) + \frac{\beta}{2}z^2\right\}.$$

Let us set $\overline{\varphi}(z) := \overline{H}(z) + \dfrac{\beta}{2}z^2$. We claim that the maxima of φ_{OS} and $\overline{\varphi}$ are the same. First we observe that

$$\begin{aligned}\mathcal{P}(t\psi) := \max_m\left\{h_m + t\int_{\Sigma_2}\psi\,dm\right\} &= \max_{z\in\mathbb{R}}\left\{\overline{H}(z) + tz\right\}\\ &= \max_{z\in\mathbb{R}}\left\{tz - (-\overline{H}(z))\right\}.\end{aligned}$$

As the function $-\overline{H}$ is convex lower semi-continuous, we deduce from the duality property of the Fenchel-Legendre transform (see for instance [7], Lemma 4.5.8) that

$$\overline{H}(z) = \inf_{t\in\mathbb{R}}\left\{\mathcal{P}(t\psi) - tz\right\}. \tag{29}$$

Lemma 2.3. — *For every z in $[\underline{A},\overline{A}]$, $\overline{\varphi}(z)\le\varphi_{\mathrm{OS}}(z)$.*

Proof. — Let $t = \beta z$. Using (29) we get

$$\overline{\varphi}(z) \leq \mathcal{P}(t\psi) - tz + \frac{\beta}{2}z^2 = \mathcal{P}(\beta z\psi) - \frac{\beta}{2}z^2 = \varphi_{\mathrm{OS}}(z).$$

□

LEMMA 2.4. — *$\overline{\varphi}(z)$ is maximal if and only if $\varphi_{\mathrm{OS}}(z)$ is maximal. In that case, $\overline{\varphi}(z) = \varphi_{\mathrm{OS}}(z)$.*

Proof. — Let z be a maximum for φ_{OS}. Then, it is a critical point for φ_{OS}. This yields

$$\beta\mathcal{P}'(\beta z) = \beta z. \tag{30}$$

Now, we recall that $t \mapsto \mathcal{P}(t\psi)$ is analytic and $\mathcal{P}'(t\psi) = \int_{\Sigma_2} \psi\, d\widetilde{\mu}_t$ (see [25]). Hence, (30) yields $\displaystyle\int_{\Sigma_2} \psi\, d\widetilde{\mu}_{\beta z} = z$. Then,

$$\begin{aligned}\overline{\varphi}(z) \geq h_{\widetilde{\mu}_{\beta z}} + \frac{\beta}{2}z^2 &= h_{\widetilde{\mu}_{\beta z}} + \beta z^2 - \frac{\beta}{2}z^2 \\ &= h_{\widetilde{\mu}_{\beta z}} + \beta z \int_{\Sigma_2} \psi\, d\widetilde{\mu}_{\beta z} - \frac{\beta}{2}z^2 \\ &= \mathcal{P}(\beta z\psi) - \frac{\beta}{2}z^2 = \varphi_{\mathrm{OS}}(z) \geq \overline{\varphi}(z).\end{aligned}$$

This means that $\overline{\varphi}(z) = \varphi_{\mathrm{OS}}(z)$ holds. On the other hand for any z',

$$\overline{\varphi}(z') \leq \varphi_{\mathrm{OS}}(z') \leq \varphi_{\mathrm{OS}}(z) = \overline{\varphi}(z),$$

which shows that z is also a maximum for $\overline{\varphi}$.

Conversely, if z is a maximum for $\overline{\varphi}$, let z' be any maximum for φ_{OS}. We get

$$\overline{\varphi}(z) \geq \overline{\varphi}(z') = \varphi_{\mathrm{OS}}(z') \geq \varphi_{\mathrm{OS}}(z) \geq \overline{\varphi}(z).$$

This shows that z is also a maximum for φ_{OS}. □

Now we are ready to finish the proof of Theorem 1.3. Indeed let m maximize

$$h_m + \frac{\beta}{2}\left(\int_{\Sigma_2} \psi\, dm\right)^2.$$

Then $z := \displaystyle\int_{\Sigma_2} \psi\, dm$ is a maximum for $\overline{\varphi}$, hence according to Lemma 2.4 z is a maximum for φ_{OS} with $\varphi_{\mathrm{OS}}(z) = \overline{\varphi}(z)$. Therefore there exists $i \in [\![1, J]\!]$ such that $z = t_i$, and

$$h_m + \frac{\beta}{2}t_i^2 = \mathcal{P}(\beta t_i\psi) - \frac{\beta}{2}t_i^2.$$

We deduce that

$$h_m + \beta t_i^2 = \mathcal{P}(\beta t_i \psi) = h_m + \beta t_i \int_{\Sigma_2} \psi dm,$$

which implies that $m = \widetilde{\mu}_{\beta t_i \psi}$, by uniqueness of the equilibrium state. It just remains to prove that each $\widetilde{\mu}_{\beta t_i \psi}$ does maximize

$$h_m + \frac{\beta}{2}\left(\int_{\Sigma_2} \psi\, dm\right)^2.$$

But this is immediate since

$$\mathcal{P}'(\beta t_i) = \int_{\Sigma_2} \psi\, d\widetilde{\mu}_{\beta t_i \psi} = t_i$$

and t_i is a maximum for $\overline{\varphi}$.

3. Proof of Theorem 1.4

In a first step we use an auxiliary function φ_P. Note that this function was already studied by Ellis and Wang in [11]. Then we deduce that $\mu_{n,\beta}(C)$ converges for any cylinder C. In a second step we identify the limit as the relevant convex combination of dynamical measures.

3.1. Auxiliary function $\boldsymbol{\varphi_P}$ and convergence for $\boldsymbol{\mu_{n,\beta}}$. — We shall need the function φ_P defined on $\mathbb{R}^q$ by

$$\varphi_P(z) = -\frac{\beta}{2}\|z\|^2 + \log \sum_{k=1}^{q} e^{\beta z_k}. \tag{31}$$

This function attains its maximum on $\mathbb{R}^q$ since $\varphi_P(z) \le -c\|z\|^2$ as $\|z\|$ tends to ∞. We recall Theorem 2.1 of [11], which describes precisely the global maximum points of φ_P.

Theorem 3.1 ((Ellis Wang [11])). — *Let $\beta_c = \frac{2(q-1)\log(q-1)}{q-2}$. For $0 < \beta < \beta_c$ set $s_\beta = 0$ and for $\beta \ge \beta_c$ let s_β be the largest solution of the equation*

$$s = \frac{e^{\beta s} - 1}{e^{\beta s} + q - 1}. \tag{32}$$

The function $\beta \mapsto s_\beta$ is strictly increasing on the interval $[\beta_c, +\infty[$, $s(\beta_c) = \frac{q-2}{q-1}$, and $\lim_{\beta\to\infty} s_\beta = 1$.

Denote by ϕ the function from $[0,1]$ into $\mathbb{R}^q$ defined by

$$\phi(s) = \left(\frac{1+(q-1)s}{q}, \frac{1-s}{q}, \cdots, \frac{1-s}{q}\right),$$

the last $(q-1)$ components all equal $\frac{1-s}{q}$. Let K_β denote the set of global maximum points of the symmetric function φ_P. Define $\nu^0 = \phi(0) = \left(\frac{1}{q}, \cdots, \frac{1}{q}\right)$.

For $\beta \geq \beta_c$, define $\nu^1(\beta) = \phi(s_\beta)$ and let $\nu^i(\beta)$, $i = 2, \cdots, q$ denote the points in $\mathbb{R}^q$ obtained by interchanging the first and ith coordinates of $\nu^1(\beta)$. Then

$$K_\beta = \begin{cases} \{\nu^0\} & \text{for } 0 < \beta < \beta_c, \\ \{\nu^1(\beta), \nu^2(\beta), \cdots, \nu^q(\beta)\} & \text{for } \beta > \beta_c, \\ \{\nu^0, \nu^1(\beta_c), \nu^2(\beta_c), \cdots, \nu^q(\beta_c)\} & \text{for } \beta = \beta_c. \end{cases}$$

For $\beta \geq \beta_c$ the points in K_β are all distinct.

We fix a finite word $\omega = \omega_0 \cdots \omega_{p-1}$ of length p and we compute the limit of $\mu_{n,\beta}([\omega])$.

LEMMA 3.2. —

$$\lim_{n\to\infty} \mu_{n,\beta}([\omega]) = \begin{cases} \frac{1}{q^p} & \text{if } \beta < \beta_c, \\ \frac{1}{q} \frac{1}{(e^{\beta s_\beta} + q - 1)^p} \sum_{k=1}^q e^{\beta s_\beta L_{p,k}(\omega)} & \text{if } \beta > \beta_c, \\ \frac{\frac{A}{q^p} + \frac{B}{(e^{\beta s_\beta}+q-1)^p} \sum_{k=1}^q e^{\beta_c s_{\beta_c} L_{p,k}(\omega)}}{A + qB} & \text{if } \beta = \beta_c. \end{cases}$$

Proof. — We want to evaluate the limit of

$$\mu_{n,\beta}([\omega]) = \sum_{\alpha,\ |\alpha|=n-p} \mu_{n,\beta}([\omega\alpha]) = \frac{\displaystyle\sum_{\alpha,\ |\alpha|=n-p} e^{\frac{\beta}{2n}\|L_n(\omega\alpha)\|^2}}{\displaystyle\sum_{\alpha,\ |\alpha|=n} e^{\frac{\beta}{2n}\|L_n(\alpha)\|^2}}.$$

With the help of the identity

$$e^{\|u\|^2} = \frac{1}{(2\pi)^{q/2}} \int_{\mathbb{R}^q} \exp\left(-\frac{1}{2}\|y\|^2 + \sqrt{2}\langle y, u\rangle\right) dy, \tag{33}$$

and noticing that $L_n(\omega\alpha) = L_p(\omega) + L_{n-p}(\alpha)$, we write

$$\begin{aligned} \sum_{\alpha,\ |\alpha|=n-p} e^{\frac{\beta}{2n}\|L_n(\omega\alpha)\|^2} &= \frac{1}{(2\pi)^{q/2}} \int_{\mathbb{R}^q} e^{-\frac{1}{2}\|y\|^2} \sum_\alpha e^{\sqrt{\frac{\beta}{n}}\langle y, L_n(\omega\alpha)\rangle} dy \\ &= \frac{1}{(2\pi)^{q/2}} \int_{\mathbb{R}^q} e^{-\frac{1}{2}\|y\|^2 + \sqrt{\frac{\beta}{n}}\langle y, L_p(\omega)\rangle} \\ &\quad \cdot \sum_\alpha e^{\sqrt{\frac{\beta}{n}}\langle y, L_{n-p}(\alpha)\rangle} dy. \end{aligned}$$

It is easily seen that

$$\sum_{\alpha, |\alpha|=n-p} e^{\sqrt{\frac{\beta}{n}}\langle y, L_{n-p}(\alpha)\rangle} = \left(\sum_{k=1}^q e^{\sqrt{\frac{\beta}{n}} y_k}\right)^{n-p},$$

therefore we get

$$\sum_{\alpha,\ |\alpha|=n-p} e^{\frac{\beta}{2n}\|L_n(\omega\alpha)\|^2} =$$

$$\frac{1}{(2\pi)^{q/2}}\int_{\mathbb{R}^q} \exp\left(-\frac{1}{2}\|y\|^2 + \sqrt{\frac{\beta}{n}}\langle y, L_p(\omega)\rangle + (n-p)\log\left(\sum_{k=1}^{q} e^{\sqrt{\frac{\beta}{n}}y_k}\right)\right) dy.$$

Now we make the change of variable $\beta z = \sqrt{\frac{\beta}{n}}y$, and we obtain

$$\sum_{\alpha,\ |\alpha|=n-p} e^{\frac{\beta}{2n}\|L_n(\omega\alpha)\|^2} = \left(\frac{n\beta}{2\pi}\right)^{q/2} \int_{\mathbb{R}^q} e^{n\varphi_P(z)} f(z)dz, \tag{34}$$

where φ_P was defined in (31) and f is defined on $\mathbb{R}^q$ by

$$f(z) = \exp\left(\beta\langle z, L_p(\omega)\rangle - p\log\left(\sum_{k=1}^{q} e^{\beta z_k}\right)\right). \tag{35}$$

Similarly, $p = 0$ yields

$$\sum_{\alpha,\ |\alpha|=n} e^{\frac{\beta}{2n}\|L_n(\alpha)\|^2} = \left(\frac{n\beta}{2\pi}\right)^{q/2} \int_{\mathbb{R}^q} e^{n\varphi_P(z)}dz,$$

hence

$$\mu_{n,\beta}([\omega]) = \frac{\sum_{\alpha,\ |\alpha|=n-p} e^{\frac{\beta}{2n}\|L_n(\omega\alpha)\|^2}}{\sum_{\alpha,\ |\alpha|=n} e^{\frac{\beta}{2n}\|L_n(\alpha)\|^2}} = \frac{\int_{\mathbb{R}^q} e^{n\varphi_P(z)} f(z)\, dz}{\int_{\mathbb{R}^q} e^{n\varphi_P(z)}dz}.$$

We denote by $D\varphi_P(z)$, respectively $\mathcal{H}(z)$, the gradient, and the Hessian matrix, of φ_P at z. It is proved in Proposition 2.2 of [11] that the Hessian matrix of φ_P is negative definite at each global maximum point of φ_P.

Now, we recall that the Laplace method states

$$\int_0 e^{n\varphi_P(z)} f(z)\, dz \sim_{n\to\infty} \frac{(2\pi)^{q/2} f(z_0) e^{n\varphi_P(z_0)}}{n^{q/2}\sqrt{|\det \mathcal{H}(z_0)|}}.$$

provided that $D\varphi_P$ vanishes at a single point z_0 in an open set O of $\mathbb{R}^q$, that $f(z_0) \neq 0$ and that the Hessian matrix $\mathcal{H}(z_0)$ is negative definite (which holds by Proposition 2.2 of [11]). We consider several cases.

If $0 < \beta < \beta_c$: according to Theorem 3.1, φ_P attains its maximum at the unique point ν^0 so applying Laplace's method yields

$$\mu_{n,\beta}([\omega]) \sim_{n\to\infty} \frac{f(\nu^0)}{1} = \frac{1}{q^p}.$$

If $\beta > \beta_c$: Theorem 3.1 states that φ_P attains its maximum at exactly q points $\nu^i(\beta)$, $i = 1, \cdots, q$, where $\nu^i(\beta)$, $i = 2, \cdots, q$ is obtained by interchanging the first and ith coordinates of $\nu^1(\beta)$. Due to the symmetry of the function

φ_P it is clear that $\det \mathcal{H}(\nu^i) = \det \mathcal{H}(\nu^1)$, $i = 2, \cdots, q$. Considering a family of disjoint open sets $(O_i)_{1\le i\le q}$ such that O_i contains ν^i and $\mathbb{R}^q = \cup_{i=1}^q O_i \cup N$, where N is a set of measure zero, Laplace's method yields

$$\mu_{n,\beta}([\omega]) \sim_{n\to\infty} \frac{1}{q} \sum_{i=1}^{q} f(\nu^i).$$

Recall that

$$f(\nu^i) = \frac{e^{\beta\langle \nu^i, L_p(\omega)\rangle}}{\left(\sum_{k=1}^q e^{\beta\nu_k^i}\right)^p}$$

with

$$\nu_k^i = \begin{cases} \frac{1-s_\beta}{q} & \text{if } k \neq i, \\ \frac{1+(q-1)s_\beta}{q} & \text{if } k = i. \end{cases}$$

As $\sum_{k=1}^q L_{p,k}(\omega) = p$ it is easily seen that

$$e^{\beta\langle \nu^i, L_p(\omega)\rangle} = \exp\left(\frac{\beta p(1-s_\beta)}{q} + \beta s_\beta L_{p,i}(\omega)\right). \tag{36}$$

As ν^i is a critical point of φ_P and $\frac{\partial \varphi_P}{\partial z_i}(z) = \frac{\beta e^{\beta z_i}}{\sum_{k=1}^q e^{\beta z_k}} - \beta z_i$, we know that

$$\sum_{k=1}^{q} e^{\beta\nu_k^i} = \frac{e^{\beta\nu_j^i}}{\nu_j^i} = \frac{q}{1-s_\beta}\, e^{\frac{\beta(1-s_\beta)}{q}}. \tag{37}$$

Putting together (36) and (37) we obtain

$$f(\nu^i) = \left(\frac{1-s_\beta}{q}\right)^p e^{\beta s_\beta L_{p,i}(\omega)},$$

which can also be written

$$f(\nu^i) = \frac{1}{(e^{\beta s_\beta} + q - 1)^p}\, e^{\beta s_\beta L_{p,i}(\omega)} \tag{38}$$

since s_β is solution of the equation (13). Therefore

$$\mu_{n,\beta}([\omega]) \sim_{n\to\infty} \frac{1}{q} \frac{1}{(e^{\beta s_\beta} + q - 1)^p} \sum_{i=1}^{q} e^{\beta s_\beta L_{p,i}(\omega)}.$$

If $\beta = \beta_c$**:** the function φ_P admits exactly $q+1$ maximum points $\nu^i(\beta)$, $i = 0, \cdots, q$ but $\det \mathcal{H}(\nu^0) \neq \det \mathcal{H}(\nu^1)$, therefore Laplace's method yields

$$\mu_{n,\beta}([\omega]) \sim_{n\to\infty} \frac{|\det \mathcal{H}(\nu^0)|^{-1/2} f(\nu^0) + |\det \mathcal{H}(\nu^1)|^{-1/2} \sum_{i=1}^q f(\nu^i)}{|\det \mathcal{H}(\nu^0)|^{-1/2} + q\, |\det \mathcal{H}(\nu^1)|^{-1/2}}. \tag{39}$$

In the proof of Proposition 2.2 of [11] it is proved that $\mathcal{H}(\nu^0)$ has a simple eigenvalue at β and an eigenvalue of multiplicity $(q-1)$ at $\beta q^{-1}(q-\beta)$ whereas

$\mathcal{H}(\nu^1)$ has simple eigenvalues at β and $\beta - \beta^2 qab$ and an eigenvalue of multiplicity $(q-2)$ at $\beta - \beta^2 b$, where $a = q^{-1}(1+(q-1)s_\beta)$ and $b = q^{-1}(1-s_\beta)$. Recalling that $s(\beta_c) = \frac{q-2}{q-1}$ we deduce that

$$|\det \mathcal{H}(\nu^0)| = \beta_c^q (1-q^{-1}\beta_c)^{q-1},$$
$$|\det \mathcal{H}(\nu^1)| = \beta_c^q (1-q^{-1}\beta_c)\left(1-\frac{\beta_c}{q(q-1)}\right)^{q-2}.$$

Reporting in (39) and recalling (38) we get the result. □

3.2. Identification of the limit. — We can already deduce from Lemma 3.2 that $\mu_{n,\beta} \underset{n\to+\infty}{\overset{w}{\longrightarrow}} \widetilde{\mu}_0$ if $\beta < \beta_c$.

LEMMA 3.3 (Computation for $\widetilde{\mu}_b^k$). — *For $k = 1, \ldots, q$,*

$$\widetilde{\mu}_b^k([\omega]) = \frac{e^{bL_{p,k}(\omega)}}{(e^b+q-1)^p}. \tag{40}$$

Proof. — The function $b\mathbb{1}_{[\theta^k]}$ depends only on the zero coordinate, therefore the supremum in (1) is attained for the product measure $(m^k)^{\otimes\mathbb{N}}$, where the probability vector $(m_j^k)_{1\le j\le q}$ on Λ maximizes the quantity

$$-\sum_{j=1}^q p_j \log p_j + bp_k$$

over all the probability vectors $(p_j)_{1\le j\le q}$ on Λ, and is given by $m_k^k = \frac{e^b}{e^b+q-1}$, $m_j^k = \frac{1}{e^b+q-1}$ if $j \neq k$ (see for instance Example 4.2.2 of [17]). The result is then clear. □

The limit in (14) is now a direct consequence of Lemmas 3.2 and 3.3.

BIBLIOGRAPHY

[1] N. BEDARIDE, P. HUBERT & R. LEPLAIDEUR – "Thermodynamic formalism and substitutions", *ArXiv e-prints* (2015).

[2] R. BOWEN – *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics, Vol. 470, Springer-Verlag, Berlin, 1975, 2nd ed. - 2008 by JR Chazottes.

[3] A. BROISE – "Transformations dilatantes de l'intervalle et théorèmes limites", *Astérisque* (1996), no. 238, p. 1–109, Études spectrales d'opérateurs de transfert et applications.

[4] H. BRUIN & R. LEPLAIDEUR – "Renormalization, freezing phase transitions and Fibonacci quasicrystals", *Ann. Sci. Éc. Norm. Supér. (4)* **48** (2015), no. 3, p. 739–763.

[5] D. Coronel & J. Rivera-Letelier – “Low-temperature phase transitions in the quadratic family.”, *Adv. Math.* **248** (2013), p. 453–494.
[6] M. Costeniuc, R. S. Ellis & H. Touchette – “Complete analysis of phase transitions and ensemble equivalence for the Curie-Weiss-Potts model”, *J. Math. Phys.* **46** (2005), no. 6, p. 063301, 25.
[7] A. Dembo & O. Zeitouni – *Large deviations techniques and applications*, 2nd éd., Springer, 2010 (English).
[8] R. S. Ellis – *Entropy, large deviations, and statistical mechanics*, Classics in Mathematics, Springer-Verlag, Berlin, 2006, Reprint of the 1985 original.
[9] R. S. Ellis & C. M. Newman – “Limit theorems for sums of dependent random variables occurring in statistical mechanics”, *Z. Wahrsch. Verw. Gebiete* **44** (1978), no. 2, p. 117–139.
[10] ———, “The statistics of Curie-Weiss models”, *J. Statist. Phys.* **19** (1978), no. 2, p. 149–161.
[11] R. S. Ellis & K. Wang – “Limit theorems for the empirical vector of the Curie-Weiss-Potts model.”, *Stochastic Processes Appl.* **35** (1990), no. 1, p. 59–79 (English).
[12] A. Erdélyi – *Asymptotic expansions*, Dover Publications, Inc., New York, 1956.
[13] H.-O. Georgii – *Gibbs measures and phase transitions*, second éd., de Gruyter Studies in Mathematics, vol. 9, Walter de Gruyter & Co., 2011.
[14] H. Hennion & L. Hervé – *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Lecture Notes in Mathematics, vol. 1766, Springer, 2001.
[15] J. Hubbard – “Calculation of partition functions”, *Phys. Rev. Lett.* **3** (1959), p. 77–78.
[16] C. T. Ionescu Tulcea & G. Marinescu – “Théorie ergodique pour des classes d’opérations non complètement continues”, *Ann. of Math. (2)* **52** (1950), p. 140–147.
[17] G. Keller – *Equilibrium states in ergodic theory.*, Cambridge University Press, 1998 (English).
[18] R. Leplaideur – “Chaos: butterflies also generate phase transitions”, *J. Stat. Phys.* **161** (2015), no. 1, p. 151–170.
[19] N. Makarov & S. Smirnov – “On thermodynamics of rational maps. II. Non-recurrent maps”, *J. London Math. Soc. (2)* **67** (2003), no. 2, p. 417–432.
[20] S. Orey – “Large deviations for the empirical field of Curie-Weiss models”, *Stochastics* **25** (1988), no. 1, p. 3–14.
[21] W. Parry & M. Pollicott – *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque, vol. 187-188, SMF, 1990.

[22] F. Rassoul-Agha & T. Seppäläinen – *A course on large deviations with an introduction to Gibbs measures*, Graduate Studies in Mathematics, vol. 162, American Mathematical Society, 2015.

[23] D. Ruelle – "Statistical mechanics of a one-dimensional lattice gas.", *Commun. Math. Phys.* **9** (1968), p. 267–278.

[24] D. Ruelle – *Statistical mechanics*, World Scientific Publishing Co. Inc., River Edge, NJ, 1999, Rigorous results, Reprint of the 1989 edition.

[25] ______, *Thermodynamic formalism*, second éd., Cambridge Mathematical Library, Cambridge University Press, 2004, The mathematical structures of equilibrium statistical mechanics.

[26] O. Sarig – "Continuous phase transitions for dynamical systems", *Comm. Math. Phys.* **267** (2006), no. 3, p. 631–667.

[27] Y. Sinaĭ – "Gibbs measures in ergodic theory,", *Uspehi Mat. Nauk* **27** (1972), no. 4(166), p. 21–64.

[28] ______, *Theory of phase transitions: rigorous results*, International Series in Natural Philosophy, vol. 108, Pergamon Press, 1982, Translated from the Russian by J. Fritz, A. Krámli, P. Major and D. Szász.

[29] R. L. Stratonovič – "A method for the computation of quantum distribution functions", *Dokl. Akad. Nauk SSSR (N.S.)* **115** (1957), p. 1097–1100.

[30] M. Thaler – "Estimates of the invariant densities of endomorphisms with indifferent fixed points", *Israel J. Math.* **37** (1980), no. 4, p. 303–314.

[31] P. Walters – *An introduction to ergodic theory. Paperback ed.*, paperback ed. éd., Springer, 2000 (English).